

Document Version

Final published version

Licence

CC BY-SA

Citation (APA)

Pierzyna, M. (2026). *Machine learning parameterizations of atmospheric optical turbulence for free-space optical communication*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:3d9d6f7e-965d-4fe3-9881-bbf9fa9aef5c>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Machine learning parameterizations of atmospheric optical turbulence for free-space optical communication

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus,
Prof. dr. ir. H. Bijl,
chair of the Board for Doctorates
to be defended publicly on
Monday, 28 September 2026, 15:00

by

Maximilian PIERZYNA

This dissertation has been approved by the (co)promoters.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. A.P. Siebesma,	Delft University of Technology, <i>promotor</i>
Prof. dr. S. Basu,	University at Albany, <i>promotor</i>
Dr. ir. R. Saathof,	Delft University of Technology, <i>copromotor</i>

Independent members:

Prof. dr. S.J. Watson,	Delft University of Technology
Prof. dr. J. Osborn,	Durham University, United Kingdom
Prof. dr. M. van Iersel,	New Mexico State University, United States of America
Prof. dr. A.M.J. van Eijk,	École centrale de Nantes, France; TNO, The Netherlands
Prof. dr. ir. B.J.H. van de Wiel,	Delft University of Technology, <i>reserve member</i>

This work is part of the project FREE – Optical Wireless Superhighways: Free photons (at home and in space) (with project number P19-13) of the research programme TTW-Perspectief which is (partly) financed by the Dutch Research Council (NWO).



Keywords: Machine learning, Optical Turbulence, Boundary Layer Meteorology, Free-Space Optical Communication
Printed by: TBD
Cover by: TBD

Copyright © 2026 by M. Pierzyna

An electronic copy of this dissertation is available at
<https://repository.tudelft.nl/>.

Für alle, die mich begleiten und unterstützen.
To all those who support and accompany me.

Contents

Summary	ix
Samenvatting	xi
Zusammenfassung	xiii
1. Introduction	1
1.1. Direct and indirect C_n^2 observations	2
1.2. Optical turbulence parameterizations	5
1.2.1. Mesoscale modelling	5
1.2.2. Machine learning modelling	6
1.3. Scope of this dissertation	6
I. NUMERICAL MESOSCALE MODELLING – THE TRADITIONAL APPROACH	9
2. Intercomparison of flux, gradient, and variance-based optical turbulence (C_n^2) parameterizations	11
2.1. Introduction	12
2.2. Methodology	13
2.2.1. C_n^2 estimation from observed and simulated meteorology	13
2.2.2. Scintillometer observations of C_n^2 at the CESAR site	17
2.2.3. Meteorological observations within the surface layer	19
2.2.4. Mesoscale simulation setup	19
2.2.5. Performance metrics	22
2.3. Results	23
2.3.1. Three-way comparison of C_n^2 parameterizations in the surface layer	24
2.3.2. Beyond the surface layer: performance of the variance-based C_n^2 parameterization	26
2.4. Conclusion	28
2.A. Technical details of gradient computation	31
3. Estimating turbulence-induced link losses in an urban environment from observed and numerically estimated C_n^2	33
3.1. Introduction	34
3.2. Example case: Reid Photonloop with collimated beam	35
3.3. Estimation of FSOC link performance under atmospheric turbulence	37
3.3.1. Integrated Turbulence Parameters	40
3.3.2. Turbulence-induced link losses	41

3.3.3. Generalized mutual information	42
3.4. Results	43
3.4.1. Meteorological assessment of the WRF simulation	45
3.4.2. Estimated link losses and link performance	47
3.5. Conclusion	52
3.A. A note on BLS uncertainties	55
3.B. Summary of simplifications and assumptions	56
II. MACHINE LEARNING MODELING – THE FUTURE?	59
4. OTClm: Generating a near-surface climatology of optical turbulence strength (C_n^2) using gradient boosting	61
4.1. Introduction	62
4.2. Relevant C_n^2 regression studies	63
4.3. Methodology	65
4.3.1. Gradient boosting regression	65
4.3.2. Baseline models	67
4.3.3. Physical consistency checking using feature importance	67
4.3.4. Performance evaluation	68
4.4. Dataset	69
4.4.1. New York State Mesonet (NYSM)	69
4.4.2. ERA5 reanalysis data	72
4.5. Results: feature importance	75
4.6. Results: OTClm performance	78
4.6.1. Temporal extrapolation	78
4.6.2. Cross-site evaluation	80
4.7. Conclusion	82
4.A. Ratio of stable to unstable conditions in dataset before and after QA/QC	84
4.B. Scaling of $\log_{10} C_n^2$ target data	84
4.C. Details about baseline models	85
4.C.1. W71 C_n^2 parameterization	86
4.C.2. In-situ GBM models	86
5. Π-ML: A dimensional analysis-based machine learning parameterization of optical turbulence in the atmospheric surface layer	87
5.1. Introduction	88
5.2. Methodology	88
5.3. Results	91
5.4. Conclusion	95
5.A. Performance evaluation using different test set intervals	96
5.B. Full list of non-dimensional Π sets	96
5.C. Performance of a Π -ML model of reduced complexity	96
5.D. Performance of an XGBoost ensemble trained with traditional features	97

6. Beyond Hufnagel-Valley: estimating high-resolution profiles of optical turbulence using deep learning	101
6.1. Introduction	102
6.2. Methods	103
6.2.1. Dataset alignment and decoupled training	103
6.2.2. Squeezeformer	106
6.2.3. Baseline models	107
6.2.4. Model evaluation	108
6.3. Datasets	111
6.3.1. WRF: year-long C_n^2 database for NL	112
6.3.2. ERA5 reanalysis	113
6.3.3. Construction of the deep learning datasets	113
6.4. Results	115
6.4.1. Estimated $\log_{10} C_n^2$ profiles	117
6.4.2. Estimates of r_0 and σ_I^2	121
6.5. Conclusion	123
6.A. Influence of quantile mapping on decoupled training	126
6.B. Additional details of Squeezeformer architecture	129

III. CONCLUSIONS AND PERSPECTIVES **133**

7. Conclusions and perspectives	135
7.1. Traditional mesoscale modelling	135
7.2. Machine learning modelling	137
7.3. Recommendations and perspectives	139
7.4. Final words	144

IV. APPENDICES **145**

A. A multi-physics ensemble modeling framework for reliable C_n^2 estimation	147
A.1. Introduction	148
A.2. Methodology	148
A.3. Results	150
A.4. Conclusion	152
B. Vertical extension of Π-ML – a data-driven similarity theory of C_T^2 for Cabauw	155
B.1. Introduction	156
B.2. Use cases	157
B.3. Results	158
B.4. Conclusion	160

Bibliography **163**

List of publications **180**

List of presentations	181
Curriculum vitae	182

Summary

Free-space optical communication (FSOC) is a key technology to meet the growing demand for high-bandwidth, secure, and energy-efficient data links. However, the atmospheric channel introduces a major challenge: optical turbulence (OT). Turbulent fluctuations of the refractive index, driven by wind shear and buoyancy in the atmosphere, distort the propagating optical beam, degrading communication performance. This dissertation investigates the modelling of optical turbulence, quantified through the refractive index structure parameter C_n^2 , along two complementary avenues: traditional numerical weather prediction (NWP) using mesoscale models and machine learning (ML) techniques.

The first part establishes the state of the art in mesoscale C_n^2 modelling for FSOC. [Chapter 2](#) presents a systematic intercomparison of the three main classes of C_n^2 parameterizations – flux-based, gradient-based, and variance-based – applied to both observed data and output from a numerical mesoscale model. Evaluated against scintillometer observations, the variance-based parameterization yields the best overall performance and, unlike the other two, is not restricted to the atmospheric surface layer. Building on this foundation, [chapter 3](#) proposes an end-to-end framework that translates C_n^2 estimates into FSOC link performance metrics, specifically turbulence-induced losses and a theoretically achievable information rate (generalized mutual information, GMI). Applied to an urban example link, the framework reveals that the sensitivity of the estimated FSOC performance to C_n^2 errors strongly depends on the specific link. Reporting C_n^2 model errors alone is therefore considered insufficient for FSOC, and end-to-end assessments of link performance are needed.

Running mesoscale models is computationally expensive, and the traditional parameterizations used to obtain C_n^2 – while physically motivated – are often limited in their applicability or accuracy. Machine learning offers a promising alternative by learning complex relationships directly from data, as explored in the second part of this dissertation. [Chapter 4](#) introduces OTcliM, a gradient boosting-based methodology that relates globally available ERA5 reanalysis data to observed near-surface C_n^2 . Across 17 diverse stations in New York State, OTcliM trained on just one year of observations accurately extrapolates C_n^2 over four held-out years. However, the resulting models are limited to near-surface C_n^2 , which is sufficient for terrestrial links but not for satellite-to-ground applications that require vertical profiles. To extend modelling into the vertical, [chapter 5](#) introduces Π -ML, a physics-inspired ML framework that combines automated dimensional analysis with gradient boosting. By expressing inputs and outputs as non-dimensional groups grounded in physical principles, Π -ML derives an interpretable, data-driven similarity theory for C_n^2 in the atmospheric surface layer. The framework achieves high accuracy under both stable and unstable atmospheric conditions and identifies scalings that are physically consistent with established theories. Finally, [chapter 6](#) addresses the

full atmospheric column by developing OTProf. This deep-learning-based model estimates C_n^2 profiles of high vertical resolution from low-resolution ERA5 pressure-level inputs. Trained on one year of mesoscale simulations, the model substantially outperforms the commonly used Hufnagel-Valley analytical model. The estimated profiles exhibit a physically realistic vertical structure, although ML-typical smoothing leads to some underestimation of integrated turbulence parameters, such as the Fried parameter and the scintillation index. Nevertheless, OTProf is regarded as a computationally efficient and considerably more accurate approach compared to traditional analytical models.

Taken together, this dissertation demonstrates that both numerical mesoscale modeling and machine learning deliver practical advances for optical turbulence estimation and FSOC applications. The overarching conclusion, however, is that the greatest potential lies in combining the two paradigms, which is viewed as a promising direction for future research.

Samenvatting

Free-space optische communicatie (engl. Free-Space Optical Communication, FSOC) is een belangrijke technologie om te voldoen aan de groeiende vraag naar veilige en energiezuinige datalinks met een hoge bandbreedte. Het atmosferische kanaal vormt echter een grote uitdaging vanwege optische turbulentie (OT). Turbulente fluctuaties van de brekingsindex, veroorzaakt door windschering en thermiek in de atmosfeer, verstoren de voortplanting van de optische straal, waardoor de communicatieprestaties verslechteren. Dit proefschrift onderzoekt de modellering van optische turbulentie, gekwantificeerd door middel van de brekingsindexstructuurparameter C_n^2 , langs twee complementaire wegen: traditionele numerieke weersvoorspelling (NWP) met behulp van mesoschaalmodellen en machine learning (ML)-technieken.

Het eerste deel legt de stand van zaken vast op het gebied van mesoschaal C_n^2 -modellering voor FSOC. Hoofdstuk 2 presenteert een systematische vergelijking van de drie belangrijkste klassen van C_n^2 -parametriseringen – flux, gradiënt en variantie gebaseerd – toegepast op zowel geobserveerde data als output van een numeriek mesoschaalmodel. De resultaten zijn vergeleken met observaties met scintillometers, waarbij de op variantie gebaseerde parameterisatie de beste algemene resultaten oplevert. Deze is, in tegenstelling tot de andere twee, niet beperkt tot de atmosferische grenzelaag. Voortbouwend op deze basis stelt hoofdstuk 3 een end-to-end-raamwerk voor dat C_n^2 -schattingen vertaalt naar FSOC-linkprestatiestatistieken. Deze schat een theoretisch haalbare informatiesnelheid (Generalized Mutual Information, GMI) die met name veroorzaakt worden door turbulentie verliezen. Toegepast op een stedelijke voorbeeldlink, laat het raamwerk zien dat de sensitiviteit van de geschatte FSOC-prestaties voor C_n^2 -fouten sterk afhankelijk is van de specifieke link. Het rapporteren van alleen C_n^2 -modelfouten wordt daarom als onvoldoende beschouwd voor FSOC, en er zijn end-to-end beoordelingen van linkprestaties nodig.

Het draaien van mesoschaalmodellen is rekenintensief en de traditionele parameterisaties die worden gebruikt om C_n^2 te verkrijgen – hoewel fysisch gemotiveerd – zijn vaak beperkt in hun toepasbaarheid of nauwkeurigheid. Machine learning biedt een veelbelovend alternatief door complexe relaties rechtstreeks uit gegevens te leren, zoals onderzocht in het tweede deel van dit proefschrift. Hoofdstuk 4 introduceert OTClIM, een op gradient boosting gebaseerde methodologie die wereldwijd beschikbare ERA5-reanalysegegevens relateert aan geobserveerde C_n^2 nabij het aardoppervlak. OTClIM traint een ML model die gebruikt maakt van één jaar aan C_n^2 -data van 17 verschillende stations in de staat New York en extrapoleert deze, over vier jaren die buiten beschouwing zijn gelaten. De resulterende modellen zijn echter beperkt tot C_n^2 nabij het aardoppervlak, wat voldoende is voor terrestrische verbindingen, maar niet voor satelliet-naar-grondtoepassingen die verticale profielen vereisen. Om de modellering uit te breiden naar de verticale richting, introduceert hoofdstuk 5 Π -ML, een op de natuurkunde ge-

inspireerd ML-raamwerk dat geautomatiseerde dimensieanalyse combineert met gradient boosting. Door inputs en outputs uit te drukken als niet-dimensionale groepen die zijn gebaseerd op fysische principes, leidt Π -ML een interpreteerbare, datagestuurde gelijkvormigheidstheorie af voor C_n^2 in de onderste atmosferische grenslaag. Het raamwerk bereikt een hoge nauwkeurigheid onder zowel stabiele als onstabiele atmosferische omstandigheden en identificeert dimensionele scaleringen die fysisch consistent zijn met gevestigde theorieën. Ten slotte behandelt hoofdstuk 6 de volledige atmosferische kolom door OTProf te ontwikkelen. Dit op deep learning gebaseerde model schat C_n^2 -profielen met een hoge verticale resolutie op basis van ERA5 pressure level data met een lage resolutie. Het model is getraind op basis van een jaar aan mesoschaalsimulaties en presteert aanzienlijk beter dan het veelgebruikte analytische model van Hufnagel-Valley. De geschatte profielen vertonen een fysisch realistische verticale structuur, hoewel de voor ML typische afvlakking leidt tot een zekere onderschatting van geïntegreerde turbulentieparameters, zoals de Fried-parameter en de scintillatie-index. Niettemin wordt OTProf beschouwd als een computationeel efficiënte en aanzienlijk nauwkeurigere benadering in vergelijking met traditionele analytische modellen.

Al met al toont dit proefschrift aan dat zowel numerieke mesoschaalmodellering als machine learning praktische voordelen bieden voor optische turbulentieschattingen en FSOC-toepassingen. De overkoepelende conclusie is echter dat het grootste potentieel ligt in het combineren van beide paradigma's, wat wordt gezien als een veelbelovende richting voor toekomstig onderzoek.

Zusammenfassung

Optischer Richtfunk (engl. Free-Space Optical Communication, FSOC) ist eine Schlüsseltechnologie, um den wachsenden Bedarf an Datenverbindungen mit hoher Bandbreite, hoher Sicherheit und hoher Energieeffizienz zu decken. Innerhalb des atmosphärischen Kanals stellt jedoch die so genannte optische Turbulenz (OT) eine große Herausforderung dar. Turbulente Schwankungen des Brechungsindex, verursacht durch Windscherung und Thermik in der Atmosphäre, verzerren den sich ausbreitenden Lichtstrahl und beeinträchtigen die Kommunikationsleistung. Diese Dissertation untersucht die Modellierung optischer Turbulenzen, quantifiziert durch den Brechungsindex-Strukturparameter C_n^2 , anhand von zwei komplementären Ansätzen: der traditionellen numerischen Wettervorhersage (NWP) unter Verwendung von Mesoskalenmodellen und mittels Machine Learnings (ML).

Der erste Teil etabliert den Stand der Technik in der mesoskaligen C_n^2 -Modellierung für FSOC. Kapitel 2 präsentiert einen systematischen Vergleich der drei Hauptklassen von C_n^2 -Parametrisierungen – flussbasiert, gradientenbasiert und varianzbasiert –, die sowohl auf Beobachtungsdaten als auch auf den Output eines numerischen Mesoskalenmodells angewandt werden. Im Vergleich zu Scintillometer-Beobachtungen liefert die varianzbasierte Parametrisierung das beste Gesamtergebnis und ist im Gegensatz zu den beiden anderen nicht auf die atmosphärische Grenzschicht beschränkt. Auf dieser Grundlage wird in Kapitel 3 ein End-to-End-Framework vorgeschlagen, das C_n^2 -Schätzungen in FSOC-Link-Leistungsmetriken übersetzt, konkret in turbulenzbedingte Verluste und eine theoretisch erreichbare Informationsrate (generalisierte gegenseitige Information, GMI). Angewandt auf einen urbanen Beispiellink zeigt das Framework, dass die Empfindlichkeit der geschätzten FSOC-Leistung gegenüber C_n^2 -Fehlern stark von der jeweiligen Verbindung abhängt. Die alleinige Angabe von C_n^2 -Modellfehlern wird daher für FSOC als unzureichend angesehen, weshalb eine End-to-End-Bewertung der Leistungskennzahlen erforderlich ist.

Die Ausführung von Mesoskalenmodellen ist rechenintensiv, und die traditionellen Parametrisierungen, die zur Ermittlung von C_n^2 verwendet werden, sind – obwohl physikalisch motiviert – oft in ihrer Anwendbarkeit oder Genauigkeit eingeschränkt. Maschinelles Lernen bietet eine vielversprechende Alternative, indem es komplexe Zusammenhänge direkt aus Daten lernt, wie im zweiten Teil dieser Dissertation untersucht wird. Kapitel 4 präsentiert OTcliM, eine auf Gradient Boosting basierende Methodik, die global verfügbare ERA5-Reanalyse-Daten mit C_n^2 -Messungen in Bodennähe in Beziehung setzt. Es wird anhand von 17 verschiedenen Messstationen im Bundesstaat New York gezeigt, dass OTcliM mit nur einem Jahr an beobachteten C_n^2 -Trainingsdaten, C_n^2 für vier weitere Testjahre mit hoher Genauigkeit zeitlich extrapolieren kann. Die resultierenden Modelle sind jedoch auf oberflächennahes C_n^2 beschränkt, was für terrestrische Verbindungen ausreichend ist. Für Satellit-Boden-Verbindungen sind jedoch vertikale Profile

erforderlich. Um die Modellierung auf die Vertikale auszuweiten, führt Kapitel 5 Π -ML ein, ein physikalisch inspiriertes ML-Framework, das automatisierte Dimensionsanalyse mit Gradient Boosting kombiniert. Durch die Darstellung von unabhängigen und abhängigen Variablen (Inputs und Outputs) als physikalisch motivierte dimensionslose Gruppen kann Π -ML eine interpretierbare, datengetriebene Ähnlichkeitstheorie für C_n^2 in der atmosphärischen Oberflächenschicht aufstellen. Das Framework erreicht sowohl unter stabilen als auch unter instabilen atmosphärischen Bedingungen eine hohe Genauigkeit und identifiziert Π -Gruppen, die physikalisch mit etablierten Theorien übereinstimmen.

Abschließend wird in Kapitel 6 die Modellierung der gesamte atmosphärische Säule behandelt. Mit Hilfe des dort vorgestellten Deep-Learning-Modells OTProf können C_n^2 -Profile mit hoher vertikaler Auflösung aus niedrigauflösenden ERA5-Druckniveaudaten abgeschätzt werden. Das Modell wurde mit einem Jahr an mesoskaligen Simulationen trainiert und übertrifft das häufig verwendete analytische Hufnagel-Valley-Modell deutlich. Die abgeschätzten Profile weisen eine physikalisch realistische vertikale Struktur auf, wobei jedoch die für ML typische Glättung zu einer gewissen Unterschätzung der integrierten Turbulenzparameter wie dem Fried-Parameter und dem Scintillationsindex führt. Dennoch kann OTProf im Vergleich zu traditionellen analytischen Modellen als rechnerisch effizienter und wesentlich genauer angesehen werden.

Insgesamt belegt diese Dissertation, dass sowohl die numerische Mesoskalenmodellierung als auch Machine Learning praktische Fortschritte für die Modellierung optischer Turbulenz und für FSOC-Anwendungen liefern. Die übergeordnete Schlussfolgerung lautet hingegen, dass das größte Potenzial in der Kombination beider Ansätze liegt, was als vielversprechende Richtung für die zukünftige Forschung angesehen wird.

1

Introduction

Free-space optical communication (FSOC) is considered a key technology to support the bandwidth and connectivity demands of an increasingly data-intensive world (Hemmati 2009). By transmitting data with optical beams propagating through the atmosphere, FSOC provides energy-efficient, secure links with high data transmission capacity at potentially low cost (Kaushal and Kaddoum 2017). Applications range from short-distance terrestrial links to satellite-to-ground communications, making FSOC relevant for both last-mile and global connectivity. The atmospheric channel, however, introduces challenges that are absent or less severe in fiber-optic or radio-frequency communication. In addition to clouds, precipitation, and aerosol scattering, turbulent fluctuations of the atmospheric refractive index form a major source of disturbance for propagating optical beams (Jahid *et al.* 2022). These fluctuations, commonly known as optical turbulence (OT), cause beam spreading, wavefront distortion, and intensity fluctuations (scintillation), all of which can significantly reduce achievable data rates or even interrupt communication links (Kaushal and Kaddoum 2017). The degradation of a received Gaussian beam that propagated through different levels of OT strength is visualized in [fig. 1.1](#). In calm OT conditions (middle panel), the beam arrives almost unperturbed in its Gaussian form, exhibiting a clear intensity maximum centered within the receiver aperture (thin circle). As turbulence strength increases, the beam breaks up into speckles (left), and the intensity maximum begins to move away from the center of the aperture. In strong turbulence (right), the beam can fully break up and move almost fully out of the aperture, likely leading to an interruption of the FSOC link. Ground-based optical astronomy is affected by the same problems, as OT limits angular resolution, making images look blurry (Hardy 1998). Adaptive optics can mitigate some OT distortions (Kaushal and Kaddoum 2017), but the ability to quantify and predict OT remains essential for the design and reliable operation of optical systems.

The strength of OT is commonly quantified by the refractive index structure parameter C_n^2 ($\text{m}^{-2/3}$), which characterizes the intensity of refractive index fluctuations in the inertial subrange of turbulence. Physically, these fluctuations are linked to turbulent fluctuations of temperature and humidity in the atmosphere, which are modulated by wind shear and buoyancy across a wide range of scales

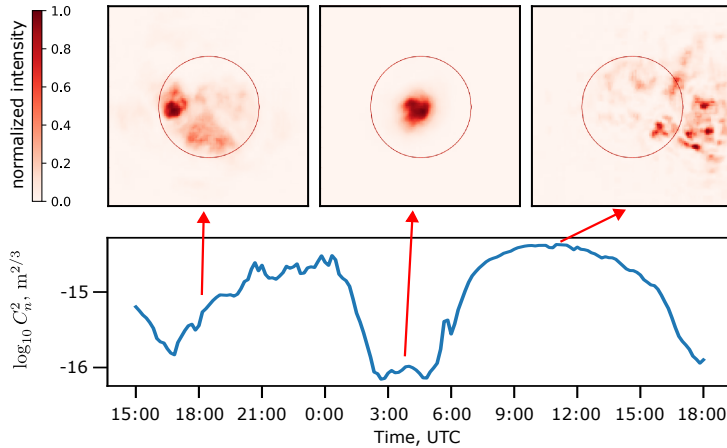


Figure 1.1.: Different degrees of distortion of a laser beam due to different optical turbulence conditions (C_n^2) along the propagation path. Speckle patterns are derived from wave optics simulation courtesy of Rudolf Saathof.

(Stull 1988). Integrating C_n^2 along a propagation path yields the Fried parameter r_0 , which describes wavefront coherence, and the Rytov variance σ_R^2 , which measures intensity fluctuations at the receiver (Andrews and Phillips 2005). Together, these parameters determine the impact of OT on a propagating optical beam and feed into link budget calculations for FSO systems. Reliable estimates of C_n^2 are therefore essential for computing turbulence-induced link losses, assessing link availability, and dimensioning communication systems.

Obtaining C_n^2 data is, however, challenging because turbulence is strongly dependent on local topography and the ever-changing meteorological conditions (Kaushal and Kaddoum 2017). For example, the C_n^2 time series in fig. 1.1 changes over almost two orders of magnitude throughout the day, resulting in the significantly different displayed distortions of the beam. This site dependence necessitates either observing or modeling C_n^2 for the specific location and conditions of interest. The observation and modeling techniques relevant to this dissertation are briefly introduced below.

1.1. DIRECT AND INDIRECT C_n^2 OBSERVATIONS

Optical turbulence strength is measured either directly with a scintillometer or indirectly with eddy covariance systems. A scintillometer consists of a transmitter and receiver separated by a horizontal distance of 0.1–10 km. The transmitter emits a beam of light at one end of a horizontal path, and the receiver receives it at the other. C_n^2 is then derived from measured fluctuations of received optical power caused by the OT along the path (Wang *et al.* 1978; Beyrich *et al.* 2021), so that $C_n^2 \sim \sigma_{\log(I)}^2$, where $\sigma_{\log(I)}^2$ denotes the variance of received log-intensity. An example of a scintillometer at the Cabauw Experimental Site for Atmospheric

Research (CESAR) in the Netherlands is shown in fig. 1.2a.

Indirect observations are obtained from eddy covariance systems, which typically consist of a 3D ultrasonic anemometer and a fast-response gas analyzer (Foken 2017) (cf. fig. 1.2b). The sonic anemometer measures high-frequency signals of 3D wind speed and temperature (sampling at rates exceeding 10Hz), while the gas analyzer measures humidity and other quantities. As refractive index fluctuations are modulated by temperature and humidity fluctuations, C_n^2 can be estimated from the strength of these fluctuations. More specifically, C_n^2 can be expressed as (Wesely 1976; Moene 2003)

$$C_n^2 = \frac{A_T^2}{\bar{T}^2} C_T^2 + 2 \frac{A_T A_q}{\bar{T} \bar{q}} C_{Tq} + \frac{A_q^2}{\bar{q}^2} C_q^2, \quad (1.1)$$

where \bar{T} (K) and \bar{q} (kg kg^{-1}) are the mean temperature and specific humidity. Following Wesely (1976), A_T and A_q are the refractivity sensitivity coefficients for temperature and specific humidity, and C_T^2 , C_q^2 , and C_{Tq} are the structure parameters of temperature, humidity, and their cross-correlation, respectively. As the sonic temperature signal already contains a humidity contribution (Kaimal and Gaynor 1991), we assume in this dissertation that $C_{T_s}^2$ based on the sonic temperature signal includes the moisture effect so that $C_n^2 \sim C_{T_s}^2$.

The structure function parameter of the temperature signal, C_T^2 ($\text{K}^2 \text{m}^{-2/3}$), is obtained from the second-order structure function given as

$$S_T^2(r) = \langle (T(x) - T(x+r))^2 \rangle. \quad (1.2)$$

The structure function describes how the ensemble average (denoted by $\langle \cdot \rangle$) of the squared temperature increments changes depending on a separation distance r (m). If r lies within the inertial range of turbulence, S_T^2 scales with a power law as

$$S_T^2(r) = C_T^2 r^{2/3}. \quad (1.3)$$

As turbulence conditions change throughout the day, C_T^2 is typically determined from short segments (5–10 min) of the temperature time series.

An example of a temperature signal segment and its corresponding structure function is given in the right branch of fig. 1.3. Taylor's hypothesis of frozen turbulence is employed to convert the sonic temperature time series into a spatial series. The hypothesis assumes that turbulent eddies are advected past the sensor by the mean wind speed \bar{M} (m s^{-1}) without significant evolution. This allows us to convert the temporal sampling frequency f_s (Hz) into spatial increments $\Delta r \approx \bar{M} / f_s$. Considering eq. (1.3), C_T^2 is found by fitting a linear function of slope 2/3 to the structure function in log-log space (cf. fig. 1.3). Alternatively, the power spectral density of the temperature signal can yield the same C_T^2 (cf. fig. 1.3, left branch) by fitting a $-5/3$ slope in log-log space. For more details on both approaches and alternative methods, see Beason *et al.* (2024). As the structure function approach is less noisy than the spectral approach and is directly linked to the definition of C_T^2 , it is the approach used to determine C_T^2 throughout this dissertation.

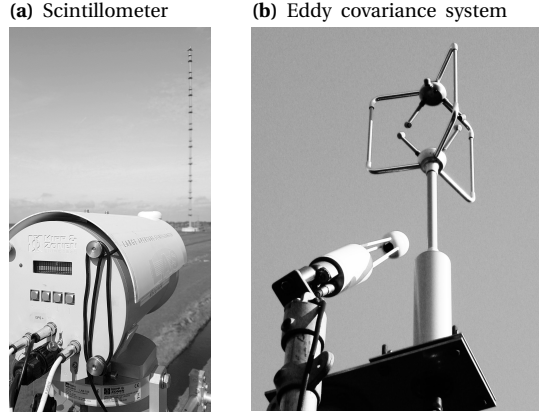


Figure 1.2.: Examples of instruments used to observe optical turbulence (C_n^2) either (a) directly or (b) indirectly.

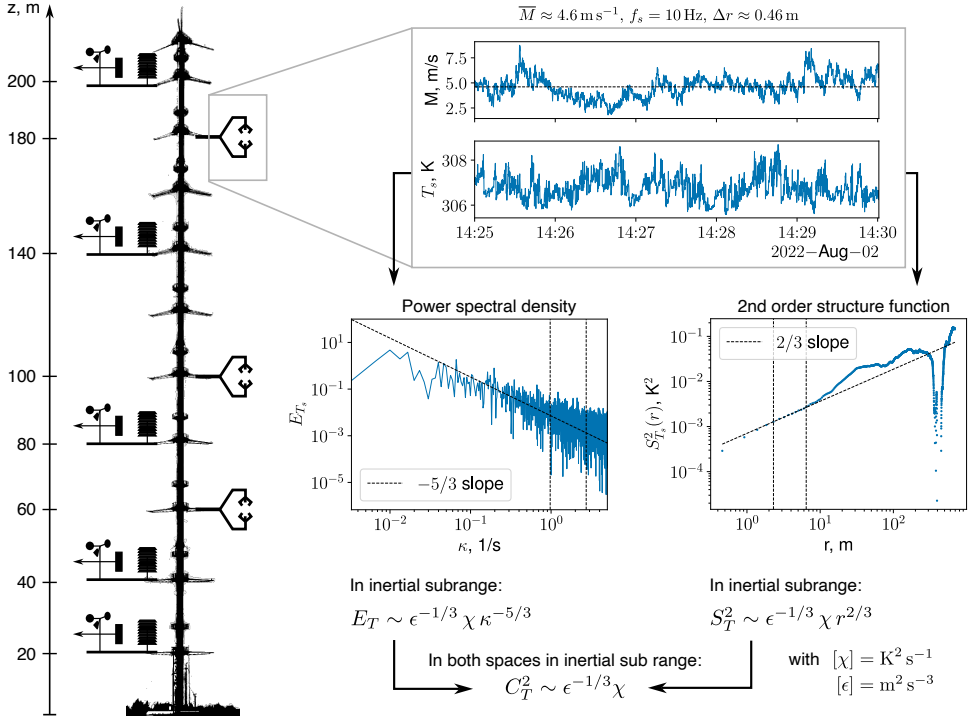


Figure 1.3.: Example of C_T^2 estimation from 5 min of observed sonic temperature using the power spectral density approach (left) or the structure function approach (right). The mean wind speed is used to employ Taylor's hypothesis of frozen turbulence.

1.2. OPTICAL TURBULENCE PARAMETERIZATIONS

The two observation strategies discussed above are reliable but require careful data processing and dedicated instruments. Consequently, long-term field campaigns to characterize OT conditions at candidate sites (see, e.g., Schöck *et al.* (2009)) become time-consuming and resource-intensive. Alternatively, C_n^2 can be estimated using parameterizations, which relate C_n^2 to more widely available standard meteorological variables that capture the processes driving OT. A range of such parameterizations exists, differing in input requirements, underlying assumptions, and applicability.

Common approaches are based on Monin-Obukhov similarity theory (MOST) (Monin and Obukhov 1954), relating C_n^2 to turbulent fluxes or vertical gradients of mean meteorological variables (see, e.g., Savage (2009) for an overview). For example, the flux-based parameterization of Wyngaard *et al.* (1971) relates the near-surface values of the kinematic sensible heat flux ($\overline{w'\theta'}$ in K m s^{-1} , proxy for buoyancy) and the friction velocity (u_* in m s^{-1} , proxy for wind shear) to C_n^2 as

$$C_n^2 \sim \left(-\overline{w'\theta'}/u_* \right)^2 z^{-2/3} g(\zeta), \quad (1.4)$$

where z (m) is the height above the surface and $g(\zeta)$ is a stability-dependent similarity function of the dimensionless stability parameter ζ .

However, MOST-based parameterizations are limited to the atmospheric surface layer, whereas space-to-ground links require C_n^2 estimates for the full atmospheric column. A more sophisticated physics-based parameterization, applicable to the full atmosphere, is the variance-based approach of He and Basu (2015). This technique, however, requires higher-order turbulence statistics as input, which are typically not readily available from standard meteorological observations. Also, high-altitude radiosonde observations are generally limited to 1 or 2 ascents per day at selected locations. That is why very simple analytical models, such as the Hufnagel-Valley (HV) model (Hufnagel 1974; Valley 1980; Ulrich 1988), are commonly employed to obtain full vertical C_n^2 profiles, but the accuracy of such estimates is very limited because the HV model barely accounts for local meteorology.

1.2.1. MESOSCALE MODELLING

Numerical weather prediction (NWP) offers a physics-based route to overcome these limitations. Mesoscale models such as the Weather Research and Forecasting (WRF) model (Skamarock *et al.* 2021) simulate the atmospheric state and the inputs required for the C_n^2 parameterizations. Instead of applying the parameterizations to observed meteorological data to obtain C_n^2 , the NWP model output is used to estimate C_n^2 in a post-processing step (Bougeault *et al.* 1995; Masciadri *et al.* 1999). The use of parameterizations remains necessary because mesoscale models operate at spatial resolutions coarser than the turbulence scale. The NWP-based approach yields spatially and temporally continuous C_n^2 fields that account for local meteorology and topography, and previous studies have demonstrated good agreement with a variety of observations (e.g., Masciadri *et al.* (1999), Frehlich *et al.* (2010), Cheinet *et al.* (2011), Basu *et al.* (2020), Ullwer *et al.* (2019), and Rafalimanana *et al.* (2022)). Yet the approach also has well-known limitations. The simulated

meteorology depends on the selected physics parameterizations, and in particular, the planetary boundary layer (PBL) scheme can significantly affect the resulting turbulence statistics (Pierzyna *et al.* 2023a). The choice of C_n^2 parameterization introduces further sensitivity. Running mesoscale models is also computationally expensive, which limits the scalability of this approach to long-term climatological studies or large numbers of sites. Moreover, complex environments such as urban areas are difficult to represent at mesoscale resolution, potentially limiting the accuracy of C_n^2 estimates for FSOC links in cities.

1.2.2. MACHINE LEARNING MODELLING

In recent years, machine learning (ML) has emerged as a complementary data-driven approach to estimating C_n^2 (Wang and Basu 2016; Jellen *et al.* 2021; Bolbasova *et al.* 2021; Su *et al.* 2021). ML models can learn complex, non-linear relationships directly from data without requiring the simplifying assumptions inherent to physics-based parameterizations. Once trained, they are computationally inexpensive to evaluate. These properties make ML attractive when computational cost or model complexity limit the use of NWP-based methods. At the same time, purely data-driven models may not generalize well to unseen conditions and can be difficult to interpret physically. Maintaining the physical consistency and interpretability of ML models is therefore a key challenge that must be addressed to enable their operational use.

1.3. SCOPE OF THIS DISSERTATION

Reliable C_n^2 estimates are essential for the design and operation of FSOC systems, yet obtaining them remains challenging due to the strong dependence of optical turbulence on local meteorology and topography. This dissertation addresses this challenge by pursuing two complementary modelling avenues. The first part establishes a baseline using numerical mesoscale modeling. The second part explores whether machine learning can overcome the limitations of physics-based approaches.

As discussed above, multiple C_n^2 parameterization approaches exist, yet they yield different results, and their relative merits remain unclear. The first part of this dissertation, therefore, presents a systematic benchmarking of the main parameterization classes, applied to both observed and NWP-simulated meteorological variables, and validated against scintillometer observations (chapter 2). However, C_n^2 estimation errors alone are difficult to interpret in an engineering context. The NWP-based C_n^2 estimation is therefore expanded into an end-to-end framework that translates C_n^2 estimates into FSOC performance metrics, such as signal-to-noise ratio and achievable data rates (chapter 3).

With this traditional modeling baseline in place, the second part turns to machine learning to improve and extend C_n^2 estimation beyond the capabilities of existing parameterizations. Here, C_n^2 derived from sonic anemometer observations serves as the primary data source. First, a gradient-boosting-based methodology is developed that combines near-surface C_n^2 observations from multiple sites with global reanalysis data to produce multi-year C_n^2 climatologies, thereby reducing the need for long

and costly measurement campaigns (chapter 4). The resulting models, however, are limited to near-surface turbulence and lack vertical extension. To incorporate vertical dependencies within the surface layer, a physics-informed ML framework is introduced that leverages multi-level sonic anemometer observations at a single site. The framework revisits the classical concept of dimensional analysis to derive an interpretable, data-driven alternative to MOST-based parameterizations (chapter 5). Finally, satellite-to-ground links require C_n^2 estimates for the full atmospheric column, not just the surface layer. Since C_n^2 observations at higher altitudes are scarce, a deep learning model is proposed that instead learns to emulate NWP-based C_n^2 estimates, following the framework established in the first part of this dissertation (chapter 6).

The key findings of the dissertation are summarized and discussed in chapter 7, where we also reflect on the broader perspective of mesoscale and ML-based C_n^2 modeling, and their interplay in future research.

I

Numerical mesoscale modeling – the traditional approach

2

Intercomparison of flux, gradient, and variance-based optical turbulence (C_n^2) parameterizations

For free-space optical communication (FSOC) or ground-based optical astronomy, abundant data of optical turbulence strength (C_n^2) is imperative but typically scarce. Turbulence conditions are strongly site-dependent, so their accurate quantification requires in-situ measurements or numerical weather simulations. If C_n^2 is not measured directly, e.g., with a scintillometer, C_n^2 parameterizations must be utilized to estimate C_n^2 from meteorological observations or model output. Even though various such parameterizations exist in the literature, their relative performance is unknown. We fill this knowledge gap by performing a systematic three-way comparison of a flux-based, a gradient-based, and a variance-based parameterization. Each parameterization is applied to both observed and simulated meteorological variables, and the resulting C_n^2 estimates are compared against observed C_n^2 from two scintillometers. The variance-based parameterization yields the overall best performance, and unlike other approaches, its application is not limited to the lowest part of the atmospheric boundary layer, i.e. the surface layer. We also show that C_n^2 estimated from the output of the Weather Research and Forecasting (WRF) model aligns well with observations, highlighting the value of mesoscale models for optical turbulence modeling.

This chapter is published as: M. Pierzyna *et al.* (2024). “Intercomparison of Flux, Gradient, and Variance-Based Optical Turbulence (C_n^2) Parameterizations”. In: *Applied Optics* 63.16. DOI: 10.1364/AO.519942

2.1. INTRODUCTION

Free-space optical communication (FSOC) is considered a key technology to support the increasing needs of our densely-connected, data-heavy world by providing energy-efficient, secure links with high-data transmission capacity at potentially low cost. In contrast to traditional radio frequency communication, FSOC transmits data with an optical beam that propagates through the atmosphere. This propagating beam is disturbed by various atmospheric phenomena, such as clouds, molecular and aerosol scattering, and fluctuations of the atmospheric refractive index, the so-called optical turbulence (OT) (Jahid *et al.* 2022). Data of these OT conditions is highly relevant for designing and deploying reliable, high-performance FSOC links, but its availability is typically scarce. Therefore, there is an urgent need to quantify the OT conditions well (Kaushal and Kaddoum 2017).

Since turbulence strongly depends on the local topography and the ever-changing meteorological conditions, site-specific estimations of the turbulence conditions are complex. As a result, the performance quantification of an FSOC link requires either measuring or modeling the OT strength (C_n^2) for the time of day and the site of (envisioned) operation. Such C_n^2 measurements are performed directly using scintillometers or indirectly through standard meteorological instruments in conjunction with a C_n^2 parameterization. While scintillometers directly measure the intensity fluctuations of light propagating through the turbulent atmosphere (Beyrich *et al.* 2021), the parameterizations of the indirect approach link C_n^2 to variables of the processes causing the OT, i.e. wind shear and buoyancy (Wyngaard *et al.* 1971). Similarly, these parameterizations also allow the estimation of C_n^2 from atmospheric mesoscale simulations in a post-processing step (Bougeault *et al.* 1995; Masciadri *et al.* 1999). In practice, direct scintillometer measurements are often not available, so the parameterizations are commonly used on observed (see, e.g., Wyngaard *et al.* (1971), Andreas (1988), Hutt (1999), and Basu (2015)) and simulated meteorological data (see, e.g., Masciadri *et al.* (1999), Frehlich *et al.* (2010), Cheinet *et al.* (2011), Basu *et al.* (2020), Ullwer *et al.* (2019), and Rafalimanana *et al.* (2022)), and a multitude of such C_n^2 parameterizations exists (see Savage (2009) and Kooijmans and Hartogensis (2016) for extensive overviews). However, their accuracy compared to scintillometer observations is unclear a-priori, and studies in literature often do not motivate their choice of one parameterization over another.

In the present work, we aim to address this issue by intercomparing the performance of C_n^2 parameterizations from three main classes. Multiple test cases representing different seasonal conditions are selected for which scintillometer-based C_n^2 time series are available. These observations serve as the best possible ground truth. The scintillometers are deployed at the Cabauw Experimental Site of Atmospheric Research (CESAR) site (Bosveld *et al.* 2020) in the Netherlands, where a recent FSOC test campaign (Broekens *et al.* 2023a) was conducted. That site is also equipped with reliable and well-maintained instruments from which we obtain standard meteorological measurements collocated and concurrent with the scintillometer data. Additionally, we run a mesoscale simulation for each test case to reproduce the meteorology at the time and location of the scintillometer observations. These observed and simulated meteorological variables are inputs to

the different C_n^2 parameterizations, and the resulting C_n^2 estimates are compared to the scintillometer measurements for performance assessment. Having two sources of meteorological data can help us to differentiate errors coming from the parameterizations themselves from errors due to limitations of instrumentation or simulation. It also enables us to systematically evaluate the feasibility of mesoscale simulations for C_n^2 estimation in an FSOC context.

2.2. METHODOLOGY

We need diverse and reliable underlying datasets to conduct a comprehensive intercomparison of C_n^2 parameterizations. As a baseline, we utilize C_n^2 observations from two scintillometers mounted at two different heights (cf. section 2.2.2), which enable us to assess height-dependencies in the performance of the C_n^2 parameterizations. Two high-quality meteorological datasets – concurrent and collocated with the scintillometer measurements – serve as parameterization inputs. One dataset is obtained experimentally at the CESAR site (cf. section 2.2.3), while the other one is a mesoscale hindcast of the same time and location produced by the Weather Research and Forecasting (WRF) (Skamarock *et al.* 2021) model (cf. section 2.2.4). Three C_n^2 parameterizations from different classes are applied to the input data for an intercomparison between estimated and observed C_n^2 as visualized in fig. 2.1. The specific parameterizations representing these three classes – flux-based, gradient-based, and variance-based – are presented in section 2.2.1.

Our methodology is applied to three test cases covering spring (18 – 22 April 2022, 4 days), summer (27 – 29 July 2022, 2 days), and autumn (8 – 10 October 2022, 2 days). By spreading the test cases throughout the year, we aim to assess whether seasonal dependencies exist in the performance of the different C_n^2 parameterizations. All test cases are subject to calm synoptic conditions and a mostly clear sky at the CESAR site. Selecting such favorable meteorological conditions is essential because the present study aims to assess the performance of the C_n^2 parameterizations, not the performance of the deployed instruments or the mesoscale models for various impactful weather events. Precipitating events, for example, are challenging to simulate, and scintillometers cannot conduct measurements when visibility is low, for instance, during heavy rain or fog. Since the Dutch weather during winter is cloudy, rainy, and/or foggy, no suitable winter test case could be identified for the CESAR site.

2.2.1. C_n^2 ESTIMATION FROM OBSERVED AND SIMULATED METEOROLOGY

The strength of OT is quantified through C_n^2 , the refractive index structure function parameter. C_n^2 describes the refractive index fluctuations caused by density fluctuations due to temperature and humidity fluctuations. If the parameter is not measured directly with a scintillometer, it can, therefore, be derived from C_T^2 , the temperature structure function parameter, utilizing the Gladstone relationship (Gladstone and Dale 1863):

$$C_n^2 = \left(A \frac{P}{T^2} \right)^2 \left(1 + \frac{0.03}{\beta} \right)^2 C_T^2. \quad (2.1)$$

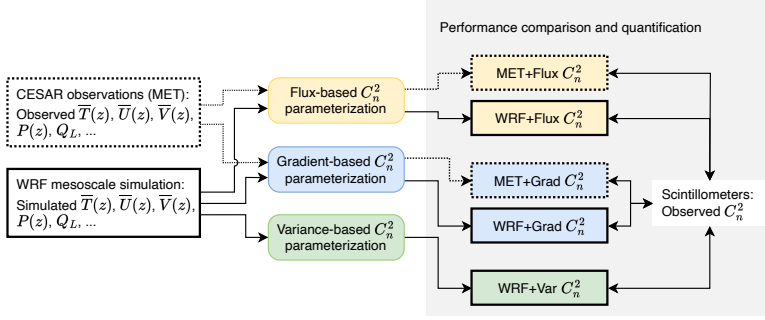


Figure 2.1.: Schematic of the three-way performance comparison and quantification applied in this study. For each test case, C_n^2 is estimated with different parameterizations (round boxes) from observed (MET) and simulated (WRF) meteorology and compared to scintillometer observations.

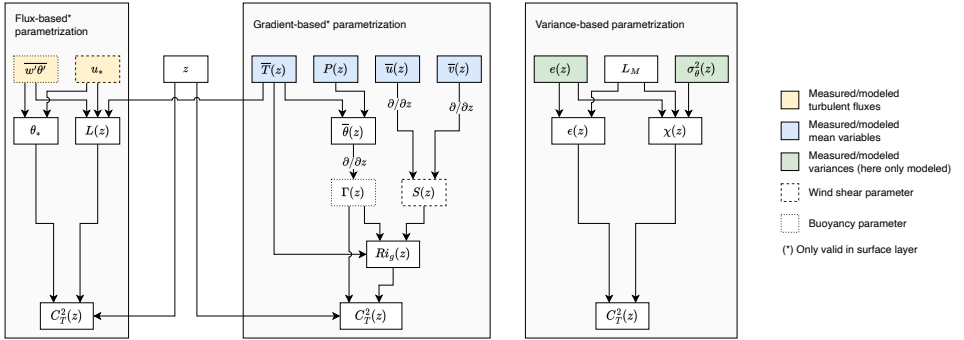


Figure 2.2.: Dependency of C_T^2 on measured or modeled input parameters for the three types of C_T^2 parameterizations.

Here, $A \approx 7.9 \times 10^{-5} \text{ KhPa}^{-1}$ is almost constant for optical wavelengths (Andrews and Phillips 2005), P is the station pressure in hPa, T is the air temperature in K, and $\beta = Q_H/Q_L$ is the Bowen ratio. By balancing the dynamic sensible heat flux Q_H with the latent heat flux Q_L , the Bowen ratio term accounts for the influence of moisture on C_n^2 (Wesely 1976). During neutral atmospheric conditions (sunrise/sunset), Q_H and β change signs twice a day, leading to unrealistically dominant moisture correction factors. This mathematical singularity is avoided by truncating β to ± 0.1 for $|\beta| < 0.1$. We also note that the $(1 + 0.03/\beta)^2$ moisture correction can introduce an overcorrection by a factor of ~ 4 around these neutral conditions due to the assumed correlation of moisture flux and heat flux ($|R_{Tq}| = 1$) in the derivation of the correction. The interested reader is referred to Moene (2003) for a detailed discussion, but for FSOc applications, the correction presented in eq. (2.1) is considered adequate. That is because neutral conditions do not prevail long and are subject to very weak turbulence, so correction errors have low practical relevance. Also, a more accurate correction requires additional atmospheric water

vapor measurements, which is impractical in an FSOC context.

Due to the discussed relation between C_n^2 and C_T^2 , the methods presented in the following only parameterize C_T^2 , which is defined as the coefficient of the second-order structure function of temperature T in the inertial range as

$$C_T^2 = \langle [T(x) - T(x+r)]^2 \rangle / r^{2/3} \quad (2.2)$$

where $\langle \square \rangle$ denotes the ensemble average. The structure-function approach requires a high-frequency temperature signal (> 10 Hz) to capture the fast-evolving turbulence, which is often infeasible to obtain through measurements or simulations. Therefore, various parameterizations were developed to estimate C_T^2 from more routinely available data. These parameterizations can be broadly classified by the type of information they use: flux-based, gradient-based, or variance-based methods, as shown in fig. 2.2.

All parameterizations aim at capturing the two processes modulating turbulence – wind shear and buoyancy (Stull 1988) – but utilize different types of physical variables (cf. dashed and dotted boxes in fig. 2.2). The difference between the variables is whether they represent the slowly varying mean or the fast-changing fluctuations of a turbulent flow field. For example, the space and time-dependent potential temperature field $\theta(t, x, y, z)$ can be decomposed into mean potential temperature $\bar{\theta}$ and turbulent fluctuations θ' such that $\theta(t, x, y, z) = \bar{\theta}(x, y, z) + \theta'(t, x, y, z)$ (Stull 1988). Similarly, the wind components u , v , and w can be decomposed into the mean field $(\bar{u}, \bar{v}, \bar{w})$ and fluctuations (u', v', w') . While gradient-based C_n^2 methods only use vertical gradients of the mean quantities, the flux and variance-based approaches require knowledge about turbulent fluxes (e.g., a momentum flux $\overline{u'w'}$) or turbulent variances (e.g., potential temperature variance $\overline{\theta'\theta'} = \sigma_\theta^2$) formed from the fluctuation components of the fields. These mean quantities, fluxes, and variances can be observed experimentally or simulated using mesoscale models, such as the WRF model.

Flux-based method One approach of linking buoyancy and wind shear to the temperature fluctuations, i.e., C_T^2 , is by utilizing the sensible heat flux $\overline{w'\theta'}$ and the friction velocity u_* , respectively. The latter has the unit ms^{-1} and is called a velocity but is a function of two components of the turbulent vertical momentum flux: $u_* = \left(\overline{u'w'^2} + \overline{v'w'^2} \right)^{1/4}$. Wyngaard *et al.* (1971) proposed a flux-based parameterization (W71-flux) based on these quantities and Monin-Obukhov similarity theory (MOST) given as

$$C_T^2 = T_*^2 z^{-2/3} g(\zeta) \quad (2.3)$$

where $T_* = -\overline{w'\theta'}/u_*$ is the temperature scale. The similarity function g

$$\begin{aligned} g(\zeta) &= 4.9(1 - 6.1\zeta)^{-2/3}, \zeta < 0 \\ g(\zeta) &= 4.9(1 + 2.2\zeta^{2/3}), \zeta \geq 0 \end{aligned} \quad (2.4)$$

depends on the stability parameter $\zeta = z/L$ with Obukhov length $L = -u_*^3 \bar{T} / (\kappa g \overline{w'\theta'})$ and height above ground z . In L , $g = 9.81 \text{ m s}^{-2}$ is the gravitational acceleration of

Earth, $\kappa = 0.4$ is the Von Kármán constant, \bar{T} is the mean air temperature in K, and $\overline{w'\theta'}$ is the kinematic sensible heat flux in mKs^{-1} . The coefficients of eq. (2.4) are empirically determined and vary in literature (Kooijmans and Hartogensis 2016). We use the coefficients presented by Andreas (1988), who adjusted the original W71-flux coefficients for $\kappa = 0.4$. Friction velocity and sensible heat flux are surface variables available from the WRF model and a sonic anemometer at the CESAR site (Bosveld *et al.* 2020).

Gradient-based method Measuring fluxes requires sonic anemometers that record high-frequency temperature and velocity signals. If such instruments are unavailable, a set of vertically separated standard thermometers and anemometers can be used for C_T^2 estimation in conjunction with a gradient-based method. These slow-response instruments yield the temporal evolution of the mean temperature and mean wind profiles. The vertical gradients of (potential) temperature and horizontal wind represent the buoyancy and shear contributions to turbulence, respectively. Wyngaard *et al.* (1971) also presented a gradient-based C_T^2 parameterization:

$$C_T^2 = z^{4/3} \Gamma^2 f(\text{Ri}_g). \quad (2.5)$$

Here, f is a similarity function depending on the gradient Richardson number $\text{Ri}_g = (g/\bar{\theta})(\Gamma/S^2)$, and the mean potential temperature gradient $\Gamma = \partial\bar{\theta}/\partial z$. The stability parameter Ri_g is a ratio balancing buoyancy, captured through Γ , and the mean wind shear $S = \sqrt{(\partial\bar{u}/\partial z)^2 + (\partial\bar{v}/\partial z)^2}$. Thus, Γ and S fulfill the same role of including buoyancy and shear modulation of the turbulence into the gradient-based formulation as u_* and $\overline{w'\theta'}$ do for W71-flux. The empirical similarity function $f(\text{Ri}_g)$ in eq. (2.5) is tabulated by Wyngaard *et al.* (1971), but He and Basu (2016b) presented an improvement of f for stably stratified conditions ($\text{Ri}_g > 0$) based on direct numerical simulations given as

$$f(\text{Ri}_g) = 0.05 + 1.02 \exp(-14.49\text{Ri}_g), \text{Ri}_g > 0. \quad (2.6)$$

We use this improved formulation for stable conditions and a spline interpolation of the original W71 values only for unstable atmospheric conditions ($\text{Ri}_g \leq 0$). This combined approach is abbreviated as W71/HB16-grad for this work.

All required mean variables (cf. fig. 2.2) are routinely available both as time-varying 3D fields from the WRF model and as time series observed at multiple heights at the CESAR site. The gradients are typically strongest close to the ground (Stull 1988), so we aim to accurately compute the lowest gradients numerically by adding the following information: For S , we assume a logarithmic wind profile close to the surface so that $u(z_0) = v(z_0) = 0$ with surface roughness length z_0 (known from WRF or experiments (Verkaik and Holtslag 2007)). For Γ , we employ the 2m-potential temperature $\bar{\theta}_2$ provided by WRF's surface layer scheme or the 2m thermometer deployed at CESAR. Section 2.A explains these technicalities in more detail.

Surface layer assumption Like the W71-flux parameterization, the gradient-based W71/HB16-grad approach is based on MOST. As a result, both approaches inherit

the MOST limitation that they are only applicable within the atmospheric surface layer (SL), which is defined as 10% of the atmospheric planetary boundary layer (PBL) height (Stull 1988). MOST assumes that fluxes within the SL are invariant with height, which is not valid anymore for the upper part of the boundary layer. The height of the PBL and, consequently, the SL changes throughout the day. Typically, the SL is shallow during the night (~ 10 m) and deep during the day (~ 100 m) (Stull 1988), so the height until which the gradient and the flux-based methods can be applied also changes.

Variance-based method A physics-based parameterization that is valid beyond the SL follows from dimensional analysis using Corrsin's (1951) (Corrsin 1951) expression and Kolmogorov's (Kolmogorov 1941) universal spectrum hypothesis as

$$C_T^2 = 3.2\chi\epsilon^{-1/3} \quad (2.7)$$

where ϵ is the turbulent energy dissipation rate and χ the destruction rate of potential temperature variance σ_θ^2 . Bougeault *et al.* (1995) and Masciadri *et al.* (1999) utilized mesoscale models with a turbulence closure scheme (Bougeault and Lacarrere 1989) that has a prognostic equation for the turbulent kinetic energy (TKE) e , which allowed the parameterization of ϵ and χ to ultimately obtain C_T^2 according to eq. (2.7). He and Basu (2015) utilized a different parameterization (Mellor 1973; Mellor and Yamada 1982) based on e and σ_θ^2 in conjunction with Eq. (2.7):

$$\epsilon = \frac{(2e)^{3/2}}{B_1 L_M}, \quad \chi = \frac{(2e)^{1/2}}{B_2 L_M} \sigma_\theta^2. \quad (2.8)$$

Here, L_M is the master length scale combining the characteristic scales of the surface, turbulence, and buoyancy, and the coefficients B_1 and B_2 were found to be equal to 24 and 15, respectively, from numerical simulations. He and Basu (2015) used the high-order MYNN 2.5 turbulence closure scheme (Nakanishi and Niino 2006; Nakanishi and Niino 2009) with a prognostic output of e and diagnosed σ_θ^2 and L_M . Physically, the TKE per unit mass is half the sum of the velocity field variances: $e = (\sigma_u^2 + \sigma_v^2 + \sigma_w^2)/2$. This variance-based approach based on He and Basu (HB15-var), combined with the MYNN 2.5 turbulence closure scheme in WRF, is also used in the present work. While the variances e and σ_θ^2 and the length scale L_M are available from the WRF model output, they are not contained in the CESAR data product available for this study. Additionally, L_M as used in MYNN-2.5 requires evaluating a vertical integral of TKE from the surface until the boundary layer height (Olson *et al.* 2019), which is difficult to obtain experimentally. Consequently, we apply HB15-var only to the simulated meteorology.

2.2.2. SCINTILLOMETER OBSERVATIONS OF C_n^2 AT THE CESAR SITE

The accuracy of the estimated C_n^2 values is assessed by comparing them against C_n^2 measurements from two scintillometers deployed at the CESAR site. The instruments are mounted at two heights: ~ 10 m and ~ 80 m. Therefore, the lower instrument is almost always situated in the SL so that it serves as a reference for all three

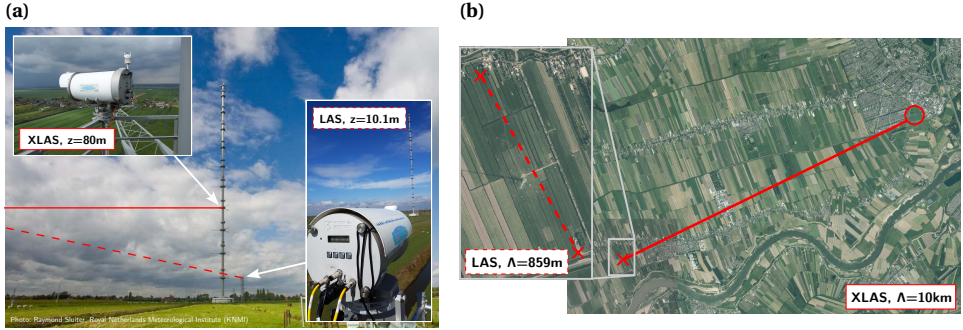


Figure 2.3.: Vertical placement and approximate measurement paths of the two scintillometers – LAS (dashed line) and XLAS (solid line) – at the CESAR site as seen (a) from the Cabauw field and (b) on an aerial photograph.

parameterizations. That is not the case for the higher instrument because the SL height regularly falls below the mounting height, for example, during the night (all seasons) and partially during the day (non-summer seasons). Consequently, we only utilize that higher scintillometer as a reference for the variance-based method.

A scintillometer consists of a transmitter and receiver displaced over a horizontal path between 0.1 – 10 km. The transmitter emits a beam of light at one end of a horizontal path, and the receiver receives it at the other. C_n^2 is then derived from measured fluctuations of received optical power caused by the OT along the path (Wang *et al.* 1978):

$$C_n^2 = 1.12D^{7/3} \Lambda^{-3} \sigma_{\ln(I)}^2, \quad (2.9)$$

where D is the scintillometer aperture size, Λ is the path length between transmitter and receiver, and $\sigma_{\ln(I)}^2$ is the log-variance statistic of the intensity fluctuations, I . The resulting C_n^2 represents a path-weighted average, where a bell-shaped path weighting function places the highest weight on turbulence at the center of the path (Beyrich *et al.* 2021); this weighting is discussed further in [section 2.2.4](#).

The two scintillometers are part of the Dutch Ruisdael observatory infrastructure and measure along two almost perpendicular paths as indicated by the solid and dashed lines in [Fig. 2.3](#). In particular, there are

1. a Kipp&Zonen MKII Large Aperture Scintillometer (LAS, $D=0.147$ m) measuring along a $\Lambda=859$ m path (dashed line) at 10.1 m height above ground, and
2. a Kipp&Zonen MKI eXtra Large Aperture Scintillometer (XLAS, $D=0.32$ m), which has been upgraded to MKII specifications, measuring along a $\Lambda=10$ km path (solid line) at ~ 80 m height.

The panels in [fig. 2.3](#) show the approximate (a) vertical and (b) horizontal scintillometer measurement paths. Both scintillometers operate at a wavelength of $\lambda=850$ nm and have a finite aperture to ensure they are saturation resistant over longer path lengths (Wang *et al.* 1978). In calculating $\sigma_{\ln(I)}^2$ from the raw

1 kHz recorded data series of I , a high pass filter of 30 s was applied to filter out absorption fluctuations, and intervals with less than 20% of available data were set to a dummy value. The intensity data are aggregated into 30 min bins from which C_n^2 is computed. In a quality control step, C_n^2 measurements are classified as unreliable and discarded if the received mean carrier power drops below a setup-dependent threshold. This happens, for example, during heavy rainfall or fog at the CESAR site, resulting in severe attenuation or loss of the scintillometer signal.

2.2.3. METEOROLOGICAL OBSERVATIONS WITHIN THE SURFACE LAYER

Several meteorological variables serve as input for the different parameterizations to estimate C_n^2 at the 10 m level (surface layer) and compare it to the corresponding scintillometer observations. These measurements are extracted as 10-min averages from the instruments mounted on and around the 200 m high Cabauw tower of the CESAR site (cf. fig. 2.3(a)). In particular, we utilize the profiles of air temperature $\overline{T}(z)$ and horizontal wind measured at 2 m (T only), 10 m, and 20 m, above ground, respectively, to estimate $\Gamma(z)$ and $S(z)$ at the 10 m level. The horizontal wind at CESAR is measured with cup anemometers and wind vanes (Bosveld *et al.* 2020) from which we obtain the horizontal components $\overline{u}(z)$ and $\overline{v}(z)$ for the computation of $S(z)$. Details about the gradient computation are presented in section 2.A. From those gradients, C_T^2 at 10 m height is estimated following W71/HB16-grad. The turbulent fluxes u_* and $\overline{w'\theta'}$ for the corresponding W71-flux C_T^2 estimates are obtained from an eddy covariance system consisting of an open-path $\text{H}_2\text{O}/\text{CO}_2$ sensor (LICOR-7500) and a sonic anemometer (Gill-R50) mounted at 3 m height (Bosveld *et al.* 2020). Following the surface layer assumption, the fluxes are considered to be height-invariant and also valid at 10 m height. As mentioned before, the variances required for HB15-var are not routinely available from CESAR and the experimental determination of the length scale L_M is challenging. Consequently, only the gradient-based and the flux-based methods are utilized in conjunction with the observed meteorology.

Converting C_T^2 to C_n^2 following eq. (2.1) is the same for all three parameterizations and requires the Bowen ratio β and the 10 m atmospheric pressure. The latter, P_{10} , is computed from the reference sea-level pressure P_0 , utilizing the hypsometric equation and the 10 m temperature. The β -based C_n^2 humidity correction is based on the dynamic sensible heat flux Q_H and the latent heat flux Q_L . Q_H (in W/m^2) is converted from the measured kinematic sensible heat flux $\overline{w'\theta'}$ (in mK/s) to its dynamic form as $Q_H = \rho c_p \overline{w'\theta'}$ with the density ρ and the specific heat capacity c_p of air (Stull 1988). The corresponding latent heat Q_L is also measured by the eddy covariance system at 3 m above ground (Bosveld *et al.* 2020). We refer to all these observed meteorological variables for C_n^2 estimation as MET data for the rest of the study.

2.2.4. MESOSCALE SIMULATION SETUP

The same variables observed at the CESAR site are numerically obtained by running one WRF simulation per test case. All simulations are forced with the ERA5 reanalysis

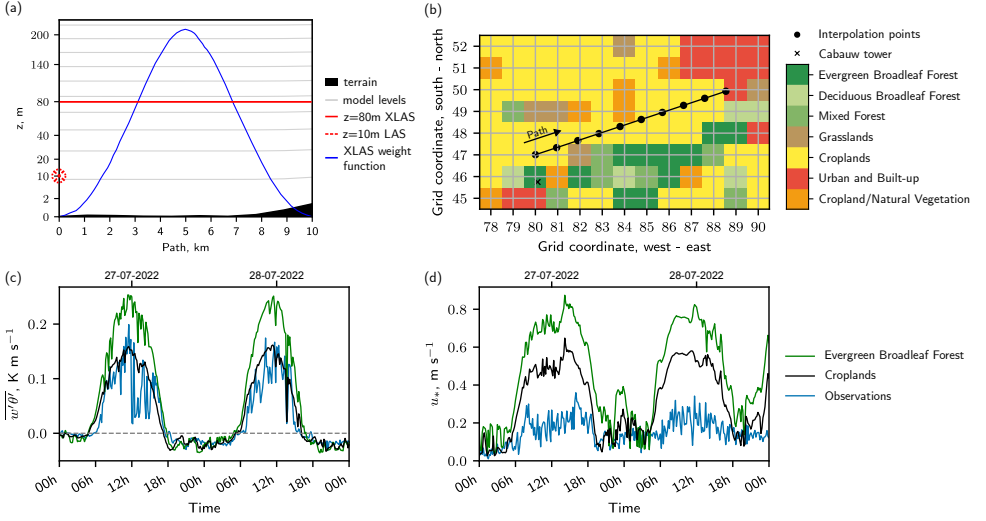


Figure 2.4.: Representation of the CESAR site in WRF. (a) Domain cross-section containing 10 km XLAS path with model levels in gray and LAS/XLAS weighting function in blue; (b) MODIS land use classes of CESAR site (cross) and its surroundings; effect of different land use classes on (c) sensible heat flux and (d) friction velocity.

Table 2.1.: Nested domain configuration used for all test cases in WRF

Domain	Horizontal resolution	# grid points	# vertical levels
D01	9 km × 9 km	128 × 100	101
D02	3 km × 3 km	124 × 127	101
D03	1 km × 1 km	130 × 127	101

data (Hersbach *et al.* 2020) provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). The horizontal grid size of ERA5 is ~ 30 km, and three nested WRF domains are used to bring the horizontal grid size down to 1 km in the finest domain (cf. table 2.1). WRF produces instantaneous snapshots of all variables every 10 minutes, matching the temporal sampling rate of the CESAR observations. Panels (a) and (b) of Fig. 2.4 depict how the CESAR site is represented as a gridded WRF domain. A vertical cross-section (cf. panel (a)) is extracted using bilinear interpolation along the path shown in panel (b). This plane (approximately) contains the Cabauw tower at the very left and the path of the XLAS (red solid line) in the horizontal direction. The average heights of the model pressure levels plotted as gray lines show that the WRF model has a vertical resolution close to the surface of ~ 25 m with the first level at ~ 9 m. The bell-shaped function plotted as solid blue line is the LAS/XLAS weighting function related to the scintillometers' path-averaging effect. The function's peak at half the path length means that turbulence in the path's center contributes strongest to the measured C_n^2 (Beyrich *et al.* 2021).

The aerial pictures in fig. 2.3(b) show that land use around the CESAR site is diverse. This heterogeneity of the surroundings – crop fields, grassland, urban areas – is represented in WRF through the 21-category IGBP-MODIS land use dataset (Broxton *et al.* 2014; NCEP 2023). Based on that dataset, one land-use class is automatically assigned to each surface grid point during WRF preprocessing as shown in fig. 2.4(b). Here, only the seven classes relevant to the selected area are shown. Depending on the assigned land-use class, the WRF model uses different values for, e.g., surface roughness (cities cause more drag than grassland) or heat capacity (water bodies have larger heat capacity than croplands), while solving for the various simulated variables. These surface-specific parameters are, for example, used by the surface layer scheme that solves for the surface fluxes based on Monin-Obukhov similarity theory (Dyer and Hicks 1970; Jiménez *et al.* 2012). The fluxes are then used in the flux-based C_n^2 parameterization, so an incorrectly assigned land-use class can cause systematic deviations between simulated values at one grid point and collocated observations. Such a land-use misclassification happens at grid point (80, 46), closest to the Cabauw tower. This grid point is classified as “Evergreen Broadleaf Forest” although the aerial photos in fig. 2.3 indicate croplands or grasslands for the CESAR site. Only the neighboring grid point (80, 47) is classified as “croplands,” which seems more realistic. The effect of this misclassification on sensible heat flux $\overline{w'\theta'}$ and friction velocity u_* is visible in panels (c) and (d) of fig. 2.4. The forest-type land-use class (green curve) leads to significantly larger values for both variables than the “croplands” class (black curve). More importantly, the lower values of the “croplands” time series agree better with observations from Cabauw (blue curve). The notably lower observed u_* in fig. 2.4(d) is likely due to the local smoothness of the CESAR site (Verkaik and Holtslag 2007), which WRF misses due to its coarse resolution. Still, u_* associated with the “croplands” class is assumed more realistic than the “forest” one. Since the W71-flux approach takes $\overline{w'\theta'}$ and u_* as inputs, these variables must be simulated realistically by WRF. We, therefore, select grid point (80, 47) as representative of the CESAR site and as the starting point of the extracted cross-section.

The selected physics schemes WRF utilizes to represent various processes in the atmosphere are listed in table 2.2, and the corresponding WRF configuration files (namelist.input files) are provided on Github¹. Most relevant for this study is the selection of the planetary boundary layer (PBL) scheme. The PBL scheme parameterizes the sub-grid turbulence, i.e., the turbulent processes within the boundary layer (first few kilometers of the atmosphere) and clean air turbulence in the free atmosphere (above the PBL). Such parameterizations are required because the grid size of mesoscale models is too coarse to explicitly resolve turbulence². The WRF model outputs different variables depending on the selected PBL scheme, so the PBL scheme selection can constrain the C_n^2 parameterizations that can be applied. We utilize the higher-order Mellor-Yamada-Nakanishi-Niino (MYNN) turbulence scheme in its 2.5-order closure version with eddy-diffusivity/mass-flux (Olson *et al.* 2019) (EDMF) option enabled. MYNN-2.5 is a high-order PBL scheme

¹https://github.com/mpierzyna/wrf_cn2_intercomp

²To explicitly resolve turbulence, one can run large-eddy simulations with $\mathcal{O}(10\text{m})$ resolution.

Table 2.2.: Physics schemes used in WRF for all test cases to parameterize atmospheric processes.

Physical process	Scheme
Surface layer	Mellor-Yamada-Nakanishi-Niino (MYNN) (Dyer and Hicks 1970)
Boundary layer	MYNN 2.5 (Nakanishi and Niino 2006; Nakanishi and Niino 2009) with eddy-diffusivity/mass-flux (Olson <i>et al.</i> 2019) (EDMF)
Microphysics	WSM-5 (Hong <i>et al.</i> 2004)
Radiation (long and short-wave)	Rapid Radiative Transfer Model for GCMs (RRTMG) (Iacono <i>et al.</i> 2008)
Land surface model	Noah LSM (Chen and Dudhia 2001) with IGMP-MODIS 21-category land use data (Broxton <i>et al.</i> 2014; NCEP 2023)
Cumulus parameterization	Modified Kain-Fritsch Scheme (Kain 2004) (D01 only)

that solves for TKE e and variance of potential temperature σ_θ^2 , which are required for variance-based C_n^2 parameterizations. Not all PBL schemes available in WRF provide TKE, and only MYNN 2.5 and MYNN 3 yield σ_θ^2 . Since we found MYNN 3 to be numerically unstable in some cases, MYNN 2.5 is the only PBL scheme suitable for this comparison. Outside this study, if only gradient-based or flux-based techniques are of interest, all other WRF PBL schemes are applicable. To determine the domain of applicability for the MOST-based C_n^2 parameterizations, we utilize the PBL height from the MYNN scheme and compute the SL height as 10% of that variable. The same estimated SL height is utilized for WRF-based and MET-based C_n^2 estimates to ensure comparability between the parameterizations.

Post-processing the WRF model outputs with the C_n^2 parameterizations yields estimates of C_n^2 at each model grid point with 1km horizontal spacing. Scintillometers, on the other hand, yield integrated C_n^2 values caused by the turbulence along the measurement path. In cases where this path crosses multiple WRF grid cells, the simulated C_n^2 estimates need to be aggregated to align them with the observations. This consideration is essential for the 10km XLAS. As mentioned, we horizontally interpolate the WRF grid to extract a vertical cross-section containing the scintillometer path. Next, we vertically interpolate the cross-section levels to the path height (here 80m AGL). Finally, a path-averaged C_n^2 estimate is computed using the scintillometer path weighting function (Wang *et al.* 1978) (see also fig. 2.4(a)). The bell-shaped weighting function of an LAS/XLAS puts the most weight on the center of the path. In the case of the LAS, where the path length of 859m is comparable to the grid resolution, we compare the observations against values from the grid point closest to the link's center, i.e., the starting point of the extracted cross-section.

2.2.5. PERFORMANCE METRICS

The agreement between any pair of observed $\log_{10} C_n^2$ time series of length n , $\{o_1, \dots, o_i, \dots, o_n\}$, and its estimated counterpart, $\{f_1, \dots, f_i, \dots, f_n\}$, is quantified in \log_{10} -space through three metrics: the constant bias between the two time series, the centered root-mean-squared error (cRMSE), and Pearson's correlation coefficient r . The bias is computed as the difference between the mean values of both signals,

\bar{f} and \bar{o} , as

$$\text{Bias} = \bar{f} - \bar{o} \quad (2.10)$$

and quantifies the constant shift between the two. The cRMSE is defined as

$$\text{cRMSE} = \frac{1}{n} \sum_{i=1}^n \left[(f_i - \bar{f}) - (o_i - \bar{o}) \right]^2 \quad (2.11)$$

and quantifies the average error throughout the signal after removing the bias. Additionally, the Pearson correlation coefficient r , defined as

$$r = \frac{\sum_{i=1}^n (f_i - \bar{f})(o_i - \bar{o})}{n \sigma_f \sigma_o}, \quad (2.12)$$

reflects how well the estimate matches the observation's pattern (e.g., diurnal cycle). Here, σ_f and σ_o are the standard deviations of the simulated and observed time series, respectively. For all metrics, the estimated C_n^2 time series are clipped to a minimum value of $C_n^2 \approx 10^{-16} \text{ m}^{-2/3}$ before computing the scores. That is done to avoid an unreasonably large contribution of estimation errors in the weak turbulence regime, which are exaggerated by the \log_{10} scaling. The $10^{-16} \text{ m}^{-2/3}$ threshold is commonly used in FSOC and optical astronomy literature to separate out weak and almost negligible OT strength (Tunick 1998; Weiss-Wrana and Balfour 2002; Masciadri *et al.* 2010; Raj and Padmavathi 2016). To match the different temporal sampling rates of the scintillometer observations (30 min) and the WRF/MET estimates (10 min) for the score computation, we downsample the 10 min data. More specifically, each 30 min scintillometer sample is matched by selecting the 10 min sample of the WRF/MET data that falls into the middle of the 30 min interval. We prefer this approach over, e.g., averaging three 10 min samples into one 30 min sample because we would like to retain the larger variability typically exhibited by flux measurements compared to, e.g., mean wind measurements.

2.3. RESULTS

In the following, the performance of each C_n^2 parameterization, with WRF and MET inputs, is compared against the scintillometer observations at 10 m and 80 m height (cf. table 2.3 for a result overview). We begin in section 2.3.1 with a surface layer-bound comparison. The MOST-based parameterizations – W71-flux and W71/HB16-grad – are compared to observations from the 10 m-high LAS. The 80 m high XLAS is not considered for this part of the analysis because the number of valid C_n^2 estimates is low due to the SL assumption. In other words, for long periods of time (especially in spring and autumn), the SL is too shallow to contain the high scintillometer, rendering the flux and gradient-based methods inapplicable. Consequently, the statistical significance of such a comparison is low and it would not be insightful. In contrast to that, the variance-based C_n^2 estimates are valid everywhere in the boundary layer, so we go beyond the SL in section 2.3.2 and assess the accuracy of the variance-based C_n^2 at both scintillometer levels. Here, only WRF data are used as inputs to HB15-var due to the unavailability of turbulent variances within the CESAR data product.

Table 2.3.: Overview of evaluated combinations of input data – observed (MET) or simulated (WRF) meteorology – and C_n^2 parameterizations for the two scintillometers mounted at height z and measuring along path of length Λ .

Scintillometer	MET input			WRF input		
	W71-flux	W71/HB16-grad	HB15-var	W71-flux	W71/HB16-grad	HB15-var
$z = 10 \text{ m}, \Lambda = 859 \text{ m}$	yes	yes	no obs. data	yes	yes	yes
$z = 80 \text{ m}, \Lambda = 10 \text{ km}$	no, beyond SL	no, beyond SL	no obs. data	no, beyond SL	no, beyond SL	yes

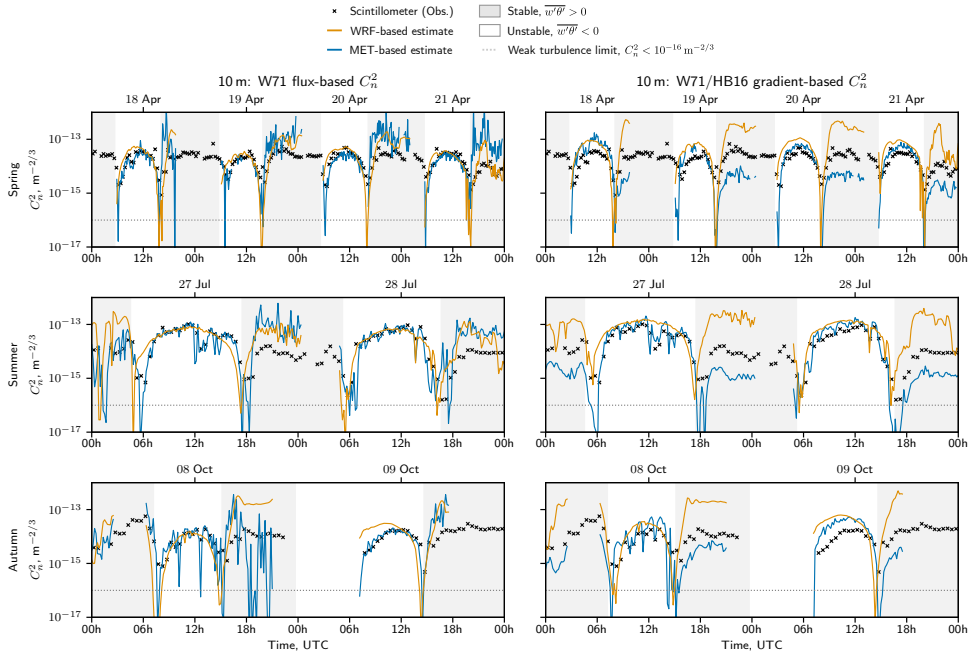
2.3.1. THREE-WAY COMPARISON OF C_n^2 PARAMETERIZATIONS IN THE SURFACE LAYER

Figure 2.5(a) presents the estimated and the observed 10 m level C_n^2 time series for all three test cases. The figure serves as a three-way comparison between observed SL C_n^2 and the effect of different input data (WRF or MET) when using the same C_n^2 parameterization or the impact of using the same data but different C_n^2 approaches (flux-based or gradient-based). The shading of the background indicates static atmospheric stability based on the simulated sensible heat flux: stable conditions ($\overline{w'\theta'} < 0$), typically present during the night, are shaded in gray and unstable conditions ($\overline{w'\theta'} > 0$), generally present during the day, are shaded in white. Most estimated curves show substantial dips to low C_n^2 values at the time of transition between stability, i.e. during neutral conditions. That is where the time series clipping to a minimum of $C_n^2 = 10^{-16} \text{ m}^{-2/3}$ (dotted lines) takes effect to avoid the unfair contribution of these negligibly low turbulence conditions. The gaps in the estimated time series stem from the SL being too shallow even for the 10 m LAS, while gaps in the observed scintillometer time series are due to low received power likely related to fog.

Bias, cRMSE, and correlation r for each pair of observed and estimated C_n^2 time series are displayed in fig. 2.6. The table follows the same structure as the time series plots, with flux-based results in columns (a) and (b) and the gradient-based results in columns (c) and (d). For ease of comparison, columns (e) and (f) contain the variance-based results for C_n^2 estimation beyond the surface layer, which are discussed in the following section. The rows break the scores down by test case and stability, and the cell colors allow for visual interpretation of the score values with respect to the color bars.

A first, overarching comparison of the W71-flux and W71/HB16-grad in fig. 2.5(a) reveals that both approaches, regardless of input data, generally match the pattern of the observations. For example, low-turbulence dips during neutral conditions or strong turbulence peaks during the day are timed mostly well, resulting in mostly high correlation coefficients in fig. 2.6. Only the MET+Flux (stable) comparison hints towards a seasonal performance dependency with lower r and higher cRMSE in spring and autumn compared to summer. The remaining cases show comparable performance throughout all seasons. Comparing the scores of the variance-based estimates (WRF+Var, (e) and (f)) against the flux and gradient-based approaches indicates an overall more balanced error behavior of WRF+Var with less variation of all metrics between day and night, which is discussed in detail in section 2.3.2.

(a) Three-way comparison of flux and gradient-based C_n^2 estimates at the 10m level only within the surface layer.



(b) Multi-level comparison of variance-based C_n^2 estimates within the surface layer (10m height, LAS) and beyond (80m height, XLAS). No variance-based estimates with MET inputs are made because of limited data availability.

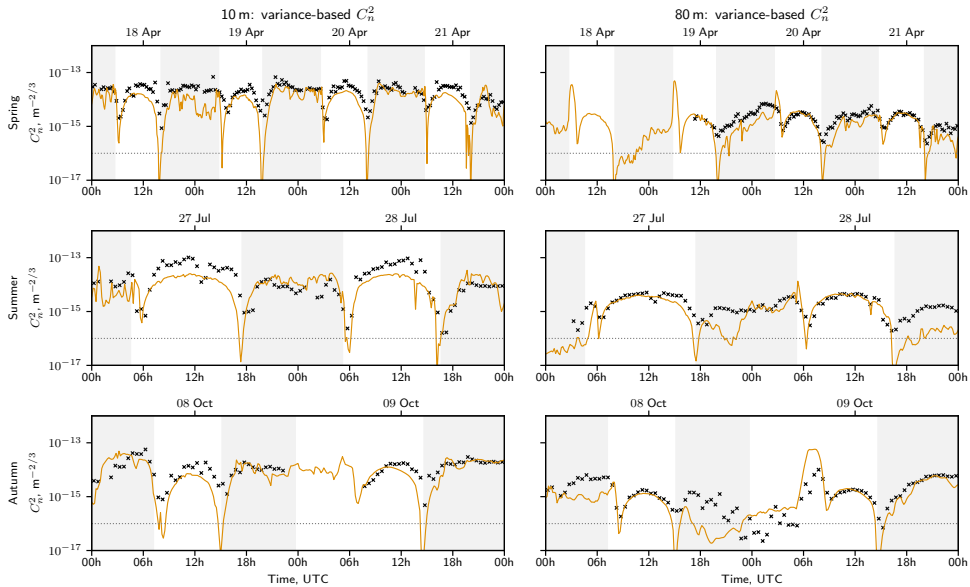


Figure 2.5.: Comparison of observed C_n^2 from scintillometers (black markers) against C_n^2 values estimated with three different parameterizations from simulated (orange) or observed (blue) meteorology. The rows in panels (a) and (b) correspond to the three test cases: spring (18 – 22 April 2022), summer (27 – 29 July 2022), and autumn (8 – 10 October 2022). 25

Despite the good overall performance, the W71-flux and W71/HB16-grad show significant biases during stable conditions upon detailed inspection. The nighttime bias is especially large for W71/HB16-grad regardless of season. With WRF inputs, the nighttime 10 m signal is overestimated by about one order of magnitude, whereas the estimates from MET input show a consistent underestimation of almost one order of magnitude. During unstable conditions, both WRF+Grad and MET+Grad slightly overestimate the C_n^2 values. These deviations could be related to too-strong (C_n^2 overestimation) or too-weak (C_n^2 underestimation) gradient estimates due to the coarse spacing of the sensors and model levels. For both input data, the gradients have to be computed numerically from levels separated by multiple meters, which is known to lead to deviations (Arya 1991). Also, the similarity relationship between C_T^2 and the gradients might be erroneous. In the case of the flux-based C_n^2 , WRF and MET inputs both lead to overestimating the nighttime values. The underlying observed and simulated nighttime fluxes show realistic values and orders of magnitude, which does not allow us to explain this behavior. The deviation seems systematic and could point to a weakness of W71-flux during the night, but a final conclusion would require an extensive analysis beyond the scope of this work.

Another difference between the C_n^2 estimates is the smoothness of the estimated signals. The gradient-based estimates are smoother, i.e., show fewer fluctuations, than the flux-based estimates, also visible in the lower cRMSE values for W71/HB16-grad. Again, the effect is more pronounced at night than during the day for all cases. Strong fluctuations of multiple orders of magnitude are observed during the 2nd night in autumn for the MET+Flux estimate. Here, the observed sensible heat flux is very close to zero, where the W71-flux parameterization is very sensitive, resulting in the depicted large jumps. Smoothness also differs between the two WRF-based estimates and their MET-based counterparts: the WRF-based results are smoother than the MET-based ones, which is expected since WRF's 1 km grid resolution is not fine enough to capture very local or short-lasting phenomena. For example, the large drop of C_n^2 during the first day of the summer case at ca. 12:00Z is entirely missed by WRF+Flux and WRF+Grad but captured well by the MET-based estimates.

Despite the discussed errors, both approaches yield consistently good estimates of daytime C_n^2 regardless of the utilized input data. Also, the generally good agreement between WRF-based and MET-based C_n^2 estimates (leaving the biases aside) is promising because it indicates that our WRF configuration reproduced the observed meteorology reasonably well.

2.3.2. BEYOND THE SURFACE LAYER: PERFORMANCE OF THE VARIANCE-BASED C_n^2 PARAMETERIZATION

So far, the discussion has focused on the SL, applying gradient and flux-based methods to 10 m observed or simulated meteorological variables. We now employ the variance-based method to estimate C_n^2 at higher altitudes and compare the results to scintillometer measurements from two heights: 10 m and 80 m above ground. Due to the difficulty to experimentally determine L_M and because of the limited availability of TKE and σ_θ^2 in the CESAR data product, we only utilize data

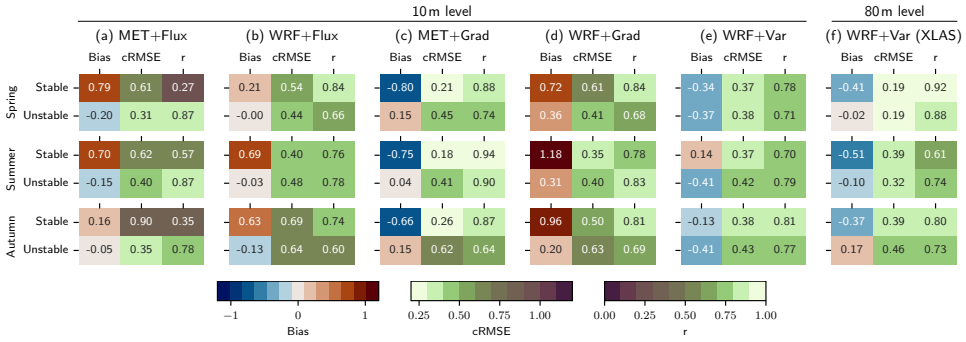


Figure 2.6.: Bias, centered RMSE, and correlation coefficient r quantifying the agreement between the observed and estimated $\log_{10} C_n^2$ time series (cf. Fig. 2.5). Columns (a) – (f) present the performance of the different combinations of input data (MET or WRF) and C_n^2 parameterizations with respect to scintillometer observations. The low values of all estimated time series are clipped to $C_n^2 = 10^{-16} \text{ m}^{-2/3}$ before computing the scores to avoid unfair dominance of negligibly weak turbulence in \log_{10} -space.

from the WRF model for this comparison. The estimated C_n^2 time series derived from WRF (WRF+Var) for both altitudes are shown in fig. 2.5(b). The 10m C_n^2 values (within the SL) are depicted in the left column and the 80m C_n^2 curves (often beyond the SL) in the right column. The previously discussed fig. 2.6 also includes the performance scores that assess agreement between the WRF+Var estimates and the observed data for easy comparison with the other approaches.

Similar to the SL results, the overall agreement of WRF+Var with observations is good. During most nights, C_n^2 matches the scintillometer data at both heights well in terms of magnitude and correlation. The daytime C_n^2 estimates also reflect the trends well by means of correlation, but WRF+Var consistently underestimates the daytime magnitude at the 10m level. This behavior cannot be observed at the 80m level where daytime C_n^2 is captured with high accuracy as quantified by the low biases. Interestingly, the trend reverses for stable conditions with larger biases at the 80m level and lower biases at the 10m level. The reason could be that WRF misses SL contributions to the turbulence during the day and some higher altitude effects during the night. Such systematic discrepancies are common in numerical weather modeling because the complex physical processes of our atmosphere have to be simplified or approximated, e.g., to keep the models' computational complexity reasonable (Dudhia 2014).

Along those lines, we attribute the three notable trend deviations at night between WRF+Var and the scintillometers (cf. fig. 2.5(b)) to complex and/or local weather phenomena. Such weather patterns are difficult to capture in WRF due to its coarse resolution. One deviation occurs in spring at the 10m level (18/19 April, 0:00Z–6:00Z) and two others at the 80m level in summer (27/28 July, 18:00Z–00:00Z) and autumn (8/9 October, 18:00Z–6:00Z). The signal degradation experienced by both scintillometers during the autumn deviation supports the assumption that a complex weather pattern is present. The observed 80m C_n^2 drops to very low values

and exhibits increased scattering, while the corresponding 10 m C_n^2 observations fail quality control and are discarded. It seems that WRF captured the weather event leading to the signal degradation but with a different timing.

Despite the discussed periods of temporarily low performance, the results are overall satisfying. There is no clear seasonal trend of the performance (visually and quantitatively), and the correlation between estimated and observed C_n^2 is high for both heights as reflected by the scores in fig. 2.6(e) and (f). When directly compared to the flux and gradient-based alternatives, the variance-based C_n^2 correlation values are lower in some cases at 10 m height but show less variation between stability conditions. We note that the scores of the MOST-based approaches are based on fewer samples for atmospheric stable conditions due to the SL constraint, so the statistical uncertainty is expected to be higher. The biases of the variance-based C_n^2 estimates are mostly negative, indicating a consistent underestimation of turbulence strength up to half an order of magnitude. This underestimation trend of WRF+Var contrasts the more variable performance of W71-flux and W71/HB16-grad, which show lower (absolute) biases during the day but significantly higher (absolute) biases during the night. In general, the error behavior of WRF+Var seems more balanced with similar r and cRMSE values throughout seasons and atmospheric stability conditions compared to the other methods. This observation is confirmed visually by comparing the WRF+Var time series (fig. 2.5(b)) with those of WRF+Grad and WRF+Flux (fig. 2.5(a)): the WRF+Var curves track the scintillometer observations more consistently for all test cases, day and night, and for both heights. Consequently, HB15-var seems to be the more reliable choice for C_n^2 estimation from WRF than the other two methods, and it has the additional advantage of being applicable anywhere in the boundary layer.

2.4. CONCLUSION

We apply three distinct C_n^2 parameterizations to observed and simulated meteorological variables and compare the resulting C_n^2 estimates to measured C_n^2 evolutions from two scintillometers mounted at different heights and measuring at paths of different lengths. The parameterizations utilize either (a) vertical gradients of the mean wind and temperature fields, (b) turbulent fluxes, or (c) turbulent variances to estimate C_n^2 . The meteorological inputs are either experimentally obtained (MET) at the Cabauw Experimental Site for Atmospheric Research (CESAR) in the Netherlands or are produced as mesoscale hindcast of the same site by the Weather Research and Forecasting (WRF) model. The variance-based method is only applied to WRF model data because observed MET variances are not routinely available and because the experimental estimation of the associated length scale is challenging. Three multi-day test cases from spring (4 days), summer (2 days), and autumn (2 days) are selected to investigate the seasonal variability of the parameterizations' performance. While our study is focused on CESAR due to its extensive instrumentation, we anticipate our results to be valid also for comparable sites.

Overall, all three parameterizations perform well during stable (nighttime) and

unstable (daytime) conditions and do not show a significant seasonal accuracy dependency. Detailed analysis revealed that the flux and gradient-based methods yield large constant biases up to one order of magnitude for stable/nighttime estimates compared to measured C_n^2 . In contrast, the unstable/daytime values are close to observations. Since both parameterizations are more sensitive to their inputs during stable than unstable conditions, this finding stresses the need to accurately model or measure the parameterization inputs, especially at night. The variance-based approach tends to underestimate the night and day values but shows high correlation. Since we do not have access to the measured turbulent variances, we cannot investigate whether the underestimation results from the variance-based formulation or the simulated data. Nevertheless, the overall error behavior is more balanced compared to the other two methods, with less variation in the performance metrics for different seasons and atmospheric stabilities. Concerning the input data, the WRF-based estimates are found to be as realistic as the MET-based ones compared to the scintillometers. The cases where WRF deviates from observed trends can typically be attributed to local or complex short-duration weather events. Such patterns are known to be challenging to model in mesoscale simulations. Due to the statistically high accuracy of the WRF-based results, we still consider them reliable and valuable for quantifying important trends.

For FSOc purposes, the selection of a C_n^2 parameterization is highly use case-dependent. The typical goal is to quantify optical turbulence (OT) strength at or around a site of interest by determining C_n^2 experimentally or numerically to feed it to FSOc downstream tasks. Such tasks can be the design of adaptive optics systems at a laser terminal (Broekens *et al.* 2023a) or the quantification of the effect that OT has on a propagating beam (Yasa *et al.* 2023). The details of the respective use cases guide the selection of an appropriate parameterization. It matters, for example, whether a simulation or a field campaign is carried out and how the FSOc link under consideration is set up (e.g., vertical or horizontal). In the case of field campaigns, the flux and gradient-based methods are most relevant due to the challenging length scale estimation required for the variance-based parameterization. The gradient-based approach only requires simple routinely available mean values of meteorological variables but from multiple instruments at different heights. The flux-based method, on the other hand, utilizes turbulent fluxes from more expensive instruments but only needs observations close to the surface. Consequently, the instruments available for a field campaign can constrain the choice of parameterization. In WRF, fluxes and the underlying fields for the gradients are routinely available. However, the numerical computation of gradients can introduce additional errors (Arya 1991), and the fluxes strongly depend on the selected land use dataset, which needs to be representative of the site being simulated. Also, the theoretical foundation of both parameterizations – Monin-Obukhov similarity theory (MOST) – only allows their application inside the constant-flux surface layer (SL), which varies in depth throughout the day. Therefore, MOST-based C_n^2 estimates for, e.g., FSOc terminals or terrestrial links are only valid when they are fully contained in the SL.

The SL restriction does not apply to the variance-based method. This approach

can be utilized for the entire atmosphere but requires outputs from a mesoscale model (e.g., WRF) configured with a high-order turbulence closure scheme. We envision that such variance-based C_n^2 estimates based on WRF (WRF+Var) can be utilized to assess the behavior of various FSOC links. For example, the long 10 km path of the 80 m high scintillometer used in this study could be viewed as an intra-city link with terminals mounted on high-rise buildings. Similarly, the lower scintillometer with 859 m path length indicates the applicability of WRF+Var also for shorter links where the length is similar to the horizontal resolution of the mesoscale model. Resolving even shorter links of $\mathcal{O}(100\text{m})$ requires to run gray-zone or large eddy simulations (Honnert *et al.* 2020). WRF supports both, and we do not see a fundamental limitation of the variance-based parameterization, which should hinder its successful application. Nevertheless, the accuracy of the C_n^2 estimates in these cases must be assessed in future work. We also see the potential to analyze OT along links between space and ground using WRF+Var. While our study did not focus on such high-altitude applications, Basu *et al.* (2020) demonstrated good agreement between variance-based C_n^2 and observed C_n^2 from thermosondes up to $\sim 25\text{km}$ height. Combined with the balanced error behavior and the overall good performance demonstrated in our work, we believe that the mesoscale simulations and variance-based C_n^2 estimates can be very useful for the optics community.

In conclusion, our study presents the differences between three main classes of C_n^2 parameterizations concerning stability-dependent performance, theoretical limitations, and relevance for FSOC applications. The selection is highly use case dependent, but it generally seems that the gradient and flux-based approaches are better suited for near-surface field campaigns, while the variance-based method, in combination with a mesoscale model, has advantages for numerical studies across the entire atmosphere.

ACKNOWLEDGMENTS

This publication is part of the project FREE – Optical Wireless Superhighways: Free photons (at home and in space) (with project number P19-13) of the research programme TTW-Perspectief which is (partly) financed by the Dutch Research Council (NWO). We are grateful for the observational data provided by Royal Netherlands Meteorological Institute (KNMI) and the Ruisdael Observatory, a scientific research infrastructure which is (partly) financed by the Dutch Research Council (NWO, grant number 184.034.015).

DATA AVAILABILITY

The post-processed WRF simulation and MET data are available at <https://doi.org/10.5281/zenodo.10966120>. The `namelist.input` files to run the simulations and the scripts to obtain the presented results are available on Github: https://github.com/mpierzyna/wrf_cn2_intercomp. The meteorological observations of the CESAR site are publically available through the data platform of the Royal Netherlands Meteorological Institute: <https://dataplatfom.knmi.nl/>. The

scintillometer validation data are not publicly available at this time but may be obtained from the authors upon reasonable request.

2.A. TECHNICAL DETAILS OF GRADIENT COMPUTATION

The two gradients $\Gamma = \partial\bar{\theta}/\partial z$ and $S = \sqrt{(\partial\bar{u}/\partial z)^2 + (\partial\bar{v}/\partial z)^2}$ required for the gradient-based C_T^2 method in eq. (2.5) are computed from the WRF-simulated 3D fields or the measured multi-height time series of the potential temperature $\bar{\theta}$ and the horizontal wind components \bar{u} and \bar{v} . All gradients need to be computed accurately, especially close to the ground where they are usually strongest (Stull 1988). We employ central differences to numerically evaluate the gradients $\partial\bar{\square}/\partial z$ at the i -th vertical model level from the values at levels $(i-1)$ and $(i+1)$. This poses a problem for the lowest and highest model levels where no $(i-1)$ or $(i+1)$ levels are available. The highest levels in WRF ($z \approx 20$ km) and on the Cabauw tower ($z = 200$ m) are far outside the SL – the area of applicability of W71/HB16-grad –, so no gradients need to be computed there. However, the lowest model level at $z \approx 10$ m is highly relevant because it is almost the same as the LAS mounting height. Therefore, we manually introduce lower-level information based on physical reasoning. Using the surface layer scheme, WRF outputs a 2 m potential temperature based on which the lowest Γ values are computed. Similarly, Cabauw has an additional 2 m temperature sensor. For the velocity gradients in z direction comprising the wind shear S , we assume a logarithmic wind profile so that $\bar{u}(z = z_0) = \bar{v}(z = z_0) = 0$. The surface roughness length z_0 is provided by WRF’s land-surface model or was determined experimentally for CESAR by Verkaik and Holtslag (2007). Contrary to the findings of Verkaik and Holtslag that z_0 is a function of wind direction, we assume a constant value of $z_0 = 0.02$ m not to introduce additional complexity. Since the first wind vane is located at $z = 10$ m, which is much higher than even $\max\{z_0\} = 0.5$ m determined by Verkaik and Holtslag (2007), we believe that this assumption does not significantly influence the gradients and is justified.

3

Estimating turbulence-induced link losses in an urban environment from observed and numerically estimated C_n^2

The parameterization intercomparison in [chapter 2](#) demonstrated that WRF can reproduce observed C_n^2 with acceptable errors. However, the implications of these estimation errors for FSOC link performance are not immediately clear. This chapter addresses this gap by developing a framework to estimate turbulence-induced link losses and communication performance metrics from C_n^2 time series for a collimated-beam FSOC link. The framework quantifies link losses due to on-axis scintillation, beam wander, and beam spread, and evaluates their impact on the generalized mutual information (GMI), a theoretically achievable information rate. We apply the framework to an example link inspired by the Reid Photonloop, a 4.6km urban FSOC testbed in Eindhoven, The Netherlands, using C_n^2 observations from a collocated boundary layer scintillometer and hindcasts from a matching WRF simulation. Results demonstrate that C_n^2 errors propagate non-linearly through the modeling chain, with GMI sensitivity to C_n^2 errors depending strongly on link design and operating conditions. For links operating with small link margins, small C_n^2 errors can produce substantial GMI errors. Consequently, reporting only C_n^2 error metrics is insufficient for FSOC applications, and end-to-end performance metrics such as GMI are necessary for meaningful system assessment.

3.1. INTRODUCTION

Previous studies (e.g., Masciadri *et al.* (1999), Frehlich *et al.* (2010), Cheinet *et al.* (2011), Basu *et al.* (2020), Ullwer *et al.* (2019), Rafalimanana *et al.* (2022), and Pierzyna *et al.* (2024); cf. chapter 2) aimed to reproduce optical turbulence (OT) observations with numerical mesoscale simulations. Common OT variables discussed in such studies include the structure parameter of the 2nd-order refractive index structure function, C_n^2 , as well as integrated parameters such as the Fried parameter r_0 and the scintillation index σ_I^2 . To assess agreement between models and observations, standard metrics such as the Pearson correlation coefficient or the root-mean-square error are typically used. While these metrics are well-suited to quantify statistical agreement between datasets, the resulting error and correlation scores are difficult to interpret practically. For example, to design a communication system, one must consider turbulence-induced link losses in dB rather than values of C_n^2 , r_0 , or σ_I^2 .

In this study, we bridge this interpretational gap by showing how C_n^2 estimates can be further processed to estimate two metrics from communication theory: signal-to-noise ratio (SNR) modulated by turbulence-induced link losses and the generalized mutual information (GMI). GMI is a theoretically achievable information rate (similar to a data rate) of a channel subject to bit-interleaved coded modulation (Caire *et al.* 1998) and is commonly used to quantify the performance of coding and modulation schemes (see, e.g., Alvarado *et al.* (2018)). While OT parameters depend only on the meteorology at the time and location of interest, communication-related metrics also depend on the characteristics of the communication system. So, while SNR and GMI are more interpretable from a communication perspective, the conclusions drawn are less general, as they may apply only to the link under consideration.

The example link used in this study is inspired by the *Reid Photonloop*, a 4.6km long horizontal link operated as a free-space optical communication (FSOC) testbed in Eindhoven, The Netherlands (Vliet *et al.* 2025; Vliet *et al.* 2026). Vliet *et al.* (2026) carried out a 3-week field campaign in summer 2025, during which C_n^2 along the path was observed with a boundary layer scintillometer (BLS). We aim to reproduce this observed C_n^2 with a numerical mesoscale model and estimate SNR and GMI from both signals to illustrate how they can be interpreted in terms of practical communication performance. The Photonloop setup used for the performance estimation is idealized in two respects to keep the optical propagation theory tractable and to increase the generalizability of the results. First, we assume a collimated beam, whereas the original setup uses a diverging beam, which allows the use of well-established closed-form expressions for turbulence-induced losses (Parenti and Sasiela 2005; Giggenbach 2008) instead of case-specific numerical integrations. Second, we neglect losses within the communication system itself, such as coupling losses or detector inefficiencies, because they are highly system-dependent and do not interact with the atmospheric propagation path that is the focus of this work. Despite these simplifications, the framework retains the key elements required for a meaningful end-to-end demonstration: realistic time-dependent meteorological conditions, independent C_n^2 observation as a quantitative baseline, and a representative urban link geometry. The

collimated-beam assumption affects the functional form of the turbulence-induced losses, so the specific error-propagation characteristics reported here are tied to this beam type. However, the overall methodology is general and can be applied with loss expressions appropriate for different beam geometries. We therefore consider the present study a solid foundation for future work involving more complex beam types.

3.2. EXAMPLE CASE: REID PHOTONLOOP WITH COLLIMATED BEAM

An experimental FSO campaign (Vliet *et al.* 2026) was conducted from 18 August 2025 until 9 September 2025 in the city of Eindhoven, The Netherlands, using the *Reid Photonloop*, a 4.6km urban FSO testbed. The setup of the link and its key characteristics as used in this study are presented in [fig. 3.1](#) and [table 3.1](#), respectively. Compared to the divergent beam of the real setup, we employ a collimated beam with a smaller divergence angle to avoid complex case-specific numerical integrations of the turbulence-induced losses in favor of simple closed-form expressions. More details on this follow in [section 3.3.2](#).

A boundary layer scintillometer (BLS; Scintec BLS900 neo) is mounted directly next to the Photonloop terminals to measure C_n^2 independently from the link (cf. [fig. 3.1\(a\)](#)). The observed path-averaged C_n^2 serves as a reference signal for our study, meaning that the turbulence-induced losses and communication performance metrics are based on the BLS C_n^2 , not on the actual received power measured during the campaign. The reason is threefold. First, the received power reported in Vliet *et al.* (2026) is not the free-space power received outside the terminal, but rather the power coupled into the fiber. This signal is subject to additional, highly system-specific losses that are not the focus of this study. Second, Photonloop employs tip-tilt correction to mitigate turbulence-induced losses partially, so the measured power inside the terminal likely does not show all effects targeted in this work. Third, simplifying the Photonloop setup by assuming a collimated beam reduces the assumed divergence compared to the real setup, leading to an underestimation of the geometric loss (i.e., loss due to natural divergence of the beam). Consequently, a direct comparison of estimated and observed received power is not meaningful. Nevertheless, the study remains valuable because the atmospheric input is identical for both the idealized and the real setup, and the overall methodology of propagating C_n^2 to link losses and communication metrics is general.

The counterpart of the observed C_n^2 is a C_n^2 hindcast made with the Weather Research and Forecasting (WRF) model (Skamarock *et al.* 2021). We used the configuration presented by Pierzyna *et al.* (2024) (cf. [chapter 2](#)) to generate a 1 km \times 1 km hindcast of the atmospheric conditions experienced by Photonloop during the campaign. The variance-based C_n^2 parameterization proposed by He and Basu (2015) (cf. [eq. \(2.7\)](#)) is employed to estimate time-varying 3D-fields of C_n^2 from WRF. Following the procedure outlined in Pierzyna *et al.* (2024) (cf. [section 2.2.4](#)), the 3D fields are horizontally interpolated to ~ 40 m above ground to match the height of

Table 3.1.: Characteristics of idealized Photonloop, following Vliet *et al.* (2026) if not marked otherwise. A summary of all simplifications is presented in section 3.B.

Parameter	Symbol	Value
Path length	R	4.6 km
Local height above ground	z	40 m
Aperture diameter (RX and TX)	D	10 cm
Beam waist radius at TX aperture	w_0	2.5 cm
Wavelength	λ	1550 nm
Beam divergence ⁽¹⁾	θ_{div}	$\sim 20 \mu\text{rad}$
Transmitted power	P_{TX}	29 dBm
RX outage probability	p_{out}	10^{-3}
Modulation format	-	4-QAM

⁽¹⁾ We assume a collimated beam with $\theta_{div} = \lambda/(\pi w_0)$, while the real setup has a diverging beam with $\theta_{div} \approx 200 \mu\text{rad}$.

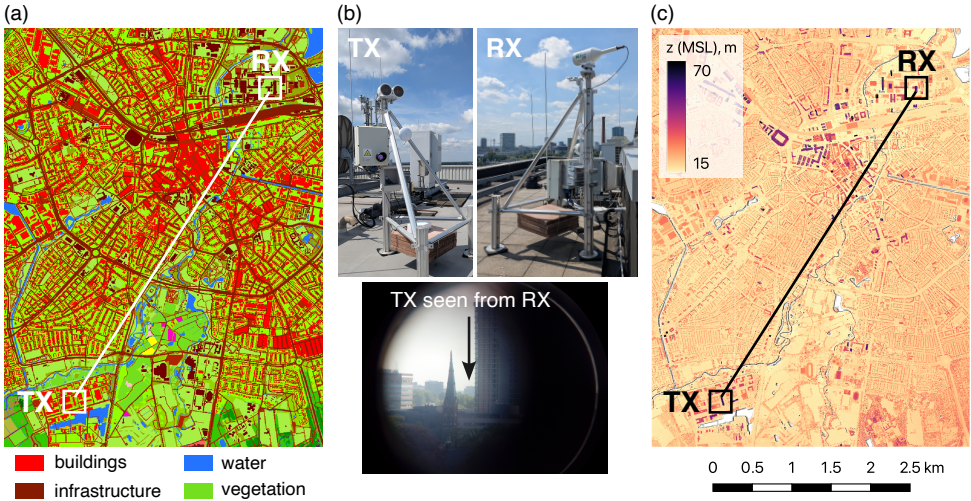


Figure 3.1.: Setup and urban operating environment of Reid Photonloop. Panel (a) displays land use from Hazeu *et al.* (2023) paired with elevation data from AHN (2021) in panel (c). Photos of transmitter (TX) and receiver (RX) side setups with laser terminals and boundary layer scintillometer are shown in (b).

Photonloop. C_n^2 extracted along the path is further path-averaged using the BLS weighting function (Beyrich *et al.* 2021), which weights contributions from the center of the path more strongly than from the edges. The resulting estimated C_n^2 time series is considered comparable to the observed C_n^2 from the BLS.

We also note that the urban environment affects optical propagation in a complex manner. For example, the closer the link is to the ground, the stronger individual shear and buoyancy sources (buildings, streets, vegetation) affect the turbulence (Oke 1984; World Meteorological Organization 2023). The wind direction determines the

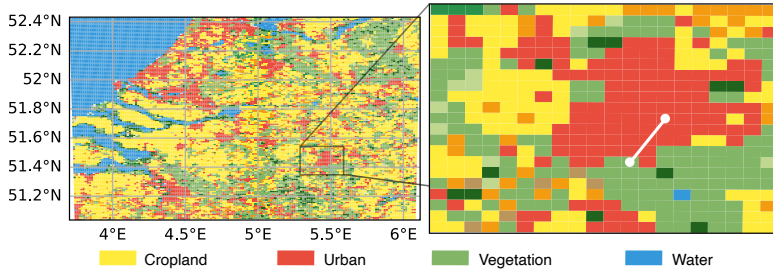


Figure 3.2.: MODIS land use classes used by WRF to represent Photonloop environment.

upwind fetch area, which can vary between open parks and dense urban canyons (cf. fig. 3.1(a)), modulating the turbulence further. These effects are also difficult to capture in mesoscale models such as WRF, which parameterizes cities only by increasing surface roughness and heat capacity by default (Chen and Dudhia 2001). This coarse description of the urban environment is shown in fig. 2.4, which depicts the IGBP-MODIS land use classes (Broxton *et al.* 2014; NCEP 2023) used by WRF. In comparison with the much finer LGN data in fig. 3.1(a), WRF cannot distinguish buildings or roads, but treats the entire city with a single land use class. This inherent mismatch can make C_n^2 time series and FSOC link performance difficult to interpret at times. Since the present study focuses on a practical modelling framework for FSOC systems, the urban environment of Eindhoven is not quantified in detail, as in e.g., Lee *et al.* (2015) or Zieliński *et al.* (2018), but such a study should be carried out in the future.

3.3. ESTIMATION OF FSOC LINK PERFORMANCE UNDER ATMOSPHERIC TURBULENCE

The performance of an FSOC link is commonly quantified in terms of the received power, P_{RX} , which is related to the transmitted power, P_{TX} , and the various losses incurred along the propagation path through the atmosphere. Losses can be constant in time (static) or time-varying (dynamic) and are typically expressed as attenuations in decibels (dB). In logarithmic decibel space, the free-space link budget balancing transmitted power, received power, and the losses is given as (Kaushal and Kaddoum 2017; ITU 2007)

$$P_{RX} = P_{TX} - L_{geo} - L_{atm} - L_{OT} - L_{sys}. \quad (3.1)$$

Here, L_{geo} is the static geometric loss due to beam divergence, L_{atm} is the loss due to atmospheric absorption and scattering, L_{OT} is the highly dynamic loss due to atmospheric optical turbulence, and L_{sys} accounts for losses inside the communication system. To establish a reliable FSOC link, the received power P_{RX} must be distinguishable from the background noise P_n , so a common design parameter for communication systems is the signal-to-noise ratio (SNR)

$$SNR = P_{RX} - P_n \quad (3.2)$$

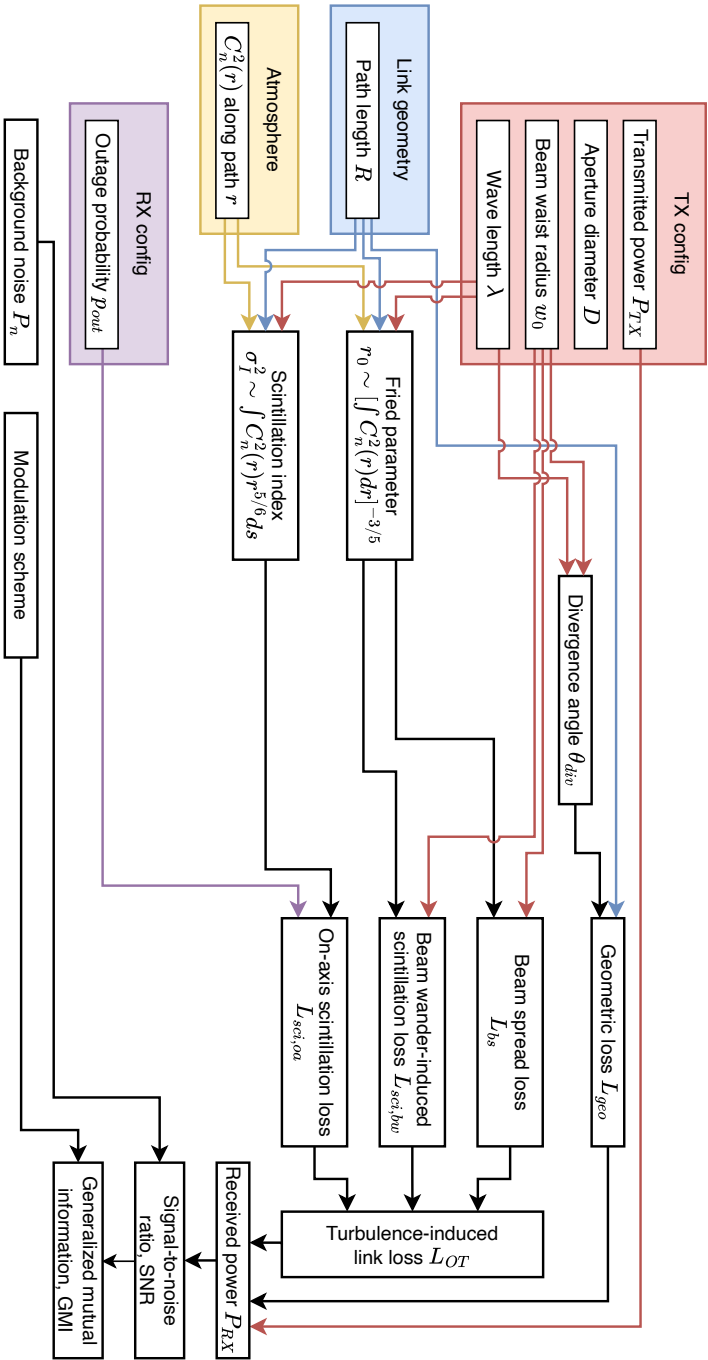


Figure 3.3: Flowchart linking C_n^2 (yellow) and communication system characteristics to optical turbulence-induced link losses and the FSOC performance metrics, signal-to-noise ratio, and generalized mutual information (GMI).

in decibel. For successful communication, the SNR must exceed a threshold, typically determined by the receiver sensitivity and other system characteristics.

The primary focus of the present study is to determine L_{OT} as a function of C_n^2 for a link employing a collimated beam. The geometric, atmospheric, and system losses are of secondary relevance to our study, as will be justified in the following. The geometric loss depends on the divergence angle θ_{div} , the propagation distance R , and the receiver aperture diameter D , which are all static. From $L_{geo} = 10 \log_{10} [(R\theta_{div})^2 / D^2]$ (ITU 2007) and the link characteristics in table 3.1, we estimate a static loss of ca. 5.2 dB for the idealized Photonloop case. In the real setup with a diverging beam, the geometric loss is expected to be higher at ca. 25 dB but still static. In both cases, the values are significant and should not be neglected, but, as they are static, they require no in-depth discussion.

Atmospheric scattering and absorption by gas molecules and aerosols constitute an additional, wavelength-dependent loss. At 1550 nm, molecular attenuation is small and essentially static at $\mathcal{O}(0.01 \text{ dB/km})$ (Kaushal and Kaddoum 2017). Aerosol attenuation is more variable: pollution particles such as black and brown carbon and mineral dust are known attenuators (Moosmüller *et al.* 2009), and their concentrations fluctuate with boundary-layer height (Stull 1988) and local sources, particularly in urban environments. Depending on visibility conditions, aerosol-induced attenuation can reach $\mathcal{O}(0.1 \text{ dB/km})$ (McClatchey and Selby 1974; Shuto 2023). The most significant contributions stem from fog, rain, and snow with losses reaching $\mathcal{O}(10)$ to $\mathcal{O}(100 \text{ dB/km})$ (Kaushal and Kaddoum 2017; ITU 2007). However, the majority of the Photonloop campaign was free of such events, so no such drastic losses are expected. Because no collocated aerosol observations are available and incorporating aerosol chemistry into the model would require a substantial additional modelling effort (Smith *et al.* 1993), we also neglect atmospheric scattering and absorption. The resulting underestimation of total path loss is expected to be of $\mathcal{O}(1 \text{ dB})$ for the Photonloop case.

Finally, we do not consider system losses in this study, as our focus is on the free-space segment of the link. System losses are highly system-specific and should be accounted for in production performance assessments, but are neglected here to increase generalizability.

Ultimately, the main focus of this study is the dynamic loss induced by atmospheric optical turbulence on a collimated beam, L_{OT} . OT losses strongly depend on meteorological conditions and can severely degrade communication performance. We consider the three dominant loss components: beam-spread loss, beam-wander-induced scintillation loss, and on-axis scintillation loss (Kaushal and Kaddoum 2017). The losses are detailed in section 3.3.2 after introducing the integrated turbulence parameters on which they depend in section 3.3.1. In section 3.3.3, we discuss how the losses affect the SNR and the generalized mutual information (GMI), which are standard FSOC performance metrics. A flowchart summarizing the relationships between system characteristics (cf. table 3.1), C_n^2 , the turbulence-induced losses, and the FSOC performance metrics is presented in fig. 3.3.

The assumption of a collimated beam considerably simplifies the computation

of the statistics of Gaussian beams under turbulence and is commonly made in literature (e.g., Andrews and Phillips (2005) and Parenti and Sasiela (2005)). For a general Gaussian beam with beam waist radius w_0 and phase front radius of curvature F_0 , the optical propagation depends on the Gaussian beam parameters in the transmitter (TX) plane $\Theta_0 = 1 - (R/F_0)$ and $\Lambda_0 = 2R/(kw_0^2)$, where $k = 2\pi/\lambda$ is the wavenumber and R the path length (Andrews and Phillips 2005). Computing the beam statistics for general Θ_0 and Λ_0 requires numerical integration of convoluted expressions, which can obfuscate the key dependencies of FSOC performance on C_n^2 . Instead, we target simplicity and interpretability by assuming a collimated Gaussian beam ($F_0 \rightarrow \infty$, $\Theta_0 \rightarrow 1$), which allows the use of simple closed-form solutions of the beam statistics. Furthermore, we assume that the beam arrives at the receiver (RX) as a spherical wave because the RX is considered far enough in the far field as indicated by $\Lambda_0 \approx 3.63$ (Andrews and Phillips 2005).

3.3.1. INTEGRATED TURBULENCE PARAMETERS

The cumulative effect of turbulence on an optical beam is commonly quantified by the integrated turbulence parameters, which are presented in the following: the Fried parameter, r_0 , and the scintillation index, σ_I^2 .

Fried Parameter The Fried parameter (Fried 1966) (in meters) is a measure to determine the strength of wavefront distortions caused by OT. If the ratio of receiver aperture diameter to Fried parameter, D/r_0 , is less than one, the optical system operates close to its theoretical optimum, i.e., is diffraction-limited. At the same time, turbulence-induced distortions degrade the performance for $D/r_0 > 1$. Assuming a spatial evolution of $C_n^2(r)$ along a propagation path r , the Fried parameter of a spherical wave is given as (Andrews and Phillips 2005)

$$r_0 = \left(0.16k^2 \int_0^R C_n^2(r) dr \right)^{-3/5}, \quad (3.3)$$

where $k = 2\pi/\lambda$ is the wavenumber with $\lambda = 1550$ nm (cf. table 3.1) and R is the path length. As scintillometers only yield path-integrated C_n^2 values, eq. (3.3) can be integrated assuming constant C_n^2 :

$$r_0 = (0.16k^2 C_n^2 R)^{-3/5}. \quad (3.4)$$

This expression shows an inverse relation between r_0 and C_n^2 , where larger r_0 values indicate weaker turbulence.

The constant C_n^2 assumption needs to be made for practical reasons to compute r_0 from path-integrated C_n^2 BLS observations. To make our WRF-based C_n^2 estimates compatible with the BLS, we also compute a path-averaged C_n^2 utilizing the BLS weighting function (Beyrich *et al.* 2021) and use eq. (3.4) instead of eq. (3.3). As the land use class along the simulated link is homogenous (cf. fig. 3.2), we do not expect much variation of simulated C_n^2 along the link.

Scintillation Index The scintillation index, σ_I^2 , expresses the normalized variance of received optical intensity fluctuations (Andrews and Phillips 2005)

$$\sigma_I^2 = \frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}, \quad (3.5)$$

where $I(t)$ is the received intensity/irradiance signal and $\langle \square \rangle$ denotes the ensemble average over a given time period. A theoretical connection between σ_I^2 and $C_n^2(r)$ along the path can be made by assuming a Kolmogorov turbulence spectrum (inner scale of zero, infinite outer scale) and Rytov theory.

We first define the Rytov variance for spherical waves as (Andrews and Phillips 2005; Andrews 2014)

$$\sigma_R^2 = 0.916k^{7/6} \int_0^R C_n^2(r)r^{5/6} dr, \quad (3.6)$$

where k and R again correspond to the laser wavenumber and propagation length. The difference to r_0 (cf. eq. (3.3)) is the additional $r^{5/6}$ term in the integrand, which weights turbulence contributions away from the receiver more strongly. This weighting is especially relevant for space-to-ground links where the path length is long and C_n^2 varies strongly with height. By considering an already path-integrated C_n^2 as discussed before, eq. (3.6) simplifies to

$$\sigma_R^2 = 0.5k^{7/6} C_n^2 R^{11/6}. \quad (3.7)$$

To account for the saturation of scintillation in strong turbulence conditions, i.e., the decoupling of C_n^2 and σ_I^2 , we relate σ_R^2 to σ_I^2 as (Andrews and Phillips 2005; Andrews 2014)

$$\sigma_I^2 = \exp \left(\frac{0.49 \sigma_R^2}{\left[1 + 0.56 (\sigma_R^2)^{6/5}\right]^{7/6}} + \frac{0.51 \sigma_R^2}{\left[1 + 0.69 (\sigma_R^2)^{6/5}\right]^{5/6}} \right) - 1. \quad (3.8)$$

This expression also includes the case of weak turbulence ($\sigma_R^2 < 1$) where $\sigma_I^2 \approx \sigma_R^2$.

3.3.2. TURBULENCE-INDUCED LINK LOSSES

The integrated turbulence parameters, r_0 and σ_I^2 , are used in the following to estimate the three primary turbulence-induced losses expected to affect the link: beam spread loss, beam wander-induced scintillation loss, and on-axis scintillation loss.

Beam spread and beam wander OT affects the beam in the RX plane through two mechanisms (Andrews and Phillips 2005; Parenti and Sasiela 2005): beam spread and beam wander. Beam spread refers to additional turbulence-induced diffraction that broadens the beam beyond its natural divergence, reducing the time-averaged intensity at the RX. Beam wander is the rapid, random deflection of the beam centroid due to refraction, introducing fast intensity fluctuations, i.e., scintillation.

When observed over a sufficiently long period, beam wander broadens the apparent beam footprint, so the *long-term* turbulence-induced beam spread has two components: the *short-term* (instantaneous) beam spread and the average beam wander. We quantify this using the Strehl ratio, i.e., the ratio of the on-axis peak intensity of the perturbed beam to that of the diffraction-limited beam. Following Parenti and Sasiela (2005), the Strehl ratio for the long-term beam spread of a collimated beam is

$$S_{bs,bw} = \left[1 + 5.56 \left(\frac{w_0}{r_0} \right)^{5/3} \right]^{-6/5}. \quad (3.9)$$

If beam wander is perfectly compensated, e.g., by ideal closed-loop beam tracking, only the short-term beam spread remains (Parenti and Sasiela 2005):

$$S_{bs} = \left[1 + \left(5.56 - \frac{4.84}{1 + 0.04 (w_0/r_0)^{5/3}} \right) \left(\frac{w_0}{r_0} \right)^{5/3} \right]^{-6/5}. \quad (3.10)$$

Since all losses in the link budget (eq. (3.1)) are defined as positive quantities, the Strehl ratios ($S \leq 1$) are converted to losses in decibels via $L = -10 \log_{10} S$, yielding $L_{bs,bw}$ and L_{bs} , respectively. To isolate the beam wander contribution, we define the beam wander-induced scintillation loss as

$$L_{sci,bw} = L_{bs,bw} - L_{bs}. \quad (3.11)$$

On-Axis Scintillation Loss On-axis scintillation refers to intensity fluctuations at the center of the beam, caused by turbulent eddies of size comparable to the beam. Compared to the previous two losses, the on-axis scintillation loss, $L_{sci,oa}$, depends on the scintillation index. The loss in decibels is given as (Giggenbach 2008)

$$L_{sci,oa} = \left(3.3 - 5.77 \sqrt{-\log_{10} p_{out}} \right) (\sigma_I^2)^{2/5}, \quad (3.12)$$

where p_{out} is the outage probability of the RX detector.

Magnitude comparison The three losses are plotted in fig. 3.4 as functions of r_0 and σ_I^2 to illustrate their magnitude and dependencies on the turbulence parameters. Note that this plot is specific to the collimated beam Photonloop setup as summarized in table 3.1. The left panel shows that the short-term beam spread loss is nearly insensitive to r_0 until r_0 falls below 10 cm, where its magnitude increases rapidly. Beam wander-induced scintillation, on the other hand, has higher relative importance for most of the r_0 range compared to L_{bs} . The on-axis scintillation loss shows high sensitivity even at low σ_I^2 values, meaning that small increases in scintillation lead to larger increases in $L_{sci,oa}$.

3.3.3. GENERALIZED MUTUAL INFORMATION

From the received power, the background noise, and the turbulence-induced losses defined in the previous section, the SNR is obtained via eq. (3.2). To translate

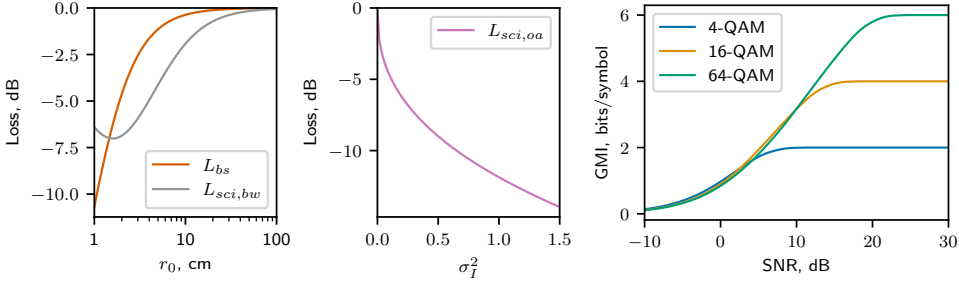


Figure 3.4.: Relation between link losses and (a) Fried parameter r_0 and (b) scintillation index σ_I^2 for idealized Photonloop setup (cf. table 3.1).

Figure 3.5.: Theoretical SNR to GMI depending on QAM modulation. 4-QAM was selected for this study.

3

the SNR into a measure of communication performance, we use the generalized mutual information (GMI), which gives the maximum achievable information rate under bit-interleaved coded modulation for a given modulation format (Caire *et al.* 1998). Figure 3.5 shows the theoretical SNR–GMI relationship for several quadrature amplitude modulation (QAM) orders. Following Vliet *et al.* (2026), we adopt 4-QAM, which reaches a GMI of 2 bits/symbol at an SNR of $\gtrsim 10$ dB. Higher-order QAM formats yield larger peak GMIs but demand correspondingly higher SNR, making them less tolerant to turbulence-induced fading. We prefer GMI to the more common bit-error rate (BER) because GMI directly quantifies the achievable information rate of the channel, whereas BER measures only the fraction of erroneously received bits.

3.4. RESULTS

In the following, we present the results of the WRF-based hindcast of optical turbulence during the Photonloop campaign and compare the WRF-based estimates of OT parameters and FSO metrics against those estimated from the BLS observations. We would like to reiterate that the reference in this work is the path-averaged C_n^2 time series observed by the BLS. The WRF-based C_n^2 estimates are path-averaged using the BLS weighting function (Beyrich *et al.* 2021) to mimic the BLS observations. From both signals, r_0 , σ_I^2 , and all downstream parameters are derived in the same way. Furthermore, a collimated beam is assumed, which is assumed to arrive as a spherical wave at the RX. A future study that considers the real divergent beam and uses measured received free-space power could validate SNR or GMI estimates more directly. Section 3.B gives a summary and an overview of all simplifications and assumptions made in this study.

To quantify the agreement between WRF-based and BLS-based estimates throughout the analysis, we employ three standard metrics. The bias is defined as the mean difference between the WRF estimate x_i and the BLS reference y_i ,

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^N (x_i - y_i), \quad (3.13)$$

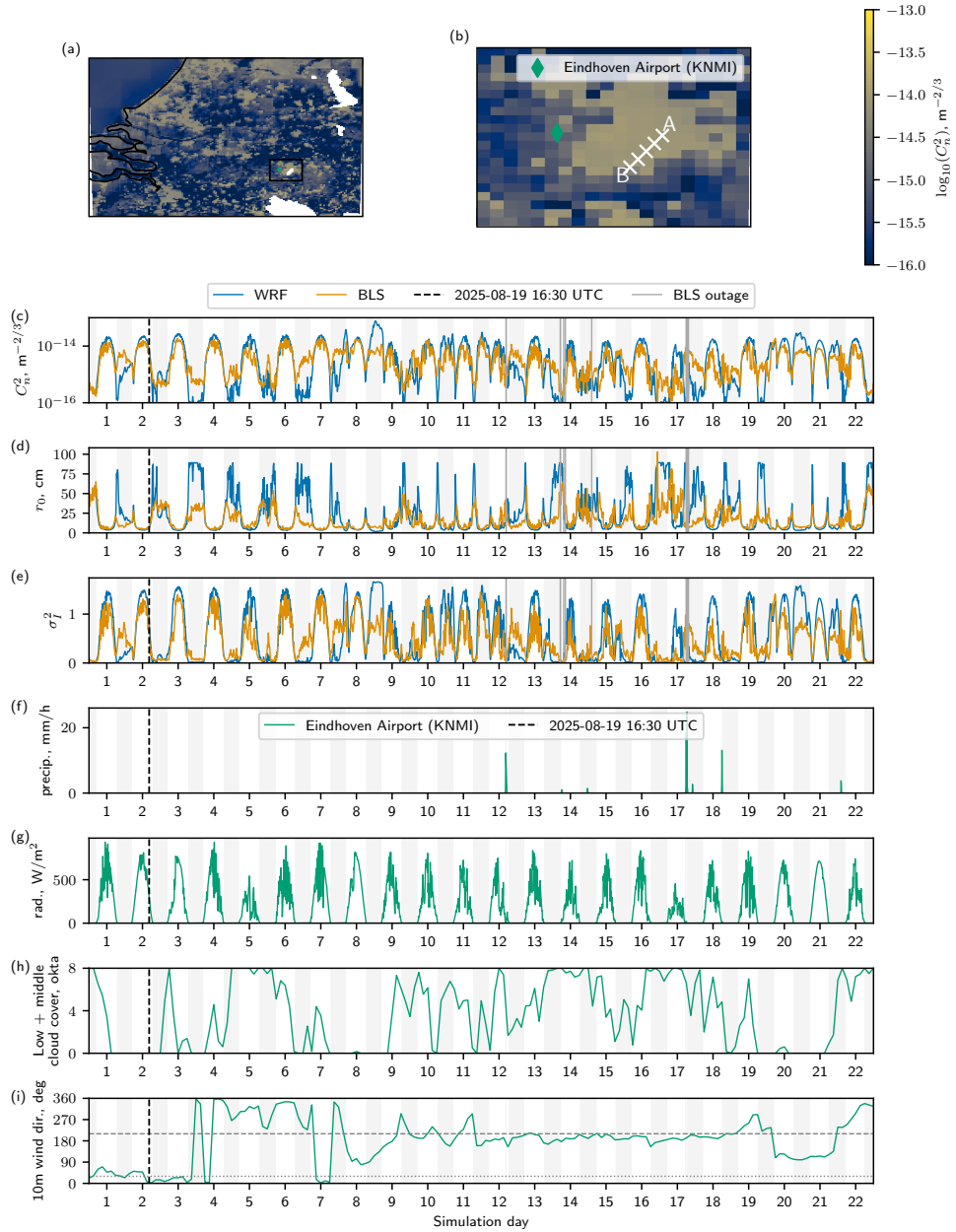


Figure 3.6.: Spatial distribution and temporal evolution of simulated and observed C_n^2 with derived optical turbulence parameters for the Photonloop urban case. Meteorological observations from the nearby weather station (green) are displayed for context. Simulation day markers indicate 12:00Z. Simulation range is 2025-08-18 until 2025-09-09, where panels (a) and (b) correspond to 2025-08-19 at 16:30Z as indicated by the dashed line in the time series.

Table 3.2.: Agreement between WRF and BLS for path-averaged $\log_{10} C_n^2$, Fried parameter r_0 , and scintillation index σ_I^2 . Positive bias values indicate overestimation by WRF

	$\log_{10}(C_n^2)$		r_0 , cm		σ_I^2	
	Day	Night	Day	Night	Day	Night
r	0.73	0.64	0.55	0.53	0.78	0.65
Bias	0.22	-0.31	-2.10	17.82	0.26	-0.01
cRMSE	0.39	0.59	12.97	24.22	0.31	0.37

so that positive values indicate overestimation by WRF. The centered root-mean-squared error (cRMSE) measures the spread of the errors around the bias,

$$\text{cRMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N [(x_i - \bar{x}) - (y_i - \bar{y})]^2}, \quad (3.14)$$

where \bar{x} and \bar{y} denote the respective means. Finally, the Pearson correlation coefficient r quantifies the linear association between the two time series,

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}}, \quad (3.15)$$

with $r = 1$ indicating perfect linear correspondence. All three metrics are reported separately for daytime and nighttime conditions. The comparison begins by assessing the meteorological quality of the WRF-based C_n^2 estimates in section 3.4.1, followed by a discussion of the estimated link losses and performance metrics in section 3.4.2.

3.4.1. METEOROLOGICAL ASSESSMENT OF THE WRF SIMULATION

Figure 3.6 shows an overview of the simulated C_n^2 and compares it against the BLS observations¹. Panels (a) and (b) display the horizontal distribution of C_n^2 in the WRF domain at $1 \text{ km} \times 1 \text{ km}$ resolution, interpolated to a constant height slice of $z_{MSL} = 60 \text{ m}$ above mean sea level (MSL). The white cutouts in panel (a) correspond to terrain heights greater than 60 m MSL. Panel (b) zooms in on Eindhoven and Photonloop, where the terrain height is $\sim 20 \text{ m}$ MSL. Consequently, the horizontal slice locally corresponds to $\sim 40 \text{ m}$ above ground, which is approximately the height at which the Photonloop loop operates. Panel (c) shows the time series of path-averaged C_n^2 estimated from WRF (blue) and observed by the BLS (orange). The corresponding estimated Fried parameter r_0 and scintillation index σ_I^2 evolutions are displayed in panels (d) and (e), respectively. Table 3.2 quantifies the agreement between WRF and BLS for all three OT variables by means of the bias, the centered root-mean-squared error, and the Pearson correlation coefficient r .

¹An animation of this figure showing the temporal evolution of the simulation is available as part of the dataset accompanying this chapter (Pierzyna 2026a).

For meteorological context, panels (f) to (i) of [fig. 3.6](#) show precipitation rate, global solar radiation flux, low and middle cloud cover, and 10 m wind direction measured at a nearby weather station. The time series of cloud cover and wind direction are block-averaged to 3 h steps, rather than the 10 min averages shown in the other panels, to smooth the signal and facilitate more straightforward interpretation. The weather station is operated by the Dutch Royal Meteorological Institute (KNMI) at Eindhoven Airport and is marked in green in panel (b). Note that the airport is located at the outskirts of the city, so the meteorological observations may not always be representative of the urban conditions experienced by the link due to the complex urban boundary layer. On the other hand, the observed BLS outages (grey lines) match well with rain events in panel (f), indicating that the weather station observations are at least qualitatively representative.

The snapshots of horizontal $\log_{10} C_n^2$ distribution in [fig. 3.6\(a\)](#) and (b) show that turbulence is heterogeneously distributed across the entire domain. The bright spot in the center of panel (b) with $C_n^2 \approx 10^{-14} \text{ m}^{-2/3}$ corresponds to Eindhoven. Comparing values along the link path suggests that there is little spatial variation, as the link is fully within the city. WRF consistently assigns the urban land-use class (not shown) to all these cells, explaining the local homogeneity. While panels (a) and (b) only show a single snapshot, this homogeneity is persistent throughout the simulation, which is why we only report the path-averaged C_n^2 in the following.

The time series of C_n^2 in panel (c) shows that the timing of the diurnal cycle between WRF and BLS matches well, as indicated by the location of neutral conditions in the time series. During daytime (white background in the time series panels), the simulated $\log_{10} C_n^2$ matches well on most days with good correlation ($r = 0.73$), but it shows systematic overestimation (positive bias, cf. [table 3.2](#)). The connection between observed radiation and both observed and simulated C_n^2 is evident. On cloudy days, such as days 5 and 17, the discrepancy between WRF and BLS becomes more severe, likely because the limited mesoscale resolution cannot resolve fine cloud fields. This limitation is also reflected in the general behavior, where fluctuations in the radiation signal and in the observed BLS C_n^2 are less pronounced in the WRF simulation.

The daytime C_n^2 overestimations translate differently to the integrated turbulence parameters r_0 and σ_I^2 (panels d and e). The scintillation index σ_I^2 behaves similarly to C_n^2 , showing positive bias during daytime and comparable correlation coefficients (cf. [table 3.2](#)). Due to the theoretical consideration of scintillation saturation (cf. [eq. \(3.8\)](#)), strong C_n^2 overestimations like during night 8/9 are weaker in σ_I^2 . In contrast, the Fried parameter r_0 shows an asymmetric response: while C_n^2 exhibits a clear positive bias during the day, the r_0 daytime bias is almost negligible. This asymmetry arises from the inverse relationship between r_0 and C_n^2 (cf. [eq. \(3.4\)](#)).

In contrast to daytime, nighttime conditions (grey shading background) show lower correlation, higher error, and negative bias in C_n^2 , indicating underestimation (cf. [table 3.2](#)). Qualitatively, observed turbulence during the night is generally high ($\geq 10^{-15} \text{ m}^{-2/3}$), as expected in an urban environment with high heat storage and complex shear due to buildings. However, WRF performance varies considerably between different nights. During some nights, such as nights 1/2, 3/4, and 4/5,

turbulence is underestimated by WRF by one order of magnitude, while other nights, such as nights 9–12, are captured surprisingly well. Notably, during these well-captured nights, turbulence is almost as high as daytime turbulence, with the practical implication of high link losses also during the night.

The nighttime C_n^2 underestimation critically affects r_0 : while C_n^2 exhibits a negative bias at night, r_0 shows a pronounced positive bias. This is problematic because an overly high nighttime r_0 leads to an underestimation of the beam spread loss and the beam wander-induced scintillation loss, which both inversely depend on r_0 . The scintillation index σ_I^2 again follows C_n^2 more closely, exhibiting negative bias at night as well.

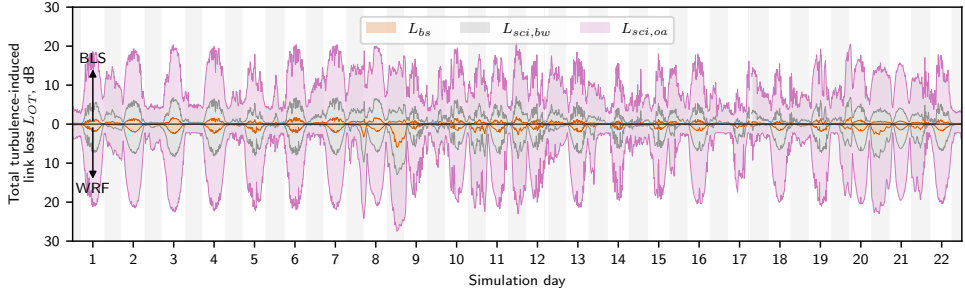
Attribution of the conditions leading to this different nightly turbulence behavior is difficult due to the complex urban environment. Cloud cover (cf. fig. 3.6(h)) likely plays a role by modulating atmospheric stability. Nights 1–3 are mostly clear sky, while nights 9–12 are mostly cloudy. Cloudiness affects surface cooling, with clear skies leading to stronger cooling and thus stronger stable stratification, while cloudy nights result in weaker stratification as clouds trap heat. However, the urban environment complicates this relationship, as buildings also release heat and generate additional shear, thereby increasing turbulence. Wind direction (cf. fig. 3.6(i)) also differs between these periods. Nights 1–3 experience wind from NE (ca. 30° , dotted line), while the wind turns 180° to SW (ca. 210° , dashed line) during nights 9–12. Thus, different fetch areas with different urban structures are located upstream of the link. More detailed disentanglement of these effects and assessment of why WRF captures some conditions better than others is beyond the scope of the present work and left for future research.

Overall, C_n^2 agreement between WRF and BLS during daytime is comparable in terms of scores to previous work (Pierzyna *et al.* 2024), while nighttime shows less agreement as expected due to the urban effects discussed. Additionally, the present study employs a continuous 3-week simulation instead of selected clear-sky benchmark days. Consequently, the present modelling results are considered satisfactory with respect to C_n^2 . The asymmetric error propagation to r_0 and the similar behavior of σ_I^2 highlight that C_n^2 errors do not translate uniformly into the integrated turbulence parameters. The upcoming analysis of link losses and GMI will determine whether performance is satisfactory from a practical FSOC perspective. For completeness, we would also like to refer the reader to a brief discussion about known uncertainties in the BLS measurements (cf. section 3.A). However, as the BLS is widely used and generally accepted as reliable, we assume it adequately represents ground truth here.

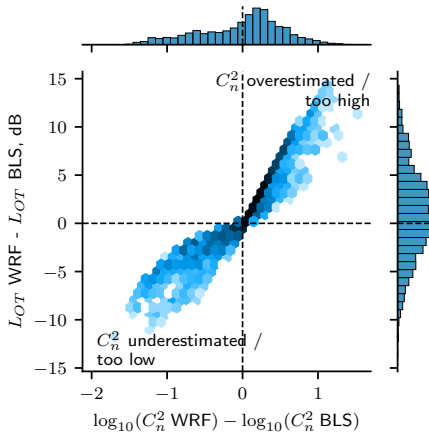
3.4.2. ESTIMATED LINK LOSSES AND LINK PERFORMANCE

Having assessed the meteorological quality of the WRF-based C_n^2 estimates, we now investigate how these estimates translate into turbulence-induced link losses and FSOC performance metrics.

(a) Stacked time series of turbulence-induced losses based on observed BLS C_n^2 (positive y direction) and estimated WRF C_n^2 (negative y direction).



(b) Joint distribution of absolute $\log_{10} C_n^2$ error vs. absolute L_{OT} error



(c) Joint distribution of relative linear C_n^2 error vs. relative L_{OT} error

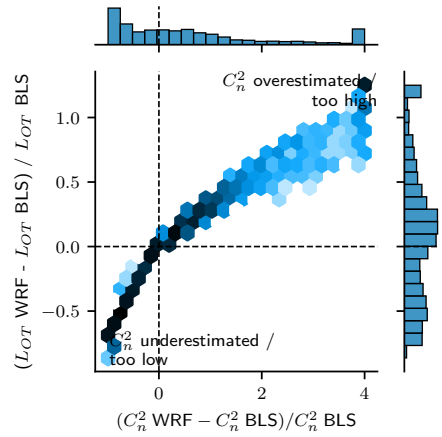


Figure 3.7.: Comparison of turbulence-induced link losses estimated from BLS observations and WRF simulation. Panel (a) depicts temporal evolution of the loss components, while panels (b) and (c) compare the total turbulence-induced link loss error against the C_n^2 estimation error.

Table 3.3.: Scores of agreement between WRF-based and BLS-based turbulence-induced link loss estimates. Positive bias values indicate overestimation by WRF. The metrics for the sum of all turbulence-induced link losses components are given as ΣL_{OT} .

	L_{bs} , dB		$L_{sci,bw}$, dB		$L_{sci,oa}$, dB		ΣL_{OT} , dB	
	Day	Night	Day	Night	Day	Night	Day	Night
r	0.76	0.53	0.78	0.65	0.76	0.67	0.78	0.66
Bias	0.37	0.14	0.91	-0.07	1.51	-1.18	2.80	-1.08
cRMSE	0.43	0.77	1.08	1.39	2.25	2.94	3.62	4.88

COMPARISON OF TURBULENCE-INDUCED LINK LOSSES

The estimated turbulence-induced link losses based on C_n^2 observed by the BLS and C_n^2 estimated from WRF are displayed in fig. 3.7. Panel (a) shows a stacked time series of the individual link loss contributions, which sum to the total turbulence-induced link loss L_{OT} . The BLS-based estimates are shown in the positive y-direction, and the WRF-based ones in the negative y-direction for easier comparison. The background shading again indicates daytime (white) and nighttime (grey). Panels (b) and (c) compare the link loss and C_n^2 estimation errors, where panel (b) shows absolute values and panel (c) relative values. Table 3.3 supports the analysis by quantifying the agreement between WRF-based and BLS-based link loss estimates per component.

Considering the time series in panel (a), the total L_{OT} estimates from WRF and BLS agree in magnitude, but, as expected, the previously observed daytime overestimations of WRF-based C_n^2 now translate into overestimations of the turbulence-induced losses. Note that the static geometric loss of ca. 5 dB for the idealized Photonloop setup is not shown. The dominant turbulence-induced loss component for our setup during both day and night is the on-axis scintillation loss $L_{sci,oa}$. This is consistent with the degressive shape of $L_{sci,oa}$ as a function of σ_I^2 (cf. fig. 3.4), which causes already small increases in σ_I^2 to produce strong increases in scintillation loss, especially at low scintillation indices. Beam spread loss L_{bs} and beam wander-induced scintillation loss $L_{sci,bw}$, on the other hand, are only relevant during strong turbulence experienced mostly during the day. During low-turbulence nights, such as nights 2/3 and 3/4, both contributions are nearly zero. As already evident from fig. 3.4, $L_{sci,bw}$ is stronger than L_{bs} for most of the r_0 range. This is practically relevant because beam wander can be partially compensated by beam tracking (Parenti and Sasiela 2005), thereby significantly reducing this loss component. The maximum combined turbulence-induced loss during daytime reaches ca. 20 dB, which is well matched by the WRF-based estimates.

The agreement between WRF-based and BLS-based loss estimates is quantified in table 3.3. For the total loss L_{OT} , the bias is positive during the day and negative during the night, which aligns with the corresponding biases in $\log_{10} C_n^2$ observed previously (cf. table 3.2). However, the relative magnitude of the biases differs between the two representations. In log space, the $\log_{10} C_n^2$ bias is 0.22 during daytime versus -0.31 during nighttime, so the daytime bias was smaller in absolute value. For the total loss, the daytime bias of 2.80 dB is more than twice the absolute nighttime bias of -1.08 dB, underscoring the importance of accurately capturing daytime turbulence. The centered root-mean-squared error, on the other hand, maintains a comparable ratio between day and night as observed for $\log_{10} C_n^2$. Correlations for total loss during day ($r = 0.78$) and night ($r = 0.66$) are slightly higher but of similar magnitude than those of $\log_{10} C_n^2$, suggesting that the $\log_{10} C_n^2$ correlation provides a reasonable estimate of the expected L_{OT} correlation. From a link-budget perspective, combining the absolute biases and cRMSE values for day and night implies that an additional link margin of ca. 7 dB would be required to accommodate the estimation uncertainty. We note, however, that these numerical values strongly depend on the idealized Photonloop system parameters and would

differ for other setups.

The non-linear loss equations (cf. fig. 3.4) cause C_n^2 errors to propagate differently to the individual loss components. For on-axis scintillation, σ_I^2 exhibits almost zero bias during the night (cf. table 3.2), yet $L_{sci,oa}$ shows a bias of -1.18 dB, which can be attributed to the high sensitivity of this loss to small changes in the scintillation index at low turbulence levels. The opposite effect is observed for L_{bs} and $L_{sci,bw}$: the Fried parameter r_0 is strongly overestimated during nighttime (cf. table 3.2), but since r_0 is generally large at night due to weak turbulence, this overestimation translates into reduced bias and error for both beam spread and beam wander-induced losses.

These non-linear effects are also reflected in panels (b) and (c) of fig. 3.7, which compare the joint distributions of absolute and relative C_n^2 estimation error and total link loss error. Panel (b) suggests an approximately linear relationship between absolute C_n^2 error and absolute L_{OT} error, but with asymmetric slopes. This asymmetry is consistent with the previous discussion: when C_n^2 is overestimated, the losses are overestimated as well, whereas underestimating C_n^2 leads to too low losses. The steeper slope on the overestimation side indicates that overestimating C_n^2 leads to larger absolute loss errors than underestimating it by the same amount. From a practical perspective, however, overestimation is less critical because it leads to overprovisioning of the link margin, whereas underestimation could result in underprovisioning and thus link outages. As C_n^2 overestimations occur mostly during high C_n^2 conditions where losses are large, the link loss errors are expected to be larger than during low turbulence conditions. Therefore, panel (c) presents a relative comparison, showing the L_{OT} error and the linear C_n^2 error normalized by their BLS-based reference values. Here, the joint distribution suggests a degressive trend: small relative underestimations of C_n^2 can lead to a substantially larger relative underestimation of the loss, whereas a comparable relative overestimation of C_n^2 produces a comparably weaker overestimation of the loss. The combined conclusion is that C_n^2 underestimations are relatively more critical than overestimations and remain practically relevant, as the absolute underestimation can reach up to 10 dB, which is half of the total maximum turbulence-induced loss.

In summary, the WRF-based C_n^2 estimates translate into turbulence-induced link losses that retain the sign and diurnal structure of the underlying C_n^2 errors. The non-linear loss equations, however, reshape the error characteristics. Daytime C_n^2 overestimations are amplified in the link loss compared to C_n^2 underestimations, but practical relevance when considering link robustness is low. During nighttime, the sensitivity of the turbulence-induced link loss on C_n^2 errors is higher. While absolute errors for low C_n^2 are smaller than those for high C_n^2 , they remain significant and practically relevant, as underestimating the loss could lead to link outages. As the final step of the analysis, we quantify, in the following, how the communication performance is affected by these link loss errors.

SNR AND GMI ESTIMATES

To assess how the estimated link losses translate into communication performance, we now estimate SNR and GMI for the idealized Photonloop setup. Computing the

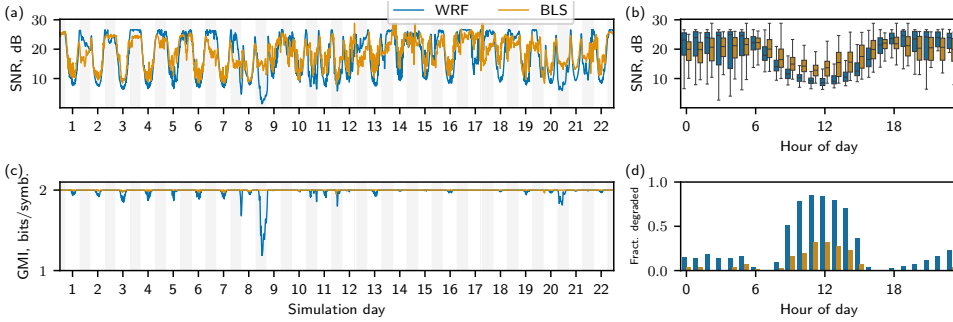


Figure 3.8.: Comparison of SNR and GMI estimated from WRF C_n^2 (blue) and BLS C_n^2 (orange). Panels (a) and (c) show time series; panel (b) compares distributions of SNR aggregated by hour of the day; panel (d) presents distribution of hourly ratios of degraded GMI rates where $\text{GMI} < \text{max. GMI}$.

SNR from the link budget (cf. eqs. (3.1) and (3.2)) requires the static geometric loss, the transmitted power, and the receiver background noise in addition to the turbulence-induced losses discussed above. Although we simplified the Photonloop setup for this study, we aim to remain close to the conditions reported by Vliet *et al.* (2026) for qualitative comparison. Because the divergent beam of the original setup incurs ca. 20 dB more geometric loss than our assumed collimated beam, we artificially add this extra loss to the SNR computation. Additionally, we set the background noise to $P_n = -25$ dBm to reproduce the ca. 20 dB of electrical SNR reported by Vliet *et al.* (2026) during weak turbulence. We acknowledge that these choices may appear arbitrary, but note that the system losses not considered in this study would add further contributions to the link budget in practice. Moreover, the extra geometric loss and noise enter the SNR additively in log space and therefore cancel when comparing WRF-based and BLS-based estimates. The qualitative conclusions drawn below, in particular when the SNR drops below the GMI saturation threshold (cf. fig. 3.5), are thus largely insensitive to the exact values chosen.

The resulting SNR and GMI estimates are shown in fig. 3.8, where GMI is computed following Alvarado *et al.* (2018) assuming 4-QAM modulation as used by Vliet *et al.* (2026). Panels (a) and (c) show the time series of SNR and GMI, respectively, while panel (b) compares the distribution of SNR values aggregated by hour of the day. Panel (d) illustrates the distribution of hourly fractions of degraded GMI rates below the maximum achievable 2 bits/symbol for 4-QAM modulation. Since the SNR with static P_{TX} , static L_{geo} , and static P_n scales only with L_{OT} , the previous discussion about the total turbulence-induced link loss applies directly to the SNR time series in panel (a) and is not repeated. The SNR data are shown primarily to aid the interpretation of the GMI data in panels (c) and (d).

Comparing GMI and SNR reveals notable differences due to the non-linear relationship between the two (cf. fig. 3.5). Because only SNR values below ~ 10 dB lead to degraded GMI rates, the GMI time series shows performance drops

exclusively during strong turbulence conditions, i.e., mostly during the day. As daytime overestimation of WRF-based C_n^2 leads to lower SNR, the WRF-based GMI exhibits more frequent degradation than the BLS-based estimates. This relative difference holds as long as the assumed static SNR contributions discussed above bring the overall SNR close to the GMI saturation threshold. The distribution of degraded fractions per hour in panel (d) provides a more quantitative view. The BLS-based estimates show a peak of degraded samples during midday, aligning qualitatively with the ca. 30% reported by Vliet *et al.* (2026), who observed less than 10% degradation during night. The WRF-based estimates, in contrast, predict substantially higher degradation rates, with almost 100% of samples degraded during midday. Despite the higher frequency of degradation, the WRF-based SNR remains close to the saturation threshold for most of these episodes, so the actual data rate drops only moderately in most cases, as evident from the GMI time series in panel (c). The transition from no degradation to degradation around 7:00Z and 17:00Z is well captured by WRF, with the morning minimum also evident in the observations of Vliet *et al.* (2026).

In conclusion, the large differences in GMI estimates between WRF and BLS highlight the need to improve C_n^2 estimates, especially under strong-turbulence conditions where the SNR approaches the GMI saturation threshold. The plateauing of GMI at high SNR values, on the other hand, suggests that moderate C_n^2 errors during weak turbulence are less critical or even negligible from a practical FSOC perspective. When relying solely on simulations without site-survey measurements, additional link margin should be provisioned to accommodate these uncertainties. Future comparisons to measured free-space received power are also needed for a more rigorous end-to-end validation.

3.5. CONCLUSION

This study presented a framework to estimate turbulence-induced link losses and communication performance metrics from a time series of C_n^2 for a free-space optical communication (FSOC) link utilizing a collimated beam. By obtaining C_n^2 from numerical mesoscale simulations for link performance estimates, we take a first step toward basing FSOC system design and performance assessment on site-specific atmospheric conditions. The example setup used in this study is inspired by the Reid Photonloop, a 4.6 km urban FSOC link in Eindhoven, Netherlands. Compared with the real Photonloop setup, we simplified the system by assuming a collimated beam and neglecting communication-system losses. For this collimated-beam version of the link, we demonstrated that losses due to beam spread, beam wander, and on-axis scintillation can be estimated from C_n^2 and that these estimates translate into estimates of the signal-to-noise ratio (SNR) and generalized mutual information (GMI). The Weather Research and Forecasting (WRF) model is employed to simulate the atmospheric conditions and estimate C_n^2 along the link path using a variance-based parameterization (He and Basu 2015). Observed C_n^2 from a collocated boundary layer scintillometer (BLS) serves as a reference for validation.

The comparison of WRF-based and BLS-based C_n^2 shows that the mesoscale

model captures the diurnal timing of optical turbulence well. During daytime, C_n^2 is reproduced with good correlation but exhibits systematic overestimation, with performance degrading on cloudy days where the 1 km model resolution likely cannot capture local cloud structure well. During nighttime, the simulation tends to underestimate C_n^2 , likely reflecting the difficulty of representing urban heat storage and shear generation at mesoscale resolution. The propagation of these C_n^2 errors to the integrated turbulence parameters r_0 and σ_I^2 is non-trivial. The scintillation index σ_I^2 generally follows the error characteristics of C_n^2 in sign and relative magnitude, with extreme overestimations dampened by the theoretical saturation of scintillation at high turbulence levels. In contrast, the Fried parameter r_0 responds asymmetrically: the daytime C_n^2 overestimation barely affects r_0 due to its inverse relationship with C_n^2 , while the nighttime underestimation produces a pronounced positive bias in r_0 that leads to underestimation of beam spread and beam wander losses.

When these integrated parameters are translated into turbulence-induced link losses, the on-axis scintillation loss $L_{sci,oa}$ emerges as the dominant dynamic loss component for both day and night. Its high sensitivity at low scintillation indices implies that even small increases in σ_I^2 lead to substantial losses. The beam-wander-induced scintillation loss exceeds the beam-spread loss under most turbulence conditions, whereas both contributions nearly vanish in weak nighttime turbulence. For the idealized Photonloop setup, the combined turbulence-induced loss reaches up to ca. 20 dB during strong daytime turbulence. The daytime overestimation of C_n^2 is amplified into larger absolute loss overestimations, but this is less critical in practice, as it corresponds to overprovisioning the link budget, i.e., a conservative design. A too conservative design, however, may increase costs, so improving C_n^2 estimates under strong-turbulence conditions remains desirable. The nighttime underestimation, while producing smaller absolute loss errors, is more concerning from a system design perspective. Too low losses could lead to an overly optimistic link budget and thus unexpected outages.

At the level of communication performance, the non-linear relationship between SNR and GMI introduces additional complexity. Because GMI saturates at its maximum value across a wide SNR range, only SNR values below a modulation-specific threshold result in degraded data rates. Consequently, if the link budget provides sufficient margin and the SNR remains well above the GMI threshold, even substantial C_n^2 errors have a negligible impact on the estimated communication performance. However, when the SNR approaches the threshold, e.g., during strong daytime turbulence, errors in C_n^2 can become significant again, as exhibited in our example case. In our example, the BLS-based GMI estimates show degraded performance for ca. 30% of midday samples, consistent with the observations of Vliet *et al.* (2026). In contrast, the WRF-based estimates predict degradation for nearly 100% of midday samples due to the overestimated losses. We stress, however, that entering the degradation regime does not imply link interruption but rather a temporary reduction in achievable data rate.

The assumption of a collimated beam in this study is a deliberate simplification to keep the loss equations in closed form and to maintain interpretability. Crucially, this

choice does not affect the general setup of the framework, as the collimated-beam assumption enters through the loss equations that translate r_0 and σ_I^2 into link losses. Because these equations differ in functional form for other beam geometries, the specific non-linear error propagation characteristics reported here may change and should be re-evaluated in future work. For example, for the divergent beam employed by the real Photonloop, larger geometric losses are likely incurred, but sensitivity to beam wander may weaken due to the broader beam footprint in the receiver plane. Additional effects arising from the curved wavefront of a diverging beam are not covered by the formalism of Parenti and Sasiela (2005) and require further investigation. The on-axis scintillation loss, however, is expected to remain applicable as the model of Giggenbach (2008) is not limited to plane or spherical waves. From a practical standpoint, the second-largest loss component, the beam wander-induced scintillation, can be (partially) mitigated through tip-tilt correction, as already implemented in the real Photonloop setup. Our estimates without tip-tilt correction, therefore, provide a conservative baseline. Finally, in the present study, we did not consider the effect of turbulence on the phase of the wavefront and its impact on link performance, which should be addressed in future work. A systematic overview of all assumptions and simplifications, and suggestions for how to address them in the future, is provided in [section 3.B](#).

In conclusion, the framework presented in this study is a first step toward meteorology-informed site-specific FSOC system design. To our knowledge, it provides the first end-to-end example connecting mesoscale atmospheric simulations to communication performance metrics for a realistic urban FSOC link. The overarching finding is that evaluating only the agreement in C_n^2 , as commonly done in the literature, is insufficient to assess the operational performance of the FSOC link in question. Although our simplifications limit comparability with the Photonloop campaign, the framework contains all key ingredients: realistic meteorological conditions, an observed C_n^2 baseline, a representative link geometry, and a transparent path from C_n^2 to SNR and GMI. Opportunities for improvement exist at all stages, from enhanced urban boundary layer modelling and more accurate C_n^2 parameterizations to optical propagation theory for realistic divergent beams and direct validation against measured free-space received power. We therefore consider this work a starting point for future research toward the operational use of numerical weather prediction for FSOC system engineering.

ACKNOWLEDGEMENTS

This publication is part of the project FREE – Optical Wireless Superhighways: Free photons (at home and in space) (with project number P19-13) of the research programme TTW-Perspectief, which is (partly) financed by the Dutch Research Council (NWO). MP thanks Oscar Hartogensis (Wageningen University, NL) for guidance on selecting the scintillometer deployed at the Photonloop site. MP thanks Vincent van Vliet (TU Eindhoven, NL) for setting up the scintillometer, sharing C_n^2 observations and characteristics of the link setup, and for countless discussions on FSOC performance metrics.

DATA AVAILABILITY

The meteorological observations at Eindhoven Airport are publicly available through the data platform of the Royal Netherlands Meteorological Institute: <https://dataplatform.knmi.nl/>. The C_n^2 observations from the scintillometer are not publicly available but can be shared upon reasonable request. The WRF simulation data and codes to reproduce the analysis are available on Zenodo (Pierzyna 2026a).

3.A. A NOTE ON BLS UNCERTAINTIES

In this work, C_n^2 observed by a Scintec BLS900 neo boundary layer scintillometer (BLS) is used as ground truth to evaluate WRF-based C_n^2 estimates. In the following, we briefly discuss some known limitations and uncertainties of this instrument, which could be responsible for some of the disagreement between the two signals discussed in section 3.4.1.

Van Kesteren *et al.* (2015) provide a detailed discussion of BLS uncertainties based on long-term C_n^2 time series with explicit consideration of the Scintec BLS 900 (non-neo). The authors report over- or underestimation of C_n^2 depending on atmospheric conditions. First, underestimation of C_n^2 of up to 5% can occur when crosswind is low, i.e., when the wind blows along the scintillometer path. As the BLS path and the Photonloop path are coincidentally almost perfectly aligned with the wind directions observed during the campaign (dotted and dashed horizontal lines), this effect might play a role.

Furthermore, scintillation measured by the BLS saturates under strong turbulence conditions ($\sigma_I^2 > 1$), leading to decoupling of C_n^2 and the observed scintillation (Andrews and Phillips 2005). Consequently, the derived C_n^2 is underestimated during periods of strong turbulence (Van Kesteren *et al.* 2015), e.g., potentially on clear sky days. Turbulence strength is expected to be generally higher in urban environments, so saturation is likely to occur more often than in rural environments. To partially compensate for saturation, the correction by Clifford *et al.* (1974) is applied in data processing (Scintec 2023), which makes it difficult to quantify the severity of the underestimation here without a more detailed investigation.

Conversely, overestimating C_n^2 of up to 30% is possible when wind shear is low, such as during strong stratification at night (clear sky, low wind) or free convection during the day (clear sky, low wind) (Van Kesteren *et al.* 2015). Again, the complex urban environment makes it difficult to attribute specific conditions to observed discrepancies between BLS and WRF.

While the above considerations serve as a disclaimer, we would like to stress that the BLS is widely used and generally accepted as reliable, so we assume it adequately represents ground truth here. Additionally, unlike meteorological field campaigns that aim for the best possible measurement accuracy, the setup of the FSO links and supporting BLSs is dictated by communication requirements, so some measurement uncertainties may be unavoidable.

3.B. SUMMARY OF SIMPLIFICATIONS AND ASSUMPTIONS

Throughout this study, several simplifications compared to the real Photonloop setup are made to maintain analytical tractability and generalizability. The simplifications, their effects, and future directions to improve the realism of the end-to-end framework are summarized below.

Beam geometry

- **Collimated beam** ($\theta_{div} \approx 20 \mu\text{rad}$ instead of $\approx 200 \mu\text{rad}$): Enables closed-form loss expressions (Parenti and Sasiela 2005; Giggenbach 2008). Affects geometric loss (both static) and functional form of beam spread/wander losses. Error propagation characteristics are beam-specific.
 - Future: extend to diverging beams and validate against free-space received power.
- **Phase effects neglected**: Focus on amplitude (scintillation) only.
 - Future: include phase for coherent detection or adaptive optics evaluation.

System components

- **System losses neglected**, $L_{sys} = 0$: Omit coupling, detector, and component losses to focus on the atmospheric path only. Prevents direct comparison with fiber-coupled power but cancels when comparing WRF vs. BLS.
 - Future: incorporate for operational design validation.
- **Tip-tilt correction not modeled**: Real setup has active tracking. Our estimates are conservative (no $L_{sci,bw}$ reduction). Real link performs better, especially during daytime.
 - Future: Reduce beam wander loss for more realistic, but less conservative estimates.
- **BLS C_n^2 as reference**: Use of BLS instead of measured power (which includes tip-tilt and system effects). Provides independent atmospheric baseline. Validates propagation model, not end-to-end system.
 - Future: validate against received free-space power before fiber coupling.
- **Artificial noise**, $P_n = -25 \text{ dBm}$: Set to match SNR by Vliet *et al.* (2026). Affects absolute GMI but cancels in WRF vs. BLS comparison. General conclusions on error propagation are unaffected.
 - Future: use measured noise for quantitative predictions.

Numerical modeling

- **Urban simulation at 1 km resolution**: No urban-specific options enabled in WRF. Cities are treated as a separate land-use class with higher heat storage and friction. Buildings/streets are not resolved. Consequently, nighttime C_n^2 is underestimated when urban effects dominate.
 - Future: use more advanced urban schemes (cf. Skamarock *et al.* (2021)) or run high-resolution large eddy simulations (LES).

Atmospheric

- **Absorption and scattering neglected**, $L_{atm} = 0$: Valid for clear-sky campaign period at 1550 nm over 4.6 km. Brief rain events resulted in complete BLS outages and were neglected.
→ Future: include rain, fog, or pollution from observations or models for more realistic absorption and scattering loss estimates.
- **Constant C_n^2 along path**: Simplifies integrals to algebraic expressions. Reasonable as BLS-weighting function is used to average WRF-based C_n^2 to mimic BLS observations. Also, WRF showed low along-path variation because of homogenous land use along the path (cf. fig. 3.2).
→ Future: for wave-optics simulations or when higher resolution models used (e.g., LES), use full C_n^2 evolution along path in all equations.

II

Machine learning modeling – the future?

4

OTClIM: Generating a near-surface climatology of optical turbulence strength (C_n^2) using gradient boosting

The numerical approach explored in the first part of this thesis is computationally expensive, sensitive to model configuration, and limited in environments that are difficult to represent at mesoscale resolution. This study takes a different approach. We introduce OTClIM (Optical Turbulence Climatology using Machine learning), a novel approach for deriving comprehensive climatologies of atmospheric optical turbulence strength (C_n^2) using gradient boosting machines. OTClIM addresses the challenge of efficiently obtaining reliable site-specific C_n^2 climatologies near the surface. Using gradient boosting machines and global reanalysis data, OTClIM extrapolates one year of measured C_n^2 into a multi-year time series. We assess OTClIM's performance using C_n^2 data from 17 diverse stations in New York State, evaluating temporal extrapolation capabilities and geographical generalization. Our results demonstrate accurate predictions of four held-out years of C_n^2 across various sites, including complex urban environments, outperforming traditional analytical models. Non-urban models also show good geographical generalization compared to urban models, which capture non-general site-specific dependencies. A feature importance analysis confirms the physical consistency of the trained models. OTClIM's ability to derive reliable C_n^2 climatologies from just one year of observations can potentially reduce resources required for future site surveys or enable studies for additional sites with the same resources.

This chapter is published as: M. Pierzyna *et al.* (2025). "OTClIM: Generating a near-Surface Climatology of Optical Turbulence Strength (C_n^2) Using Gradient Boosting". In: *Artificial Intelligence for the Earth Systems* 4.2. DOI: 10.1175/AIES-D-24-0076.1

4.1. INTRODUCTION

Atmospheric optical turbulence is highly relevant for optical ground-based astronomy and future free-space optical communication (FSOC). Both applications suffer from light getting distorted when propagating through the turbulent atmosphere. In astronomy, turbulent fluctuations of the atmospheric refractive index, known as optical turbulence (OT), cause blurry images and limit the detection of small objects (Hardy 1998). FSOC links, which use optical beams to transmit data instead of traditional radio waves, experience reduced data rates or even link interruptions due to OT (Kaushal and Kaddoum 2017; Jahid *et al.* 2022). Therefore, the optical turbulence strength (C_n^2 , where index n denotes the refractive index) must be carefully considered in the design and operation of such optical systems. This requires robust statistical evaluation of C_n^2 over time for the operational sites of interest, known as a C_n^2 climatology. Such a climatology, derived from long-term C_n^2 data, portrays trends, seasonal variations, or potential anomalies in OT strength.

Obtaining C_n^2 climatologies is challenging because turbulence strongly depends on the local topography and varying meteorological conditions. That is why long-term on-site surveys measuring local OT conditions are crucial. For instance, before a new optical observatory is built, OT strength is typically measured for multiple years at a few carefully selected locations to identify the best one (see, e.g., Hill *et al.* (2006) and Schöck *et al.* (2009)). When envisioning a global FSOC network, site surveys have so far been focused on cloud cover (see, e.g., Fuchs and Moll (2015), Poulernard *et al.* (2015), and Pham *et al.* (2023)), but similar site surveys targeting OT are expected to become highly relevant in the future.

However, conducting long-term site surveys at various locations of interest is time-consuming and resource-intensive. While C_n^2 can be obtained by post-processing numerical weather model output (e.g., Masciadri *et al.* (1999)), running such models for multiple years is computationally expensive, and the accuracy of the resulting C_n^2 is very sensitive to the model configuration (Pierzyna *et al.* 2023b) and the selected C_n^2 parameterization (Pierzyna *et al.* 2024). We address these issues by proposing a novel machine learning-based (ML) approach called OTCLiM (Optical Turbulence Climatology using Machine learning). OTCLiM aims to extrapolate just one year of measured near-surface C_n^2 into a multi-year time series, enabling the generation of a comprehensive site-specific C_n^2 climatology with less data than a conventional site survey. Our approach does not yet provide full vertical C_n^2 profiles as often sought in astronomy or FSOC. But for near-surface conditions, which are highly relevant for, e.g., horizontal near-surface FSOC links, site survey costs can potentially be reduced, climatologies can be obtained faster, and more sites can be surveyed within a given time frame. By leveraging ML, OTCLiM can also model complex input-output relations, offering an advantage over traditional empirical C_n^2 models, which are discussed in the following section.

This study assesses OTCLiM's performance using an extensive C_n^2 dataset containing measurements from 17 diverse stations. We evaluate the temporal extrapolation capabilities of the trained models and analyze their potential for geographical generalization to other sites. Additionally, we examine the importance of each input variable for predicting C_n^2 to probe if the ML models have learned physically

plausible relations. That analysis increases confidence in the ML models and could uncover new dependencies of C_n^2 from the data.

4.2. RELEVANT C_n^2 REGRESSION STUDIES

The experimental determination of C_n^2 is challenging because it requires expensive instruments, such as scintillometers (Beyrich *et al.* 2021) or meticulous post-processing of high-frequency temperature measurements (Beason *et al.* 2024). Consequently, there has been a sustained effort over several decades to develop parameterizations of C_n^2 using more readily available meteorological variables. Conventional approaches to C_n^2 parameterization typically involve the application of physics-based analytical equations in conjunction with empirically derived coefficients or regression functions. A popular example is the formulation proposed by Wyngaard *et al.* (1971), based on Monin-Obukhov similarity theory (MOST) (Monin and Obukhov 1954). That parameterization – called W71 in the following – relates sensible heatflux $\overline{w'\theta'}$ and friction velocity u_* to the strength of temperature fluctuations (C_T^2) at height z as

$$C_T^2 = \left(-\overline{w'\theta'}/u_* \right)^2 z^{-2/3} g(\zeta). \quad (4.1)$$

Sensible heat flux $\overline{w'\theta'}$ and friction velocity $u_* = \left(\overline{u'w'^2} + \overline{v'w'^2} \right)^{1/4}$ describe the vertical (i.e., surface-normal) turbulent transport of heat (θ') and momentum (u' , v') due to fluctuations of the vertical wind component (w') (Stull 1988). As these parameters capture the two effects modulating turbulence – buoyancy (sensible heat flux) and wind shear (friction velocity) –, they are commonly used in turbulence parameterizations. Since refractive index fluctuations (C_n^2) are driven by density fluctuations due to temperature and moisture fluctuations, C_n^2 is linked to C_T^2 as (Wesely 1976; Moene 2003)

$$C_n^2 = \left(A \frac{P}{\bar{T}^2} \right)^2 \left(1 + \frac{0.03}{\beta} \right)^2 C_T^2 \quad (4.2)$$

with $A \approx 7.9 \times 10^{-5} \text{ KhPa}^{-1}$ for optical wavelengths (Andreas 1988), station pressure P in hPa, and mean air temperature \bar{T} in K. The contribution of moisture fluctuations to C_n^2 is captured by the Bowen ratio β , the ratio of the sensible heat flux and the latent heat flux due to evaporation. If C_n^2 is estimated above the boundary layer in the free atmosphere, a modified formulation of eq. (4.2) can be used (Cherubini and Businger 2013).

Function $g(\zeta)$ in eq. (4.1) is a similarity function, empirically determined from observations,

$$\begin{aligned} g(\zeta) &= 4.9(1 - 6.1\zeta)^{-2/3}, \zeta < 0 \\ g(\zeta) &= 4.9(1 + 2.2\zeta^{2/3}), \zeta \geq 0 \end{aligned} \quad (4.3)$$

with stability parameter $\zeta = z/L$ where L is the Obukhov length $L = -u_*^3 \bar{T} / (\kappa g \overline{w'\theta'})$ and z the height above ground. In L , $g = 9.81 \text{ m s}^{-2}$ is the gravitational acceleration

of Earth and $\kappa = 0.4$ is the Von Kármán constant. Various alternative regression-based similarity functions $g(\zeta)$ were proposed in literature aimed at improving the accuracy and applicability of C_n^2 estimates across diverse meteorological and topographic conditions (see, e.g., Savage (2009) for an extensive review).

In recent years, different machine learning (ML) techniques have been utilized to derive C_n^2 parameterizations. Similar to W71, ML-based approaches aim to obtain regression models that estimate C_n^2 from routine meteorological variables (temperature, pressure, wind speed), gradients (potential temperature gradient or wind shear), or heat and radiation fluxes. However, instead of deriving physics-based functional expressions (cf. eqs. (4.1) and (4.2)), which are fitted to observations (cf. eq. (4.3)), ML directly models the relation between observed C_n^2 and the input variables. The power of such ML regression models is that they can model complex multi-variate relations, yielding potentially better C_n^2 estimates compared to traditional approaches.

The ML-based C_n^2 parameterizations in literature differ in their input variables, the ML regression technique employed, and if C_n^2 is measured at a single or multiple vertical levels. Much work has been devoted to single-level near-surface C_n^2 obtained as point measurements (e.g., from sonic anemometer) or via path-averaging (e.g., from scintillometer). Wang and Basu (2016), for example, used fully connected neural networks to estimate single-level C_n^2 from routine meteorological data and gradients. Vorontsov *et al.* (2020) also utilized neural networks but aimed at estimating path-averaged C_n^2 from images of a received laser beam that is broken up into speckles due to turbulence. Jellen *et al.* (2020) and Jellen *et al.* (2021) employed random forest and gradient boosting machines (GBM) to estimate path-averaged C_n^2 in a maritime environment from routine parameters and radiation fluxes. Most ML-based models do not yield analytical equations of the fitted model, unless they explicitly target it as shown by Bolbasova *et al.* (2021). The authors obtained a complex site-specific equation for single-level C_n^2 , similar to non-ML empirical models proposed by, e.g., Sadot (1992), Wang *et al.* (2015), or Arockia Bazil Raj *et al.* (2015). The difference between these empirical models and MOST-based models is that they are not derived physically but aim to capture the local optical turbulence strength well. Going beyond single-level C_n^2 estimates, Pierzyna *et al.* (2023a) proposed a physics-inspired framework to derive a data-driven similarity theory of optical turbulence that combines C_n^2 observations from multiple levels within the surface layer into a single non-dimensional model. In contrast to the near-surface studies, Su *et al.* (2021) proposed a method to estimate a vertical C_n^2 profile by modeling OT strength as a function of radiosonde measurements. Cherubini *et al.* (2021) presented an ML framework to estimate seeing, which is the vertical integral of the C_n^2 profile, close to the surface and in the free atmosphere, but does not explicitly resolve vertical levels. Milli *et al.* (2020) tackled yet another estimation angle by using ML to forecast future OT conditions for the next two hours based on past observations and auxiliary features. However, all these approaches require in-situ data, rendering them unsuitable for temporal extrapolation as targeted in the present study. Additionally, in-situ-based models need sensors deployed at the site of interest for C_n^2 estimation, which hinders their applicability to sites without

instrumentation. To alleviate this issue, we proposed OTCliM, which parameterizes C_n^2 based on reanalysis data, which are available globally and for multiple decades into the past.

4.3. METHODOLOGY

Our proposed OTCliM¹ approach aims to extrapolate one year of observed C_n^2 to multiple years to obtain a robust statistical description of OT strength at a particular site. An overview of OTCliM is given in fig. 4.1, which we base on the measure-correlate-predict (MCP) framework popular in wind energy (Carta *et al.* 2013; Kartal *et al.* 2023). In the first step, C_n^2 is measured for one year (golden yellow) and then correlated to a reference dataset. For OTCliM, we utilize variables from the ERA5 reanalysis (Hersbach *et al.* 2020) extracted at the location of the C_n^2 measurements as reference data and gradient boosting machines (GBM) to regress the ERA5 time series to C_n^2 where they overlap (blue). In the second step, the trained model is utilized to extrapolate C_n^2 based on ERA5 to multiple years (prediction step in MCP), which can then be used to obtain site-specific seasonal statistics as presented in step 3.

We utilize GBMs for the regression step (cf. section 4.3.1) because they are known for their high performance in non-linear tabular regression tasks. For comparison, we also train GBM models that use in-situ observations instead of ERA5 as input data, which serve as performance baselines for OTCliM. Similarly, the traditional W71 model is evaluated with ERA5 inputs to provide a second baseline (cf. section 4.3.2). To ensure our GBM models capture physically reasonable dependencies, we quantify the feature importance assigned to each ERA5 input variable (cf. section 4.3.3). Finally, the performance of the trained models is assessed in terms of their temporal extrapolation and geographical generalization capabilities (cf. section 4.3.4).

4.3.1. GRADIENT BOOSTING REGRESSION

Following the MCP framework presented in fig. 4.1, we aim to train one ML model for one site using training data from one year. The corresponding training data are the input data $\mathbf{X}_{s,t} \in \mathbb{R}^{n \times p}$ and the target vector $\mathbf{y}_{s,t} \in \mathbb{R}^n$, where $s \in \mathcal{S}$ is one out of multiple sites \mathcal{S} and $t \in \mathcal{T}$ is one out multiple training years \mathcal{T} available at that site. The ML task is to learn a regression function $\hat{f}_{s,t}$ that approximates a sample $y_i \in \mathbf{y}_{s,t}$ from the target vector based on a sample $\mathbf{x}_i \in \mathbf{X}_{s,t}$ of the input data: $\hat{f}_{s,t}(\mathbf{x}_i) = \hat{y}_i \approx y_i$. $\mathbf{X}_{s,t}$ and $\mathbf{y}_{s,t}$ contain a matching number of n samples, i.e., the samples of the one-year time series, where $\mathbf{X}_{s,t}$ is composed of p features, i.e., concurrent time series of different meteorological variables, and $\mathbf{y}_{s,t}$ contains the scaled $\log_{10} C_n^2$.

The OT strength C_n^2 is \log_{10} -transformed because it varies over multiple orders of magnitude throughout the day, which is challenging to capture in non-log space. Also, the range of yearly C_n^2 variation varies between sites, so $\log_{10} C_n^2$ is scaled before training to make the performance scores and feature importance values

¹The Python code for training and the trained models are available on GitHub: <https://github.com/mpierzyna/otclim>.

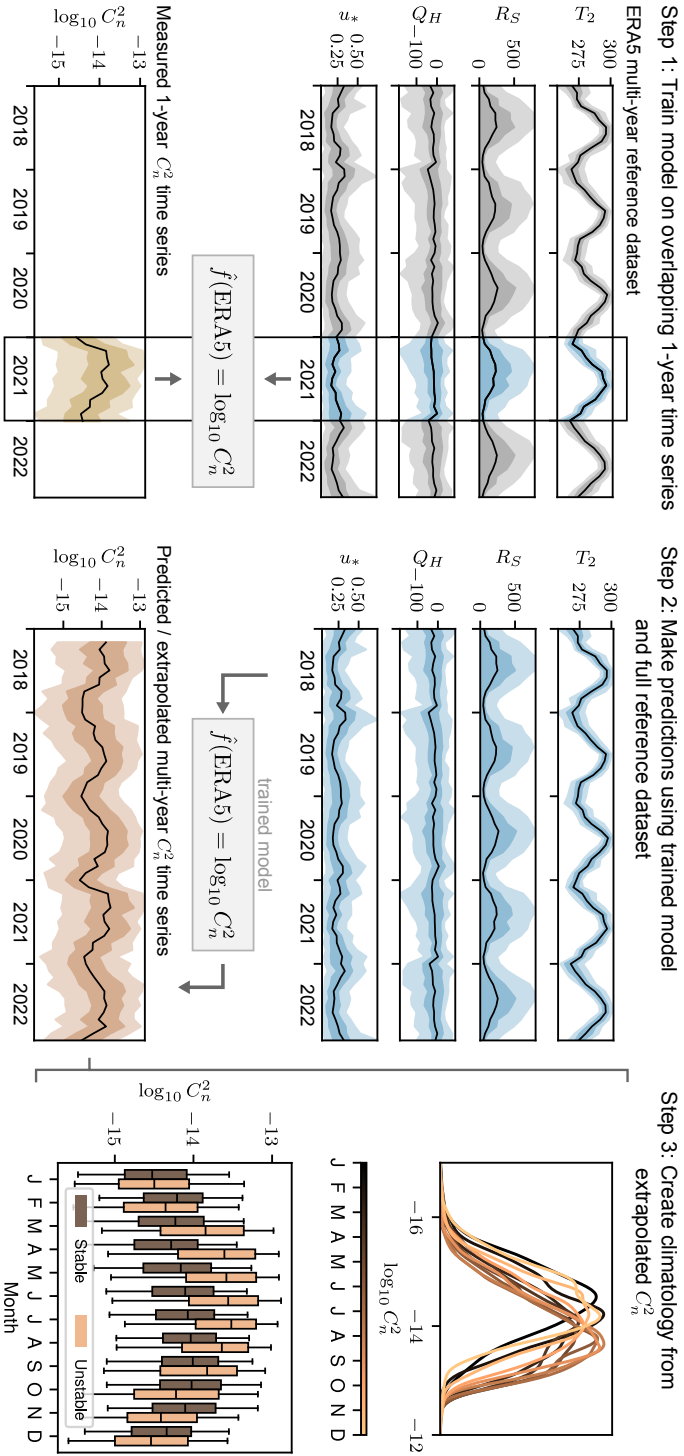


Figure 4.1.: Proposed OTCLiM approach to extrapolate a measured 1-year time series of optical turbulence strength (golden yellow) to multiple years (orange) based on ERA5 reference data (blue). Robust yearly C_n^2 statistics can be obtained from the extrapolated data.

comparable between sites. More details about the C_n^2 pre-processing are presented in section 4.4.1.

The GBM regression models $\hat{f}_{s,t}$ are trained using the AutoML library FLAML (Wang *et al.* 2021). FLAML is a time-constrained hyperparameter optimizer that aims to find the optimal GBM model configuration within a specified time budget. FLAML optimizes not only the hyperparameters of a single algorithm but also explores switching between algorithms, here, the two popular algorithms XGBoost (Chen and Guestrin 2016) and LightGBM (Ke *et al.* 2017). The internal loss functions of XGBoost and LightGBM are also treated as hyperparameters to be optimized by FLAML with L_1 -norm, L_2 -norm, and Huber norm as options. The hyperparameter optimization uses 5-fold cross-validation based on the training data and aims to minimize the root-mean-squared error between the predicted and hold-out \mathbf{y} . Each model is trained for 45 min on 8 CPU cores of an Intel Xeon 2648R, resulting in 6 core hours per model.

4.3.2. BASELINE MODELS

We employ two baseline models to put the performance of OTClIM into perspective. The first model is the traditional empirical W71 model as given by eqs. (4.1) and (4.3). Since the similarity function $g(\zeta)$ does not contain site-specific model coefficients, the accuracy of the W71 C_n^2 estimates is expected to be lower than the site-specific OTClIM models. Secondly, we utilize GBM models trained with in-situ observations as input instead of reanalysis. Reanalysis datasets are coarse-resolution model outputs (spatial resolution typically $\mathcal{O}(10\text{km})$), which are expected to miss some local and complex processes and patterns modulating turbulence. Such processes would be contained in the in-situ data, so the performance of in-situ GBM models is expected to be higher than that of ERA5-based OTClIM models. These baselines allow us to disentangle the influence of the modeling approach (traditional vs. GBM) from that of the input data (ERA5 vs. in-situ). Section 4.C gives more technical details about both baseline models.

4.3.3. PHYSICAL CONSISTENCY CHECKING USING FEATURE IMPORTANCE

To assess if the trained models captured physically consistent relationships and to potentially discover currently unknown dependencies between features and the regression target, we quantify the importance of each feature for the prediction using SHapely Additive exPlanation (SHAP) values (Lundberg and Lee 2017). SHAP values explain how a trained model $\hat{f}_{s,t}$ arrives at its predictions $\hat{\mathbf{y}}$ on a per-sample basis. For each sample \mathbf{x}_i , the prediction \hat{y}_i is explained as

$$\hat{f}_{s,t}(\mathbf{x}_i) = \hat{y}_i = \mathbb{E}[\hat{\mathbf{y}}_i] + \sum_{j=1}^p \phi_{j,i} \quad (4.4)$$

where $\phi_{j,i}$ are the SHAP values of the i -th sample and $\mathbb{E}[\hat{\mathbf{y}}_i]$ is the expected value of all predictions. In plain words, the j -th SHAP value $\phi_{j,i}$ describes the contribution of the j -th feature to the prediction assuming a local linear model. A large step

$\phi_{j,i}$ compared to the other SHAP values indicates that the j -th feature contributes strongly to that prediction and must be important. Consequently, a global view of the importance of the j -th feature can be obtained by averaging the j -th SHAP values for all n samples of the dataset (Molnar 2022) such that

$$I_j = \frac{1}{n} \sum_{i=1}^n |\phi_{j,i}|. \quad (4.5)$$

Note that SHAP values explain a feature's contribution to a model's prediction \hat{y}_i , not to the true value y_i . Accurate feature importance (FI) estimates, thus, require well-performing models $\hat{f}_{s,t}$ where \hat{y}_i closely approximates y_i for training and testing data. As predictions typically align more closely with true values in training data than testing data, we compute SHAP values using the former. This approach allows us to examine the model's internal workings without the influence of the generalization error. Moreover, it aligns with practical scenarios where all available data is used for training the final model, meaning that the FI analysis must be performed using the training dataset (Molnar 2022).

4

4.3.4. PERFORMANCE EVALUATION

The performance of the trained OTCLiM models is quantified through the Pearson correlation coefficient r and the root-mean-squared error (RMSE) ϵ :

$$r = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{n \sigma_y \sigma_{\hat{y}}}, \quad (4.6)$$

$$\epsilon = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (4.7)$$

In eq. (4.6), the overbar denotes the mean and σ the standard deviation of true (y) and estimated (\hat{y}) scaled C_n^2 , respectively. Two validation strategies are employed to assess the temporal and geographical extrapolation capabilities of the OTCLiM models. Note that only the testing data used to compute the scores differ between the evaluation strategies – the training data and the trained model remain the same.

Temporal extrapolation For the MCP application – the temporal extrapolation –, the model $\hat{f}_{s,t}$ trained on data from year $t \in \mathcal{T}$ needs to accurately predict OT strength at the same site for the held-out testing years $\mathcal{T} \setminus t$: $\hat{f}_{s,t}(\mathbf{x}_i) = \hat{y}_i$ with $\mathbf{x}_i \in \mathbf{X}_{s,\mathcal{T} \setminus t}$ and $\hat{y} \in \hat{\mathbf{y}}_{s,\mathcal{T} \setminus t}$. The scores are computed between the true held-out data $\mathbf{y}_{s,\mathcal{T} \setminus t}$ and the predictions $\hat{\mathbf{y}}_{s,\mathcal{T} \setminus t}$ and are denoted $\square_{s,t}$, where \square is a placeholder for the performance metrics r and ϵ . Training happens in a round-robin fashion, so one model is trained per training year and evaluated for the hold-out years. This process is repeated until all years are used for training once. The average performance of these models for a specific station s , called the MCP score, is obtained by averaging the scores across all training years: $\bar{\square}_s = \langle \square_{s,t} \rangle_{t \in \mathcal{T}}$.

Cross-site evaluation The second evaluation strategy probes the geographical generalization capability of a model trained on station s when applied to another station $\tilde{s} \in \mathcal{S} \setminus s$. The trained model is tasked to predict OT strength at \tilde{s} given the full multi-year input data $\mathbf{X}_{\tilde{s}, \mathcal{T}}$ at \tilde{s} from which the cross-site scores using r and RMSE can be obtained: $\square_{(s,t) \rightarrow (\tilde{s}, \mathcal{T})}$. For readability, we average the cross-site scores achieved by the individual models per training site s : $\square_{s \rightarrow \tilde{s}} = \langle \square_{(s,t) \rightarrow (\tilde{s}, \mathcal{T})} \rangle_{t \in \mathcal{T}}$.

4.4. DATASET

To evaluate OTClIM’s performance, we train GBM models for 17 diverse locations across New York State (cf. section 4.4.1). Five years of C_n^2 are available at each site for which we select collocated and concurrent ERA5 data (cf. section 4.4.2).

4.4.1. NEW YORK STATE MESONET (NYSM)

The New York State Mesonet (NYSM) comprises 127 standard weather stations as of 2024 spread across New York State, United States of America, and has been fully operational since 2018 (Brotzge *et al.* 2020). These weather stations measure routine meteorological parameters such as 2 m-temperature, 10 m-wind speed and direction, surface pressure, and several other variables. The sampling rate of these measurements is on the order of seconds, and final values are reported as 10 min aggregates (mean and variance). However, measuring C_n^2 requires higher-frequency observations to resolve the inertial range of turbulence. Such high-frequency measurements are available at a subset of 17 stations – the flux stations – additionally equipped with Campbell Scientific CSAT3A sonic anemometers mounted at 9 m height. The NYSM sonic anemometers measure the three wind components and the sonic temperature at $f_s = 10$ Hz (Brotzge *et al.* 2020), which is high enough to obtain C_n^2 .

We utilize the data from these NYSM flux stations to obtain training targets for the OTClIM models. The flux stations are placed in diverse topographical and climatological environments as listed in table 4.1. The stations BKLN, QUEE, and STAT are located on rooftops in urban environments where measurements are strongly influenced by their immediate surroundings. Neighboring buildings can, for example, cast shadows or cause wakes, which influence radiation, wind, and, therefore, also the local turbulence (World Meteorological Organization 2023). Since we expect these urban stations to behave differently than the rural stations, they are marked with (*) in the table. For each of the 17 stations, five years of measurements are available. Following the notation introduced previously, the set of flux stations is denoted \mathcal{S} , and \mathcal{T} represents the set of five training years 2018 – 2022 available at each site. The corresponding target vector $\mathbf{y}_{s,t}$ is obtained by estimating C_n^2 from sonic anemometer measurements, applying quality assurance and control (QA/QC) steps, and scaling the OT data to make them comparable between different sites. These stages are described in detail below.

Structure function approach As noted in the context of eq. (4.2), C_n^2 quantifies the strength of refractive index fluctuations due to density fluctuations caused by

Table 4.1.: Flux stations of the New York State Mesonet used to benchmark the OTCLIM approach. The three urban stations are marked with (*).

ID	Nearest City	Latitude, ° N	Longitude, ° W	Environment
BELL	Belleville	43.79	76.11	Lake
BKLN*	Brooklyn	40.63	73.95	Coastal / Urban
BURT	Burt	43.32	78.75	Lake
CHAZ	Chazy	44.90	73.46	Valley
FRED	Fredonia	42.42	79.37	Lake
ONTA	Ontario	43.26	77.37	Lake
OWEG	Owego	42.03	76.26	Plateau
PENN	Penn Yan	42.66	76.99	Mountainous / Lake
QUEE*	Queens	40.73	73.82	Coastal / Urban
REDF	Redfield	43.62	75.88	Lake
REDH	Red Hook	42.00	73.88	Valley
SCHU	Schuylerville	43.12	73.58	Valley
SOUT	Southold	41.04	72.47	Coastal
STAT*	Staten Island	40.60	74.15	Coastal / Urban
VOOR	Voorheesville	42.65	73.98	Valley
WARS	Warsaw	42.78	78.21	Mountainous / Lake
WHIT	Whitehall	43.49	73.42	Valley

turbulent temperature and moisture fluctuations. The temperature T_s measured by sonic anemometers also contains a humidity contribution as $T_s = T(1 + 0.51q)$ with specific humidity q (Kaimal and Gaynor 1991). We assume that using the simplified version

$$C_n^2 \approx \left(AP / \bar{T}^2 \right)^2 C_{T_s}^2 \quad (4.8)$$

of eq. (4.2) implicitly accounts for moisture if the strength of these sonic temperature fluctuations, $C_{T_s}^2$, is used. That is because $C_{T_s}^2$ can be decomposed as $C_{T_s}^2 \approx C_T^2 + 1.02\bar{T}C_{Tq} + 0.26\bar{T}^2C_q^2$ with C_T^2 and C_q^2 representing the strength of pure temperature and moisture fluctuations and C_{Tq} the cross-term between the two. For readability, we will refer to $C_{T_s}^2$ simply as C_T^2 but would like to reiterate that a moisture contribution is included implicitly.

In the inertial range, C_T^2 is the coefficient of the second-order structure function of the (sonic) temperature defined as

$$S_T^2(\Delta x) = \langle (T(x) - T(x + \Delta x))^2 \rangle \stackrel{\Delta x \in \text{inertial range}}{\approx} C_T^2 \Delta x^{2/3}. \quad (4.9)$$

The structure function is computed from 5 min non-overlapping windows of the sonic temperature signal, and C_T^2 is obtained by fitting a 2/3 slope to the inertial range of S_T^2 in log-log coordinates.

Quality assurance and control (QA/QC) Since our GBM models are purely data-driven, the quality of the fitted model strongly depends on the quality of the training data, so we apply several QA/QC steps to obtain a clean dataset. First, C_T^2 has to be determined from the slope of the inertial range, i.e., slopes of $2/3$. Since the inertial range is not always resolved or well captured in every 5 min window, only C_T^2 values with well-fitted slopes are kept ($2/3 \pm 5\%$ slope, $R^2 > 0.95$). A similar procedure was followed by He and Basu (2016a) for the analysis of simulated data. Second, S_T^2 in eq. (4.9) is defined spatially as a function of separation distance Δx , but the sonic anemometer measures T temporally. Taylor’s frozen turbulence assumption is commonly made to convert the temporal sonic signal into a spatial form using the horizontal mean wind speed \overline{M} and $\Delta x = \overline{M}/f_s$. However, Taylor’s assumption might break down in low wind conditions, so all C_T^2 values for $\overline{M} < 1 \text{ m s}^{-1}$ are discarded. Finally, precipitation strongly affects sonic temperature measurements (Zhang *et al.* 2016), so only C_T^2 values for dry conditions can be used. This filter is also justified from an operational perspective because optical telescopes are typically not operated in the rain, and FSOC links are strongly attenuated by snow (ITU 2007).

Being an observed dataset, the sonic temperature signal contains gaps due to, e.g., power or communication outages or instrument malfunctioning (NYS Mesonet 2023), leading to $\sim 16\%$ of missing values on average. That leaves $\sim 84\%$ of the 5-min non-overlapping windows gap-free and suitable for the computation of C_T^2 . The three QA/QC steps further reduce the average number of valid samples, leaving $\sim 66\%$ after inertial range fitting, $\sim 55\%$ after ensuring the applicability of the Taylor assumption, and, $\sim 53\%$ after discarding precipitation events. Although only half of the whole dataset is considered valid, it still contains ca. fifty thousand samples per site and training year spread throughout diverse meteorological conditions. This claim is supported in section 4.A, where we present that the QA/QC procedure generally maintains the ratio between stable and unstable atmospheric conditions before and after QA/QC. Thus, the final dataset is considered adequate for training the OTclim models.

Distribution scaling The quality-controlled C_T^2 time series is converted to C_n^2 using the Gladstone equation (eq. (4.2)) and then scaled. The scaling is needed because the yearly distribution of OT strength differs between sites. Some locations show stronger OT (shifted distribution) or a larger range of C_n^2 values (scaled distribution) than other sites. To make the performance scores and SHAP FI values comparable between sites, we scale the site-specific $\log_{10} C_n^2$ before training based on the 50% log-range such that

$$[\log_{10} C_n^2(t)]_{\text{scaled}} = \frac{\log_{10} C_n^2(t) - [\log_{10} C_n^2]_{p25}}{[\log_{10} C_n^2]_{p75} - [\log_{10} C_n^2]_{p25}}. \quad (4.10)$$

Here, the subscripts p25 and p75 indicate the 25th and 75th percentiles of the original distribution used for scaling. More details are given in section 4.B.

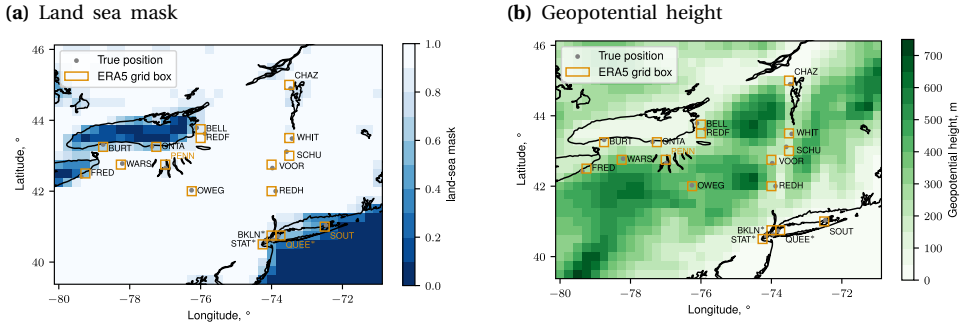


Figure 4.2.: ERA5 representation of NYSM domain with true locations of NYSM flux stations (grey) and corresponding ERA5 grid box (orange) containing the stations. Urban sites are marked with (*).

4

4.4.2. ERA5 REANALYSIS DATA

The ERA5 reanalysis (Hersbach *et al.* 2020) serves as the input dataset for OTCLiM from which C_n^2 shall be estimated. ERA5 is available globally from 1950 until the present and, thus, includes our NYSM training locations and times. We extract ERA5 input data for all NYSM stations and training years at the ERA5 grid points closest to the respective stations. ERA5 grid points represent 1/4-degree grid boxes (ca. 30 km by 30 km), so by selecting the closest point, we obtain data representative for the grid box containing the station. Figure 4.2 depicts the location of the flux stations (grey markers) and their collocated grid boxes (orange square) and also gives an impression of how the coarse ERA5 represents (a) land, sea, and (b) terrain. The temporal resolution/sampling rate also differs between ERA5 (1 h) and the observed C_n^2 dataset (5 min). ERA5 data are instantaneous snapshots available at each full hour, so we match each ERA5 sample with the average of the six 5-min C_n^2 values ± 15 min around the hour. This 30 min average centered around the hour smoothes the C_n^2 signal to counteract potential temporal misalignment between observations and reanalysis.

We aim to incorporate all variables linked to the two processes driving atmospheric turbulence: wind shear and buoyancy. Commonly used variables include sensible heat flux and friction velocity, as utilized in W71. However, an advantage of using ML over deriving analytical equations from theory is the ability to include variables that may indirectly influence turbulence. For example, the ERA5 gravity wave dissipation rate (GWD) could be significant in complex mountainous regions where orographic gravity wave drag is known to modulate momentum fluxes (Lilly 1972; Palmer *et al.* 1986). If there is a relationship between GWD and C_n^2 , ML models will likely identify and utilize it, potentially revealing new dependencies.

A complete list of the ERA5 variables selected as features is presented in table 4.2. Many listed features are (partially) redundant and/or (partially) correlated. Since it is initially unknown which features are most suitable for estimating near-surface C_n^2 , we aim for a broad feature space. During training, the GBM algorithm will

Table 4.2.: ERA5 variables and variables derived thereof serve as features for the OTClmI approach. Derived features do not have ERA5 variable names and are marked with (-).

Category	Symbol	Grouping	ERA5 variable name	Description	Unit
Wind	M_{100}	-	-	100m horizontal wind speed, see eq. (4.11)	m s^{-1}
Wind	M_{10}	-	-	10m horizontal wind speed, see eq. (4.11)	m s^{-1}
Wind	G_{10}	i10fg	-	Instantaneous 10m wind gust	m s^{-1}
Shear	u_*	zust	-	Friction velocity	m s^{-1}
Shear	α	-	-	Shear coefficient of power-law wind profile, see eq. (4.12)	-
Shear	β	-	-	Directional shear, $ X_{10} - X_{100} $	$^\circ$
Fetch	$\cos(X_{100})$	ΣX	-	Cosine of 100m wind direction, see eq. (4.13)	-
Fetch	$\cos(X_{10})$	ΣX	-	Cosine of 10m wind direction, see eq. (4.13)	-
Fetch	$\sin(X_{100})$	ΣX	-	Sine of 100m wind direction, see eq. (4.13)	-
Fetch	$\sin(X_{10})$	ΣX	-	Sine of 10m wind direction, see eq. (4.13)	-
Buoyancy	Q_H	ishf	-	Instantaneous surface sensible heat flux	W m^{-2}
Temperature	T_2	t2m	-	2m temperature	K
Temperature	T_0	skt	-	Skin temperature	K
Temperature	T_{soil}	stl1	-	Soil temperature	K
Stability	ΔT_0	$\Sigma \Delta T$	-	Temperature difference $T_{\text{soil}} - T_0$	K
Stability	ΔT_1	$\Sigma \Delta T$	-	Temperature difference $T_0 - T_2$	K
Radiation	$R_{S1,\text{diff}}$	ΣR	msdwsurf	Mean diffuse downwelling short-wave radiation flux at surface	W m^{-2}
Radiation	$R_{S1,\text{dir}}$	ΣR	msdrswrf	Mean direct downwelling short-wave radiation flux at surface	W m^{-2}
Radiation	R_S	ΣR	msnswrf	Mean net short-wave radiation flux at surface	W m^{-2}
Radiation	R_{L1}	ΣR	msdwlwrf	Mean downwelling long-wave radiation flux at surface	W m^{-2}
Radiation	R_L	ΣR	msnlwrf	Mean net long-wave radiation flux at surface	W m^{-2}
Cloud	lcc	Σcc	lcc	Low cloud cover fraction (0 - 1)	-
Cloud	tcc	Σcc	tcc	Total cloud cover fraction (0 - 1)	-
Moisture	ΔT_d	-	-	Dew point spread, $T_{d2} - T_2$	K
Moisture	T_{d2}	d2m	-	2m dew point temperature	K
Moisture	Q_q	ie	-	Instantaneous moisture flux	$\text{kg m}^{-2} \text{s}^{-1}$
Moisture	Q_L	mslhf	-	Mean surface latent heat flux	W m^{-2}
Auxiliary	h_i	blh	-	Boundary layer height	m
Auxiliary	BLD	bld	-	Boundary layer dissipation rate	J m^{-2}
Auxiliary	GWD	gwd	-	Gravity wave dissipation	J m^{-2}
Auxiliary	CAPE	cape	-	Convective available potential energy	J kg^{-1}
Auxiliary	P_0	msl	-	Mean sea level pressure	Pa
Auxiliary	$\partial n / \partial z_a$	dndza	-	Mean vertical gradient of refractivity inside trapping layer	m^{-1}
Auxiliary	h_{db}	dctb	-	Duct base height	m
Auxiliary	h_{tpb}	tplb	-	Trapping layer base height	m
Auxiliary	h_{tpt}	tplt	-	Trapping layer top height	m
Time	$\cos \text{hr}'$	$\Sigma \text{hr}'$	-	Cosine of normalized hour of the day	-
Time	$\sin \text{hr}'$	$\Sigma \text{hr}'$	-	Sine of normalized hour of the day	-
Time	$\cos \text{day}'$	$\Sigma \text{day}'$	-	Cos of normalized day of the year	-
Time	$\sin \text{day}'$	$\Sigma \text{day}'$	-	Sin of normalized day of the year	-
Time	$\cos \text{month}'$	$\Sigma \text{month}'$	-	Cos of normalized month of the year	-
Time	$\sin \text{month}'$	$\Sigma \text{month}'$	-	Sin of normalized month of the year	-

identify and base its predictions on the most important features. Through the post-training feature importance analysis, we can identify which features the trained model deems relevant and assess their physical relevance. It is usually preferred in ML to minimize the number of features and reduce model complexity. However, this concern does not apply to tree-based algorithms like GBM used in this study because such algorithms only consider features during fitting, which increase the model accuracy. Consequently, features contributing not or little to the model's accuracy are not selected (for splitting) while the trees are grown (Chen and Guestrin 2016), so their impact on model complexity is low (Spiliotis 2022).

The features in table 4.2 without ERA5 variable names are so-called engineered features, meaning that they are derived from one or more ERA5 variables. In the following, we detail the variable selection and the feature engineering.

Shear-related features Table 4.2 lists multiple wind-related features aimed at capturing wind shear, i.e., the vertical change of the wind. Like traditional C_n^2

parameterizations, such as W71, we select the friction velocity u_* as the first shear-related feature. We also include the horizontal wind speed obtained from the zonal and meridional wind fields of ERA5 – U_z and V_z – as

$$M_z = \sqrt{U_z^2 + V_z^2} \quad (4.11)$$

at $z = 10$ m and $z = 100$ m above ground. Assuming a power-law wind profile of the form $M(z) = M_{\text{ref}}(z/z_{\text{ref}})^\alpha$ between the 10 m and the 100 m level, we utilize the exponent of the power law α as an additional shear-related feature:

$$\alpha(z_1, z_2) = \frac{\log M_{z_2} - \log M_{z_1}}{\log z_2 - \log z_1}. \quad (4.12)$$

The directional shear, i.e., wind turning with height, is included through the absolute angular difference between 10 m and 100 m wind direction defined as $\beta = |X_{10} - X_{100}|$ where the wind direction X_z is given as

$$X_z = \arctan2(-U_z, -V_z). \quad (4.13)$$

The wind direction also serves as a proxy for upstream effects, or fetch, that might influence the observed turbulence. For example, atmospheric turbulence measured at a station close to the coast can be very different if the wind blows from land to sea or sea to land. The periodicity of X_z is accounted for by including the sines and cosines of X_z as fetch features instead of X_z directly.

Buoyancy and stability The prime candidate to reflect the influence of buoyancy on C_n^2 is the sensible heat flux Q_H . Sensible heat flux is also featured in W71 because it captures buoyancy and static atmospheric stability (in ERA5's convention, $Q_H > 0$ during stable/nighttime conditions and $Q_H < 0$ during unstable/daytime conditions). Additionally, static stability can be estimated from temperature gradients, so we utilize the absolute values of 2 m-temperature, skin temperature, and soil temperature, as well as the differences between them as complementary features.

Surface Energy Budget The strength of buoyancy depends on the radiation that reaches the surface, so we also include ERA5 radiation fluxes and cloud cover as features to complement the buoyancy information from Q_H . The simplified surface energy budget describes how radiation forces the surface fluxes. If the small soil heat flux into the ground is neglected, the steady-state energy balance at the surface is (Stull 1988)

$$Q_L + Q_H \approx R_S + R_L, \quad (4.14)$$

where Q_L is the latent heat flux due to evaporation and R_S and R_L are the net shortwave and longwave surface radiation fluxes. The net fluxes $R_\square = R_{\square\downarrow} - R_{\square\uparrow}$ are the difference between incoming/downwelling (index \downarrow) and the reflected/upwelling radiation (index \uparrow) of shortwave ($\square = S$) and longwave ($\square = L$) radiation. ERA5 contains the net and downward radiation fluxes as variables. Further, it splits the

shortwave downwelling radiation $R_{S\downarrow}$ into a direct component $R_{S\downarrow,\text{dir}}$ and a diffuse component $R_{S\downarrow,\text{diff}}$ resulting from scattering by clouds. Since clouds affect how much shortwave radiation reaches the surface and how much longwave radiation from the surface is reflected, we include ERA5's low cloud cover ($z < \text{ca. } 2 \text{ km}$) and total cloud cover into the dataset. Close to water bodies, Q_L is expected to be higher due to stronger evaporation leading to lower Q_H , and, thus, lower C_n^2 . Therefore, Q_L and other moisture-related ERA5 variables are listed as moisture-related features in [table 4.2](#).

Auxiliary features Finally, we include several auxiliary features more loosely related to wind shear and buoyancy, such as boundary layer dissipation rate (BLD), convective available potential energy (CAPE), or the aforementioned gravity dissipation rate (GWD). Also, certain daily and seasonal patterns exist in meteorology, which we aim to capture through synthetic time-dependent features. Based on the timestamp of the data point, we compute the sines and cosines (for periodicity) of the normalized hour of the day ($\text{hr}' = 2\pi \text{ hr}/24 \text{ h}$), the normalized day of the year ($\text{day}' = 2\pi \text{ day}/365 \text{ d}$), and the normalized month of the year ($\text{month}' = 2\pi \text{ month}/12 \text{ mo}$).

4.5. RESULTS: FEATURE IMPORTANCE

The physics governing optical turbulence are the same everywhere – in urban or rural environments or for in-land stations or coastal sites. The modulating processes are always buoyancy and wind shear. However, the ERA5 features that best reflect and predict these processes locally can differ between sites. To assess and quantify potential differences, we present the SHAP-based feature importance values for all trained models in [fig. 4.3](#). To make the FI analysis less verbose, we make use of the linearity of the SHAP values (cf. [eq. \(4.4\)](#)) and group the SHAP values of features related to similar physical processes as presented in [table 4.2](#). These groups are prefixed with Σ .

Panel (a) depicts the importance distributions of all features for all models. The distributions are represented with one box plot per feature where the box shows the inter-quartile range (IQR), i.e., the 25%–75% range. The whiskers show 1.5 IQR above and below the 25% and 75% quartiles, and the circle indicate stations where the FI values leave that range. This global view reveals that buoyancy-related features such as radiation ΣR , sensible heat flux Q_H , and cloud cover Σcc are key variables for C_n^2 prediction. Also, the shear exponent α is picked up, indicating that the OTClIM models identify wind shear as a modulating factor of turbulence. The dependency on wind direction ΣX suggests that upstream effects influence C_n^2 prediction at many sites, while atmospheric stability is likely captured next to Q_H by the temperature differences $\Sigma \Delta T$ and the boundary layer height h_i . All of the above aligns with our physical knowledge of atmospheric turbulence, indicating that the models picked up physically meaningful relations between ERA5-generated variables and C_n^2 . However, the boxplots also show outliers for several features, indicating that not all models agree with the global view.

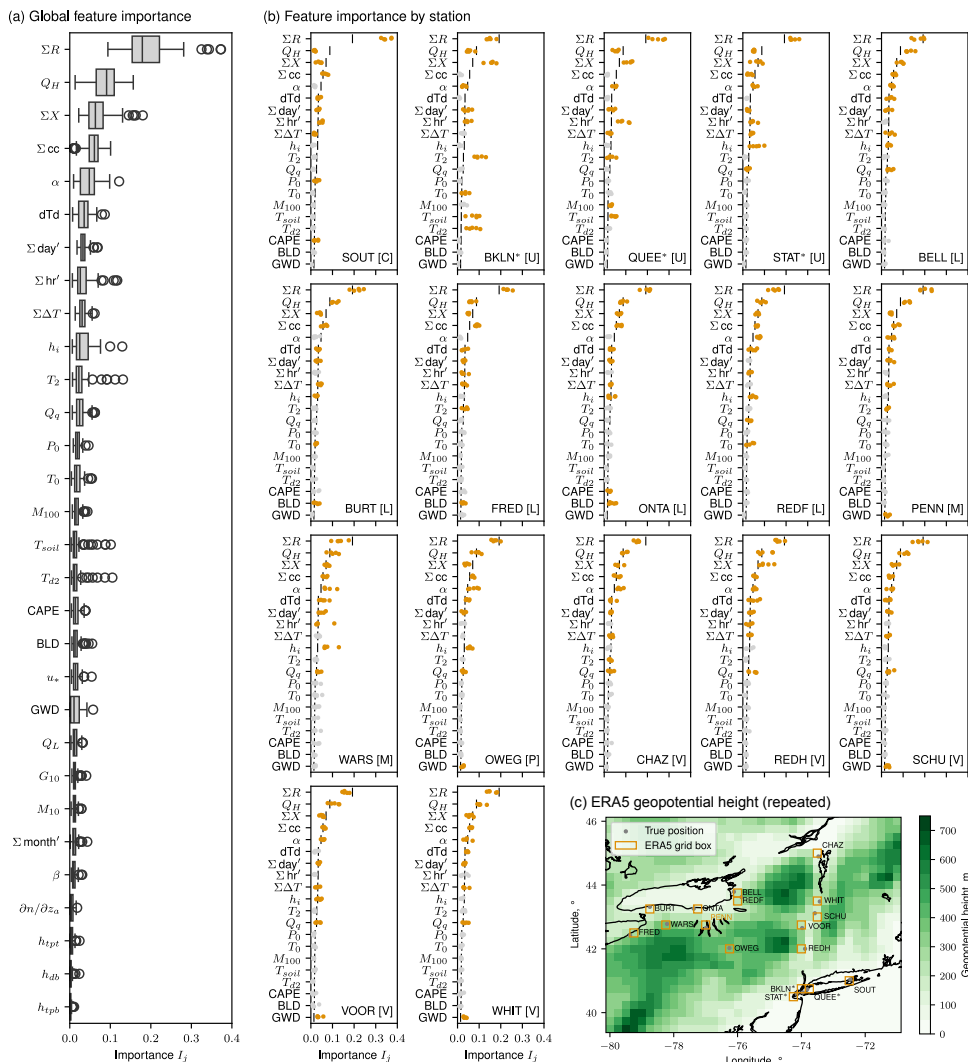


Figure 4.3.: SHAP value-based feature importance (FI) of all OTCLiM models (a) aggregated and (b) per training site. In (b), urban stations are marked with (*), the top 10 features of each station are highlighted in orange, and the global FI averages are indicated as black dashes. Panel (c) repeats the geopotential height from fig. 4.2 to aid geographical interpretation of the results.

The FI values are displayed per station in fig. 4.3(b), allowing us to assess the outliers in more detail. The ten features with the highest average FI for each location are highlighted in orange. Since different stations exhibit different top ten features, 20 features are shown in total, but their markers are colored in light grey if the features are not in the top ten of the site in question. For reference, the black

lines for each feature represent the global average FI values according to panel (a). Comparing all per-station plots on a high level reveals two key points. First, although the FI values associated with the five models trained for each site exhibit some scatter, the overall importance of features at one site is consistent between models. Therefore, we conclude that the models captured features representative of that site throughout the training years. Secondly, the color coding highlights that different features are important for different sites, suggesting that the processes dominating turbulence vary between locations.

The three urban stations, BKLN, QUEE, and STAT, and the coastal SOUT site deviate most obviously from the mean FI distribution. In particular, BKLN and QUEE show an above-average dependency of estimated C_n^2 on wind direction. Both stations are located on rooftops and have some tall buildings in their neighborhood (Brotzge *et al.* 2020). Such inhomogeneous urban conditions are known to influence local measurements (World Meteorological Organization 2023), so it seems reasonable that turbulence strength observed at BKLN and QUEE depends more strongly on wind direction compared to other more rural locations. All four stations (urban + SOUT) also show below-average dependency on Q_H , which we view as a feature capturing the diurnal cycle. Instead, the models seem to have picked $\Sigma hr'$ or h_i as predictors for this information. This behavior differs from all other models, where Q_H is typically the second most important input, in line with more traditional physics-based C_n^2 models, such as W71. This shift, however, need not be viewed as unfavorable but as a demonstration of the power of ML-based modeling to shift from traditional to “unconventional” features if complex flow conditions require it. The downside is that we expect such models to perform more poorly when applied to other sites where the more traditional features are relevant. Especially, BKLN is expected to generalize poorly because it also depends strongly on the temperature features T_2 , T_{d2} , and T_{soil} , deemed irrelevant by almost all other models.

Less drastic but distinct differences are also visible between the non-urban models. Wind shear α , for example, has lower than average importance assigned for models trained on lake shores (FRED, BURT, ONTA). The models of another set of stations (WHIT, VOOR, SCHU, PENN, OWEG) picked up the gravity-wave dissipation rate (GWD). These sites are located in valleys or mountainous areas where gravity waves could modulate near-surface turbulence (Lilly 1972; Palmer *et al.* 1986). Still, the dependency is small, and stations CHAZ and REDH, located at the end of valleys but still surrounded by mountains, do not depend on GWD. A more detailed study of the sites’ climatologies would be needed, which is beyond the scope of this work.

Overall, the FI values of most models represent the known physical dependencies of atmospheric turbulence. That supports our confidence that our OTClIM approach is well suited for MCP. The “unconventional” features picked up by some models are viewed as an advantage of ML-based methods to arrive at accurate predictions even in complex environments. However, we assume that geographical generalization will be more difficult for such models, which is addressed later in this paper.

4.6. RESULTS: OTCLIM PERFORMANCE

After establishing that all OTCLIM models picked up physically meaningful dependencies, we turn our analysis to quantifying their performance. The foundation for this analysis are 85 models trained individually for the 17 NYSM stations and the five training years. The temporal extrapolation capabilities of each model are quantified in [section 4.6.1](#) to assess the suitability of OTCLIM for MCP. To put the MCP scores achieved by our approach into perspective, we compare them to the two baseline models, W71 and the in-situ-based GBM models. Since we have a network of stations available, we also assess the geographical generalizability of the OTCLIM models by applying models trained on one site to all other sites. The results of this cross-site evaluation study are discussed in [section 4.6.2](#).

4

4.6.1. TEMPORAL EXTRAPOLATION

The temporal extrapolation performance of each model $\hat{f}_{s,t}$ with respect to the correlation coefficient r and the scaled RMSE ϵ is presented in panels (a) and (b) of [fig. 4.4](#), respectively. The heatmaps for both metrics show the MCP scores $\square_{s,t}$ with their site-averaged values $\bar{\square}_s$ being compared to the baseline models in the accompanying scatter plots. The heatmap reveals similar patterns for both metrics: the score variance across training years (rows) is low for almost all stations, while the site-specific performance (columns) varies notably. In other words, regardless of the year of C_n^2 observations used for training, the MCP performance stays consistent, whereas C_n^2 at some locations (e.g., QUEE, BKLN, SOUT) is easier to predict from ERA5 than at others (e.g., VOOR, WHIT). It is unexpected but impressive that the urban models marked with (*) are the highest performing ones due to the typically complex urban climatology, which makes traditional modeling difficult (Rotach *et al.* 2005). Panels (c) and (d) show a few randomly drawn batches of observed C_n^2 with their corresponding predictions. These curves highlight that the overall performance at SOUT (c) is better than that of WHIT (d) for two reasons. First, the observed C_n^2 (black) at WHIT exhibits more high-frequency oscillations than SOUT, which are missing in the predictions (red). Second, complex short-duration events seem to occur more frequently for WHIT, which are only partially captured. Both effects indicated that WHIT is subject to more complex flows than SOUT and that OTCLIM misses these details. We attribute this to missing details in ERA5, as will be discussed later. For now, we conclude that the models for both locations capture the general trends of C_n^2 well and that the observed intra-site performance variation is primarily due to smoothed predictions. For practical applications such as obtaining a C_n^2 climatology, the large-scale behavior of C_n^2 is the most relevant, so the smoothing is of limited concern.

The insight that the training year selection has a low influence on the model performance has high practical relevance. It suggests that models can be trained on archive data and do not necessarily require recent observations if and only if the site's climatology remains comparable between the training period and the envisioned period of model application. Drastic changes in the station's surroundings, e.g., through construction projects or climate change, would likely break this assumption and, thus, the model's applicability.

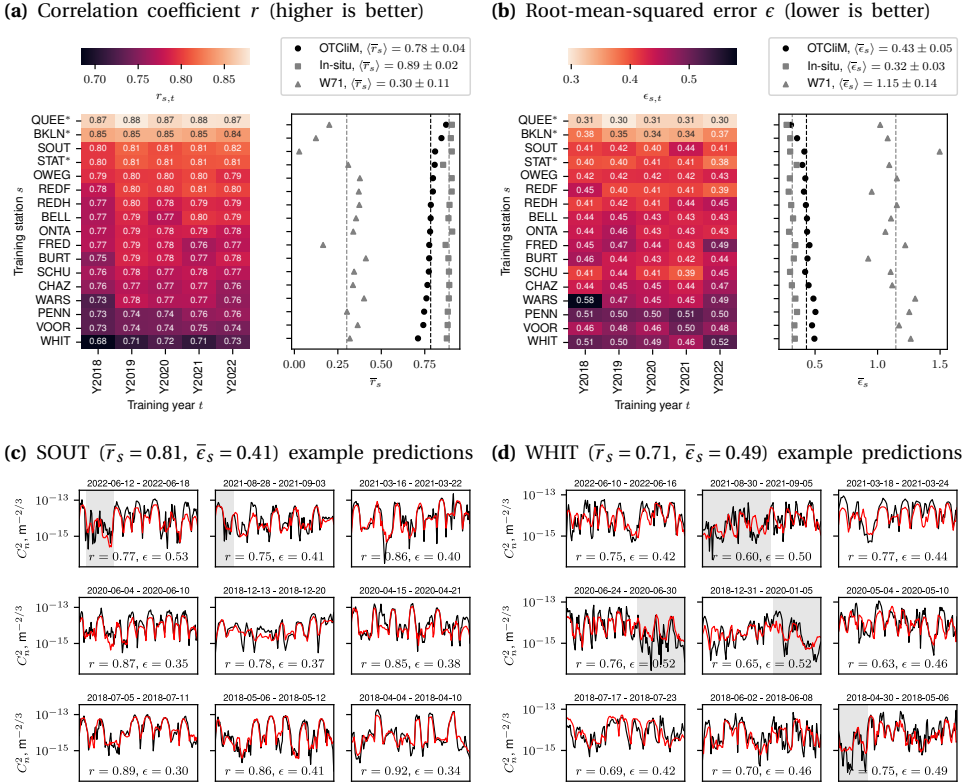


Figure 4.4.: Performance of OTCLiM models compared to the baseline models. The heatmaps show the performance of individual OTCLiM models trained with one year of C_n^2 observations from site s when evaluated on the four hold-out years of the same site. The scatter plots compare the site-averaged performance against the baseline estimates. Seven-day batches of C_n^2 are randomly drawn to compare observations (black) against predictions (red) for a high-accuracy (c) and a lower-accuracy OTCLiM model (d), with complex turbulence conditions shaded in grey.

The variability across models trained for different stations is likely due to ERA5 not fully capturing all local effects modulating the turbulence. For example, some landscape features, such as small lakes around PENN and WARS, are smaller than the ERA5 resolution, so they are missed. The relevance of these local features is highlighted when the OTCLiM performance (black circles) is compared to that of the in-situ baseline models in the \bar{r}_s and $\bar{\epsilon}_s$ scatter plots. While the in-situ models exhibit a small performance scatter around their mean accuracy (dashed lines), a generally widening performance gap is visible, which sets off small (QUEE) and grows toward the lower end of the performance ranking. Remember that only the input data differs between OTCLiM and in-situ (ERA5 vs. in-situ observations); the GBM training framework remains the same. Consequently, the growing gap between

the model variants can be attributed to ERA5 holding less predictive power for some locations than others. Expectedly, the models with lower accuracy (e.g., WHIT, VOOR, PENN) are located close to lakes or in mountainous areas, which are usually subject to complex flows that ERA5 might not well capture. Surprisingly, the ERA5-based urban models perform almost on par with the in-situ-based models. This result is impressive, given that urban flows are complex and likely not well represented in ERA5. Nevertheless, as discussed earlier, GBM seemingly circumvented this issue by shifting from traditional to unconventional features in urban cases. These features still seem to hold enough predictive power for the high performance observed here.

A final comparison between OTCLIM and the traditional W71 parameterization in the scatter plots of panels (a) and (b) shows that OTCLIM clearly outperforms W71. That is because the underlying similarity function (cf. eq. (4.3)) is empirically determined based on the flow over a flat, unobstructed plain (Wyngaard *et al.* 1971) and does not adapt to local topography or climatology. This limitation becomes especially evident for the urban sites QUEE and BKLN where \bar{r}_s is much below the W71 average (dashed line). Our OTCLIM models, on the other hand, perform well at these locations, highlighting the advantage of accounting for local climatology in complex atmospheric conditions. In summary, the presented scores demonstrate the capability of our OTCLIM approach to accurately extrapolate a 1-year C_n^2 time series to multiple years with ERA5 input for a diverse set of locations.

4.6.2. CROSS-SITE EVALUATION

Next, we assess the geographical generalization capabilities of the OTCLIM models by assessing their performance when being evaluated across different sites. Each model $\hat{f}_{s,t}$ trained on one site s is asked for C_n^2 predictions based on five years of ERA5 input data of all other sites ($\bar{s} \in \mathcal{S} \setminus s$). The resulting cross-site (c/s) correlation coefficient and RMSE scores are presented in fig. 4.5. Each cell (s, \bar{s}) of the heatmaps corresponds to the average score achieved by the five OTCLIM models trained on site s (rows) and evaluated on site \bar{s} (columns). The heatmaps on the left present the scores normalized by the MCP scores, reflecting the performance degradation of a model when applied to other sites compared to performance on the hold-out years of the original training site. The heatmaps on the right show the absolute scores, and the histograms in the middle display the distribution of both relative and absolute scores with diagonals excluded, i.e., without the MCP performance. The rows and columns of all heatmaps are ordered by mean generalization performance quantified as $\langle r'_{s \rightarrow \bar{s}} \rangle_{\bar{s} \in \mathcal{S} \setminus s}$.

Overall, the geographical generalization works very well for the non-urban sites. For half of all cases, correlation and RMSE performance degrades by not more than 16% for r and 25% for ϵ compared to the original MCP scores. Considering 75% of the cases, corresponding to the c/s scores of almost all non-urban sites, the absolute performance falls as low as $r = 0.56$ and $\epsilon = 0.63$, which is still significantly better than the W71 baseline scores of $\langle \bar{r}_s \rangle = 0.3$ and $\langle \bar{\epsilon}_s \rangle = 1.15$ (cf. fig. 4.4). In other words, any non-urban OTCLIM model applied to any other non-urban site yields significantly better performance than the traditional approach.

As expected from the FI analysis, the urban models do not perform well in

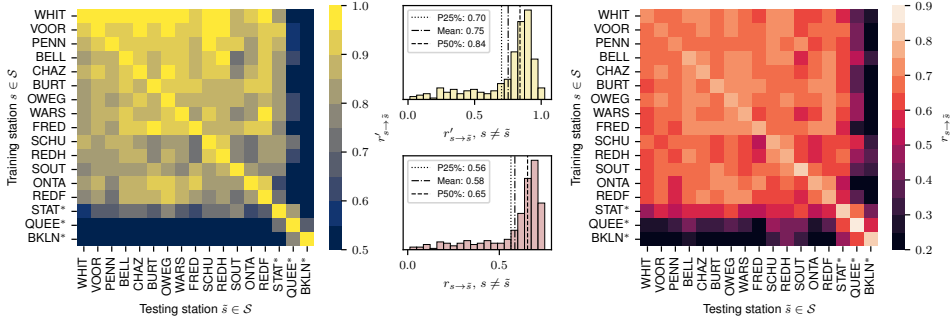
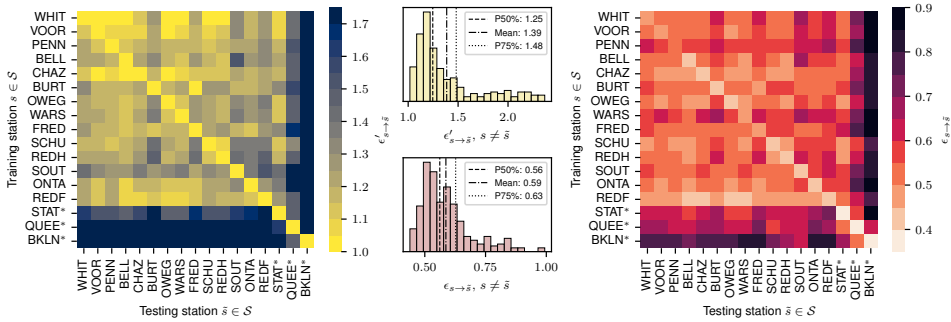
(a) Relative and absolute correlation coefficient r (higher is better)(b) Relative and absolute root-mean-squared error ϵ (lower is better)

Figure 4.5.: Cross-site evaluation performance of OTCLiM models. The rows of the heatmaps present the performance of the models trained on site s when evaluated on all other sites $\tilde{s} \in \mathcal{S}$. The left heatmaps show the performance degradation compared to the MCP case, i.e., the relative scores $\square'_{s \rightarrow \tilde{s}} = \square_{s \rightarrow \tilde{s}} / \square_{s \rightarrow s}$ with \square representing r or ϵ , and the right ones show the absolute scores for reference. The histograms depict the distributions of both heatmaps.

non-urban locations. Both relative and absolute c/s scores in fig. 4.5 are low for STAT, QUEE, and BKLN, resulting in long tails in the c/s histograms. However, the urban models perform better at other urban sites, as indicated by the small square of higher performance in the bottom right corner of the c/s matrices. It seems that the performance of urban generalization could be linked to the degree of urbanization. BKLN is the most urban site, QUEE has more vegetation than BKLN, and STAT has more vegetation than QUEE, which is reflected in the progression of c/s performance: BKLN generalizes most poorly and depends on the most uncommon features, QUEE performs slightly better, and STAT performs even better with FI values similar to the average case. Nevertheless, this observation might be unique to these New York City stations, which are all relatively close to each other and experience similar mesoscale conditions.

4.7. CONCLUSION

This study presents OTCLiM, a gradient boosting-based measure-correlate-predict approach to obtain climatologies of optical turbulence strength from one year of C_n^2 observations and multi-year ERA5 reference data. A feature importance (FI) analysis based on SHAP values revealed that most models, especially the non-urban ones, learn similar dependencies. The dominating features reflect the known and expected physical dependencies, suggesting that the OTCLiM models captured physically meaningful relations between ERA5 and C_n^2 . The FI results also suggest features such as boundary layer dissipation rate or gravity wave dissipation rate as C_n^2 dependencies, which are not considered in traditional models. This result demonstrates how machine learning models could hint at new insights into turbulent processes. In contrast to most stations, the three urban locations differed significantly from the average FI distribution. We assume that such models learned very different input-output relations, hindering them from generalizing well to other sites. In a temporal extrapolation study, we demonstrate that the OTCLiM models accurately predict four held-out years of C_n^2 from ERA5 inputs across 17 diverse sites in New York State. The models' performances stay consistent regardless of the year of data selected for training, suggesting archive data can also be used for training. The trained models are also found suitable for stations located in complex environments such as valleys or cities, which are often difficult to model traditionally. Minor accuracy variations between models from different sites are visible but can be attributed to local details missing in the coarse ERA5 input data. For some regions in the world, regional reanalysis datasets are available at higher resolution (e.g., Europe (Schimanke *et al.* 2021) or Australia (Su *et al.* 2019)) or with more assimilated data (e.g., North America (Mesinger *et al.* 2006)), which might help to reduce this problem. Already with ERA5 input, OTCLiM models significantly outperform traditional analytical C_n^2 models and achieve scores close to those of GBM models trained on in-situ inputs. A geographical generalization study also showed that applying non-urban models to other non-urban sites yields high scores in many cases, indicating that these models generalize well. The performance only degrades significantly when the urban models are evaluated at non-urban sites. While these models performed very well in the temporal extrapolation test case, the low scores for cross-site application indicate that the urban models learned very site-specific relations.

The key conclusion for practice from our work is that one year of C_n^2 observations is sufficient to obtain reliable site-specific C_n^2 climatologies using OTCLiM. Measuring C_n^2 at sites of interest for only one year instead of multiple years would free up instruments sooner. As a result, site survey costs could be reduced, or more locations could be surveyed for the same costs. OTCLiM's high performance in urban environments is highly relevant for FSOC, where terminals can be expected to be located in both rural and urban locations. The good geographical generalization results of the non-urban models indicate that future OTCLiM models trained on a few stations can potentially be used to estimate the spatial distribution of C_n^2 over larger areas. Such C_n^2 maps could provide a first indication of locations suitable for FSOC or astronomy with respect to optical turbulence. However, the region-specific

features identified in the FI analysis for lake or mountain environments also suggest that the models must account for local effects. Developing such multi-region models requires more work before large-scale C_n^2 estimates can be considered reliable. As a first step, users can perform an FI analysis on their trained models and use the results as guidance for expected geographical generalization. For example, if the models pick up unconventional features, they will likely not generalize to other sites, but if features are more traditional, they might.

Finally, we believe that OTClIM can be highly relevant also for C_n^2 forecasts. The rapidly advancing ML weather prediction (MLWP) models, such as Graphcast (Lam *et al.* 2023) or AIFS (Lang *et al.* 2024a), are typically trained on ERA5 and, therefore, produce ERA5-like forecasts. Currently, not all variables required for OTClIM are available from MLWP, but in principle, OTClIM can be used to translate the MLWP forecasts to C_n^2 forecasts. Similarly, one could train OTClIM on historical data from the Global Forecasting System (GFS) (National Centers for Environmental Prediction 2015), a traditional numerical weather prediction system, and convert the GFS forecasts to C_n^2 forecasts.

Besides all these advantages, the critical assumption of OTClIM is that the observed year of C_n^2 represents the temporal extrapolation range. If the surroundings change drastically through, e.g., construction, this assumption breaks down, and OTClIM predictions can no longer be considered valid. Also, our proposed approach currently only predicts near-surface C_n^2 , but for astronomy and FSOC, higher-level C_n^2 or full profiles are also relevant. In a previous study, we presented an ML-based framework that can combine C_n^2 from multiple levels into one physics-inspired model (Pierzyna *et al.* 2023a) for multi-level C_n^2 predictions. How this approach can be integrated into OTClIM remains open for future work. In summary, our presented OTClIM approach is shown to be highly accurate in diverse meteorological conditions with the potential for geographical generalization. We believe that OTClIM is a relevant tool for the optical turbulence community for future site surveys and related studies.

ACKNOWLEDGMENTS

This publication is part of the project FREE – Optical Wireless Superhighways: Free photons (at home and in space) (with project number P19-13) of the research programme TTW-Perspectief, which is (partly) financed by the Dutch Research Council (NWO). This research is made possible by the New York State (NYS) Mesonet. Original funding for the NYS Mesonet (NYSM) buildup was provided by Federal Emergency Management Agency grant FEMA-4085-DR-NY. The continued operation and maintenance of the NYSM is supported by National Mesonet Program, University at Albany, Federal and private grants, and others. Sukanta Basu is grateful for financial support from the State University of New York's Empire Innovation Program.

DATA AVAILABILITY

The Python code for training and the trained models are available on GitHub: <https://github.com/mpierzyna/otclim>.

4.A. RATIO OF STABLE TO UNSTABLE CONDITIONS IN DATASET BEFORE AND AFTER QA/QC

A three-step quality assurance and control (QA/QC) procedure is described in [section 4.4.1](#), which aims at filtering unphysical values from our C_n^2 dataset to lay the foundation for high-quality OTCLIM models. As discussed in the respective section, the QA/QC procedure discards approximately half of the data points across all stations and all years. We aim to not significantly change the ratio of unstable to stable atmospheric conditions by QA/QC to expose the OTCLIM models to turbulence conditions representative of the respective site. In [fig. 4.6](#), we present the distribution of atmospheric stability before and after QA/QC for each location. Stability is quantified through the bulk potential temperature gradient $\Gamma = (\theta_9 - \theta_2)/7\text{ m}$, which is computed from the potential temperature θ measured by two thermometers at 2 m and 9 m above ground (NYS Mesonet 2023). These instruments differ from the sonic anemometers used to estimate C_n^2 . They are more reliable, resulting in only 1.5% of missing values on average due to instrumentation problems compared to 16% for the sonic anemometers. The only exception is the QUEE site, which did not have any 9 m observations during our time of interest, which is why it is omitted in this discussion. [Figure 4.6](#) shows that the unstable/stable ratio (see bar chart insets) does not change much at most sites due to quality control. In general, the number of stable condition samples decreases because stable conditions occur more frequently in winter when instrument malfunctioning due to rain, ice, or snow is more common. Additionally, the inertial range shrinks with increasing stability and can even disappear (Grachev *et al.* 2013), resulting in more failed fits of the 2/3 slope (QA/QC step one). Since the balance between stable and unstable cases remains better than approx. 30/70 for all stations, we view our dataset as reasonably well-balanced and adequate for training.

4.B. SCALING OF $\log_{10} C_n^2$ TARGET DATA

As described in [section 4.4.1](#) of the main text, the $\log_{10} C_n^2$ data are normalized using the 25th and 75th percentiles of the site-specific $\log_{10} C_n^2$ training data (cf. [eq. \(4.10\)](#)). The effect of this scaling is illustrated in [Figure 4.7](#), which presents the unscaled (left) and scaled (right) data colored by location. The plots demonstrate that the scaling makes the range of $\log_{10} C_n^2$ more comparable between sites, which is crucial to obtain comparable performance and FI scores. That is especially important for the urban BKLN site (dark blue), where the unscaled right tail of the distribution is shifted to the right by half an order of magnitude (higher OT strength) compared to the other sites. To keep the scaling robust and general (i.e., applicable to non-gaussian distributions), we scale using the inter-quartile range and not the commonly used standard deviation.

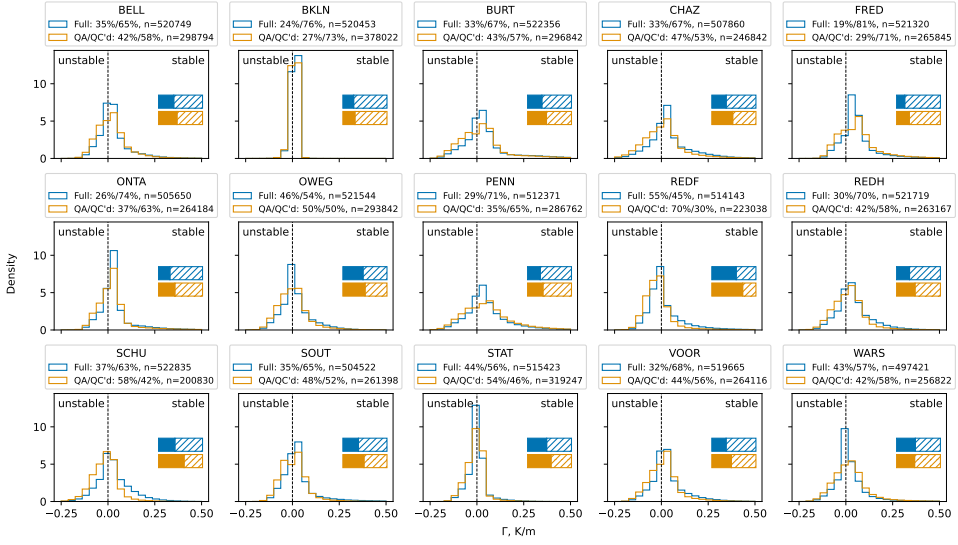


Figure 4.6.: Distribution of bulk potential temperature gradient $\Gamma = (\theta_3 - \theta_2)/7m$ before (blue) and after (orange) applying the quality assurance and control (QA/QC) steps presented in section 4.4.1. The split between unstable ($\Gamma < 0$, solid bar) and stable ($\Gamma > 0$, hatched bar) atmospheric conditions is visualized by the bar charts in each panel or quantified in the respective legend. The QUEE site is omitted fully because of a malfunctioning instrument.

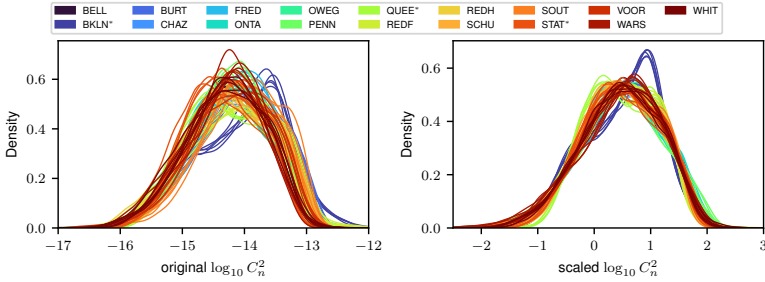


Figure 4.7.: Comparison between distributions of unscaled (left) and scaled (right) $\log_{10} C_n^2$. Colors indicate the different sites at which C_n^2 is observed with urban sites marked by (*).

4.C. DETAILS ABOUT BASELINE MODELS

This section details how the traditional W71 C_n^2 parameterization and the in-situ-based GBM models are setup and utilized as performance baselines for our proposed OTcliM approach.

4.C.1. W71 C_n^2 PARAMETERIZATION

All variables on which the W71 equations (cf. eqs. (4.1) to (4.3)) depend are available from ERA5 (cf. table 4.2). Only the dynamical sensible heat flux Q_H from ERA5 needs to be converted to its kinematic form $\overline{w'\theta'}$ as $\overline{w'\theta'} = -Q_H/(\rho c_p)$, where ρ and c_p are the density and specific heat capacity of air (Stull 1988). Also, the sign of Q_H is flipped because ERA5 treats an upward Q_H as negative, whereas W71 expects the upward $\overline{w'\theta'}$ to be positive. After this conversion, we utilize the same ERA5 data extracted for the 17 NYSM stations that we use to train the OTCLiM models to evaluate the W71 equations. The results are 17 five-year C_n^2 time series, which can be compared to the observed C_n^2 evolution at each station. The corresponding RMSE and r scores are presented in fig. 4.4.

4

4.C.2. IN-SITU GBM MODELS

These upper-bound models are similar to the OTCLiM models but use different input data. The OTCLiM model of a specific site s is a GBM model (cf. section 4.3.1) that employs ERA5 input data extracted from the grid box containing site s and C_n^2 observed at site s . Instead of ERA5, one can utilize in-situ weather data observed at s . The resulting GBM in-situ models are quite similar, for example, to work presented by Wang and Basu (2016), Jellen *et al.* (2021), or Pierzyna *et al.* (2023a). As for selecting ERA5 variables (cf. section 4.4.2), we aim to capture buoyancy and wind shear with multiple features based on the instruments deployed at the NYSM stations. As buoyancy features, we compute the temperature gradient $\Gamma = (T_9 - T_2)/(9\text{ m} - 2\text{ m})$ from observed 2 m and 9 m temperature, and utilize the observed 30 min sensible heat flux and observed incoming radiation. Assuming the velocity at the ground to be close to zero, we obtain a crude estimate of the bulk wind shear $S = M_{10}/10\text{ m}$ where M_{10} is the horizontal windspeed at 10 m. Additionally, observed friction velocity u_* , latent heat flux $\overline{w'q'}$, and the dew point spread $\Delta T_d = T_2 - T_{d2}$ are included as features, where T_{d2} is the 2 m dew point temperature. To account for atmospheric stability, Γ and S are combined into a bulk Richardson number as $\text{Ri} = (g/\overline{T})\Gamma/S^2$ with $g = 9.81\text{ m s}^{-2}$. Upstream effects are captured by adding sine and cosine of the 10 m wind direction. All these variables form the in-situ dataset, which is used to train GBM models in the same fashion as OTCLiM: each in-situ model is trained on one year of observations from site s and evaluated on the hold-out years of the same site. The resulting scores are presented as upper baseline in fig. 4.4.

5

Π -ML: A dimensional analysis-based machine learning parameterization of optical turbulence in the atmospheric surface layer

While OTcliM (chapter 4) demonstrated that near-surface C_n^2 can be estimated well from reanalysis data, it treats the underlying input-output relationship as a black box and is limited to a single level. This chapter proposes Pi-ML, a physics-informed machine learning (ML) methodology based on dimensional analysis and gradient boosting. By expressing inputs and outputs as non-dimensional groups grounded in physical principles, the resulting models remain interpretable. Π -ML is applied to C_n^2 observations recorded at 6 m, 15 m, and 25 m height at the Mauna Loa Observatory (MLO) on Hawai'i. Through a systematic feature importance analysis, we identify the normalized variance of potential temperature as the dominating feature for predicting C_n^2 . For statistical robustness, we train an ensemble of models which yields high performance on the out-of-sample data of $R^2 = 0.958 \pm 0.001$. As the MLO dataset is limited in duration and measurement height, we demonstrate in appendix B that Π -ML can be extended to larger heights and more extensive datasets, while maintaining high performance and interpretability.

This chapter is published as: M. Pierzyna *et al.* (2023a). " Π -ML: A Dimensional Analysis-Based Machine Learning Parameterization of Optical Turbulence in the Atmospheric Surface Layer". In: *Optics Letters* 48.17. DOI: 10.1364/OL.492652

5.1. INTRODUCTION

Free-space optical communication (FSOC) between satellites and ground or between multiple ground terminals is among emerging applications in which an optical beam propagates through the atmosphere. FSOC can have a major societal impact, increasing data throughput, data security, and global internet coverage while potentially reducing the cost per bit per second (Hemmati 2009). However, some challenges need to be addressed; besides precipitation, clouds, fog, and aerosol scattering, turbulent fluctuations of the atmospheric refractive index form a major source of disturbance (Kaushal and Kaddoum 2017). The strength of these fluctuations – called optical turbulence – is quantified by the refractive index structure parameter C_n^2 . Good knowledge about the behavior of C_n^2 in diverse locations and meteorological conditions is required to design and deploy reliable future FSOC links. However, measuring C_n^2 is difficult and typically needs elaborate post-processing of high-frequency observations (Wyngaard *et al.* 1971). As a result, a wide range of empirical C_n^2 models and parameterizations have emerged, which aim to relate C_n^2 to more easily obtainable variables (Smith *et al.* 1993).

Conventional physics-based C_n^2 parametrizations typically make use of Monin-Obukhov similarity theory (MOST) (Monin and Obukhov 1954) and associated empirically determined similarity relationships. One of the earliest parameterizations was proposed by (Wyngaard *et al.* 1971) and utilizes turbulent fluxes to estimate C_n^2 . Several other competing formulations exist (refer to (Savage 2009) for a comprehensive review).

Recently, multiple studies (Wang and Basu 2016; Jellen *et al.* 2021; Bolbasova *et al.* 2021; Su *et al.* 2021) showed that machine learning (ML) models can be used to parametrize C_n^2 based on routinely-available meteorological inputs. These ML approaches parametrize the underlying physical processes from data through sophisticated regression, but they do not explicitly incorporate physical knowledge. In this letter, we propose an alternative physics-inspired ML framework. We present Π -ML, a dimensional analysis-based ML framework, which strives to improve conventional MOST-based surface layer parameterizations with the power of ML. We utilize dimensional analysis (DA) constrained with domain knowledge to expand the set of traditional MOST variables and an ensemble of gradient-boosting ML regression models to learn similarity relationships from observations. In DA, the relevant dimensional variables of a physical process are combined into non-dimensional groups which describe that process equally well (Stull 1988). DA is compelling to use in practice because the non-dimensional variables enable us to combine observational data from different field campaigns around the world. More importantly, when using ML, DA can change the extrapolation problem in dimensional variables to an interpolation problem in non-dimensional variables (Kashinath *et al.* 2021).

5.2. METHODOLOGY

To investigate the strengths and weaknesses of the proposed methodology, we use measurements collected during a seeing study at the Mauna Loa Observatory (MLO)

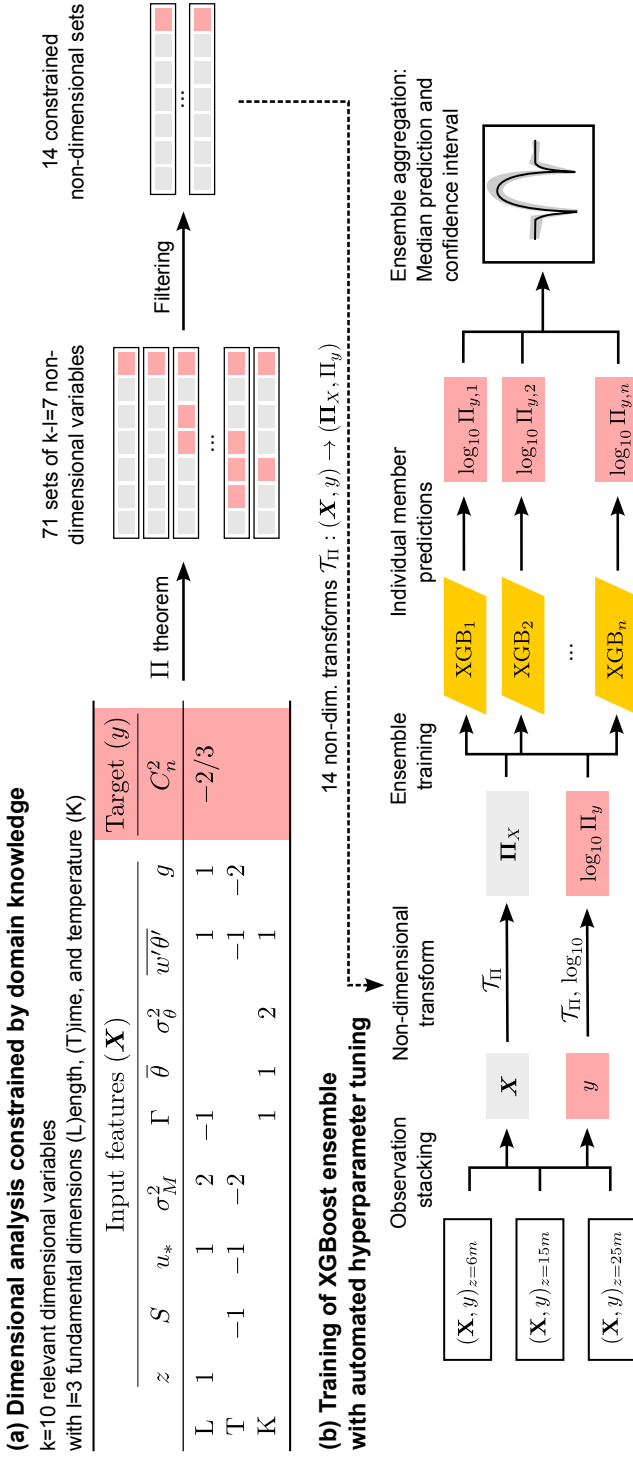


Figure 5.1.: Our Π -ML methodology consists of two components. (a) The dimensional analysis based on the Buckingham Π theorem combines observed dimensional variables into Π sets of normalized non-dimensional variables. (b) These sets are used to transform the observed data into a stacked non-dimensional dataset to train an ensemble of XGBoost regression models.

on the island of Hawai'i. The MLO study was conducted by the National Center for Atmospheric Research (NCAR) from 9 June 2006 until 8 August 2006 (~8 weeks). The publicly available dataset contains measurements of mean meteorological quantities, turbulent fluxes, and turbulent variances obtained from three sonic anemometers deployed at ca. 6m, 15m, and 25m altitude. C_n^2 values were estimated by NCAR via inertial-range scaling of temperature spectra (Oncley and Horst 2013). We compute two gradients from the mean horizontal wind components \bar{u} and \bar{v} and the mean potential temperature $\bar{\theta}$: mean wind shear $S = \sqrt{(\partial\bar{u}/\partial z)^2 + (\partial\bar{v}/\partial z)^2}$ and mean potential temperature gradient $\Gamma = \partial\bar{\theta}/\partial z$. Atmospheric turbulence is modulated by thermal buoyancy and wind shear effects, which are captured by the sensible heat flux $\overline{w'\theta'}$ and the friction velocity $u_* = \left(\overline{u'w'^2} + \overline{v'w'^2}\right)^{1/4}$, respectively; where $\overline{u'w'}$ and $\overline{v'w'}$ are momentum flux components. Additionally, we incorporate the variances of potential temperature and horizontal wind magnitude, σ_θ^2 and $\sigma_M^2 = \sigma_u^2 + \sigma_v^2$. The altitude z serves as length scale because we aim for a surface layer C_n^2 parametrization, but to model OT at higher altitudes, the suitable length scale is expected to differ. All relevant variables forming the input for our Π -ML methodology are summarized in the table of fig. 5.1a with their respective fundamental dimensions. Earth's gravitational acceleration $g = 9.81\text{ms}^{-2}$ is also included because it is required for the atmospheric force balance. Given the dry atmospheric conditions at the MLO sites, moisture variables were ignored. To later assess the C_n^2 estimation performance of the trained model, the first two weeks of July 2006 are set aside as test data. Although using data out of the middle as test data might seem unconventional, it is used so that the ML models can capture the seasonal change from June to August (see section 5.A for details).

The two key components of our proposed Π -ML methodology are illustrated in fig. 5.1: the DA constrained with domain knowledge (a) and the ensemble of gradient-boosting ML models, which perform regression on the stacked, non-dimensionalized observations (b). We set off with the table in fig. 5.1a and the Buckingham Π theorem (Stull 1988), popular in DA. The theorem states that our $k = 10$ dimensional variables with their $l = 3$ fundamental dimensions (length, time, temperature) can be expressed as a set of $(k - l) = 7$ independent non-dimensional Π groups. Multiple options exist to form these sets, so we employ the Π theorem implementation of (Karam and Saad 2021), which generates 71 sets with 7 Π groups each. Using domain knowledge, we conceive three constraints to reduce the number of sets from 71 to 14. First, each set can only contain one dependent Π group that is function of C_n^2 (cf. pink highlights in fig. 5.1a). All other Π groups should only be functions of the independent dimensional variables \mathbf{X} . Second, C_n^2 and its normalized variant Π_y vary over multiple orders of magnitude, so the ML models are trained on $\log_{10}\Pi_y$. Since the logarithm is not defined for negative arguments, only Π sets where Π_y is strictly positive are valid. Third, the dimensional variables Γ and $\overline{w'\theta'}$ can be positive and negative, so raising them to fractional or even-integer powers can result in complex values or a loss of sign. Therefore, valid Π sets cannot contain such expressions. Each of the 14 constrained Π sets is used to scale and non-dimensionalize the dimensional observations \mathbf{X} and $y = C_n^2$ to yield Π_X and Π_y ,

respectively, as illustrated in [fig. 5.1b](#). The non-dimensionalized observations from all three levels can be stacked into a combined dataset from which ML learns the non-dimensional black box similarity relationship $f(\mathbf{\Pi}_X) \approx \log_{10} \Pi_y$. For each Π set, we train one ensemble of $n = 25$ member models to make robust C_n^2 predictions with uncertainty estimates using the gradient boosting algorithm XGBoost (XGB) and the AutoML library FLAML (Wang *et al.* 2021). FLAML performs time-constrained hyperparameter tuning of the XGB models using 5-fold cross-validation. For each ensemble member, FLAML was given a 10-minute time budget on 8 cores of a 3 GHz Intel Xeon E5-6248R CPU. Such a time-constrained optimization is crucial to keep the overall training costs reasonable (~ 34 core hours per ensemble). We employ the Monte-Carlo resampling strategy to generate a different 4-week subset of the 6-week training data for each member. Two non-overlapping sets of seven consecutive days are randomly removed from the training data, so each subset covers different meteorological conditions. As depicted in [fig. 5.1b](#), each of the n trained members produces a prediction that is robustly aggregated into an ensemble prediction using the median.

The prediction accuracy and model complexity of each trained Π -ML ensemble is assessed to decide which Π set is best suited for our ML-based parameterization. The root-mean-squared error (RMSE) $\epsilon = \sqrt{\langle (y - \hat{y})^2 \rangle}$ in the log-space is utilized to quantify accuracy as the deviation between the observed $\log_{10} C_n^2 = y$ from the test set (July 1 – 14) and the corresponding Π -ML prediction $\hat{y} = \log_{10} \hat{C}_n^2$. We also evaluate the complexity of the Π sets and their trained ML ensembles. That is essential because ML models should only be as complex as necessary to increase their ability to perform well on new unseen data (Vapnik 1998). One Π set is considered simpler than another set if its Π groups are constructed from fewer dimensional variables. Similarly, one trained ensemble is considered simpler than another one if fewer Π groups are important for the ML prediction, i.e. the modeled C_n^2 is sensitive to fewer input features. The importance of input features of the trained Π -ML models is quantified by the permutation feature importance technique (PFI) (Molnar 2022). For each feature Π_i , PFI yields a ratio $(\epsilon'_i - \epsilon)/\epsilon$, which describes how the RMSE ϵ_i of a trained model magnifies when the model gets shuffled data for Π_i compared to the baseline RMSE ϵ where the correlation is intact. That means a highly important feature results in a large error magnification.

5.3. RESULTS

The performance and complexity of the 14 Π -ML ensembles are shown in [fig. 5.2](#). The boxplots in panel (a) display the ϵ distributions for each ensemble. While all ensembles show median RMSEs of the same order of magnitude, Π sets 9 to 14 outperform the others. Panel (b) visualizes complexity through the number of dimensional variables constituting each Π group (left) together with the sum per set (right). This plot reveals that sets 9 to 12 of the well-performing ensembles are the only ones consisting of Π groups formed from no more than three dimensional variables. These four low-error, low-complexity candidates are further assessed based on their PFI score distributions displayed in [fig. 5.3](#). Remember that the DA

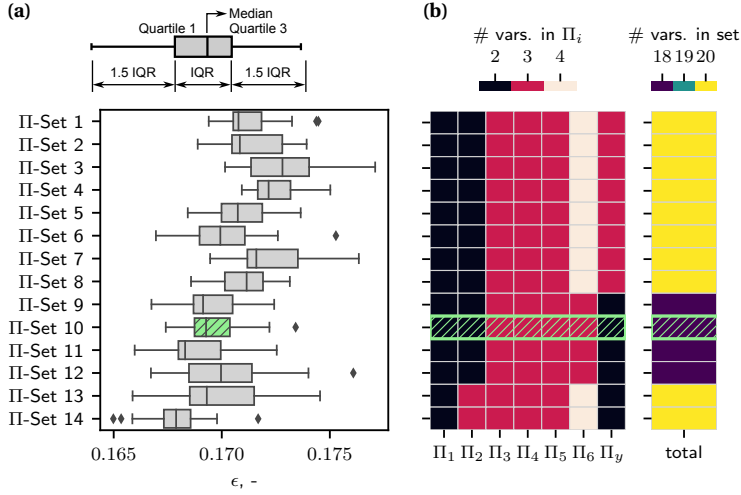


Figure 5.2.: Comparison of (a) ensemble performance and (b) Π set complexity for our 14 different Π sets, where winning set 10 (green/hatched) balances performance and complexity well.

yields different functional expressions for Π_i for each set, which is why each set shows different PFI distributions. The boxplots reveal that Π sets 9 and 11 yield more complex Π -ML ensembles compared to 10 and 12 because they significantly rely on two Π groups (Π_2 and Π_4 , see inset) for C_n^2 estimation instead of one (Π_2). Consequently, only sets 10 and 12 remain candidates for our ML-based similarity theory of optical turbulence. From these, we ultimately select Π set 10 because of the lower ϵ spread in fig. 5.2a with $\Pi_1 = \sigma_M^2/u_*^2$, $\Pi_2 = \bar{\theta}/\sqrt{\sigma_\theta^2}$, $\Pi_3 = (Sz)/u_*$, $\Pi_4 = \overline{w'\theta'}/(u_*\sigma_\theta)$, $\Pi_5 = (gz)/u_*^2$, $\Pi_6 = (\Gamma z)/\sigma_\theta$, and $\Pi_y = (C_n^2)^{3/2}z$. The expressions for the other 13 Π sets are listed in section 5.B. The observation that Π_2 – the inverse normalized potential temperature variance – is the only dominating feature of our parameterization could have practical implications: First, temperature variances can be measured with thermocouples (Albertson *et al.* 1995), which are cheaper than sonic anemometers. Second, the low relevance of the gradients (Π_3 and Π_6) indicates that even single-level measurements might be sufficient to estimate C_n^2 accurately. Therefore, our approach might lead to simpler C_n^2 measurement setups. In section 5.C, we confirm that retraining the models with Π_2 as the sole input feature still yields highly accurate predictions.

The performance of the final Π -ML ensemble is illustrated in more detail in figs. 5.4 and 5.5. The observed (red) and the predicted median evolutions of C_n^2 (black) for the test data are shown in fig. 5.4. The evolutions are plotted for the three original sonic heights individually for visualization. The agreement between prediction and observation is high for all levels, although the level-specific ϵ slightly increases with height. For nighttime conditions, the surface layer depth is typically shallower than 10–20 m. Thus, the topmost sonic anemometer at 25 m might be outside the surface

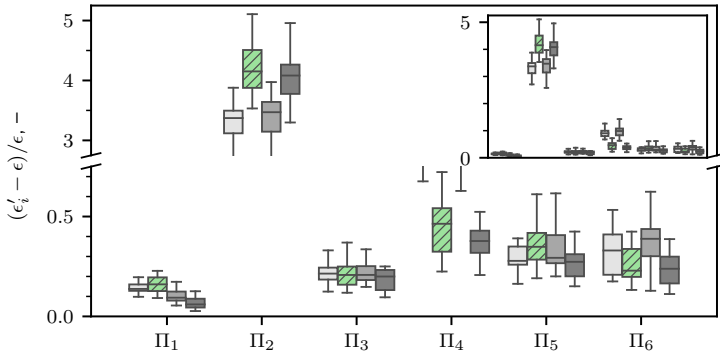


Figure 5.3.: Importance of non-dimensional Π groups (i.e., Π_1 to Π_6) based on the permutation feature importance strategy. For easy intercomparison, for each non-dimensional Π -group, four boxplots representing ML ensembles corresponding to Π -sets 9, 10, 11, and 12 (left to right) are plotted side-by-side. The best performing ensemble 10 is marked in green (hatched).

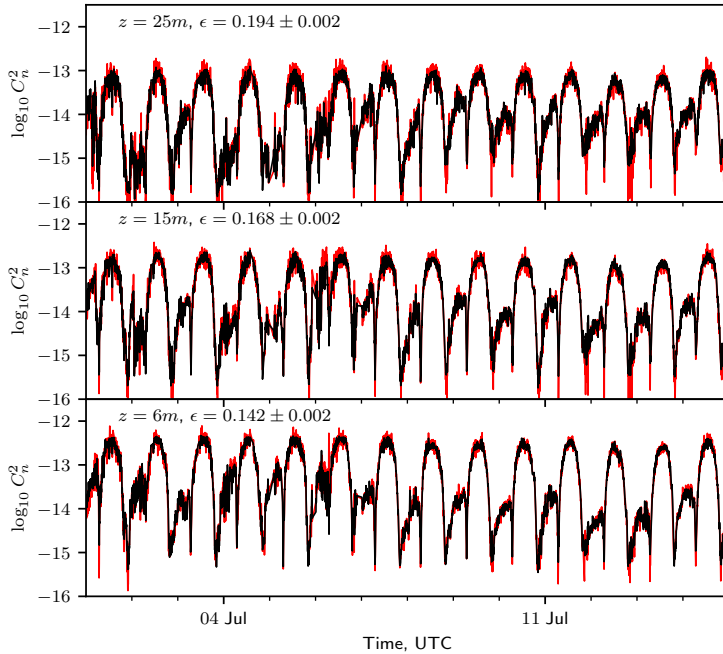


Figure 5.4.: Median predictions of $\log_{10} C_n^2$ based on test data (black) using the selected Π set 10 ensemble. The observed values (red) are shown for reference.

layer. In addition, outer layer effects such as wave-induced bursting events can force the turbulence underneath (Mahrt 2014). In such cases, cause (forcing) and effect (turbulence) are vertically separated, so the sonic signal only contains the effect

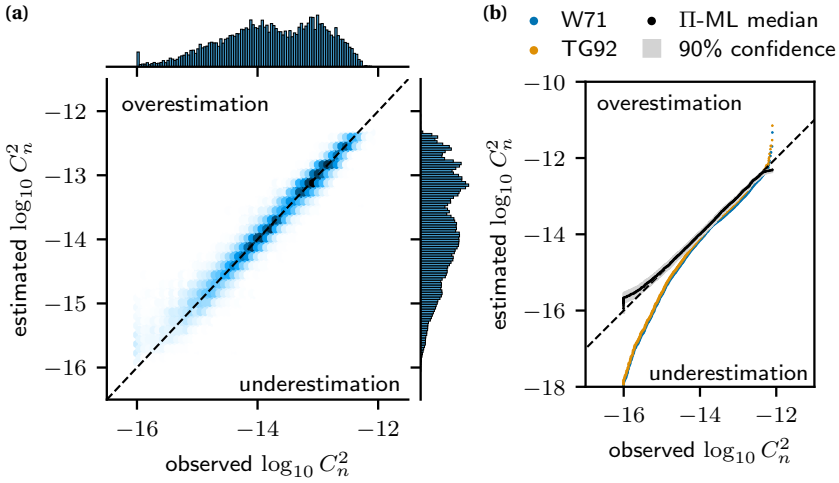


Figure 5.5.: Correlation histogram and quantile-quantile plot for Π set 10 ensemble showing (a) high correlation ($R^2 = 0.958 \pm 0.001$) and (b) well-captured C_n^2 distributions compared to traditional models from literature (blue, orange).

5

but not the cause. Thus, prediction accuracy decreases without additional upper-air information. Notable errors on all levels mostly occur during atmospheric neutral conditions shortly after sunrise and sunset where the observed C_n^2 drops as low as $10^{-16} \text{ m}^{-2/3}$. These drops are overestimated by our ensemble, which is also visible in the 2D correlation histogram of [fig. 5.5a](#) and the quantile-quantile (QQ) plot in [fig. 5.5b](#). Panel (a) directly compares observed C_n^2 samples to their ML-estimated counterpart, while panel (b) plots the cumulative density functions of observed and estimated C_n^2 against each other. The overestimation of neutral conditions is visible in both panels as the deviation of the histogram/curve from the ideal 1:1 line (dashed) for $C_n^2 < 10^{-15} \text{ m}^{-2/3}$. Simultaneously, the grey 90% confidence band in (b) grows, which indicates increasing disagreement between the predictions of the ensemble members. However, less than 8% of C_n^2 measurements are smaller than $10^{-15} \text{ m}^{-2/3}$, so the regularization of the ML training results in models that favor the center of the C_n^2 distribution, not its tails. Also, the lower signal-to-noise ratio of the sonic anemometers in weak turbulence conditions increases the measurement uncertainty ([Rannik *et al.* 2016](#)). Since very low turbulence conditions are also not critical for optical applications such as optical links or astronomy, we argue that little emphasis should be put on these deviations. The regularization mentioned above also explains the minor underestimation visible in panel (b) for observations with $C_n^2 > 10^{-12.5} \text{ m}^{-2/3}$, which make up less than 3.5% of the data. Leaving the tails of the distributions aside, both panels of [fig. 5.5](#) show excellent performance of our ensemble for most data. Most points in (a) the histogram and (b) the QQ plot are close to the ideal 1:1 line as quantified by the coefficient of determination of $R^2 = 0.958$ computed on all test data, including the deviating tails. The spread of the correlation distribution around the 1:1 line is symmetric for $C_n^2 > 10^{-15} \text{ m}^{-2/3}$.

That means the ensemble predictions are well-balanced and not biased towards over or underestimation for most of the C_n^2 range. A brief comparison of Π -ML with two conventional MOST-based C_n^2 parameterizations (W71 (Wyngaard *et al.* 1971) and TG92 (Thiermann and Grassl 1992)) in fig. 5.5b illustrates the potential of improvement by utilizing ML. While W71 and TG92 have the operational advantage of being formulated as analytical equations, they lack the flexibility to capture complex behavior where ML excels. This results in the larger over and underestimations shown in the QQ plots for these popular approaches. Comparing Π -ML to a more traditional ML approach based on Wang and Basu (2016) (see section 5.D) also shows significantly higher performance of Π -ML.

5.4. CONCLUSION

In summary, we demonstrated how dimensional analysis constrained with domain knowledge yields non-dimensional surface layer scaling expressions, which enable us to train accurate XGBoost regression models. Our approach has two advantages over C_n^2 parametrizations from the literature. First, the final ensemble produced highly accurate predictions for both daytime and nighttime, while previous models are often limited to one or the other (Smith *et al.* 1993). Second, we expect that the non-dimensional formulation allows making predictions with a pre-trained ensemble for new sonics set up at different heights or locations if the new non-dimensionalized data fall into the original non-dimensional training ranges. The data scaling should enable our ensemble to stay in the interpolation regime longer, i.e. cover a larger dimensional space, compared to traditional ML-based models. At this point, these claims are speculative in nature and need extensive validation. Our final Π -ML ensemble was shown to perform well, regardless of the complex meteorology of Hawaii (Businger *et al.* 2002) and the limited measurement duration of only two months. While the complexity and data sparsity of the MLO campaign limits the applicability of the trained ensemble to other sites, the good performance leads us to posit that our Π -ML methodology might perform well in more favorable setups. Additionally, we observed a strong dependency of C_n^2 on σ_θ^2 (Π_2), suggesting that relatively inexpensive single-level variance measurements might be sufficient for accurate C_n^2 estimation in the surface layer. In conclusion, we presented a powerful, statistically robust physics-informed machine learning methodology (Π -ML) to estimate C_n^2 from turbulence measurements.

ACKNOWLEDGMENTS

We are grateful to NCAR for making the MLO C_n^2 data publicly available. MP is funded by the FREE project (P19-13) of the TTW-Perspectief research program partially financed by the Dutch Research Council (NWO).

DATA AVAILABILITY

The code implementing the Π -ML methodology is available on GitHub: <https://github.com/mpierzyna/piml>.

5.A. PERFORMANCE EVALUATION USING DIFFERENT TEST SET INTERVALS

In the main text, a 14-day test set (July 1-14, 2006) was used to evaluate all the Π -ML and conventional models. This test set was extracted from the middle of the MLO study dataset; the rest of the dataset was utilized in training. This train-test split strategy enables the models to ‘learn’ seasonal trends in local weather patterns. Instead, if the test data were taken from the beginning or toward the end of the dataset, the trained Π -ML would need to extrapolate the trends rather than interpolate them. In fig. 5.6, we report the sensitivity of the model performance with respect to different test set intervals. Subplots (a) and (c) represent the ensemble performance if test data is taken from the beginning or toward the end of the dataset, respectively. Subplot (b) is a repetition of the boxplot from figure 2 of the main text for completeness. These plots reveal that the absolute root-mean-squared-error ϵ increases by $\sim 20\%$ for cases (a) and (c) from corresponding values in case (b). Similarly, the coefficient of determination R^2 of the selected Π -set 10 drops marginally from ca. 0.96 in case (b) to ca. 0.93 in cases (a) and (c). These degradations in performance are expected due to extrapolation beyond the training range.

5

5.B. FULL LIST OF NON-DIMENSIONAL Π SETS

Table 5.1 supplements fig. 5.2a and provides a full list of the non-dimensional expressions of the 14 Π sets. The Π_i groups are sorted by complexity just like in figure 2b, and the normalized C_n^2 target is denoted by Π_y .

5.C. PERFORMANCE OF A Π -ML MODEL OF REDUCED COMPLEXITY

Based on feature importance analysis, shown in fig. 5.3, we concluded that $\Pi_2 = \bar{\theta}/\sigma_\theta$ is the most dominant input feature predicting C_n^2 scaled as $\Pi_y = (C_n^2)^{3/2} z$. To further support this claim, we trained a new reduced-order Π -ML model with only Π_2 as input. The results are displayed in figure 5.7. These plots should be compared against figures 4 and 5 of the main text. The performance of the new reduced-order model deteriorates slightly compared to the baseline model, which includes all the Π groups from Π -set 10. The mean RMSE increases from $\epsilon = 0.170 \pm 0.001$ (baseline model) to $\epsilon = 0.220 \pm 0.000$ (reduced-order model), and mean correlation decreases from $R^2 = 0.958 \pm 0.001$ (baseline model) to $R^2 = 0.929 \pm 0.000$ (reduced-order model). Such drops in performance are expected because the contribution of Π_i other than Π_2 , in the baseline model is small but not negligible. Nonetheless, the overall

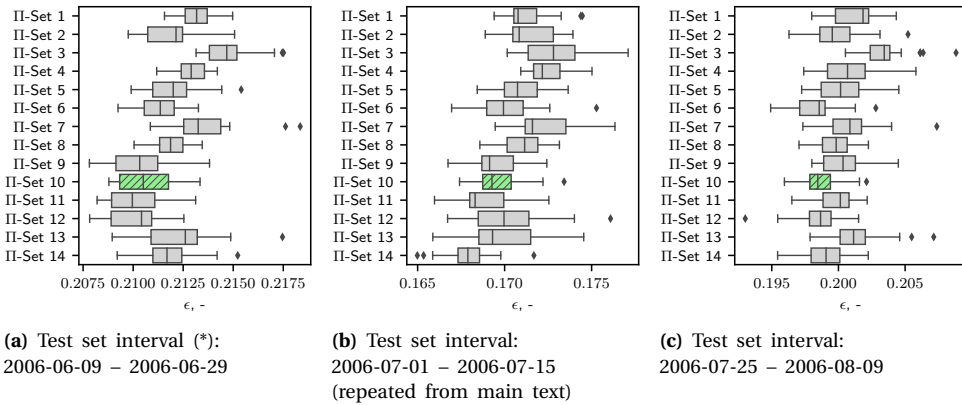


Figure 5.6.: Comparison of model performance when three different test set intervals are used. Each interval contains 14 days of data for testing. The first interval marked with (*) contains a few days of missing data, so the window is large enough to contain 14 full valid days.

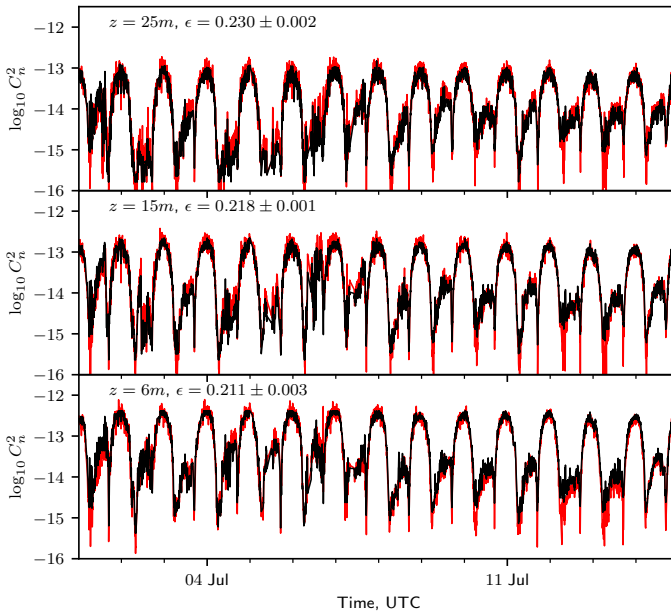
performance of the new, reduced-order Π -ML model is very satisfactory; it is surprising that the XGBoost ensemble can predict surface layer C_n^2 values from just σ_θ and $\bar{\theta}$ as inputs. We would like to emphasize that this result does not mean that the other Π groups carry no physical importance or meaning. It only demonstrates that Π_2 alone already contains enough information for our Π -ML model to accurately predict C_n^2 .

5.D. PERFORMANCE OF AN XGBOOST ENSEMBLE TRAINED WITH TRADITIONAL FEATURES

To corroborate our claim that non-dimensional features are better suited for C_n^2 prediction than dimensional ones, we trained an XGBoost ensemble based on traditional features using our XGBoost ensemble strategy. The traditional dimensional inputs are 2m-temperature, 2m-relative humidity, 2m-pressure, 15m- Γ , and 15m- S to predict the 15m- C_n^2 as proposed by Wang and Basu (2016). The performance of the resulting model is shown in fig. 5.8. Instead of the multilayer perceptron used by the original authors, we use our XGBoost ensemble methodology, which we also use for Π -ML. Although this approach leads to a reasonable result, the root-mean-square error ϵ using the traditional features is almost twice as high as for our best Π -ML ensemble, and R^2 is also notably lower. Additionally, the non-dimensional scaling enables Π -ML to predict C_n^2 independent of the z level, while the traditional approach presented in fig. 5.8 requires to train one model per z level.

Table 5.1.: Expressions of all non-dimensional Π sets for input features (Π_i) and respective target variable Π_y (scaled C_n^2). Note that σ_X and σ_X^2 denote the standard deviation and variance of quantity X , respectively. The final set selected for its balance of ensemble prediction accuracy and complexity is marked with (*).

Set	Π_1	Π_2	Π_3	Π_4	Π_5	Π_6	Π_y
Set 1	$\frac{\sigma_M^2}{u_*^2}$	$\frac{\sigma_\theta^2}{\theta^2}$	$\frac{S\bar{z}}{u_*}$	$\frac{\overline{w'\theta'}}{\theta u_*}$	$\frac{g}{S u_*}$	$\frac{\Gamma u_*}{S\theta}$	$\frac{u_* (C_n^2)^{3/2}}{S}$
Set 2	$\frac{\sigma_M^2}{u_*^2}$	$\frac{\bar{\theta}}{\sigma_\theta}$	$\frac{S\bar{z}}{u_*}$	$\frac{\overline{w'\theta'}}{u_* \sigma_\theta}$	$\frac{g}{S u_*}$	$\frac{\Gamma u_*}{S\sigma_\theta}$	$\frac{u_* (C_n^2)^{3/2}}{S}$
Set 3	$\frac{\sigma_M^2}{u_*^2}$	$\frac{\sigma_\theta^2}{\theta^2}$	$\frac{g\bar{z}}{u_*^2}$	$\frac{S u_*}{g}$	$\frac{\overline{w'\theta'}}{\theta u_*}$	$\frac{\Gamma u_*^2}{g\theta}$	$\frac{u_*^2 (C_n^2)^{3/2}}{g}$
Set 4	$\frac{\sigma_M^2}{u_*^2}$	$\frac{\bar{\theta}}{\sigma_\theta}$	$\frac{g\bar{z}}{u_*^2}$	$\frac{S u_*}{g}$	$\frac{\overline{w'\theta'}}{u_* \sigma_\theta}$	$\frac{\Gamma u_*^2}{g\sigma_\theta}$	$\frac{u_*^2 (C_n^2)^{3/2}}{g}$
Set 5	$\frac{u_*}{\sigma_M}$	$\frac{\sigma_\theta^2}{\theta^2}$	$\frac{S\bar{z}}{\sigma_M}$	$\frac{\overline{w'\theta'}}{\theta \sigma_M}$	$\frac{g}{S\sigma_M}$	$\frac{\Gamma \sigma_M}{S\theta}$	$\frac{\sigma_M (C_n^2)^{3/2}}{S}$
Set 6	$\frac{u_*}{\sigma_M}$	$\frac{\bar{\theta}}{\sigma_\theta}$	$\frac{S\bar{z}}{\sigma_M}$	$\frac{\overline{w'\theta'}}{\sigma_\theta \sigma_M}$	$\frac{g}{S\sigma_M}$	$\frac{\Gamma \sigma_M}{S\sigma_\theta}$	$\frac{\sigma_M (C_n^2)^{3/2}}{S}$
Set 7	$\frac{u_*}{\sigma_M}$	$\frac{\sigma_\theta^2}{\theta^2}$	$\frac{g\bar{z}}{\sigma_M^2}$	$\frac{S\sigma_M}{g}$	$\frac{\overline{w'\theta'}}{\theta \sigma_M}$	$\frac{\Gamma \sigma_M^2}{g\theta}$	$\frac{\sigma_M^2 (C_n^2)^{3/2}}{g}$
Set 8	$\frac{u_*}{\sigma_M}$	$\frac{\bar{\theta}}{\sigma_\theta}$	$\frac{g\bar{z}}{\sigma_M^2}$	$\frac{S\sigma_M}{g}$	$\frac{\overline{w'\theta'}}{\sigma_\theta \sigma_M}$	$\frac{\Gamma \sigma_M^2}{g\sigma_\theta}$	$\frac{\sigma_M^2 (C_n^2)^{3/2}}{g}$
Set 9	$\frac{\sigma_M^2}{u_*^2}$	$\frac{\sigma_\theta^2}{\theta^2}$	$\frac{S\bar{z}}{u_*}$	$\frac{\overline{w'\theta'}}{\theta u_*}$	$\frac{g\bar{z}}{u_*^2}$	$\frac{\Gamma \bar{z}}{\theta}$	$(C_n^2)^{3/2} \bar{z}$
Set 10*	$\frac{\sigma_M^2}{u_*^2}$	$\frac{\bar{\theta}}{\sigma_\theta}$	$\frac{S\bar{z}}{u_*}$	$\frac{\overline{w'\theta'}}{u_* \sigma_\theta}$	$\frac{g\bar{z}}{u_*^2}$	$\frac{\Gamma \bar{z}}{\sigma_\theta}$	$(C_n^2)^{3/2} \bar{z}$
Set 11	$\frac{u_*}{\sigma_M}$	$\frac{\sigma_\theta^2}{\theta^2}$	$\frac{S\bar{z}}{\sigma_M}$	$\frac{\overline{w'\theta'}}{\theta \sigma_M}$	$\frac{g\bar{z}}{\sigma_M^2}$	$\frac{\Gamma \bar{z}}{\theta}$	$(C_n^2)^{3/2} \bar{z}$
Set 12	$\frac{u_*}{\sigma_M}$	$\frac{\bar{\theta}}{\sigma_\theta}$	$\frac{S\bar{z}}{\sigma_M}$	$\frac{\overline{w'\theta'}}{\sigma_\theta \sigma_M}$	$\frac{g\bar{z}}{\sigma_M^2}$	$\frac{\Gamma \bar{z}}{\sigma_\theta}$	$(C_n^2)^{3/2} \bar{z}$
Set 13	$\frac{\sigma_\theta^2}{\theta^2}$	$\frac{u_*}{\sqrt{g}\sqrt{\bar{z}}}$	$\frac{S\sqrt{\bar{z}}}{\sqrt{g}}$	$\frac{\sigma_M^2}{g\bar{z}}$	$\frac{\Gamma \bar{z}}{\theta}$	$\frac{\overline{w'\theta'}}{\sqrt{g}\theta\sqrt{\bar{z}}}$	$(C_n^2)^{3/2} \bar{z}$
Set 14	$\frac{\bar{\theta}}{\sigma_\theta}$	$\frac{u_*}{\sqrt{g}\sqrt{\bar{z}}}$	$\frac{S\sqrt{\bar{z}}}{\sqrt{g}}$	$\frac{\sigma_M^2}{g\bar{z}}$	$\frac{\Gamma \bar{z}}{\sigma_\theta}$	$\frac{\overline{w'\theta'}}{\sqrt{g}\sigma_\theta\sqrt{\bar{z}}}$	$(C_n^2)^{3/2} \bar{z}$



(a) Median predictions (black) compared to test set data (red) per z-level.

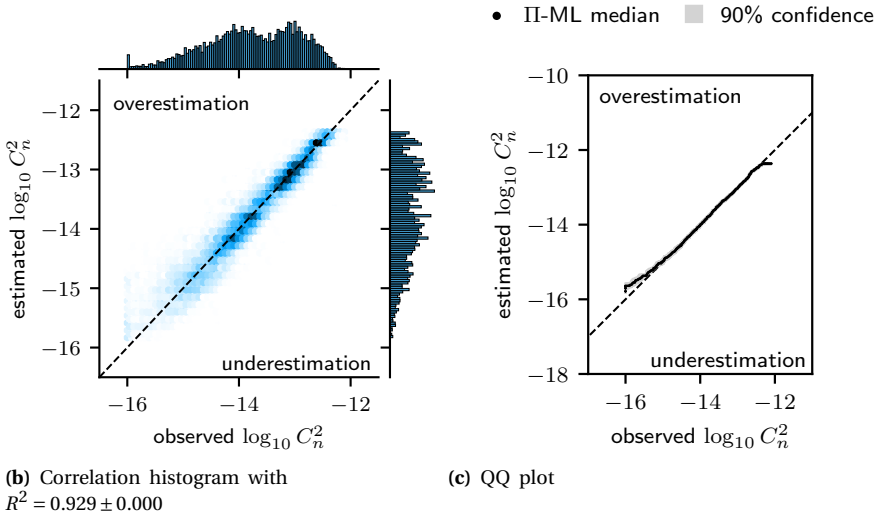


Figure 5.7.: Performance of Π -ML model trained with $\Pi_2 = \bar{\theta}/\sigma_\theta$ as sole feature and $\Pi_y = (C_n^2)^{3/2} z$ as target. Refer to figs. 5.4 and 5.5 where the model performance of the full Π -set 10 model is displayed in equivalent formats.

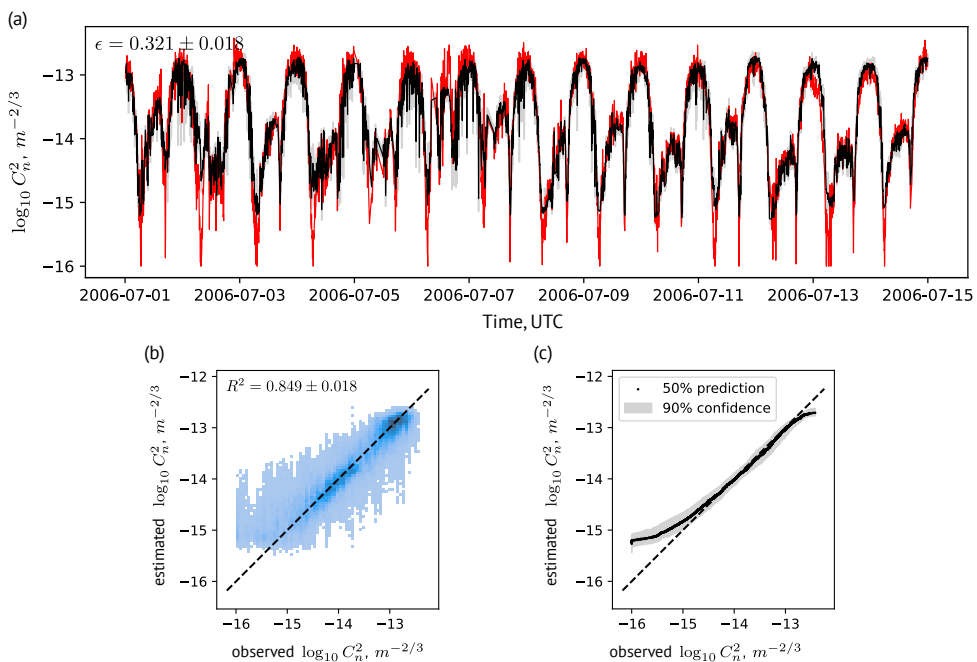


Figure 5.8.: XGBoost ensemble trained on traditional features (2m-temperature, 2m-relative humidity, 2m-pressure, 15m- Γ , 15m- S) as proposed by Wang and Basu (2016) to predict 15m- C_n^2 . The predicted (black) and observed (red) 15m time series are shown in (a) with correlation plot and quantile-quantile plot in panels (b) and (c).

6

Beyond Hufnagel-Valley: estimating high-resolution profiles of optical turbulence using deep learning

The machine learning (ML) approaches presented so far are not applicable beyond the lower atmospheric boundary layer. For satellite-to-ground links, however, vertical C_n^2 profiles through the full atmosphere are needed. This chapter develops a deep learning model dubbed OTProf, which estimates high-resolution C_n^2 profiles (100 levels) from coarse-resolution ERA5 reanalysis data (30 pressure levels), performing both regression and vertical super-resolution. Training data are generated from a year-long mesoscale simulation, so the model learns from physics-based turbulence estimates while being cheap to evaluate. The estimated profiles substantially outperform the widely used Hufnagel-Valley (HV) model. In contrast to HV, the estimated profiles from OTProf have a physically realistic vertical structure, which is highly relevant for estimating integrated parameters such as the Fried parameter r_0 or the scintillation index σ_I^2 . However, the ML-typical smoothing of fine-scale features in the estimated profiles compared to the training data is also visible in OTProf, leading to overestimated (underestimated) r_0 (σ_I^2).

6.1. INTRODUCTION

Vertical profiles of optical turbulence, quantified by the refractive index structure parameter C_n^2 , are essential for characterizing the performance of free-space optical communication (FSOC) systems and ground-based astronomical observations. These profiles determine key parameters such as the Fried parameter r_0 and the scintillation index σ_I^2 , which directly impact beam spreading, wavefront distortion, and signal fading (Andrews and Phillips (2005), Kaushal and Kaddoum (2017), and chapter 3). To predict optical turbulence conditions, the community has traditionally relied on empirical models such as the non-parametric SLC models (Miller and Zieske 1979), the parametric Hufnagel-Valley (HV) model (Hufnagel 1974; Valley 1980; Ulrich 1988), and others (cf. Good *et al.* (1988) or Smith *et al.* (1993) for an overview). However, these models have a limited ability to account for local meteorology. The SLC models, for example, are non-parametric, so their C_n^2 estimates do not depend on location, time, or local meteorology at all. The HV model has a weak dependence on local meteorology but still prescribes a largely generic vertical structure. Efforts are still being made to adjust HV-based models better to local boundary layer conditions (e.g., Comeron *et al.* (2005), Andrews *et al.* (2009), Stotts and Andrews (2023), and Dasgupta *et al.* (2026)), but the generic exponential shape of the models remains a fundamental limitation. Alternatively, numerical weather prediction models like the Weather Research and Forecasting (WRF) model (Skamarock *et al.* 2021) can provide physically consistent high-resolution C_n^2 profiles, which do account for local conditions (e.g., Masciadri *et al.* (1999), Cherubini *et al.* (2008), Giordano *et al.* (2014), Basu *et al.* (2020), and Rafalimanana *et al.* (2022)), but they are computationally expensive to obtain. Given these limitations, there is a clear need for more efficient and physically consistent approaches to predict optical turbulence profiles from meteorological data.

This study addresses this need by proposing OTProf, a machine learning approach to estimate high-resolution (high-res) vertical C_n^2 profiles from coarse-resolution (low-res) meteorological data, such as the globally available ERA5 reanalysis. This approach requires solving two tasks simultaneously: regression and super-resolution. The regression task is to predict optical turbulence (C_n^2) from meteorological variables, while super-resolution aims to achieve a vertical output resolution of the predicted profiles that exceeds the vertical input resolution of the meteorological data. As a proof of concept, this study focuses on the Netherlands, for which we generate a year-long high-resolution WRF model-based training dataset. The ERA5 reanalysis (Hersbach *et al.* 2020) pressure-level dataset is used as low-res meteorological input data. The Squeezeformer architecture (Kim *et al.* 2022) is used to learn a mapping from low-res ERA5 meteorological profiles to the high-res WRF C_n^2 profiles. We suggest applying quantile mapping (QM) (Cannon *et al.* 2015) to statistically align the low-res ERA5 data with the WRF training data prior to inference, thereby improving the accuracy of the predicted C_n^2 profiles. The model's performance is benchmarked in three experiments of increasing complexity: (a) regression, (b) regression and super-resolution, (c) regression, super-resolution, and distribution shift from ERA5 to WRF. All experiments are compared against the traditional HV model (Hufnagel 1974) with boundary layer extension (Valley 1980;

Ulrich 1988). For practical interpretation, we also quantify performance as the ability of the estimated profiles to capture derived integrated parameters r_0 and σ_T^2 .

6.2. METHODS

Estimating fine, high-resolution (high-res) vertical C_n^2 profiles from coarse, low-resolution (low-res) meteorological profiles is a regression and super-resolution task. For a single time and location, we consider a set of p profile variables (features) $\mathbf{X} \in \mathbb{R}^{p \times m}$ given at m levels from which we aim to estimate r profile variables (targets) $\mathbf{Y} \in \mathbb{R}^{r \times n}$ at n levels where $m < n$. The input \mathbf{X} contains meteorological profiles, such as temperature and wind speed, while the target profiles \mathbf{Y} contain $C_n^2(z)$, which is not part of the inputs \mathbf{X} . As OT close to the surface is primarily modulated by surface parameters, such as surface heat and momentum flux, we also consider a vector of q surface variables, $\mathbf{x}_s \in \mathbb{R}^q$, resulting in $p+q$ total input variables. The prediction task can then be formulated as

$$f(\mathbf{X}, \mathbf{x}_s) = \hat{\mathbf{Y}} \approx \mathbf{Y} \quad (6.1)$$

at every time and location. Here, the spatio-temporal dependency is implicit as f does not aim to model a temporal evolution of \mathbf{Y} .

In practice, \mathbf{X} and \mathbf{Y} are typically not produced by the same model—for example, \mathbf{X} may be obtained from a global reanalysis (e.g., ERA5) while \mathbf{Y} comes from high-resolution regional simulations (e.g., WRF). This leads to spatio-temporal misalignment between datasets, which we address through a decoupled training approach combined with quantile mapping (Cannon *et al.* 2015). This decoupled pipeline is detailed in section 6.2.1. The Squeezeformer (Kim *et al.* 2022) DL architecture utilized in this study to obtain f is described in section 6.2.2. The two baseline models used to benchmark the Squeezeformer, including the Hufnagel-Valley model as a lower baseline, are introduced in section 6.2.3. Finally, section 6.2.4 summarizes the evaluation metrics and strategy used to assess model performance.

6.2.1. DATASET ALIGNMENT AND DECOUPLED TRAINING

Training machine learning models on datasets generated by different numerical models at varying resolutions poses challenges due to spatiotemporal misalignment. Regional models, such as the WRF model, drift with increasing simulation time relative to their forcing data because the forcing is applied only at the boundaries, while the model evolves independently inside the domain. Additionally, differences in physics parameterizations and resolutions across models lead to structural differences. For example, the wind field around complex terrain can differ significantly between a coarse global model and a high-resolution regional simulation. An example is given in fig. 6.1, where the 10 m wind field around islands along the Dutch coast is presented. In the high-res wind field (panel b), a low-velocity wake is visible in the lee of the islands and the coast, together with terrain-induced changes of wind direction indicated by the white stream lines. While

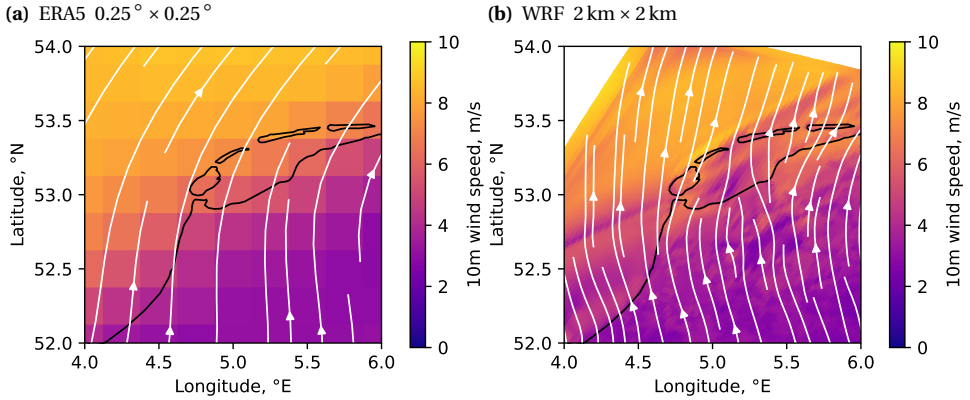


Figure 6.1.: Example of mismatching 10m wind fields due to higher spatial resolution of WRF compared to ERA5.

the low-res wind field (panel a) captures the general pattern of wind magnitude evolution and direction, these fine-scale features are absent. Such misalignments are inherent and cannot be fully avoided, but we aim to partially mitigate them through decoupled training and bias correction, as detailed below.

To disentangle regression and super-resolution from these misalignment challenges, we employ a decoupled training approach as shown in [fig. 6.2a](#). Instead of directly training the Squeezeformer on ERA5 data as input and WRF data as the target, we train the model only on WRF data. Specifically, we vertically coarsen the high-resolution WRF dataset to a low-res dataset that emulates ERA5 in vertical resolution and variable selection. Coarsening is achieved by vertically interpolating high-res WRF profiles onto the low-res ERA5 pressure levels in log space. During training, the model learns to map low-res WRF inputs to native high-res WRF targets, thereby learning the regression and super-resolution tasks on a consistent dataset.

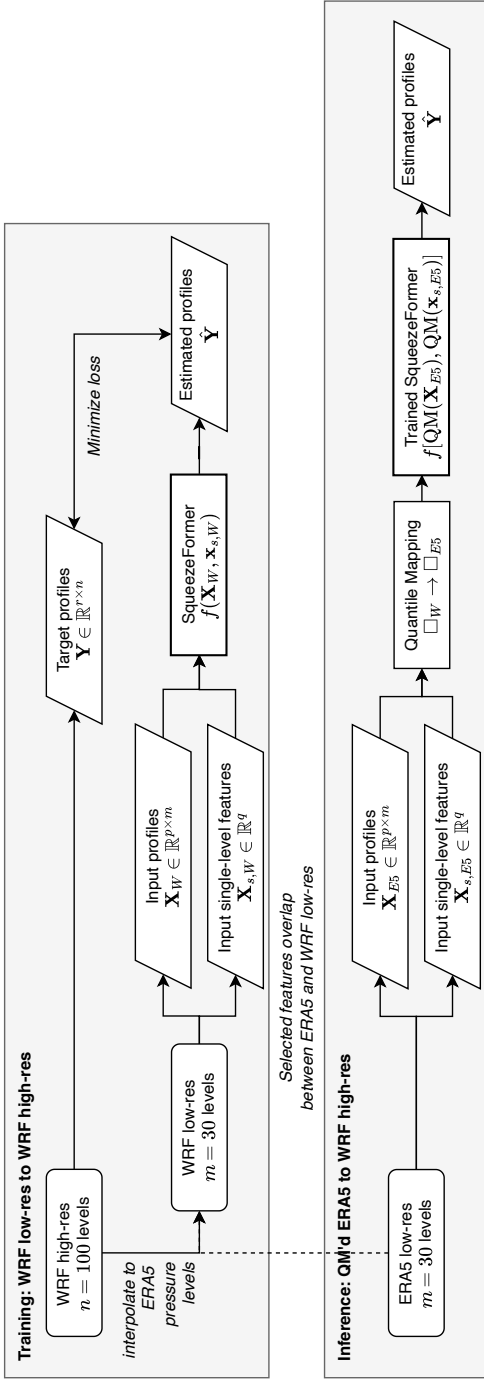
For inference with ERA5, we apply quantile mapping (QM) ([Cannon *et al.* 2015](#)) to statistically align the ERA5 input distributions to those of the low-res WRF training inputs. QM transforms an ERA5 variable x by matching its quantiles to the WRF distribution:

$$\text{QM}(x) = F_W^{-1}(F_{E5}(x)), \quad (6.2)$$

where F_{E5} and F_W are the empirical cumulative distribution functions of ERA5 and WRF, respectively, and F^{-1} denotes the inverse CDF. To avoid data leakage, both CDFs are computed only on training data. During inference, each ERA5 variable is transformed globally across all locations and times using the pre-determined quantiles. Formally, we consider two datasets $(\mathbf{X}_W, \mathbf{x}_{s,W}, \mathbf{Y})$ and $(\mathbf{X}_{E5}, \mathbf{x}_{s,E5})$ based on WRF and ERA5, respectively. During training, the DL model uses WRF data and is trained as $f: \mathbf{X}_W, \mathbf{x}_{s,W} \rightarrow \mathbf{Y}$. During inference, the trained model makes predictions based on quantile-mapped ERA5 data as $f: \text{QM}(\mathbf{X}_{E5}), \text{QM}(\mathbf{x}_{s,E5}) \rightarrow \hat{\mathbf{Y}}$.

This statistical alignment enables the model trained on WRF to make predictions

(a) Decoupled training pipeline – training on WRF data, inferring on quantile-mapped ERA5 data



(b) Squeezeformer architecture with added super-resolution. The Squeezeformer convolution and transformer blocks are abbreviated here and presented in detail in fig. 6.9.

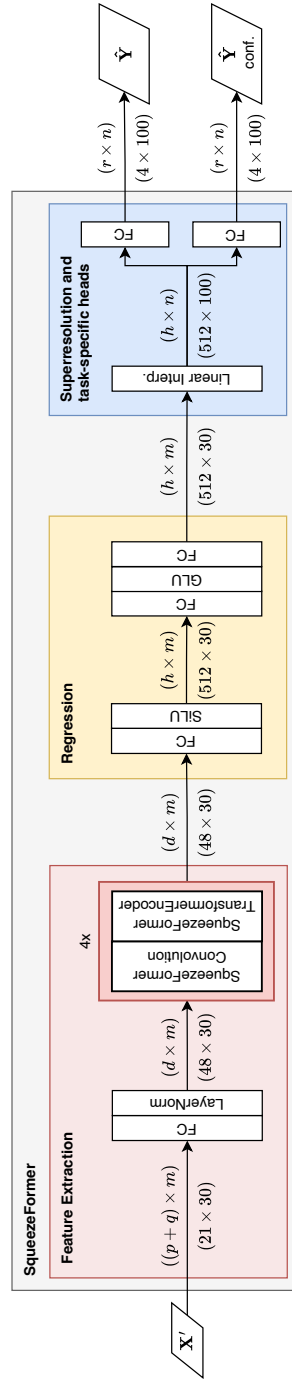


Figure 6.2.: Overview of training and inference pipeline of OTProf utilizing the Squeezeformer architecture.

on ERA5 data in a feature space similar to the training distribution. The effectiveness of this approach compared to using uncorrected ERA5 or direct training on ERA5 is demonstrated in supplement 6.A. While this alignment improves model performance inside the training domain, it may reduce performance for unseen locations. We recommend future work to study the generalization of trained OTRProf models to new domains and assess the effectiveness of QM in these situations.

6.2.2. SQUEEZEFORMER

The Squeezeformer architecture used in this work was originally proposed by Kim *et al.* (2022) and has been successfully used in multiple Kaggle machine learning competitions (Henkel 2023; Sohn 2024; Ron 2024). The core concept of the Squeezeformer is to combine the local feature extraction capabilities of convolutional neural networks with the long-range dependency modeling of transformer architectures (Kim *et al.* 2022). We consider this combination particularly well-suited to modeling turbulence in the atmospheric column because turbulence is modulated by both local gradients and larger-scale atmospheric structures (Stull 1988). For the present study, the Squeezeformer variant of Ron (2024) is utilized, which modified the architecture from previous competitions for the regression of atmospheric variables. We extend the architecture for super-resolution by performing interpolation in the feature space, as detailed below.

A schematic of the extended architecture is presented in fig. 6.2b, visualizing how low-res input profiles are processed to yield high-res outputs with corresponding confidence estimates. Following Ron (2024), we incorporate the q surface variables \mathbf{x}_s by expanding them along the vertical dimension into profiles of constant values before concatenating them to the meteorological profiles \mathbf{X} resulting in $\mathbf{X}' \in \mathbb{R}^{(p+q) \times m}$. The inputs are first embedded into a higher-dimensional space $\mathbb{R}^{d \times m}$ with $d > (p+q)$ using a fully connected layer (FC) with layer normalization before being processed by the sequence of Squeezeformer blocks composed of convolution and transformer blocks (red). A detailed description of the Squeezeformer blocks is given in section 6.B but is skipped here for brevity. The extracted features are embedded again into a high-dimensional regression space $\mathbb{R}^{h \times m}$ by several FC layers with sigmoid linear unit (SiLU) (Ramachandran *et al.* 2017) and gated linear unit (GLU) activations (Dauphin *et al.* 2017) (yellow). The gated activation enables the model to learn to suppress irrelevant information and focus on important features. Finally, the high-res output profiles are obtained by linearly interpolating the regression output from m to n levels in feature space. Two parallel task-specific regression heads (FC layers) bring the dimensionality of h down to the r target variables. The output heads yield the values of the target profiles (\mathbf{Y}) and corresponding confidence estimates (\mathbf{Y}_{conf}).

Super-resolution via linear interpolation is performed as late as possible in the architecture, since we aim to extract as much information as possible from the low-res inputs. Interpolating earlier could smooth out important features, such as sharp gradients in wind speed and temperature, that are relevant to OT. Also, atmospheric model data is often given at constant pressure levels instead of constant height. These pressure levels are relative to the local surface pressure, so the vertical

spacing between levels varies with location and time. By interpolating after the regression module, we enable the model to process these variable spacings before interpolating to the target grid, which often also has variable spacing. The late interpolation is also computationally efficient because the early layers do not need to process high-res data. While linear, we stress that the interpolation is performed in a high-dimensional feature space. Consequently, the FC layers forming the output heads learn to utilize the interpolated high-res features for projecting into the final high-res target spaces.

TRAINING PROCEDURE

The Squeezeformer is trained to minimize a two-part loss function comprising a regression loss and a confidence loss that serves as regularization (Ron 2024). Both loss components are based on the root-mean-squared error (RMSE), where the regression loss minimizes the squared error between predicted and reference profiles, $\boldsymbol{\epsilon}_i = (\mathbf{Y}_i - \hat{\mathbf{Y}}_i)^2 \in \mathbb{R}^{r \times n}$:

$$L_{rgr} = \sqrt{\langle \boldsymbol{\epsilon}_i \rangle}. \quad (6.3)$$

Here, the index i indicates a single sample of target profiles and $\langle \cdot \rangle$ denotes the mean over all target variables, levels, and samples. The task of the confidence head is to estimate the squared error $\boldsymbol{\epsilon}_i$ as a confidence estimate, so the confidence loss is defined as the RMSE between $\boldsymbol{\epsilon}_i$ and $\mathbf{Y}_{conf,i}$:

$$L_{conf} = \sqrt{\langle (\boldsymbol{\epsilon}_i - \mathbf{Y}_{conf,i})^2 \rangle}. \quad (6.4)$$

The total loss is the sum of regression and confidence loss, $L = L_{rgr} + L_{conf}$, and is minimized using the AdamW (Loshchilov and Hutter 2019) optimizer. Cosine annealing with warm restarts (Loshchilov and Hutter 2017) is used to schedule the learning rate during training, and early stopping based on the validation loss is applied to prevent overfitting.

6.2.3. BASELINE MODELS

To benchmark the performance of the Squeezeformer, we consider two baseline models that represent the lower and upper performance bounds. The upper baseline is a Squeezeformer trained and evaluated on high-res data only, representing the best-case performance of the architecture without super-resolution. For this baseline, the input variables (profiles and surface) are the same as for the main model, so the regression task to estimate high-res C_n^2 profiles from standard meteorological variables is identical.

As a lower baseline, we consider the Hufnagel-Valley (HV) model (Hufnagel 1974; Valley 1980; Ulrich 1988) combined with the Wyngaard *et al.* (1971) parameterization for surface C_n^2 . The HV model is a simple analytical, empirical, and parametric model widely used in the optical turbulence community (e.g., Andrews *et al.* (2009), Dimitrov *et al.* (2016), Camboulives *et al.* (2018), Osborn *et al.* (2021), and Walsh and

Schediwy (2023)). Following Smith *et al.* (1993), the HV model is given as

$$C_n^2(\bar{z}) = c_1 W^2 \bar{z}^{10} \exp(-\bar{z}) + c_2 \exp\left(-\frac{\bar{z}}{1.5}\right) + \exp\left(-\frac{\bar{z}}{0.1}\right) (C_n^2|_{z=0} - c_2) \quad (6.5)$$

where \bar{z} is the height above ground in km, and c_1 and c_2 are constants with $c_1 = 8.2 \times 10^{-26} \text{ m}^{-2/3}$ and $c_2 = 2.7 \times 10^{-16} \text{ m}^{-2/3}$, respectively. The parameter W represents the root-mean-square (RMS) wind speed between 5 km and 20 km height. Considering a profile of absolute wind speed $m(\bar{z})$, W is computed as

$$W = \left(\frac{1}{15} \int_5^{20} m(\bar{z})^2 d\bar{z} \right)^{1/2}. \quad (6.6)$$

Equation (6.5) also requires a surface estimate of C_n^2 which we obtain following the flux-based approach of Wyngaard *et al.* (1971) (cf. eq. (2.3)):

$$C_n^2|_{z=0} = \left(A \frac{p_0}{T_2^2} \right)^2 T_*^2 z^{-2/3} g(\zeta). \quad (6.7)$$

The temperature scale $T_* = -\overline{w'\theta'}/u_*$ is computed using the friction velocity u_* and the kinematic sensible heat flux $\overline{w'\theta'}$ at the surface, while p_0 is the surface pressure in hPa, T_2 is the temperature at 2 m height in K, and z is the measurement height in m. $A \approx 7.9 \times 10^{-5} \text{ K hPa}^{-1}$ is almost constant for optical wavelengths (Andrews and Phillips 2005) and $g(\zeta)$ is a stability dependent similarity function with $\zeta = z/L$ and Obukhov length L (cf. eq. (2.4)). We call this combination HV+W71 for the rest of the manuscript.

All meteorological variables required by the HV model are typically available from atmospheric datasets. Except for the wind profile to compute W , only surface variables are needed. To fairly compare the Squeezeformer against HV+W71, we utilize HV+W71 in a “superresolution mode” where W is computed from low-res profiles of $m(\bar{z})$ but, eq. (6.5) is evaluated at the \bar{z} positions of the high-res levels.

6.2.4. MODEL EVALUATION

The performance of the fitted models is assessed with respect to their ability to capture the high-res reference C_n^2 profiles and two integrated astroclimatic parameters derived from these profiles, the Fried parameter r_0 and the scintillation index σ_I^2 . Both parameters are employed for their practical relevance to the OT community (Smith *et al.* 1993; Hardy 1998) and because they weight different parts of the C_n^2 profiles differently. The different weighting allows us to assess how well different parts of the profiles are captured. To quantify the agreement between predictions and reference data, we employ four metrics: bias, centered root-mean-square error (cRMSE), Pearson correlation coefficient (r), and coefficient of determination (R^2). Additionally, a structure-function analysis is performed to assess how well the models capture the vertical variability of the C_n^2 profiles across different vertical scales. All parameters, metrics, and the structure-function analysis are detailed below.

INTEGRATED ASTROCLIMATE PARAMETERS

Fried Parameter The Fried parameter (Fried 1966) (in meters) is a measure to determine the strength of wavefront distortions caused by OT. If the ratio of receiver aperture diameter to Fried parameter, D/r_0 , is less than one, an optical system operates close to its theoretical optimum, i.e., is diffraction-limited. At the same time, turbulence-induced distortions degrade the performance for $D/r_0 > 1$. Assuming a vertical profile of C_n^2 , $C_n^2(z)$, the Fried parameter is given as (Andrews and Phillips 2005)

$$r_0 = \left(0.423 k^2 \sec(\theta_z) \int_0^H C_n^2(z) dz \right)^{-3/5}, \quad (6.8)$$

where $k = 2\pi/\lambda$ is the wavenumber with $\lambda = 1550$ nm and H is the propagation distance. The zenith angle θ_z accounts for slant paths deviating from a direct zenith path of $\theta_z = 0$. As OTProf aims at estimating C_n^2 along a one-dimensional vertical column, we always compute r_0 for $\theta_z = 0$ for evaluation. Alternatively, r_0 along a slant path can also be obtained by estimating a C_n^2 cross-section as multiple columns containing the line-of-sight and integrating along that path.

Scintillation Index The scintillation index, σ_I^2 , expresses the normalized variance of received optical intensity fluctuations (Andrews and Phillips 2005)

$$\sigma_I^2 = \frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}, \quad (6.9)$$

where $I(t)$ is the received intensity/irradiance signal and $\langle \square \rangle$ denotes the ensemble average over a given time period.

A theoretical connection between σ_I^2 and the $C_n^2(z)$ profile can be made by assuming a Kolmogorov turbulence spectrum and Rytov theory. We first define the Rytov variance for plane waves as (Andrews and Phillips 2005)

$$\sigma_R^2 = 2.25 k^{7/6} \int_0^H C_n^2(z) z^{5/6} dz, \quad (6.10)$$

where k and H again correspond to the laser wavenumber and propagation distance, and the receiver is located at $z = 0$ (ground level), corresponding to a zenith ($\theta_z = 0$) downlink scenario. To account for the saturation of scintillation in strong turbulence conditions, i.e., the decoupling of C_n^2 and σ_I^2 in strong turbulence, we relate σ_R^2 to σ_I^2 as (Andrews and Phillips 2005)

$$\sigma_I^2 = \exp \left(\frac{0.49 \sigma_R^2}{\left[1 + 1.11 (\sigma_R^2)^{6/5} \right]^{7/6}} + \frac{0.51 \sigma_R^2}{\left[1 + 0.69 (\sigma_R^2)^{6/5} \right]^{5/6}} \right) - 1. \quad (6.11)$$

This expression also includes the case of weak turbulence ($\sigma_R^2 < 1$) where $\sigma_I^2 \approx \sigma_R^2$.

Comparing the integrands of r_0 and σ_R^2 reveals that different parts of the $C_n^2(z)$ profile are weighted differently in the integration. Compared to r_0 , the expression for σ_R^2 contains an additional $z^{5/6}$ term in the integrand, weighting high-altitude

turbulence more strongly, compared to the uniform weighing of r_0 . As C_n^2 is typically strongest closest to the surface, r_0 is primarily sensitive to near-surface turbulence, whereas σ_I^2 is more sensitive to upper-air turbulence. These different weightings allow us to assess how well different parts of the $C_n^2(z)$ profiles are captured with a practical interpretation attached.

PERFORMANCE METRICS

Four metrics are used to quantify the performance of the different models in estimating high-res C_n^2 profiles, r_0 , and σ_I^2 : bias, centered root-mean-square error (cRMSE), Pearson correlation coefficient (r), and coefficient of determination (R^2).

The bias between an estimated profile variable $\hat{y} \in \mathbb{R}^n$ and the true profile variable $y \in \mathbb{R}^n$ is computed as

$$\text{Bias} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i), \quad (6.12)$$

where y_i and \hat{y}_i are the values of the true and estimated profiles at level i , respectively. The cRMSE removes the bias from the error metric and is computed as

$$\text{cRMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n ((\hat{y}_i - \bar{\hat{y}}) - (y_i - \bar{y}))^2}, \quad (6.13)$$

where \bar{y} and $\bar{\hat{y}}$ are the mean values of the true and estimated profiles, respectively. By separating bias and cRMSE, static biases that would otherwise inflate the RMSE are removed.

Finally, the Pearson correlation coefficient r is given by

$$r = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6.14)$$

and the coefficient of determination R^2 as

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}. \quad (6.15)$$

These per-profile scores are averaged over all profiles in the test dataset to obtain overall performance metrics for each model. As the integrated parameters r_0 and σ_I^2 are scalars per profile, bias, cRMSE, and r are computed directly from the predicted and true values across all profiles in the test dataset.

STRUCTURE FUNCTION ANALYSIS

The structure function (SF) analysis is a method for assessing how well models capture the vertical variability of C_n^2 profiles across different vertical scales (Lovejoy and Schertzer 2012). The structure function of a signal is related to its power spectrum obtained through the Fourier transform (Frisch 1995; Lovejoy and Schertzer 2012), so one can loosely think of SF analysis as a spectral analysis in physical

space instead of frequency space. The advantage of SFs over power spectra is that they can be computed more easily from non-uniformly spaced data, as is typical in atmospheric profiles. Additionally, interpreting SFs in physical space is often more intuitive than interpreting power spectra in frequency space. It should be emphasized that SFs are employed in this study solely as a diagnostic tool to assess vertical variability across different scales. We do not aim to link the SFs to turbulence theory, where certain scaling laws are expected, or additional interpretation constraints apply.

The 2nd order structure function of an ensemble of vertical profiles $y(z)$ is given as

$$S_2(\Delta z) = \langle (y(z + \Delta z) - y(z))^2 \rangle, \quad (6.16)$$

where $S_2(\Delta z)$ is the average squared increment between two points in the profile separated by a vertical distance Δz . Often, SFs follow power laws of the form $S_2(\Delta z) = c\Delta z^\zeta$ for parts of the Δz range, where ζ is called the scaling exponent. Comparing ζ and c of the SFs of the predicted and reference profiles allows us to assess, e.g., if the profiles differ in smoothness or if vertical variability (i.e., features) at certain scales is missing. Therefore, SF analysis is a powerful diagnostic tool for assessing the quality of predictions across different scales.

Practically, we compute, e.g., the vertical SF of $C_n^2(z)$ from a large ensemble of discrete profiles, where $[C_n^2]_{i,j}$ is the value of the i -th profile at the j -th level. As meteorological datasets often use levels of constant pressure rather than constant height z in the vertical direction, the z spacing between levels is typically non-uniform and time-dependent. Therefore, each $[C_n^2]_{i,j}$ value has an associated height above ground $z_{i,j}$. The SF is then computed in an ensemble manner by first computing the squared increments $\Delta C_n^2(\Delta z) = \left([C_n^2]_{i,j} - [C_n^2]_{k,l} \right)^2$ and corresponding vertical distances $\Delta z = |z_{i,j} - z_{k,l}|$ for all pairs of profiles and levels. The SF is then obtained by binning the ΔC_n^2 values according to their corresponding Δz values and averaging the ΔC_n^2 values within each Δz bin.

With this methodological foundation in place, the following section describes the datasets used for training and inference, including details on the WRF simulations that provide the high-res training targets and the ERA5 reanalysis that provides the low-res input data.

6.3. DATASETS

The two main datasets used in this work are a high-resolution (2 km \times 2 km) mesoscale dataset generated using the Weather Research and Forecasting (WRF) model (Skamarock *et al.* 2021) and the lower-resolution (0.25° \times 0.25° corresponding to ca. 17 km \times 17 km at 52°N in the Netherlands) ERA5 reanalysis dataset (Hersbach *et al.* 2020). The WRF data not only have higher horizontal resolution but also higher vertical resolution, with 100 vertical levels between the surface and ca. 20 km height, compared to 37 pressure levels in ERA5 up to ca. 30 km height. The extent of the domains in horizontal and vertical direction is illustrated in [fig. 6.3](#). Both datasets cover the same geographic region, here, the Netherlands, and the same time periods.

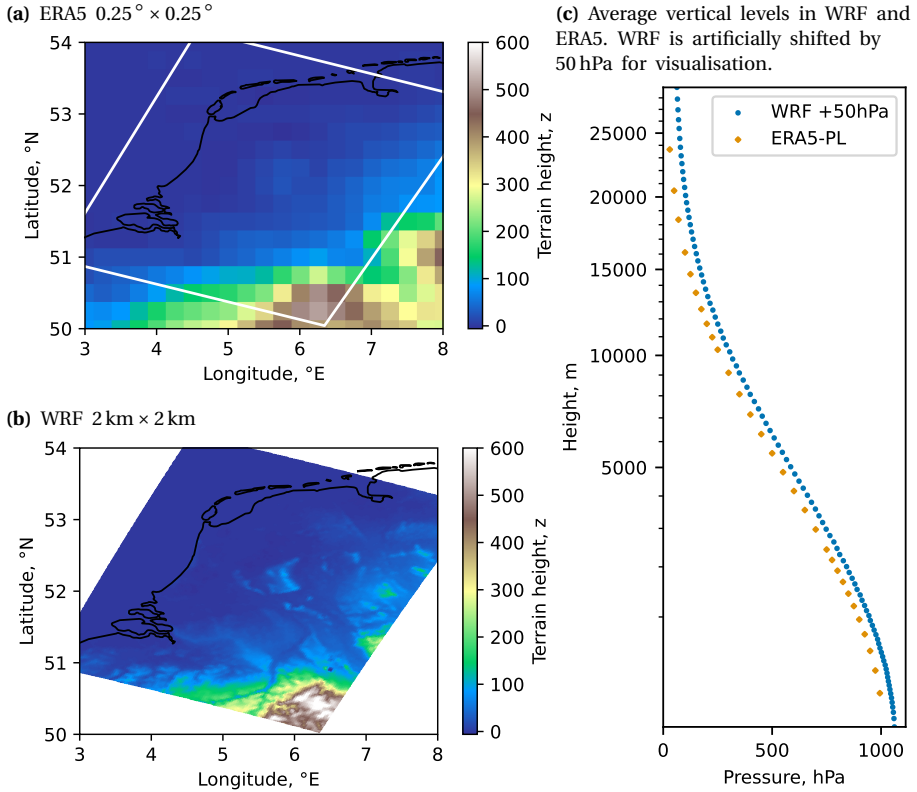


Figure 6.3.: Horizontal (a and b) and vertical (c) extent of ERA5 and WRF domains used in this work.

Details about the generation of the WRF dataset are presented in [section 6.3.1](#), while ERA5 is introduced in [section 6.3.2](#). [Section 6.3.3](#) summarizes the construction of the datasets used for training and inference of the Squeezeformer model.

6.3.1. WRF: YEAR-LONG C_n^2 DATABASE FOR NL

The Weather Research and Forecasting (WRF) model (Skamarock *et al.* 2021) is used to generate a statistically representative¹, high-resolution training dataset of C_n^2 over the Netherlands. The WRF configuration is based on Pierzyna *et al.* (2024) (cf. [chapters 2](#) and [3](#)) where C_n^2 is estimated following the variance-based parameterization of He and Basu (2015) (cf. [eq. \(2.7\)](#)).

The aim is to simulate a full year of hourly output at $2\text{ km} \times 2\text{ km}$ horizontal resolution and at 100 vertical levels, reaching up to ca. 20km in height. To keep this task manageable in terms of computational costs and storage requirements,

¹We refer to temporal representativeness for the Netherlands here. Geographical representative requires more extensive WRF simulations.

we employ two tricks. First, we use the Copernicus European Regional ReAnalysis (CERRA) (Schimanke *et al.* 2021) dataset, which has higher horizontal resolution than the global ERA5, to force WRF. Simulations forced with ERA5 typically require computing and storing 3 nested domains, whereas employing CERRA ($5\text{ km} \times 5\text{ km}$) enables us to reach the target resolution of $2\text{ km} \times 2\text{ km}$ with a single domain (Baki *et al.* 2025). We assume that the 2 km resolution is still sufficient to capture the relevant processes modulating optical turbulence.

The second trick is not to simulate a continuous year-long trajectory, but instead to run a series of shorter simulations staggered over multiple years to increase the statistical representativeness of meteorological conditions. In particular, we run 73 5.5-day simulations with 12 hours of warmup each, yielding 365 days net. The simulations are staggered in a round-robin fashion over four years (2017, 2018, 2019, 2020) to avoid sampling a particularly hot, cold, wet, or dry year. Consequently, simulations for the same year are always 15 days apart, with 18 simulations per year (19 for 2017 to cover the full 365 days). This short simulation length also limits model drift relative to the reanalysis forcing, as discussed in section 6.2.1.

6.3.2. ERA5 REANALYSIS

With C_n^2 profiles from WRF forming the high-res target data, we require collocated low-res meteorological input profiles and surface features. These are obtained from the ERA5 reanalysis dataset (Hersbach *et al.* 2020), which is available globally at a horizontal resolution of $0.25^\circ \times 0.25^\circ$. The vertical resolution of ERA5 depends on the variant used. For this work, we use the ERA5 pressure-level dataset (ERA5-PL), which contains 37 pressure levels between 1000 hPa and 1 hPa and is commonly used and readily available. Additionally, ERA5 provides a variety of surface variables at each grid point and has a temporal resolution of 1 hour, matching the WRF output frequency. Data are selected to cover the same geographic region and time periods as the WRF simulations. A list of the selected ERA5 variables is provided in table 6.1a.

6.3.3. CONSTRUCTION OF THE DEEP LEARNING DATASETS

The data from WRF and ERA5 are combined into a single dataset by collocating each ERA5 grid point with the closest WRF grid point. Due to the much higher horizontal resolution of WRF, this collocation results in a significant under-sampling of WRF data, retaining only ca. 1% of all WRF grid points. However, because the effective resolution of WRF is $\sim 7\Delta x$ (Skamarock *et al.* 2021), i.e., $\sim 14\text{ km}$ for this study, the under-sampling yields a more independent dataset with less redundant information. Vertically, we retain only the lower 30 pressure levels from ERA5, as our WRF simulations stop at 10 hPa (ca. 20 km height), whereas ERA5 extends up to 1 hPa (ca. 30 km height).

Features The feature and target variables utilized in this study are summarized in table 6.1. We consider $p = 7$ vertical profiles and $q = 14$ surface variables, which are known to modulate the primary regression target C_n^2 . Since (optical)

Table 6.1.: Overview of $p + q = 21$ features and $r = 4$ targets contained in the datasets. Variables marked with (*) are derived from other variables. All variables are normalized to the range $[0, 1]$ using the minimum and maximum values specified in the normalization column, where the prefix P indicates percentiles.

(a) Low-resolution meteorological input features from ERA5-PL or WRF-PL

Variable	Description	Unit	Type	Normalization
$u(z)$	Zonal wind speed	m s^{-1}	Profile	[P1, P99]
$v(z)$	Meridional wind speed	m s^{-1}	Profile	[P1, P99]
$\theta(z)$	Potential temperature	K	Profile	[min, max]
$p(z)$	Pressure	hPa	Profile	[min, max]
z	Height above ground	m	Profile	[min, max]
S	Wind shear, cf. eq. (6.17)	s^{-1}	Profile (*)	[0, P99]
Γ	Potential temperature gradient, cf. eq. (6.18)	Km^{-1}	Profile (*)	[P1, P99]
u_*	Friction velocity	m s^{-1}	Surface	[0, P99]
Q_H	Surface sensible heat flux	W m^{-2}	Surface	[P1, P99]
Q_L	Surface latent heat flux	W m^{-2}	Surface	[P1, P99]
$[\log_{10} C_T^2]_0$	C_T^2 at surface (Wyngaard <i>et al.</i> 1971)	-	Surface (*)	[-8, P99.5]
h	Boundary layer height	m	Single Level	[0, P99]
u_{10}	Zonal wind speed, 10 m above ground	m s^{-1}	Single Level	[P1, P99]
v_{10}	Meridional wind speed, 10 m above ground	m s^{-1}	Single Level	[P1, P99]
T_2	Air temperature, 2 m above ground	K	Single Level	[P1, P99]
p_0	Mean sea level pressure	hPa	Single Level	[min, max]
LSM	Land-sea mask	-	Single Level	-
$\cos(\text{hr}')$	Cosine of normalized hour of the day	-	Single Level (*)	[min, max]
$\sin(\text{hr}')$	Sine of normalized hour of the day	-	Single Level (*)	[min, max]
$\cos(\text{doy}')$	Cosine of normalized day of the year	-	Single Level (*)	[min, max]
$\sin(\text{doy}')$	Sine of normalized day of the year	-	Single Level (*)	[min, max]

(b) High-resolution turbulence target variables from WRF

Variable	Description	Unit	Type	Normalization
$\log_{10} [C_n^2(z)]$	C_n^2 profile, acc. to eq. (2.7)	-	Profile (*)	[P0.5, P99.5]
$\log_{10} [2e(z)]$	Twice the turbulent kinetic energy from WRF	-	Profile	[P0.5, P99.5]
$\log_{10} [L_m(z)]$	Master length scale from WRF	-	Profile	[P0.5, P99.5]
$\log_{10} [\sigma_\theta^2(z)]$	Potential temperature variance from WRF	-	Profile	[P0.5, P99.5]

atmospheric turbulence is modulated by wind shear and buoyancy (Stull 1988), the vertical profiles of wind components u and v as well as potential temperature θ are included. As not the absolute values of u , v , and θ but their vertical gradients drive turbulence, we additionally compute the mean wind shear S and the potential temperature gradient Γ between adjacent levels as

$$S = \left((\Delta u / \Delta z)^2 + (\Delta v / \Delta z)^2 \right)^{1/2} \quad (6.17)$$

and

$$\Gamma = \Delta \theta / \Delta z. \quad (6.18)$$

These engineered features are marked with (*) in table 6.1a.

As for the profiles, we select surface variables related to wind shear and buoyancy. These variables include the friction velocity u_* , the sensible and latent heat fluxes

Q_H and Q_L , as well as the W71-based estimate of surface layer C_T^2 (cf. eq. (2.3)). To allow the model to learn diurnal and seasonal patterns, temporal features based on the hour of the day (hr) and the day of the year (doy) are included. Due to their periodic nature, hr and doy are encoded using their sine and cosine components. Both features are also normalized to a 2π period as $hr' = (2\pi \text{hr})/24$ and $doy' = (2\pi \text{doy})/365$ before computing the sine and cosine.

Target variables The $r = 4$ target variables listed in table 6.1b are the \log_{10} -transformed C_n^2 from WRF at the 100 vertical levels and the physical variables used to parameterize C_n^2 (cf. eq. (2.7)): turbulent kinetic energy (TKE) e , potential temperature variance σ_θ^2 , and master length scale L_m . We incorporate these additional targets as auxiliary variables to support the model training, but the main focus of this study is on predicting C_n^2 .

Normalization All features and targets are normalized to the range $[0,1]$ before training. The method for determining the normalization bounds differs between features and targets as indicated in table 6.1. For well-behaved variables with clear minimum and maximum values, such as height above ground or pressure, the absolute minimum and maximum values across all vertical columns (time and location) are used. For variables with occasional outliers, such as u and v , the 1st and 99th percentiles are used instead to avoid placing too much emphasis on outliers. Note that the normalization does not involve clipping, so values outside the normalization bounds still contribute to the training. If variables such as u_* or S have physically meaningful lower bounds at zero, these are enforced during normalization.

Dataset summary Following the decoupled training approach described in section 6.2.1, three datasets are utilized in this work. First, the high-res WRF dataset with 100 vertical levels serves as the training target. Second, the WRF-PL dataset, a coarsened version of the WRF data with 30 pressure levels, is used as the low-res input during training to emulate ERA5. Third, the ERA5-PL dataset, also on 30 pressure levels, is quantile-mapped to match the WRF-PL distributions and is used as the low-res input for inference.

The following section presents results from experiments that systematically evaluate the Squeezeformer across these data configurations, isolating the contributions of regression, super-resolution, and distributional shift to overall prediction performance.

6.4. RESULTS

This section presents the results of experiments conducted to assess the performance of the Squeezeformer model in estimating high-resolution C_n^2 profiles from low-resolution meteorological data. The different experiments are summarized in table 6.2, where the experiments are set up to increase in complexity step by step to illuminate different aspects of the model's performance. The simplest experiment,

Table 6.2.: Overview of experiments differing by performed tasks and data utilized for training and inference. $\mathbf{X} \in \mathbb{R}^{p \times m}$ and $\mathbf{Y} \in \mathbb{R}^{r \times n}$ represent p input profiles and r output profiles, respectively, at $m < n$ levels.

Experiment	Regression $f(\mathbf{X}) \approx \mathbf{Y}$	Superresolution $m \rightarrow n$	Training data $\mathbf{X} \rightarrow \mathbf{Y}$	Inference data $f(\mathbf{X}) = \hat{\mathbf{Y}}$
▲ (Wn)	WRF native	✓	N/A	WRF high-res
● (Wpl)	WRF-PL	✓	30 → 100	WRF high-res
● (E5pl QM)	ERA5-PL inference, quantile-mapped (QM'd)	✓	30 → 100	WRF-PL
▼ (HV+W71)	Hufnagel-Valley with Wyngaard <i>et al.</i> (1971)	✓	30 → 100	same as (Wpl)
			W from WRF-PL	ERA5-PL QM'd
			z from WRF high-res	

Wn , uses the native high-res WRF data for input and as targets (cf. sections 6.3.1 and 6.3.3 and fig. 6.2a). As the input variables are still only those listed in table 6.1a, the Wn model performs regression toward C_n^2 but does not perform super-resolution or account for spatio-temporal shifts. This model also serves as an upper-performance baseline for the subsequent analysis. The second experiment, Wpl , employs the WRF-PL data as input and the high-res WRF data as target, thus performing both regression and super-resolution. This experiment is key to our decoupled training approach, in which the WRF-PL dataset aims to emulate ERA5-PL without spatiotemporal misalignment. The trained Wpl model is then used in the third experiment, $E5pl$ QM, where predictions are made using quantile-mapped ERA5 data to simulate the operational situation in which WRF-PL data are no longer available. In supplement 6.A, we demonstrate that statistically aligning ERA5-PL to the WRF-PL training data using QL improves prediction accuracy over using uncorrected ERA5-PL data. That section also compares our proposed decoupled training approach against directly training a Squeezeformer on ERA5-PL inputs and WRF high-res targets, showing that the decoupled approach yields predictions that more accurately and realistically reflect the vertical structure of C_n^2 profiles. The final experiment in table 6.2 is the lower baseline, in which the HV+W71 model is evaluated on WRF-PL data.

The Squeezeformer models in all experiments are configured and evaluated identically. The $p + q = 21$ input features (cf. table 6.1a) are expanded to $d = 48$ embedding dimensions in the first layer before being processed by 4 Squeezeformer blocks. To facilitate the extraction of small-scale features from the input profiles, a convolutional kernel size of 3 is used in the convolutional layers of the Squeezeformer blocks. The transformer blocks contain 2 attention heads each, and the regression space has size $h = 512$. The prediction and confidence heads map this high-dimensional space, interpolated from 30 to 100 levels, back to the $q = 4$ target variables. The resulting model has 1 million trainable parameters. The temporal staggering of the WRF simulations enables splitting the DL dataset into training, validation, and testing sets by selecting different years. For training, data from 2017 and 2018 are used, while 2019 is reserved for validation and 2020 for testing. All datasets cover all seasons to ensure statistical representativeness, but being from different years avoids data leakage.

The results presented below are based on the independent 2020 test set. The

analysis begins with a visual comparison of predicted and reference $\log_{10} C_n^2$ profiles, followed by a statistical quantification of agreement in [section 6.4.1](#). Through a structure function (SF) analysis, we zoom in on the ability of different models to capture the vertical structure of C_n^2 . The practical impact of the performance differences between models is discussed in the context of the integrated astroclimate parameters r_0 and σ_I^2 in [section 6.4.2](#).

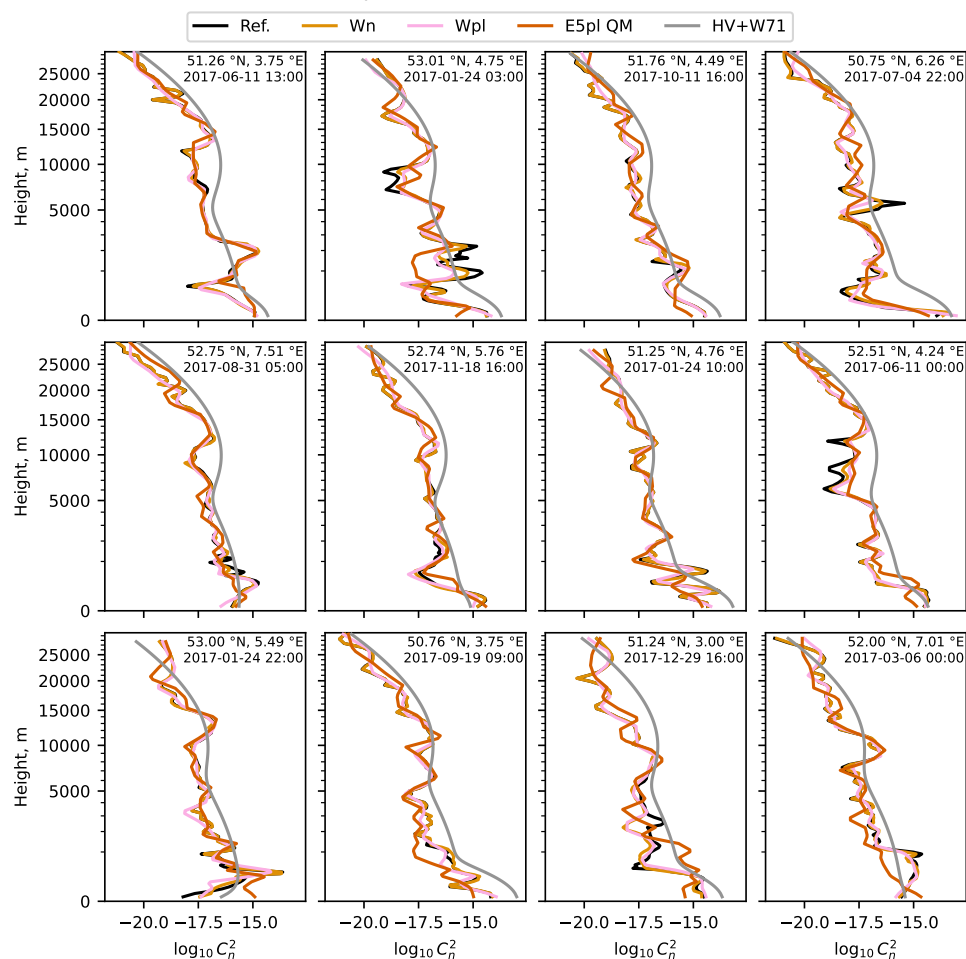
6.4.1. ESTIMATED $\log_{10} C_n^2$ PROFILES

A visual impression of the performance of the different models is given in [fig. 6.4a](#) with scores for the full test set displayed in [fig. 6.4b](#). The panels in [fig. 6.4a](#) display $\log_{10} C_n^2$ profiles at randomly selected time instances and locations, where the reference high-res WRF profiles are shown in black and the predictions of the four experiments with *Wn* in orange, *Wpl* in pink, *E5pl QM* in red, and *HV+W71* in grey. The upper baseline *Wn* profiles closely resemble the reference profiles and are often indistinguishable from the reference profiles. Large-scale trends across the atmospheric boundary layer (ABL) and local small-scale features confined to a few 100 m are well captured, with only a few underestimations of sharp edges relative to the ground truth. The performance scores reflect this visual observation with a bias close to zero, small cRMSE, and very high r and R^2 values. The *HV+W71* profiles representing the lower baseline match the general trend surprisingly ($r = 0.883$), although they clearly miss small-scale features present in all other profiles (low R^2) and consistently overestimate the reference ($\log_{10} C_n^2$ bias of 0.587). In conclusion, the *HV+W71* and *Wn* predictions serve as effective and simple upper and lower baselines before super-resolution and spatio-temporal misalignment are introduced as additional challenges in the *Wpl* and *E5pl QM* experiments.

Panels (a) and (b) of [fig. 6.4](#) show that the *Wpl* predictions are visually and statistically situated between the upper baseline *Wn* and the lower baseline *HV+W71*. *Wpl* exhibits greater smoothing than *Wn*, yet still captures the reference profiles well, given the significant super-resolution task from 30 to 100 levels. Compared to *Wn*, *Wpl* shows similar bias and correlation but higher cRMSE and lower R^2 due to smoothing and occasionally misplaced features. The *E5pl QM* profiles visually show larger discrepancies to the WRF-based profiles. This discrepancy is expected as the model did not learn to compensate for the spatio-temporal shift. Therefore, we limit ourselves to a general visual assessment here, with more quantitative statistical assessments presented later. Visually, *E5pl QM* does not seem to differ in smoothness compared to the *Wpl* predictions but mostly shows misplaced features and too low variance, i.e., too little overall variability. A small positive bias and reduced correlation metrics confirm this impression. Compared to *HV+W71*, however, *E5pl QM* still shows clear improvements in all scores despite the spatio-temporal misalignment.

Statistical characteristics [Figure 6.5](#) provides an overview of the statistical characteristics of the model predictions compared to the WRF-based reference profiles. Panel (a) compares the distributions of $\log_{10} C_n^2$ in two ways, with probability density functions (PDFs) on the left and quantile-quantile (qq) plots on the right.

(a) Randomly drawn examples of $\log_{10} C_n^2$ profiles.



(b) Global performance scores of $\log_{10} C_n^2$, where arrows indicate if lower scores (↓) or higher scores (↑) are better. The table on the right presents numerical values corresponding to the scatter plot.

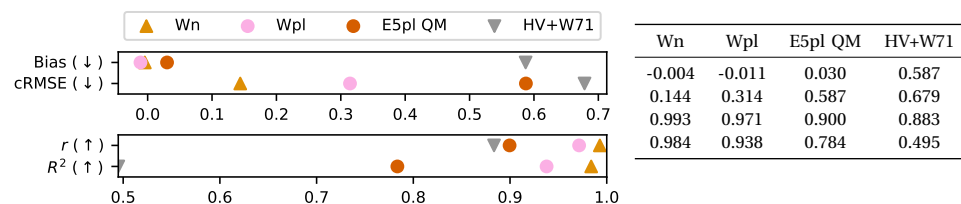
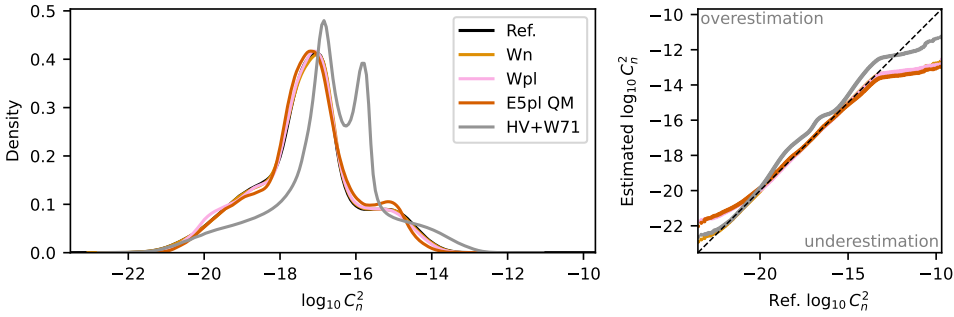


Figure 6.4.: Visual and statistical comparison of predicted $\log_{10} C_n^2$ profiles against WRF-based reference profiles. Scores are aggregated over all levels, locations, and time.

(a) Probability density functions and quantile-quantile plots of $\log_{10} C_n^2$ predictions against WRF-based reference profiles.



(b) Vertical second-order structure functions of $\log_{10} C_n^2$ profiles.

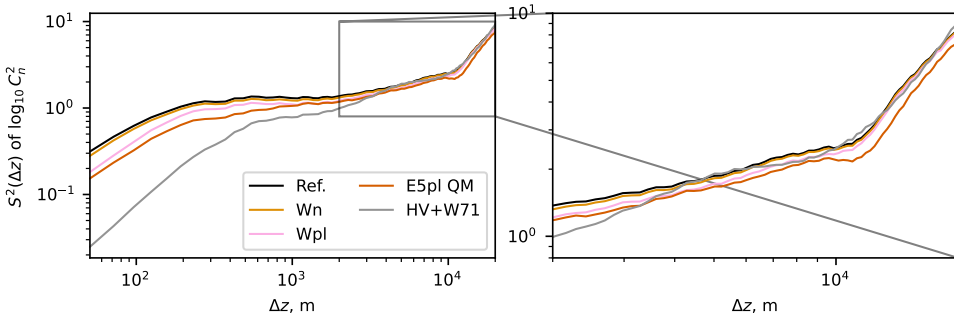


Figure 6.5.: Statistical characteristics of predicted $\log_{10} C_n^2$ profiles compared to WRF-based reference profiles. All curves are based on data aggregated over all levels, locations, and time.

The qq plots compare the empirical density functions of the predictions and the reference profiles, where a perfect match would result in a straight 1:1 line. Panel (b) shows the vertical second-order SFs of $\log_{10} C_n^2$ profiles. Compared to the previous discussion, the global statistical characteristics are not affected by spatio-temporal misalignment, as no profile-by-profile comparison is made. Instead, the statistical characteristics reflect the ability of different models to produce profiles with correct vertical structure and correct magnitude.

The PDF of *Wn* matches the reference very well, consistent with the high scores reported earlier. However, deviations become visible in the qq-plot in the tails of the distribution. Very low values ($10^{-22} \text{ m}^{-2/3}$) are overestimated, likely due to regularization during training that prevents overfitting to the few potential outliers present in this range. A pronounced kink is visible at high values ($C_n^2 \approx 10^{-14} \text{ m}^{-2/3}$), indicating underestimation of the high tail. The underestimation of the high tail is practically more critical than the low-tail overestimation, as high C_n^2 values are most relevant for FSOC applications due to the possibility of causing link degradation or outages. Nevertheless, as only 0.5% of conditions are affected and the bulk of the distribution is represented well, we consider the overall performance of *Wn* good.

Wpl and $E5pl$ QM track the reference PDFs well in the center but show larger mismatches than Wn at both tails. Both experiments already overestimate the lower tail at $C_n^2 < 10^{-20} \text{ m}^{-2/3}$ compared to $C_n^2 < 10^{-22} \text{ m}^{-2/3}$ for Wn . For the high tail, the underestimation already observed for Wn is more pronounced for $E5pl$ QM, with its qq curve deviating from the 1:1 line earlier than Wn or Wpl . The similarity between the two PL experiments supports the conclusion that quantile-mapping ERA5-PL to WRF-PL is effective. In supplement 6.A, we also show that using uncorrected ERA5-PL data leads to stronger overestimation of the lower tail. These results also validate the decoupled training approach, demonstrating that statistically consistent estimates can be obtained even when using the model with ERA5-PL data. The HV+W71 model, by contrast, does not need much detailed consideration. The simplicity of its analytical formulation does not allow it to capture the statistical characteristics well, as evident from the shifted bimodal distribution and mismatched qq-plot.

Representation of vertical structure The capability of the different experiments to capture the vertical structure of C_n^2 is assessed by the SF in fig. 6.5b. Assessing how well vertical structure is captured is important because smoothing of small-scale features leads to overestimation (underestimation) of r_0 (σ_I^2), as will be discussed in section 6.4.2. The Wn curve follows the WRF-based reference SF closely over the entire range of scales with only a small widening gap toward smaller scales. This gap reflects the occasional smoothing of local/small-scale features observed in fig. 6.4a, but overall vertical structure is reproduced very well. The HV+W71 model, on the other hand, shows clear deficits in capturing vertical structure, as was already evident from visual inspection. The HV+W71 SF is close to the reference at very large scales, matching the visual impression that HV+W71 reflects the overall trend of C_n^2 . However, the HV+W71 SF quickly and strongly departs from the WRF-based reference toward smaller scales, indicating oversmoothing. Again, this observation is expected based on the simple analytical nature of the HV+W71 model.

As before, Wpl and $E5pl$ QM fall between the two baselines Wn and HV+W71. However, the SF reveals a clearer difference between Wpl and $E5pl$ QM than was apparent in the distributions. At large scales (zoom of panel (b)), $E5pl$ QM shows a constant offset relative to all other curves, indicating that its C_n^2 profiles exhibit less overall vertical variability. This offset likely stems from a remaining mismatch between the WRF-PL and ERA5-PL input distributions. Supplement 6.A confirms this interpretation by showing that uncorrected ERA5-PL data produce an even stronger offset. Wpl closely follows the reference at large scales but shows a widening gap relative to both the reference and Wn when moving toward smaller scales, reflecting the smoothing of small-scale features already observed in the profile comparisons. $E5pl$ QM exhibits this same scale-dependent smoothing in addition to the large-scale offset. In summary, all Squeezeformer-based experiments remain substantially closer to the WRF reference than HV+W71, indicating good overall performance. The SF analysis confirms that while Wpl and $E5pl$ QM smooth smaller-scale features compared to Wn , they still capture the overall vertical structure well and demonstrate the model's capability to estimate realistic high-res C_n^2 profiles from ERA5-PL data.

The remaining differences between *E5pl QM* and *Wpl* are likely due to inherent differences in the dataset stemming from different physics parameterizations and resolution, which cannot be fully resolved by quantile mapping.

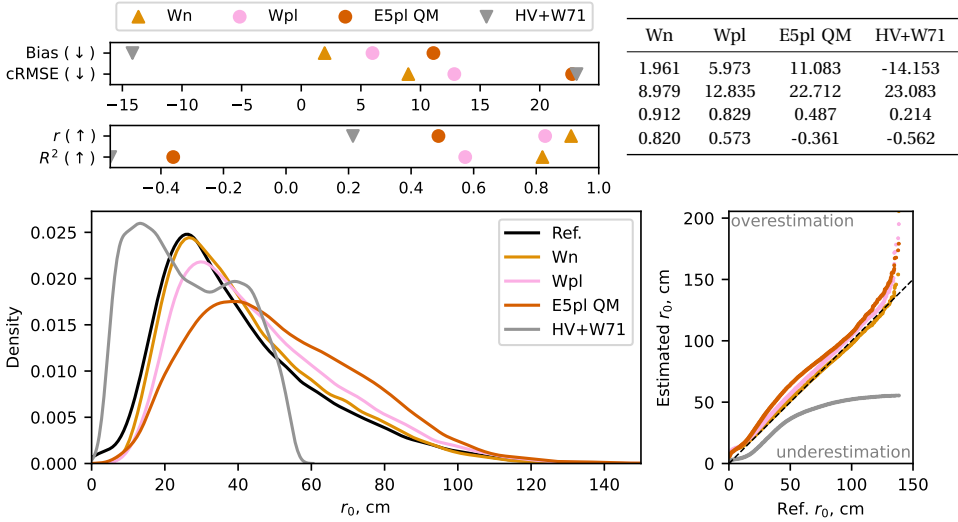
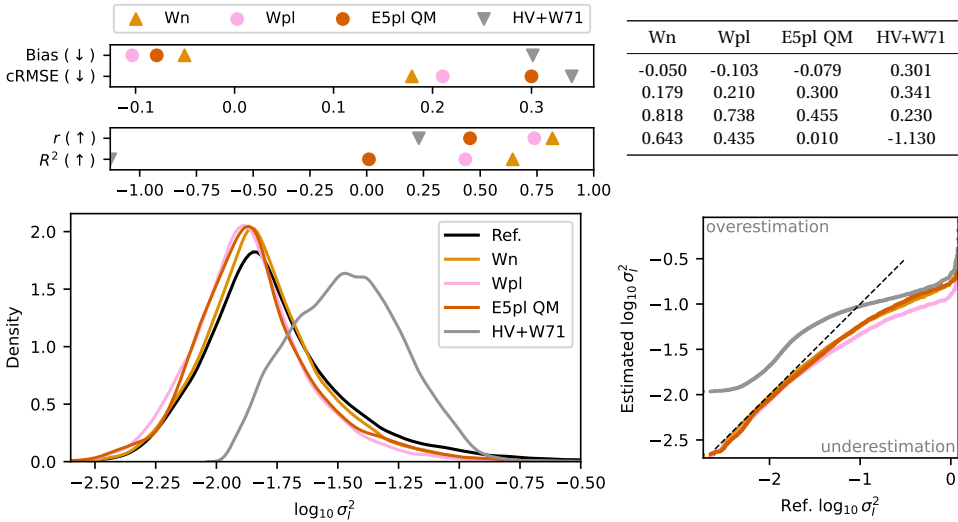
6.4.2. ESTIMATES OF r_0 AND σ_I^2

Following the detailed assessment of the predicted C_n^2 profiles, we now discuss the impact of the observed smoothing and distributional deviations on the integrated astroclimate parameters r_0 and σ_I^2 , both evaluated at zenith ($\theta_z = 0$). In analogy to the profile analysis, performance scores and distributions for r_0 and σ_I^2 computed from the predicted C_n^2 profiles are presented in figs. 6.6a and 6.6b, respectively. Before proceeding, we emphasize that the Fried parameter r_0 and C_n^2 are inversely related, so a low r_0 corresponds to high turbulence/ C_n^2 and vice versa, whereas the scintillation index σ_I^2 increases with increasing turbulence and C_n^2 .

Fried Parameter Considering r_0 in fig. 6.6a, the *Wn* experiment again performs best with high scores compared to the reference. The distributions also match well for most of the range, with only small overestimations visible for both small ($r_0 < 10$ cm) and large ($r_0 > 130$ cm) values. Interestingly, no underestimation of r_0 is observed in the qq-plot, which can be explained by linking the r_0 analysis back to the earlier C_n^2 profile assessment. When C_n^2 is overestimated, turbulence is too high, so r_0 should be too low, i.e., underestimated. However, the C_n^2 overestimations discussed previously occurred for very small values in log-space, whereas r_0 is based on the integral of C_n^2 in linear space. Consequently, the overestimation of low C_n^2 values has a negligible effect on the integrated r_0 , explaining why no underestimation is visible in the qq-plot. The opposite case reveals the impact more clearly: when high C_n^2 values are underestimated, the error in linear space becomes very pronounced, leading to overestimated r_0 . This overestimation is attributed to the smoothing of previously observed profiles, both in the structure function analysis and visually in fig. 6.4a.

The behavior of *Wpl* and *E5pl QM* is similar to *Wn* in that they also show skewed single-mode distributions with long tails toward high r_0 . However, the distributions of *Wpl* and *E5pl QM* are progressively more skewed to higher r_0 values, indicating increasing overestimation compared to the reference. This trend is also reflected in the increasing bias and cRMSE, with biases of approximately 6 cm for *Wpl* and approximately 11 cm for *E5pl QM*. While the PDF of *Wpl* still follows the reference relatively closely, the difference for *E5pl QM* is quite pronounced. Recalling the earlier structure-function analysis, we attribute this behavior to the observed offset in the SE, which reflected a reduced overall variance in the C_n^2 profiles. This reduced variance now manifests as overly optimistic (i.e., too high) r_0 estimates for *E5pl QM*, a bias that should be addressed in future work.

Scintillation Index The performance characteristics of σ_I^2 in fig. 6.6b differ from those of r_0 due to the height-dependent weighting in the integrand compared to the uniform weighting for r_0 . Since turbulence strength and C_n^2 generally decrease

(a) Fried parameter r_0

 (b) Logarithm of scintillation index $\log_{10} \sigma_I^2$

Figure 6.6.: Performance scores, histograms, and quantile-quantile plots for the integrated FSOC parameters r_0 and σ_I^2 computed from the predicted C_n^2 profiles at zenith ($\theta_z = 0$, $\lambda = 1550$ nm).

with height, the effect of models underestimating high C_n^2 values becomes less pronounced for σ_I^2 . This is visible in the lower spread of the PDFs compared to r_0 with all distributions except HV+W71 being close to the reference. Consistent

with previous discussions, the PDFs remain skewed toward lower turbulence (lower σ_I^2 and higher r_0) due to the observed smoothing in the C_n^2 profiles. The overall agreement with the reference σ_I^2 is better than for r_0 for all experiments, as is evident from the higher r and R^2 values. However, the qq-plots also show strong underestimation for higher scintillation indices. We assume that this underestimation corresponds to rare high-altitude, high C_n^2 events (e.g., associated with jet streams (Basu *et al.* 2020)), which strongly affect σ_I^2 due to its emphasis on upper-air. Surprisingly, even Wn suffers from this issue, suggesting that the model generally struggles to reconstruct these rare high-turbulence events. The reason could be insufficient representation in the training data or architecture limitations, which are left for future work to investigate.

FSOC Implications The conclusion from the r_0 and σ_I^2 analysis is that results are much better than the HV+W71 baseline, but both parameters are affected by the turbulence underestimation due to smoothing. The practical implications of this underestimation for FSOC are discussed below. Considering a communication terminal for space-to-ground communication with a telescope aperture diameter D , the turbulence-induced degradation of the optical wavefront is determined by the ratio D/r_0 (Andrews and Phillips 2005; Sasiela 2007). More specifically, degradations occur for $D/r_0 > 1$, so $r_0 < D$. Space-to-ground terminals typically have aperture diameters between 50 cm and 100 cm (e.g. Saathof *et al.* (2017) and Klop *et al.* (2021)), so degradations are expected for r_0 values below this range. According to the reference distribution in fig. 6.6a, r_0 values below 100 cm occur in approximately 98.5% of cases, and values below 50 cm occur in approximately 72.5% of cases, whereas the E5pl QM predictions suggest 97.1% and 52.8% of cases, respectively. This comparison suggests that the relevance of the r_0 underestimation statistically decreases with increasing aperture diameters, as large apertures tend to be affected by turbulence in most cases, regardless of the error. Nevertheless, this underestimation should be addressed in future work, as the crossover point at $D \approx r_0$ remains relevant for small apertures. The scintillation index σ_I^2 modulates link losses due to intensity fluctuations. Following Giggenbach (2008), the corresponding link loss is proportional to $(\sigma_I^2)^{2/5}$, implying that the sensitivity to errors in σ_I^2 decreases for higher scintillation indices. Consequently, the observed good match in statistics for low σ_I^2 , where the sensitivity of link loss is high, is a positive result. However, high σ_I^2 values are also relevant as communication outages can occur under these conditions. Since these extreme conditions are not captured well by any of the models, future work is needed to improve performance in this regime. In conclusion, OTProf's performance is promising across many FSOC operating conditions, but further improvements are needed to ensure reliable performance under extreme turbulence conditions.

6.5. CONCLUSION

This study proposed OTProf, a deep learning approach to estimate high-resolution vertical C_n^2 profiles (100 levels) from coarse-resolution reanalysis data (30 pressure

levels). A variant of the Squeezeformer architecture (Kim *et al.* 2022), adapted for atmospheric processes (Ron 2024), is modified to also enable vertical super-resolution. Training data were generated using a year-long mesoscale simulation over the Netherlands at $2\text{ km} \times 2\text{ km}$ horizontal resolution, with C_n^2 computed following the variance-based parameterization of He and Basu (2015). The simulations were performed using the Weather Research and Forecasting (WRF) model (Skamarock *et al.* 2021) and are staggered in time over 4 years to ensure temporal statistical representativeness. ERA5 reanalysis data (Hersbach *et al.* 2020) on pressure levels (ERA5-PL) was used as coarse-resolution input data. The selection of input features was motivated by physical principles to capture the processes that modulate atmospheric turbulence: wind shear and buoyancy. Consequently, 7 vertical variables, such as wind, temperature, and engineered gradients, were used, along with 14 surface variables related to wind shear and buoyancy. Training on datasets produced by different numerical models (e.g., ERA5-PL and WRF) can be problematic due to slight spatiotemporal misalignment and differences in model resolution and physics parameterizations. We propose to address this misalignment issue in two steps. First, we employ a decoupled training approach in which the Squeezeformer is trained on vertically coarsened WRF data (WRF-PL) as input, which emulates ERA5-PL. Second, ERA5-PL data is quantile-mapped to WRF-PL to reduce distributional differences and then used for inference with the trained model. In supplement 6.A, we show that this approach is effective compared to using uncorrected ERA5-PL data for inference or training directly on ERA5-PL data. The analytical Hufnagel-Valley model (Hufnagel 1974; Ulrich 1988; Valley 1980), commonly used in the optical turbulence community, served as a lower baseline throughout the study.

Three experiments of increasing complexity were conducted to isolate the contributions of regression, super-resolution, and distributional shift to the overall prediction error. The simplest experiment (Wn) uses native high-resolution WRF data for both input and target, the second (Wpl) introduces super-resolution from 30 to 100 vertical levels using coarsened WRF input, and the third (E5pl QM) applies the trained model to quantile-mapped ERA5 data to simulate operational conditions.

The results show that the Squeezeformer accurately reconstructs C_n^2 profiles in the Wn setting, confirming that the selected meteorological input variables carry sufficient information for this regression task. Introducing super-resolution (Wpl) leads to some smoothing of small-scale features, as quantified by the structure function analysis, but retains the overall vertical structure of the profiles. Applying the trained model to quantile-mapped ERA5 data (E5pl QM) introduces additional differences compared to the WRF-based experiments. In particular, the structure function analysis reveals a reduction in the captured vertical variability across all scales, which leads to a systematic overestimation of r_0 and an underestimation of σ_I^2 . This degradation is attributed to fundamental differences between the ERA5 and WRF datasets. ERA5 and WRF are produced by different numerical models with different physics parameterizations, different horizontal resolutions, and different representations of, e.g., orography. Quantile mapping can reduce the statistical mismatch in input distributions, but it cannot compensate for the structural differences between the two modelling systems. The remaining gap after

quantile mapping, therefore, reflects an inherent challenge of combining datasets from different sources.

The practical relevance for FSOC was assessed through the Fried parameter r_0 and scintillation index σ_I^2 , both at zenith ($\theta_z = 0$). Both parameters improve considerably over the HV+W71 baseline but are affected by the underestimation of turbulence due to smoothing. For space-to-ground terminals with aperture diameters D between 50 and 100 cm, wavefront degradation is determined by the ratio D/r_0 , with perturbations occurring when $r_0 < D$. A statistical comparison shows that the impact of the r_0 error decreases with increasing aperture size, as larger apertures are more often affected by turbulence, regardless of the error. Nevertheless, the crossover region at $D \approx r_0$ remains relevant for smaller apertures, so model errors in that regime should be addressed in future work. For σ_I^2 , scintillation-induced link loss scales with $(\sigma_I^2)^{2/5}$, so error sensitivity decreases at higher scintillation indices. The well-matched statistics at low σ_I^2 , where sensitivity is highest, are therefore encouraging. However, rare high σ_I^2 conditions are also relevant as they can cause communication outages. Since these extreme turbulence cases are not well captured by any model, further improvements are needed to ensure reliable performance under such conditions. Finally, we note that the r_0 and σ_I^2 results are given for zenith paths because OTProf estimates C_n^2 of vertical columns. However, realistic FSOC paths are slant paths, which can be evaluated by assembling a cross-section of columns along the line-of-sight and integrating the FSOC parameters along the slant path.

Several other aspects warrant further investigation. The present study is limited to the Netherlands, so the spatial generalizability to regions with more complex orography or different climates remains to be assessed. However, we do not see any barrier that would prevent the successful application of OTProf in other regions. The coarse vertical spacing of ERA5 pressure levels within the boundary layer, where turbulence is typically strongest, may limit the available information in this critical region. Using ERA5 model levels for training could help in this regard, but would also partially reduce practicability because pressure-level data are more readily available than model-level data. Finally, observational data from, e.g., scintillometers (C_n^2), SCIDARs (C_n^2 profiles), or DIMMs (r_0) could further improve the model estimates in the future. The possibilities range from simple site-specific bias corrections of the estimated r_0 using the observed r_0 to more complex fine-tuning of the Squeezeformer on observed C_n^2 profiles or on integrated parameters.

Despite these remaining challenges, OTProf represents a considerable step forward compared to the Hufnagel-Valley model, which remains widely used in the FSOC and optical turbulence communities. The HV model is constrained to a fixed exponential profile shape, which cannot capture the diversity of real atmospheric conditions, leading to a persistent positive bias and a poor representation of vertical variability at all but the largest scales. OTProf, in contrast, produces profiles that are more realistic in both shape and magnitude while requiring only modest data resources: a single year of regional WRF simulations for training and globally available ERA5 pressure level reanalysis data for inference. Therefore, we are confident that OTProf provides a promising pathway toward more accurate and realistic estimates of

vertical C_n^2 profiles at manageable computational cost.

ACKNOWLEDGMENTS

MP is funded by the FREE project (P19-13) of the TTW-Perspectief research program, partially financed by the Dutch Research Council (NWO). The WRF training dataset was generated using the Dutch national e-infrastructure with the support of the SURF Cooperative using grant no. EINF-15953. GPU resources for training the deep learning models were provided by University at Albany.

DATA AVAILABILITY

The WRF training dataset and the Squeezeformer code are available on Zenodo (Pierzyna 2026b). Additionally, the codes can be found on GitHub: <https://github.com/mpierzyna/otprof>. The raw WRF model output is available from the authors upon reasonable request.

6.A. INFLUENCE OF QUANTILE MAPPING ON DECOUPLED TRAINING

6

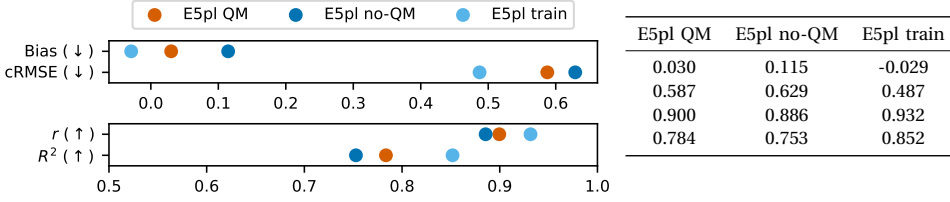
Table 6.3.: Overview of additional experiments to demonstrate effect of quantile mapping and decoupled training. Experiment E5pl QM is discussed in the main text and repeated here for reference.

Experiment		Regression $f(\mathbf{X}) \approx \mathbf{Y}$	Superresolution $m \rightarrow n$	Training data $\mathbf{X} \rightarrow \mathbf{Y}$	Inference data $f(\mathbf{X}) = \hat{\mathbf{Y}}$
● (E5pl QM)	ERA5-PL inference, quantile-mapped (QM'd)	✓	30 \rightarrow 100	same as (Wpl)	ERA5-PL QM'd
● (E5pl no-QM)	ERA5-PL inference	✓	30 \rightarrow 100	same as (Wpl)	ERA5-PL native
● (E5pl train)	ERA5-PL training	✓	30 \rightarrow 100	ERA5-PL \rightarrow WRF native	ERA5-PL native

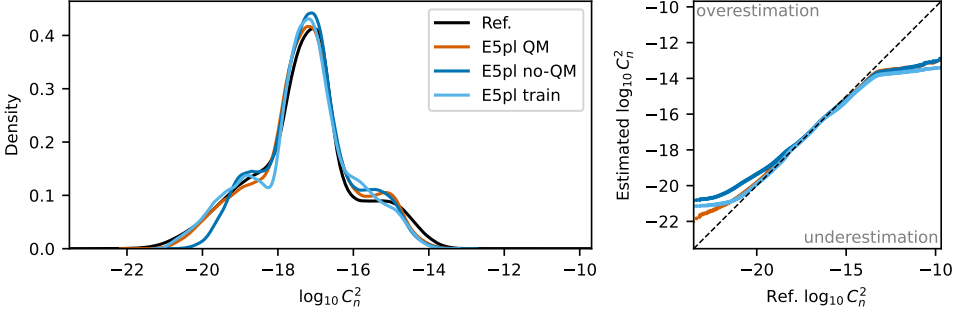
This supplement examines the effect of quantile mapping (QM) on the decoupled training approach by comparing three additional experiments summarized in table 6.3. The first experiment, *E5pl QM*, was already discussed in the main text and serves here as the reference point for comparison. The second experiment, *E5pl no-QM*, uses the same model trained on WRF-PL data (*Wpl*) but applies it to native, unmodified ERA5-PL data during inference to isolate the effect of QM. The third experiment, *E5pl train*, takes a different approach by directly training a Squeezeformer on native ERA5-PL inputs and WRF high-res targets, allowing to learn the mapping between the two datasets explicitly rather than relying on statistical alignment. Performance scores and statistical characteristics for all three experiments are presented in fig. 6.7, with example profiles shown in fig. 6.8.

Quantile Mapping versus no correction Comparing *E5pl QM* against *E5pl no-QM* reveals the benefit of statistically aligning ERA5 data to the WRF training distribution.

(a) Global performance scores, where arrows indicate if lower scores (↓) or higher scores (↑) are better. The table on the right presents numerical values corresponding to the scatter plot.



(b) Histograms and quantile-quantile plots of $\log_{10} C_n^2$ predictions against WRF-based reference profiles.



(c) Vertical second-order structure functions of $\log_{10} C_n^2$.

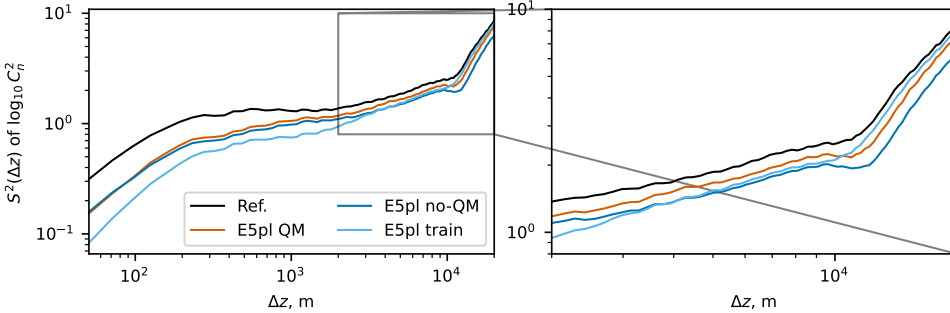


Figure 6.7.: Performance scores and statistical characteristics of predicted $\log_{10} C_n^2$ profiles compared to WRF-based reference profiles. Scores and curves are aggregated over all levels, locations, and time. This figure supplements fig. 6.5.

The performance scores in fig. 6.7a show that quantile mapping reduces both bias and cRMSE while improving the correlation coefficient r and the coefficient of determination R^2 . These improvements are also reflected in the distributions shown in fig. 6.7b. Here, the quantile-mapped predictions have tails that are closer to the reference compared to the uncorrected case. The structure function analysis reveals even more pronounced differences between the two approaches. The offset in the SF observed for *E5pl QM* in the main text becomes noticeably larger for *E5pl no-QM*, indicating that less overall variance is captured when ERA5 data are not statistically

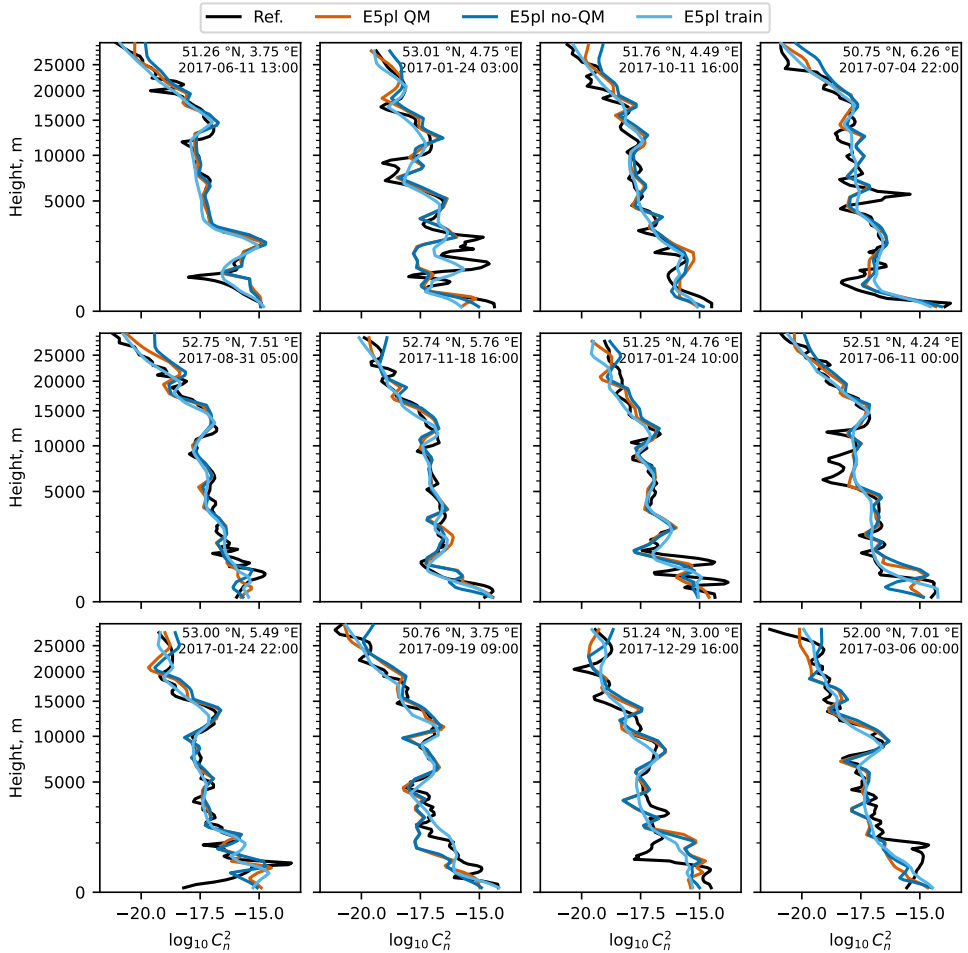


Figure 6.8.: Randomly drawn examples of $\log_{10} C_n^2$ profiles. This figure supplements fig. 6.4a.

aligned. Furthermore, the *QM* curve consistently lies above the *no-QM* curve across all scales. This demonstrates better representation of vertical variability at all scales when quantile mapping is applied.

Decoupled training versus direct training Rather than using quantile mapping to bridge the gap between WRF-PL and ERA5-PL, the *E5pl train* experiment explores whether directly training on ERA5-PL inputs and WRF high-res targets is beneficial. This approach allows the model to learn the direct mapping between the two datasets, potentially compensating for the dataset misalignment during training. The performance scores show that *E5pl train* achieves a bias with different sign but similar magnitude compared to *E5pl QM*, which is lower than that of *E5pl no-QM*.

Additionally, the cRMSE and correlation metrics of *E5pl train* are better than both *E5pl QM* and *E5pl no-QM*. This suggests that the model has indeed learned to partially compensate for the dataset misalignment through direct training.

However, the benefit of the decoupled training approach combined with quantile mapping becomes evident when examining the statistical characteristics. In the distributions shown in fig. 6.7b, *E5pl train* performs better than *E5pl no-QM* at the lower end of the distribution but does not match the performance of *E5pl QM*. Moreover, *E5pl train* underestimates the upper tail more strongly than either of the other two ERA5-based approaches. Considering the structure function analysis in panel (c), *E5pl train* shows a smaller gap to the reference at large scales compared to both *E5pl QM* and *E5pl no-QM*. This is expected since the direct mapping has been learned during training. However, the model fails to capture smaller-scale variability. For vertical separations $\Delta z < 3000$ m, *E5pl train* exhibits stronger smoothing with a pronounced widening gap relative to the reference. This indicates that the direct training approach fails to represent small scale features.

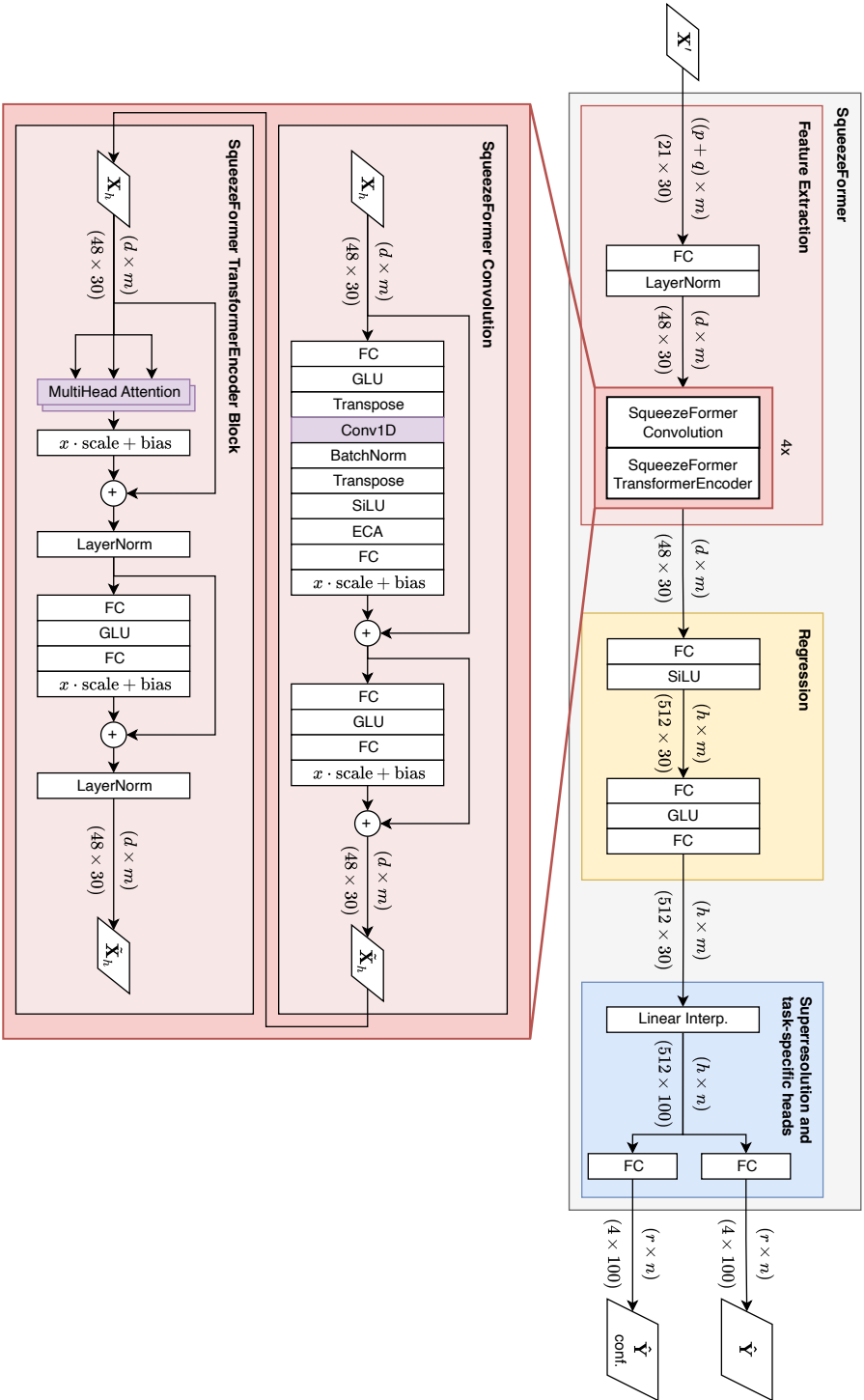
The comparison of these three approaches reveals that quantile mapping finds a beneficial middle ground. While QM cannot fully compensate for the differences between WRF and ERA5 datasets, it consistently improves predictions compared to using uncorrected ERA5 data. The remaining misalignment likely stems from different physical parameterizations and ERA5's coarser horizontal resolution capturing less terrain effects. Moreover, the decoupled training approach with quantile mapping yields better statistical characteristics than direct training. This is particularly true with respect to the representation of small-scale vertical variability. Based on these findings, we chose *E5pl QM* as the preferred approach for this work. It provides the best balance between overall performance and realistic representation of vertical variability across all scales.

6.B. ADDITIONAL DETAILS OF SQUEEZEFORMER ARCHITECTURE

Figure 6.9 provides a more detailed schematic of the Squeezeformer architecture compared to fig. 6.2b presented in the main text. Primarily, the figure illustrates the composition of the Squeezeformer convolution and transformer blocks. The convolution block takes the embedded hidden input (index h) and expands the number of channels by a factor of 4, followed by a depthwise convolution. Efficient channel attention (ECA, Wang *et al.* (2020)) is applied to capture cross-channel interactions followed by a projection back to the original number of channels. Learnable scaling and bias parameters are applied per channel before adding \times_h via a residual connection. Another block of GLU-activated embedding is applied, followed by the application of learnable scaling and bias parameters, before adding a second residual connection. The gated GLU activations throughout the architecture allow the network to learn to suppress irrelevant information and focus on the important features.

The output of the convolution block is fed into the transformer block. Two attention heads are used to capture interactions across the vertical dimension,

Figure 6.9.: Detailed schematic of the Squeezeformer architecture extending fig. 6.2b of the main text.



followed by another learnable channel scaling. The output is normalized after adding the residual connection and embedded through a GLU-activated fully connected network like in the convolution block. After adding a second residual connection, another LayerNorm is applied to yield the final output of the transformer block.

III

Conclusions and perspectives

7

Conclusions and perspectives

This dissertation has explored the modelling of optical turbulence, primarily targeting the coefficient of the second-order refractive index structure function, C_n^2 , along two parallel avenues: traditional numerical mesoscale modelling and state-of-the-art machine learning (ML) techniques. The advances made in both domains will be briefly summarised in [sections 7.1](#) and [7.2](#). The overarching picture emerging from this dissertation, however, is the potential and value of combining physics-based and data-driven modelling. Recommendations for future work and perspectives on free-space optical communication (FSOC), optical turbulence modelling, and modelling of boundary layer turbulence in general are discussed in [section 7.3](#).

7.1. TRADITIONAL MESOSCALE MODELLING

While the Weather Research and Forecasting (WRF) model (Skamarock *et al.* 2021) is well-established, and a multitude of C_n^2 parameterizations have been proposed in the literature, this dissertation has systematically assessed and extended the applicability of these techniques to FSOC.

Intercomparison of C_n^2 parameterizations The foundation was laid in [chapter 2](#) by systematically comparing three distinct types of C_n^2 parameterizations: flux-based, gradient-based, and variance-based. Each parameterization was applied to both observed and simulated meteorological variables, and the resulting C_n^2 estimates were compared against observed C_n^2 from two scintillometers. The aim was to disentangle the performance of the parameterizations from the capability of the WRF model to simulate the meteorology. The variance-based parameterization (He and Basu 2015) yields the overall best performance. Unlike other parameterizations, its application is not limited to the lowest part of the atmospheric boundary layer, i.e. the surface layer. We showed that C_n^2 estimated from WRF output aligns well with observations, highlighting the value of mesoscale models for optical turbulence modeling.

The variance-based C_n^2 parameterization requires high-order statistics and a turbulent length scale. The difficulty in accurately deriving this length scale makes the parameterization challenging to apply directly to observational data. For numerical modeling, this requirement also limits the choice of compatible planetary boundary layer

(PBL) schemes in WRF. However, this restriction has minimal practical impact, as suitable high-order schemes are readily available. For instance, the MYNN 2.5 scheme (Nakanishi and Niino 2006; Nakanishi and Niino 2009), utilized by He and Basu (2015) and Pierzyna *et al.* (2024), is already well-tested and operational (Olson *et al.* 2019), so combining it with the variance-based approach provides a robust framework for FSOC applications.

Ensemble estimates of C_n^2 Not only is the choice of C_n^2 parameterization critical for accurate C_n^2 estimates, but so is the choice of WRF physics parameterizations, as presented in appendix A. We demonstrated that a significant spread in C_n^2 estimates arises when rerunning the same case with different PBL schemes. While this spread underscores the need to select a WRF configuration carefully, it also provides an opportunity to treat the spread as an uncertainty estimate and the different configurations as ensemble members. We found that the ensemble statistically agrees better with scintillometer observations than individual members. Given the aforementioned limited choice of PBL schemes when using the variance-based C_n^2 approach, physics ensembles remain possible by varying other physics options, such as microphysics or radiation schemes. It is also worth noting that the computational effort scales linearly with the ensemble size, since each member requires a separate WRF run.

Estimation of FSOC link performance To keep computational effort manageable, we continued with the variance-based parameterization and a deterministic WRF run using a high-order turbulence scheme, and focused on making C_n^2 estimates more practically interpretable for FSOC applications. Numerical models are never perfect, so some degree of C_n^2 estimation error is unavoidable. However, FSOC system designers are arguably more interested in knowing the estimation errors with respect to link performance or expected communication data rates rather than in C_n^2 errors. Therefore, chapter 3 proposed an end-to-end framework to obtain turbulence-induced FSOC link losses and a theoretically achievable information rate from WRF-simulated meteorology. The example link for that study was inspired by the 4.6 km long urban FSOC test bed in Eindhoven, The Netherlands, the Reid Photonloop (Vliet *et al.* 2025; Vliet *et al.* 2026).

To keep the framework interpretable, we limited the study to a collimated beam, deviating from the divergent beam of the original Photonloop setup. We found that errors propagate asymmetrically from C_n^2 to the link losses and to the information rate, due to non-linear relationships between C_n^2 , integrated turbulence parameters, link losses, and the estimated data rate. We concluded that system design choices, such as the selected modulation scheme and link margins, strongly influence the impact of C_n^2 errors on the final link performance estimates. For instance, if the WRF-based link loss estimates, including errors, do not exceed the threshold at which data rate degradation occurs, the link performance is unaffected. Conversely, if the link budget is already tight, even small C_n^2 errors can lead to significantly optimistic or pessimistic estimates of data rate degradation. Consequently, it is argued that reporting C_n^2 model errors alone is insufficient in an FSOC context and that end-to-end modelling for specific target links is necessary. Our work provides a starting point that prioritises interpretability, but future work should also account for more complex beam types, such as divergent or focused

beams.

In summary, the first part of this dissertation has led to the development of a comprehensive end-to-end framework to aid in the design and operation of FSOC links while accounting for local meteorology. We have identified and partially addressed multiple sources of uncertainty in the modelling chain, each of which can serve as a starting point for future work.

7.2. MACHINE LEARNING MODELLING

Despite the successful application of numerical models, inherent challenges remain. Numerical models are computationally expensive and embedded in complex multi-step workflows that require a large number of user inputs and decisions, all of which influence the final results. Formulating C_n^2 parameterizations requires extensive domain expertise, and limiting assumptions are typically made along the way. Such assumptions are necessary to keep the problem tractable but lead to deviations from reality and restrict applicability. Therefore, the second part of this work explored the potential to circumvent some of these limitations with data-driven machine learning techniques. Care was taken throughout to maintain physical interpretability and consistency to develop trustworthy ML models that can be relied on in production settings.

Temporal extrapolation of near-surface C_n^2 The starting point for the ML avenue was OTcliM presented in [chapter 4](#). Traditional mesoscale models solve a complex set of partial differential equations forced by global reanalysis data, such as ERA5 (Hersbach *et al.* 2020), to simulate the local meteorology and to estimate turbulence statistics from which C_n^2 is derived (cf. part I of this dissertation). OTcliM, however, directly relates ERA5 reanalysis data to observed C_n^2 values using gradient boosting machine (GBM) regression models. Unlike traditional approaches, OTcliM requires no manual selection of C_n^2 parameterizations or physics schemes, and the remaining sensitivity of GBM models to user-defined hyperparameters was eliminated by employing automated hyperparameter tuning. Across 17 diverse locations, we demonstrated that GBMs trained on just one year of observations can estimate near-surface C_n^2 with high accuracy from ERA5 inputs. The resulting models performed site-specific temporal extrapolation of C_n^2 to obtain a robust site-specific climatology, which can be used to design optical systems. As site surveys typically take multiple years (e.g., Schöck *et al.* (2009)), the ability to reliably extrapolate from limited observations can substantially reduce both the duration and cost of site surveys. A feature importance analysis revealed that the trained models had learned to rely on features related to wind shear and buoyancy, which are known to modulate turbulence. This finding suggested that the models picked up physically consistent relations. We also demonstrated that OTcliM generalises to locations beyond the training sites, provided the climatology remains comparable, which is promising for applications at sites without any observations. Future work should explore whether generalisability to other sites can be further improved and whether the training data requirement can be reduced to less than one year.

A key limitation of OTcliM, however, is that it is restricted to estimating near-surface C_n^2 . Such models are very useful for quantifying near-surface turbulence, as required

for terrestrial links, but for satellite-to-ground links, vertical profiles of C_n^2 are needed. Consequently, this dissertation took two steps towards vertical modelling of C_n^2 by developing the data-driven similarity-theory framework Π -ML in [chapter 5](#) and OTProf, a deep-learning-based vertical superresolution model in [chapter 6](#).

Vertical C_n^2 models for the lower boundary layer Π -ML is a general framework for deriving data-driven similarity theories using automated dimensional analysis and gradient boosting machines for regression. We demonstrated in [chapter 5](#) that Π -ML is capable of obtaining a surface layer similarity theory for C_n^2 that remains inherently interpretable, since the resulting Π groups are grounded in physical principles. A limitation of that study was that it used data only up to 25 m height and covered only a few summer weeks on Hawai'i. Considering the complex terrain and meteorology of Hawai'i, the generalizability of the results was unclear. Therefore, we applied Π -ML to a much more extensive six-year dataset covering altitudes much higher than the original study in [appendix B](#). Observations were obtained from instruments mounted at 60 m, 100 m, and 180 m on the 200 m-high Cabauw tower in the Netherlands, which is situated in a flat, homogeneous, and thus more standard environment. All conclusions from the original Π -ML work were confirmed with Cabauw data. The resulting Π groups were physically consistent and aligned with scalings known from the literature. The fitted models were found to extrapolate vertically by training on two levels and testing on the left-out level. Some scatter remained, which was attributed to a missing length scale in the scaling and should be addressed in future work. Nevertheless, these results reinforce the conclusion that Π -ML is a promising and robust framework, capable of deriving an accurate data-driven similarity theory of optical turbulence.

While the instrumentation of the Cabauw tower, reaching up to almost 200 m, is exceptional, it still confines Π -ML to the lower atmospheric boundary layer. Satellite-to-ground links require C_n^2 profiles up to 20 km altitude, which are difficult to obtain observationally. Mesoscale models such as WRF can also be used for this purpose (Masciadri *et al.* 1999; Basu and Holtslag 2022) but are computationally expensive. Simple analytical models, such as the Hufnagel-Valley (HV) model (Hufnagel 1974; Valley 1980; Ulrich 1988), offer a computationally inexpensive alternative, but they barely account for local meteorology and cannot capture details of the vertical structure of C_n^2 .

Modelling C_n^2 in the full atmospheric column In the final chapter of this dissertation, [chapter 6](#), we aimed to address this gap by drawing on the success of OTClIM and WRF-based C_n^2 estimation. We developed OTProf, a deep-learning-based superresolution and regression model. OTProf takes coarse-resolution vertical profiles of mean meteorological variables from ERA5 as input and outputs high-resolution C_n^2 profiles as produced by the WRF model. One year of WRF data is still required for training, but it can be generated efficiently by forcing WRF with a fine regional reanalysis. We showed that OTProf can estimate C_n^2 profiles from standard ERA5 inputs, and also demonstrated successful superresolution from 30 ERA5 pressure levels to 100 WRF output levels. The estimated C_n^2 profiles are smoother than the reference WRF data but retain a realistic vertical structure. This smoothing is evident when deriving integrated turbulence parameters, such as the Fried parameter and the scintillation index, which show an underestimation of es-

timated turbulence. While these underestimations should be addressed in future work, the resulting estimates remain notably better than those from the traditional HV model, which is still widely used in the FSOC community. Our results highlight that simply applying the HV model to ERA5 profiles is insufficient and that more sophisticated models, such as OTProf, are needed.

OTProf achieved good results because we employ a decoupled training strategy, in which the model is first trained on artificially coarsened WRF profiles and subsequently applied to bias-corrected ERA5 profiles during inference. This strategy is designed to compensate for spatio-temporal and distributional shifts between datasets produced by different numerical models, which can prevent the model from learning details of the profile structure. The performance still degrades slightly when ERA5-based inputs are used compared to coarsened WRF inputs. The degradation is attributed to different physics schemes and spatial resolutions between the two producing models, which cannot be compensated for by the simple quantile mapping (Cannon *et al.* 2015) applied here. Future research should therefore focus on more advanced domain adaptation techniques. Future work should also explore the application of OTProf to different topographies and climates to assess the performance of our approach beyond the Netherlands.

In summary, we believe that the second part of this dissertation has demonstrated the great potential of machine learning for optical turbulence modelling, FSOC applications, and boundary layer meteorology as a whole.

7.3. RECOMMENDATIONS AND PERSPECTIVES

This dissertation has established physics-based and ML-based modelling frameworks for C_n^2 in an FSOC context and has demonstrated their practical value for the field. At the same time, we consider all the developed ML frameworks – Π -ML, OTClIM, and OTProf – applicable to other turbulence statistics and, often, to modelling problems in boundary layer meteorology in general. In the following section, we present recommendations for future research that could build on the work presented here and discuss perspectives on broader applications of the developed techniques. The subsections focus on overarching modelling challenges identified in this dissertation and can serve as individual starting points for future work.

Ensemble and probabilistic modelling Most studies in this dissertation relied on collocated observations for validation, but such reference data are often unavailable in practice. Combining the FSOC performance framework from [chapter 3](#) with the WRF ensemble approach from [appendix A](#) could yield uncertainty estimates of link performance metrics even without collocated scintillometer observations. For close-to-surface terrestrial links such as the Photonloop, applying surface-layer-only C_n^2 parameterisations — the flux-based and gradient-based approaches — may be justified to increase the ensemble size at a manageable computational cost. Additionally, WRF is sensitive to the selected forcing dataset (Baki *et al.* 2025), e.g., ERA5 or CERRA (Schimanke *et al.* 2021), which provides another way to create an ensemble.

A WRF physics ensemble could also benefit OTProf. One challenge identified in [chapter 6](#) was a remaining dataset mismatch when training on WRF data and inferring with ERA5, attributed to differences in physics parameterisations and spatial resolution. While the resolution mismatch cannot be resolved by ensemble diversity alone, training on a diverse multi-physics WRF ensemble would expose the model to a broader range of atmospheric states, potentially making it more robust to distribution shifts and reducing the need for domain adaptation techniques.

Machine learning weather prediction has recently moved toward probabilistic outputs, achieved through diffusion-based architectures (Price *et al.* 2023) or probabilistic loss functions (Lang *et al.* 2024b; Alet *et al.* 2025). Probabilistic models not only provide uncertainty estimates of predictions, but also reduce the smoothing of fine-scale structures that is common in deterministic models (Lang *et al.* 2024b). Given that smoothing of strong turbulence events was identified as a limitation of OTProf, probabilistic modelling is a promising direction for its future development. The advantages of ensemble estimates are well established in weather prediction systems (Gneiting and Raftery 2005), so their adoption in the FSOC context is highly encouraged.

Spatial generalisation A recurring theme across part II of this dissertation is the question of spatial generalisation: does a model trained at one site produce accurate predictions at a new, unseen location? Answering this question is non-trivial because turbulence is strongly modulated by local topography and climate, so training and target sites may differ substantially.

The author recommends that, in addition to proposing potential solutions, the community should develop a formal metric to quantify the meteorological and topographic similarity between sites. In OTClIM, we resorted to a rudimentary classification of sites as rural or urban, but even within the rural category, nearby mountains or bodies of water are expected to modulate the local climatology substantially. A rigorous and quantifiable “site similarity metric” is envisioned to be beneficial in two ways. First, it could enable the quantification of the expected generalisability of already proposed models by applying the metric to their training data, thereby supporting informed decisions about production adoption. Second, it could guide dataset construction by explicitly targeting out-of-distribution regions, ensuring training sets are sufficiently diverse. Such a metric is expected to be valuable for all data-driven Earth system models, not just for optical turbulence and FSOC. A starting point could be the site classification standard of the World Meteorological Organization (2014), which currently does not consider vertical profiles or turbulence statistics.

Spatial resolution and continuous fields Although OTClIM models have been shown to generalize well in rural areas, they remain point models, whereas continuous gridded fields of C_n^2 are often desired in practice, e.g., for selecting a ground station location. It is possible to train a single model on data from multiple sites, but explicitly accounting for spatial heterogeneity is expected to yield better results. Therefore, the author suggests that future work should explore spatial interpolation techniques for weather station networks, with a focus on turbulence statistics. In Baki, Pierzyna, *et al.* (2025), a deep learning-based sparse-to-gridded interpolation framework (S2G-DI) was pro-

posed, which we recommend to adapt for C_n^2 . We do not foresee any fundamental limitations, but expect some challenges, as observations of turbulence statistics are typically even sparser than those of standard meteorological variables.

A distinct but related spatial limitation of OTCLiM is resolution. OTCLiM predictions are only available on the coarse ERA5 grid, whereas mesoscale models like WRF explicitly resolve local details through physics-based downscaling. With the increasing availability of high-resolution regional reanalysis datasets (e.g., CERRA (Schimanke *et al.* 2021), DANRA (Yang *et al.* 2025), or ICON-DREAM (Valmassoi *et al.* 2025)), ML-based downscaling is becoming a viable alternative. Steps in that direction have recently been taken, e.g., by Mardani *et al.* (2025), who presented CorrDiff, and by Saccardi, Pierzyna, *et al.* (2025), who assessed and improved the physical consistency of CorrDiff downscaling results. These studies do not yet address turbulence statistics, which we recommend as a direction for future work.

Local calibration and fine-tuning In contrast to targeting spatial generalisation, many practical applications prioritize high local accuracy. For example, the design of a specific FSOC link or the selection of a new astronomical observatory site (Schöck *et al.* 2009) requires optical turbulence estimates that are well calibrated to the conditions at that location. ML models have a natural advantage in this respect because they can extract site-specific characteristics from local observations, which are often difficult to represent in generic numerical models. OTCLiM demonstrates this clearly by achieving high accuracy (but poor generalization) at three complex urban locations, contrasting the difficulty of simulating such environments with a mesoscale model as illustrated by the WRF discrepancies noted when simulating the Photonloop site (cf. chapter 3). It is therefore suggested that an OTCLiM model be trained on the existing Photonloop scintillometer data, which is expected to improve C_n^2 estimates and all downstream link performance metrics.

Training OTCLiM from scratch, however, currently requires at least one year of local observations, which may not be available or feasible to obtain for many sites in practice. Instead, fine-tuning of pre-trained ML models could be a promising solution. In fine-tuning, an existing model, pre-trained at one or more reference sites, is updated with a small amount of local data to improve local accuracy while retaining the general physical relationships learned during pre-training. The same concept applies to OTProf, where we envision that limited in-situ observations from radiosondes (Trinquet and Vernin 2007) or SCIDARs (Masciadri *et al.* 2010) can constrain the model to produce more accurate local C_n^2 profiles.

On the numerical side, the accuracy of simulations of urban environments can potentially also be improved. The WRF model offers a range of urban parameterisations (Skamarock *et al.* 2021) that were not activated in chapter 3 to reduce sources of uncertainty. Future work should assess their benefit in an FSOC context. We attributed some of the observed discrepancies in chapter 3 to the limited resolution of WRF, so finer-resolution models, such as urban large-eddy simulations (LES) (Maronga *et al.* 2020; Owens *et al.* 2024), seem promising alternatives. However, LES is more computationally demanding and requires accurate, high-resolution land-use and building-geometry data, thereby increasing the complexity of the workflow. Validating such simulations

is also expected to become more difficult, as obtaining representative urban measurements already requires careful consideration (World Meteorological Organization 2023). Future work should explore the potential of urban LES for FSOC applications, while being mindful of the increased complexity and computational cost.

Unconventional and end-to-end applications A broader opportunity that ML opens up is the freedom to think unconventionally about both model inputs and model outputs. In appendix B, we demonstrated that a similarity theory for C_T^2 can be derived using only 1 Hz temperature measurements, without invoking Taylor’s hypothesis or requiring wind observations. This is an illustrative case in which a traditional physics-based derivation would be difficult, whereas a data-driven approach yields a practical, interpretable result. Similarly, the feature importance analysis of OTcliM revealed dependencies on wind direction and cloud cover, which we interpreted as physically consistent proxies for surface fetch and solar radiation, respectively. Recent studies have also proposed to reconstruct C_n^2 – integrated (Vorontsov *et al.* 2020) or as a vertical profile (Arvis *et al.* 2025) – from received speckle patterns of a laser beam affected by turbulence. Incorporating such non-standard features into traditional numerical models is not straightforward. ML, on the other hand, enables such approaches as long as a physically motivated connection between inputs and outputs can be expected. The author, therefore, encourages future work to explore the ingestion of non-standard observational inputs.

Machine learning also allows for the easy development of end-to-end models. Chapter 3 demonstrated the length and complexity of the traditional modelling pipeline required to estimate FSOC link performance metrics from global reanalysis data. Additionally, several assumptions and simplifications were made along the way, necessitating extensive justification and discussion. As FSOC infrastructure expands, however, more observed link performance data, such as real achieved data rates, are expected to become available. This creates the opportunity to train OTcliM-like models that skip the intermediate steps entirely and directly relate reanalysis variables to observed link performance. The expected benefit of such end-to-end models is not limited to FSOC. In many fields, the influence of meteorology must be considered (e.g., renewable energy, agriculture, transportation), but meteorological variables such as temperature, wind speed, or turbulence statistics are rarely the outputs relevant to the end user. It is therefore recommended to explore ML-based end-to-end models that directly estimate the outputs required in the target domains.

Improving physics-based models with ML While much of the previous discussion focused on recommendations to improve the ML models, machine learning can also help to improve physical models. In general, ML techniques are effective and automated tools for pattern discovery that can uncover new relations in data, potentially leading to improved process understanding or even process discovery.

II-ML illustrates this potential well. Through automated dimensional analysis, dependencies among non-dimensional groups may be discovered that were not previously captured by conventional similarity theories. Similarly, OTcliM’s feature importance analysis suggested a small dependency of C_n^2 on a gravity-wave proxy in complex terrain,

which is not currently accounted for in standard C_n^2 formulations. More broadly, existing high-fidelity numerical datasets from large-eddy or direct numerical simulations could be re-analysed with Π -ML or similar methods to systematically improve and extend existing physical parameterisations.

The interpretability of ML models, such as Π -ML, can be further enhanced by employing symbolic regression (Udrescu and Tegmark 2020). Symbolic regression yields compact analytical expressions rather than black-box models. Appendix B already showed that the relationship between the dominant Π groups is close to linear in log-log space, suggesting that a simple analytical fit may be sufficiently accurate. Such analytical expressions from symbolic regression could directly inform turbulent length scale formulations, which remain an open research question in both theory (Basu and Holtslag 2022) and operational NWP (Olson *et al.* 2019).

Finally, physics-based models and machine learning can be combined into hybrid ML models. Such models employ a traditional, typically simplified, physics-based numerical solver and couple it with an ML model acting as a “super-parameterisation”. The ML model is trained and applied *online*, meaning that it contributes to each integration time step of the numerical solver. NeuralGCM (Kochkov *et al.* 2024), for example, numerically solves equations for momentum, continuity, and thermodynamics on a global scale while using a neural network to parameterise sub-grid-scale processes. This division of labour works well because the ML component only needs to represent processes not yet captured by the physical equations, while the solver handles the well-understood governing dynamics. The resulting models are computationally more efficient than purely numerical models and more constrained than purely data-driven models, combining advantages from both worlds.

Beyond computational efficiency, we also see potential for such hybrid models to serve as tools for process discovery. Current hybrid models are typically trained on reanalysis data, so they only act as efficient emulators of the numerical system producing the reanalysis. The author believes, however, that hybrid models can be trained (or fine-tuned) on high-quality numerical outputs from LES (Schneider *et al.* 2017) or directly on observations. The resulting deep-learning parameterization can potentially be interpreted and compared against existing physical parameterisations to identify missing physics and guide future numerical model development. The recently proposed one-dimensional atmospheric model *JAX-SCM* (Pierzyna 2026c) could serve as a starting point for such work. *JAX-SCM* implements the high-order MYNN-2.5 turbulence closure (Nakanishi and Niino 2009), making it well suited for boundary-layer turbulence research. Because the model is implemented in JAX (Bradbury *et al.* 2018), it is, in principle, ready for physics–ML coupling, although Pierzyna (2026c) acknowledges that further work on the coupling details is needed.

The author strongly encourages the community to draw inspiration from the above-mentioned ideas, as more accurate physics-based models will also enable more accurate dataset generation for ML.

7.4. FINAL WORDS

This dissertation has demonstrated what physics-based mesoscale modelling and ML modelling can each achieve independently, and it also suggested that a combination of both approaches can be beneficial. Π -ML, for example, has demonstrated that revisiting classical physical frameworks, such as dimensional analysis, can lead to interpretable and physically consistent ML models. OTProf has demonstrated the potential to exploit synergies between NWP and ML, in which NWP generates high-quality turbulence data when observations are scarce, and ML efficiently estimates turbulence from more readily available inputs. The author regards this combination of physical knowledge and data-driven modelling as a promising direction for future boundary-layer turbulence models. The author argues that numerical models will remain relevant for the foreseeable future, particularly for turbulence statistics that are difficult to observe directly and are not yet fully captured by any single modelling approach. At the same time, the discussed diverse capabilities of ML enable us to approach traditional problems from new angles. The author is therefore optimistic that the future of boundary-layer meteorology will be shaped by the synergy of the two paradigms, in which the strengths of each approach are exploited to overcome their respective limitations.

IV

Appendices

A

A multi-physics ensemble modeling framework for reliable C_n^2 estimation

Chapter 2 showed that the choice of C_n^2 parameterization matters, but when setting up a WRF simulation, also the selected physics configuration affects the simulated meteorology and, consequently, the estimated C_n^2 . This appendix addresses that sensitivity by running WRF with ten different planetary boundary layer (PBL) schemes to form a multi-physics ensemble. The ensemble statistically outperforms the individual deterministic runs compared to scintillometer observations and is robust against outlier configurations. The ensemble spread also provides an uncertainty estimate of C_n^2 , which is valuable when validation data are not available.

A.1. INTRODUCTION

Free-space optical communication (FSOC) is considered a key technology to support the increasing needs of our connected data-heavy world by providing energy-efficient, secure links with high-data transmission capacity at potentially low cost. In contrast to traditional radio frequency communication, FSOC transmits data with an optical beam that propagates through the atmosphere. This propagating beam is disturbed by the atmosphere through, for example, clouds, particle/molecular scattering, and fluctuations of the atmospheric refractive index, the so-called optical turbulence (Jahid *et al.* 2022). Knowledge of these optical turbulence conditions is highly relevant for designing and deploying reliable, high-performance FSOC links, and there is an urge to quantify them well (Kaushal and Kaddoum 2017). Turbulence strongly depends on the local topography and the ever-changing meteorological conditions. Hence, performance quantification of an FSOC link requires measuring or modeling the optical turbulence conditions for the time and site of (envisioned) operation. One modeling approach is to employ well-established mesoscale models to simulate the relevant atmospheric conditions of the link and estimate the optical turbulence strength C_n^2 in a post-processing step (Bougeault *et al.* 1995; Masciadri *et al.* 1999). Mesoscale models, such as the Weather Research and Forecasting (WRF) model (Skamarock *et al.* 2021), come with a multitude of physics schemes to parameterize different atmospheric processes of the atmosphere, such as radiation and turbulence. For each process, multiple options, i.e., multiple physics schemes, are offered that all aim to solve the same general problem but vary in underlying assumptions or simplifications and, consequently, in complexity and accuracy. As a result, the simulated meteorology can differ significantly between different model configurations. Notably, the choice of the Planetary Boundary Layer (PBL) scheme, which parameterizes turbulent processes within the first ~ 1 km of the atmosphere, is expected to impact C_n^2 estimation significantly. At the same time, the PBL is the most relevant for FSOC because it distorts satellite uplinks and downlinks and is the region where terrestrial links operate (Jahid *et al.* 2022). However, the effect of different PBL schemes on C_n^2 estimation is challenging to assess a-priori. We propose to solve this issue by utilizing an ensemble of WRF models configured with different PBL schemes, similar to techniques successfully used for, e.g., precipitation forecasting (Vannitsem *et al.* 2019). Many PBL schemes have specific strong points, such as good performance of non-local schemes (e.g., the ACM2 scheme) in the daytime PBL, so a diverse PBL ensemble is expected to benefit from these specializations while being robust against outliers. Quantifying the ensemble spread yields information about the quality of the simulated C_n^2 and, subsequently, the expected quality of derived FSOC calculations, such as the link budget.

A.2. METHODOLOGY

To demonstrate the capability of our probabilistic C_n^2 simulation, we use the WRF model to simulate a 2-day test case with quiescent atmospheric conditions around the Cabauw Experimental Site for Atmospheric Research (Bosveld *et al.* 2020) (CESAR) in the Netherlands (cf. fig. A.1(a)). CESAR is selected because of the many

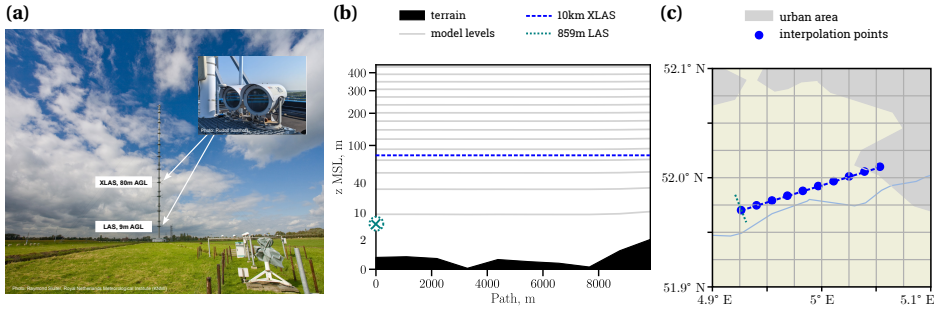


Figure A.1.: (a) Cabauw Experimental Site for Atmospheric Research (CESAR) with two scintillometers used to validate C_n^2 obtained from our probabilistic WRF simulation. The paths of both scintillometers are marked as dashed (blue) and dotted (green) lines in the sketches of (b) the vertical cross-section and (c) the top-down view of the CESAR site.

instruments deployed, which serve for accuracy assessment of our WRF ensemble, and because it has been used as a test site for FSOC links (Broekens *et al.* 2023b). Two scintillometers that measure C_n^2 along different paths are mounted at CESAR at different heights (see inset of fig. A.1(a)). Like FSOC links, scintillometers emit a laser beam at one end of the path and receive it at the other. C_n^2 is then derived by measuring the fluctuations in received power caused by optical turbulence along the path. At CESAR, a Large Apperture Scintillometer (LAS) is mounted at 9 m above ground, measuring along a 859 m path, and an EXtra Large Apperture Scintillometer (XLAS) operates at 80 m above ground along a 10 km path. Both scintillometer paths are marked in panels (b) and (c) of fig. A.1, which represent a vertical cross-section and a top-down view of the CESAR site, respectively.

The WRF ensemble is made up of ten individual WRF runs, so-called ensemble members, where each member is configured with a different PBL scheme. To connect the surface with the overlying atmosphere, WRF also requires configuring a surface layer (SL) scheme that depends on the selected PBL scheme, so we vary PBL and SL as PBL/SL pairs as listed in table A.1. It is beyond the scope of this work to compare the schemes in detail, so the reader is referred to Skamarock *et al.* (2021) for an overview or to the respective original publications of each scheme for in-depth discussions. Other physics options, such as microphysics or radiation schemes, are kept unchanged for all the simulations. In particular, we use WSM-5 (Hong *et al.* 2004) for the microphysics, the Rapid Radiative Transfer Model for GCMs (Iacono *et al.* 2008) (RRTMg) for long-wave and short-wave radiation, the Noah (Chen and Dudhia 2001) land surface model, and the Modified Kain-Fritsch Scheme (Kain 2004) to parameterize cumulus clouds in the largest domain. Each WRF simulation is forced with the same ERA5 reanalysis data provided by the European Centre for Medium-Range Weather Forecasts and set up with three nested domains with a final horizontal resolution of 1 km on the finest domain. The vertical resolution close to the surface ranges between 20 m and 30 m as can be seen in fig. A.1(b) depicting the

model levels (grey lines).

Each run yields the temporal evolution of the 3D fields of mean potential temperature θ , mean wind components (\bar{u}, \bar{v}) , relative humidity, and atmospheric pressure, from which we estimate C_n^2 following the gradient-based parametrization of Wyngaard *et al.* (1971), W71 hereafter. W71 is selected because it only requires the aforementioned meteorological variables, readily available from any PBL scheme. However, W71 is only applicable in the lower 10% of the PBL, the surface layer (SL), which changes in depth throughout the day. Typically, the SL is shallow during the night (~ 10 m) and deep during the day (~ 100 m) (Stull 1988), so the height until which W71 can be applied also changes. Therefore, our current approach cannot be used to estimate C_n^2 during the night for FSOC links or instruments mounted at large heights. Alternative C_n^2 parameterizations (Masciadri *et al.* 1999; He and Basu 2015), which are applicable anywhere in the PBL, depend on higher-order turbulence variables, such as turbulent kinetic energy, which not all PBL schemes yield. Selecting these parameterizations would limit the number of applicable PBL schemes and, thus, the ensemble diversity. Since in this study we target diversity, we accept the current surface layer constraint.

The agreement between probabilistic $\log_{10} C_n^2$ estimated from the WRF ensemble and the respective scintillometer observations is quantified on two levels: the individual member level and the ensemble level. To compare each individual model to the observations, the mean absolute error (MAE) and the Pearson correlation coefficient r are used. The MAE can be interpreted as a distance between simulated and observed time series, and the correlation coefficient quantifies how well the pattern of the time series (e.g., the diurnal cycle) is matched where $r = 0$ means no correlation and $r = 1$ perfect correlation. The Continuous Ranked Probability Score (Hersbach 2000) (CRPS) is employed to compare the predicted $\log_{10} C_n^2$ distribution, i.e., the entire ensemble, against the observations. The CRPS can be considered as a probabilistic distance and is comparable to the MAE (Hersbach 2000).

A.3. RESULTS

The simulated probabilistic C_n^2 is compared to the observed scintillometer data in fig. A.2. Both time-series plots show that the median ensemble estimates of $\log_{10} C_n^2$ in panel (a) match observations from XLAS and LAS well in magnitude and the diurnal cycle. Note that no estimates are made for the XLAS during the night because the SL is expected to be shallower than the mounting height of the XLAS. The dark 50% uncertainty bands of the probabilistic simulations are relatively narrow, indicating consistency between the ensemble members. Only the 90% uncertainty band of the simulated C_n^2 time series at the XLAS location expands significantly during daytime. Responsible for the enhancement of uncertainty are two poorly performing members that are identified in panel (d) as BouLac/Eta (\blacktriangleleft) and $E-\epsilon$ /MYNN (\star) due to their considerable overestimation of the daytime C_n^2 values at the XLAS location. This exemplifies how some PBL schemes can perform poorly in one condition but well in others (e.g., for the C_n^2 values at the LAS location). The strength of the ensemble is its robustness against such

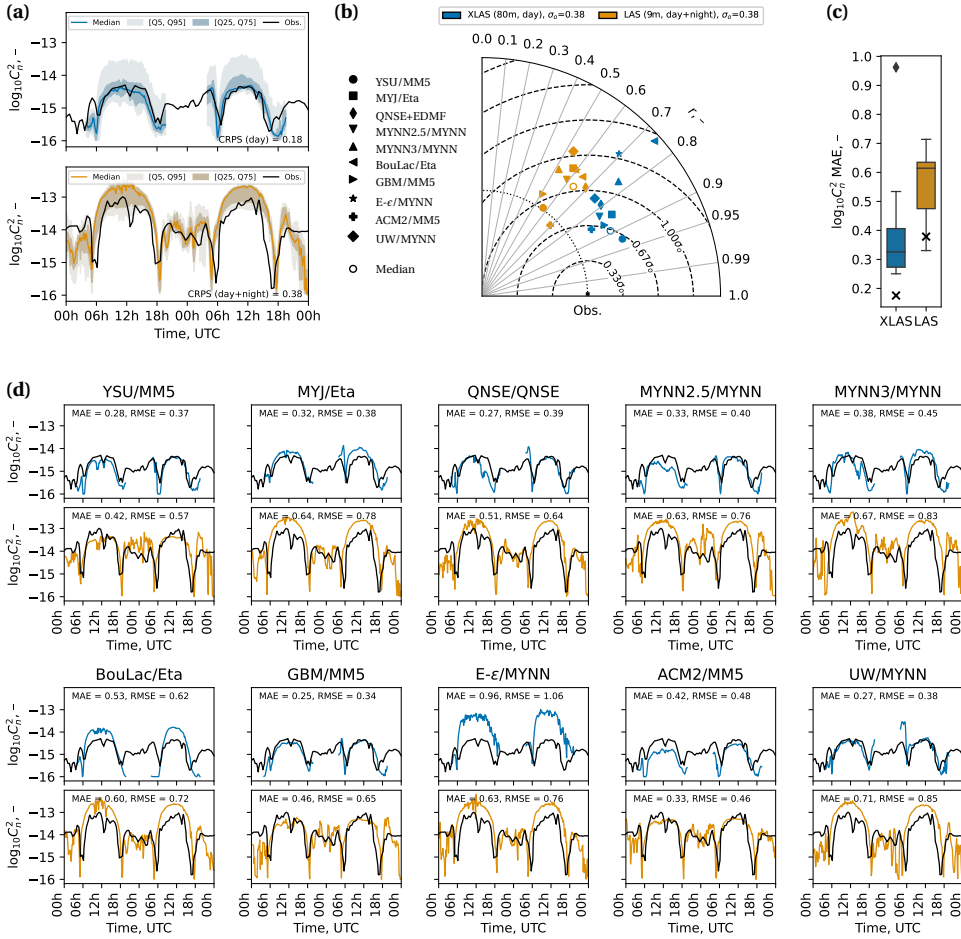


Figure A.2.: Comparison of $\log_{10} C_n^2$ estimates from the WRF simulations against scintillometer observations as (a) ensemble time series and (b) individual member model performance quantified by correlation (azimuthal) and relative error (concentric circles around observations). Distributions of members' MAE are compared to CRPS of the full ensemble in (c), and individual C_n^2 time series produced by each ensemble member are shown in (d). The titles in (d) correspond to the PBL/SL configurations listed in table A.1. Night-time estimates for the XLAS location are excluded from the entire analysis because the surface layer is expected to be too shallow for W71 to be applicable.

Table A.1.: Planetary boundary layer (PBL) and surface layer (SL) schemes used for the 10 WRF models forming our ensemble. The numbers in the square brackets correspond to the WRF configuration option for `bl_pbl_physics` and `sf_sfclay_physics`.

PBL scheme	SL scheme
● Yonsei University (YSU) (Hong and Pan 1996; Hong <i>et al.</i> 2006) [1]	Revised MM5 MO (Jiménez <i>et al.</i> 2012) [1]
■ Mellor-Yamada-Janjic (MYJ) (Mellor and Yamada 1982; Janjić 2002) [2]	Eta MO similarity (Janjic 1996; Janjić 2002) [2]
◆ QNSE with EDMF (Sukoriansky <i>et al.</i> 2005) [4]	QNSE (Sukoriansky <i>et al.</i> 2005) [4]
▼ Mellor-Yamada-Nakanishi-Niino (MYNN) 2.5 (Nakanishi and Niino 2006; Nakanishi and Niino 2009) [5]	MYNN (Dyer and Hicks 1970) [5]
▲ MYNN 3 (Nakanishi and Niino 2006; Nakanishi and Niino 2009) [6]	MYNN (Dyer and Hicks 1970) [5]
▶ Bougeault-Lacarrere (BouLac) (Bougeault and Lacarrere 1989) [8]	Eta MO similarity (Janjic 1996; Janjić 2002) [2]
► Grenier-Bretherton-McCaa (GBM) (Grenier and Bretherton 2001) [12]	Revised MM5 MO (Jiménez <i>et al.</i> 2012) [1]
★ TKE / TKE dissipation rate scheme (E- ϵ) (Langland and Liou 1996; Zhang <i>et al.</i> 2020) [16]	MYNN (Dyer and Hicks 1970) [5]
⊕ Asymmetrical Convective Model version 2 (ACM2) (Pleim 2007) [7]	Revised MM5 MO (Jiménez <i>et al.</i> 2012) [1]
◆ University of Washington (UW) (Bretherton and Park 2009) [9]	MYNN (Dyer and Hicks 1970) [5]

conditionally low performance of a few members. BouLac/Eta and $E-\epsilon$ /MYNN are also visible as outliers in the Taylor diagram (Taylor 2001) in panel (b). The Taylor diagram shows the correlation coefficient r in azimuthal direction and increasing relative error as increasing distance between the configuration markers and the observation on the x-axis. Most of the markers are right of the dotted circle, indicating that most ensemble members overestimate $\log_{10} C_n^2$ on average, as visible also in the time series in (a). Except for the two outlier configurations, most ensemble members form a cluster around the median (\circ), showing inter-model agreement. The overall performance of the ensemble is compared to the individual performance of its members in panel (c). The two box plots present the distribution of the MAE computed between simulated $\log_{10} C_n^2$ and the observations for each model, and the black crosses represent the mean CRPS. Since CRPS and MAE are comparable metrics, the CRPS values lower than most MAE values demonstrate that the probabilistic C_n^2 estimate outperforms almost all the individual ones.

A.4. CONCLUSION

In summary, we demonstrated that our probabilistic C_n^2 WRF ensemble not only captures measurements in the surface layer but also outperforms almost all the individual deterministic model runs. We also showed the more general advantages of ensemble approaches, such as their robustness against outliers and the possibility of using the ensemble spread as a reliability measure. We believe these advantages to be also highly relevant in the FSOC context when FSOC calculations are based on a probabilistic simulation rather than a deterministic one. For example, in situations where validation data are not available – at new sites or for forecasts into the future –, we envisage that the ensemble’s robustness and its uncertainty bands will make it more trustworthy than single runs. Also, downstream tasks, such as link budget calculations or link availability estimates, become potentially more reliable if they are based on robust C_n^2 estimates. The disadvantage of our methodology is increased computation time, which should be weighed with the advantages on a per-case

basis. We also note that a more extensive study is needed that covers more than two days to support our results further. Future ensembles could also vary other physics options, such as the microphysics scheme, which is expected to be significant under precipitating conditions. To extend the applicability beyond the SL, more advanced C_n^2 parametrizations (Masciadri *et al.* 1999; He and Basu 2015; Basu *et al.* 2020) can be used, but at the cost of a reduced ensemble diversity due to fewer applicable PBL schemes.

ACKNOWLEDGMENTS

This publication is part of the project *FREE – Optical Wireless Superhighways: Free photons (at home and in space)* of the research programme TTW-Perspectief, which is (partly) financed by the Dutch Research Council (NWO, project number P19-13). The scintillometer data used for validation were generated in the Ruisdael Observatory, a scientific research infrastructure, which is (partly) financed by the Dutch Research Council (NWO, grant number 184.034.015).

B

Vertical extension of Π -ML – a data-driven similarity theory of C_T^2 for Cabauw

The Π -ML framework developed in [chapter 5](#) was successfully applied to C_n^2 data obtained during an 8-week field campaign at Mauna Loa with observations up to 25 m height. This appendix extends the analysis to six years of data from the 200 m Cabauw tower, covering heights of 60 m, 100 m, and 180 m. The original conclusions are confirmed: the framework identifies physically consistent scalings, and the fitted models can extrapolate vertically by training on two levels and predicting the third. An additional use case demonstrates that C_T^2 can be estimated from only 1 Hz temperature data, without invoking Taylor’s hypothesis. While these results reinforce the robustness of Π -ML, the surface layer instrumentation still limits the approach to the lower boundary layer, motivating the full-column model in [chapter 6](#).

This chapter was presented as a poster titled “Tailor-Made Data-Driven Similarity Theories for the Temperature Structure Parameter in the Lower Atmospheric Boundary Layer” during the *Boundary Layers and Turbulence Meeting* of the *American Meteorological Society* held in June 2025 in Turin, Italy. MP received the *Best Student Poster Presentation* award (1st place).

B.1. INTRODUCTION

Pierzyna *et al.* (2023a) developed a data-driven similarity theory framework dubbed Π -ML and successfully applied it to the refractive index structure parameter C_n^2 . The authors utilized data from an 8-week seeing study performed at the Mauna Loa Observatory (MLO) on the island of Hawai'i (cf. chapter 5). During that campaign, C_n^2 was observed at 6 m, 15 m, and 25 m altitude. While this data was sufficient to demonstrate the potential of the Π -ML framework, the limited duration and vertical coverage of the dataset left room for further exploration. Therefore, we extend the analysis to a more comprehensive dataset from the Cabauw Experimental Site for Atmospheric Research (CESAR) in the Netherlands.

The 200 m tall Cabauw tower at the CESAR site is equipped with sonic anemometers at three levels (60 m, 100 m, and 180 m) and with mean wind and temperature sensors at five levels (20 m, 40 m, 80 m, 140 m, and 200 m) (Bosveld 2020) as visualized in fig. B.1. The sonic anemometers provide fast 10 Hz measurements of the three-dimensional wind vector and temperature, which are needed to compute structure function parameters. Although a fourth sonic anemometer at 3 m height is present, we exclude it from this analysis because it is expected to be influenced by local fetch rather than the larger regional fetch experienced at higher levels (Verkaik and Holtslag 2007). We view this decision as justified because our previous work (Pierzyna *et al.* 2023a) already focused on near-surface observations. Nevertheless, we acknowledge that omitting the 3 m sonic may not allow the application of the similarity theories derived in this work to the near-surface layer. The Cabauw dataset spans six years (2018 – 2024, with 2020 missing due to technical reasons), providing a rich and diverse set of atmospheric conditions for analysis. The structure function parameter of the refractive index structure function, C_n^2 , is typically determined from the structure function parameter of the sonic temperature, C_T^2 , as $C_n^2 \sim (p/T^2)^2 C_T^2$ (Gladstone and Dale 1863), unless direct scintillation observations, e.g., from a scintillometer, are available (Beyrich *et al.* 2021). As the conversion from C_T^2 to C_n^2 using local pressure p and temperature T is straightforward, we focus on C_T^2 as the target variable in the following analysis. Figure B.1 illustrates how C_T^2 is found from the second order structure function in the inertial range where $\langle (T(x + \Delta x) - T(x))^2 \rangle = C_T^2 (\Delta x)^{2/3}$.

To demonstrate the flexibility of the Π -ML framework, we apply it to two versions of the Cabauw dataset to derive two use-case-specific, i.e., tailor-made, similarity theories (ST) for C_T^2 . The first version follows traditional parameterizations utilizing vertical gradients, turbulent fluxes, and turbulent variances. Such parameterizations can be used to estimate C_T^2 from observed or simulated¹ mean variables and turbulence statistics (cf. chapter 2 and Pierzyna *et al.* (2024)). The second version takes a more practical but unconventional approach by using only features derived from a 1 Hz temperature time series. An ST based solely on 1 Hz temperature data would eliminate the need for collocated wind measurements typically required to employ Taylor's hypothesis, thereby simplifying field campaigns. Furthermore, such a parameterization could enable C_T^2 estimation using modern sensors, such as

¹To simulate turbulent statistics, large-eddy simulations or mesoscale simulations with a high-order turbulence closure scheme must be employed

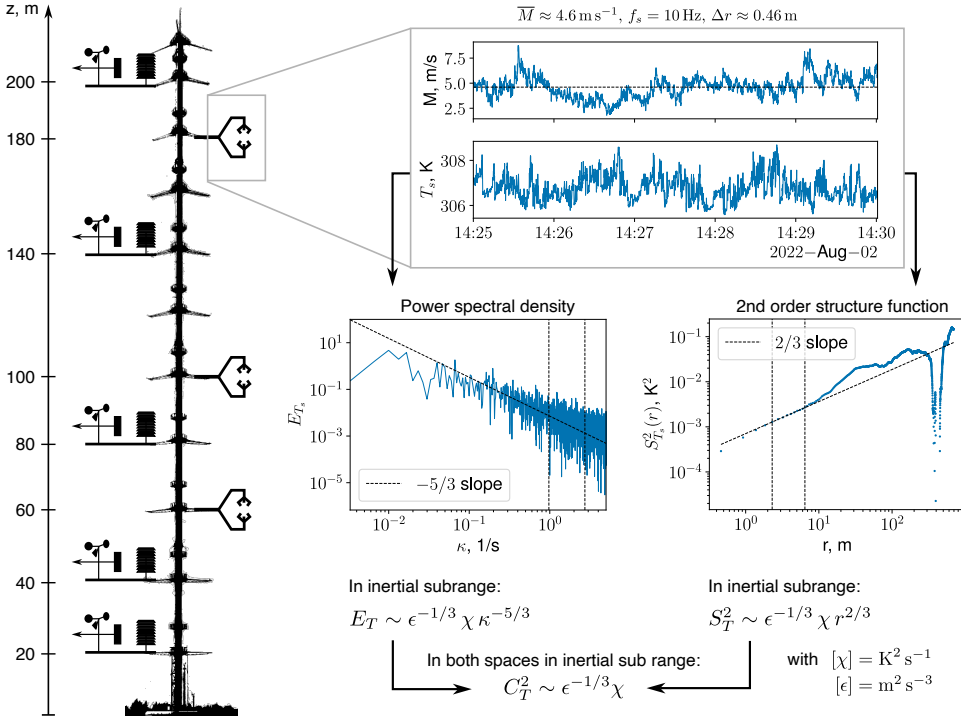


Figure B.1.: Overview of the instrumentation of the Cabauw tower (left) and the process of determining the structure function parameter C_T^2 from 10 Hz sonic temperature observations. The parallel approaches of fitting the inertial range in the power spectral density or the 2nd-order structure function (SF) of the sonic temperature are presented. Due to less noise, the SF approach is used in this study.

distributed temperature sensing (DTS) systems. Such DTS setups were recently used to obtain vertical temperature profiles at very high vertical resolution ($\Delta z \approx 0.25 \text{ m}$) (Thomas *et al.* 2012; Dai *et al.* 2023). Paired with a matching C_T^2 ST, DTS could enable the investigation of C_T^2 at similarly high resolution.

B.2. USE CASES

The specific variables and their dimensions of the two cases are summarized in table B.1. The traditional use case is similar to that of Wyngaard *et al.* (1971), who based their C_T^2 parameterization on Monin-Obukhov similarity theory (MOST). The gradients of mean potential temperature θ and mean wind speed u_i are obtained from the mean temperature and wind observations at Cabauw levels adjacent to the sonics (cf. fig. B.1). The turbulent statistics are obtained from the 10 Hz sonic observations as 5 min averages, where $(u_*)_L$ is the local friction velocity, $(\overline{w'\theta'})_L$ is the local sensible heat flux, and $(\overline{u_i'^2})_L$ and $(\overline{\theta'^2})_L$ are the local velocity and

(a) Traditional features: gradients, fluxes, and variances (5 min statistics)

	Input features (\mathbf{X})						Target (y)
	$\partial\bar{u}_i/\partial z$	$\partial\bar{\theta}/\partial z$	$(u_*)_L$	$(\overline{w'\theta'})_L$	$(\overline{u_i^2})_L$	$(\overline{\theta'^2})_L$	C_T^2
K		1		1		2	2
L		-1	1	1		1	-2/3
T	-1		-1	-1	-2		

(b) Practical features: 1 Hz temperature features only

	Input features (\mathbf{X})						Target (y)
	$\text{var}[\tilde{T}]$	$\text{skew}[\tilde{T}]$	$\text{mean}[\Delta_\tau\tilde{T}]$	$\text{var}[\Delta_\tau\tilde{T}]$	$\text{skew}[\Delta_\tau\tilde{T}]$	z	C_T^2
K	2	3	1	2	3		2
L						1	-2/3

Table B.1.: Variables and their dimensions used to derive data-driven similarity theories for a traditional (a) and a practical (b) use case.

temperature variances, respectively.

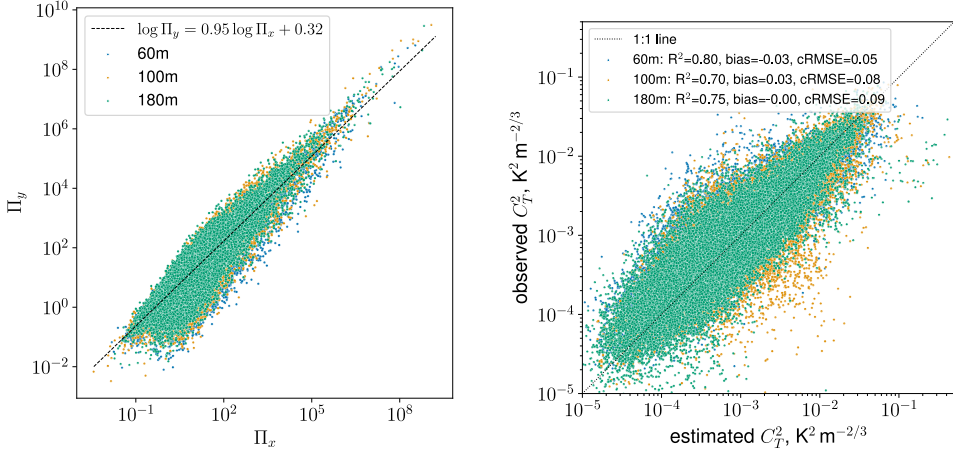
The practical 1 Hz use case relies solely on features derived from a downsampled 1 Hz version $\tilde{T}(t)$ of the 10 Hz sonic temperature signal $T(t)$. Additionally, we compute features from the increment signal of $\tilde{T}(t)$ as $\Delta\tilde{T} = \tilde{T}(t+\tau) - \tilde{T}(t)$, with a constant $\tau = 1$ s time lag. As presented in table B.1(b), the five features are the variance (var) and skewness (skew) of $\tilde{T}(t)$ over 5 min bins and the mean, variance, and skewness of the $\Delta\tilde{T}$ signal over the same 5 min bins. To capture vertical dependencies, the height above ground z is included as a feature in both use cases.

B.3. RESULTS

The variables listed in table B.1 can be combined into non-dimensional Π groups in various ways, as discussed by Pierzyna *et al.* (2023a). In the present study, we only present the final best-performing Π sets for each use case. The Π -ML selection process to arrive at these sets is based on cross-validation: each year of the six-year dataset is left out once for testing, while the remaining five years are used for training. The selected Π -ML-based similarity theories are then additionally evaluated for their ability to extrapolate vertically, i.e., to predict C_T^2 at heights not included in the training data. This is done by training a model on data from two levels and testing on the third level, using the final selected Π set. That process is repeated three times to cover all combinations of training and testing levels. The final Π sets and the performance of vertical extrapolation are presented below.

Panels (a) and (b) of fig. B.2 show the collapse of data in Π space for the best Π sets (left) and the scatter plots of vertical extrapolation (right) for the traditional and practical use cases, respectively. The two dominant Π groups obtained in the traditional setting are $\Pi_x = \sigma_{\theta,L}^2 / \theta_{*,L}^2$, where the local potential temperature variance is scaled with the local temperature scale $\theta_{*,L} = -\overline{w'\theta'}_L / u_{*,L}$, and $\Pi_y = C_T^2 z^{2/3} / \theta_{*,L}^2$, where C_T^2 is scaled with the same temperature scale and height z . Note that more Π

(a) Flux and variance-based features



(b) 1 Hz Temperature

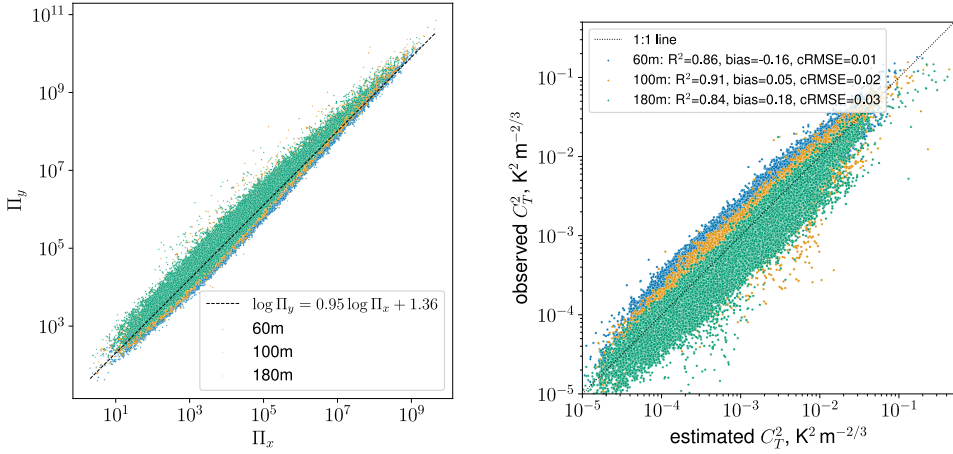


Figure B.2.: Left: Data collapse in Π space for the best-performing Π sets for the traditional (a) and practical (b) use cases. Right: Scatter plots of vertical extrapolation performance for the same use cases.

groups can be formed from the listed variables, but only Π_x is found to be relevant for prediction. Notably, Π groups containing the vertical gradients did not show high importance. This observation aligns with our earlier findings (Pierzyna *et al.* 2023a) and with Basu and Holtslag (2022) who derived $C_T^2 \sim \sigma_\theta^2 L^{-2/3}$ with the length scale L being proportional to z under neutral conditions. The Π_y expression is also similar to that of Wyngaard *et al.* (1971), suggesting physical consistency of our results.

The left panel of fig. B.2(a) shows that the observed data from all three levels collapse well in Π space, indicating that Π_x and Π_y are well suited to form a

similarity theory. The relationship between the two variables is even close to linear in log-log space, which allows us to fit a simple linear model, as indicated in the legend of the plot. The right panel indicates that vertical extrapolation performs well across levels, with low biases and root-mean-square errors. However, both panels of (a) show significant scatter, which we attribute to the absence of a stability-dependent length scale L in our formulations. Despite the scatter, our findings validate the applicability of Π -ML to larger heights and demonstrate its potential to uncover physically meaningful scalings that align with theory.

The practical use case is presented in fig. B.2(b). The best-identified Π groups are very similar to those in previous findings for the traditional case. Here, also two dominant Π groups are identified with $\Pi_x = \text{var}(\Delta\tilde{T})/\text{mean}(\Delta\tilde{T})^2$ and $\Pi_y = C_T^2 z^{2/3}/\text{mean}(\Delta\tilde{T})^2$, where the mean of the increment signal acts as a temperature scale and $\text{var}(\Delta\tilde{T})$ captures the temperature fluctuation strength similar to σ_θ^2 . Again, the observations collapse well in Π space with an even more linear relation between Π_x and Π_y . The scatter in both panels of (b) is also reduced, but a noticeable bias is present in the vertical extrapolation performance. Only the bias for the 100 m extrapolation, where training data from 60 m and 180 m is used, is close to zero, which suggests that the practical model is suitable for interpolation but not for extrapolation.

B.4. CONCLUSION

The application of Π -ML to six years of Cabauw data demonstrates the robustness of the framework across sites, height ranges, and atmospheric conditions, with all our previous findings and conclusions (Pierzyna *et al.* 2023a) being reinforced. The traditional use case demonstrated the framework's capability to identify physically consistent features and scalings consistent with established theories (e.g., Wyngaard *et al.* (1971) and Basu and Holtslag (2022)), and the practical use case presented a promising approach to more easily estimate C_T^2 in future field campaigns. Future work should address the remaining scatter of the traditional case by explicitly investigating the missing length scale. We believe that Π -ML can be a powerful tool for such work as well, potentially uncovering improved length-scale formulations. Furthermore, validation of the 1 Hz-temperature-only parameterizations with, e.g., DTS observations and collocated sonic measurements would be a valuable next step toward simplifying future field campaigns. Overall, the Π -ML framework shows great promise for deriving tailor-made similarity theories for various use cases in atmospheric turbulence research.

ACKNOWLEDGMENTS

MP thanks the Royal Netherlands Meteorological Institute (KNMI) for providing the sonic anemometer data used in this study.

DATA AVAILABILITY

The code implementing the Π -ML methodology is available on GitHub: <https://github.com/mpierzyna/piml>. The sonic anemometer training data from the Cabauw tower is not publicly available but can be obtained from the Royal Netherlands Meteorological Institute (KNMI) upon reasonable request.

Bibliography

- AHN (2021). *Actueel Hoogtebestand Nederland (AHN) 4*. Version 4. ahn.nl. URL: ahn.nl (visited on 01/10/2026).
- Albertson, J. D., M. B. Parlange, G. G. Katul, C.-R. Chu, H. Stricker, and S. Tyler (1995). "Sensible Heat Flux From Arid Regions: A Simple Flux-Variance Method". In: *Water Resources Research* 31.4, pp. 969–973. DOI: [10.1029/94WR02978](https://doi.org/10.1029/94WR02978). (Visited on 03/20/2023).
- Alet, F., I. Price, A. El-Kadi, D. Masters, S. Markou, T. R. Andersson, J. Stott, R. Lam, M. Willson, A. Sanchez-Gonzalez, and P. Battaglia (2025). *Skillful Joint Probabilistic Weather Forecasting from Marginals*. DOI: [10.48550/arXiv.2506.10772](https://doi.org/10.48550/arXiv.2506.10772). URL: <http://arxiv.org/abs/2506.10772> (visited on 08/01/2025). Pre-published.
- Alvarado, A., T. Fehenberger, B. Chen, and F. M. J. Willems (2018). "Achievable Information Rates for Fiber Optics: Applications and Computations". In: *Journal of Lightwave Technology* 36.2, pp. 424–439. DOI: [10.1109/JLT.2017.2786351](https://doi.org/10.1109/JLT.2017.2786351). URL: <http://ieeexplore.ieee.org/document/8240991/> (visited on 01/14/2026).
- Andreas, E. L. (1988). "Estimating C_n^2 over Snow and Sea Ice from Meteorological Data". In: *JOSA A* 5.4, pp. 481–495. DOI: [10.1364/JOSAA.5.000481](https://doi.org/10.1364/JOSAA.5.000481). (Visited on 10/13/2023).
- Andrews, L. C., R. L. Phillips, D. Wayne, T. Leclerc, P. Sauer, R. Crabbs, and J. Kiriazes (2009). "Near-Ground Vertical Profile of Refractive-Index Fluctuations". In: *Atmospheric Propagation VI*. Atmospheric Propagation VI. Vol. 7324. SPIE, pp. 11–22. DOI: [10.1117/12.820369](https://doi.org/10.1117/12.820369). URL: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/7324/732402/Near-ground-vertical-profile-of-refractive-index-fluctuations/10.1117/12.820369.full> (visited on 02/25/2026).
- Andrews, L. C. (2014). *Field Guide to Atmospheric Optics*. Second Edition. Vol. FG41. SPIE Field Guides. Bellingham, Washington, USA: SPIE Press.
- Andrews, L. C. and R. L. Phillips (2005). "Laser beam propagation through random media". In: *Laser Beam Propagation Through Random Media: Second Edition*.
- Arockia Bazil Raj, A., J. Arputha Vijaya Selvi, and S. Durairaj (2015). "Comparison of Different Models for Ground-Level Atmospheric Turbulence Strength (C_n^2) Prediction with a New Model According to Local Weather Data for FSO Applications". In: *Appl. Opt.* 54.4, pp. 802–815. DOI: [10.1364/AO.54.000802](https://doi.org/10.1364/AO.54.000802). (Visited on 07/29/2024).
- Arvis, M., S. Ciarella, J. Loicq, and R. Saathof (2025). "Single Shot Line-of-Sight Atmospheric Turbulence Profiling for Laser Satellite Communications with STORM". In: *Applied Optics* 64.32, pp. 9586–9597. DOI: [10.1364/AO.572341](https://doi.org/10.1364/AO.572341). URL: <https://opg.optica.org/ao/abstract.cfm?uri=ao-64-32-9586> (visited on 03/02/2026).
- Arya, S. P. (1991). "Finite-Difference Errors in Estimation of Gradients in the Atmospheric Surface Layer". In: *J. Appl. Meteor.* 30.2, pp. 251–253. DOI: [10.1175/1520-0450\(1991\)030<0251:FDEIEO>2.0.CO;2](https://doi.org/10.1175/1520-0450(1991)030<0251:FDEIEO>2.0.CO;2). (Visited on 11/30/2023).
- Baki, H., S. Basu, and G. Lavidas (2025). "Modeling Frontal Low-Level Jets and Associated Extreme Wind Power Ramps over the North Sea". In: *Wind Energy Science* 10.8,

- pp. 1575–1609. DOI: 10.5194/wes-10-1575-2025. URL: <https://wes.copernicus.org/articles/10/1575/2025/> (visited on 02/26/2026).
- Baki, H., M. Pierzyna, and S. Basu (2025). “S2G-DI: A Deep Learning Framework for Generating High-Resolution Wind Gust Fields from Sparse Mesonet Observations”. In: *EES Open Archive*. DOI: 10.22541/essoar.176031304.41826798/v2. Under Review.
- Basu, S., J. Osborn, P. He, and A. W. DeMarco (2020). “Mesoscale Modelling of Optical Turbulence in the Atmosphere: The Need for Ultrahigh Vertical Grid Resolution”. In: *Monthly Notices of the Royal Astronomical Society* 497.2, pp. 2302–2308. DOI: 10.1093/MNRAS/STAA2010.
- Basu, S. (2015). “A Simple Approach for Estimating the Refractive Index Structure Parameter (C_n^2) Profile in the Atmosphere”. In: *Opt. Lett., OL* 40.17, pp. 4130–4133. DOI: 10.1364/OL.40.004130. (Visited on 08/09/2022).
- Basu, S. and A. A. M. Holtslag (2022). “Revisiting and Revising Tatarskii’s Formulation for the Temperature Structure Parameter (C_T^2) in Atmospheric Flows”. In: *Environmental Fluid Mechanics* 22.5, pp. 1107–1119. DOI: 10.1007/s10652-022-09880-3. (Visited on 01/04/2023).
- Beason, M., G. Potvin, D. Sprung, J. McCrae, and S. Gladysz (2024). “Comparative Analysis of C_n^2 Estimation Methods for Sonic Anemometer Data”. In: *Appl. Opt.* 63.16, E94–E106. DOI: 10.1364/AO.520976. (Visited on 05/24/2024).
- Beyrich, F., O. K. Hartogensis, H. A. de Bruin, and H. C. Ward (2021). “Scintillometers”. In: *Springer Handbook of Atmospheric Measurements*. Ed. by T. Foken. Springer Handbooks. Cham: Springer International Publishing, pp. 969–997. DOI: 10.1007/978-3-030-52171-4_34. (Visited on 07/14/2023).
- Bolbasova, L. A., A. A. Andrahanov, and A. Y. Shikhovtsev (2021). “The Application of Machine Learning to Predictions of Optical Turbulence in the Surface Layer at Baikal Astrophysical Observatory”. In: *Monthly Notices of the Royal Astronomical Society* 504.4, pp. 6008–6017. DOI: 10.1093/mnras/stab953. (Visited on 03/20/2023).
- Bosveld, F. C. (2020). *The Cabauw In-situ Observational Program 2000 - Present: Instruments, Calibrations and Set-up*. Technical Report TR-384. De Bilt, The Netherlands: KNMI, p. 79.
- Bosveld, F. C., P. Baas, A. C. M. Beljaars, A. A. M. Holtslag, J. V.-G. De Arellano, and B. J. H. Van De Wiel (2020). “Fifty Years of Atmospheric Boundary-Layer Research at Cabauw Serving Weather, Air Quality and Climate”. In: *Boundary-Layer Meteorology* 177.2-3, pp. 583–612. DOI: 10.1007/s10546-020-00541-w. (Visited on 07/12/2023).
- Bougeault, P. and P. Lacarrere (1989). “Parameterization of Orography-Induced Turbulence in a Mesobeta-Scale Model”. In: *Monthly Weather Review* 117.8, pp. 1872–1890. DOI: 10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2. (Visited on 07/05/2023).
- Bougeault, P., C. D. Hui, B. Fleury, and J. Laurent (1995). “Investigation of Seeing by Means of an Atmospheric Mesoscale Numerical Simulation”. In: *Applied Optics* 34.18, pp. 3481–3488. DOI: 10.1364/AO.34.003481. (Visited on 08/29/2023).
- Bradbury, J., R. Frostig, P. Hawkins, M. J. Johnson, Y. Katariya, C. Leary, D. Maclaurin, G. Nacula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang (2018). *JAX: Composable Transformations of Python+NumPy Programs*. Version 0.9.0.1. URL: <http://github.com/jax-ml/jax>.

- Bretherton, C. S. and S. Park (2009). “A New Moist Turbulence Parameterization in the Community Atmosphere Model”. In: *Journal of Climate* 22.12, pp. 3422–3448. DOI: 10.1175/2008JCLI2556.1. (Visited on 07/11/2023).
- Broekens, K. A., N. J. Doelman, W. A. Klop, F. Silvestri, G. C. do Amaral, Y. Vos, E. P. Veldhuis, T.-C. Bui, C. W. Korevaar, I. Ferrario, and R. Saathof (2023a). “Field Test Demonstration of Adaptive Optics Pre-Correction for a Terabit Optical Communication Feeder Link”. In: *International Conference on Space Optical Systems (Accepted)*. Vancouver, BC, Canada.
- Broekens, K., W. Klop, T. Moens, M. Eschen, G. C. do Amaral, F. Silverstri, A. Oosterwijk, M. Visser, N. Doelman, L. Kaffa, *et al.* (2023b). “Adaptive optics pre-correction demonstrator for terabit optical communication”. In: *International Conference on Space Optics—ICSO 2022*. Vol. 12777. SPIE, pp. 758–766.
- Brotzge, J. A., J. Wang, C. D. Thorncroft, E. Joseph, N. Bain, N. Bassill, N. Farruggio, J. M. Freedman, K. Hemker, D. Johnston, E. Kane, S. McKim, S. D. Miller, J. R. Minder, P. Naple, S. Perez, J. J. Schwab, M. J. Schwab, and J. Sicker (2020). “A Technical Overview of the New York State Mesonet Standard Network”. In: *Journal of Atmospheric and Oceanic Technology* 37.10, pp. 1827–1845. DOI: 10.1175/JTECH-D-19-0220.1.
- Broxton, P. D., X. Zeng, D. Sulla-Menashe, and P. A. Troch (2014). “A Global Land Cover Climatology Using MODIS Data”. In: *Journal of Applied Meteorology and Climatology* 53.6, pp. 1593–1605. DOI: 10.1175/JAMC-D-13-0270.1. (Visited on 12/06/2023).
- Businger, S., R. McLaren, R. Ogasawara, D. Simons, and R. J. Wainscoat (2002). “Starcasting”. In: *Bulletin of the American Meteorological Society* 83.6, pp. 858–871. DOI: 10.1175/1520-0477(2002)083<0858:S{>}2.3.CO;2. (Visited on 11/07/2022).
- Caire, G., G. Taricco, and E. Biglieri (1998). “Bit-Interleaved Coded Modulation”. In: *IEEE Transactions on Information Theory* 44.3, pp. 927–946.
- Camboulives, A.-R., M.-T. Velluet, S. Poulenard, L. Saint-Antonin, and V. Michau (2018). “Statistical and Temporal Irradiance Fluctuations Modeling for a Ground-to-Geostationary Satellite Optical Link”. In: *Applied Optics* 57.4, pp. 709–721. DOI: 10.1364/AO.57.000709. URL: <https://opg.optica.org/ao/abstract.cfm?uri=ao-57-4-709> (visited on 02/25/2026).
- Cannon, A. J., S. R. Sobie, and T. Q. Murdock (2015). “Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes?” In: *Journal of Climate* 28.17, pp. 6938–6959. DOI: 10.1175/JCLI-D-14-00754.1. URL: <https://journals.ametsoc.org/view/journals/clim/28/17/jcli-d-14-00754.1.xml> (visited on 02/04/2026).
- Carta, J. A., S. Velázquez, and P. Cabrera (2013). “A Review of Measure-Correlate-Predict (MCP) Methods Used to Estimate Long-Term Wind Characteristics at a Target Site”. In: *Renewable and Sustainable Energy Reviews* 27, pp. 362–400. DOI: 10.1016/j.rser.2013.07.004. (Visited on 06/11/2024).
- Cheinet, S., A. Beljaars, K. Weiss-Wrana, and Y. Hurtaud (2011). “The Use of Weather Forecasts to Characterise Near-Surface Optical Turbulence”. In: *Boundary-Layer Meteorol* 138.3, pp. 453–473. DOI: 10.1007/s10546-010-9567-z. (Visited on 04/08/2024).
- Chen, F. and J. Dudhia (2001). “Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and

- Sensitivity". In: *Monthly Weather Review* 129.4, pp. 569–585. DOI: 10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2. (Visited on 07/13/2023).
- Chen, T. and C. Guestrin (2016). "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. New York, NY, USA: Association for Computing Machinery, pp. 785–794. DOI: 10.1145/2939672.2939785. (Visited on 05/24/2023).
- Cherubini, T. and S. Businger (2013). "Another Look at the Refractive Index Structure Function". In: *Journal of Applied Meteorology and Climatology* 52.2, pp. 498–506. DOI: 10.1175/JAMC-D-11-0263.1. (Visited on 12/06/2024).
- Cherubini, T., S. Businger, and R. Lyman (2008). "Modeling Optical Turbulence and Seeing over Mauna Kea: Verification and Algorithm Refinement*". In: *Journal of Applied Meteorology and Climatology* 47.12, pp. 3033–3043. DOI: 10.1175/2008JAMC1839.1. (Visited on 12/05/2024).
- Cherubini, T., R. Lyman, and S. Businger (2021). "Forecasting Seeing for the Maunakea Observatories with Machine Learning". In: *Monthly Notices of the Royal Astronomical Society* 509.1, pp. 232–245. DOI: 10.1093/mnras/stab2916. (Visited on 12/05/2024).
- Clifford, S. F., G. R. Ochs, and R. S. Lawrence (1974). "Saturation of Optical Scintillation by Strong Turbulence*". In: *JOSA* 64.2, pp. 148–154. DOI: 10.1364/JOSA.64.000148. URL: <https://opg.optica.org/josa/abstract.cfm?uri=josa-64-2-148> (visited on 01/20/2026).
- Comeron, A., F. Dios, A. Rodriguez, J. A. Rubio, M. Reyes, and A. Alonso (2005). "Modeling of Power Fluctuations Induced by Refractive Turbulence in a Multiple-Beam Ground-to-Satellite Optical Uplink". In: *Optics & Photonics 2005*. Ed. by D. G. Voelz and J. C. Ricklin. San Diego, California, USA, 589200. DOI: 10.1117/12.617219. URL: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.617219> (visited on 01/21/2025).
- Corsin, S. (1951). "On the Spectrum of Isotropic Temperature Fluctuations in an Isotropic Turbulence". In: *Journal of Applied Physics* 22.4, pp. 469–473. DOI: 10.1063/1.1699986. (Visited on 10/29/2023).
- Dai, Y., J. Boeke, B. Schilperoord, M.-C. ten Veldhuis, and B. J. H. van de Wiel (2023). "Wind Machines for Frost Damage Mitigation: A Quantitative 3D Investigation Based on Observations". In: *Agricultural and Forest Meteorology* 338, p. 109522. DOI: 10.1016/j.agrformet.2023.109522. URL: <https://www.sciencedirect.com/science/article/pii/S0168192323002137> (visited on 02/12/2026).
- Dasgupta, A., C. Cicalla, B. Mendoza, and D. Foti (2026). "Optical Turbulence Profile Modeling in the Atmospheric Boundary Layer: A Random Forest Regression Approach". In: *Radio Science* 61.2, e2025RS008369. DOI: 10.1029/2025RS008369. URL: <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2025RS008369> (visited on 02/17/2026).
- Dauphin, Y. N., A. Fan, M. Auli, and D. Grangier (2017). "Language Modeling with Gated Convolutional Networks". In: *Proceedings of the 34th International Conference on Machine Learning*. International Conference on Machine Learning. PMLR, pp. 933–941. URL: <https://proceedings.mlr.press/v70/dauphin17a.html> (visited on 02/09/2026).

- Dimitrov, S., R. Barrios, B. Matuz, G. Liva, R. Mata-Calvo, and D. Giggenbach (2016). “Digital Modulation and Coding for Satellite Optical Feeder Links with Pre-Distortion Adaptive Optics”. In: *International Journal of Satellite Communications and Networking* 34.5, pp. 625–644. DOI: 10.1002/sat.1163. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sat.1163> (visited on 02/25/2026).
- Dudhia, J. (2014). “A History of Mesoscale Model Development”. In: *Asia-Pacific J Atmos Sci* 50.1, pp. 121–131. DOI: 10.1007/s13143-014-0031-8. (Visited on 04/04/2024).
- Dyer, A. J. and B. B. Hicks (1970). “Flux-Gradient Relationships in the Constant Flux Layer”. In: *Quarterly Journal of the Royal Meteorological Society* 96.410, pp. 715–721. DOI: 10.1002/qj.49709641012. (Visited on 07/11/2023).
- Foken, T. (2017). *Micrometeorology*. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-25440-6. (Visited on 10/05/2023).
- Frehlich, R., R. Sharman, F. Vandenberghe, W. Yu, Y. Liu, J. Knivel, and G. Jumper (2010). “Estimates of Cn2 from Numerical Weather Prediction Model Output and Comparison with Thermosonde Data”. In: *Journal of Applied Meteorology and Climatology* 49.8, pp. 1742–1755. DOI: 10.1175/2010JAMC2350.1. (Visited on 11/07/2022).
- Fried, D. L. (1966). “Optical Resolution Through a Randomly Inhomogeneous Medium for Very Long and Very Short Exposures”. In: *JOSA, Vol. 56, Issue 10, pp. 1372-1379*. DOI: 10.1364/JOSA.56.001372. URL: <https://opg.optica.org/josa/abstract.cfm?uri=josa-56-10-1372> (visited on 01/20/2026).
- Frisch, U. (1995). *Turbulence: The Legacy of A.N. Kolmogorov*. Cambridge, [Eng.] ; New York: Cambridge University Press. 296 pp.
- Fuchs, C. and F Moll (2015). “Ground Station Network Optimization for Space-to-Ground Optical Communication Links”. In: *J. Opt. Commun. Netw.* 7.12, pp. 1148–1159. DOI: 10.1364/JOCN.7.001148. (Visited on 07/13/2024).
- Giggenbach, D. (2008). “Fading-Loss Assessment in Atmospheric Free-Space Optical Communication Links with on-off Keying”. In: *Optical Engineering* 47.4, p. 046001. DOI: 10.1117/1.2903095. URL: <http://opticalengineering.spiedigitallibrary.org/article.aspx?doi=10.1117/1.2903095> (visited on 11/17/2025).
- Giordano, C., J. Vernin, H. Trinquet, and C. Muñoz-Tuñón (2014). “Weather Research and Forecasting Prevision Model as a Tool to Search for the Best Sites for Astronomy: Application to La Palma, Canary Islands”. In: *Monthly Notices of the Royal Astronomical Society* 440.3, pp. 1964–1970. DOI: 10.1093/mnras/stu356. (Visited on 03/14/2023).
- Gladstone, J. H. and T. P. Dale (1863). “XIV. Researches on the Refraction, Dispersion, and Sensitiveness of Liquids”. In: *Philosophical Transactions of the Royal Society of London* 153, pp. 317–343. DOI: 10.1098/rstl.1863.0014. (Visited on 01/30/2024).
- Gneiting, T. and A. E. Raftery (2005). “Weather Forecasting with Ensemble Methods”. In: *Science* 310.5746, pp. 248–249. DOI: 10.1126/science.1115255. URL: <https://www.science.org/doi/full/10.1126/science.1115255> (visited on 02/21/2026).
- Good, R. E., R. R. Beland, E. A. Murphy, J. H. Brown, and E. M. Dewan (1988). “Atmospheric Models Of Optical Turbulence”. In: *Modeling of the Atmosphere*. 1988 Technical Symposium on Optics, Electro-Optics, and Sensors. Vol. 0928. Orlando, FL, United States: SPIE, pp. 165–186. DOI: 10.1117/12.975626. URL: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.975626> (visited on 02/25/2026).

- Grachev, A. A., E. L. Andreas, C. W. Fairall, P. S. Guest, and P. O. G. Persson (2013). “The Critical Richardson Number and Limits of Applicability of Local Similarity Theory in the Stable Boundary Layer”. In: *Boundary-Layer Meteorology* 147.1, pp. 51–82. DOI: 10.1007/s10546-012-9771-0. (Visited on 12/04/2024).
- Grenier, H. and C. S. Bretherton (2001). “A Moist PBL Parameterization for Large-Scale Models and Its Application to Subtropical Cloud-Topped Marine Boundary Layers”. In: *Monthly Weather Review* 129.3, pp. 357–377. DOI: 10.1175/1520-0493(2001)129<0357:AMPPFL>2.0.CO;2. (Visited on 07/05/2023).
- Hardy, J. W. (1998). *Adaptive Optics for Astronomical Telescopes*. Vol. 16. Oxford University Press, USA.
- Hazeu, G., D. Thomas, M. Vittek, and I. Staritsky (2023). *Landelijk Grondgebruik Nederland 2022 (LGN2022)*. Geotif file. Version 1. 4TU.ResearchData. DOI: 10.4121/688363CC-8C79-439F-BB0E-FE5D0DEB3161. URL: <https://data.4tu.nl/datasets/688363cc-8c79-439f-bb0e-fe5d0deb3161> (visited on 01/07/2026).
- He, P. and S. Basu (2015). “Mesoscale Modeling of Optical Turbulence (C_n^2) Utilizing a Novel Physically-Based Parameterization”. In: *SPIE Optical Engineering + Applications*. Ed. by A. M. J. van Eijk, C. C. Davis, and S. M. Hammel. San Diego, California, United States, 96140K. DOI: 10.1117/12.2188227. (Visited on 10/10/2022).
- (2016a). “Development of Similarity Relationships for Energy Dissipation Rate and Temperature Structure Parameter in Stably Stratified Flows: A Direct Numerical Simulation Approach”. In: *Environ Fluid Mech* 16.2, pp. 373–399. DOI: 10.1007/s10652-015-9427-y. (Visited on 03/03/2023).
- (2016b). “Extending a Surface-Layer C_n^2 Model for Strongly Stratified Conditions Utilizing a Numerically Generated Turbulence Dataset”. In: *Opt. Express, OE* 24.9, pp. 9574–9582. DOI: 10.1364/OE.24.009574. (Visited on 02/17/2023).
- Hemmati, H. (2009). *Near-Earth Laser Communications*. CRC Press.
- Henkel, C. (2023). *Google - ASL Fingerspelling Recognition, 1st Place Solution*. URL: <https://github.com/ChristofHenkel/kaggle-asl-fingerspelling-1st-place-solution> (visited on 01/28/2026).
- Hersbach, H. (2000). “Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems”. In: *Weather and Forecasting* 15.5, pp. 559–570. DOI: 10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2. (Visited on 07/14/2023).
- Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J.-N. Thépaut (2020). “The ERA5 Global Reanalysis”. In: *Quarterly Journal of the Royal Meteorological Society* 146.730, pp. 1999–2049. DOI: 10.1002/qj.3803. (Visited on 05/11/2023).
- Hill, E., J. Beckers, P. Brandt, J. Briggs, T. Brown, W. Brown, M. Collados, C. Denker, S. Fletcher, S. Hegwer, T. Horst, M. Komsa, J. Kuhn, A. Lecinski, H. Lin, S. Oncley, M. Penn, R. Radick, T. Rimmele, H. Socas-Navarro, and K. Streater (2006). “Site Testing for the Advanced Technology Solar Telescope”. In: *SPIE Astronomical Telescopes +*

- Instrumentation*. Ed. by L. M. Stepp. Vol. 6276. Orlando, Florida, USA, 62671T. DOI: 10.1117/12.673677. (Visited on 06/01/2023).
- Hong, S.-Y., J. Dudhia, and S.-H. Chen (2004). "A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation". In: *Monthly Weather Review* 132.1, pp. 103–120. DOI: 10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2. (Visited on 07/13/2023).
- Hong, S.-Y., Y. Noh, and J. Dudhia (2006). "A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes". In: *Monthly Weather Review* 134.9, pp. 2318–2341. DOI: 10.1175/MWR3199.1. (Visited on 07/05/2023).
- Hong, S.-Y. and H.-L. Pan (1996). "Nonlocal Boundary Layer Vertical Diffusion in a Medium-Range Forecast Model". In: *Monthly Weather Review* 124.10, pp. 2322–2339. DOI: 10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2. (Visited on 07/05/2023).
- Honnert, R., G. A. Efstathiou, R. J. Beare, J. Ito, A. Lock, R. Neggers, R. S. Plant, H. H. Shin, L. Tomassini, and B. Zhou (2020). "The Atmospheric Boundary Layer and the "Gray Zone" of Turbulence: A Critical Review". In: *Journal of Geophysical Research: Atmospheres* 125.13, e2019JD030317. DOI: 10.1029/2019JD030317. (Visited on 04/08/2024).
- Hufnagel, R. E. (1974). "Propagation Through Atmospheric Turbulence". In: *The Infrared Handbook*. 1st ed. Washington, D.C.: USGPO, Chap. 6.
- Hutt, D. L. (1999). "Modeling and Measurements of Atmospheric Optical Turbulence over Land". In: *Opt. Eng* 38.8, p. 1288. DOI: 10.1117/1.602188. (Visited on 03/22/2023).
- Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins (2008). "Radiative Forcing by Long-Lived Greenhouse Gases: Calculations with the AER Radiative Transfer Models". In: *Journal of Geophysical Research: Atmospheres* 113.D13. DOI: 10.1029/2008JD009944. (Visited on 07/13/2023).
- ITU (2007). *Prediction Methods Required for the Design of Terrestrial Free-Space Optical Links*. Recommendation Rec. ITU-R P.1814. International Telecommunication Union.
- Jahid, A., M. H. Alsharif, and T. J. Hall (2022). "A Contemporary Survey on Free Space Optical Communication: Potentials, Technical Challenges, Recent Advances and Research Direction". In: *Journal of Network and Computer Applications* 200, p. 103311. DOI: 10.1016/j.jnca.2021.103311. (Visited on 08/16/2023).
- Janić, Z. I. (2002). *Nonsingular Implementation of the Mellor-Yamada Level 2.5 Scheme in the NCEP Meso Model*. Office Note 437. National Centers for Environmental Prediction (U.S.) (Visited on 07/05/2023).
- Janjic, Z. I. (1996). "The Surface Layer in the NCEP Eta Model". In: *Eleventh Conference on Numerical Weather Prediction*. Norfolk, VA: Amer. Meteor. Soc, pp. 354–355.
- Jellen, C., J. Burkhardt, C. Brownell, and C. Nelson (2020). "Machine Learning Informed Predictor Importance Measures of Environmental Parameters in Maritime Optical Turbulence". In: *Appl. Opt.* 59.21, pp. 6379–6389. DOI: 10.1364/AO.397325. (Visited on 10/26/2022).
- Jellen, C., M. Oakley, C. Nelson, J. Burkhardt, and C. Brownell (2021). "Machine-Learning Informed Macro-Meteorological Models for the near-Maritime Environment". In: *Appl. Opt.* 60.11, pp. 2938–2951. DOI: 10.1364/AO.416680. (Visited on 10/26/2022).
- Jiménez, P. A., J. Dudhia, J. F. González-Rouco, J. Navarro, J. P. Montávez, and E. García-Bustamante (2012). "A Revised Scheme for the WRF Surface Layer Formulation". In:

- Monthly Weather Review* 140.3, pp. 898–918. DOI: 10.1175/MWR-D-11-00056.1. (Visited on 07/11/2023).
- Kaimal, J. C. and J. E. Gaynor (1991). “Another Look at Sonic Thermometry”. In: *Boundary-Layer Meteorol* 56.4, pp. 401–410. DOI: 10.1007/BF00119215. (Visited on 11/15/2024).
- Kain, J. S. (2004). “The Kain–Fritsch Convective Parameterization: An Update”. In: *Journal of Applied Meteorology and Climatology* 43.1, pp. 170–181. DOI: 10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2. (Visited on 07/13/2023).
- Karam, M. and T. Saad (2021). “BuckinghamPy: A Python Software for Dimensional Analysis”. In: *SoftwareX* 16, p. 100851. DOI: 10.1016/j.softx.2021.100851. (Visited on 01/31/2023).
- Kartal, S., S. Basu, and S. J. Watson (2023). “A Decision-Tree-Based Measure–Correlate–Predict Approach for Peak Wind Gust Estimation from a Global Reanalysis Dataset”. In: *Wind Energy Science* 8.10, pp. 1533–1551. DOI: 10.5194/wes-8-1533-2023. (Visited on 05/27/2024).
- Kashinath, K., M. Mustafa, A. Albert, J.-L. Wu, C. Jiang, S. Esmailzadeh, K. Azizzadenehsheli, R. Wang, A. Chattopadhyay, A. Singh, A. Manepalli, D. Chirila, R. Yu, R. Walters, B. White, H. Xiao, H. A. Tchelepi, P. Marcus, A. Anandkumar, P. Hassanzadeh, and n. Prabhat (2021). “Physics-Informed Machine Learning: Case Studies for Weather and Climate Modelling”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379.2194, p. 20200093. DOI: 10.1098/rsta.2020.0093. (Visited on 09/07/2022).
- Kaushal, H. and G. Kaddoum (2017). “Optical Communication in Space: Challenges and Mitigation Techniques”. In: *IEEE Communications Surveys & Tutorials* 19.1, pp. 57–96. DOI: 10.1109/COMST.2016.2603518. (Visited on 10/28/2022).
- Ke, G., Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu (2017). “LightGBM: A Highly Efficient Gradient Boosting Decision Tree”. In: *Advances in Neural Information Processing Systems*. Vol. 30. Curran Associates, Inc. (Visited on 05/24/2023).
- Kim, S., A. Gholami, A. Shaw, N. Lee, K. Mangalam, J. Malik, M. W. Mahoney, and K. Keutzer (2022). *Squeezeformer: An Efficient Transformer for Automatic Speech Recognition*. DOI: <https://arxiv.org/abs/2206.00888>. URL: <https://arxiv.org/abs/2206.00888>. Pre-published.
- Klop, W., R. Saathof, N. Doelman, M. Gruber, T. Moens, C. I. O. Tamayo, and C. Duque (2021). “QKD Optical Ground Terminal Developments”. In: *International Conference on Space Optics — ICSO 2020*. International Conference on Space Optics — ICSO 2020. Vol. 11852. SPIE, pp. 388–403. DOI: 10.1117/12.2599217. URL: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11852/118520Y/QKD-optical-ground-terminal-developments/10.1117/12.2599217.full> (visited on 02/26/2026).
- Kochkov, D., J. Yuval, I. Langmore, P. Norgaard, J. Smith, G. Mooers, M. Klöwer, J. Lottes, S. Rasp, P. Düben, S. Hatfield, P. Battaglia, A. Sanchez-Gonzalez, M. Willson, M. P. Brenner, and S. Hoyer (2024). *Neural General Circulation Models for Weather and Climate*. DOI: 10.48550/arXiv.2311.07222. URL: <http://arxiv.org/abs/2311.07222> (visited on 05/07/2024). Pre-published.

- Kolmogorov, A. N. (1941). “The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers”. In: *C. R. Acad. Sci. URSS* 30, pp. 301–305. (Visited on 08/04/2022).
- Kooijmans, L. M. J. and O. K. Hartogensis (2016). “Surface-Layer Similarity Functions for Dissipation Rate and Structure Parameters of Temperature and Humidity Based on Eleven Field Experiments”. In: *Boundary-Layer Meteorol* 160.3, pp. 501–527. DOI: 10.1007/s10546-016-0152-y. (Visited on 02/02/2024).
- Lam, R., A. Sanchez-Gonzalez, M. Willson, P. Wirnsberger, M. Fortunato, F. Alet, S. Ravuri, T. Ewalds, Z. Eaton-Rosen, W. Hu, A. Merose, S. Hoyer, G. Holland, O. Vinyals, J. Stott, A. Pritzel, S. Mohamed, and P. Battaglia (2023). “Learning Skillful Medium-Range Global Weather Forecasting”. In: *Science* 382.6677, pp. 1416–1421. DOI: 10.1126/science.adi2336. (Visited on 11/15/2023).
- Lang, S., M. Alexe, M. Chantry, J. Dramsch, F. Pinault, B. Raoult, M. C. A. Clare, C. Lessig, M. Maier-Gerber, L. Magnusson, Z. B. Bouallègue, A. P. Nemesio, P. D. Dueben, A. Brown, F. Pappenberger, and F. Rabier (2024a). *AIFS - ECMWF's Data-Driven Forecasting System*. DOI: 10.48550/arXiv.2406.01465. (Visited on 06/04/2024).
- Lang, S., M. Alexe, M. C. A. Clare, C. Roberts, R. Adewoyin, Z. B. Bouallègue, M. Chantry, J. Dramsch, P. D. Dueben, S. Hahner, P. Maciel, A. Prieto-Nemesio, C. O'Brien, F. Pinault, J. Polster, B. Raoult, S. Tietsche, and M. Leutbecher (2024b). *AIFS-CRPS: Ensemble Forecasting Using a Model Trained with a Loss Function Based on the Continuous Ranked Probability Score*. DOI: 10.48550/arXiv.2412.15832. URL: <http://arxiv.org/abs/2412.15832> (visited on 07/01/2025). Pre-published.
- Langland, R. H. and C.-S. Liou (1996). “Implementation of an E- ϵ Parameterization of Vertical Subgrid-Scale Mixing in a Regional Model”. In: *Monthly Weather Review* 124.5, pp. 905–918. DOI: 10.1175/1520-0493(1996)124<0905:IOAPOV>2.0.CO;2. (Visited on 07/05/2023).
- Lee, S.-H., J.-H. Lee, and B.-Y. Kim (2015). “Estimation of Turbulent Sensible Heat and Momentum Fluxes over a Heterogeneous Urban Area Using a Large Aperture Scintillometer”. In: *Advances in Atmospheric Sciences* 32.8, pp. 1092–1105. DOI: 10.1007/s00376-015-4236-2. URL: <https://doi.org/10.1007/s00376-015-4236-2> (visited on 01/19/2026).
- Lilly, D. K. (1972). “Wave Momentum Flux—A GARP Problem”. In: *Bulletin of the American Meteorological Society* 53.1, pp. 17–24. DOI: 10.1175/1520-0477-53.1.17. (Visited on 07/30/2024).
- Loshchilov, I. and F. Hutter (2017). *SGDR: Stochastic Gradient Descent with Warm Restarts*. DOI: 10.48550/arXiv.1608.03983. URL: <http://arxiv.org/abs/1608.03983> (visited on 02/09/2026). Pre-published.
- (2019). *Decoupled Weight Decay Regularization*. DOI: 10.48550/arXiv.1711.05101. URL: <http://arxiv.org/abs/1711.05101> (visited on 02/09/2026). Pre-published.
- Lovejoy, S. and D. Schertzer (2012). “Haar Wavelets, Fluctuations and Structure Functions: Convenient Choices for Geophysics”. In: *Nonlinear Processes in Geophysics* 19.5, pp. 513–527. DOI: 10.5194/npg-19-513-2012. URL: <https://npg.copernicus.org/articles/19/513/2012/> (visited on 02/10/2026).
- Lundberg, S. M. and S.-I. Lee (2017). “A Unified Approach to Interpreting Model Predictions”. In: *Proceedings of the 31st International Conference on Neural Information Pro-*

- cessing Systems*. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., pp. 4768–4777. (Visited on 08/15/2022).
- Mahrt, L. (2014). “Stably Stratified Atmospheric Boundary Layers”. In: *Annual Review of Fluid Mechanics* 46.1, pp. 23–45. DOI: 10.1146/annurev-fluid-010313-141354. (Visited on 04/06/2023).
- Mardani, M., N. Brenowitz, Y. Cohen, J. Pathak, C.-Y. Chen, C.-C. Liu, A. Vahdat, M. A. Nabian, T. Ge, A. Subramaniam, K. Kashinath, J. Kautz, and M. Pritchard (2025). “Residual Corrective Diffusion Modeling for Km-Scale Atmospheric Downscaling”. In: *Communications Earth & Environment* 6.1, p. 124. DOI: 10.1038/s43247-025-02042-5. URL: <https://www.nature.com/articles/s43247-025-02042-5> (visited on 06/29/2025).
- Maronga, B., S. Banzhaf, C. Burmeister, T. Esch, R. Forkel, D. Fröhlich, V. Fuka, K. F. Gehrke, J. Geletić, S. Giersch, T. Gronemeier, G. Groß, W. Heldens, A. Hellsten, F. Hoffmann, A. Inagaki, E. Kadasch, F. Kanani-Sühring, K. Ketelsen, B. A. Khan, C. Knigge, H. Knoop, P. Krč, M. Kurppa, H. Maamari, A. Matzarakis, M. Mauder, M. Pallasch, D. Pavlik, J. Pfafferott, J. Resler, S. Rissmann, E. Russo, M. Salim, M. Schrempf, J. Schwenkel, G. Seckmeyer, S. Schubert, M. Sühring, R. von Tils, L. Vollmer, S. Ward, B. Witha, H. Wurps, J. Zeidler, and S. Raasch (2020). “Overview of the PALM Model System 6.0”. In: *Geoscientific Model Development* 13.3, pp. 1335–1372. DOI: 10.5194/gmd-13-1335-2020. URL: <https://gmd.copernicus.org/articles/13/1335/2020/> (visited on 02/22/2026).
- Masciadri, E., J. Stoesz, S. Hagelin, and F. Lascaux (2010). “Optical Turbulence Vertical Distribution with Standard and High Resolution at Mt Graham”. In: *Monthly Notices of the Royal Astronomical Society* 404.1, pp. 144–158. DOI: 10.1111/j.1365-2966.2010.16313.x. (Visited on 03/25/2024).
- Masciadri, E., J. Vernin, and P. Bougeault (1999). “3D Mapping of Optical Turbulence Using an Atmospheric Numerical Model - I. A Useful Tool for the Ground-Based Astronomy”. In: *Astron. Astrophys. Suppl. Ser.* 137.1, pp. 185–202. DOI: 10.1051/aas:1999474. (Visited on 03/14/2023).
- McClatchey, R. A. and J. E. A. Selby (1974). *Atmospheric Attenuation of Laser Radiation from 0.76 to 31.25 μm*. AFCRL-TR-74-0003. Bedford, MA, USA: Air Force Cambridge Research Laboratories. URL: <https://apps.dtic.mil/sti/html/tr/AD0779726/> (visited on 03/10/2026).
- Mellor, G. L. (1973). “Analytic Prediction of the Properties of Stratified Planetary Surface Layers”. In: *Journal of the Atmospheric Sciences* 30.6, pp. 1061–1069. DOI: 10.1175/1520-0469(1973)030<1061:APOTPO>2.0.CO;2. (Visited on 01/30/2024).
- Mellor, G. L. and T. Yamada (1982). “Development of a Turbulence Closure Model for Geophysical Fluid Problems”. In: *Reviews of Geophysics* 20.4, pp. 851–875. DOI: 10.1029/RG020i004p00851. (Visited on 07/05/2023).
- Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P. C. Shafran, W. Ebisuzaki, D. Jović, J. Woollen, E. Rogers, E. H. Berbery, M. B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi (2006). “North American Regional Reanalysis”. In: *Bulletin of the American Meteorological Society* 87.3, pp. 343–360. DOI: 10.1175/BAMS-87-3-343. (Visited on 07/04/2024).

- Miller, M. G. and P. L. Zieske (1979). *Turbulence Environment Characterization*. Technical report RADC-TR-M9131. Rome Air Development Center.
- Milli, J., T. Rojas, B. Courtney-Barrer, F. Bian, J. Navarrete, F. Kerber, and A. Otarola (2020). *Turbulence Nowcast for the Cerro Paranal and Cerro Armazones Observatory Sites*. (Visited on 04/06/2023).
- Moene, A. F. (2003). “Effects of Water Vapour on the Structure Parameter of the Refractive Index for Near-Infrared Radiation”. In: *Boundary-Layer Meteorology* 107.3, pp. 635–653. DOI: 10.1023/A:1022807617073. (Visited on 09/08/2023).
- Molnar, C. (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Second edition. Munich, Germany: Christoph Molnar.
- Monin, A. S. and A. M. Obukhov (1954). “Basic Laws of Turbulent Mixing in the Surface Layer of the Atmosphere”. In: *Contrib. Geophys. Inst. Acad. Sci.* 151.163.
- Moosmüller, H., R. Chakrabarty, and W. Arnott (2009). “Aerosol Light Absorption and Its Measurement: A Review”. In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 110.11, pp. 844–878. DOI: 10.1016/j.jqsrt.2009.02.035. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0022407309000879> (visited on 03/09/2026).
- Nakanishi, M. and H. Niino (2006). “An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog”. In: *Boundary-Layer Meteorology* 119.2, pp. 397–407. DOI: 10.1007/s10546-005-9030-8. (Visited on 07/05/2023).
- (2009). “Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer”. In: *Journal of the Meteorological Society of Japan. Ser. II* 87.5, pp. 895–912. DOI: 10.2151/jmsj.87.895. (Visited on 07/05/2023).
- National Centers for Environmental Prediction (2015). *NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive*. DOI: 10.5065/D65D8PWK. (Visited on 07/30/2024).
- NCEP (2023). *Modified IGBP MODIS 20-Category Vegetation (Land-Use) Data*. (Visited on 12/15/2023).
- NYS Mesonet (2023). *New York State Mesonet Flux Network Data*. README. University at Albany. URL: https://www.nysmesonet.org/documents/NYSM_Readme_Flux.pdf (visited on 12/04/2024).
- Oke, T. R. (1984). “Methods in Urban Climatology”. In: *Applied Climatology* 14, pp. 19–29.
- Olson, J. B., J. S. Kenyon, W. A. Angevine, J. M. Brown, M. Pagowski, and K. Sušelj (2019). *A Description of the MYNN-EDMF Scheme and the Coupling to Other Components in WRF-ARW*. Technical Report OAR GSD-61. National Ocean and Atmospheric Administration. (Visited on 08/16/2023).
- Oncley, S. and T. Horst (2013). *Calculation of Cn2 for Visible Light and Sound from CSAT3 Sonic Anemometer Measurements*. Tech. rep., p. 15. (Visited on 09/06/2022).
- Osborn, J., M. J. Townson, O. J. D. Farley, A. Reeves, and R. M. Calvo (2021). “Adaptive Optics Pre-Compensated Laser Uplink to LEO and GEO”. In: *Optics Express* 29.4, pp. 6113–6132. DOI: 10.1364/OE.413013. URL: <https://opg.optica.org/oe/abstract.cfm?uri=oe-29-4-6113> (visited on 02/25/2026).
- Owens, S. O., D. Majumdar, C. E. Wilson, P. Bartholomew, and M. Van Reeuwijk (2024). “A Conservative Immersed Boundary Method for the Multi-Physics Urban Large-Eddy

- Simulation Model uDALES v2.0". In: *Geoscientific Model Development* 17.16, pp. 6277–6300. DOI: 10.5194/gmd-17-6277-2024. URL: <https://gmd.copernicus.org/articles/17/6277/2024/> (visited on 03/12/2025).
- Palmer, T. N., G. J. Shutts, and R. Swinbank (1986). "Alleviation of a Systematic West-erly Bias in General Circulation and Numerical Weather Prediction Models through an Orographic Gravity Wave Drag Parametrization". In: *Quart J Royal Meteorol Soc* 112.474, pp. 1001–1039. DOI: 10.1002/qj.49711247406. (Visited on 07/30/2024).
- Parenti, R. and R. Sasiela (2005). *Distribution Models for Optical Scintillation Due to At-mospheric Turbulence*. Technical Report TR-1108. Lexington, MA, USA: Lincoln Labo-ratory.
- Pham, T. V., H. Yamano, and I. Susumu (2023). "A Placement Method of Ground Stations for Optical Satellite Communications Considering Cloud Attenuation". In: *IEICE Com-mun. Express* 12.10, pp. 568–571. DOI: 10.23919/comex.2023XBL0092. (Visited on 07/13/2024).
- Pierzyna, M. (2026a). *Dataset: Estimating turbulence-induced link losses in an urban environment from observed and numerically estimated C_n^2* . Zenodo. DOI: 10.5281/zenodo.18835717. URL: <https://doi.org/10.5281/zenodo.18835717>.
- (2026b). *Dataset: OTProf: Estimating High-Resolution Profiles of Optical Turbulence (C_n^2) from Reanalysis Using Deep Learning*. Zenodo. DOI: 10.5281/zenodo.18938684. URL: <https://doi.org/10.5281/zenodo.18938684>.
- (2026c). *JAX-SCM v1.0: A Modern Atmospheric Single-Column Model for Boundary Layer Research*. DOI: 10.48550/arXiv.2605.24544. Preprint.
- Pierzyna, M., S. Basu, and R. Saathof (2025). "OTClIM: Generating a near-Surface Cli-matology of Optical Turbulence Strength (C_n^2) Using Gradient Boosting". In: *Artificial Intelligence for the Earth Systems* 4.2. DOI: 10.1175/AIES-D-24-0076.1.
- (2026). *OTProf: Estimating High-Resolution Profiles of Optical Turbulence (C_n^2) from Reanalysis Using Deep Learning*. DOI: 10.48550/arXiv.2604.09346. Under Review.
- Pierzyna, M., O. Hartogensis, S. Basu, and R. Saathof (2024). "Intercomparison of Flux, Gradient, and Variance-Based Optical Turbulence (C_n^2) Parameterizations". In: *Applied Optics* 63.16. DOI: 10.1364/AO.519942.
- Pierzyna, M., R. Saathof, and S. Basu (2023a). "PI-ML: A Dimensional Analysis-Based Ma-chine Learning Parameterization of Optical Turbulence in the Atmospheric Surface Layer". In: *Optics Letters* 48.17. DOI: 10.1364/OL.492652.
- (2023b). "A Multi-Physics Ensemble Modeling Framework for Reliable C_n^2 Estimation". In: *Proceedings of Environmental Effects on Light Propagation and Adaptive Systems VI*. Vol. 12731. SPIE, pp. 185–191. DOI: 10.1117/12.2680997.
- Pleim, J. E. (2007). "A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing". In: *Journal of Applied Mete-orology and Climatology* 46.9, pp. 1383–1395. DOI: 10.1175/JAM2539.1. (Visited on 07/11/2023).
- Poulenard, S., M. Crosnier, and A. Rissons (2015). "Ground Segment Design for Broad-band Geostationary Satellite With Optical Feeder Link". In: *J. Opt. Commun. Netw.* 7.4, pp. 325–336. DOI: 10.1364/JOCN.7.000325. (Visited on 06/27/2022).
- Price, I., A. Sanchez-Gonzalez, F. Alet, T. Ewalds, A. El-Kadi, J. Stott, S. Mohamed, P. Battaglia, R. Lam, and M. Willson (2023). *GenCast: Diffusion-based Ensemble Forecasting for Medium-*

- Range Weather*. URL: <http://arxiv.org/abs/2312.15796> (visited on 01/31/2024). Pre-published.
- Rafalimanana, A., C. Giordano, A. Ziad, and E. Aristidi (2022). “Optimal Prediction of Atmospheric Turbulence by Means of the Weather Research and Forecasting Model”. In: *PASP* 134.1035, p. 055002. DOI: 10.1088/1538-3873/ac6536. (Visited on 08/29/2022).
- Raj, A. A. B. and S. Padmavathi (2016). “Quality Metrics and Reliability Analysis of Laser Communication System”. In: *Def. Sc. Jl.* 66.2, p. 175. DOI: 10.14429/dsj.66.9707. (Visited on 03/26/2024).
- Ramachandran, P., B. Zoph, and Q. V. Le (2017). *Swish: A Self-Gated Activation Function*. Version 1. DOI: 10.48550/arXiv.1710.05941. URL: <http://arxiv.org/abs/1710.05941> (visited on 02/09/2026). Pre-published.
- Rannik, Ü., O. Peltola, and I. Mammarella (2016). “Random Uncertainties of Flux Measurements by the Eddy Covariance Technique”. In: *Atmospheric Measurement Techniques* 9.10, pp. 5163–5181. DOI: 10.5194/amt-9-5163-2016. (Visited on 06/21/2023).
- Ron, S. (2024). *LEAP - Atmospheric Physics Using AI (ClimSim), 1st Place Solution*. URL: <https://github.com/shlomoron/LEAP-solution> (visited on 01/28/2026).
- Rotach, M. W., R. Vogt, C. Bernhofer, E. Batchvarova, A. Christen, A. Clappier, B. Feddersen, S.-E. Gryning, G. Martucci, H. Mayer, V. Mitev, T. R. Oke, E. Parlow, H. Richner, M. Roth, Y.-A. Roulet, D. Ruffieux, J. A. Salmond, M. Schatzmann, and J. A. Voogt (2005). “BUBBLE – an Urban Boundary Layer Meteorology Project”. In: *Theor. Appl. Climatol.* 81.3-4, pp. 231–261. DOI: 10.1007/s00704-004-0117-9. (Visited on 05/30/2024).
- Saathof, R., R. den Breeje, W. Klop, S. Kuiper, N. Doelman, F. Pettazzi, A. Vosteen, N. Truysens, W. Crowcombe, J. Human, I. Ferrario, R. M. Calvo, J. Poliak, R. Barrios, D. Giggenbach, C. Fuchs, and S. Scalise (2017). “Optical Technologies for Terabit/s-Throughput Feeder Link”. In: *2017 IEEE International Conference on Space Optical Systems and Applications (ICSOS)*. 2017 IEEE International Conference on Space Optical Systems and Applications (ICSOS), pp. 123–129. DOI: 10.1109/ICSOS.2017.8357221. URL: <https://ieeexplore.ieee.org/document/8357221> (visited on 02/26/2026).
- Saccardi, C., M. Pierzyna, H. S. d. O. Borde, S. Monaco, C. Meo, P. Liò, R. Saathof, G. Joseph, and J. Dauwels (2025). *Assessing the Geographic Generalization and Physical Consistency of Generative Models for Climate Downscaling*. DOI: 10.48550/arXiv.2510.13722. Preprint.
- Sadot, D. (1992). “Forecasting Optical Turbulence Strength on the Basis of Macroscale Meteorology and Aerosols: Models and Validation”. In: *Opt. Eng.* 31.2, pp. 200–212. DOI: 10.1117/12.56059. (Visited on 07/29/2024).
- Sasiela, R. J. (2007). *Electromagnetic Wave Propagation in Turbulence*. 2nd ed. SPIE.
- Savage, M. J. (2009). “Estimation of Evaporation Using a Dual-Beam Surface Layer Scintillometer and Component Energy Balance Measurements”. In: *Agricultural and Forest Meteorology* 149.3, pp. 501–517. DOI: 10.1016/j.agrformet.2008.09.012. (Visited on 03/06/2023).
- Schimanke, S., M. Ridal, P. Le Moigne, L. Berggren, P. Undén, R. Randriamampianina, U. Andrea, E. Bazile, A. Bertelsen, P. Brousseau, P. Dahlgren, L. Edvinsson, A. El Said, M. Glinton, S. Hopsch, L. Isaksson, R. Mladek, E. Olsson, A. Verrelle, and Z. Wang (2021). *CERRA Sub-Daily Regional Reanalysis Data for Europe on Single Levels from 1984 to Present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.

- 24381/CDS.622A565A. URL: <https://cds.climate.copernicus.eu/doi/10.24381/cds.622a565a> (visited on 06/13/2024).
- Schneider, T., J. Teixeira, C. S. Bretherton, F. Brient, K. G. Pressel, C. Schär, and A. P. Siebesma (2017). "Climate Goals and Computing the Future of Clouds". In: *Nature Climate Change* 7.1, pp. 3–5. DOI: 10.1038/nclimate3190. URL: <https://www.nature.com/articles/nclimate3190> (visited on 03/11/2026).
- Schöck, M., S. Els, R. Riddle, W. Skidmore, T. Travouillon, R. Blum, E. Bustos, G. Chanan, S. G. Djorgovski, P. Gillett, B. Gregory, J. Nelson, A. Otárola, J. Seguel, J. Vasquez, A. Walker, D. Walker, and L. Wang (2009). "Thirty Meter Telescope Site Testing I: Overview". In: *PASP* 121.878, pp. 384–395. DOI: 10.1086/599287. (Visited on 08/03/2022).
- Scintec (2023). *Theory Manual*. v1.06. Manual.
- Shuto, Y. (2023). "Effect of Water and Aerosols Absorption on Laser Beam Propagation in Moist Atmosphere at Eye-Safe Wavelength of 1.57 μm ". In: *Journal of Electrical and Electronic Engineering*. DOI: 10.11648/j.jeeee.20231101.12. URL: <http://www.sciencepublishinggroup.com/journal/paperinfo?journalid=239&doi=10.11648/j.jeeee.20231101.12> (visited on 03/10/2026).
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, Z. Liu, J. Berner, W. Wang, J. G. Powers, M. G. Duda, D. M. Barker, and X.-Y. Huang (2021). *A Description of the Advanced Research WRF Model Version 4.3*. Tech. rep. UCAR/NCAR. DOI: 10.5065/1DFH-6P97. (Visited on 10/11/2022).
- Smith, F. G., J. S. Accetta, and D. L. Shumaker (1993). *Atmospheric Propagation of Radiation*. Vol. 2. The Infrared & Electro-Optical Systems Handbook. Infrared Information Analysis Center. (Visited on 08/10/2022).
- Sohn, H. (2024). *Stanford - Ribonanza RNA Folding, 2nd Place Solution*. URL: <https://github.com/hoyso48/Stanford---Ribonanza-RNA-Folding-2nd-place-solution> (visited on 01/28/2026).
- Spiliotis, E. (2022). "Decision Trees for Time-Series Forecasting". In: *Foresight: The International Journal of Applied Forecasting* 2022.Q1, pp. 30–44.
- Stotts, L. B. and L. C. Andrews (2023). "Improving the Hufnagel-Andrews-Phillips Refractive Index Structure Parameter Model Using Turbulent Intensity". In: *Optics Express* 31.9, p. 14265. DOI: 10.1364/OE.488544. URL: <https://opg.optica.org/abstract.cfm?URI=oe-31-9-14265> (visited on 02/27/2026).
- Stull, R. B. (1988). *An Introduction to Boundary Layer Meteorology*. Dordrecht: Kluwer Academic Publishers.
- Su, C., X. Wu, S. Wu, Q. Yang, Y. Han, C. Qing, T. Luo, and Y. Liu (2021). "In Situ Measurements and Neural Network Analysis of the Profiles of Optical Turbulence over the Tibetan Plateau". In: *Monthly Notices of the Royal Astronomical Society* 506.3, pp. 3430–3438. DOI: 10.1093/mnras/stab1792. (Visited on 03/20/2023).
- Su, C.-H., N. Eizenberg, P. Steinle, D. Jakob, P. Fox-Hughes, C. J. White, S. Rennie, C. Franklin, I. Dharssi, and H. Zhu (2019). "BARRA v1.0: The Bureau of Meteorology Atmospheric High-Resolution Regional Reanalysis for Australia". In: *Geoscientific Model Development* 12.5, pp. 2049–2068. DOI: 10.5194/gmd-12-2049-2019. (Visited on 07/04/2024).
- Sukoriansky, S., B. Galperin, and V. Perov (2005). "Application of a New Spectral Theory of Stably Stratified Turbulence to the Atmospheric Boundary Layer over Sea Ice". In:

- Boundary-Layer Meteorology* 117.2, pp. 231–257. DOI: 10.1007/s10546-004-6848-4. (Visited on 07/05/2023).
- Taylor, K. E. (2001). “Summarizing Multiple Aspects of Model Performance in a Single Diagram”. In: *Journal of Geophysical Research: Atmospheres* 106.D7, pp. 7183–7192. DOI: 10.1029/2000JD900719. (Visited on 07/10/2023).
- Thiermann, V. and H. Grassl (1992). “The Measurement of Turbulent Surface-Layer Fluxes by Use of Bichromatic Scintillation”. In: *Boundary-Layer Meteorology* 58.4, pp. 367–389. DOI: 10.1007/BF00120238. (Visited on 03/06/2023).
- Thomas, C. K., A. M. Kennedy, J. S. Selker, A. Moretti, M. H. Schroth, A. R. Smoot, N. B. Tuffillaro, and M. J. Zeeman (2012). “High-Resolution Fibre-Optic Temperature Sensing: A New Tool to Study the Two-Dimensional Structure of Atmospheric Surface-Layer Flow”. In: *Boundary-Layer Meteorology* 142.2, pp. 177–192. DOI: 10.1007/s10546-011-9672-7. URL: <https://doi.org/10.1007/s10546-011-9672-7> (visited on 02/12/2026).
- Trinquet, H. and J. Vernin (2007). “A Statistical Model to Forecast the Profile of the Index Structure Constant C_N^2 ”. In: *Environmental Fluid Mechanics* 7.5, pp. 397–407. DOI: 10.1007/s10652-007-9031-x. URL: <https://doi.org/10.1007/s10652-007-9031-x> (visited on 08/08/2022).
- Tunick, A. D. (1998). *The Refractive Index Structure Parameter/Atmospheric Optical Turbulence Model: CN2*: tech. rep. Fort Belvoir, VA: Defense Technical Information Center. DOI: 10.21236/ADA341685. (Visited on 03/26/2024).
- Udrescu, S.-M. and M. Tegmark (2020). “AI Feynman: A Physics-Inspired Method for Symbolic Regression”. In: *Science Advances* 6.16, eaay2631. DOI: 10.1126/sciadv.aay2631. URL: <https://www.science.org/doi/10.1126/sciadv.aay2631> (visited on 11/21/2022).
- Ullwer, C., D. Sprung, E. Sucher, T. Kociok, P. Grossmann, A. M. J. van Eijk, and K. Stein (2019). “Global Simulations of Cn2 Using the Weather Research and Forecast Model WRF and Comparison to Experimental Results”. In: *Laser Communication and Propagation through the Atmosphere and Oceans VIII*. Vol. 11133. SPIE, pp. 126–136. DOI: 10.1117/12.2530280. (Visited on 08/23/2022).
- Ulrich, P. B. (1988). *Hufnagel-Valley Profiles for Specified Values of the Coherence Length and Isoplanatic Angle*. MA-TN-88-013. W. J. Schafer Associates.
- Valley, G. C. (1980). “Isoplanatic Degradation of Tilt Correction and Short-Term Imaging Systems”. In: *Applied Optics* 19.4, pp. 574–577. DOI: 10.1364/AO.19.000574. URL: <https://opg.optica.org/ao/abstract.cfm?uri=ao-19-4-574> (visited on 01/30/2026).
- Valmassoi, A., H. Anlauf, A. Becker, J. D. Keller, S. Krebber, G. Zängl, R. Potthast, A. Cress, F. Fundel, T. Hanisch, M. Lange, T. Steinert, and F. Kaspar (2025). *ICON-DREAM: ICON-Dual resolution Reanalysis for Emulators, Applications and Monitoring*. open-data, GRIB. Version v1.0. Deutscher Wetterdienst. DOI: 10.5676/DWD/ICON-DREAM_V1. URL: https://opendata.dwd.de/climate_environment/CDC/help/landing_pages/doi_landingpage_ICON-DREAM_v1-en.html (visited on 03/10/2026).
- Van Kesteren, B., F. Beyrich, O. K. Hartogensis, and M. Braam (2015). “Long-Term Evaluation of the Scintec Boundary-Layer Scintillometer and the Wageningen Large-Aperture

- Scintillometer: Implications for Scintillometer Users”. In: *Boundary-Layer Meteorology* 156.2, pp. 303–323. DOI: 10.1007/s10546-015-0023-y. URL: <http://link.springer.com/10.1007/s10546-015-0023-y> (visited on 09/02/2024).
- Vannitsem, S., D. Wilks, and J. Messner, eds. (2019). *Statistical Postprocessing of Ensemble Forecasts*. Amsterdam, Netherlands: Elsevier.
- Vapnik, V. N. (1998). *Statistical Learning Theory*. Adaptive and Learning Systems for Signal Processing, Communications, and Control. New York: Wiley.
- Verkaik, J. W. and A. A. M. Holtslag (2007). “Wind Profiles, Momentum Fluxes and Roughness Lengths at Cabauw Revisited”. In: *Boundary-Layer Meteorol* 122.3, pp. 701–719. DOI: 10.1007/s10546-006-9121-1. (Visited on 09/05/2022).
- Vliet, V. van, M. van den Hout, K. Gümüş, E. Tangdiongga, and C. Okonkwo (2025). *5.7 Tbs Transmission Over a 4.6 Km Field-Deployed Free-Space Optical Link in Urban Environment*. DOI: 10.48550/arXiv.2503.21058. URL: <http://arxiv.org/abs/2503.21058> (visited on 04/01/2025). Pre-published.
- (2026). *Coherent Free-Space Optical Communications with Concurrent Turbulence Characterization in a Terrestrial Urban Link*. DOI: 10.48550/arXiv.2601.03772. URL: <http://arxiv.org/abs/2601.03772> (visited on 01/13/2026). Pre-published.
- Vorontsov, A. M., M. A. Vorontsov, G. A. Filimonov, and E. Polnau (2020). “Atmospheric Turbulence Study with Deep Machine Learning of Intensity Scintillation Patterns”. In: *Applied Sciences* 10.22, p. 8136. DOI: 10.3390/app10228136. (Visited on 07/29/2024).
- Walsh, S. and S. Schediwy (2023). “Adaptive Optics LEO Uplink Pre-Compensation with Finite Spatial Modes”. In: *Optics Letters* 48.4, pp. 880–883. DOI: 10.1364/OL.482550. URL: <https://opg.optica.org/ol/abstract.cfm?uri=ol-48-4-880> (visited on 02/25/2026).
- Wang, C., Q. Wu, M. Weimer, and E. Zhu (2021). “FLAML: A Fast and Lightweight AutoML Library”. In: *Proceedings of Machine Learning and Systems*. Vol. 3, pp. 434–447. (Visited on 11/23/2022).
- Wang, H., B. Li, X. Wu, C. Liu, Z. Hu, and P. Xu (2015). “Prediction Model of Atmospheric Refractive Index Structure Parameter in Coastal Area”. In: *Journal of Modern Optics* 62.16, pp. 1336–1346. DOI: 10.1080/09500340.2015.1037801. (Visited on 07/29/2024).
- Wang, Q., B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu (2020). *ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks*. DOI: 10.48550/arXiv.1910.03151. URL: <http://arxiv.org/abs/1910.03151> (visited on 02/25/2026). Pre-published.
- Wang, T.-i., G. R. Ochs, and S. F. Clifford (1978). “A Saturation-Resistant Optical Scintillometer to Measure Cn²”. In: *J. Opt. Soc. Am., JOSA* 68.3, pp. 334–338. DOI: 10.1364/JOSA.68.000334. (Visited on 12/14/2023).
- Wang, Y. and S. Basu (2016). “Using an Artificial Neural Network Approach to Estimate Surface-Layer Optical Turbulence at Mauna Loa, Hawaii”. In: *Opt. Lett.* 41.10, pp. 2334–2337. DOI: 10.1364/OL.41.002334. (Visited on 10/26/2022).
- Weiss-Wrana, K. and L. S. Balfour (2002). “Statistical Analysis of Measurements of Atmospheric Turbulence in Different Climates”. In: *International Symposium on Remote Sensing*. Ed. by A. Kohnle, J. D. Gonglewski, and T. J. Schmutge. Toulouse, France, p. 93. DOI: 10.1117/12.454395. (Visited on 03/26/2024).
- Wesely, M. L. (1976). “The Combined Effect of Temperature and Humidity Fluctuations on Refractive Index”. In: *Journal of Applied Meteorology and Climatology* 15.1, pp. 43–

49. DOI: 10.1175/1520-0450(1976)015<0043:TCEOTA>2.0.CO;2. (Visited on 09/04/2023).
- World Meteorological Organization (2014). *Siting Classification for Surface Observing Stations on Land*. Standard ISO 19289:2014(E).
- (2023). *Volume III – Observing Systems*. Guide WMO-No. 8. Geneva: WMO, p. 428. (Visited on 05/30/2024).
- Wyngaard, J. C., Y. Izumi, and S. A. Collins (1971). “Behavior of the Refractive-Index-Structure Parameter near the Ground”. In: *J. Opt. Soc. Am.* 61.12, pp. 1646–1650. DOI: 10.1364/JOSA.61.001646. (Visited on 01/16/2023).
- Yang, X., C. Peralta, B. Amstrup, K. S. Hintz, S. B. Thorsen, L. Denby, S. K. Christiansen, H. Schulz, S. Pelt, and M. Schreiner (2025). *DANRA: The Kilometer-Scale Danish Regional Atmospheric Reanalysis*. DOI: 10.48550/arXiv.2510.04681. URL: <http://arxiv.org/abs/2510.04681> (visited on 10/13/2025). Pre-published.
- Yasa, U. G., M. Z. Keskin, Z. G. Figen, E. Lippert, and H. Fønnum (2023). “Measuring and Modeling the Influence of Atmospheric Turbulence on a 2-Mm Laser Beam”. In: *Environmental Effects on Light Propagation and Adaptive Systems VI*. Ed. by K. Stein and S. Gladysz. Amsterdam, Netherlands: SPIE, p. 8. DOI: 10.1117/12.2678501. (Visited on 11/07/2023).
- Zhang, C., Y. Wang, and M. Xue (2020). “Evaluation of an E- ϵ and Three Other Boundary Layer Parameterization Schemes in the WRF Model over the Southeast Pacific and the Southern Great Plains”. In: *Monthly Weather Review* 148.3, pp. 1121–1145. DOI: 10.1175/MWR-D-19-0084.1. (Visited on 07/05/2023).
- Zhang, R., J. Huang, X. Wang, J. A. Zhang, and F. Huang (2016). “Effects of Precipitation on Sonic Anemometer Measurements of Turbulent Fluxes in the Atmospheric Surface Layer”. In: *J. Ocean Univ. China* 15.3, pp. 389–398. DOI: 10.1007/s11802-016-2804-4. (Visited on 02/01/2024).
- Zieliński, M., K. Fortuniak, W. Pawlak, and M. Siedlecki (2018). “Long-Term Turbulent Sensible-Heat-Flux Measurements with a Large-Aperture Scintillometer in the Centre of Łódź, Central Poland”. In: *Boundary-Layer Meteorology* 167.3, pp. 469–492. DOI: 10.1007/s10546-017-0331-5. URL: <https://doi.org/10.1007/s10546-017-0331-5> (visited on 01/19/2026).

List of publications

FIRST-AUTHOR PUBLICATIONS

- Pierzyna, M.**, R. Saathof, and S. Basu (2023a). “ Π -ML: A Dimensional Analysis-Based Machine Learning Parameterization of Optical Turbulence in the Atmospheric Surface Layer”. In: *Optics Letters* 48.17. DOI: [10.1364/OL.492652](https://doi.org/10.1364/OL.492652).
- (2023b). “A Multi-Physics Ensemble Modeling Framework for Reliable C_n^2 Estimation”. In: *Proceedings of Environmental Effects on Light Propagation and Adaptive Systems VI*. Vol. 12731. SPIE, pp. 185–191. DOI: [10.1117/12.2680997](https://doi.org/10.1117/12.2680997).
- Pierzyna, M.**, O. Hartogensis, S. Basu, and R. Saathof (2024). “Intercomparison of Flux, Gradient, and Variance-Based Optical Turbulence (C_n^2) Parameterizations”. In: *Applied Optics* 63.16. DOI: [10.1364/AO.519942](https://doi.org/10.1364/AO.519942).
- Pierzyna, M.**, S. Basu, and R. Saathof (2025). “OTClIM: Generating a near-Surface Climatology of Optical Turbulence Strength (C_n^2) Using Gradient Boosting”. In: *Artificial Intelligence for the Earth Systems* 4.2. DOI: [10.1175/AIES-D-24-0076.1](https://doi.org/10.1175/AIES-D-24-0076.1).
- Pierzyna, M.** (2026c). *JAX-SCM v1.0: A Modern Atmospheric Single-Column Model for Boundary Layer Research*. DOI: [10.48550/arXiv.2605.24544](https://doi.org/10.48550/arXiv.2605.24544). Preprint.
- Pierzyna, M.**, S. Basu, and R. Saathof (2026). *OTProf: Estimating High-Resolution Profiles of Optical Turbulence (C_n^2) from Reanalysis Using Deep Learning*. DOI: [10.48550/arXiv.2604.09346](https://doi.org/10.48550/arXiv.2604.09346). Under Review.

CONTRIBUTING-AUTHOR PUBLICATIONS

- Baki, H., **M. Pierzyna**, and S. Basu (2025). “S2G-DI: A Deep Learning Framework for Generating High-Resolution Wind Gust Fields from Sparse Mesonet Observations”. In: *EES Open Archive*. DOI: [10.22541/essoar.176031304.41826798/v2](https://doi.org/10.22541/essoar.176031304.41826798/v2). Under Review.
- Saccardi, C., **M. Pierzyna**, H. S. d. O. Borde, S. Monaco, C. Meo, P. Liò, R. Saathof, G. Joseph, and J. Dauwels (2025). *Assessing the Geographic Generalization and Physical Consistency of Generative Models for Climate Downscaling*. DOI: [10.48550/arXiv.2510.13722](https://doi.org/10.48550/arXiv.2510.13722). Preprint.

List of presentations

- **ESA/ECMWF Machine Learning Workshop**, Bologna, Italy April 2025
Presented OTProf (Pierzyna *et al.* 2026), talk.
- **Geosphere Austria**, Vienna, Austria (virtual) November 2025
“S2G-DI: A Deep Learning Framework for Generating High-Resolution Wind Gust Fields from Sparse Mesonet Observations” (Baki, Pierzyna, *et al.* 2025), invited talk.
- **IEA Wind Task 51 Workshop**, DWD, Offenbach, Germany September 2025
“S2G-DI: A Deep Learning Framework for Generating High-Resolution Wind Gust Fields from Sparse Mesonet Observations” (Baki, Pierzyna, *et al.* 2025).
- **AMS Boundary Layers and Turbulence Meeting**, Turin, Italy June 2025
“II-ML at Cabauw” (cf. appendix B), poster, received Best Student Poster Presentation Award.
- **Johns Hopkins University**, Baltimore, MD, USA October 2024
Presented OTCliM (Pierzyna *et al.* 2025) and II-ML (Pierzyna *et al.* 2023a).
- **Airforce Institute of Technology**, Dayton, OH, USA (virtual) August 2024
Presented OTCliM (Pierzyna *et al.* 2025), invited talk.
- **Optica Imaging Congress 2024**, Toulouse, France July 2024
Presented OTCliM (Pierzyna *et al.* 2025), received Best Student Paper Award.
- **ESA/ECMWF ML4EOPS**, Frascati, Italy May 2024
KelpNet: Probabilistic multi-task learning for satellite-based kelp forest monitoring, poster.
- **Dutch Meteorological Society**, Utrecht, The Netherlands November 2023
Presented II-ML (Pierzyna *et al.* 2023a), invited talk.
- **TMT International Observatory**, Pasadena, CA, USA (virtual) September 2023
Presented II-ML (Pierzyna *et al.* 2023a), invited talk.
- **SPIE Remote Sensing 2023**, Amsterdam, The Netherlands September 2023
A multi-physics ensemble modeling framework for reliable C_n^2 estimation (Pierzyna *et al.* 2023b), received Best Student Paper Award.
- **COAT 2023**, Durham, UK March 2023
Parametrizing optical turbulence (C_n^2) in the atmospheric surface layer with gradient boosting

Curriculum vitae

Maximilian Pierzyna

born 1996-04-03 in Georgsmarienhütte, Germany

EDUCATION

Doctor of Philosophy	2022-08 – 2026-10
Delft University of Technology	Delft, The Netherlands
Master of Science, Aerospace Engineering	2019-10 – 2022-07
Technical University of Braunschweig	Braunschweig, Germany
Erasmus+ Exchange, Aerospace Engineering	2020-08 – 2021-01
KTH Royal Institute of Technology	Stockholm, Sweden
Bachelor of Science, Mechanical Engineering	2015-10 – 2019-09
Technical University of Braunschweig	Braunschweig, Germany

RESEARCH VISITS

Visiting Researcher	2024-09 – 2024-11
University at Albany	Albany, NY, USA
Visiting Researcher	2024-01-10 – 2024-01-12
Fraunhofer IOSB	Ettlingen, Germany
Summer Colloquium 2023	2023-07
NCAR Advanced Study Program	Boulder, CO, USA

AWARDS AND HONOURS

Best Student Poster Presentation Award	2025-06
AMS Boundary Layer and Turbulence Meeting	
Best Student Paper Award	2024-07
Optica Imaging Congress 2024	
Editors' pick for Pierzyna <i>et al.</i> (2023a)	2023-09
Optica Publishing Group	
Best Student Paper Award	2023-09
SPIE Remote Sensing 2023	

GRANTS

Compute grant EINF-15953 "5 years of Dutch Optical Turbulence"	2025-10
300 000 core hours on the Dutch national supercomputer Snellius	SURE, Netherlands