Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Gunawardena, K. (2025). *Wind signal smoothing: in EnergyPlus (EPW) or EnviMET (csv) weather files*. Delft University of Technology, Faculteit Bouwkunde. https://doi.org/10.5281/zenodo.21419641

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.

This work is downloaded from Delft University of Technology.

Method Statement

Wind signal smoothing: in EnergyPlus (EPW) or EnviMET (csv) weather files

Version: 2.2 | December 2025 | Use freely with attribution.

© K. R. Gunawardena Ph.D (*Cantab*), TU Delft, Delft, Netherlands.

1.0 Purpose and scope

This document describes the complete methodology implemented in the MATLAB function *Smooth_Wind_Signal (V2.2),* for smoothing wind speed (WS) and wind direction (WD) inputs. The workflow reads EnergyPlus/IWEC weather files (EPW), or EnviMET CSV formatted weather input files, applies robust smoothing to the wind signals within a user-defined calendar window, and exports an Excel workbook, diagnostic plots and a revised EPW or EnviMET CSV formatted weather input file, with only the in-window WS and/or WD values modified.

This method includes a configurable wind-speed (WS) moderating/clamping step that occurs strictly within the defined smoothing window and prior to applying smoothing. When enabled via the Excel input file, WS values inside the selected window that fall below a specified minimum are uplifted to that minimum, and values above a specified maximum are capped to that maximum. The moderated series then feeds the smoothing step, while outside the window the data remains unaltered. Results and logs record whether moderation was applied and how many data points were uplifted/capped.

This revised approach facilitates simulation stability by limiting meteorological forcing WS signal extremes, while preserving diurnal variability and mitigating discontinuities at the 0°/360° signal boundary by smoothing the underlying vector components. The workflow/tool was developed to mitigate meteorological forcing errors in EnviMET or similar CFD approaches, where flow change (WS and WD) limits are needed to maintain simulation stability.

2.0 Software environment and inputs

- Implementation language: MATLAB (version-agnostic).
- Required inputs:
  - ✓ EPW or IWEC file (*.epw* or *.txt*).
  - ✓ Excel workbook containing *Parameter* | *Value* settings (sheet "*Inputs*", i.e., first sheet).
  - ✓ Output folder (prompted interactively; if cancelled, defaults to the Excel file folder).

3.0 Configuration via Excel (*Parameter* | *Value*)

The WS moderating parameters (*Apply WS Moderating, WS Moderating Minimum, WS Moderating Maximum*), govern pre-smoothing clamping inside the selected window; they default to disabled and do not affect data outside the window.

Required parameters:

- *Wind Data Smoothing* : "Windspeed Only" | "Windspeed & Direction"
- *Smoothing_StartDate* : e.g., "JAN-01, 00:00" (year ignored)
- *Smoothing_EndDate* : e.g., "JAN-01, 23:00" (year ignored)

Optional parameters (defaults in parentheses):

- *Apply WS Moderating:* “yes” | “no” (no)
- *WS Moderating Minimum:* numeric [m/s]
- *WS Moderating Maximum:* numeric [m/s]
- *WindowHours:* (7)
- *FrameLength_samples* (N): (7 – odd, ≥3 and > *PolyOrder*)
- *PolyOrder* (p): (2)
- *PadMode:* (MIRROR | REPLICATE; default MIRROR)
- *UseLocalWindowOnly:* (FALSE)
- *CalmThreshold_mps:* (0.1 m/s)
- *Output_Folder:* (<prompt>; default Excel folder)

4.0 Weather file semantics (EPW / IWEC)

EPW is a comma-separated, text-based format with eight header lines followed by hourly data records. Wind direction (WD) and wind speed (WS) are contained in core data fields #21 and #22, respectively. EPW hours are stored as the end of hour (1…24); internally, the method converts to start-of-hour timestamps (hour-1) for consistency in plotting and alignment with other tools.

5.0 Smoothing methodology

‘Smoothing’ here describes applying a workflow to suppress high-frequency noise while preserving the shape and amplitudes of slower trends and peaks of a uniformly sampled time series.

Filter selection:

1. **Primary smoother:** is a Savitzky-Golay (SG) filter with polynomial order p =PolyOrder, and frame length N = *FrameLength_samples*.
2. **Fallback smoother:** When SG requirements are not satisfied, the fall back to a moving-average (boxcar) filter with the same window length N is applied.

Windowing strategy:

Two strategies are supported:

a) Filter the full annual series and apply smoothed values only inside the selected calendar window (recommended for continuity).

b) Filter only inside the window using padded local context (*UseLocalWindowOnly* = TRUE). The window may wrap year-end (e.g., DEC-31 → JAN-05).

All moderation, smoothing, calm handling and rounding are confined to the selected window only. No values outside this window are altered, preserving original forcing outside the period of interest.

Edge handling:

To mitigate end-point bias introduced by finite windows, the series is padded before filtering using MIRROR padding by default (symmetric reflection of the data), or REPLICATE padding (edge value repetition) as an alternative.

Vector representation of wind:

To avoid angular wrap artefacts at 0°/360°, wind is converted from meteorological convention (direction *FROM*) to Cartesian components (u, v): $u = -WS \cdot sin(WD)$, $v = -WS \cdot cos(WD)$, with angles in degrees. After smoothing, WS and WD are reconstructed as $WS = hypot(\hat{u}, \hat{v})$ and $WD = mod(atan2(-\hat{u}, -\hat{v}), 360)$.

5.1 Background to Savitzky-Golay (SG) smoothing

Savitzky-Golay (SG) filtering performs a local least-squares fit of a polynomial of degree p to the samples within a sliding window of length N (typically odd). The smoothed value at the window's centre is the fitted polynomial evaluated at the central sample. This yields a linear, time-invariant finite-impulse-response (FIR) filter with symmetric convolution coefficients h_k (zero-phase when centered), determined entirely by p and N for uniformly spaced data (Press *et al.* 2007; Savitzky & Golay 1964; Schafer 2011).

Key properties

- *Shape conservation:* SG filters exactly preserve polynomials up to degree p: if the underlying signal within a window is any polynomial of degree $\leq p$, the SG output reproduces it at the center sample. Practically, this gives high passband flatness near DC (direct current, "maximally flat" behaviour) and helps retain peak heights and widths better than simple averaging at the same window length (Savitzky & Golay 1964; Schafer 2011).
- *Linear phase* (centred form): With an odd N and symmetric coefficients, the filter is zero-phase in offline use (no net time shift). For streaming/causal use, a group delay of $(N-1)/2$ samples applies (Oppenheim & Schafer 2010; Schafer 2011).
- *Noise attenuation:* For white noise with variance σ^2, the output variance is $\sigma^2 \sum_k h_k^2$. Compared to a moving-average of the same length, SG can achieve similar or better noise reduction without as much peak distortion, especially for $p \in [2,4]$ (Press *et al.* 2007; Savitzky & Golay 1964; Schafer 2011).
- *Constraints:* Standard SG smoothing requires N to be odd and $N > p+1$ (practically, $N \geq p+3$) to allow a good fit and symmetric coefficients (Savitzky & Golay 1964; Schafer 2011).

Implementation outline:

The window should be centred at sample index n with offsets $m \in \{-(N-1)/2, \ldots, (N-1)/2\}$. The Vandermonde design matrix A has entries $A_{m,j} = m^j$ for $j = 0, \ldots, p$. The convolution coefficient row for smoothing equals the centre-evaluation row extracted from the projection $A(\overline{A}\ A)^{-1}\overline{A}$, yielding a fixed FIR kernel h , reused at all positions (Press *et al.* 2007; Savitzky & Golay 1964).

Parameter guidance:

Typical choices are $p \in \{2,3,4\}$; N should be large enough to average noise but short enough to avoid biasing local curvature (rule-of-thumb: span N to cover $\approx 1-3$ periods of the highest frequency to be preserved, not attenuated) (Oppenheim & Schafer 2010; Savitzky & Golay 1964; Schafer 2011).

5.2 Background to moving-average fallback (boxcar filter)

The moving-average of length N replaces each sample by the arithmetic mean of the N samples in a sliding, typically centred, window:

$$y[n] = \frac{1}{N} \sum_{k=-(N-1)/2}^{(N-1)/2} x[n-k].$$

This is an FIR filter with uniform weights and symmetric linear phase when centred (Oppenheim & Schafer 2010; Schafer 2011).

A moving-average is equivalent to an SG filter application of order $p = 0$ (i.e., constant fit). Therefore, the fallback approach is a special case of the primary approach using the same window length N, but with minimal polynomial order (Press *et al.* 2007; Savitzky & Golay 1964; Schafer 2011).

Frequency response and noise reduction:

The discrete-time frequency response magnitude is:

$$|H(\omega)| = \frac{1}{N}\left|\frac{sin(N\omega/2)}{sin(\omega/2)}\right|,$$

...which exhibits main-lobe smoothing with sidelobe ripple. For white noise with variance σ^2, the output variance is σ^2/N (since $\sum h_k^2 = 1/N$). While effective at noise suppression, the boxcar's passband ripple can distort peaks and edges more than SG at the same N (Oppenheim & Schafer 2010; Schafer 2011).

5.3 Padding approach to limit end-point bias

Finite windows require samples beyond the data boundaries for centred convolution. Without padding, truncated windows near the ends cause amplitude bias and increased variance.

- **MIRROR** (the default approach). Reflect the series symmetrically about the boundary $(\ldots, x_2, x_1 \mid x_0, x_1, x_2, \ldots)$. Reflection preserves continuity of edge slopes for smoothly varying signals, reducing bias and edge artifacts for both SG and moving average (Oppenheim & Schafer 2010; Schafer 2011).
- **REPLICATE.** Extend with the nearest boundary value $(, \ldots, x_0, x_0 \mid x_0, x_1, \ldots)$. This enforces a zero-slope condition at the edge and can introduce bias when the true signal trends upward or downward at the boundary, but is sometimes preferred for monotone segments or when reflection would create unrealistic edge structure (Oppenheim & Schafer 2010; Schafer 2011).

Effect on phase and length

With symmetric padding and centred windows, the output is length-preserving and zero-phase in offline processing. Padding affects only the first and last $(N-1)/2$ samples; interior samples are unaffected (Oppenheim & Schafer 2010; Schafer 2011).

6.0 Practical notes

- *Normalisation:* All finite-impulse-response (FIR) kernels are normalised to unit direct current (DC) gain $\sum_k h_k = 1$, preserving signal mean.
- *Complexity:* Once SG coefficients are precomputed for (p, N), both SG and moving-average are $O(N)$ per output sample; moving-average can be implemented in $O(1)$ amortised time with running sums. For typical N, both are inexpensive (Oppenheim & Schafer 2010; Press *et al.* 2007; Savitzky & Golay 1964; Schafer 2011).
- *When to prefer which:* SG is used (with $p \geq 2$) when retaining peak height/width and local curvature matters. The moving average fallback is used when SG coefficients are unavailable on a platform or library, accepting more passband ripple for the same N (Oppenheim & Schafer 2010; Press *et al.* 2007; Savitzky & Golay 1964; Schafer 2011).
- *Calm winds and rounding of modified outputs:* Inside the smoothing window, wind direction is set to 0° for hours classified as calm, i.e. when the smoothed wind speed falls below *CalmThreshold_mps* (default 0.1 m/s). For reporting and EPW writing, modified values are rounded: WS to two decimals and WD to integer degrees.

7.0 End-to-end workflow

1. Prompt user for EPW or EnviMET formatted CSV, Excel input file, and to select an output folder (fallback to Excel input file folder if cancelled).
2. Read Parameter | Value settings from Excel (sheet "*Inputs*"; i.e., first sheet); and resolve the modifying window (*idxRange*).
3. If Apply *WS Moderating* = '*yes'*, create a moderated WS series where, within *idxRange* only, WS < Minimum is set to Minimum and WS > Maximum is set to Maximum.
4. Parse EPW header and core records; build hourly datetime vector at start-of-hour.
5. Locate start-end / *idxRange* markers (month–day–hour); support wrapping across year-end.
6. Compute valid SG frame length (odd, ≥3, > *PolyOrder*).
7. Convert (WS, WD) → (u, v); smooth using SG or moving-average; reconstruct (WS, WD).
8. Apply calm rule and rounding (WS: 2 decimal places; WD: integer) inside the window only.
9. Export Excel: *Results* (row-wise; Datetime in one column), *InputsLog* (Parameter | Value).
10. Generate and save PNG plots (WS and WD: Original vs Smoothed) and embed them in an Excel sheet "*Charts*" (Windows).
11. Write revised EPW next to the Excel output; only fields #21 (WD) and #22 (WS) changed inside the window.

8.0 EnviMET CSV workflow

- Read EnviMET formatted CSV (Version 5.9 compatible), and map inputs (explicit column positions): Date = column 1; Time = column 2; Windspeed = column 8; Wind direction = column 9.
- Output CSV preserves exactly 10 columns (same order as input, and Version 5.9 compatible).
- Excel workbook (for both EPW and CSV workflows) is named dynamically: *WindSmoothing_Output_<Start>_to_<End>_CSV.xlsx*.
- Plots saved next to the workbook with fixed names: *WindSpeed_Original_vs_Smoothed.png* and *WindDirection_Original_vs_Smoothed.png*.
- Smoothing, calm rule (WS < *CalmThreshold_mps* ⇒ WD = 0°), and rounding (WS → 2 decimals, WD → integer), remain as specified.

9.0 Quality assurance and diagnostics

- *Outside-window invariance:* Confirm no differences exist outside the selected window in EPW/CSV outputs.
- *InputsLog enhancements:* Log whether moderation was applied and counts of uplifted (below-min) and capped (above-max) samples; verify counts match violations within the window.
- *Input validation:* EPW token counts (≥22 fields), Excel parameters present or defaulted.
- *Parameter sanity:* SG frame length odd and > *PolyOrder*; window markers found once; wrap logic verified.
- *Output differences:* ΔWS and signed ΔWD columns in *Results* for the selected window.
- *Plot inspection:* WS/WD overlays to confirm whether smoothing preserves diurnal structure without phase shift.
- *EPW integrity:* header preserved verbatim; only WD/WS edited in-window; CSV row counts unchanged.

10.0 Limitations and preferences

Savitzky-Golay (1964) assumes approximately even sampling. For strongly intermittent records or substantial data gaps, moving-average fallback may be preferable. Vector smoothing assumes horizontal-only winds; vertical components (w) are not considered. The calm rule (WD = 0°) follows common reporting practice but may be adjusted as per project needs.

11.0 Reproducibility and file outputs

- **EPW:** <EPWBase>*_WSWD_smoothed.epw* – revised EPW with rounded WS (2 decimals) and WD (integer) in-window.
- **EXCEL:** *WindSmoothing_Output_*<range>.xlsx – *Results*, *InputsLog*, *Charts*.
  The Results workbook's *InputsLog* records the moderation settings and the counts of uplifted/capped samples (when enabled).
- **Plots:** *WindSpeed_Original_vs_Smoothed*.png; *WindDirection_Original_vs_Smoothed*.png – saved next to the workbook.

12.0 Smoothing sensitivity

Parameter effects and recommended adjustments for greater smoothing:

Parameter	Default	Effect	For greater smoothing
WindowHours	7	Used to derive frame length if *FrameLength_samples* invalid. Larger window → heavier smoothing.	**12-24** (e.g., 12 for moderate, 24 for strong smoothing)
***FrameLength_samples* (N)**	7	Actual Savitzky-Golay frame length (must be odd, > *PolyOrder*). Larger → smoother but less responsive to short-term changes.	**15-25** (e.g., 15 for moderate, 21 for strong smoothing)
***PolyOrder* (p)**	2	Polynomial order for SG fit. Lower order → smoother (but less shape preservation).	Keep **2** or reduce to **1** for extra smoothing
PadMode	MIRROR	Edge handling: MIRROR is reliable; REPLICATE is slightly flatter at edges.	Keep **MIRROR** (recommended)
UseLocalWindowOnly	FALSE	FALSE = filter entire year (better continuity). TRUE = only inside window (less context, can distort edges).	Keep **FALSE**
CalmThreshold_mps	0.10	Threshold for WD=0° when WS is calm. Does not affect smoothing strength.	Keep **0.10**

- ***FrameLength_samples* (N)** is the most direct control: increasing from 7 → 15 or 21 makes the filter average over a much longer period, reducing short-term fluctuations.
- ***WindowHours*** matters only if *FrameLength_samples* is invalid or left blank, but setting both consistently is good practice.
- ***PolyOrder* (p)**: Lowering from 2 → 1 makes the filter behave more like a moving average (stronger smoothing, but less curvature preservation).
- ***PadMode:*** MIRROR is best for continuity; REPLICATE can flatten edges, but does not change smoothing strength extensively.

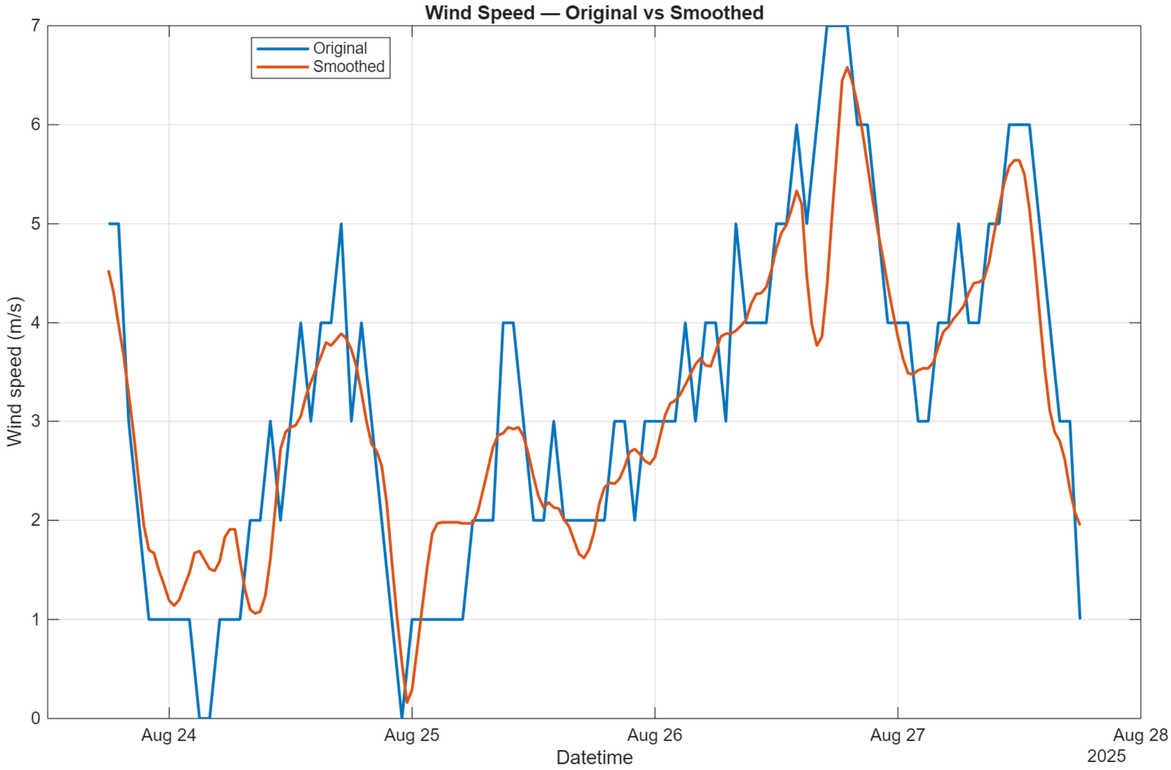

Figure 1. Example profiles of WS original vs smoothed.

13.0 Batch processing (v2.2)

The batch update extends the wind-smoothing workflow by introducing a new Excel-controlled *ProcessMode* parameter, where "Single" preserves the original one-file processing behaviour and "Batch" enables automated processing of multiple EnviMET CSV files from a user-selected folder. When *ProcessMode* = "Batch", the MATLAB script dynamically prompts the user to select the folder containing the CSV files, scans all *.csv files in that directory, and processes each file using the same smoothing settings defined in the revised Excel input workbook, including *InputFileType*, *AppliedParams*, smoothing date range, Savitzky-Golay settings, padding method, calm threshold, and optional wind-speed moderating limits. For each CSV, the script reads Date from column 1, Time from column 2, WS from column 8, and WD from column 9, applies the selected wind-speed-only or wind-speed-and-direction smoothing method within the defined calendar window, and writes a revised CSV while preserving the original column structure.

To avoid overwriting outputs during batch execution, each processed CSV is saved in its own output subfolder together with its results workbook and plots, while a separate batch summary workbook is created in the main output folder to record the processing status, output paths, number of smoothed records, and any error messages for each file. Redundant single-file CSV processing logic was consolidated into a reusable helper function so that both *Single* and *Batch* modes use the same core processing method, improving maintainability and reducing the risk of inconsistent outputs.

14.0 Python implementation (v2.2)

The new Python implementation converts the MATLAB-based wind smoothing workflow into a standalone, cross-platform tool while preserving the main processing logic. The script reads the Excel input workbook and follows the same MATLAB logic to generate the same outputs (as above).

15.0 Open sharing license with attribution

Copyright © 2025 Kanchane Gunawardena.

Permission is hereby granted, free of charge, to any person obtaining a copy of this work and associated documentation files (the "Work"), to use, reproduce, modify, publish, distribute, and create derivative works of the Work, subject to the following conditions:

1. **Attribution**
   You must give appropriate credit, provide a link to the license, and indicate if changes were made. Attribution must be clearly visible in any public use, publication, or redistribution of the Work. Suggested citation format:

   "*Wind signal smoothing: in EnergyPlus (EPW) or EnviMET (csv) weather files,* by K. Gunawardena, licensed under the Open Sharing License with Attribution".

   Gunawardena, K. R. (2025). *Wind signal smoothing: in EnergyPlus (EPW) or EnviMET (csv) weather files*. Delft: TU Delft.

2. **No Warranty**
   The Work is provided "as is", without warranty of any kind, express or implied, including but not limited to the warranties of merchantability, fitness for a particular purpose, and noninfringement.

3. **Redistribution**
   Copies or substantial portions of the Work must retain this license and attribution notice.

4. **Modifications**
   You may modify the Work, but must clearly indicate that changes were made and retain attribution to the original author.

Acknowledgment

The work described here is part of the Multifunctional Quay Wall project, funded by the Dutch National Growth Fund (NGF), under the *Future-proof Living Environment* programme, and facilitated by the Amsterdam Institute for Advanced Metropolitan Solutions (AMS).

References

ASHRAE (n.d.). *International Weather Files for Energy Calculations 2.0 (IWEC2)*. Available at: https://www.ashrae.org/technical-resources (Accessed 11 Nov 2025).

ECMWF (2024). *ERA5: How to calculate wind speed and wind direction from u and v components of the wind?* Copernicus Knowledge Base. Available at: https://confluence.ecmwf.int (Accessed 11 Nov 2025).

EnergyPlus (2015). *EnergyPlus Weather File (EPW) Data Dictionary*. Big Ladder Software (Online). Available at: https://bigladdersoftware.com (Accessed 11 Nov 2025).

ENVI-met (n.d.). *Forcing Manager*. Available at: https://envi-met.info (Accessed 11 Nov 2025).

ENVI-met (n.d.). *Meteorological Conditions & Full Forcing* (Tutorial). Available at: https://envi-met.com/tutorials (Accessed 11 Nov 2025).

Gunawardena, K. (2018a). *Fundamentals of Urban Heat Islands: Concise guide for architects and urban planners.* Cambridge: University of Cambridge.

Gunawardena, K. (2018b). *Methodologies for assessing urban microclimates.* In University of Cambridge, Department of Architecture, PhD Symposium, Cambridge: University of Cambridge, pp. 1-20.

Jia, H. (n.d.). *Read and modify an EnergyPlus Weather File (EPW) - eplusr*. Available at: https://hongyuanjia.github.io/eplusr/reference/Epw (Accessed 11 Nov 2025).

NCAR/EOL (n.d.). *Wind Direction Quick Reference*. Earth Observing Laboratory. Available at: https://www.eol.ucar.edu/content/wind-direction-quick-reference (Accessed 11 Nov 2025).

Oppenheim, A. V, & Schafer, R. W. (2010). *Discrete-Time Signal Processing*, 3rd edn, Upper Saddle River, NJ: Pearson.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical Recipes: The Art of Scientific Computing*, 3rd edn, Cambridge: Cambridge University Press.

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8), 1627–1639.

Schafer, R. W. (2011). What is a Savitzky-Golay filter? *IEEE Signal Processing Magazine*, 28(4), 111–117.