Lightning forecast using the Lightning Potential Index in the HARMONIE-AROME model

D.D. Gur

Graduation committee:

W. de Rooy
Dr. M.J. Schmeits
Prof. Dr. A.P. Siebesma
Dr. M.A. Schleiss

This page has intentionally been left blank.

Cover picture credit: Barešić, Š. (2021)

Lightning forecast using the Lightning Potential Index in the HARMONIE-AROME model

by

D.D. (Daniëlla) Gur

To obtain the degree of Master of Science in Geoscience and Remote Sensing
at Delft University of Technology

To be defended in public on Tuesday November 23, 2021

Graduation Committee

Daily supervisor	W. (Wim) de Rooy	Royal Dutch Meteorological Institute (KNMI)
Daily supervisor	Dr. M.J. (Maurice) Schmeits	Royal Dutch Meteorological Institute (KNMI)
Supervisor	Prof. Dr. A.P. (Pier) Siebesma	Geoscience and Remote Department (TU Delft)/ Royal Dutch Meteorological Institute (KNMI)
2nd corrector	Dr. M.A. (Marc) Schleiss	Geoscience and Remote Department (TU Delft)

Abstract

Rain and lightning associated with thunderstorms can cause damage to crops, cars, houses and other properties. As the climate is warming, it is expected that the frequency of lightning will increase, but it is highly uncertain to what degree. For the KNMI'23 climate scenarios, it becomes essential to make reliable statements on future thunderstorm scenarios. This research evaluates a new thunderstorm predictor, the LPI, proposed by Yaïr et al. (2010), that uses the vertical velocity and presence of hydrometeors for thunderstorm prediction. Both are important for the formation of a thunderstorm. This study generates LPI results for the Netherlands using hourly output data from the operational weather prediction model HARMONIE. The results are compared to the results of other widely used indices, the Modified Jefferson index, the K-index, the MUCAPE (Most Unstable Convective Available Potential Energy) and evaluated with historical observations. The period over which the study is done is the 15th of April to the 15th of October of 2016 and 2017. The different indices are compared using case studies and objectively by deriving logistic regression equations using the indices as a single predictor. The resulting probabilistic forecasts are compared using the Brier Skill Score and reliability plots. Besides, scatter plots have been made to study the relationship between the indices and the number of discharges.

The LPI as a thunderstorm predictor was found to perform worse than the other indices. The case study showed that the instability of the atmosphere is modelled correctly by HARMONIE but that precipitation and vertical velocity are often modelled at the wrong place by the HARMONIE model. The LPI output is strongly related to this model output of the precipitation and vertical velocity. Therefore, the LPI is wrong if the model output is wrong. Furthermore, the case study showed that the LPI gives a very local, discrete result, which makes it prone to uncertainties. The other indices have a much smoother and continuous result. This smoothness makes them less prone to, for example, phase errors. Although this result was only seen in the case study, it is very well possible it is true for more cases. Therefore, apart from improving the LPI formulation itself, a future study could use the LPI as an extra potential predictor to improve probabilistic forecasts of thunderstorms.

Preface

This nine-month project was filled with ups and downs. Moments of happiness and breakthroughs alternated with moments of insecurity. Doing this project helped me by learning to program, presenting and clearly expressing myself. During my Master's studies, I already learned a lot about these as well as about atmospheric processes, but using these skills in an actual project has helped develop the skills further.

This project would not exist if others did not help. I first want to thank Wim, Pier and Maurice for their cheerful, motivating and kind words when I was overcome by stress and insecurity. I would also like to thank Wim for his patience every time I asked a question and answering it extensively, for thinking with me when the programming did not work or when I was lost in the data and for being there during the whole process. I would also like to thank Pier for bringing me in contact with the KNMI. It was always a dream to work with people who feel the same passion for the weather as I do, for the feedback and supervision of my Thesis's grammar and theoretical background. I would like to thank Maurice for his patience when answering my questions, pointing out when some plot was peculiar, and for the introduction in verification. I liked it more than I thought. I would like to thank Marc for his fresh eyes in the project and solid feedback at important moments which helped bring my thesis to a higher level.

Furthermore, I want to thank some people of the KNMI. Without them, it would have been much harder to have finished my thesis on time, and they saved me much stress with their patience and tips.
Jelle Assink, thank you a lot for helping me open the GRIB-files and understand how to process them.
Jos de Kloe, many thanks for looking at my Python code and changing the run time from 1.5 hours to 1.2 seconds. The code was suddenly workable. And for all other data/Python/workstation related cases you helped with.
Rob van Schaik, you saved my summer. When I could not get the data from the MOS, you solved the problem, even though you were on holiday. I am really grateful.
Ruoyi Cui, who is not from the KNMI, but a PhD at ETH Zürich, gave insight into the ins and outs of the LPI implemented in COSMO. Thank you!
Sander Tijm, for answering questions on how to use the data inside the GRIB-files and the definition of some variables.
Kiri Whan, thank you for helping open the reforecasted indices file in a readable way.
Toon Moene, for the support, when there were problems with the MOS and searching for a solution.
Maarten Sneep for sending tips and tricks to increase the speed of my Python code despite his lack of time.
Sebastian Contreras, for the first introduction to the digital world of the KNMI and the pleasant walks we had during this period. I want to thank Rudolf van Westrhenen, Bert van Ulf and Bram van 't Veen for their help opening and processing the first GRIB-file in Python.
Moreover, I want to thank all KNMI'ers for their general enthusiasm about the weather and their work.

Lastly, I would like to thank my family for their belief in me, even when I do not and for their support through thick and thin in the past, present and future. For dabbling through my thesis and give me feedback. My friends and my housemates gave me energy and a pep talk when I needed it and did fun things when we wanted to. For the great time I had as a student and for help me becoming the person I am now. I want to thank Pim, for his patience, for the valleys and peaks we went through together and for the time we already had and the time that will come. And finally, a big thanks to everyone who was there for me during this time, even though I had some trouble staying in contact.

This project about thunderstorms and lightning feels like a perfect ending to a really great and educational time as a student, where I grew from youth to adult. I am grateful to end this great time with a project related to the weather, a subject I have been passionate about for 16 years. The next step is a time filled with new challenges. Challenges I look forward to with much joy and hope to use my qualities and grow.

List of Figures

Table of contents

1 Introduction

Thunderstorms, often accompanied by hail and strong wind gusts, cause damage to crops, cars, houses and other properties. Thunderstorms disrupt society, cause millions of euros in damages each year in the Netherlands, and are a significant liability for insurance companies. Lightning is a characteristic of this severe weather and a natural hazard by itself. It injures humans and damages nature and human-made structures. A recent example is a thunderstorm on the 18th of June, 2021 in Brabant nearby Eindhoven. Due to the wind gusts, many trees were uprooted and fell on houses and cars. Due to the heavy rain, the capacity of the sewage system was insufficient, which resulted in flooding. It caused a lot of nuisance (HartvanNederland, 2021).

As the climate is warming, it is expected that the weather will become more extreme, and the frequency and intensity of thunderstorms (as well as hailstorms) in the Netherlands will increase (Klein Tank et al., 2015). However, it is highly uncertain to what degree. With the arrival of the KNMI'23 climate scenarios in mind, it becomes essential to make reliable statements on the future scenarios for hail- and thunderstorms and inform the public about the effect of climate change on the frequency and intensity of both.

Traditionally, statistical models for predicting thunderstorms were mainly based on precipitation and the thermal instability of the atmosphere (Schmeits et al., 2008). Such an approach can be improved as such approaches do not account for factors such as the existence of convective clouds, vertical updraft and the presence of hydrometeors (supercooled water, graupel, ice) (Siebesma et al., 2019). Therefore, it is essential to investigate other approaches for predicting thunderstorms that account for these physical mechanisms.

Understandably, the traditional indices that have been used for a long time, as weather models were not as advanced. However, the development of weather models continued, and for some years now, new numerical weather prediction (NWP-) models have been used. These models have a high resolution and can resolve deep convection, which is relevant for thunderstorms, explicitly. HARMONIE (used by the Dutch KNMI) and COSMO (used by the German Deutscher Wetterdienst) are examples of such new NWP models.

For some years, the Royal Netherlands Meteorological Institute (KNMI) has used HARMONIE, a storm-permitting model for numerical weather prediction (NWP). HARMONIE explicitly resolves deep convection (De Wit & Tijm, 2010), i.e. is not parameterized. Hydrometeors, such as cloud-water, cloud-ice, rain, graupel, and snow, are given as prognostic variables. HARMONIE makes it possible to use more accurate predictors for lightning than the ones only based on atmospheric instability and precipitation. An example of such an accurate predictor is the Lightning Potential Index (LPI) introduced by Yair et al. 2010, which uses the vertical updraft and the presence of hydrometeors for lightning prediction. The LPI uses the added value of the storm-resolving model HARMONIE and is expected to provide an improved estimate of thunderstorm frequency and intensity.
Recent studies done by Brisson et al. (2021) using the operational NWP model COSMO have demonstrated that using more sophisticated indicators like the LPI gives more accurate results and an improved prediction of the spatiotemporal distribution of lightning. Brisson et al. (2021) also used the LPI to investigate the effect of climate change on the lightning frequency. According to the LPI, there will be a 4.8% decrease in lightning flash rate by the end of the century, in contrast to traditional predictors, which predicted an increase of 17.4%.

It becomes apparent when looking at the result of Brisson et al. (2016) that we will not be able to make reliable statements about the thunderstorm frequency in the future climate, without an innovative step. The innovative step proposed here is combining a new predictor, the LPI (Siebesma et al., 2019) together with an NWP model such as HARMONIE. Therefore, it is necessary to dive deeper into the definition and performance of the LPI by using HARMONIE data and see how it will affect the thunderstorm predictions.

It is helpful to compare the LPI with indices currently used by KNMI to predict thunderstorms to see whether or not predictions improve by using the LPI. Such comparison is a known way of getting insight into the performance of different predictors. For example, Schmeits (2002) compared 19 different indices using binary and continuous scatter plots to show which predictors are most suitable for predicting thunderstorms. Schmeits (2002) concluded that a combination of two or more indices may be preferable when predicting thunderstorms. Different indices focus on different atmospheric characteristics, so a combination may result in a more complete picture of the atmosphere and, therefore, a more accurate prediction than when only one predictor is used. Such a combination of predictors was used in logistic regression equations by Schmeits et al. (2008). More recently, Groot (2019) used logistic regression and quantile regression forests to search for the optimal combination of predictors. The LPI is not investigated in Groot (2019). Even though a combination of different predictors may be desirable, the new Lightning Potential Index is not combined with other indices in this thesis, because the LPI is relatively new and the performance of LPI using data produced by HARMONIE is hitherto unknown, the performance of the LPI alone is first evaluated.

Accordingly, the scope of this thesis is to give a first evaluation of the performance of the LPI in the Netherlands by testing and evaluating the LPI offline using output diagnostics of HARMONIE cycle 40 and comparing it with historical observed lightning data. Furthermore, this paper gives a first insight into variables that affect the Lightning Potential Index. Further, this thesis compares the performance of the LPI to traditional predictors using statistical forecast tools and verification metrics to see the effect of using not only thermal instabilities as traditional predictors but also vertical updraft and the presence of hydrometeors being included in the LPI.

In this research, the following research question is answered:

Could the prediction of lightning in HARMONIE become more accurate by the use of a new predictor, the Lightning Potential Index?

Including the following sub-question:
To what extent does the use of the Lightning Potential Index have an added value for climate predictions?

To answer these research questions, the Lightning Potential Index will be compared with traditional predictors. The predictors used for the comparison are briefly stated here and discussed in-depth in section 3. The first two indices are the Modified Jefferson index and the K-index, both based on vertical profiles of atmospheric temperatures and dew point temperatures but using a different formulation. The third predictor is the MUCAPE, which is related to the instability of the atmosphere. These three indices previously gave the best result for the prediction of thunderstorm occurrence, and lightning intensity (Groot, 2019).

In this thesis, the three predictors above and the LPI are evaluated for the Netherlands, as this area is

the main focus of KNMI and, therefore, of this research. To answer the different research questions, different domains are used. In the case of the general research question about LPI as a predictor for short-term weather forecasting, it is essential to forecast where lightning might occur. Therefore, in this paper, the Netherlands is divided into 20 subdomains, each of 5625 km^2, which cover the whole country, plus parts of Germany and Belgium. The subdomains are pooled into land, shore and sea, based on their dominant land type. In this way, the effect of water and land surface cover can be evaluated.
For climate runs, the specific location of lightning is not relevant, but frequency distribution is relevant. Therefore, in that case, the whole of the Netherlands is used without dvision into subdomains.
It is crucial to have enough data on thunderstorms to get a good result of the evaluations of the predictors. Therefore, in this research, historical data about the Netherlands thunderstorm season (April through October) of 2016 and 2017 are used. Most thunderstorms occur during that time of the year.

In this study logistic regression equations are derived using the different indices as single predictors. The resulting probabilistic forecasts are compared using two different verification metrics. The first metric is the Brier Score (BS, (Wilks, 2011)) that is calculated for all forecasts using the maximum or mean value of an index, for all 20 subdomains (or over the whole domain), and for all different lead times, separately. The use of separate leadtimes is to evaluate the effect of a shorter or longer lead time on the performance of the LPI. The Brier Score is translated into a Brier Skill Score (BSS) with the mean LPI as a reference, so that the performance of the other indices can be easily compared to that of the LPI.
The second method is using a reliability plot (Wilks, 2011; Jiang, 2020), which gives more information than the BSS. For the reliability plot, different lead times and all subdomains are used.
In the third method, scatter plots (as in Schmeits, 2002) that plot the number of observed discharges against the mean predicted index value are used to investigate the relation between the number of discharges and the different indices.

Furthermore, two case studies are considered to get a deeper insight into the performance of the LPI and the other indices. The case studies show the details of the forecasted indices, giving a better understanding of the differences between the indices and their dependence on the model.

The structure of this thesis is as follows:
In section 2, a brief theoretical background is given. Section 3 describes the methods that were used to give the results. In section 4, the two case studies are discussed. In section 5, logistic regression examples and results of the verification are shown. In section 6, the results are discussed. Section 7 presents conclusions and recommendations.

2 Theoretical background

In this section, information on the physical processes behind thunderstorms and lightning is provided. Thereafter, HARMONIE is discussed in-depth. Lastly, a theoretical introduction of the Lightning Potential Index follows.

2.1 Physics behind Thunderstorms

In the atmosphere, different kinds of clouds can form, based on the condition of the atmosphere. Cumulonimbus clouds, or thunderclouds, are one kind of cloud that can form and are the only cloud type that produces lightning. The formation of a cloud is the result of cooling air parcels due to a rising motion (Siebesma & De Roode, 2020; Poelman, 2010; Rakov & Uman, 2003). Three mechanisms could be responsible for rising motion (Siebesma & De Roode, 2020):

1. Large scale lifting; Two air masses with different conditions in temperature or moisture meet each other and form fronts

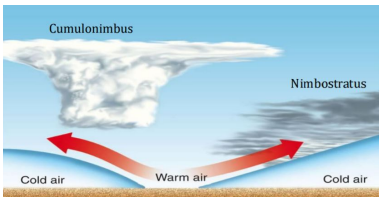


Figure 1: A schematic representation of the front lifting rising mechanism (National Weather Service, 2007).

2. Orography; Mountains force the air mass to rise

Figure 2: A schematic representation of the orographic rising mechanism (Canada Avalanche, 2021).

3. Convection; Air movement induced by temperature differences in atmospheric layers: cold air is denser and moves downward, while warm air is less dense and rises. It is the most common mechanism. Thunderclouds are formed by the convective rising mechanism.

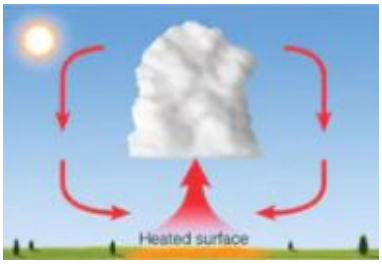


Figure 3: A schematic representation of the convection rising mechanism (Canada Avalanche, 2021).

In an unstable atmosphere, the temperature of the environment (T_V) decreases rapidly with altitude (Arasa, 2010). Surface air parcels may start to accelerate upward (Arasa, 2010; Siebesma & De Roode, 2020) due to this instability. A parcel will rise until the level of neutral buoyancy (LNB), at which the parcel and the environment have the same virtual potential temperature (De Roode, 2020; Arasa, 2010). Above this level, a stable layer is present. Since such a parcel has a positive (upward) vertical velocity, it overshoots the LNB into the stable layer (De Roode, 2020). In the stable layer, the parcel gradually decelerates (De Roode, 2020). Above the LNB the parcel adiabatically expands and its the temperature decreases (Poelman, 2010; Rakov & Uman, 2003). As the parcel becomes colder, its water carrying capacity decreases. When the temperature of the parcel reaches its dew point temperature, which depends on the relative humidity of the parcel, the saturation point is reached (Poelman, 2010; Rakov & Uman, 2003; De Roode, 2020). Before this saturation point, the parcel is unsaturated, and the relative humidity is less than 100%. After the saturation point is exceeded, the parcel becomes saturated, having a relative humidity of 100 %, and water starts to condense (Arasa, 2010) on cloud condensation nuclei (CCN). This point is called the lifting condensation level (LCL) and forms the cloud base (De Roode, 2020). As many of these small water droplets together make a visible cloud (Poelman, 2010; Rakov & Uman, 2003). During condensation, latent heat is released. This warms the parcel. The rate of adiabatic cooling of the parcel is reduced so that the temperature difference between the parcel and the environment stays large. This increases buoyancy.
Above the LCL, the saturated air parcel follows the moist-adiabatic lapse rate (De Roode, 2020). Here, the parcel follows one of two options. First, if the vertical temperature gradient of the environment is larger than the moist adiabatic lapse rate, the parcel remains negatively buoyant with respect to the environment and the vertical motion is damped (De Roode, 2020). In such condition, the height of such clouds is limited. The amount of negatively buoyant energy that suppresses the upward motion is indicated by CIN (convective inhibition) and defined by the formula:

$$CIN = \int_{LNB}^{LFC} g \frac{T_{v,p} - T_{v,env}}{T_{v,env}} dz \tag{1}$$

where LNB is the level of neutral buoyancy, LFC is the level of free convection, i.e. where $T_{v,p} = T_{v,env}$, g is 9.81 [m/s^2], $T_{v,p}$ is the virtual temperature of the parcel and $T_{v,env}$ is the virtual temperature of the environment.
Second, if the vertical temperature gradient of the environment is less than the moist-adiabatic lapse rate, the vertical velocity of the parcel is not damped, and the parcels (the cloud) reaches the level of free convection (LFC) (De Roode, 2020). At the LFC, the buoyancy of the parcel is neutral with respect to the environment. Above this level, the parcel becomes positively buoyant, and the vertical upward motion starts again (De Roode, 2020). The parcel rises until the limit of free convection (LOC) is reached. Due to the positive vertical velocity, the parcel overshoots the LOC and therefore, the cloud top is often found above this level (De Roode, 2020). The positively buoyant energy that air parcels experience is indicated by CAPE and defined by:

$$CAPE = \int_{LFC}^{LOC} g \frac{T_{v,p} - T_{v,env}}{T_{v,env}} dz \tag{2}$$

where LOC is the limit of free convection, i.e. the maximum level where $T_{v,p} = T_{v,env}$, LFC is the level of free convection, i.e. where $T_{v,p} = T_{v,env}$ below the CAPE layer, g is 9.81 [m/s^2], $T_{v,p}$ is the virtual temperature of the parcel and $T_{v,env}$ is the virtual temperature of the environment.
The process of the rising parcel for the second case is represented schematically in Figure 4 showing how a cumulus cloud is formed. If there is enough atmospheric instability, moisture and lifting, a strong vertical

updraft will let the cumulus cloud grow into a thundercloud.

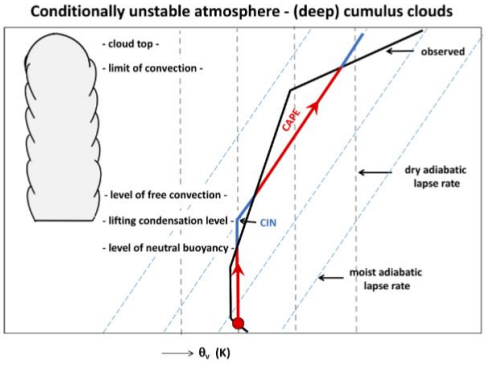


Figure 4: A schematic representation of the path that an air parcel follows during ascent when the vertical gradient of the atmosphere is less than the moist adiabatic lapse rate. The virtual potential temperature is shown on the x-axis. The virtual potential temperature is defined by: $\theta = T + \frac{c_p}{g} z$. The grey dashed lines are the profiles of dry-adiabatic parcels. The blue dashed lines are the profiles of moist-adiabatic parcels. The solid black line is the temperature profile of the environment. Below the cloud base, parcels follow the dry-adiabatic lapse rate. Above the cloud base, the virtual potential temperature increases due to latent heat release (De Roode, 2020).

As stated above, thunderclouds are the only clouds that produce lightning. This lightning production is the result of different processes in the cloud. The next part of this section focuses on the origin of lightning inside a cumulonimbus cloud.

In the troposphere, the temperature decreases with increasing height at about 6.5 °C per 1000 m (Poelman, 2010). With a height that varies from 10 km to 18 km over the earth (KNMI, 2021; Rakov & Uman, 2003), there is a point in the atmosphere where a rising parcel reaches a temperature below 0 °C. At this point, some water droplets will freeze, while others remain liquid, called supercooled particles (Poelman, 2010; Rakov & Uman, 2003). A mixed-phase region forms due to the co-existence of supercooled water particles and ice particles, in this mixed-phase region, graupel forms. It forms when ice particles collide with supercooled water particles (Palmer, 2017; Greitemann, 2014). The supercooled water particles freeze instantly (Greitemann, 2014) on the hail particles and form a layer of ice particles covering the hail particles (Palmer, 2017). During the freezing, the supercooled water droplets release latent heat to the hail particle, so that the temperature of the hail particle increases (Greitemann, 2014). The result is a soft hail particle (Greitemann, 2014) with a lower density (Palmer, 2017), this is known as graupel (Greitemann, 2014).

The upper limit of this mixed-phase layer is the -40 °C isotherm. It was found that above this temperature, all particles will freeze and become ice (Poelman, 2010; Rakov & Uman, 2003). It is in this mixed-phase layer that most electrification of thunderstorms takes place (Poelman, 2010; Rakov Uman, 2003; Deierling et al., 2008).

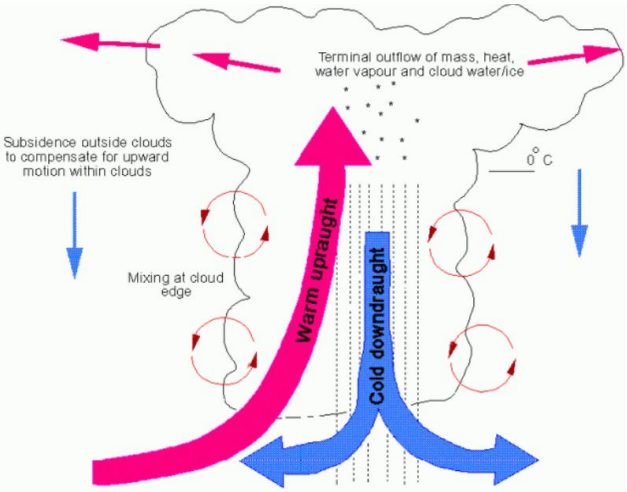


Figure 5: A schematic representation of the forming of a thunderstorm. The updraft takes a warm, moist air parcel to higher levels in the atmosphere. The parcel cools down due to adiabatic expansion to below its dew point temperature forming water droplets. The water droplets constitute a cloud. Above the 0 °C isotherm, some, but not all, of the water droplets freeze. A downdraft is present to compensate for the updraft (Roose, 2015).

The mechanism that dominates thunderstorm electrification is non-conductive charging (Brisson et al., 2021; Yair et al., 2010; Saunders, 1993; Latham et al., 1978), which involves the collision between graupel and small cloud-ice particles as well as deep convection (Poelman, 2010; Yair et al., 2010).

First, the most dominant mechanism is electric charge separation through the collision of graupel and small cloud-ice particles (Poelman, 2010; Brisson et al., 2021). Here significant charging occurs because graupel, ice crystals, and supercooled water droplets co-exist (Deierling et al., 2008). Graupel takes the electron of the smaller ice crystals while descending (Saunders, 1993). As a result, the graupel particles become negatively charged, and the ice particles become positively charged (Poelman, 2010). The lighter, positively charged ice crystals ascend, while the heavier negatively-charged graupel particles descend due to gravity (Poelman, 2010; Saunders, 1993; Brisson et al., 2021). On a macroscale, the cloud top is more positive, while the cloud base is more negative. Field and laboratory evidence points out the importance of interactions between particles in the ice phase in the presence of supercooled liquid water droplets, for separating electric charge in thunderstorms (Saunders, 1993; Yair et al., 2021; Deierling et al., 2008).

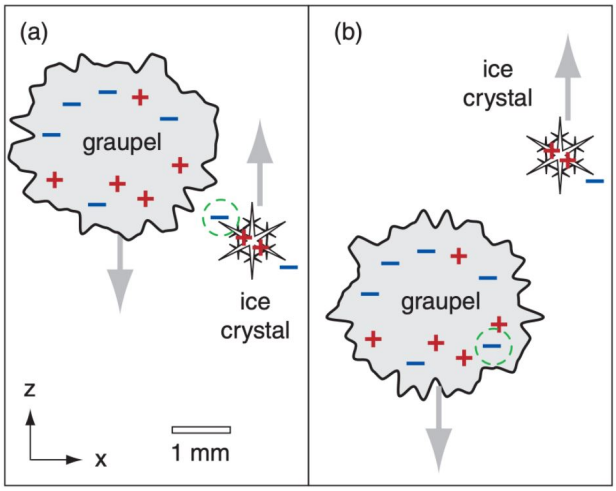


Figure 6: An illustration of the process of charge transfer between graupel and small cloud-ice particles. The green dashed circled electron is transferred. The downward-moving graupel takes an electron from the rising ice particle. As a result, graupel is negatively-charged, while the ice particle is positively-charge (Stull, 2020).

The second mechanism is deep convection. Positively-charged particles near the ground surface are brought to the top of the cumulus by vertical updrafts (Poelman, 2010). Negatively-charged particles produced by cosmic rays are attracted and attached at the positive cloud boundary (Poelman, 2010). The downdrafts outside the cloud, resulting from convective circulation, are thought to carry the negative charges down to the cloud's base (Poelman, 2010). Again this results in a positive charge at the top and a negative charge at the bottom of the cloud. This convection mechanism is less critical for cloud electrification. The charge between the positive earth surface and the negative bottom of a cloud base has a magnitude comparable to a single lightning discharge. However, thunderstorms often produce several lightning discharges and not only one (Poelman, 2010). However, convection is an essential condition that affects the collisions in the charging zone (0 °C to -40 °C) (Yair et al., 2010). Many field observations suggest that a strong vertical updraft in the mixed-phase region is necessary to produce lightning (Deierling et al., 2008). The strong vertical updraft lift water droplets to above the 0 °C isotherm, such that the mixed-phase layer is replenished with supercooled particles. The strong vertical updraft should lift graupel to above the reversal zone (-15 [°C]) (Yair et al., 2010), below which it takes negatively charged particles (Saunders, 1993; Poelman, 2010), and it should disperse the positively charged ice particles upward. Further, a larger vertical updraft volume produces more hydrometeors in the charging zone, leading to more collisions between graupel and ice crystals. As a result, a more significant charge separation will be present, and lightning activity increases (Deierling et al., 2008).

Although there is some clarity about the physics of thunderstorms, uncertainties about details on the formation of thunderstorms are still present. For example how the water/ice collisions occur (Poelman, 2010) and how charge is transferred between the two interacting particles (Saunders, 1993). Despite these uncertainties, the literature seems to agree on the fact that the occurrence of thunderstorms and lightning discharges is related to deep convection, with a strong vertical updraft in the mixed-phase layer and the co-existence of solid and liquid hydrometeors, such as graupel, ice, snow and water droplets. All these effects are captured in the LPI.

2.2 HARMONIE

Weather models are computer calculations that describe the state of the atmosphere. Using recent data of the atmosphere, such as temperature, pressure, and wind, the model calculates what is expected to happen based on physical laws (KNMI, n.d.-b). This model calculation results in a weather prediction.
Since 1993 KNMI had used the HIRLAM model for the short-term weather forecasting (48 hours), but more recently has used a new model named HARMONIE (KNMI, 2020, n.d.-b). HARMONIE was developed by the ALADIN (Aire Limiteé Adaption Dynamique Développement International), and HIRLAM (High-Resolution Limited Area Model) consortia, which together hold 26 countries from north and south Europe (Bengtsson et al., 2017; KNMI, 2020). The prime objectives of the consortia is 'to provide all members with a numerical weather prediction (NWP) model for short and very short-range forecasting including nowcasting for both research and development activities and operational usage' (Malcorps & Agren, 2005). One main focus of the consortia is adapting the French weather model AROME in such a way that it can be used in the standard ALADIN-HIRLAM NWP system to make it accessible to all the countries (Bengtsson et al., 2017). The implementation and the optimisation of the AROME model for both the northern and the southern European conditions led to improvements to the model's physical parameterisations (Bengtsson et al., 2017). As a result, the existing biases were reduced, and the physical description of clouds and land surface fluxes improved (Bengtsson et al., 2017). The model configuration with the updated physical parameterisations is called HARMONIE-AROME, from hereon HARMONIE, and is different from the French model definition AROME (Bengtsson et al., 2017).

As noted above, for some years KNMI has used HARMONIE instead of HIRLAM. HARMONIE brings new and vital improvements concerning general weather forecasting and specifically thunderstorm predictions. One of the differences is in the resolution of both models. The resolution of the HIRLAM model is 5 kilometres to 10 kilometres, which means that the physical parameters simulated by the model represent a large grid box volume (KNMI, 2020).
Furthermore, HIRLAM is a hydrostatic model, which uses a hydrostatic approximation for simulations. It uses the momentum equations, but vertical velocities are neglected. Therefore, it cannot capture features on the vertical length scales, which becomes valid for thunderstorm up- and downdrafts.

HARMONIE, on the other hand, has a resolution of 2.5 kilometres (KNMI, 2020), which makes it much more suitable for the description and the prediction of small scale processes. HARMONIE is a non-hydrostatic model (Van den Brink, 2018; Bengtsson et al., 2017), of which the core is based on the fully compressible Euler equations (Bengtsson et al., 2017). The evolution of the equations inside the model is based on the semi-Langrangian advection scheme (Bengtsson et al., 2017; Gregow et al., 2020). The horizontal diffusion is based on semi-Langrangian horizontal diffusion (Bengtsson et al., 2017). The lateral boundary conditions are given by the ECMWF model (Bengtsson et al., 2017).

The HARMONIE model consists of 65 vertical levels, with the top (level = 0) at 10 [hPa] and the lowest level (level = 64) at 12 [m] above the ground (Bengtsson et al., 2017). The levels are not all of the same thickness. Closer to the surface, the model levels are thinner and have smaller mutual distances, which increases with height. The lowest levels follow the terrain, and the upper levels follow the pressure levels. The model consists of half levels and full levels. For the full levels, temperature, moisture, and wind are computed. For the half levels, the fluxes are computed (Bengtsson et al., 2017). Figure 7 shows an example of the half and full levels in the HARMONIE model. This figure shows a model that consists of 60 levels. This is different from the 65-levels of the HARMONIE model.

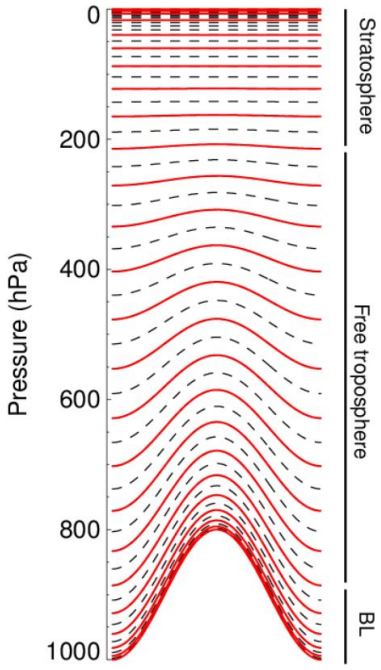


Figure 7: Example of full (solid, red) and half (striped, black) levels in the operational ECMWF model. The HARMONIE model looks the same but has 65 levels (Krol et al., 2005).

The physics in the HARMONIE model consists of many different aspects, such as radiation, cloud microphysics and turbulence. These physical processes are combined in an AROME package (Ivanov et al., 2021). Different schemes parameterize these processes (Gregow et al., 2020). Below an overview of the HARMONIE model physics is given. First, zoom in on the radiation used in the HARMONIE model. As a default for shortwave radiation parameterisation, the Morcrette radiation scheme from ECMWF is used, which contains six spectral intervals ranging from 0.185 to 4.00 [m] (Bengtsson et al., 2017; Gregow et al., 2020). For longwave radiation, the default settings are based on the rapid radiative transfer model of Mlawer et al. (1997) (Bengtsson et al., 2017). These settings contain 16 spectral bands ranging between 3.33 and 1000 [m] (Bengtsson et al., 2017; Gregow et al., 2020). The diagnostic output from the radiation parameterisations includes different factors, namely, the accumulated spectrally averaged downwelling shortwave globally, direct and indirect normal irradiances at the surface and the net shortwave and longwave radiative fluxes at the top of the atmosphere, at the surface and on each model level (Bengtsson et al., 2017).

The cloud microphysics scheme used in HARMONIE is a one-moment bulk scheme (Bengtsson et al., 2017; Ivanov et al., 2021). It uses a three-class ice parameterisation, which contains the following solid hydrometeors as prognostic variables: cloud ice, snow, and a combination of graupel and hail (Bengtsson et al., 2017; Ivanov, Michaelides, Ruban, Charalambous, & Tymvios, 2021). The liquid hydrometeors that are present as prognostic variables in the microphysics scheme are water vapour, cloud liquid water, and rain (Bengtsson et al., 2017; Ivanov et al., 2021). All hydrometeors are advected horizontally by a semi-Langrangian scheme and vertically by a sedimentation scheme (Bouteloup et al., 2011).
In the HARMONIE model, the cloud fraction is not a prognostic variable. It is modelled using a cloud and condensation scheme (Gregow et al., 2020; Bengtsson et al., 2017; Ivanov et al., 2021).

The HARMONIE model uses the momentum equations for simulating the vertical wind speed and calculating vertical movements, (Van den Brink, 2018). This enables the model to resolve deep convection and explicitly represent the model's non-hydrostatic dynamics, i.e. there is no parameterisation of deep convection (Bengtsson et al., 2017).
Shallow convection, however, needs to be parameterised (Bengtsson et al., 2017). In HARMONIE, a dual mass-flux framework is used for this, which consists of two updrafts that transport heat, moisture, and momentum (Bengtsson et al., 2017; Ivanov et al., 2021). One of the two updrafts is an unsaturated updraft that never reaches the lifting condensation level, and the other is a saturated updraft that condenses and becomes a cloud (Bengtsson et al., 2017). The convective transport is proportional to the difference between the updraft properties and the environment, times the amount of mass transported by the updraft (Bengtsson et al., 2017; Ivanov et al., 2021). The upward mass flux can be described by an eddy diffusivity mass-flux framework (Ivanov et al., 2021; Bengtsson et al., 2017).

The eddy-diffusivity, i.e. the turbulence part, is given by the turbulence scheme HARATU (Bengtsson et al., 2017), which initially was developed for the regional climate model RACMO (Bengtsson et al., 2017). It is based on a prognostic equation for turbulent kinetic energy (TKE) equation combined with a diagnostic mixing length scale (Bengtsson et al., 2017; Ivanov et al., 2021). The mixing length scales are different for heat, and momentum (Bengtsson et al., 2017).

The mass-flux component describes transport by updrafts that are saturated as well as unsaturated (Bengtsson et al., 2017). The mass-flux scheme starts with the initialisation of the excess of the updrafts, whereafter a vertical velocity equation (De Rooy & Siebesma, 2010) is used to determine the updraft vertical velocity and the height to which the updraft can come (Bengtsson et al., 2017). This gives the cloud-base and -top height, and using these heights, the profiles of entrainment rate could be defined as it is a function of these heights (Bengtsson et al., 2017). For the unsaturated updraft, the entrainment formulation of Siebesma et al. (2007) is used. This formulation is based on large eddy simulation results for the dry convective boundary layer (Bengtsson et al., 2017). In the case of the saturated updraft, the entrainment profile consists of two parts. The first part goes from the surface to the cloud base and has the formulation for unsaturated updrafts (Bengtsson et al., 2017). The second part goes past the cloud base and represents larger, faster-rising thermals that are associated with lower entrainment rates in comparison to the unsaturated updrafts (Bengtsson et al., 2017). The value for entrainment at the cloud base in saturated updrafts are scaled with the inverse of the cloud base height, which is suggested by LES (Bengtsson et al., 2017). This relation shows that higher vertical velocities and larger thermals are related to deeper mixed layers (De Rooy & Siebesma, 2010). Furthermore, the considerable variation of detrainment and the strong relation with mass flux is considered in the model (Bengtsson et al., 2017). In the model, environmental conditions influence both the cloud layer depth and the detrainment (Bengtsson et al., 2017). The relation represents that if relative humidity and buoyancy excess of the updraft increases, detrainment decreases and the mass flux decrease with altitude. This is again related to large clouds with high vertical updraft velocities (Bengtsson et al., 2017).

The surface physics is described by an external surface scheme called SURFEX (Ivanov et al., 2021), which was developed by Météo-France (Bengtsson et al., 2017). It consists of a set of physical models for different surface types (Ivanov et al., 2021; Bengtsson et al., 2017), such as natural land surfaces, sea and inland water bodies, and urban environments (Ivanov et al., 2021; Bengtsson et al., 2017). SURFEX is a model on its own and can be used in standalone or in combination with an atmospheric model (Bengtsson et al., 2017).

Four surface tiles represent each grid box: seas, inland water bodies, urban area and nature (Bengtsson et al., 2017; Ivanov et al., 2021). Each surface tile is modelled with a specific surface model, and the total flux of the grid box results from the addition of the individual fluxes weighted by their respective fraction (Bengtsson et al., 2017). The default topography is based on Global Multi-resolution Terrain Elevation Data (Bengtsson et al., 2017).

As became clear in this section, many parameterisation schemes are used to solve and describe the atmospheric processes in HARMONIE. The high resolution and the non-hydrostatic properties of HARMONIE give a more realistic description of the atmospherical processes (Van den Brink, 2018) compared to HIRLAM. Furthermore, HARMONIE not only resolves deep convection explicitly, but it also uses a cloud microphysics scheme which contains the following prognostic variables: cloud ice, snow, graupel, hail, water vapour, cloud liquid water, and rain (Bengtsson et al., 2017). It simulates microphysical mixed-phase precipitation formation processes. The presence and amount of these hydrometeors are essential indicators of thunderstorms.

It is apparent that HARMONIE could improve the prediction of thunderstorms due to the high resolution, the explicit simulation of deep convection at a small scale, and hydrometeors' presence as prognostic variables.

2.3 Lightning Potential Index

The connection between cloud microphysics and thunderstorms and lighting is apparent. Still, commonly used thunderstorms and lightning potential forecasting indices mostly rely on the stability and thermodynamical properties of the atmosphere (Yair et al., 2010). They may be not related to the electrification process (Brisson et al., 2021). In this respect, the Lightning Potential Index proposed by Lynn and Yair (2010) is a new and innovative thunderstorm predictor, as it focuses on the microphysics of the clouds and it accounts for the presence of both solid and liquid hydrometeors (Brisson, Blahak, Lucas-Picher, Purr, & Ahrens, 2021), inherent to the origin of lightning and thunderstorms. As such it could use the thunderstorm-predicting qualities of HARMONIE.

The LPI is a measure of the kinetic energy of an updraft in a developing thundercloud, scaled by the potential for charge separation based on ratios of solid and liquid hydrometeors in the main charging zone of the cloud (Yair & Lynn, 2008; Yair et al., 2010). The LPI portrays the potential for charge generation and separation that leads to lightning flashes in developing convective thunderclouds based on model outputs of vertical updraft and mass mixing ratios of hydrometeors (Fiori et al., 2016; Yair & Lynn, 2008) in a relatively simple and computationally cheap way (Brisson et al., 2021). It leads to the ability to determine areas prone to lightning (Sokol & Minářová, 2020), but it does not calculate the resulting electric field or its evolution (Yair & Lynn, 2008; Yair et al., 2010).

The original LPI definition by Yair and Lynn (2010) is derived from the vertical updraft and the mass mixing ratios of liquid water, cloud ice, snow and graupel (Yair et al., 2010), integrated from the freezing level to the -20 [°C] isotherms. The definition is the following.

$$LPI = \frac{1}{V} \int \int \int \epsilon w^2 dx dy dz \tag{3}$$

It has its most significant values in the presence of strong vertical updraft when liquid and solid hydrometeors coexist (Fiori et al., 2016). The approach of this derivation is in line with the theory showing the critical relation of the vertical updraft, hydrometeors and lightning (Lagasio et al., 2017). Brisson et

al. (2021) changed the original definition by adding a filter for large-amplitude gravity waves caused by orographic features.

Others already researched the performance of the LPI, and the obtained results were promising. The Mediterranean case studies done by Yair et al. (2010) and Lynn and Yair (2010) showed 'superiority of the LPI against other thermodynamic indices'. Brisson et al. (2021) changed the definition of the LPI by adding the CAPE, which resulted in the following formulation.

$$LPI = f_{w,max} f_{CAPE} \frac{1}{H_{-20^\circ C} - H_{0^\circ C}} \int_{H_{0^\circ C}}^{H_{-20^\circ C}} \epsilon g_(w) dz \tag{4}$$

Brisson at al. (2021) showed that the LPI is much better at deriving a correct representation of temporal probability densities and latitudinal dependencies in the COMSO-model in comparison with another index.(Brisson et al., 2021). In addition, another research done in the COSMO model done by Sokol et al. (2020) confirms that the LPI is a suitable tool for implicit forecasting of lightning. Lagasio et al. (2017) concluded that the LPI shows a high correlation with the observed lightning and is helpful for real-time forecasting, if a discrimination is made between cases leading to deep moist convective and persistent storms and shallower and chaotic convective situations, lacking significant ground effects.
Furthermore, it could be helpful as a hindcast tool for more in-depth investigations into the physical parameters that form lightning activity. In addition the LPI could reproduce essential parameters responsible for high-impact events over complex topography areas (Lagasio et al., 2017). In another research, the LPI is considered as a good forecasting tool whenever the model can reproduce the convective process (Fiori et al., 2016) as the main problem seems to be that the model has difficulty in correctly reproducing the convective field characterising the event. In this research (Fiori et al., 2016), the effect of orographic up- and downdrafts on the LPI becomes clear and shows that the LPI will result in a non-homogeneous signal.

To conclude this section, the LPI uses vertical updraft and the mass mixing ratios of liquid and solid hydrometeors, which are the most important conditions of the initation of lightning and thunderstorms. Furthermore, the LPI takes advantage of one of the main features of HARMONIE, the explicitly resolved deep convection. In addition, the results obtained by previously-done research shows the LPI's outperforms traditional indices. Therefore, it is expected that the LPI used together with HARMONIE will give a considerable improvement on the prediction of the probability of thunderstorm occurrence in the Netherlands compared to the currently used indices.

3 Methods

3.1 Model Setup

3.1.1 HARMONIE

All data used in this research is instantaneous hourly output data of HARMONIE cycle 40. HARMONIE is a non-hydrostatic, area limited, storm permitting model and a terrain-following pressure-based system (Bengtsson et al., 2017). It contains 65 levels in the vertical, with the lowest at 12 metres and the top at circa ten [hPa].
The lateral boundary conditions are given by the ECMWF model (Bengtsson et al., 2017). Surface physics are simulated by the surface scheme named SURFEX, developed by Météo France (Masson, 2013). The cloud microphysics scheme used in HARMONIE is a one-moment bulk scheme containing prognostic variables for hydrometeors (Bengtsson et al., 2017). Deep convection is solved explicitly, i.e. it is not parameterized. The horizontal resolution is 2.5 kilometres, and the model time step is 75 seconds (Bengtsson et al., 2017; personal communication De Rooy, W., 2021). The area of the HARMONIE domain that is available for this study is shown in Figure 8. The computational area of HARMONIE is larger than the area used in this study.

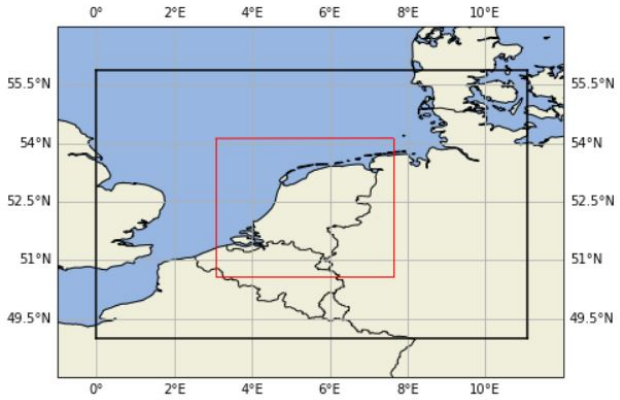


Figure 8: The black rectangle shows the HARMONIE domain that is available for this study. The red rectangle inside is the area of interest for this research.

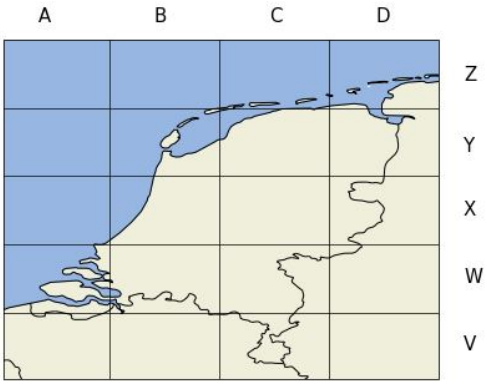


Figure 9: Map of the Netherlands and surroundings divided into 20 subdomains. This area is the area of interest for this research.

3.1.2 Area and time of interest

Figure 9 shows the area used in this research. It covers the Netherlands and parts of the North Sea, Belgium and Germany. The reason for choosing only the Netherlands and not the whole domain of HARMONIE is that this is the area of interest for KNMI, and historical observations of this area are available. The area is divided into 20 subdomains, each 5625 [km^2], which is 30 x 30 times larger than the grid box size but still small enough to make forecasts for an area that is not too large. The subdomains are pooled into land, shore, and sea to get more data per group and insight into the effect of land cover on the LPI performance. The area division is as follows:

land: AV, BV, BW, CV, CW, CX, DV, DW, DX

shore: AW, BX, BY, CY, CZ, DY, DZ

sea: AX, AY, AZ, BZ

The periods used for the research are April 15 to October 15 for the years 2016 and 2017. These periods correspond to the thunderstorm season in the Netherlands (KNMI, n.d.-a, 2016). The years 2016 and 2017 are chosen because observations and reforcasted indices are already available in the KNMI database. Furthermore, 2016 was a thunderstorm-rich year (Météorage, n.d.; Groot, 2019) and the 2017 KLDN data showed to have many lightning detections (Groot, 2019).

3.2 Lightning parameterizations

3.2.1 Lightning Potential Index

The lightning parameterisation used in this research is based on the LPI formulation of Yair et al. (2010). However, as some details in the formulation were unclear, the parameterisation proposed by Brisson et al. (2021) is also used. Output data of HARMONIE of two thunder seasons was used to evaluate the performance of the LPI. The LPI has the units of [J/kg] and is defined by:

$$LPI = f_1 \frac{1}{H_{-20} - H_0} \int_{H_0}^{H_{-20}} \epsilon w^2 dz \tag{5}$$

where w is the vertical updraft in [m/s]. The integral is taken over the primary charging zone in one grid point, between the 0 °C and -20 °C isotherm, as most effective charge separation by collisions of ice and graupel in the presence of supercooled water, occurs within this zone (Fiori et al., 2016). The units for height of these levels is metres above the earth surface. ϵ is a dimensionless number having a value between 0 and 1 related to relative amounts of liquid and solid hydrometeors, defined by the formula:

$$\epsilon = \frac{2\sqrt{R_L * R_F}}{R_L + R_F} \tag{6}$$

with R_L the total liquid mass mixing ratio [kg/kg], containing all liquid species and R_F is the ice fractional mixing ratio [kg/kg], containing all frozen hydrometeors. These are defined by the formulae:

$$R_L = r_c + r_r \tag{7}$$

$$R_F = \frac{r_g}{2}\left[\frac{2\sqrt{r_i r_g}}{r_i + r_g} + \frac{2\sqrt{r_s r_g}}{r_s + r_g}\right] \tag{8}$$

where r_c, r_r, r_g, r_i and r_s are the mass mixing ratios of cloud water, rain water, graupel, cloud ice and snow [kg/kg], respectively. The model output was specific humidity [kg/kg]. Before adding it to the formula, the specific humidity was converted to mass mixing ratios using $r = \frac{q}{1-q}$, where r is the mass mixing ratio [kg/kg] and q is the specific humidity [kg/kg].
The ϵ reaches its maximum value when the mass mixing ratios of the liquid and solid hydrometeors are equal. It shows the importance of the simultaneous existence of both hydrometeors for the initialisation of lightning. A minimum is reached when the hydrometeors do not exist simultaneously. No charge separation will occur so that thunderclouds are not electrified, so no lightning discharge can happen.

Based on the findings of Yair et al. (2010) and Brisson et al. (2021), a filter is added to correct for updrafts that do not represent physically coherent convective updrafts. Km-scale models, such as HARMONIE, do not fully resolve small scale convective updrafts, resulting in the modelling of isolated single-grid-column

updrafts. As a result, the LPI signal becomes weak, noisy and unrealistic as a result. To correct for this, the LPI for a particular grid point is only non-zero when a majority of the neighbouring horizontal cells within a 5 grid-radius (10 x 10 [km^2]) of that grid point have a vertical updraft above a specific threshold (Yair & Lynn, 2008; Brisson et al., 2021). For this research, the threshold is set to 0.5 m/s (Yair et al., 2010). The 5 grid-radius surroundings represent a 12.5 x 12.5 [km^2] area in HARMONIE. The reason is that because of the resolution of HARMONIE in combination with the demands of the LPI definition, a 10 [km] x 10 [km] was not possible. This correction filter is called f_1, and it is not related to a physical mechanism of charge separation in a thunderstorm. The correction filter resembles like:

$$f_1 = \begin{cases} 1 & \text{, if } C > 0.5 * TG \\ 0 & \text{, else} \end{cases}$$

TG is the total number of grid points in the 5 grid radius, and C is the number of grid points with a vertical updraft greater than the threshold in [m/s].

The comparison with observations is not made by converting the LPI into lightning flash rate, as was proposed by Brisson et al. (2021). Instead, a threshold value based on the spatial distribution of the maximum and mean value of the LPI is set, above which it is assumed lightning and thunderstorms occur.

3.2.2 Modified Jefferson Index

In this research, a comparison between the LPI and currently used indices is made. The first index that will be discussed is the Modified Jefferson Index, hereinfurther MJI. The MJI describes the instability between the 925 [hPa] and the 500 [hPa] levels and is defined by the formula:

$$MJI = 1.6\theta_{w,925} - T_{500} - 0.5DD_{700} \tag{9}$$

where $\theta_{w,925}$ is the wet-bulb potential temperature at 925 [hPa], T_{500} is the temperature at 500 [hPa], and DD_{700} is the dew point depression, which is the difference between the temperature and the dewpoint temperature at 700 [hPa].
The MJI used nowadays is a modification of the original Jefferson Index established by Jefferson in 1963. The modification includes an addition of the 700 [hPa] dewpoint depression, which increases with decreasing relative humidity (Haklander & Van Delden, 2013; Groot, 2019). This accounts for the fact that a rising air parcel can be diluted or cool due to entrainment of drier air (Haklander & Van Delden, 2013; Groot, 2019). The MJI does not give a direct relation between a parcel with latent heat release and its environment. The MJI does discriminate well between stable and unstable conditions, which lead to convective storms (Groot, 2019).

3.2.3 K-index

The second index used for comparison is the K-Index defined by the formula:

$$KI = T_{850} - T_{500} + T_{d,500} - DD_{700} \tag{10}$$

where T_{850} is the temperature at 850 [hPa] and $T_{d,500}$ is the dew point temperature at 500 [hPa].
The K-index is a measure of thunderstorm potential. A higher value indicates a higher thunderstorm probability. A higher value is reached if the stability between 850 [hPa] and 500 [hPa] decreases, if the moisture increases at 500 [hPa], and relative humidity increases at 700 [hPa].

3.2.4 Most Unstable CAPE

The third and last index is the MUCAPE or most unstable CAPE. MUCAPE is the CAPE (convective available potential energy), where the parcel used for initialisation is the one that gives the largest value of CAPE. The MUCAPE is found by lifting a parcel from every one of the lowest pressure levels in the troposphere. For every parcel, the CAPE will be calculated, and the parcel that produces the highest CAPE is the MUCAPE and is defined by the formula:

$$MUCAPE = \int_{LFC}^{LOC} g \frac{T_{v,p} - T_{v,env}}{T_{v,env}} dz \tag{11}$$

where LOC is the limit of free convection, i.e. the maximum level where $T_{v,p} = T_{v,env}$, LFC is the level of free convection, i.e. where $T_{v,p} = T_{v,env}$ below the CAPE layer, g is 9.81 [m/s^2], $T_{v,p}$ is the virtual temperature of the parcel and $T_{v,env}$ is the virtual temperature of the environment.
CAPE describes the instability of the atmosphere by measuring the amount of energy a parcel has gained after rising to the LNB. CAPE provides an approximation of updraft strength within a thunderstorm. The more potential energy an air parcel possesses, the higher the CAPE value is. A higher CAPE value means a more unstable atmosphere and a stronger updraft. A sufficiently strong updraft means a parcel can be lifted high enough in the atmosphere to produce a convective weather systems.

3.3 Validation Data set

A data set of observed number of lightning discharges is needed as a reference data set to compare the various indices and assess how well they perform. In this research, the reference data set consists of lightning discharges as detected by Météorage, the operator of the European lightning detection network headquartered in Pau, Cedex 9, France. Météorage has detected lighting since 1987 and since 2010 part of the data is delivered to KNMI. More recently, Météorage started to operationally deliver data to KNMI, which is referred to as KLDN. In the past, FLITS was used for the lightning detection (Groot, 2019; De Vos, 2015). The full period data set of KLDN contains some inhomogeneities due to the improvement and addition of extra sensors in the UK (autumn of 2015) and the Netherlands (2017, June) (Groot, 2019). Unfortunately, there is no homogeneous detection system so it is impossible to correct these inhomogeneities (Groot, 2019). The reference data sets used herein were gathered using KLDN. KLDN detects lightning using electromagnetic signals and can distinguish and classify cloud-to-ground (CG), and cloud-to-cloud (CC) discharges, but not unambiguously. KLDN can detect CC lightning, but such a detection can consist of different segments of the same discharge. For CG lightning, KLDN can distinguish between the first return stroke and the subsequent return strokes that originate from the same lightning event(De Vos, 2015). In this research both CC and CG discharges are used and no distinction is made between the two.
In KLDN some discharges are missed, but more often, extra detections from another source are counted as discharges. This is explained by the fact that KLDN measures electromagnetic signals to detect lightning discharges but there are other sources of electromagnetic signals such as a flying aeroplane. The lightning detection system could catch the signal of an airplane and register this as a discharge, while there was not any. An additional source of a false lightning strike detected by KLDN is a horizontal lightning which is not counted as a single event, but as at least two different events at different locations. This, too, increases the number of detected lightning discharges, while only one discharge was present. In this study, a threshold was used to decrease the probability of incorporating the non-lightning signals. An event of lightning discharges is only accepted as such if two or more detections were made in a specific time and area.

3.4 Data description

3.4.1 Reforecasted data used for LPI

The LPI is a new index, and therefore, unlike the other three indices, the LPI predictions are not available in the archive of KNMI. In this study, new LPI output is generated using instantaneous hourly reforecasted data of HARMONIE cycle 40 on a 2.5 [km] grid, using the analysis time 00 with the corresponding lead times +00, +01, ..., to +48.
The reason behind the analysis time of 00 UTC is twofold. Lightning and thunderstorms are known to happen in the afternoon and evening. The model needs some time to develop, which means that the first lead times are not useable for forecasting. By using 00 UTC, the model is fully developed at the time lightning and thunderstorms are generally present.

3.4.2 Reforecasted Indices

The other three indices, MUCAPE, K-Index and the MJI, are currently used by KNMI for weather forecasting. The data set of these three indices used in this research is computed from the same HARMONIE cycle 40 reforecast data for previous research. Corresponding to the LPI data, the analysis time used is 00, with lead times +00 to +48.

3.5 Forecast and verification methods

The observation output, which originally is the number and the current of the discharges at specific locations and times, is changed to a binary data set containing zeros (no thunderstorms) and ones (thunderstorms). The current of the discharges is in this case not taken into account. The forecast output is taken as the maximum and the mean value of each index per time span and subdomain. As noted above, the area of interest is divided into 20 subdomains of 75 [km] x 75 [km] (31 by 31 grid points) each. The subdivision of these subdomains was done for the observations and the indices, such that all cover the same area and subdomains. Regarding the time, the observations and the indices are binned in 6-hour non-overlapping periods, such that the comparison is done for the same time span and that the daily cycle is present and could be considered. The averaging over time and area is done to reduce location errors and increase predictability.

3.5.1 Logistic regression

The forecast probability is determined using logistic regression and produces a value in the range of 0 to 1. Before going deeper into the verification metrics, first an explanation of logistic regression will be given. Part of the logistic regression examples will be shown in the *Results* section the rest can be found in the *Appendix*.

Logistic regression is used to make probability forecasts and estimates the relationship between one dependent binary variable and one or more independent variables (StatisticsSolutions, 2021). Single variate logistic regression is the most straightforward one, in which there is only one independent variable, also called predictor. Here, one set of input-output pairs exist (Stojiljković, 2019). Logistic regression finds the weights (regression coefficients) b_0, b_1, b_n that corresponds to the maximum likelihood estimation (Stojiljković, 2019). It is an iterative method that maximizes the product of all computed probabilities of the (non)occurrence

of the event in the dependent data set (Schmeits et al., 2008). This method gives the probability P that an event y occurs by (Schmeits, Kok, Vogelezang, & Van Westrhenen, 2008; Stojiljković, 2019)

$$P(y) = \frac{1}{1 + exp(b_0 + b_1 x_1 + a_2 x_2 + ... + b_n x_n)}. \tag{12}$$

The predictors $(x_1, x_2, ..., x_n)$ are selected via the forward stepwise selection method (Schmeits et al., 2008). It means that for each step, the predictor that produces the best regression in co-occurrence with the previously chosen predictors is selected (Schmeits et al., 2008). In this research, single variate logistic regression is used, so only 1 predictor (i.e. 1 index) is used.

3.5.2 Verification metrics

There are numerous verification measures to compare forecasts based on the different indices. An often used verification measure for verification of probabilistic forecasts is the Brier Score (BS). It directly verifies a probability forecast, which is more preferable than a categorical forecast, which only has two options, 0 (no) and 1 (yes), i.e. probability forecasts have more skill than categorical forecasts. The Brier Score is given as the mean square error (MSE) of the probability forecasts and is defined by (Wilks, 2011):

$$BS = \frac{1}{n} \sum_{k=1}^{n} (p_k - o_k)^2 \tag{13}$$

In this equation, p_k is the forecast probability, o_k is the binary observation, and n is the number of observations in the data set. The binary observation can have a value of either 0 (no event) or 1 (event).

Because the Brier Score is an MSE, the lower the value, the higher the performance of the forecast. This means that BS = 0 indicates that the forecast is perfect, while BS=1 indicates that none of the predictions corresponds to the observations.
The Brier Score can be decomposed into three components: resolution, reliability and uncertainty. The resolution indicates how well a forecast discriminates between thunderstorms and no thunderstorms. A high-resolution model will result in no thunderstorms for a probability of 0%, and the probability will be 100% if a thunderstorm is observed. It becomes clear that it is better when this resolution term is higher. In the worst case, the resolution is zero. The reliability indicates how close the forecast probabilities are to the observed frequencies. If the reliability is perfect, the forecasted probability always corresponds to the thunderstorm observation frequency. For example, if a forecast probability is 25%, for a perfectly reliable model, it means the event would happen 25% of the time. The uncertainty indicates to what extent one can trust the outcome. If the probability is 50% for a reference forecast, the data distribution is uniformly between 0 to 1, and the outcome is highly uncertain. A probability 0f 50% will give a maximum in the reference Brier Score, which will be 0.25. For reference probabilities that are high or low, the outcome becomes more certain. It results in a lower Brier Score than 0.25.

The Brier Score indicates the accuracy of a forecast in terms of resolution, reliability and uncertainty but does not compare it with a reference. In this research, the goal was to compare different indices. Thus, it makes more sense to use a skill score that makes it possible to compare the different indices. Therefore, the Brier Skill Score (BSS) is used, which is defined as:

$$BSS = \frac{BS_{ref} - BS}{BS_{ref}} = 1 - \frac{BS}{BS_{ref}} \tag{14}$$

The BS_{ref} is the Brier Score of a reference forecasting method, and BS is the Brier Score of the forecasting method that one wants to compare with the former score. The BSS shows the improvement of the forecasting method compared to the reference forecasting method. The climatological forecast is often used as the reference forecast, which is issued using historical records. In this research, the calculated mean value of the LPI ($BS_{LPI,mean}$) was used instead of the climatological forecast. This was done because the most important goal of this research is to get a clear view of the performance of the LPI in comparison with the currently used indices. The BSS can have a value between -∞ and 1. If BSS > 0, the forecasting method's performance is better than the reference. If BSS < 0, the performance is worse.

The Brier Skill Scores are presented in Box and Whisker plots that visualize the data distribution using the five-number summary of the data: 'minimum', first quartile, median, third quartile, 'maximum'. Box and Whisker plots give a good indication of the spread of the values in the data. Below in Figure 10, an example is presented using random numbers (not based on the data) to explain how Box and Whisker plots are read.

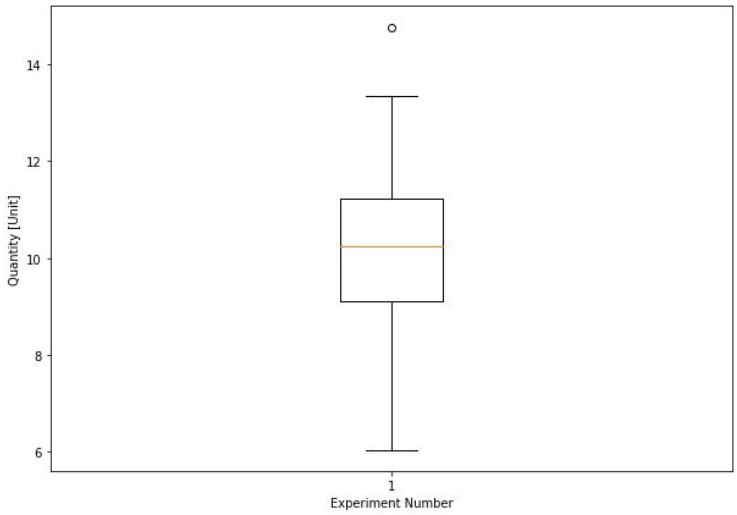


Figure 10: Example of a Box and Whisker plot using random numbers. It shows some quantity versus the experiment number.

Starting with the box, where the lowest line of the box represents the 1st quartile, and the upper line of the box represents the 3rd quartile. The red line inside the box represents the median (2nd quartile). The interquartile range (IQR) is the distance between the 1st and third quartile ($Q_3 - Q_1$). The IQR describes the variability of the data. The larger the IQR, the larger the variability of the data set. The vertical lines extending parallel from the boxes are called the whiskers. The whisker at the top is measured as 1.5 times the IQR and drawn up to the largest observed point from that data set that falls in that distance ($Q_3 + 1.5 * IQR$). From below, a distance of 1.5 times the IQR is drawn from the lower quartile down to the lowest point that falls in this range ($Q_1 - 1.5 * IQR$). The points not included in the whiskers or the box are called outliers.

In addition to the BSS, reliability plots are used to get more insight into the performance of the probability forecasts based on the indices. An example of a reliability plot is shown in Figure 11, using random data non-related to this project.

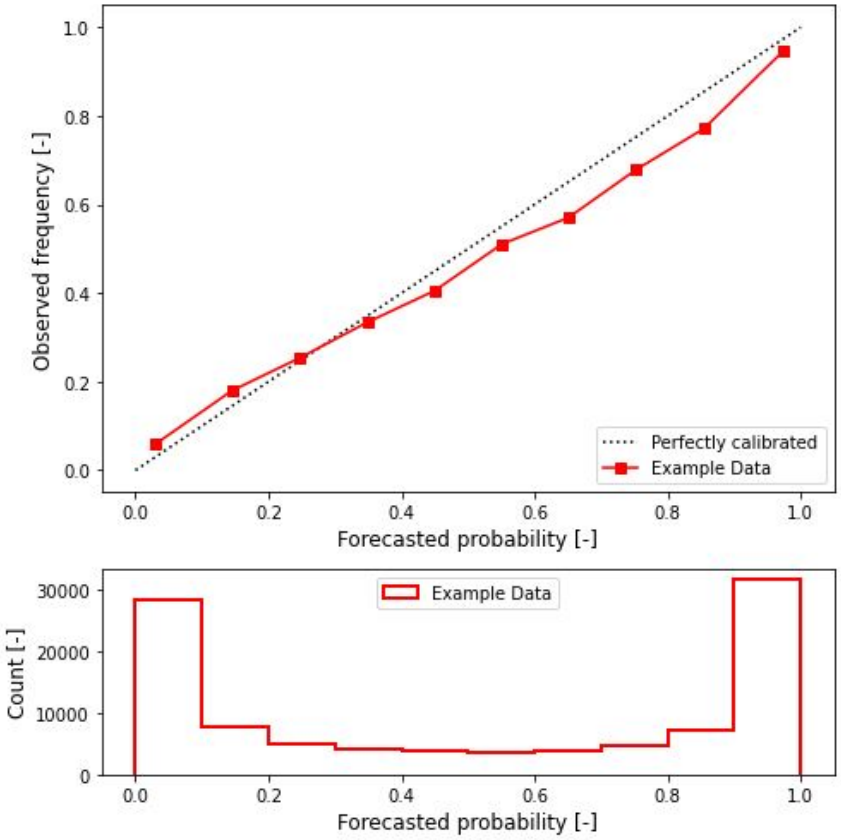


Figure 11: Upper panel: Example of a reliability plot using random numbers. On the x-axis it shows the binned forecast probabilities of a specific event and on the y-axis the observed frequency of the same specific event. The dashed black line shows a perfectly reliable model. The red line represents some data. Lower panel: The graph shows the number of observations per bin. On the x-axis there is the forecasted probability of an event and on the y-axis the count or the number of observations of that specific event. The data used is not related to this study.

In a reliability plot, the forecast probabilities are binned. These binned probabilities are plotted (x-axis) against the observed relative frequency (y-axis). The diagonal black line represents the line that a perfectly reliable forecast would follow. A perfectly reliable forecast would mean that, for example, if a forecast states that an event will occur with a probability of 40%, then the event will occur 40% of the time. As the plotted curve deviates more from this black line, the reliability becomes less. If the curve lies above the diagonal line, the forecast probabilities are too low (under-forecast). If the curve lies below the line, the forecast probability is too high (over-forecast). There may be only a few probabilities issued that lie within a specific bin. It then becomes more likely that the result will deviate from the diagonal. As a result, the diagram will look very unstable. A second plot gives insight in the number of issued probabilities per bin (Figure 11, lower panel).

For the above described forecast and verification methods Python packages have been used. The logistic regression was performed using *sklearn.linear_model_LogisticRegression*. The Brier Score was calculated using *sklearn.metrics.brier_score_loss*, the Box and Whisker plots are made using *matplotlib.pyplot.boxplot* and for the reliability plots *sklearn.calibration.calibration_curve* was used.

3.5.3 Continuous predictand

The skill scores described so far were based on a binary predictand (thunderstorm or not). To make the analysis complete and get a better impression of the performance of the LPI, another analysis is done using a continuous predictand. This predictand is the total number of discharges in the specific time frame in a specific subdomain. The result is plotted in a scatter plot, where the y-axis represents the total number of discharges and the x-axis represents the mean value of the index. The mean is calculated inside the time frame for every specific subdomain. The mean can be used to find a threshold index value above which there is a higher chance of many discharges. An example of such a scatter plot is shown in Figure 12.

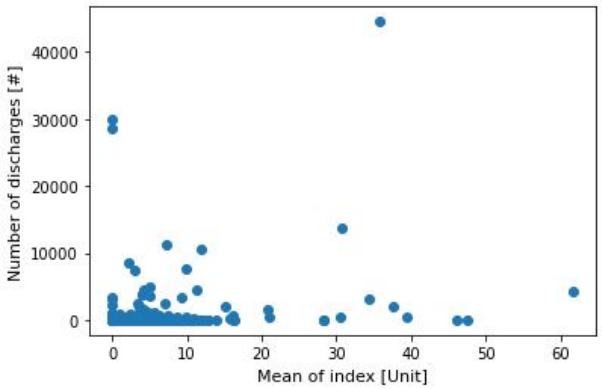


Figure 12: Example of a scatter plot which shows the number of discharges as a function of the mean of a specific index. The data in this scatter plot is random and does not describe the data used in this research.

4 Case Studies

In this chapter, two case studies at different lead times will be considered in-depth. The case studies will focus on two specific days when many lightning discharges were detected. They give insight into the performance of the LPI and the other indices for a shorter and longer lead time. They demonstrate how the LPI behaves in these cases for different lead times, and the behaviour of diverse components of the LPI. This allows a deeper insight into the relationship of the LPI and the other indices with the physical and dynamic atmospheric processes. Moreover, the case studies will help explain and understand the results.

As said, two case studies will be discussed in the following subsections. The first case study represents a situation in which the LPI operates reasonably. In the second case study the LPI provided predictions that were far from the actual observed lightning discharges.
The subsections start with an introduction to the atmospheric state and weather. Then the analyses for two different lead times will be performed based on a visual result of the different indices. Lastly, the performance of the LPI is evaluated.

4.1 23 June 2016

On the 22nd and 23rd of June, 2016, the thunderstorms that crossed the Netherlands led to much nuisance. Flooding, strong wind, large hail, and gusts damaged nature and human properties. On the 23rd, an exceptionally high number of lightning discharges was registered. This heavy thunderstorm related to a south southwest air flow containing hot, moist air that was highly unstable, which arose due to a high-pressure area above the Baltic states and a low-pressure area to the west of the United Kingdom (Sluijter, 2016). Other thunderstorms resulted from this airflow, but the focus of this case study will be the thunderstorm cluster that passed the south-east of the Netherlands at the end of 23rd June 2019. Radar observations of this storm are found in the *Appendix*.

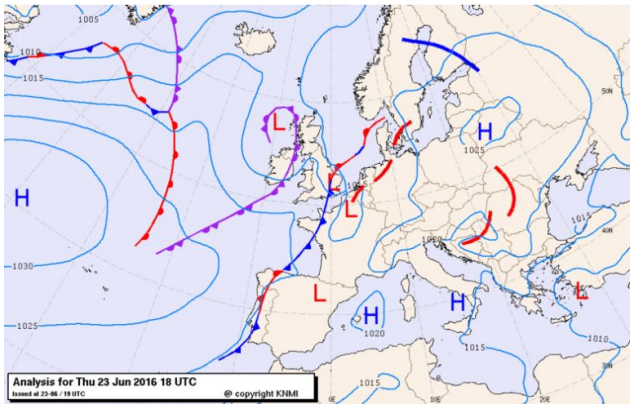
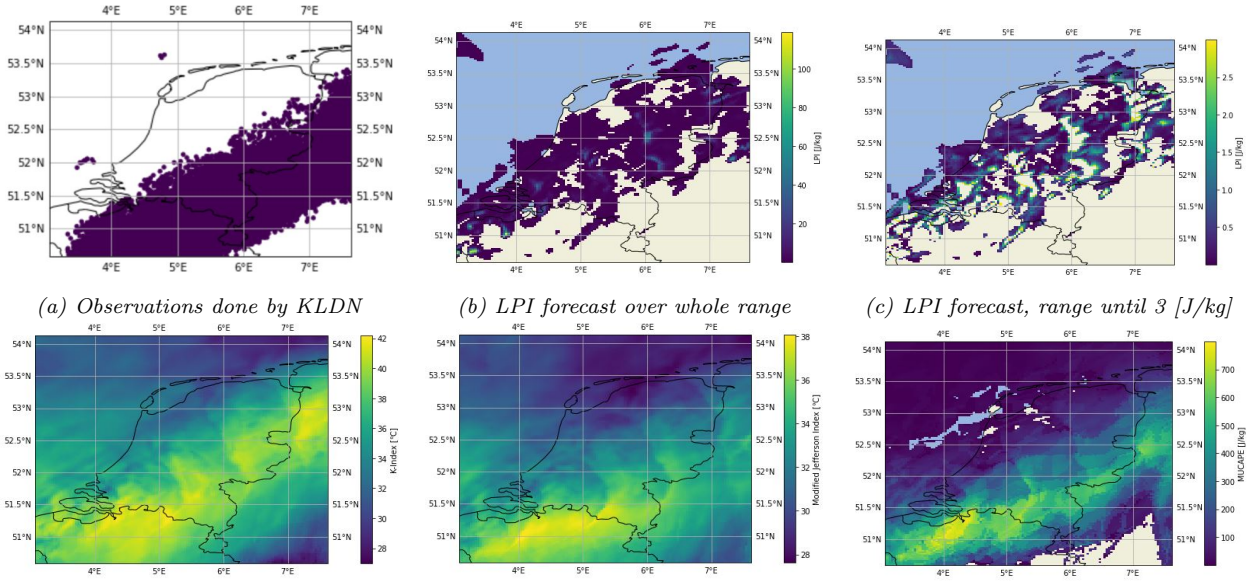


Figure 13: A map of Europe with the locations of the low and high pressure system on 23 June, 2016. The low pressure system is found west of the UK and the high pressure system is located above the Baltic states (Sluijter, 2016).

At the end of the morning of 23 June 2016, after the previous thunderstorms were gone, the sky cleared up, resulting in very high temperatures. In the east and southeast, the temperatures passed 30 °C, and the humidity was high. The dewpoint temperature reached 22 °C and, at some places, even higher. Two mechanisms probably caused air to move upward. The first mechanism was the effect of the diurnal cycle:

because the surface temperature becomes hotter than the air temperature, the air close to the ground warms and starts to rise. The second mechanism was the presence of the low-pressure system, in which air converges at the surface and starts to rise. Thunderstorms start to develop in the unstable air in the early evening in the east. Later on, thunderstorms that developed in France and Belgium in the evening moved to the Netherlands.

4.1.1 Lead time: 23 June 2016, 00 UTC +18 until 00 UTC +24



(a) Observations done by KLDN *(b) LPI forecast over whole range* *(c) LPI forecast, range until 3 [J/kg]*

(d) K-index forecast *(e) Modified Jefferson index forecast* *(f) MUCAPE forecast*

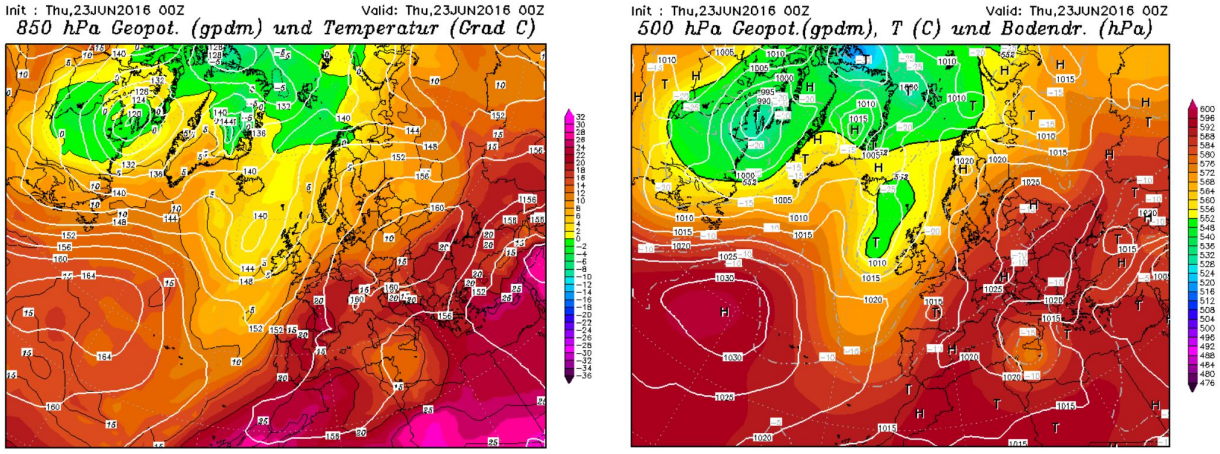
Figure 14: The observations and the result of the four indices are shown. The observations originate from 23 June 2016 18.00 until 24.00. The indices are simulated for 23 June 2016 for the 00 UTC +18 to 00 +24 run and averaged over this period. The LPI is shown twice, both representing the same situation. The upper-middle plot shows the result for the whole LPI range. The right upper plot shows the LPI values for the range until 3 [J/kg], values above 3 [J/kg] are not shown in the plot. This was done to see the lower LPI values better.

Figure 14 shows the observations of 23 June 2016 18.00 until 24.00 and the predictions of the indices using HARMONIE output for the lead time of 00 UTC +18 until 00 UTC +24. The discharges are observed by KLDN and are detected in the South, south-east and east of the Netherlands, in Belgium, and in part of Germany.

It can be seen that the LPI prediction of lightning discharges spans the correct domain shape, as it is oriented southwest to northeast. However, it is shifted to the north compared to the observations.
Interestingly, the K-Index shows high values precisely at the location where discharges are observed. The domain and the exact location seem to be correct.
The MJI also shows to have higher values in the areas where discharges were observed. The highest values are predicted in northern Belgium at the border with the Netherlands.
The higher values predicted by the MJI could be explained by the difference in temperature of the different atmospheric layers. The lower atmospheric layer (950 [hPa] and 850 [hPa]) have different temperatures for different locations in the Netherlands. Figure 15 shows that the temperature at 850 [hPa] for the area of Limburg is higher than for the area closer to the sea (west and north Netherlands). The temperature of

the 500 [hPa] layers seems to have the same temperature range. This means that for the more inland areas the K-index and the MJI, the values are expected to be higher as the differences in temperature are larger, based on the observations.

The highest values predicted by MUCAPE are found in the area where the discharges are actually observed. The general area span of the high values of MUCAPE corresponds to the observations.
This result seems to show that the HARMONIE model predicts the potential of thunderstorms well. As the K-index, the MJI and the MUCAPE represent the thunderstorm occurrence well. The question is therefore why does the LPI not predict the correct location of the thunderstorms. The different input parameters of the LPI are plotted to investigate the relation between the LPI inputs and resulting LPI predictions.

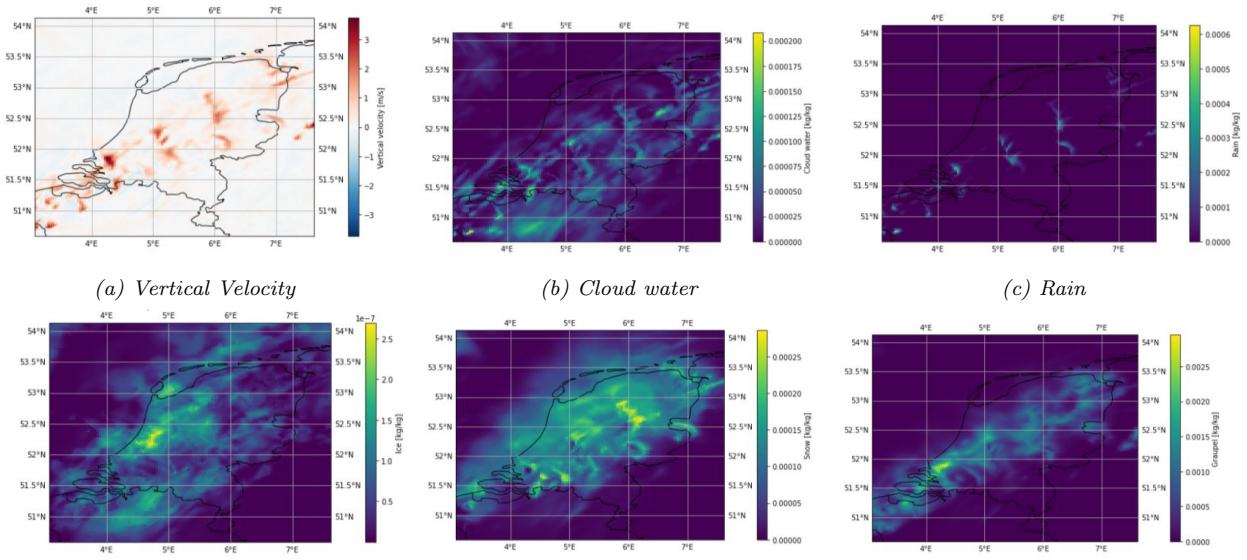


(a) The geopotential height (white lines) and the temperature (the colourbar) at a height of 850 hPa in Europe for 23 June 2016, 00z.

(b) The geopotential height (white lines), the temperature (grey, dashed lines and the surface pressure (colourbar) in Europe for 23 June 2016, 00z.

Figure 15: In the figures above the geopotential height and the temperature are shown. In the right on the surface pressure is present too (Müller and Floors, 2021).

In Figure 16 the different input parameters used in the calculation of the LPI are plotted in six different images. The vertical up- and downdraft and the different hydrometeors are plotted as an average over 6 hours from 00 UTC +18 until 00 UTC +24. The instantaneous values per hour are used. The values for hydrometeors that are plotted are averaged over height and time. This is not the same as an accumulation over time, but it gives an impression of where the hydrometeors are present and in what amount. This method is chosen because the unit of the hydrometeors given in the output file is in [kg/kg].



(a) Vertical Velocity *(b) Cloud water* *(c) Rain*

(d) Ice *(e) Snow* *(f) Graupel*

Figure 16: The vertical velocity [m/s] and the hydrometeors [kg/kg] as an average over 18.00 until 24.00 for 23 June 2016. The instantaneous values per hour are used to generate the average.

Figure 16 shows, is that for all hydrometeors and the vertical velocity, the highest values are present more to the north than where the discharges were detected. The figure shows that the locations of high vertical velocity and rain values correspond to high LPI signals. Furthermore, there is a correlation between the extent of the graupel and the extent of the LPI. These results indicate, that in this case, the two most essential hydrometeors are rain and graupel. This is consistent with the literature, in which it is stated that there is a high relation between graupel, heavy precipitation and lightning. The LPI signal fits the area of where these two hydrometeors and the vertical velocity exist, perfectly, as is expected due to the formulation of the LPI.

The results in Figure 14 and Figure 16 show that there is a difference in the model between the potential for thunderstorms due to instability and the actual consequence of the instability, the occurrence of precipitation and vertical updrafts. This effect will be researched by using the Most Unstable CIN (MUCIN). The MUCIN for this case is plotted in Figure 17. The MUCIN is the same as the other indices, averaged for 23 June 2016 over the six hours 00 UTC +18 until 00 UTC +24.

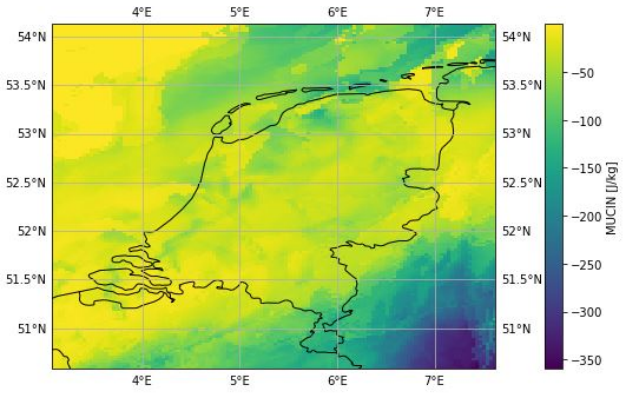


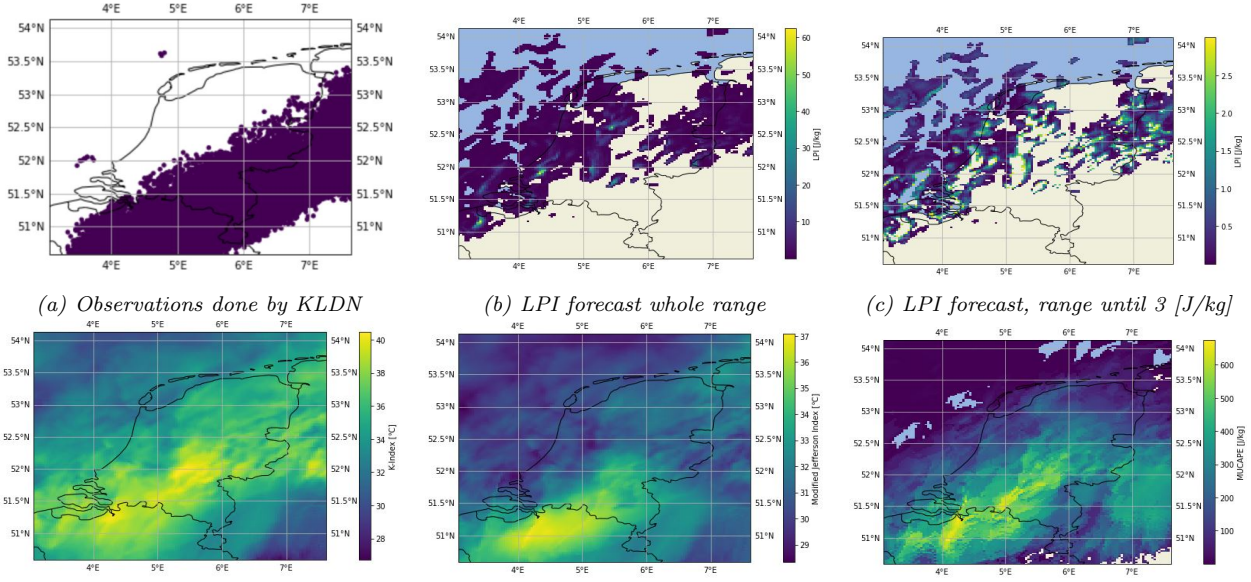
Figure 17: The most unstable CIN for 23 June 2016 averaged over the period 00 UTC +18 until 00 UTC +24. The values are negative.

Figure 17 shows the MUCIN for this case. MUCIN is the most unstable convective inhibition and is the energy that parcels need to overcome to rise higher in the atmosphere. It serves as a cap in the atmosphere, which hinders vertical updrafts from producing convective systems. If the MUCIN has a larger negative value, the negative buoyancy is more significant. This means that it becomes more difficult for the parcels to break through the CIN, and the probability of a thunderstorm is low. A first evaluation of the performance of the LPI can made using this theory.

When looking at Figure 17, it can be seen that at the location where discharges were observed, the instability was large and the MUCIN was large too, i.e. the MUCIN has large negative values. These high values of MUCAPE could be a possible explanation for the fact that no precipitation was modelled by HARMONIE. However, it could also be that when the model generates precipitation the MUCIN is reduced.

4.1.2 Lead time: 22 June 2016, 00 UTC +42 until 00 UTC +48

This subsection shows the case for 23 June 2016 from 18.00 to 24.00, but uses longer lead times (+24 hours) than the abovementioned subsection. To see the effect of the lead time on the performances. The observations and the indices are presented in Figure 18. The indices are made by averaging 22 June 2016 over 00 UTC +42 until 00 +48.



(a) Observations done by KLDN *(b) LPI forecast whole range* *(c) LPI forecast, range until 3 [J/kg]*

(d) K-index forecast *(e) Modified Jefferson index forecast* *(f) MUCAPE forecast*

Figure 18: The observations and the result of the four indices are shown. The observations originate from 23 June 2016 18.00 until 24.00. The indices are simulated for 22 June 2016 for the 00 UTC +42 to 00 +48 run and averaged over this period. The LPI is shown twice, both times representing the same situation. The upper-middle plot shows the result for the whole LPI range. The right upper plot shows the LPI values for the range until 3 [J/kg], values above 3 [J/kg] are not shown in the plot. This was done to better see the lower LPI values

Figure 18 shows the forecast for a longer lead time. All indices performed less well than when a shorter lead time was used, as expected. All four indices have their highest values shifted to the north west compared to the observations and with the previous shown short lead time forecast.
When comparing the different indices with each other, it can be seen that the LPI seems to diverge most from the observations. Almost all high values of the LPI do not correspond to where the discharges were

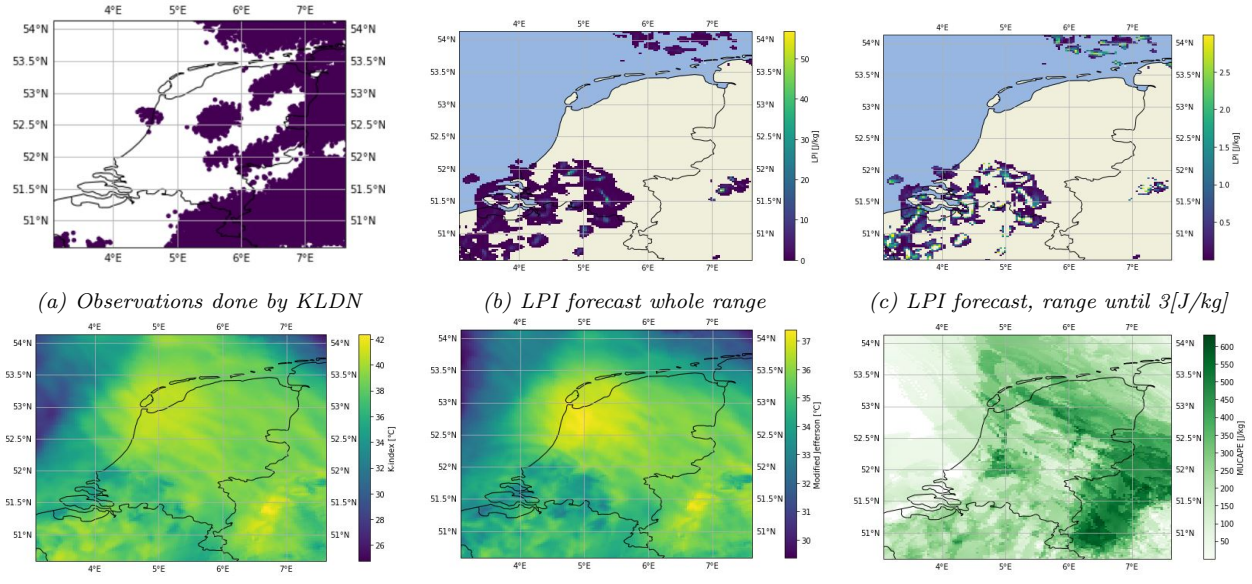
observed. Only in the north east, the LPI covers the observations, but that area contains mostly low LPI values. The K-index predicts part of the observed area, but fails to predict discharges over Limburg and parts of Germany. The MJI fails to predict discharges over a large part of the observed thunderstorm area. MJI does predict high values in northern Belgium and Brabant that do correspond to the thunderstorm area. Further to the north east, less high values are predicted. Lastly, the MUCAPE discharge prediction covers a more significant part of the observed discharge area, but the area with high predicted MUCAPE values is much larger than the actual area where discharges were observed.

4.2 19 July 2017

4.2.1 Lead time: 19 July 2017, 00 UTC + 12 until 00 UTC +18

The second case that will be studied is 19 July 2017 00 UTC +12. Observations show that a low-pressure system was present above the United Kingdom, and a high-pressure system north of the Netherlands moved east. A warm, moist and unstable airflow moved to the Netherlands from the South between these two pressure systems. During the day, in some parts of the country, the temperature reached 30 °C or higher. For some periods, the sun shined, and sometimes there were clouds. Although most of the thunderstorms resulting from this air flow were over Germany, during the day some thunderstorms reached the Netherlands (Sluijter, 2017).

As with the first case, for this second case the actual observations and the result of predictions for different indices are presented. The results are shown in Figure 19. The observations represent 19 July 2017 from 12.00 until 18.00. The indices are averaged for 19 July 2017 over 00 UTC +12 until +18.



(a) Observations done by KLDN *(b) LPI forecast whole range* *(c) LPI forecast, range until 3[J/kg]*

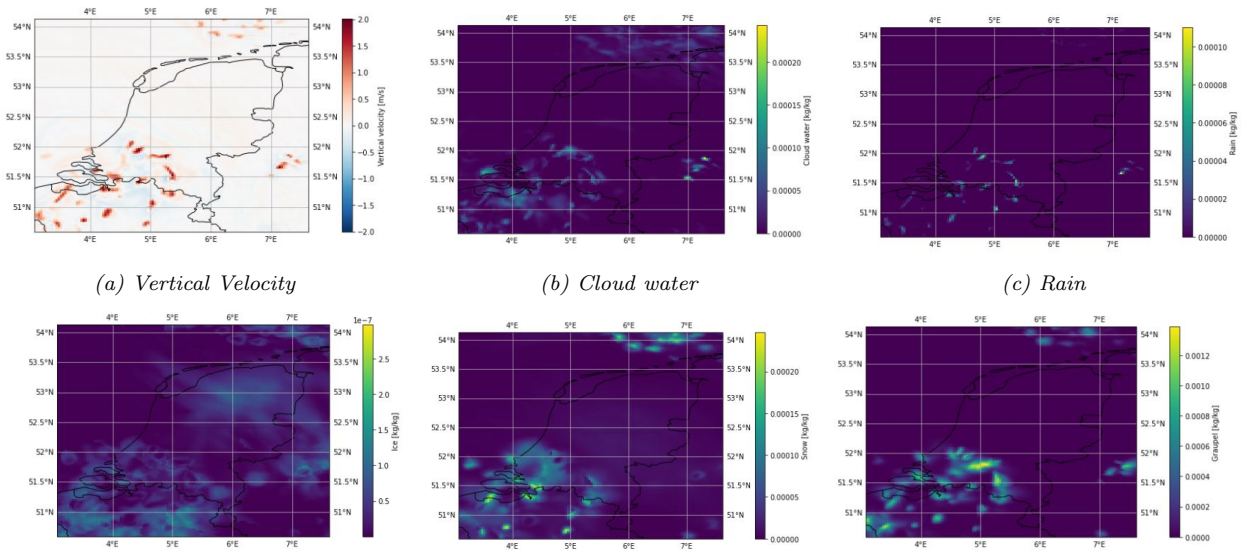
(d) K-index forecast *(e) Modified Jefferson index forecast* *(f) MUCAPE forecast*

Figure 19: The observations and the result of the four indices are shown. The observations originate from 19 July 2017 12.00 until 18.00. The indices are simulated for 19 July 2017 for the 00 +12 to 00 +18 run and averaged over this entire period. The LPI is shown twice both representing the same situation. The upper-middle plot shows the result for the whole LPI range. The right upper plot shows the LPI values for the range until 3 [J/kg], values above 3 [J/kg] are not shown in the plot. This was done to better see the lower LPI values.

Figure 19 shows that the lightning discharges are primarily registered in the east of the Netherlands and in

Germany. Some discharges are registered in the middle and western parts of the Netherlands. The forecast of the LPI is inaccurate: most values are in the south of the Netherlands, some in the north of the Netherlands and a tiny part in Germany. The northern part of the LPI does correspond to the observations, but the most significant part does not (observations in the east while the LPI predictions were in the south). The K-index shows the highest values from the east, over the center towards the northwest of the Netherlands with some additional high values in the southwest over Belgium. In this whole area, thunderstorms did indeed occur. However, the actual discharge detection is not related one-to-one to high values of the K-index. The K-index failed to predict discharges in the northeast corner and had low values around Limburg in the southeast, actual discharges were detected there. The MJI has higher values in the east, middle and northwest of the Netherlands, approximately where thunderstorms were actually observed. The northern and southern part of the thunderstorm area were not predicted by the MJI. The MUCAPE predicts that the highest values are present in the east, which is the location where many discharges are measured. The MUCAPE predicts high values in the northern part of the Netherlands and a part in the middle, which corresponds with the observations. The MUCAPE predicted lower values for Limburg, where discharges were detected.
In general the figures show that CAPE, K-index and MJI, i.e. the instability, are low at the location where the precipitation was generated by HARMONIE. This is expected as the convection in the model consumes the CAPE.
For this second case, too, HARMONIE seems to predict the potential for thunderstorms at the right location, as the three indices, excluding LPI, predict the general area of the thunderstorms.

As in this research, we are most interested in the details of the LPI performance, the model outputs of the hydrometeors and the vertical velocity are plotted. This is done to get insight into how the LPI reacts to these model outputs. Figure 20 presents the results. The hydrometeors and vertical velocities are visualised in the same way as was done above.



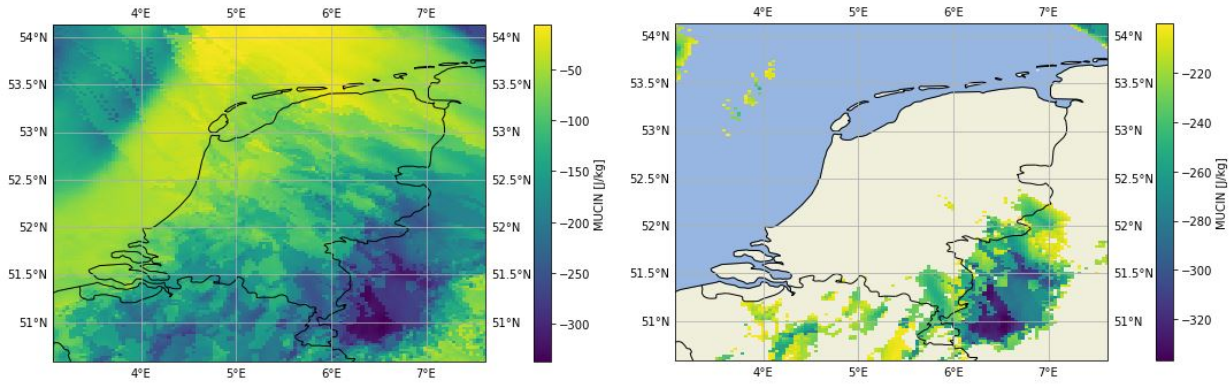
(a) Vertical Velocity *(b) Cloud water* *(c) Rain*

(d) Ice *(e) Snow* *(f) Graupel*

Figure 20: The vertical velocity [m/s] and the hydrometeors [kg/kg] as an average over 12.00 until 18.00 for 19 July 2017. The instantaneous values per hour are used to generate the average.

Figure 20 shows that locations of high hydrometeors and high vertical velocity values are related to the

LPI. All highest values of the hydrometeors from HARMONIE are located in the south of the Netherlands and in the north above the North Sea. These values corresponds to the location of high values of the LPI. The important hydrometeors are again graupel and rain, but in this case, cloud water is significant too. This case again makes it clear that the ratio of hydrometeors is responsible for the general location of the LPI and the vertical updraft is responsible for the detailed location and the magnitude of the LPI.
Again the plots show that HARMONIE seems to give a fair prediction on the potential for thunderstorms, given by the K-index, the MJI and the MUCAPE. However, the actual rainfall seems to be predicted in the wrong location, which could be the effect of the MUCIN. For the completeness of the analysis, MUCIN will be plotted for this case.



(a) MUCIN over the whole range

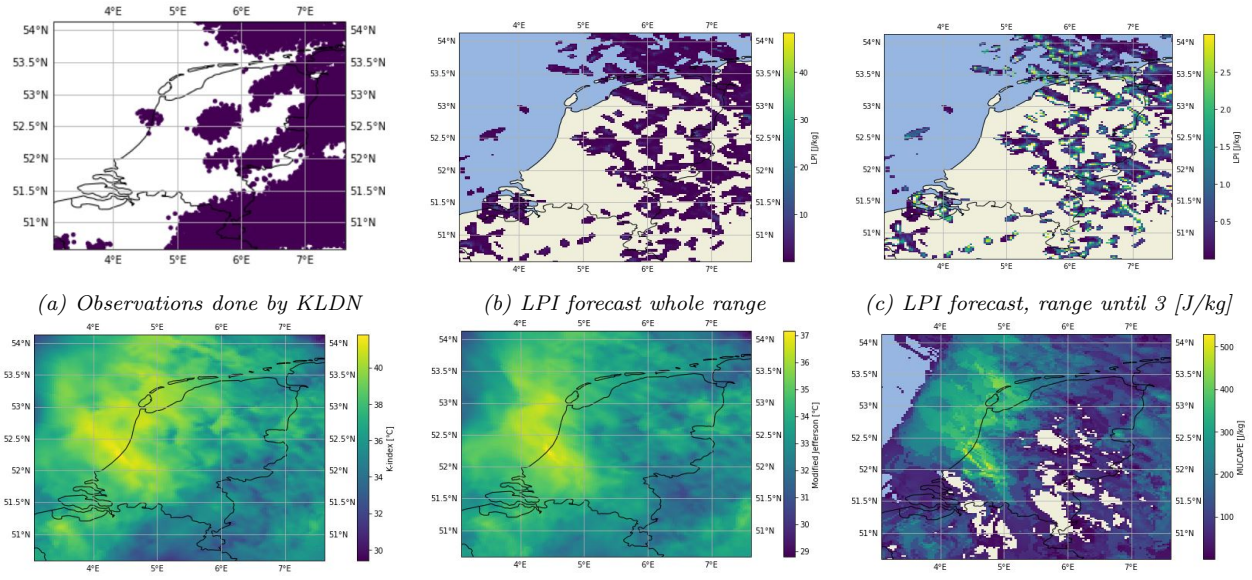
(b) MUCIN with a range starting from -200 [J/kg]

Figure 21: The most unstable CIN for 19 July 2017 averaged over the period 00 UTC +12 until 00 UTC +18. The figure left is the MUCIN plotted over its whole range. The figure on the right represent the MUCIN only over a range which starts from -200 [J/kg]. The -200 [J/kg] threshold is used, as after this values a significant amount of energy is needed to break through the CIN (Hallar, 2017; Pier, n.d.).

Figure 21 shows the most unstable CIN for the case of 19 July 2017. The left plot is the MUCIN over the whole range, but a second plot is made as it is hard to see the details. The plot on the right represents the MUCIN over the range starting from -200 [J/kg]. Using these figures, it becomes clear that small negative MUCIN values are present at the locations where the LPI predicts higher values. However, the plots also show that only a part of the area where the other indices show a high value has a larger negative value for the MUCIN, which suggests that maybe the MUCIN is not the only factor affecting the result. Another possibility is that the model generates the MUCIN correctly, but that it does not have enough capacity to overcome the MUCIN.

4.2.2 Lead time: 18 July 2017, 00 UTC + 36 until 00 UTC +42

Another example is given in this subsection that shows the effect of a different lead-time. In this example, a +24 hours lead time is taken compared to the previous example, 18 July 2017 00 UTC +36 until 00 UTC +42. The observations still represent 19 July 12.00 until 18.00.



(a) Observations done by KLDN *(b) LPI forecast whole range* *(c) LPI forecast, range until 3 [J/kg]*

(d) K-index *(e) Modified Jefferson index forecast* *(f) MUCAPE forecast*

Figure 22: The observations and the result of the four predictors are shown. The observations originate from 19 July 2017 12.00 until 18.00. The predictors are simulated for 18 July 2017 and averaged over 00 UTC +36 until 00 UTC +42 run. The LPI is shown two times. They represent the same situation. The upper-middle plot shows the result for the whole LPI range. The right upper plot shows the LPI values for the range until 3 [J/kg], values above 3 [J/kg] are not shown in the plot. This was done to see the lower LPI values better.

The difference in performance for the different lead times is represented in Figure 22. The K-index, MJI, and the MUCAPE are all shifted. The MUCAPE and the MJI have highest values shifted to west and northwest. The MJI is shifted to the south along the coastline. For these three indices, the performance is worse for a longer lead time, as expected. However, the LPI prediction shifted the forecast to the east. It means that in contrast to the other indices and unexpectedly, the LPI predicts the thunderstorm location better for the longer lead time than for the shorter lead time. The LPI location relative to the location of hydrometeors and vertical updraft as forecasted by HARMONIE. Although there is no clear explanation for this result, the most logical explanation is that incidentally the LPI at this longer lead time relates better to the observations than at the shorter lead time.

4.3 Concluding summary of the two case studies

Based on the two case studies discussed, three short conclusions are stated. First, overall, the K-Index and the MUCAPE perform very well when predicting lightning discharge location. The MJI performance is not bad but sometimes misses parts of the area. The LPI performance is the worst. HARMONIE accurately represents the potential of thunderstorms, but for some reason, the actual precipitation and strong convective activity are modelled at the wrong place. Because the LPI is directly related to the HARMONIE's vertical updraft and rain (and graupel) output, the LPI predicts the wrong discharge area if the precipitation is modelled in the wrong area. A reason for this wrong prediction could be the forecast of the MUCIN is incorrect. Another reason could be that the prediction of the MUCIN is accurate, but that the modelled convection is not strong enough. However, the effect of the MUCIN on the modelled precipitation and convection after it occurred is not known. This unknown effect impedes the certainty on which one is the cause and which one the consequence.

Second, the effect of the components of the LPI is better understood. The ratio of the hydrometeors is responsible for the general area of the LPI. The vertical velocity is responsible for the area span details and the magnitude of the LPI.

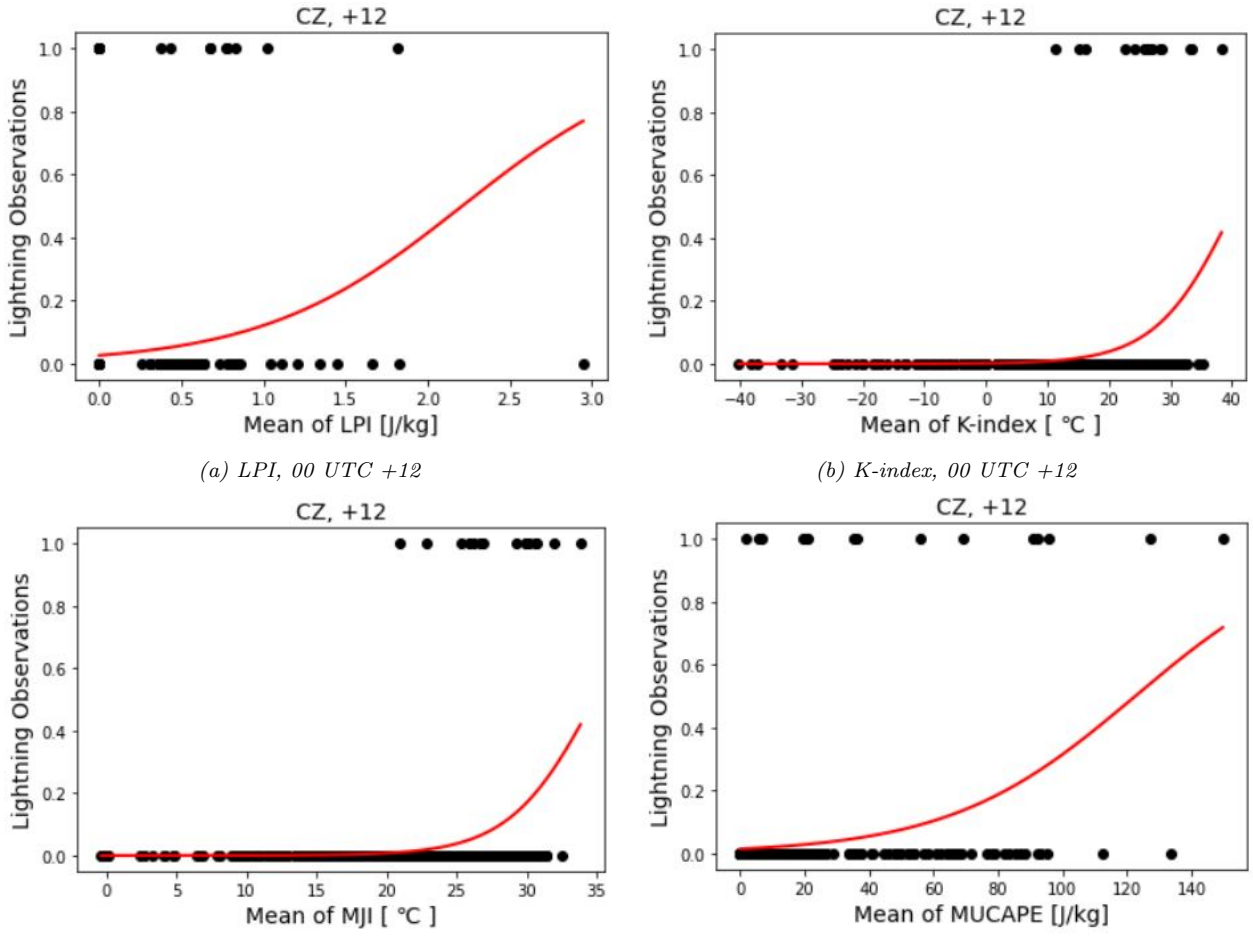
Third, the LPI is shown to generate a more local and more variable field. As a result errors will be punished more than if a smooth field was generated, as is done by the other indices. If the LPI could be changed to give a less variable more homogenous field, the performance could improve.

5 Logistic regression examples and verification results

Following the case studies that gave insight into the LPI performance for two specific cases, this chapter aims to provide insight into the performance of the probabilistic forecasts based on the different indices for a more extensive data set (the thunderstorms season of two different years). The performances of the individual indices are compared to answer the research questions. First, a number of logistic regression examples are shown. Subsequently, as explained in the *Methods* section, three different analysis methods are presented. All three analysis methods will be presented here: Brier Skill Score, reliability plots and scatter plots. At the end of this section, verification results will be presented when the whole domain is used, i.e. when the area is not subdivided.

5.1 Logistic regression examples

In this first subsection, logistic regression examples will be presented. Figure 23 shows the logistic regression fit for the four indices for one lead time and one subdomain. Other results are presented in the *Appendix*. Logistic regression is done for every mean and maximum index at all lead times, separately for every subdomain.



(a) LPI, 00 UTC +12

(b) K-index, 00 UTC +12

(c) MJI, 00 UTC +12

(d) MUCAPE, 00 UTC +12

Figure 23: Scatter plots with the results of the binary lightning observations versus four index values (black dots). Including a logistic regression fit (red line) over a shore area (CZ). The mean index value over a specific time and area is used.

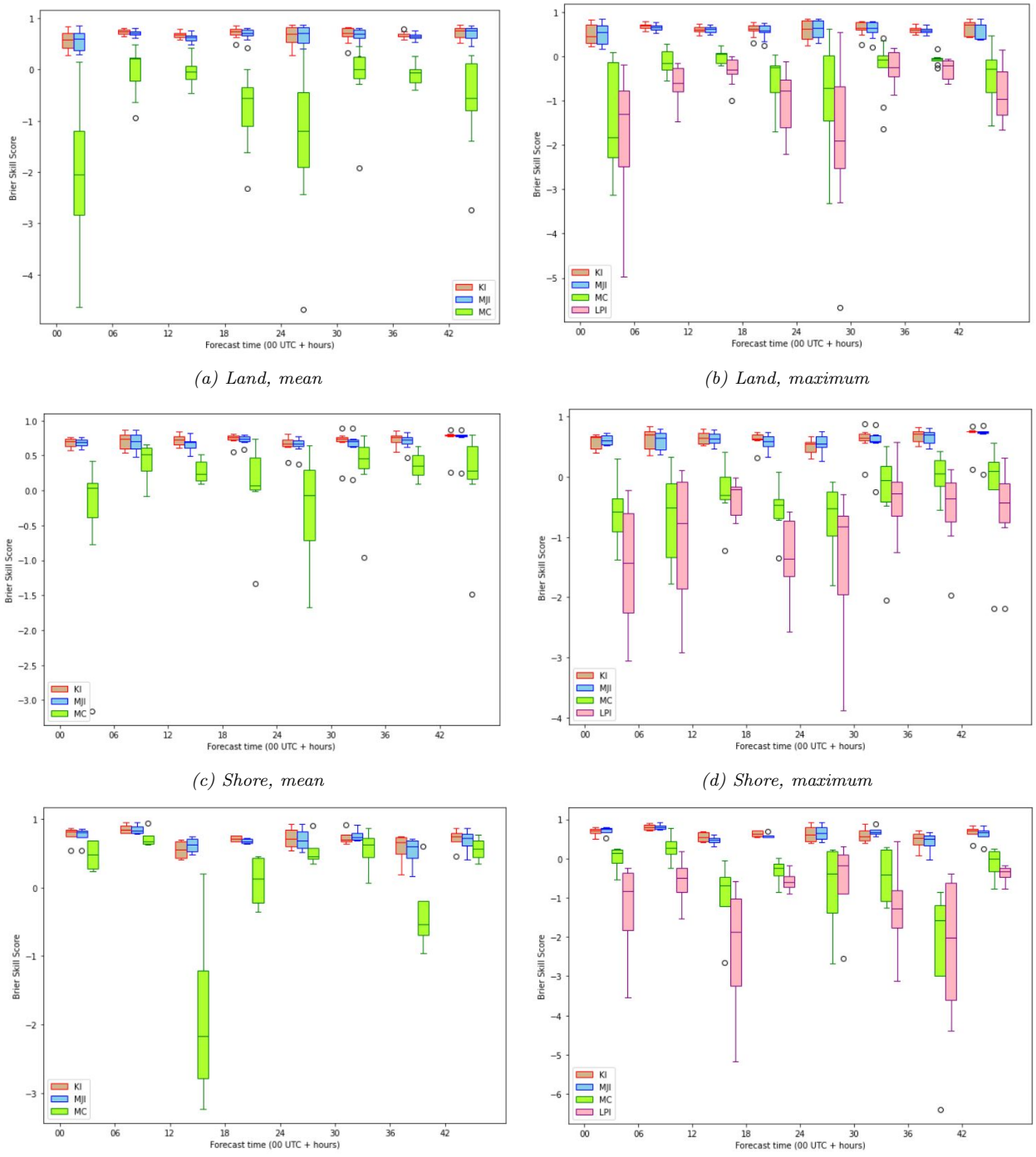
For the LPI there is not a clear separation of cases where thunderstorms occurred (black dots at y = 1 in Figure 23 and where they did not (black dots at y = 0). The logistic regression fit (the red line) starts a little bit above a probability of 0 and ends at a probability of 0.8.

The K-index and the MJI data both show a different spread than the LPI data. In this case, values of 1 are only found for high values of the mean index. Values of 0 are found over the whole mean index value range. The logistic regression fit starts to rise only at higher MJI and KI values and reaches a probability of 0.5.

The spread of the MUCAPE data is larger than for the other three indices. The plots show that both values of 1 and 0 are present at high and low mean MUCAPE values. The logistic regression fit reaches a probability of around 0.7. This is higher than is the case for the K-index and the MJI.

5.2 Brier Skill Score

The Brier Skill Score (BSS) box plots results are displayed in Figure 24. The BSS is evaluated for the maximum and the mean of all four indices and all lead times. Both the mean and maximum are calculated in each time frame and subdomain. The BSS results from combining the subdomains by taking sea, shore and land subdomains together, respectively. The reference (zero-line) is the BSS for the mean value of the LPI. The plots where the results for the K-index and the MJI are zoomed in can be found in the *Appendix*.



(a) Land, mean

(b) Land, maximum

(c) Shore, mean

(d) Shore, maximum

(e) Sea, mean

(f) Sea, maximum

Figure 24: Box and Whisker plots containing the Brier Skill Score on the y-axis. The x-axis represents the lead times of the forecasts. Left the BSS values of the mean index values are plotted and at the right the BSS values of the maximum index values are plotted. Each colour represents a different one of the four indices, as indicated in the legend.

The first noteworthy feature is that the BSS values for the maximum values of almost all indices and lead times are lower than for the mean values. This means that the mean value for all indices in general

performs better as a thunderstorm predictor than the maximum value.

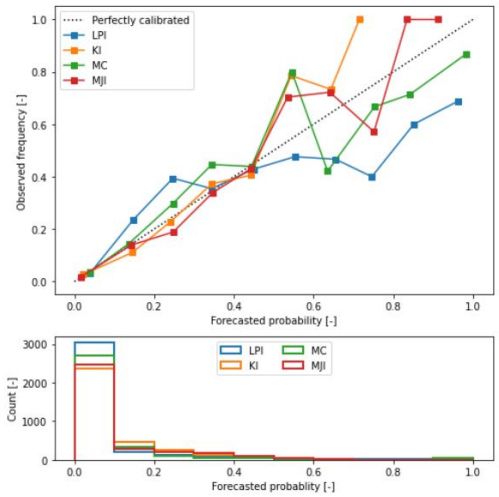
The second noteworthy feature that the figure shows is that the K-index and the MJI have a higher BSS than the LPI and the MUCAPE for every lead time and every area. The MUCAPE often shows a lower BSS than the LPI for the mean values, but sometimes it falls in the same range. Therefore, the plot shows that the K-index and the MJI perform better than the LPI and the MUCAPE. For the maximum values, the LPI performs the worst. The differences in BSS for the K-index and the MJI are minor. This is not surprising, as the definition of both is similar. In general, the K-index seems to have a slightly higher BSS, but the MJI has a higher BSS value for some lead times.
Furthermore, the Box and Whisker plots give an indication about the spatial variability of the BSS data. The variability is much more significant for the MUCAPE and the maximum LPI: the results are less consistent than the results of the K-Index and the MJI. This data spread is partly the result of unlimited negative values, but most of the time, the KI and MJI indices seem not the be affected by the positive limit. Nothing can be said about the variability of the mean of the LPI using this plot, as it is taken as the zero reference.

The difference in BSS for land, shore and sea is not large for all four indices. The BSS values at different lead times for the different area types are about the same. The K-index, MJI and the maximum LPI do not show a clear trend in BSS concerning the different area types compared to the mean LPI. For the mean MUCAPE, there seems to be a difference in BSS values between the area types: the BSS of the mean MUCAPE for the shore area seems to be the highest when compared to the mean LPI.
In addition, Figure 24 shows an unexpected trend of the BSS as a function of lead time. The expectation would be that shorter lead times show better skill, but the result shows a different outcome. This is because the reference is not climatology, but the mean value of the LPI. Therefore, based on this plot, nothing can be said about the performance of the indices at different lead times compared to the climatology, only in comparison to the mean LPI.

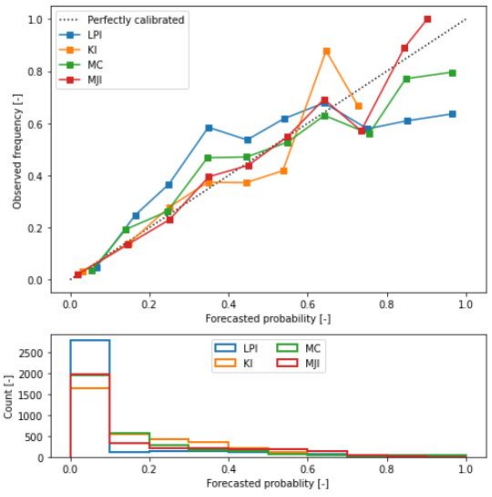
It can be concluded that forecasts based on the mean index values perform better than those based on the maximum values. Further, between the indices these results show that the K-index performs best, followed closely by the MJI. The mean and maximum MUCAPE, in general, perform worse than the mean LPI index. Moreover, in general, the maximum values of the LPI perform the least. There is not a clear pattern of the BSS of the indices through the lead times, as the BSS changes for indices and the areas. The reason for this could be that there is too little data to make a clear pattern.

5.3 Reliability Plots

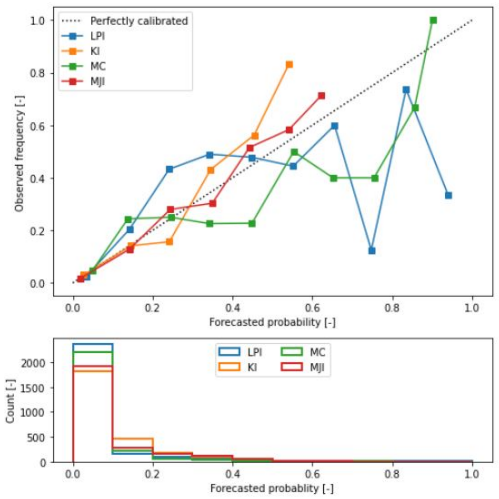
Reliability plots are the second method used to analyse the LPI performance compared to the other indices. The reliability plots were made separately for different lead times, and for land, shore and sea areas. Figure 25 shows part of the resulting reliability plots that help to answer the research question. More reliability plots can be found in the *Appendix*.



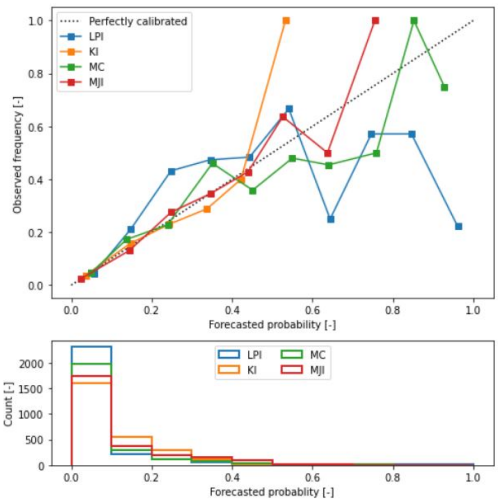
(a) Land, 00 UTC +18



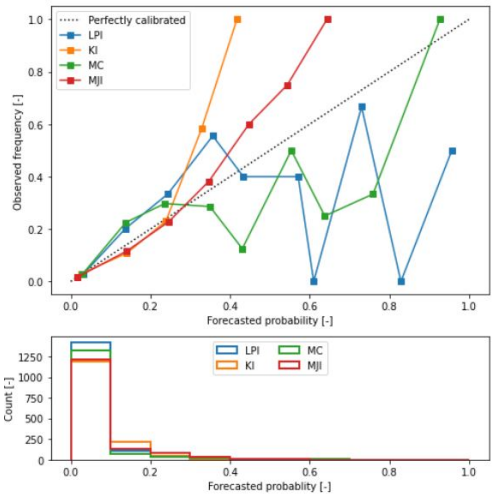
(b) Land, 00 UTC +36



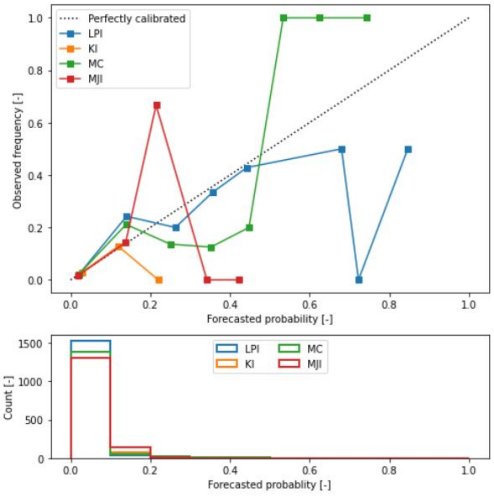
(c) Shore, 00 UTC +18



(d) Shore, 00 UTC +36



(e) Sea, 00 UTC +18



(f) Sea, 00 UTC +36

Figure 25: In the upper part of the plots, reliability plots are presented where the forecasted probability given by logistic regression versus the observed frequency is shown. In the lower part it is shown how often logistic regression issues a probability in the same bin. The black, dotted line is a perfectly reliable model. The other lines represent the forecasts based on the different indices, as indicated.

The histograms in the figure show a difference in the number of thunderstorm cases between land, sea and shore. Most cases are observed above land, least cases are observed above the sea, and for the shore, the amount of observed cases lies somewhere in between. The reliability plots that represent the areas above sea show a significant deviation from the diagonal for all indices. Next, the figure shows that sometimes the reliability curve for some of the indices does not reach a probability of 1. This happens more often for the graphs that represent the areas above the sea. It happens the least for the curves that represent the areas above land. Both can be explained by the fact that fewer observed thunderstorm cases are present above the sea than above land. As a result, it is hard to make a sharper forecast above sea areas.

In addition, the figure shows a clear distribution for high and low probabilities. For all indices and area types, the maximum number of cases is near low probabilities and progressively decreases towards the high probabilities. For this reason, the reliability curves show a larger deviation from the black line for higher probabilities than for lower probabilities and not all curves reach the probability of one.
However, there is a difference: for land, more high probabilities are present than for shore and sea. As a result, the curves representing the land areas deviate less from the black line, while the curves representing the sea areas start to deviate already for small probabilities. Based on all the described features of the plot, it seems that the reliability is higher for lower probabilities than for higher probabilities for this data.

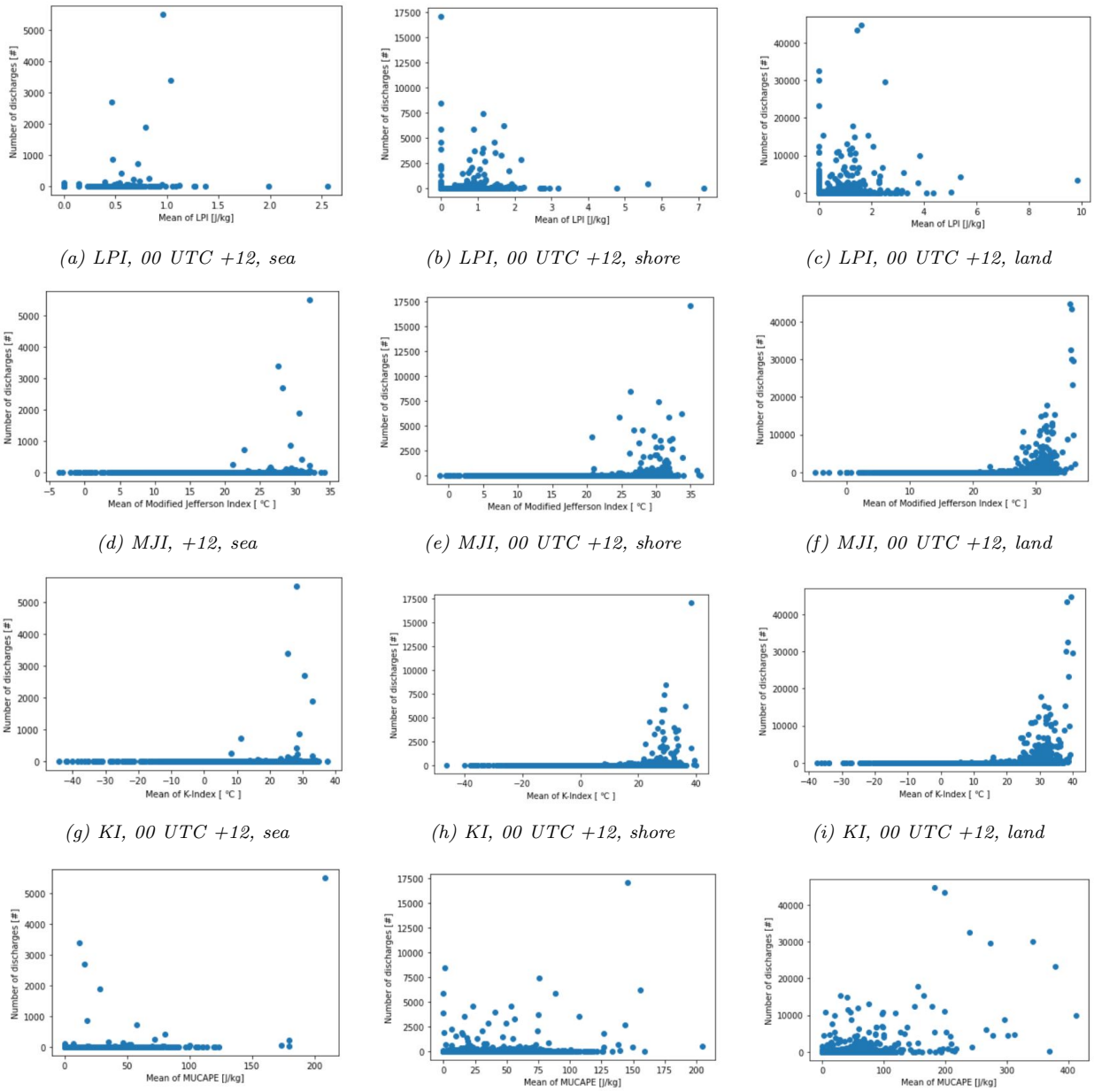
The figure further presents variations in probability distribution and reliability for the different indices. The figure indicates that the LPI forecast shows the largest occurrence of low probabilities. After the LPI, the MUCAPE is in second place in the bin for the lowest probabilities. The K-index and the MJI alternate between the third and fourth places. Figure 25 shows that the representation of higher probabilities differs for the different indices. If a sufficient amount of high probabilities is present for all indices, then the MJI and K-index have higher reliability than the LPI. However, for the shore and sea areas, the K-index and the MJI do not show any value for the higher probabilities, while the MUCAPE and LPI do. Although they are predicting high probabilities of a lightning event, there is no skill in those predictions as their observed frequency is much too low.

In the reliability plots, the reliability curves show some other differences between the indices. The curves of the K-index and the MJI follow a course that ends above the diagonal (perfect-reliability-line), implying under-forecasting. The curves of the MUCAPE and the LPI follow a course that ends below the perfect-reliability line, implying over-forecasting, i.e. the forecasted probability is too high. Additionally, it is seen that the reliability curve of the MJI, in general, deviates from the perfect-reliablility-line the least of the four indices. The K-index follows this diagonal second best, the MUCAPE third and the LPI generally deviates most. Based on these plots, the LPI is most often the least reliable and the MJI is the most reliable. This conclusion applies to the sea and shore areas as well as for the land areas. In some cases, the LPI contains fewer values in a bin, which could explain smaller reliability. However, in general, the reliability curves of the LPI deviate more than the reliability curves of the other indices, even when they contain more probabilities in a specific bin.

5.4 Scatter plots

The third analysis method uses scatter plots, but as already discussed in the *Methods* section, the scatter plots are not based on the binary predictand but on the continuous predictand, i.e. the number of discharges. The scatter plots are made separately for each index and lead time, and the subdomains are again subdivided

into the land, shore and sea. Figure 26 shows the resulting scatter plots. More scatter plots can be found in the *Appendix*.



(a) LPI, 00 UTC +12, sea *(b) LPI, 00 UTC +12, shore* *(c) LPI, 00 UTC +12, land*

(d) MJI, +12, sea *(e) MJI, 00 UTC +12, shore* *(f) MJI, 00 UTC +12, land*

(g) KI, 00 UTC +12, sea *(h) KI, 00 UTC +12, shore* *(i) KI, 00 UTC +12, land*

(j) MUCAPE, 00 UTC +12, sea *(k) MUCAPE, 00 UTC +12, shore* *(l) MUCAPE, 00 UTC +12, land*

Figure 26: Scatter plots containing the number of discharges versus the mean value of the different indices over land, sea and shore for one lead time (00 UTC +12).

This figure shows that more discharges happen over land than over shore, and more discharges happen over shore than over sea. This finding is in accordance with the results of the reliability plots. The presence of a relation between the number of discharges and the value of an index seems to be independent of the land cover (sea, shore or land). If for a given index the number of discharges is found to be related to the value of the index for land, such relationship is also found for that index for shore and sea. If for a given

index no such relationship is found for land, then none is found for shore and sea also.

It is important to see whether or not there is a correlation between the number of discharges and the values of an index.
For the LPI, the first thing that stands out is the distribution of the LPI values, which shows that low values of the mean LPI relate to low and high number of discharges and that higher values of the mean relate to low and, to a lesser extent, to high values of number of discharges. This characteristic seems to make it harder to find a correlation. In addition, the results show that the range of the mean value of the LPI is not related to the area (sea/shore/land) over which it is calculated. Furthermore, Figure 26 shows that over land low LPI values more often correspond to high number of discharges than over sea and shore. Over sea, less high number of discharges are plotted for low LPI values.

In contrast to the LPI, the MJI seems to show a correlation between high mean values and high number of discharges: The figure shows that higher mean values of MJI are related to higher numbers of discharges. No high numbers of discharges are found at lower mean values. The result corresponds to what was found by Groot (2019). This property allows one to decide on a threshold MJI value, below which threshold value it can be said that there is a low chance of high number discharges and above which threshold value it can be said there is a high chance of high number of discharges. The value of this threshold does not seem to be related to the area type over which it is taken, because discharges start at approximately the same value for land, shore and sea type areas.

The K-index shows a similar correlation to the observed lightning strikes as the MJI, where a higher mean value correlates with a higher number of discharges. This result is confirmed by Groot (2019) and expected as the definitions of the indices are very similar. Like with the MJI, this allows one to decide on a threshold MJI value, below which threshold value it can be said there is a low chance of a high number of discharged and above which threshold value it can be said there is a high chance of high number of discharges. The range of the values and the threshold value are not dependent on the areas above which the K-Index is computed.

In contrast to the MJI and K-index, the MUCAPE does not show a clear relationship between high mean values and a high number of discharges. There are situations in which a high number of discharges corresponds to a high mean value, but there are many situations where a low mean value relates to a high number of discharges. This spread is found in the sea, shore, and land areas. Setting a threshold based on these plots is therefore not possible. In addition, when going from sea to shore to land, low and high number of discharges are seen in the scatter plots. In contrast to the scatter plots of the previous two indices, in which only low or high numbers of discharges were seen.

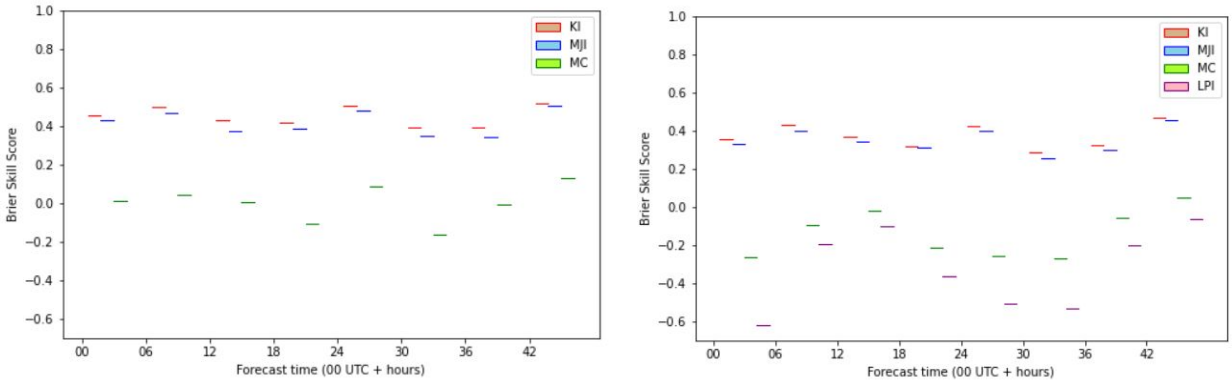
All four indices, when implemented in a simple framework as used here, could lead to false alarms, which means that a high mean value of the index relates to a small number of discharges.

5.5 LPI potential for climate runs

The case studies showed that phase errors are the most significant problem in the usefulness of the LPI as a predictor for thunderstorms. These phase errors are present because the predictions are made over small, local subdomains, although the temporal aggregation over 6 hours diminishes those errors. KNMI wants to use the LPI also for climate runs. The specific area where lightning can happen is less relevant for climate

runs, but the lightning frequency is important. Therefore, it makes sense to focus on the performance of the LPI for large areas when studying the potential of the LPI for a climate run. In addition, the effect of phase errors becomes less significant when a prediction is done over a larger area. To see the effect on the LPI performance when having less phase errors and the potential performance in climate runs, the area of interest taken as a whole is verified using the same methods as above. In this subsection, the verification results of the three different methods are shown.
The indices are binned in 6-hour, non-overlapping bins. The mean and maximum of the indices are taken over the whole area in each time bin. The observations, too, are evaluated over the whole area in a specific 6 hour time bin. The reason for using a 6-hour bin is the fact that the presence of discharges is influenced by the daily cycle. When using 6-hour bins, the daily cycle is still visible. The first result that will be shown is the Box and Whisker plots containing the Brier Skill Score.

5.5.1 Box and Whisker plot for data over the whole domain



(a) Mean of indices *(b) Maximum of indices*

Figure 27: Box and Whisker plots containing the Brier Skill Score on the y-axis. The x-axis represents the lead times of the forecasts. Left the BSS values of the mean index values are plotted and at the right the BSS values of the maximum index values are plotted. Every colour stands for an index, as indicated in the legend.

Figure 27 shows the plots containing the Brier Skill Score for the four different indices for all lead times. The left plot shows the Brier Skill Scores when the mean index value was used for the calculation, and the right plot shows the result of when the maximum index value was used. The zero reference line is given by the BSS of the LPI mean values.

The plots show that the BSS for the mean index values is lower than for the maximum index values. The BSS of the mean values of the K-index and the MJI is mostly above 0.4. The BSS of the maximum values of the K-index and the MJI is mostly below 0.4. The BSS of the mean MUCAPE is around 0. The BSS of the maximum value of the MUCAPE and the LPI are almost always lower than 0.

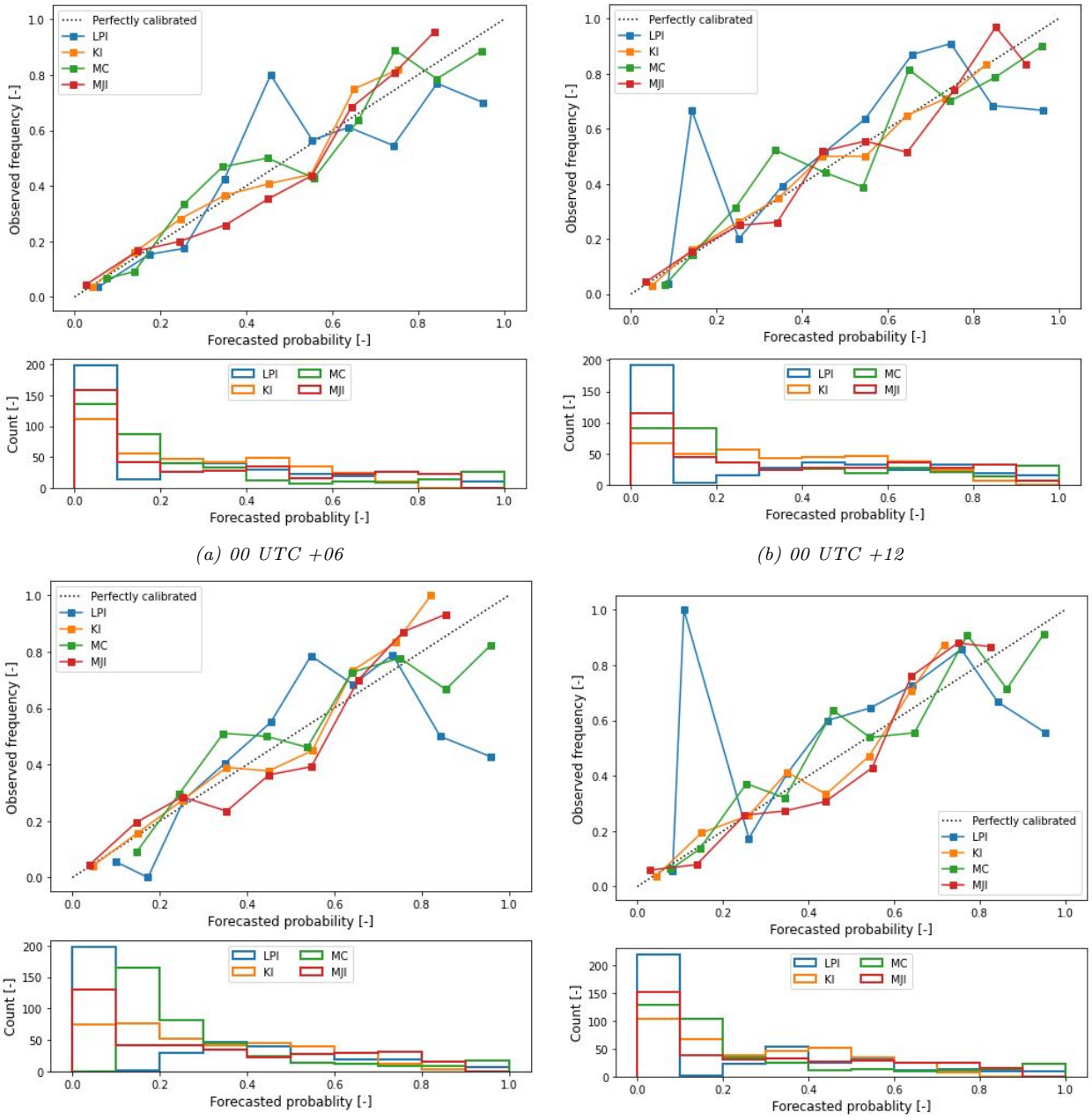
A number of differences in BSS for the four indices can be seen. First, looking at the BSS of the mean values: the BSS of the mean K-index and MJI is higher than the BSS of the mean value of the MUCAPE and LPI. The mean K-index has a slightly higher BSS than the MJI, and therefore performs best. The MJI performs second best based on this plot. The BSS of the mean MUCAPE and the mean LPI seems to have around the same value. The BSS of the mean MUCAPE fluctuates around 0.
When looking at the BSS of the maximum index values, the results are partly is the same. The maximum

values of the K-index and of the MJI have a higher BSS than the maximum values of the LPI and the MUCAPE. The maximum value of the K-index has a slightly higher BSS value than that of the MJI. The right plot shows that the BSS of the maximum LPI is lower than the BSS of the maximum MUCAPE.

When comparing Figure 24 and Figure 27, the results show that the difference in the BSS for the four indices is smaller for the whole area than for the subdomains. The difference is smaller between the mean values and the maximum values and between all indices.

5.5.2 Reliability curve for data over the whole domain

As with the verification methods of the data over the subdomains, the second method for the verification of the whole area is the reliability curve. Part of the resulting reliability diagrams is shown in Figure 28. The other results are included in the *Appendix*.



(a) 00 UTC +06

(b) 00 UTC +12

(c) 00 UTC +18

(d) 00 UTC +30

Figure 28: In the upper part of the plots, reliability diagrams are presented where the forecasted probability given by logistic regression versus the observed frequency is shown. In the lower part it is shown how often logistic regression issues a probability in the same bin. The black, dotted line is a perfectly reliable model. The other lines represent the forecasts based on the different indices, as indicated. This plot represents the result calculated over the whole area and using the mean values of the indices.

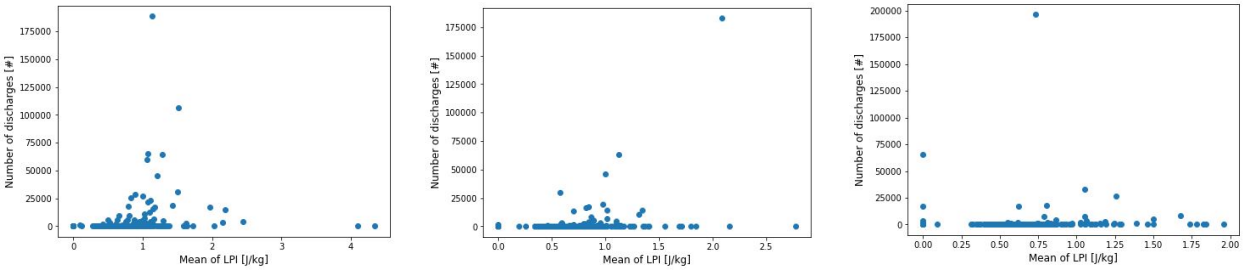
Figure 28 shows that the curves of the K-index follow the diagonal most closely in most cases in comparison with the other indices. This result seems to show that the K-index is most reliable. The curves of the MJI deviate from the diagonal slightly more than the curves of the K-index but seem to be the second-best. The curves of the LPI seem to deviate most from the diagonal in most cases. For some bins and lead times, the LPI curves deviate from the diagonal the least. However, the number of cases in which this happens is

much smaller than the number of cases in which the curves of the LPI deviate the most. This result seems to show that the LPI is the least reliable for predicting lightning over the whole area. The curves of the MUCAPE deviates from the diagonal less than the curve of the LPI but more than either the curves of the K-index or the MJI.

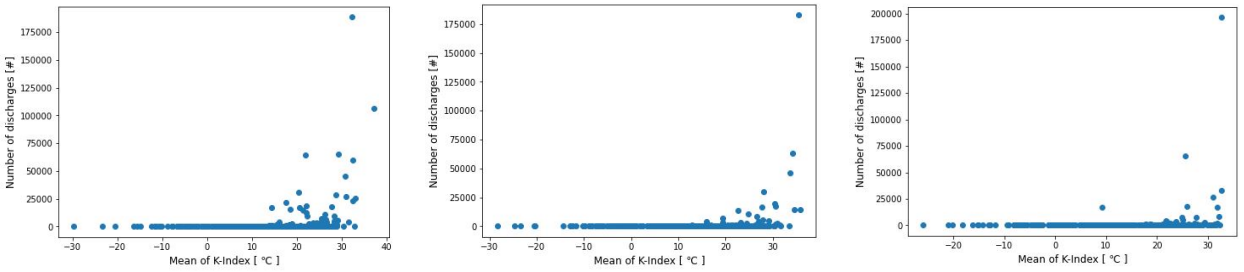
Comparing Figure 25 and Figure 28, the difference in performance is not obvious. The deviation for lower probabilities does not seem to increase when taking this larger area. This statement is true for all indices. However, the deviation is much smaller for higher probabilities for all indices. All indices are more capable of predicting high probabilities for the large area than is the case for the subdivided area.

5.5.3 Scatter plots for data over whole domain

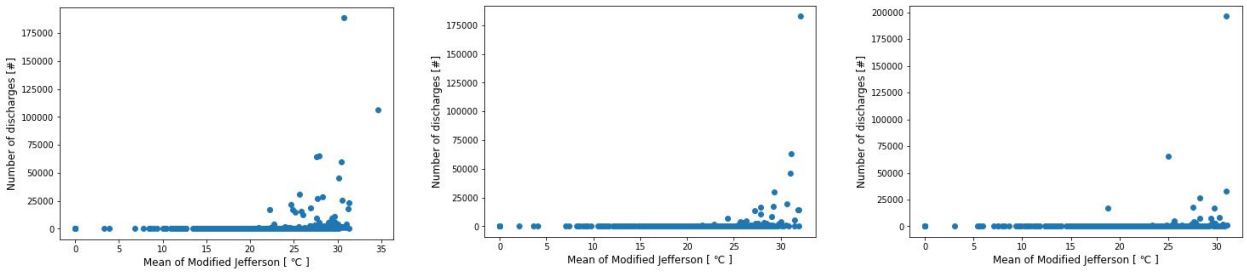
The third method used for the analysis of the data over the whole domain is the scatter plots. This method is used to investigate the relation between the number of discharges and the mean value of the index. Part of the results is shown in Figure 29. The rest of the results are shown in the *Appendix*.



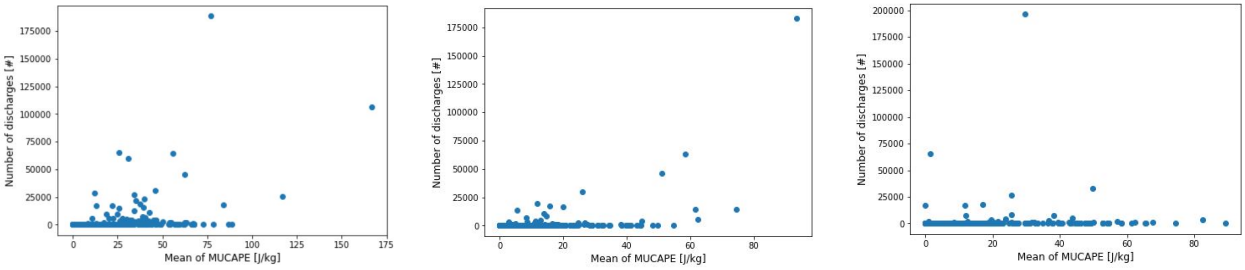
(a) LPI, 00 UTC +12 *(b) LPI, 00 UTC +18* *(c) LPI, 00 UTC +30*



(d) K-index, 00 UTC +12 *(e) K-index, 00 UTC +18* *(f) K-index, 00 UTC +30*



(g) Modified Jefferson index, 00 UTC +12 *(h) Modified Jefferson index, 00 UTC +18* *(i) Modified Jefferson index, 00 UTC +30*



(j) MUCAPE, 00 UTC +12 *(k) MUCAPE, 00 UTC +18* *(l) MUCAPE, 00 UTC +30*

Figure 29: Scatter plots containing the number of discharges versus the mean value of the different indices over the whole area for three lead times (00 UTC +12, +18 and +30).

Based on the results, partly shown in Figure 29, it can be said that a clear relationship between a high number of actually observed discharges and a high mean LPI value is not found. In many cases a high number of discharges are found at low LPI values, which makes it hard to define a clear relation. However, for some cases, it seems that a high number of discharges is present only above a value of 0.5 [J/kg]. This could indicate that there is some correlation but this correlation is not clear based on all the results. However, comparing Figure 29 with Figure 26, the result shows that the LPI values have a more substantial correlation with the number of discharges for the larger area than for the subdivided areas.

The mean value of the K-index shows to have a clear correlation with the number of discharges. Low mean K-index values correlate with a low number of discharges while high mean values correlate with a large number of discharges. For all cases, many discharges are detected when the K-index has a value above 10 °C.

For the MJI, a correlation is clear as well. High mean MJI values correlate with many discharges, and low mean MJI values correlate with a small numbers of lightning discharges. The threshold above which the larger number of discharges are observed is 20 °C.

The results show that the distribution for the MUCAPE is large. No correlation is seen relation the number of discharges and the mean MUCAPE value. Large numbers of discharges are mainly present at low mean values of the MUCAPE. However, some cases show that high mean values of MUCAPE only relate to large numbers of discharges. Compared with the results over the subdivided domains, the MUCAPE using the whole domain shows to have a somewhat stronger relationship with the number of discharges.

All four indices, when implemented in a simple framework as used here, could lead to false alarms, i.e. a high mean index value, but no discharges are observed.

6 Discussion

In this study, a first insight into the performance of the Lightning Potential Index using HARMONIE output data is given. The performance of the LPI was already studied by other researchers, but not when using HARMONIE output data. This research tried to evaluate the performance of the LPI in comparison with three other traditional indices and to provide insight into how the LPI works. This section will discuss and interpret the results of this research. They are put in perspective with the results of previous studies. The significance of the results will be discussed, and the limits of the research will be reviewed.

This research shows that the LPI performs worse than the K-Index and the MJI. The MUCAPE performance fluctuates from lead time to lead time in being better than, worse than or around the same as the performance of the LPI. The result that instability indices are performing well for thunderstorm prediction supports the results of Groot (2019), in which he stated that instability indices are the most important for thunderstorm prediction. The here-observed result of the worse performing LPI is based on verification of logistic regression equations using each index as a single predictor. The used verification metrics are the Brier Skill Score and reliability plots, in which the LPI showed to not perform very well. Besides case studies have been performed, from which it became clear that for the cases considered the LPI often predicts thunderstorms at the wrong places. The result of the LPI is affected by the definition of the filter function elaborated on earlier in this thesis. However, the definition is changeable, and the effect on the results could be studied more in further research.
The case study showed that a possible reason for the incorrect thunderstorm forecasts of the LPI is that the HARMONIE model predicts the location of strong vertical updrafts and heavy precipitation at the wrong place. However, this is not researched in depth for the whole time series. As expected, the LPI is strongly related to a model's prediction of vertical updraft and hydrometeors. As a result, no LPI signal is produced when the model does not predict heavy precipitation. The other indices show that the HARMONIE model performs well when modelling the potential for thunderstorms. Thus, for some reason, the thunderstorm potential is modelled correctly, but the actual precipitation is modelled slightly off. One of the explanations is that the MUCIN is incorrect and that modelled convection is not strong enough to break trough. It is a known feature of this model. If there is no break through, convection does not start, and precipitation will not occur. As a result, the LPI is 0. However, the effect of precipitation on the breaking through of convection is not studied in this thesis and is therefore unknown. It could be that if heavy precipitation occurs, the MUCIN is consumed, which will result in low modelled convection. This effect makes it hard to decide which factor is the cause and which the consequence: does precipitation happen because MUCIN is low? Or does the MUCIN disappear after precipitation has happened?

The case studies gave insight into the effect of the vertical velocity and the hydrometeors on the distribution and value of the LPI. It seems that the ratio of different hydrometeors is responsible for the broad span of the LPI. The vertical velocity is responsible for the details of the LPI's location and the magnitude of the LPI. The first result can be explained by the fact that the filter function depends on the vertical updraft. The second would be that if the vertical velocity is large, the air is more unstable (UniveristyOfIllinois, 2010). Parcels will reach deeper in the atmosphere, and more interchanging of charges will be present (Deierling, Petersen, Latham, Ellis, & H.J., 2008). As a result, more enormous charge differences emerge in the cloud, and the thunderstorm will be more severe, the LPI value will be higher.

Furthermore, the case study showed the LPI to be a very local and variable indicator, which is the result

of how the LPI is defined. As a result, the LPI is prone to phase errors. The other studied indices generate a much smoother field. If the LPI field would become smooth too, part of the problem for the LPI as a thunderstorm predictor would probably be solved.

Assumptions were made in the definition of the LPI and the filter function, such as a threshold for the vertical velocity of 0.5 [m/s] and an area over which the majority of the surrounding grid points is taken. These assumptions are based on the definition of Lynn and Yair (2010) but are changeable. This thesis has not investigated different assumptions for the LPI definition, as it was out of the scope. However, it could affect the performance of the LPI and would therefore be interesting to investigate in subsequent research. A first try could be to decrease the vertical velocity threshold to 0.1 [m/s], which will increase the area over which the LPI is defined. It could increase the performance of the LPI.

The purpose of this study was to study the performance of the LPI compared to three other indices and not to develop the best statistical model for probabilistic forecasts of (severe) thunderstorms as in Schmeits et al. (2008) and Groot (2019). When a statistical model is trained, tested and validated, different data sets should be used for this, namely, a training data set and an independent validation data set. The training data set consists of all the available data in this study. This means that for the training data set, 14 months of data have been used. Using this data logistic regression equations (using the LPI and the other indices as single predictors), have been derived and tested. As the difference between the performance of the LPI on one hand and that of the K-index and the MJI on the other hand is so large, the conclusions would be the same if an independent validation set would have been used. The most optimal parameters for the LPI have not been studied here. Finding the optimal parameters and using the adjusted LPI in a statistical model as an extra potential predictor is recommended for further research.

In this research, the original data was modified before it was used for the analysis. The data were binned in 6 hours non-overlapping bins, and the area was divided into 20 subdomains.
The subdivision of the large area makes the result more prone to phase errors, as the prediction becomes very localised. The phase errors can be reduced by taking the whole area and not subdividing it into subdomains. However, in this part of the research, the main focus was not the frequency distribution in climate runs but a first insight into the performance of the LPI as a thunderstorm and lightning predictor. When predicting lightning in the short term (+48 hour forecast), the probabilities of lightning in relatively small areas is important. Therefore, the subdivision of the larger area in this research makes sense.

The whole forecast period of 48 h is used in this research. The results show that for all lead times the LPI does not perform well. It is known that forecasts for short lead times are generally better than for long lead times. It could be questioned if it would not have been better to focus on one short lead time and try to increase the performance of the LPI based on this one short lead time. However, the model is affected by spin-up. It means the model needs some time to develop fully.
Furthermore, it is preferable to include the daily cycle in thunderstorms in the analysis which means that at least 24 hours of the forecast is needed.

Furthermore, this research investigated the correlation between high LPI values and a high number of discharges. However, according to the results of this research, this correlation does not exist. This became clear from the scatter plots. The same lack of correlation was found for the MUCAPE. For the K-index and the MJI, the correlation was much better. For both indices a threshold value can be deduced, above which

almost all large numbers of discharges occur.

The results show that, in general, the LPI does not work as well, altough that was expected based on the studies of Brisson et al. (2021) and Yair and Lynn (2008). They both found that the LPI outperforms other indices and shows a robust relationship with observational data. The results obtained in this research show that the LPI does not correlate to the observations so well, and it does not outperform the traditional indices when using HARMONIE output.

The differences are partly explainable. In their research, Yair and Lynn (2008) focus only on two specific cases, which are two days. It is very well possible that the LPI works perfectly for these two days. However, to make a more reliable statement about the overall performance of the LPI, a more extensive data set should be used, as we have done.
The data set used in this research spans seven months of two years, which means that data of 14 months is used.

Brisson et al. (2021) used a longer and more extensive data set (13 years). Their results showed that the LPI performed well, but they did not verify the performance of the LPI in a forecasting context. They looked at the relation between the LPI value and the discharge frequency distribution. They also used a different definition of the LPI. Their definition included the buoyancy effect in the form of CAPE by adding an extra filter function. This different definition could affect the performance of the LPI. However, suppose a look is taken at the results of the case studies and the definition of the extra filter function. In that case, it is not plausible that it explains the significant difference in the outcome. The extra filter function is added to prevent false LPI signals that result from deep orographic wave clouds. First, in the Netherlands, there are not many orographical features that could be responsible for these false LPI signals. Maybe, it could happen in Maastricht, but when looking at the case studies, no LPI signal is detected in Maastricht three out of four times.
Furthermore, the case studies show that the problem of the wrong forecast of the LPI is not related to too many LPI signals but to phase errors. Therefore, when looking at the case studies, the added filter function does not seem to be the reason for the difference in the outcome. However, the effect on the whole data set is not tested so that nothing can be said about the effect on the overall outcome.

Another reason for the differences could be the definition of the filter function. Starting with the area over which the majority is taken: Brisson et al. (2021) used an area of 10 [km] x 10 [km], in this research an area of 12.5 [km] x 12.5 [km] is used. The effect of taking a different area is not studied here. The results could become different if the LPI is evaluated over a larger area because it becomes less local. However, when looking at the case studies, this condition does not seem to solve the problem.
In addition, there is another difference in the definition of the filter function. Brisson et al. (2021) state that valid LPI values are restricted to grid points for which the majority of grid columns in a specific horizontal neighbourhood (10 [km] x 10 [km]) exhibit a maximum updraft speed of the column above a certain threshold. In this research, not the maximum of the surrounding columns is used, but the value of each grid point in the horizontal plane of the grid point of interest should be larger than a threshold.
The last difference for the definition of the filter is this certain threshold. The wind speed that is used as the threshold by Brisson et al. (2021) is 1.1 [m/s], while in this study, 0.5 [m/s] is used.

Although the results show that the LPI is not performing as well as a traditional thunderstorm predictor

using HARMONIE, it does not mean the LPI can not be used for climate runsa and on the short-term KNMI would like to use the LPI for climate runs. Based on the case study, the most important reason for the LPI to perform bad seems to be the wrong location of thunderstorms at the wrong time. For climate runs, the location of the thunderstorms is less critical. More important is the frequency of the thunderstorms that occur. Therefore, in this research, a first look is taken into the performance of the LPI over a large area, not using subdomains.

However, the verification results showed that the LPI does not perform well when evaluated over the whole domain. The reliability plot and the Brier Skill Score showed that the mean K-index, in general, performs best and that the mean Modified Jefferson index performs second best. The mean MUCAPE and mean LPI perform around the same. The maximum values of all indices perform worse than the mean values of the same index. Among the maximum index values, the maximum value of the LPI performed the worst, and the maximum values of the K-index performed best.
The differences in performance between the four indices did become smaller for a larger area than for the subdomains.

The reliability plots show that, in general, the indices are better at predicting high probabilities for the large domain than for the smaller subdomains. Furthermore, the LPI seems to predict small probabilities well. However, larger probabilities are harder to predict using the LPI. The last thing that becomes apparent when looking at the reliability plots is having fewer issued probabilities. The effect is mainly visible for the LPI at low probabilities. If fewer probabilities are issued, the reliability line of the index can more easily deviate from the hypothetical perfectly reliable model.

Because the whole domain is used, fewer combinations of observations and index values are present. As a result, the statistical significance of the result of the whole domain will be less than for the results of the subdomains. However, as it was intended to use the data to give a first impression of the performance of the LPI on a larger area, the results give relevant information. If research into the performance of the LPI on a larger area is done in the future, a higher statistical significance is essential. Then more data is needed.

7 Conclusion and Recommendations

The performance of the LPI has been evaluated and compared with traditional thunderstorm indices. Based on model verification results and the case studies, a conclusion can be drawn. Furthermore, recommendations for further research will be mentioned.

7.1 Conclusion

Previous research showed that the new LPI, based on microphysics and resolved vertical motions in the atmosphere, gives an accurate representation of the location of lightning. This research focused on the LPI's performance using output data of HARMONIE cycle 40 compared to traditionally used indices. Based on the results of the case studies, the BSS and the reliability plots, I found that the LPI performs worse than the traditionally used indices. The K-index and Modified Jefferson index perform relatively well when predicting lightning. It means that I would not recommend the LPI as a lightning predictor using the formulation assumed in this research. Furthermore, the LPI was found to be dependent on the location of the vertical updraft and the presence of hydrometeors. The definition of the LPI explains this, but it makes it very prone to errors in the model. If the precipitation is placed at the wrong location, nothing can be expected from the LPI. The hydrometeors are responsible for the area span of the LPI. The vertical velocity determines the details of the LPI and the magnitude of the LPI signal. Further, the HARMONIE model predicts the potential of lightning relatively well, shown by the smooth fields generated by the instability indices. The HARMONIE model has more difficulties in predicting the exact location of the thunderstorm, i.e. the location where the instability is consumed, which is predicted by the LPI. Additionally, the LPI produces a local, strongly fluctuating field and is more prone to phase errors than when the field would be smooth, as is the case for the other investigated indices.

The conclusions above hint to at how the LPI could be improved as a thunderstorm predictor. However, one of the most important goals of future research is to see how the LPI will perform during climate runs, it makes sense to see how the LPI performs on a larger area, where phase errors have a lesser impact. Therefore, a first, small study was done on the performance of the LPI over the whole area of interest. Based on the verification results, some conclusions can be drawn. First, the mean value of every index works better than the maximum value of the same index. Second, the differences in performance of the indices are smaller when taking the whole domain than when using the subdomains. However, the LPI is not the best performing index, not even when we diagnose the performance on a very large domain. The K-index performs best, and the Modified Jefferson index performs second best. The performance of the MUCAPE and the LPI seems to be around the same. The last thing the results show is that the LPI is best in predicting low probabilities. High probabilities, however, are predicted less accurately.

Lastly, as this thesis gives a first impression of the performance of the LPI with HARMONIE, it does not include different formulations of the LPI, which is essential for optimizing its performance. Therefore, I would recommend further research on the LPI to investigate how it could perform better.

7.2 Recommendations

As stated, profound research is needed, as still a lot is unknown about the LPI. Furthermore, although the performance of the LPI as a thunderstorm predictor in this research was not well, the potential for the LPI when the HARMONIE model would not have phase errors is large.

The performance of the LPI is evaluated in this research only using one set of formulations of the filter function and the LPI. More definitions are possible and could affect the performance of the LPI. Therefore, the first recommendation is to test the LPI with other formulations or assumptions. Examples of assumptions or formulations that could be changed are:

1. the minimum vertical updraft in the filter function
2. a larger area, wherein the majority of surrounding grid points should have this minimum vertical updraft
3. the definition of the filter function such that not a majority of the parcels, but the majority of the columns have a minimum vertical velocity

The data used in this research is given as instantaneous hourly output. The result is that the LPI is calculated with instantaneous hydrometeor content data for every hour. This amplifies the local characteristics of the LPI. In the following research, I would recommend using data output that is accumulated for over a whole hour or use output data of the hydrometeors for every 15 minutes or even test the LPI online. Such that a calculation is done for every moment. This would probably give a smoother field. I think it is interesting to study the effect of a smoother field on the performance of the LPI.

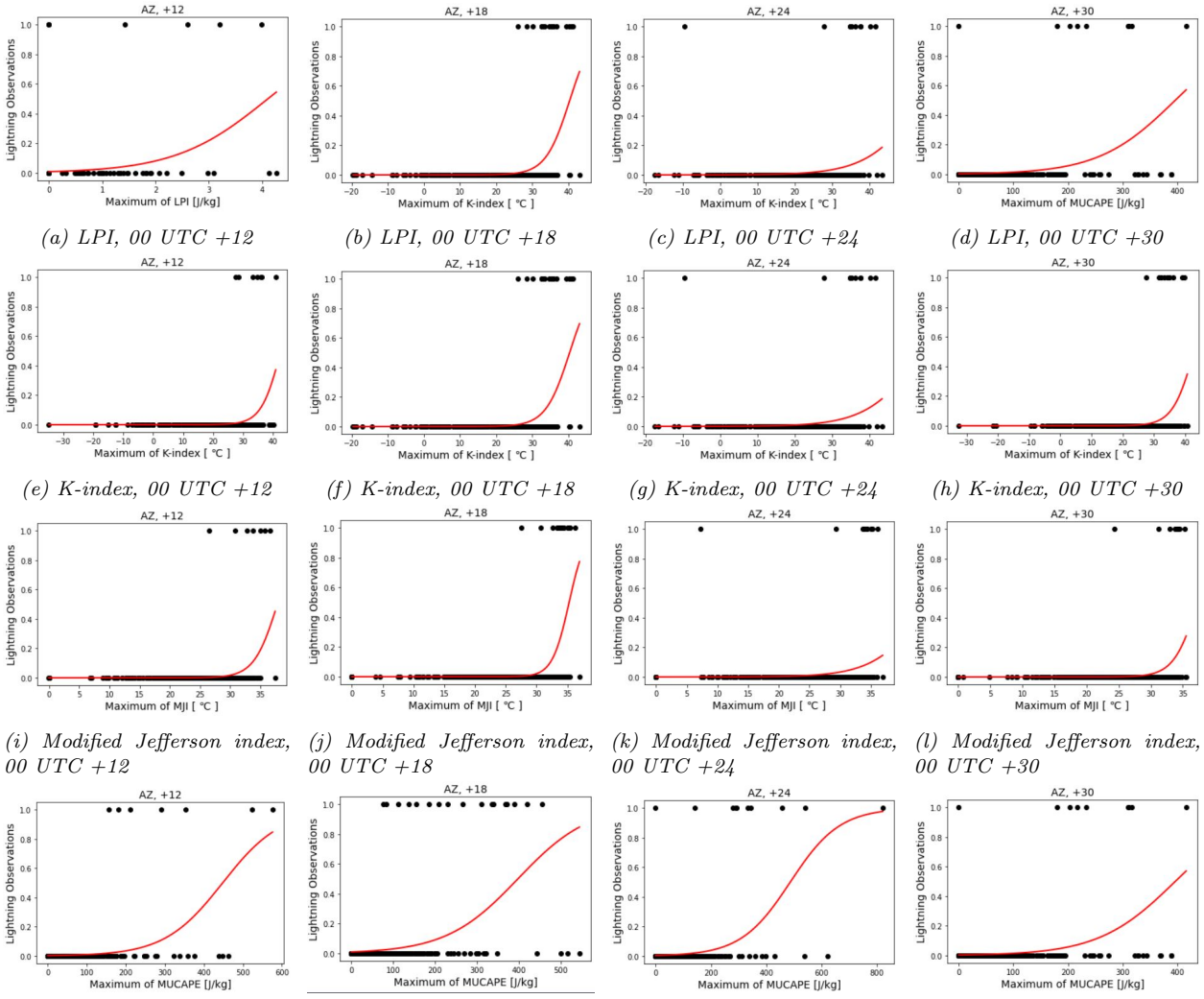
It is well known that the individual performance of predictors is less than when multiple predictors are used. Therefore, it would be interesting to investigate the performance of the LPI together with other predictors. I would recommend using the mean value of the LPI as it has been shown to perform better than the maximum value. A statistical forecast model can be developed using post-processing techniques. I would recommend using quantile regression forests (QRF) as this was found to be the best performing post-processing method for probabilistic forecasts of thunderstorm occurrence (Groot, 2019). Combining the benefits of the instability indices and the theoretical benefits of the LPI, which possibly give independent information, will likely improve the performance.

It is known that the LPI response is strongly intertwined with the NWP-model and that the new HARMONIE cycles perform better than cycle 40. Even though the microphysical scheme will not affect the MUCIN and the resolved convection, it could impact the performance of the LPI and the precipitation. Due to a stronger build-up of moisture under the inversion at the top of the boundary layer in the new version of the HARMONIE model, strongly resolved convection is supported. It would mean that the LPI would work better too if using the output of the new model. Therefore, it is interesting to test the LPI using data of the new cycle with different microphysical schemes and see if there is an improvement in the performance of the LPI.

The recommendations above are mainly related to increasing the performance of the LPI as a thunderstorm indicator. This is interesting and could be significant, in particular, if the new HARMONIE cycle achieves a higher hit rate. However, in the long term, KNMI also wants to use the LPI for climate runs. Therefore it would make sense to also dive deeper into that. I recommend that subsequent research focuses on the formulation of the LPI to find the preferred formulation that gives the best representation of lightning appearance. Furthermore, it would make sense to investigate the relation between the Lightning Potential Index value and the discharge frequency distribution as in Brisson et al. (2021). At last, after the preferred

formulation is found, it is recommended to study the discharge frequency distribution based on lightning observations and compare it with the distribution for other indices for the past, present, and future.

A - Logistic regression fit



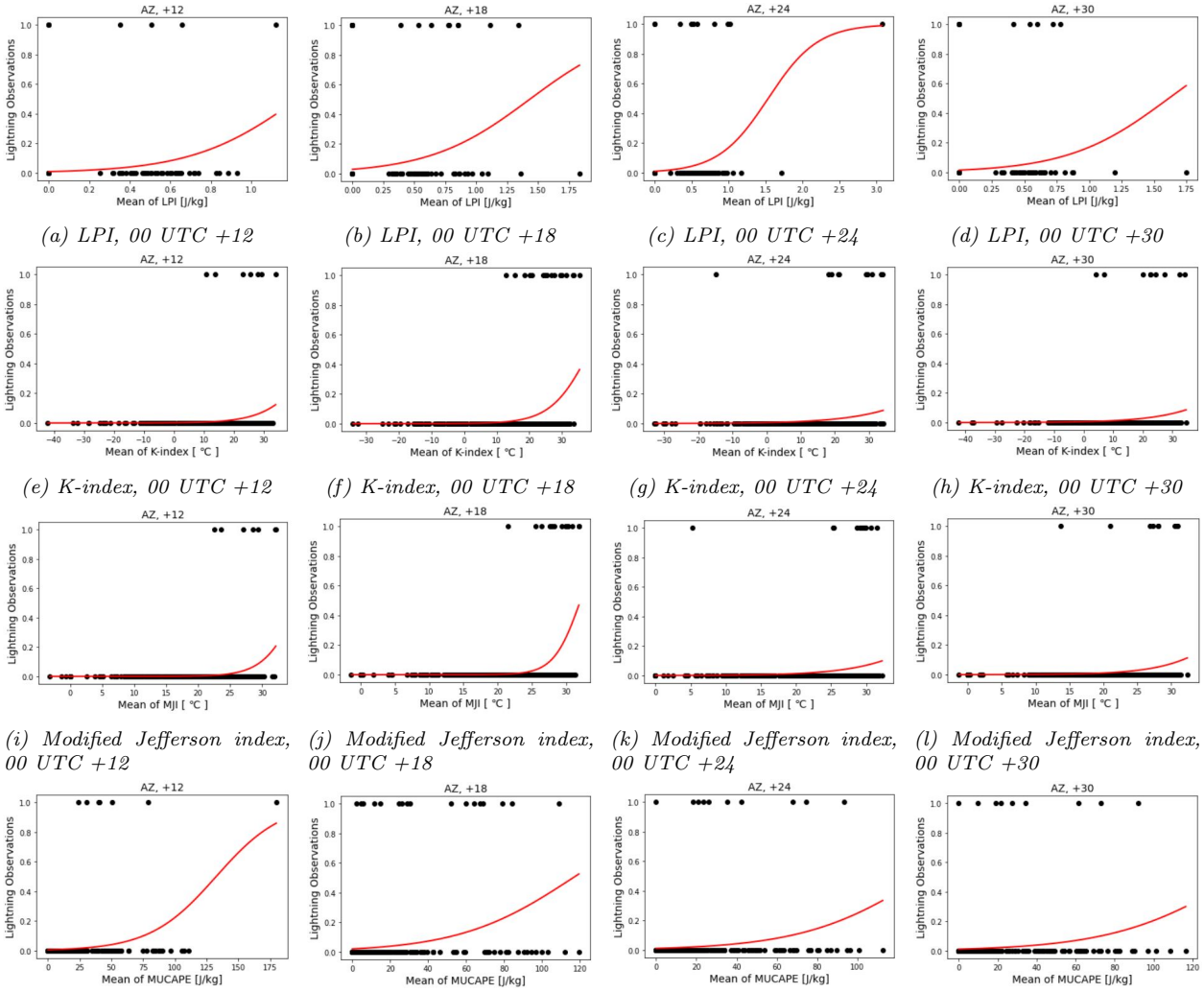
(a) LPI, 00 UTC +12 *(b) LPI, 00 UTC +18* *(c) LPI, 00 UTC +24* *(d) LPI, 00 UTC +30*

(e) K-index, 00 UTC +12 *(f) K-index, 00 UTC +18* *(g) K-index, 00 UTC +24* *(h) K-index, 00 UTC +30*

(i) Modified Jefferson index, 00 UTC +12 *(j) Modified Jefferson index, 00 UTC +18* *(k) Modified Jefferson index, 00 UTC +24* *(l) Modified Jefferson index, 00 UTC +30*

(m) MUCAPE, 00 UTC +12 *(n) MUCAPE, 00 UTC +18* *(o) MUCAPE, 00 UTC +24* *(p) MUCAPE, 00 UTC +30*

Figure 30: Scatter plots with the results of the binary lightning observations versus four index values. Including a logistic regression fit (red line) over a sea area (AZ). The maximum index value over a specific time and area.



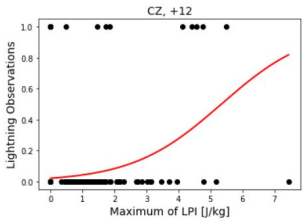
(a) LPI, 00 UTC +12 *(b) LPI, 00 UTC +18* *(c) LPI, 00 UTC +24* *(d) LPI, 00 UTC +30*

(e) K-index, 00 UTC +12 *(f) K-index, 00 UTC +18* *(g) K-index, 00 UTC +24* *(h) K-index, 00 UTC +30*

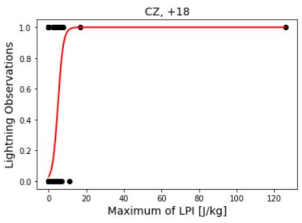
(i) Modified Jefferson index, 00 UTC +12 *(j) Modified Jefferson index, 00 UTC +18* *(k) Modified Jefferson index, 00 UTC +24* *(l) Modified Jefferson index, 00 UTC +30*

(m) MUCAPE, 00 UTC +12 *(n) MUCAPE, 00 UTC +18* *(o) MUCAPE, 00 UTC +24* *(p) MUCAPE, 00 UTC +30*

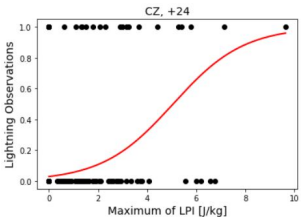
Figure 31: As Figure 30 but for the mean index value over a specific time and area.



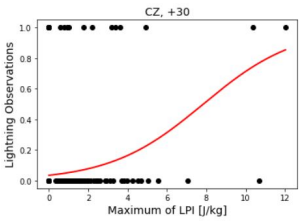
(a) LPI, 00 UTC +12



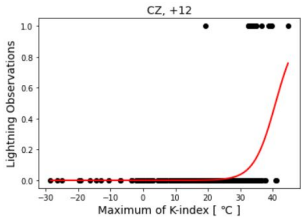
(b) LPI, 00 UTC +18



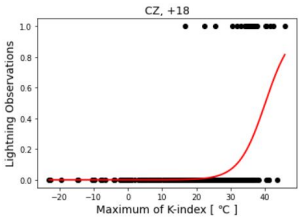
(c) LPI, 00 UTC +24



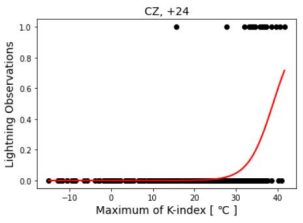
(d) LPI, 00 UTC +30



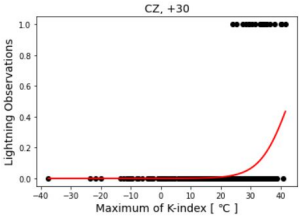
(e) K-index, 00 UTC +12



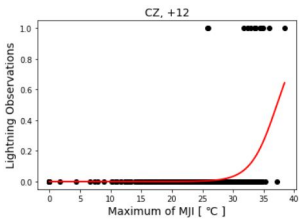
(f) K-index, 00 UTC +18



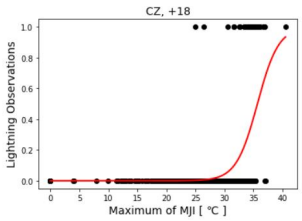
(g) K-index, 00 UTC +24



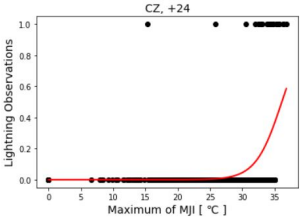
(h) K-index, 00 UTC +30



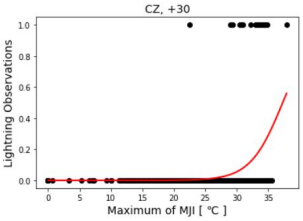
(i) Modified Jefferson index, 00 UTC +12



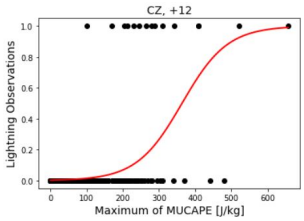
(j) Modified Jefferson index, 00 UTC +18



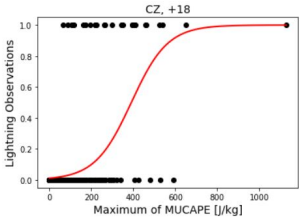
(k) Modified Jefferson index, 00 UTC +24



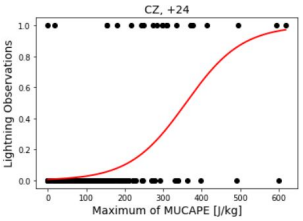
(l) Modified Jefferson index, 00 UTC +30



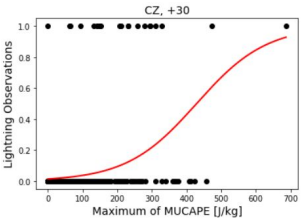
(m) MUCAPE, 00 UTC +12



(n) MUCAPE, 00 UTC +18

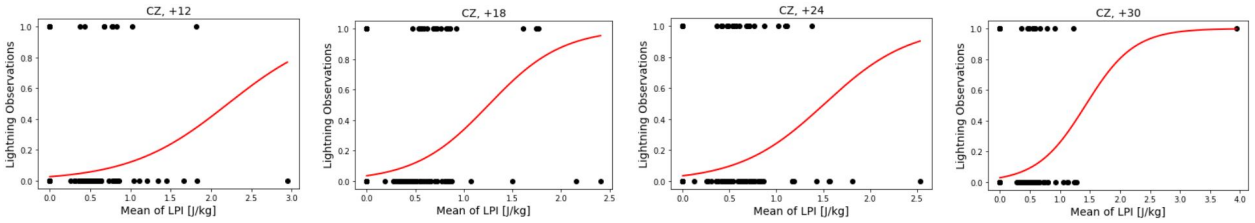


(o) MUCAPE, 00 UTC +24

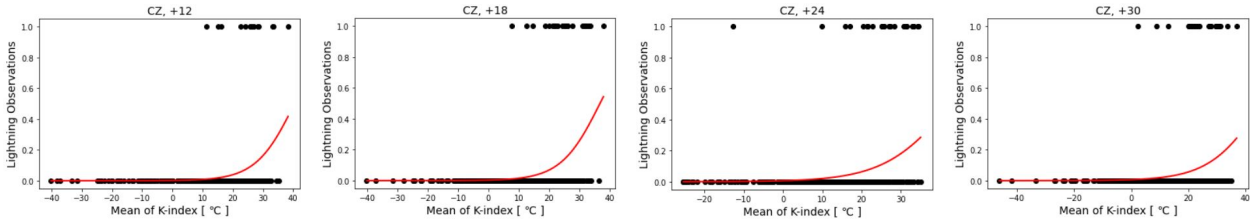


(p) MUCAPE, 00 UTC +30

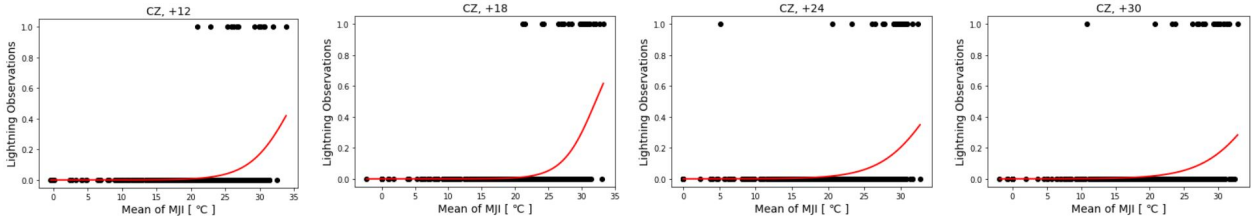
Figure 32: As Figure 30 but for the shore area (CZ).



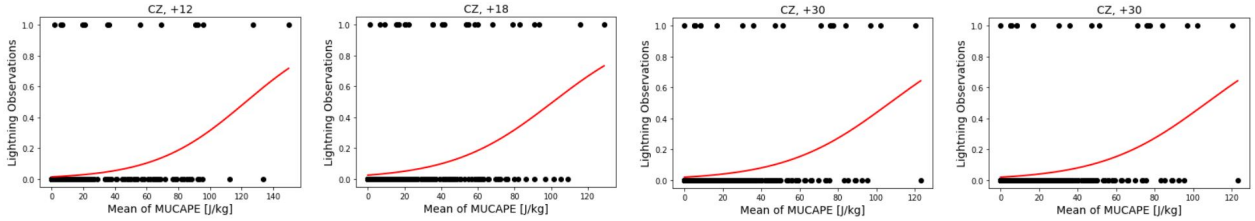
(a) LPI, 00 UTC +12 *(b) LPI, 00 UTC +18* *(c) LPI, 00 UTC +24* *(d) LPI, 00 UTC +30*



(e) K-index, 00 UTC +12 *(f) K-index, 00 UTC +18* *(g) K-index, 00 UTC +24* *(h) K-index, 00 UTC +30*

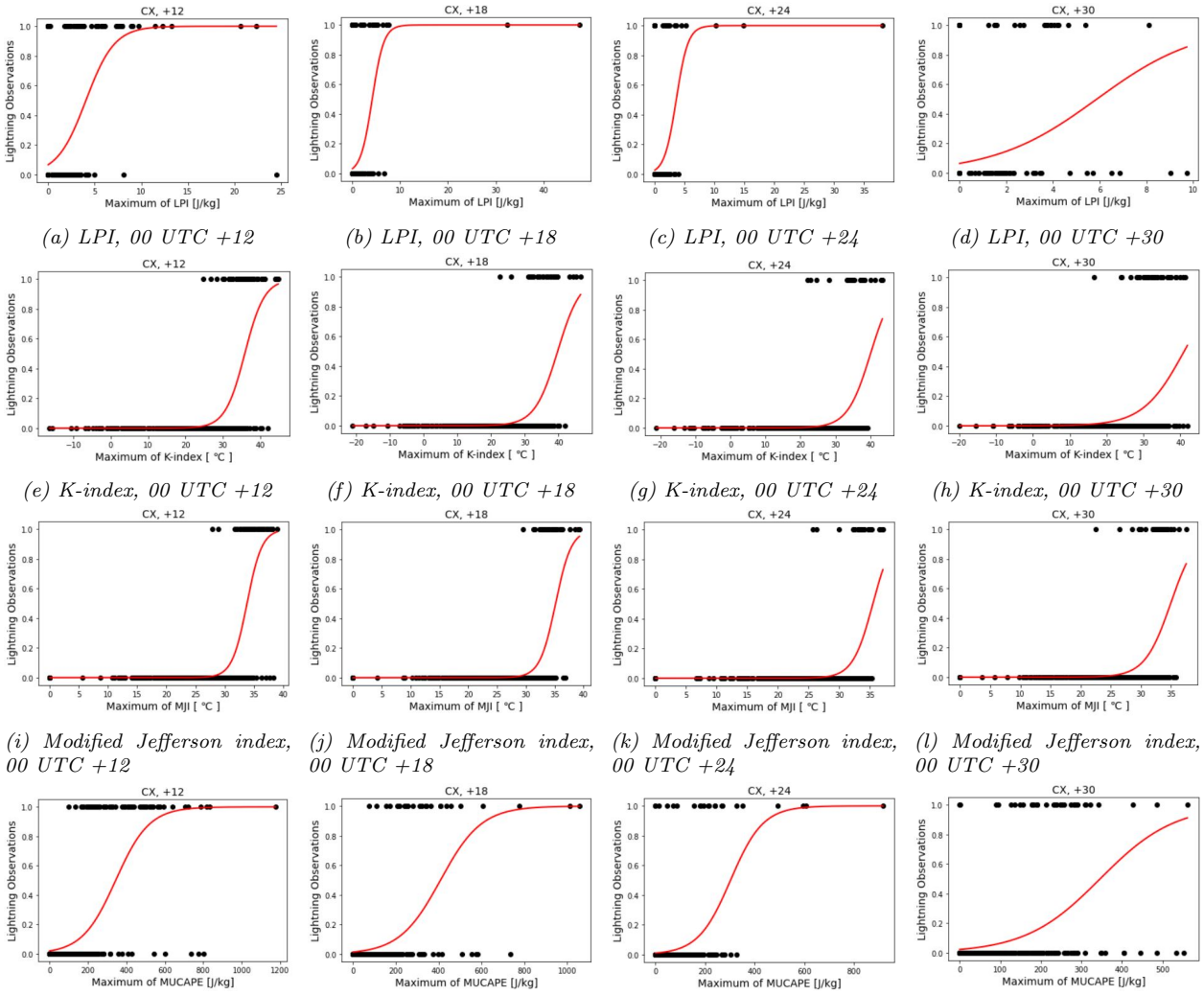


(i) Modified Jefferson index, 00 UTC +12 *(j) Modified Jefferson index, 00 UTC +18* *(k) Modified Jefferson index, 00 UTC +24* *(l) Modified Jefferson index, 00 UTC +30*



(m) MUCAPE, 00 UTC +12 *(n) MUCAPE, 00 UTC +18* *(o) MUCAPE, 00 UTC +24* *(p) MUCAPE, 00 UTC +30*

Figure 33: As Figure 30 but for the shore area (CZ) and for the mean index value over a specific time and area.



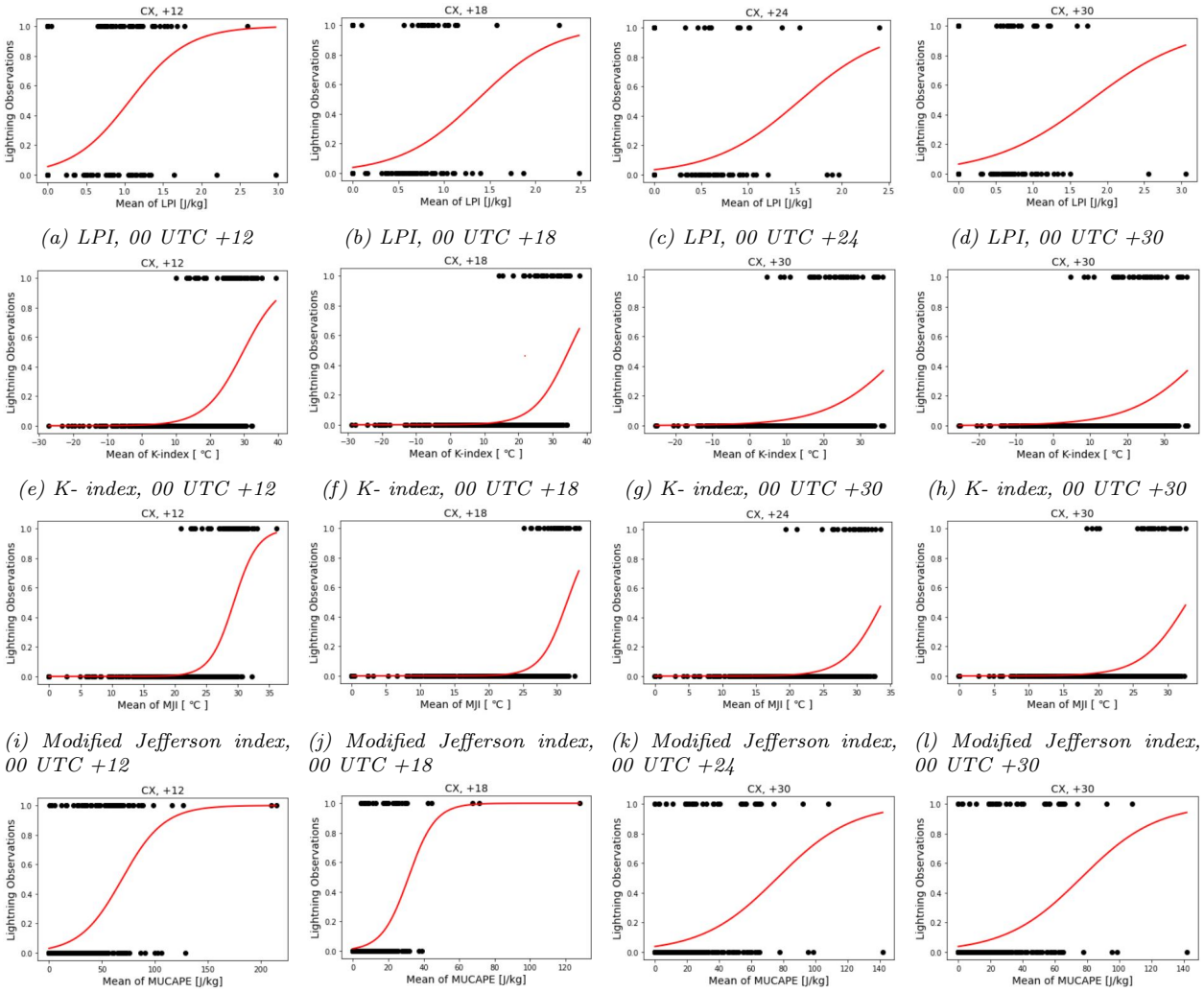
(a) LPI, 00 UTC +12 *(b) LPI, 00 UTC +18* *(c) LPI, 00 UTC +24* *(d) LPI, 00 UTC +30*

(e) K-index, 00 UTC +12 *(f) K-index, 00 UTC +18* *(g) K-index, 00 UTC +24* *(h) K-index, 00 UTC +30*

(i) Modified Jefferson index, 00 UTC +12 *(j) Modified Jefferson index, 00 UTC +18* *(k) Modified Jefferson index, 00 UTC +24* *(l) Modified Jefferson index, 00 UTC +30*

(m) MUCAPE, 00 UTC +12 *(n) MUCAPE, 00 UTC +18* *(o) MUCAPE, 00 UTC +24* *(p) MUCAPE, 00 UTC +30*

Figure 34: As Figure 30 but for the land area (CX).



(a) LPI, 00 UTC +12 *(b) LPI, 00 UTC +18* *(c) LPI, 00 UTC +24* *(d) LPI, 00 UTC +30*

(e) K- index, 00 UTC +12 *(f) K- index, 00 UTC +18* *(g) K- index, 00 UTC +30* *(h) K- index, 00 UTC +30*

(i) Modified Jefferson index, 00 UTC +12 *(j) Modified Jefferson index, 00 UTC +18* *(k) Modified Jefferson index, 00 UTC +24* *(l) Modified Jefferson index, 00 UTC +30*

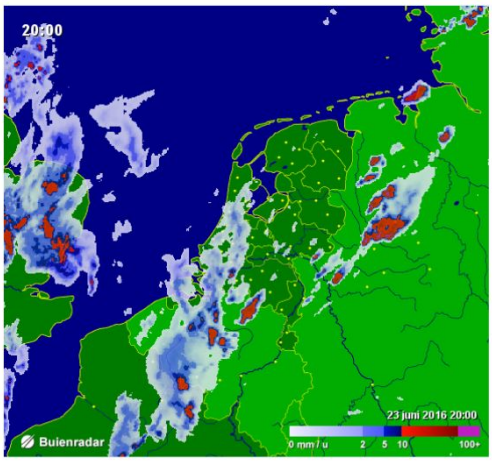
(m) MUCAPE, 00 UTC +12 *(n) MUCAPE, 00 UTC +18* *(o) MUCAPE, 00 UTC +30* *(p) MUCAPE, 00 UTC +30*

Figure 35: As Figure 30 but for the land area (CX) and for the mean index value over a specific time and area.

B - Radar observations of 23 June 2016

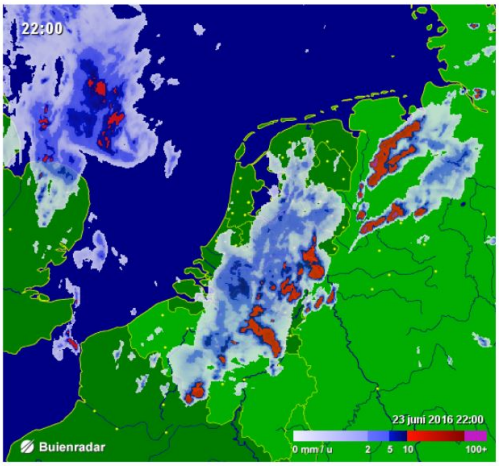
(a) Radar observations at 18.00 h

(b) Radar observations at 19.00 h

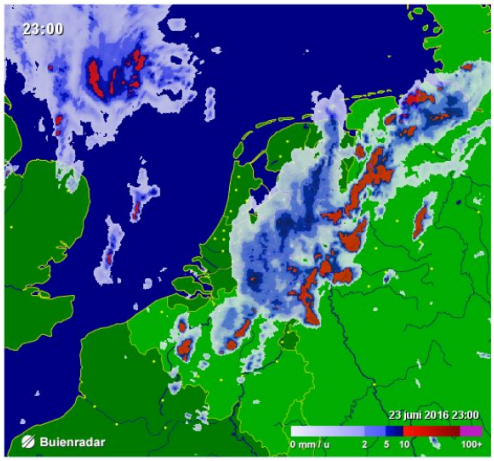


(c) Radar observations at 20.00 h

(d) Radar observations at 21.00 h



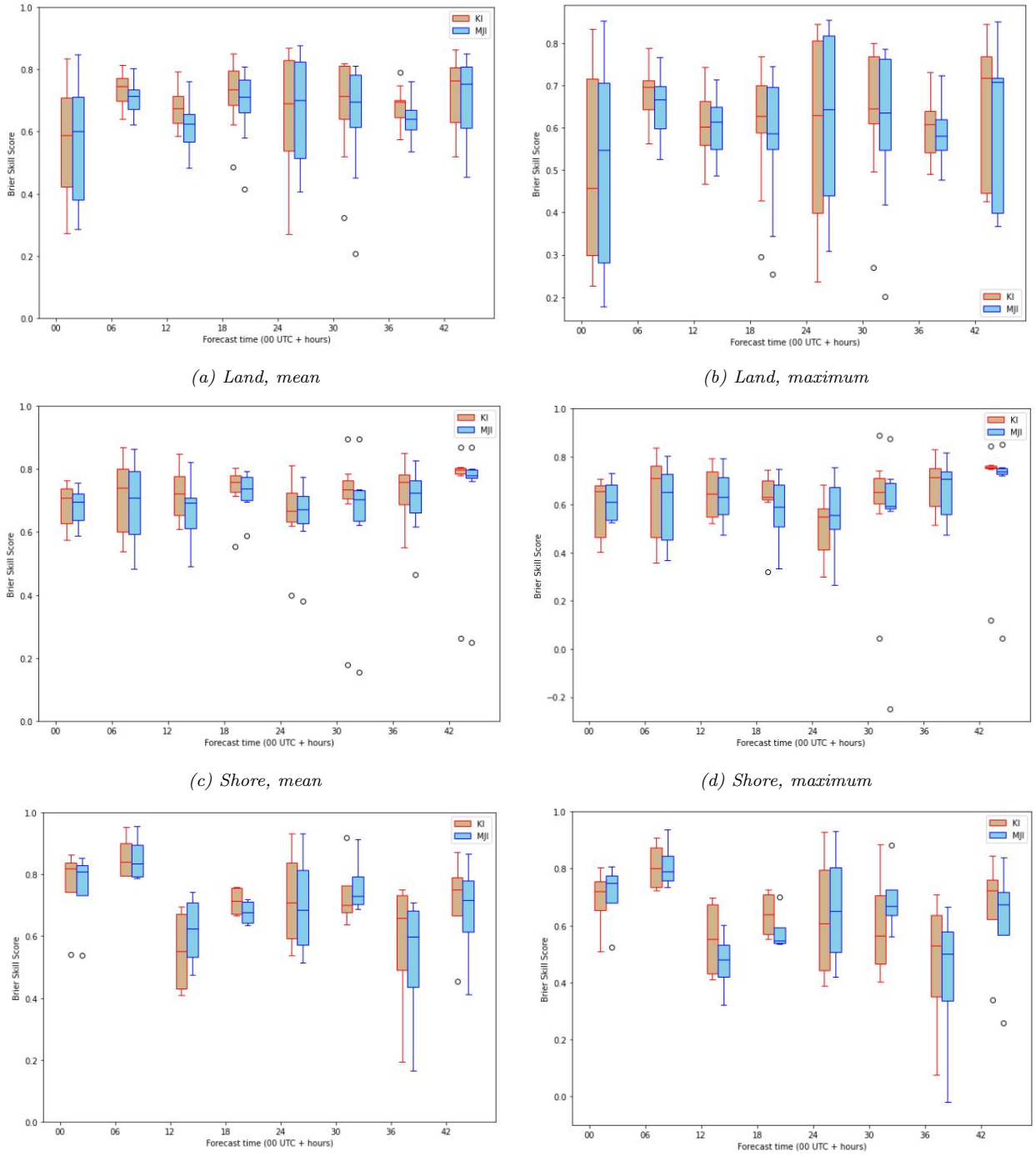
(e) Radar observations at 22.00 h



(f) Radar observations at 23.00 h

Figure 36: Historical radar observations done by Buienradar over the time span of interest (Buienradar, n.d.).

C - Box and Wishker plots



(a) Land, mean

(b) Land, maximum

(c) Shore, mean

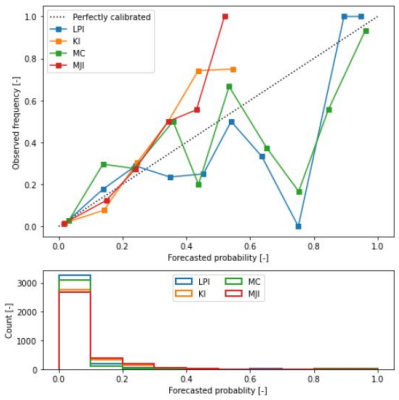
(d) Shore, maximum

(e) Sea, mean

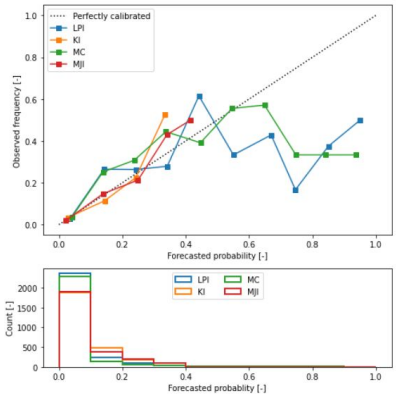
(f) Sea, maximum

Figure 37: Box and Whisker plots containing the Brier Skill Scores of only the K-index (red) and the Modified Jefferson index (blue). The Brier Skill Score are presented on the y-axis. The x-axis represents the lead times of the forecasts. Left are the mean values plotted and at the right the maximum values are plotted.

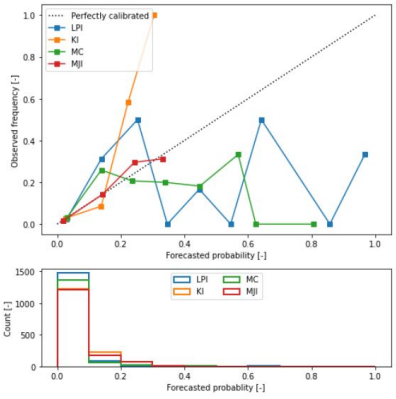
D - Reliability plots



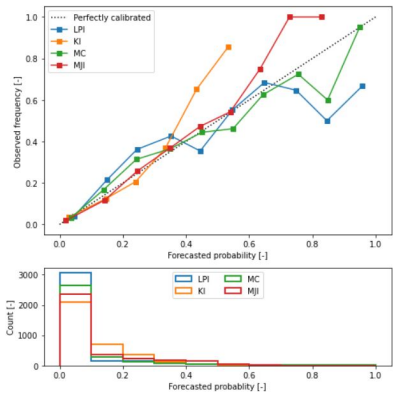
(a) Land, 00 UTC +00



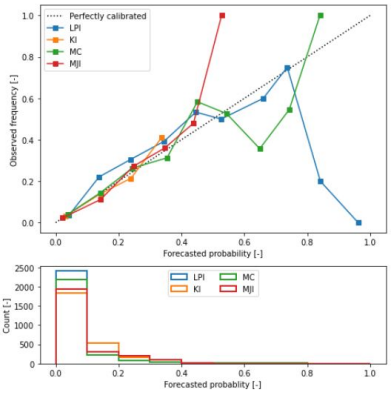
(b) Shore, 00 UTC +00



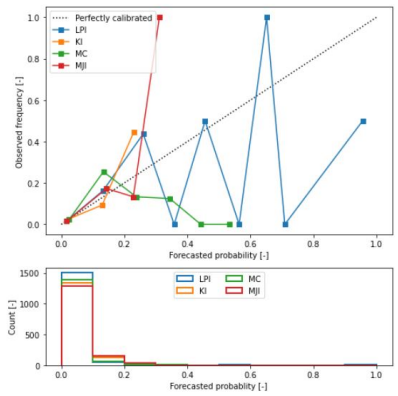
(c) Sea, 00 UTC +00



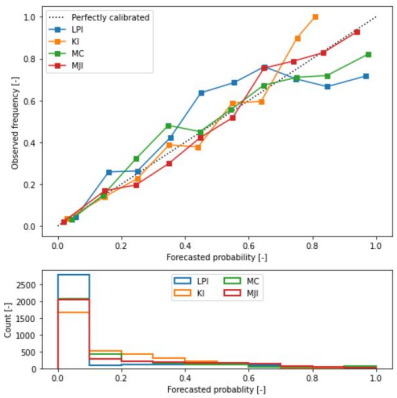
(d) Land, 00 UTC +06



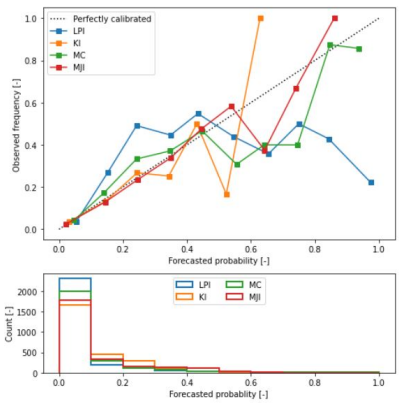
(e) Shore, 00 UTC +06



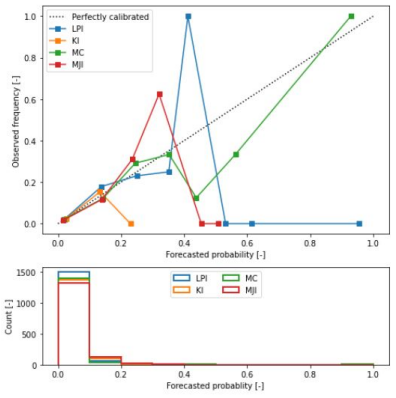
(f) Sea, 00 UTC +06



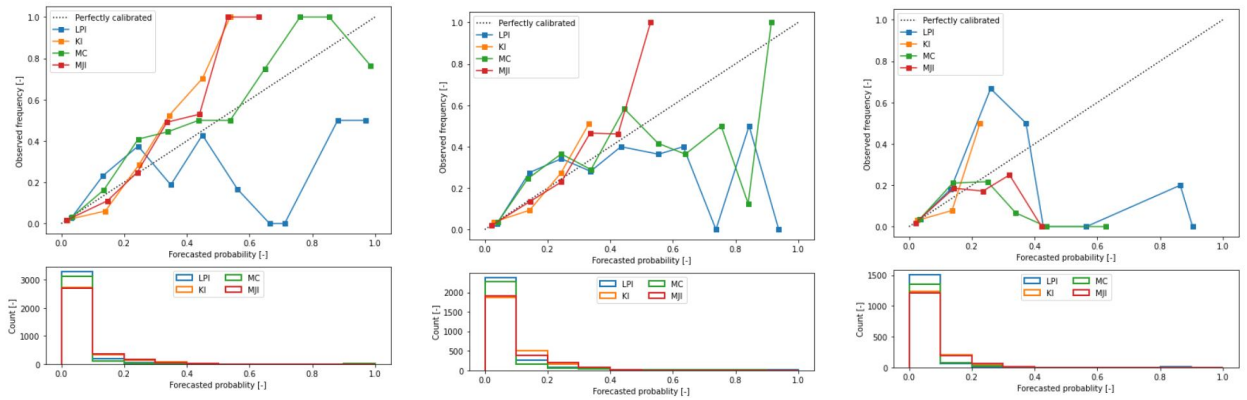
(g) Land, 00 UTC +12



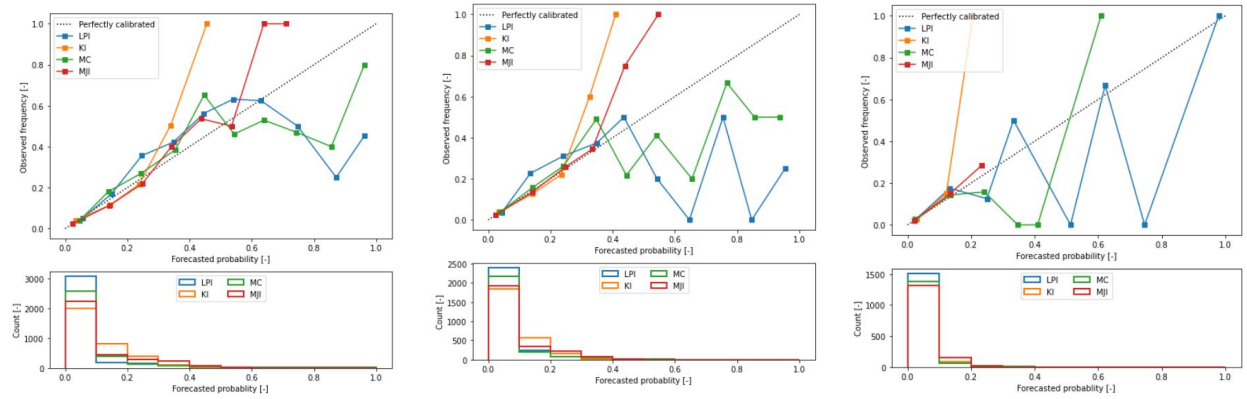
(h) Shore, 00 UTC +12



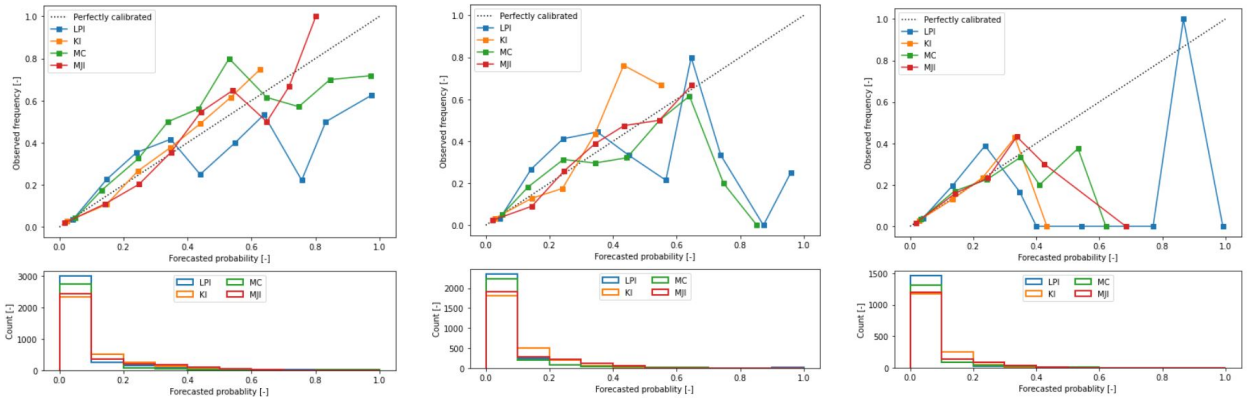
(i) Sea, 00 UTC +12



(a) Land, 00 UTC +24 *(b) Shore, 00 UTC +24* *(c) Sea, 00 UTC +24*



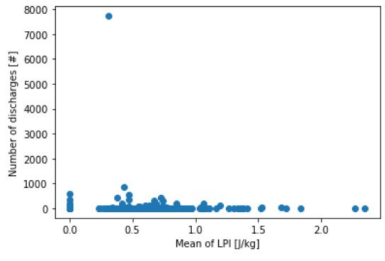
(d) Land, 00 UTC +30 *(e) Shore, 00 UTC +30* *(f) Sea, 00 UTC +30*



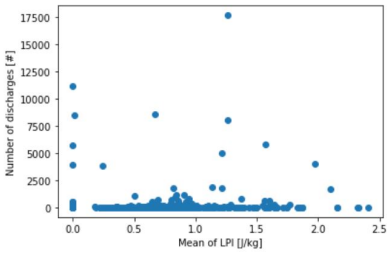
(g) Land, 00 UTC +42 *(h) Shore, 00 UTC +42* *(i) Sea, 00 UTC +42*

Figure 39: Reliability plots with on the x-axis the forecasted probability and on the y-axis the observed frequency. The black striped line is for a perfect reliable model. The lines with colours represent the indices. In the lower part of the plots the number of issued probabilities in each bin is shown.

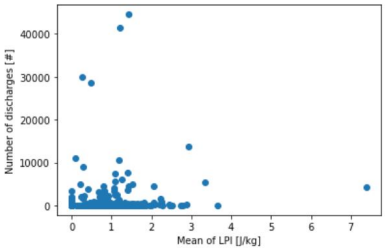
E - Scatter plots



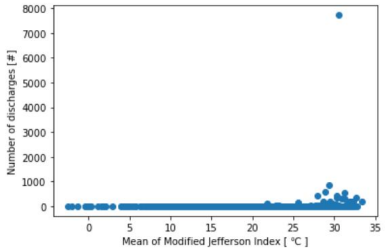
(a) LPI, 00 UTC +18, sea



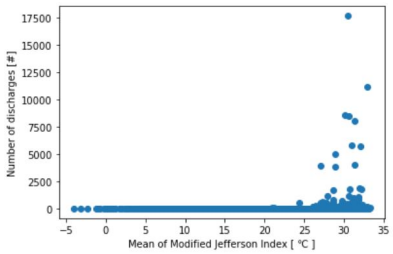
(b) LPI, 00 UTC +18, shore



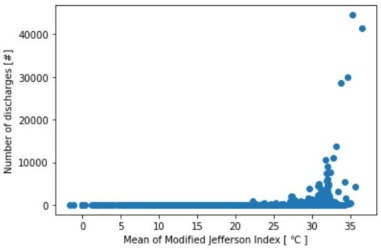
(c) LPI, 00 UTC +18, land



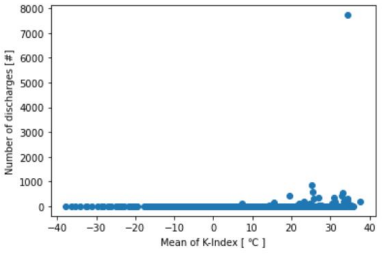
(d) MJI, 00 UTC +18, sea



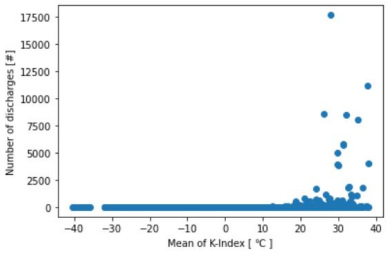
(e) MJI, 00 UTC +18, shore



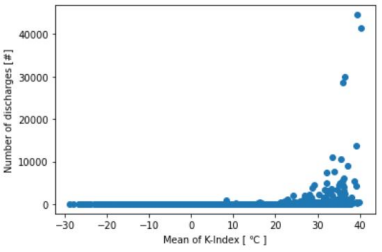
(f) MJI, 00 UTC +18, land



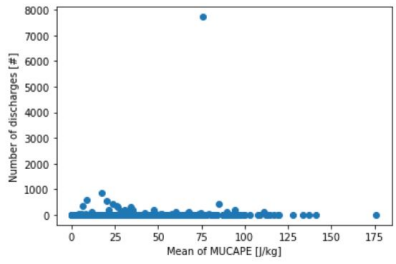
(g) KI, 00 UTC +18, sea



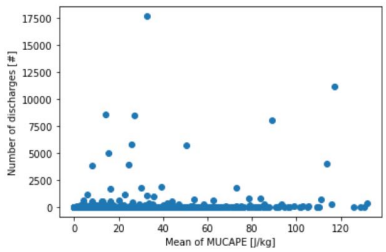
(h) KI, 00 UTC +18, shore



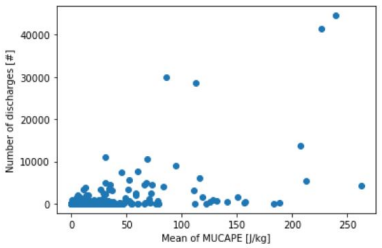
(i) KI, 00 UTC +18, land



(j) MUCAPE, 00 UTC +18, sea

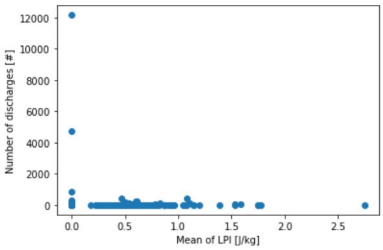


(k) MUCAPE, 00 UTC +18, shore

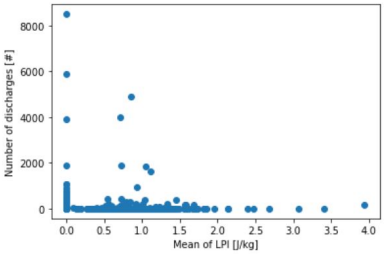


(l) MUCAPE, 00 UTC +18, land

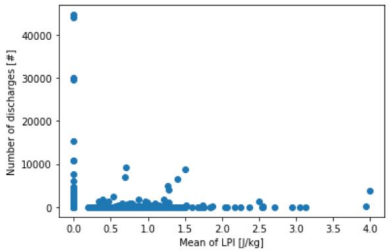
Figure 40: Scatter plots containing the number of discharges versus the mean value of the different indices over land, sea and shore for 00 UTC +18.



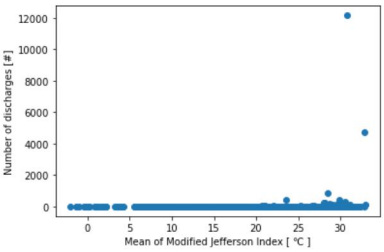
(a) LPI, 00 UTC +30, sea



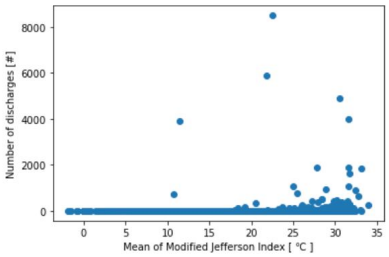
(b) LPI, 00 UTC +30, shore



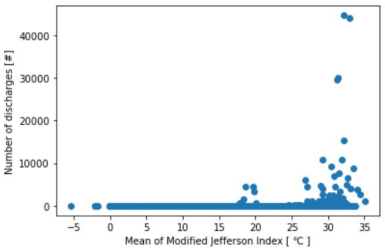
(c) LPI, 00 UTC +30, land



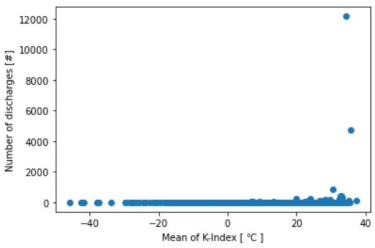
(d) MJI, 00 UTC +30, sea



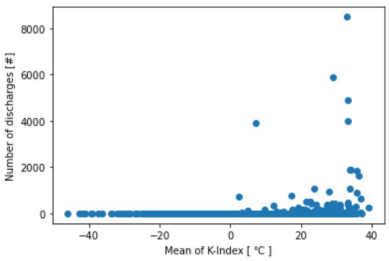
(e) MJI, 00 UTC +30, shore



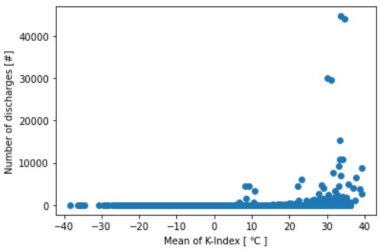
(f) MJI, 00 UTC +30, land



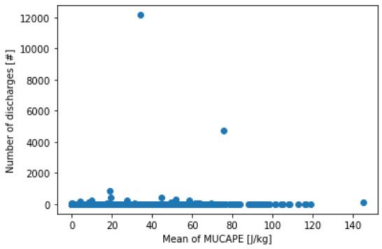
(g) KI, 00 UTC +30, sea



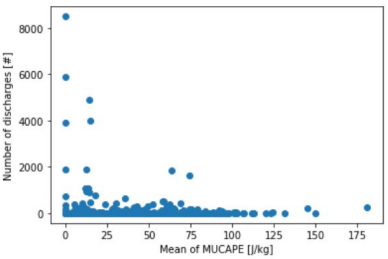
(h) KI, 00 UTC +30, shore



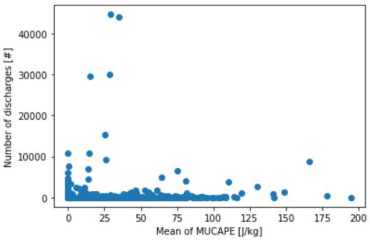
(i) KI, 00 UTC +30, land



(j) MUCAPE, 00 UTC +30, sea

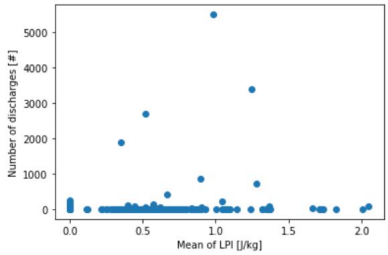


(k) MUCAPE, 00 UTC +30, shore

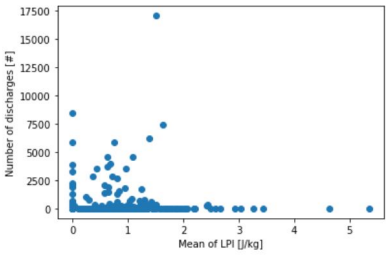


(l) MUCAPE, 00 UTC +30, land

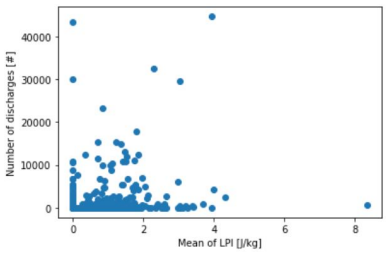
Figure 41: As Figure 40 but for 00 UTC +30.



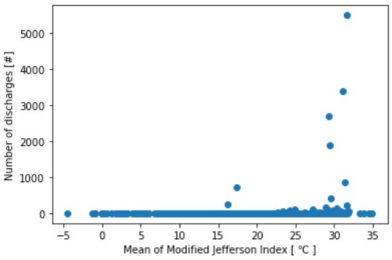
(a) LPI, 00 UTC +36, sea



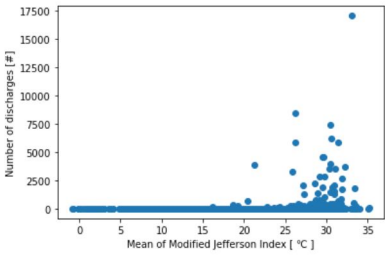
(b) LPI, 00 UTC +36, shore



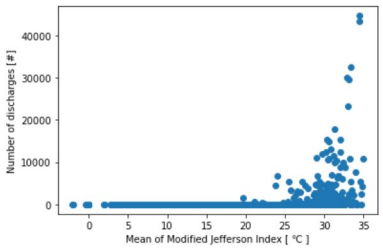
(c) LPI, 00 UTC +36, land



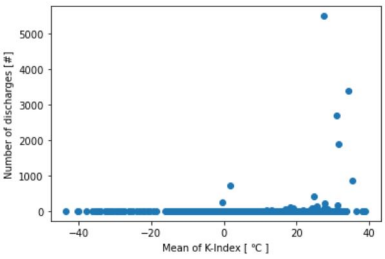
(d) MJI, 00 UTC +36, sea



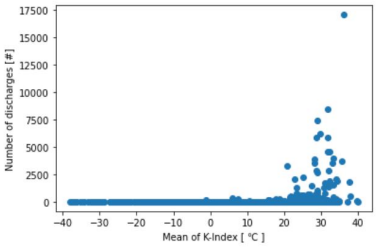
(e) MJI, 00 UTC +36, shore



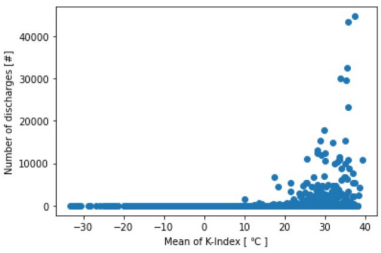
(f) MJI, 00 UTC +36, land



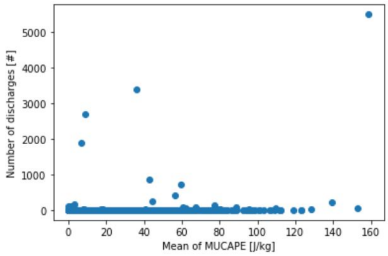
(g) KI, 00 UTC +36, sea



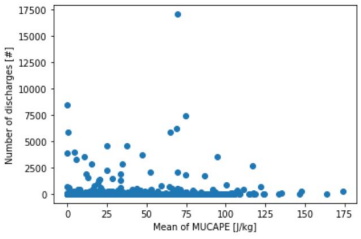
(h) KI, 00 UTC +36, shore



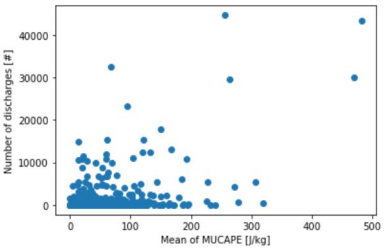
(i) KI, 00 UTC +36, land



(j) MUCAPE, 00 UTC +36, sea



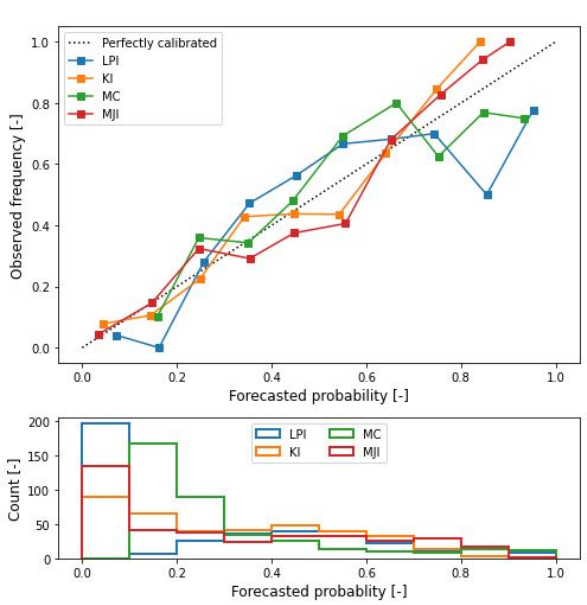
(k) MUCAPE, 00 UTC +36, shore



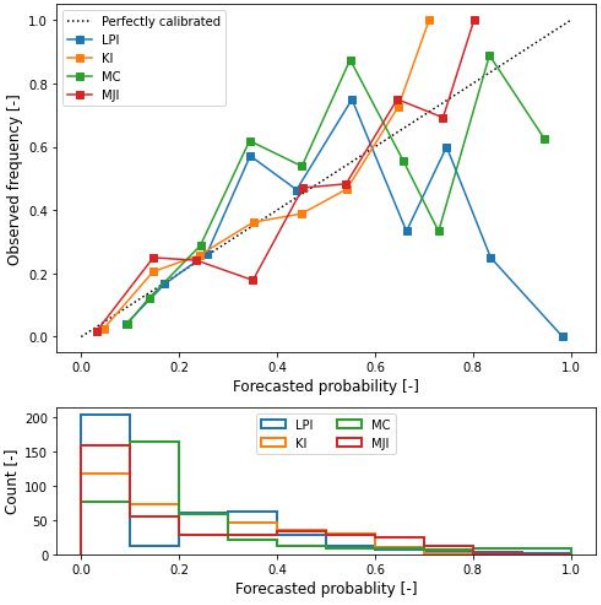
(l) MUCAPE, 00 UTC +36, land

Figure 42: As Figure 40 but for 00 UTC +36.

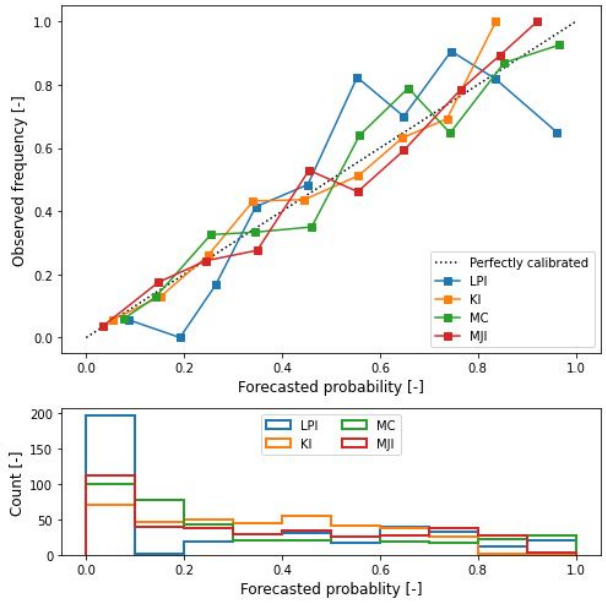
F - Reliability plots over whole area



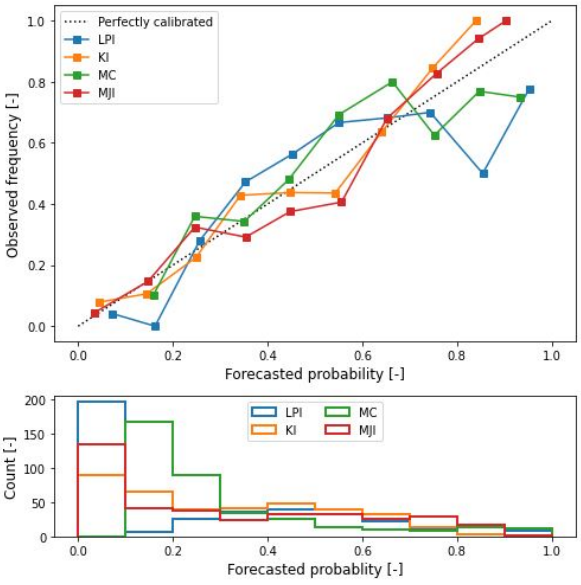
(a) 00 UTC +00



(b) 00 UTC +24



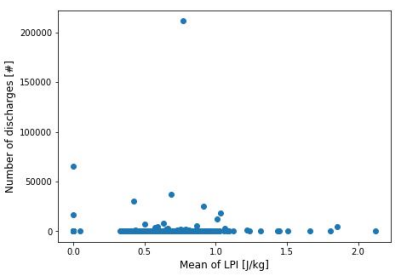
(c) 00 UTC +36



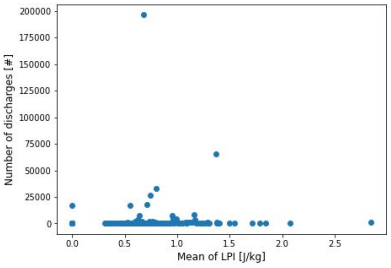
(d) 00 UTC +42

Figure 43: The reliability diagrams present the forecast probability given by the model on the y-axis versus the observed frequency on the x-axis. The lower part of the plot shows how often a model issues a probability in the same bin. The black striped line is a perfectly reliable model. The other lines represent the indices. This plot shows different lead times and represents the result over the whole area.

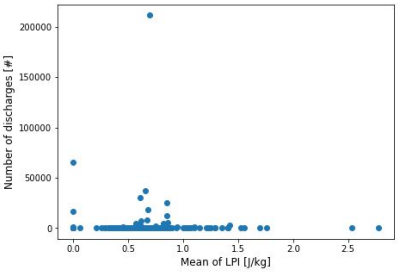
G -Scatter plots over whole area



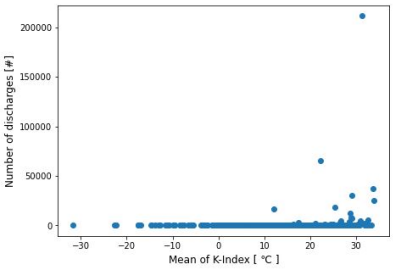
(a) LPI, 00 UTC +00



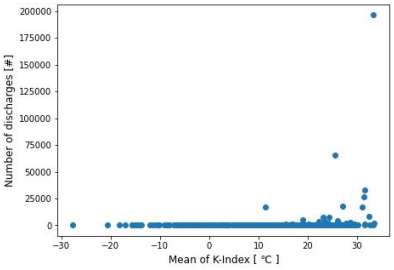
(b) LPI, 00 UTC +06



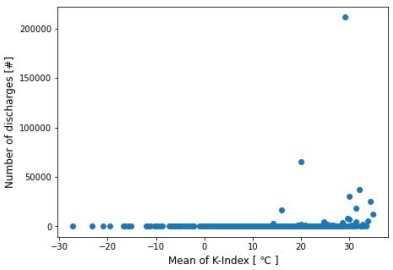
(c) LPI, 00 UTC +24



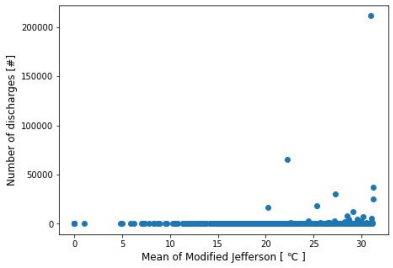
(d) K-index, 00 UTC +00



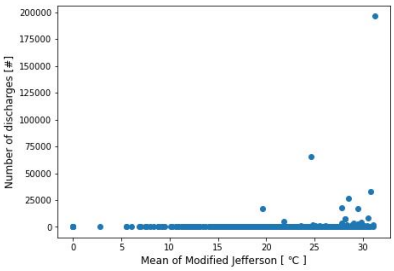
(e) K-index, 00 UTC +06



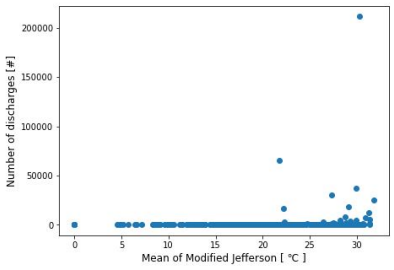
(f) K-index, 00 UTC +24



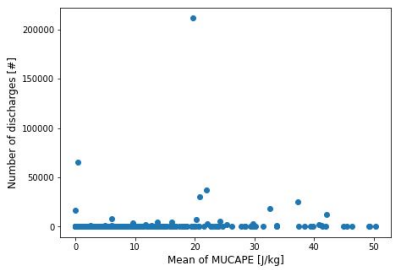
(g) Modified Jefferson index, 00 UTC +00



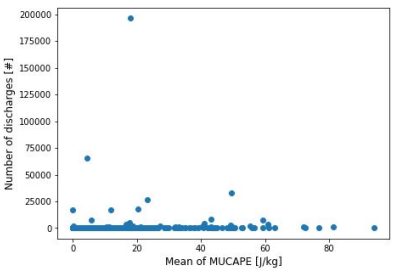
(h) Modified Jefferson index, 00 UTC +06



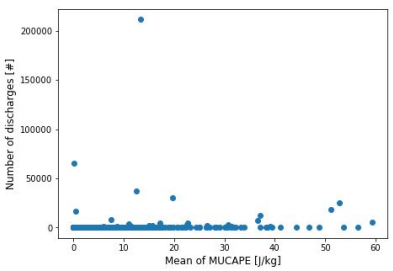
(i) Modified Jefferson index, 00 UTC +24



(j) MUCAPE, 00 UTC +00

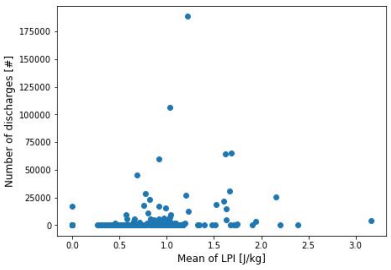


(k) MUCAPE, 00 UTC +06

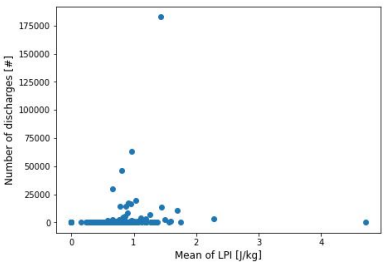


(l) MUCAPE, 00 UTC +24

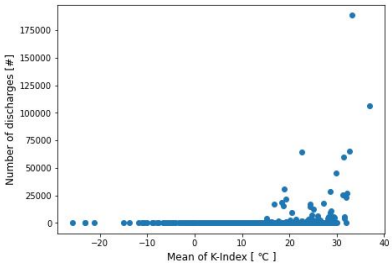
Figure 44: Scatter plots containing the number of discharges versus the mean value of the different indices over the whole area for three lead times (00 UTC +00, +06, +24).



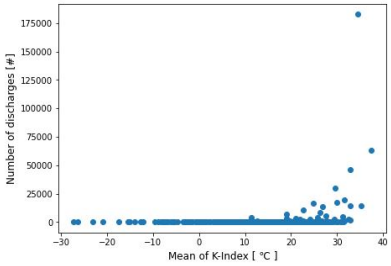
(a) LPI, 00 UTC +36



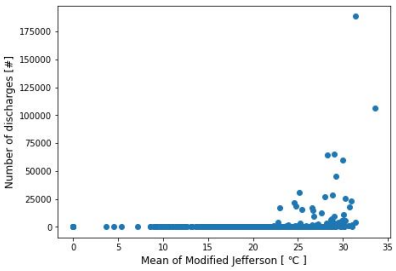
(b) LPI, 00 UTC +42



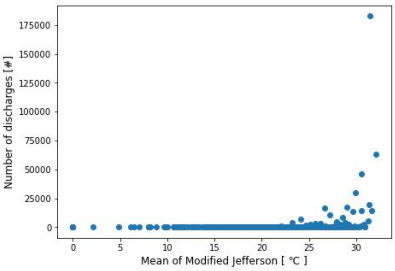
(c) K-index, 00 UTC +36



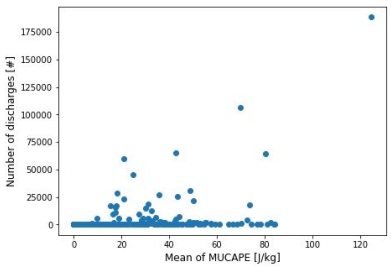
(d) K-index, 00 UTC +42



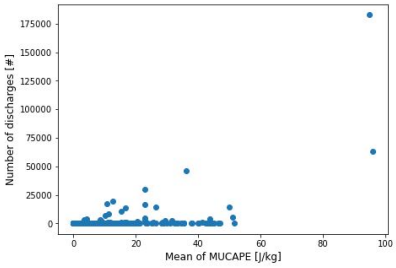
(e) Modified Jefferson index, 00 UTC +36



(f) Modified Jefferson index, 00 UTC +42



(g) MUCAPE, 00 UTC +36



(h) MUCAPE, 00 UTC +42

Figure 45: Scatter plots containing the number of discharges versus the mean value of the different indices over the whole area for two lead times (00 UTC +36, +42).

References

Arasa, N. (2010). *Development of an object oriented thunderstorm nowcasting system for barcelona airport* (Unpublished master's thesis).

AvalancheCanada. (2021). *Atmospheric lift.* Retrieved 2021-10-21, from `https://www.avalanche.ca/glossary?q=convect`

Barešić, (2021). Retrieved 2021-10-15, from `https://www.flickr.com/photos/183939796@N07/51282814772/in/photolist-2m8FQaA-r9pLHZ-dYReio-euTzD4-zSEBwX-shqjwv-QhRZqG-RUkcbn-euWHSE-RsKYEU-2jCQoFH-2hJr3hs-KbprTf-euTAca-X6KC8b-2kAeS1g-2hR4D4P-euTz4V-2hnvHi6-eayRvH-2kr9GnP-21sWXay-2gQCoSX-27ZpeTr-ZvafQ5-27D6FX6-2kiZk5d-f7FsjV-2fpyca8-2eiBbbW-q8utMZ-2isfsPz-SoXu-74tuvF-2m6QdDQ-7ZzqqF-2m6XLRo-2m6TWGA-2m6XLwF-2m6Zj2y-2m6UYkq-2m6Xg16-2kwAy4G-2m6QeHt-2m6Qe8L-2m6UZAr-2m6UZcf-2m6QevQ-2m6UZnq-2m6XMkp`

Bengtsson, L., Andrea, U., Aspelien, T., Batrak, Y., Calvo, J., De Rooy, W., ... Odegaard Koltzow, M. (2017, May). The harmonie-arome model configuration in the aladin-hirlam nwp system. *Monthly Weather Review*, *145*, 1919-1935. doi: https://doi.org/10.1175/MWR-D-16-0417.1

Brisson, E., Blahak, U., Lucas-Picher, P., Purr, C., & Ahrens, B. (2021, June). Contrasting lightning projection using the lightning potential index adapted in a convection-permitting regional climate model. *Climate Dynamics*. doi: 10.1007/s00382-021-05791-z

Buienradar. (n.d.). *Buienradar terugkijken.* Retrieved 2021-09-20, from `https://www.buienradar.nl/nederland/neerslag/buienradar-terugkijken/archief/201606231800`

Deierling, W., Petersen, W., Latham, J., Ellis, S., & H.J., C. (2008, August). The relationship between lightning activity and ice fluxes in thunderstorms. *Journal of Geophysical Research*, *113*, 1-17. doi: 10.1029/2007JD00970

De Roode, S. (2020, May). *Syllabus atmospheric physics.* Delft University of Technology.

De Rooy, W., & Siebesma, A. (2010). Analytical expressions for entrainment and detrainment in cumulus convection. *Quarterly Journal of the Royal Meteorological Society*, *136*, 1216 - 1227. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.640` doi: https://doi.org/10.1002/qj.640

De Vos, L. (2015). *Intercomparison of the flits and kldn with atdnet lightning detection systems in the netherlands, with a case study on the potential of infrasound lightning detection* (Unpublished master's thesis).

De Wit, R., & Tijm, S. (2010, March). Voorspelling van bliksemintensiteit met behulp van graupel. *Meteorologica*, *19*(2), 21-24.

Fiori, E., Lagasio, M., Parodi, A., Procopio, R., Smorgonskiy, A., Rachidi, F., & Diendorfer, G. (2016). Implementation and performance analysis of the lightning potential index as a forecasting tool..

Gregow, E., Lindfors, A., Van der Veen, S., Schoenach, D., De Haan, S., & Lindskog, M. (2020). The use of satellite and surface observations for initializingclouds in the harmonie nwp model. *Meteorological Applications*, 1-11. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/epdf/10.1002/met.1965?__cf_chl_jschl_tk__=pmd_Jx21m4ZjyDyAw9g2q3B3EkA7ZtlMDK35QGJ5AaubHNU-1635254310-0-gqNtZGzNAjujcnBszQ59` doi: 10.1002/met.1965

Greitemann, J. (2014). *Concept of separation of charges in lightning clouds.* Retrieved 2021-10-22, from `https://physics.stackexchange.com/questions/122210/concept-of-separation-of-charges-in-lightning-clouds`

Groot, E. (2019). *Probabilistic thunderstorm forecasts using statistical post-processing: Comparison of logistic regression and quantile regression forest and an investigation of physical predictors.* Retrieved from

https://dspace.library.uu.nl/handle/1874/383079

Haklander, A., & Van Delden, A. (2013). Thunderstorm predictors and their forecast skill for the netherlands. *Elsevier*, *67-68*, 273-299. Retrieved from https://webspace.science.uu.nl/~delde102/Thunderst.Predictors.pdf doi: 10.1016/S0169-8095(03)00056-5

Hallar, G. (2017). *Atmos5130, lecture 14.* https://home.chpc.utah.edu/ hallar/Thermo/Lectures/Lecture14.pdf.

HartvanNederland. (2021). *Onweer houdt huis in nederland: bomen uit de grond gerukt en overstromingen.* Retrieved 2021-06-18, from https://www.hartvannederland.nl/nieuws/weer/bomen-uit-de-grond-wateroverlast-nachtelijke-storm

Ivanov, S., Michaelides, S., Ruban, I., Charalambous, D., & Tymvios, F. (2021). Impact of radar data assimilation on simulations of precipitable water with the harmonie model: A case study over cyprus. *Atmospheric Research*, *253*, 105473. Retrieved from https://www.sciencedirect.com/science/article/pii/S0169809521000259 doi: https://doi.org/10.1016/j.atmosres.2021.105473

Jiang, X. (2020). *A brief introduction to uncertainty calibration and relability diagrams.* Retrieved 2021-10-01, from https://towardsdatascience.com/introduction-to-reliability-diagrams-for-probability-calibration-ed785b3f5d44

KNMI. (n.d.-a). *Uitleg over onweersdagen.* Retrieved 2021-10-5, from https://www.knmi.nl/kennis-en-datacentrum/uitleg/onweersdagen

KNMI. (n.d.-b). *Weermodellen.* Retrieved from https://www.knmi.nl/kennis-en-datacentrum/uitleg/weermodellen

KNMI. (2016). *Knmi neemt nieuw meetsysteem voor bliksem in gebruik.* Retrieved 2021-10-5, from https://www.knmi.nl/over-het-knmi/nieuws/knmi-neemt-nieuw-meetsysteem-voor-bliksem-in-gebruik

KNMI. (2020). *Knmi stopt met weermodel hirlam.* Retrieved from https://www.knmi.nl/over-het-knmi/nieuws/knmi-stopt-met-weermodel-hirlam

KNMI. (2021). *Uitleg over dampkring.* Retrieved from https://www.knmi.nl/kennis-en-datacentrum/uitleg/dampkring

Krol, M., Houweling, S., Bregman, B., Van den Broek, M., Segers, A., Velthoven, P., ... Dentener, F. (2005, February). The two-way nested global chemistry-transport zoom model tm5: Algorithm and applications. *Atmospheric Chemistry and Physics*, *5*. doi: 10.5194/acpd-4-3975-2004

Malcorps, H., & Agren, M. (2005). *Hirlam–aladin agreement.* Retrieved from http://www.hirlam.org/index.php/component/docman/doc_view/24-aladin-hirlam-a-cooperation-agreement?Itemid=70

Masson, V. (2013). The surfexv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. *Geoscientific Model Development*, *6*, 929-96.

Météorage. (n.d.). *Onweer in europa.* Retrieved 2021-10-5, from https://www.meteorage.com/nl/onweer-europa

Müller, G., & Floors, R. (2021). *Archive, daten: Gfs-modell des amerikanischen wetterdienstes.* Retrieved from https://www.wetterzentrale.de/reanalysis.php?map=1&model=avn&var=2&jaar=2016&maand=6&dag=23&h=0&nmaps=24

NationalWeatherService. (2007). *Chapter 7 stability and cloud development.* Retrieved 2021-10-21, from https://www.weather.gov/media/zhu/ZHU_Training_Page/clouds/stability_clouds/stability_clouds.pdf

Palmer, P. (2017). *The atmosphere: A very short introduction* (1st ed.). New York, USA: Oxford Univeristy press.

Pier, H. (n.d.). *Convective inhibition (cin): Definition role in forecasting.* Retrieved 2021-09-28, from https://study.com/academy/lesson/convective-inhibition-cin-definition-role-in-forecasting.html

Poelman, D. (2010). On the science of lightning: An overview. *Royal Meteorological Institute of Belgium Report*, *115*, 4-6.

Rakov, V., & Uman, M. (2003). *Lightning: Physics and effects.* New York: Cambridge University Press.

Roose, N. (2015). *Wat is een shelfcloud?* Retrieved 2021-09-30, from https://www.noodweer.be/wat-is-een-shelfcloud/

Saunders, C. (1993, April). A review of thunderstorm electrification processes. *Journal of Applied Meteorology*, 642-655. doi: https://doi.org/10.1175/1520-0450(1993)032⟨0642:AROTEP⟩2.0.CO;2

Schmeits, M. (2002). *Indices voor extreem convective situaties (fase i). de meest geschikte indicatoren voor onweer en zware neerslag.* Retrieved from http://info.knmi.nl/wm-am/stat/exp/schmeits/indecs.htm

Schmeits, M., Kok, C., Vogelezang, D., & Van Westrhenen, R. (2008, December). Probabilistic forecasts of (severe) thunderstorms for the purpose of issuing a weather alarm in the netherlands. *Weather and Forecasting*, *23*, 1253 - 1267. doi: https://doi.org/10.1175/2008WAF2007102.1

Siebesma, A., & De Roode, S. (2020). *Cumulus cloud dynamics.* Lecture slides.

Siebesma, A., De Rooy, W., Schmeits, M., De Vries, H., Lenderink, G., & Van Ulft, B. (2019). Improving hail and thunderstorm prediction for weather and climate. In *Abstract mso proposal.*

Siebesma, A., Soares, P., & Teixeira, J. (2007). A combined eddy-diffusivity mass-flux approach for the convective boundary layer. *Journal of Atmosperic Sciences*, *64*, 1230–1248. Retrieved from https://www.knmi.nl/kennis-en-datacentrum/publicatie/a-combined-eddy-diffusivity-mass-flux-approach-for-the-convective-boundary-layer doi: 10.1175/JAS3888.1

Sluijter, R. (2016). *Zware onweersbuien juni 2016 vol extremen.* Retrieved from https://www.knmi.nl/kennis-en-datacentrum/achtergrond/zware-onweersbuien-op-22-en-23-juni-vol-extremen

Sluijter, R. (2017). *Zeer zware onweersbuien op 19 juli 2017.* Retrieved from https://www.knmi.nl/kennis-en-datacentrum/achtergrond/zeer-zware-onweersbuien-op-19-juli-2017

Sokol, Z., & Minářová, J. (2020). Impact of 1- and 2-moment cloud microphysics and horizontal resolution on lightning potential index within cosmo nwp model. *Atmospheric Research*, *237*, 104862. Retrieved from https://www.sciencedirect.com/science/article/pii/S0169809519306982 doi: https://doi.org/10.1016/j.atmosres.2020.104862

StatisticsSolutions. (2021). *What is logistic regression?* Retrieved 2021-10-19, from https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/what-is-logistic-regression/

Stojiljković, M. (2019). *Logistic regression in python.* Retrieved 2021-10-19, from https://realpython.com/logistic-regression-python/

Stull, R. (2020). *Lightning and thunder.* Retrieved 2021-09-30, from https://geo.libretexts.org/Bookshelves/Meteorology_and_Climate_Science/Book%3A_Practical_Meteorology_(Stull)/15%3A_Thunderstorm_Hazards/15.02%3A_Section_3-

UniveristyOfIllinois. (2010). *Updrafts/downdrafts, rising and sinking air.* Retrieved 2021-09-29, from http://ww2010.atmos.uiuc.edu/(Gh)/guides/mtr/svr/comp/up/home.rxml

Van den Brink, H. (2018). *Extreme wind en druk in de ecmwf seizoensverwachtingen* (Report No. TR-364). KNMI.

Wilks, D. (2011). *Statistical methods in the atmospheric sciences* (3rd ed.). Oxford, UK: Elsevier.

Yair, Y., & Lynn, B. (2008, November). Lightning potential index: A new tool for predicting the lightning density and the potential for extreme rainfall. *Geophysical Researsch Abstracts*, *10*.

Yair, Y., Lynn, B., Price, C., Kotroni, V., Lagouvardos, K., Morin, E., ... Llasat, M. (2010, February). Predicting the potential for lightning activity in mediterranean storms based on the weather research and forecasting (wrf) model dynamic and microphysical fields. *Journal of Geophysical Research*, *115*. doi: 10.1029/2008JD010868