

Ⓞ Detecting Severe Weather Trends Using an Additive Regressive Convective Hazard Model (AR-CHaMo)

ANJA T. RÄDLER

Munich Re, and Ludwig-Maximilians University, Munich, Germany

PIETER GROENEMEIJER

European Severe Storms Laboratory, Wessling, Germany

EBERHARD FAUST

Munich Re, Munich, Germany

ROBERT SAUSEN

Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany

(Manuscript received 10 May 2017, in final form 2 December 2017)

ABSTRACT

A statistical model for the occurrence of convective hazards was developed and applied to reanalysis data to detect multidecadal trends in hazard frequency. The modeling framework is based on an additive logistic regression for observed hazards that exploits predictors derived from numerical model data. The regression predicts the probability of a severe hazard, which is considered as a product of two components: the probability that a storm occurs and the probability of the severe hazard, given the presence of a storm [$P(\text{severe}) = P(\text{storm}) \times P(\text{severe}|\text{storm})$]. The model was developed using lightning data as an indication of thunderstorm occurrence and hazard reports across central Europe. Although it uses only two predictors per component, it is capable of reproducing the observed spatial distribution of lightning and yields realistic annual cycles of lightning, hail, and wind fairly accurately. The model was applied to ERA-Interim (1979–2016) across Europe to detect any changes in lightning, hail, and wind hazard occurrence. The frequency of conditions favoring lightning, wind, and large hail has increased across large parts of Europe, with the exception of the southwest. The resulting predicted occurrence of 6-hourly periods with lightning, wind, and large hail has increased by 16%, 29%, and 41%, respectively, across western and central Europe and by 23%, 56%, and 86% across Germany and the Alps during the period considered. It is shown that these changes are caused by increased instability in the reanalysis rather than by changes in midtropospheric moisture or wind shear.

1. Introduction

Severe convective weather is an important hazard to both people and property. Estimated mean annual direct economic losses of severe storms over the last decade (2007–16) are \$19.6 billion (U.S. dollars) and EUR 3.8 billion. The costliest year in the United States was 2011,

with ~\$40 billion, and in Europe it was 2013, with ~EUR 8 billion (J. Eichner 2017, personal communication).

Relative to other types of extreme weather phenomena, the confidence is low that changes in the occurrence of convective hazards can be attributed to climate change (National Academies of Sciences Engineering and Medicine 2016). The relationship between the convective hazards and climate change has become a topic of study only relatively recently. A natural starting

Ⓞ Denotes content that is immediately available upon publication as open access.

Corresponding author: Anja T. Rädler, anja.raedler@physik.lmu.de

Publisher's Note: This article was revised on 19 March 2018 to include the open access designation that was missing when originally published.

point for studying this relationship is to investigate changes that have occurred in recent times. Authors have noted, however, that observational datasets of severe weather, such as tornadoes, hail, and wind reports, are often inhomogeneous in both time and place (Verbout et al. 2006; Doswell 2007; Brooks et al. 2014; Groenemeijer and Kühne 2014; Allen and Tippett 2015), to an extent that often prevents the detection of trends. It is, therefore, necessary to develop methods for using more homogeneous datasets, such as reanalyses, for this purpose. Likewise, predictions for the future also require the extraction of convective hazard occurrence from climate model simulations.

Most reanalyses and climate models have resolutions that are too coarse to fully resolve severe weather hazards. Several approaches have been taken into account to link the occurrence of small-scale severe weather phenomena to spatial scales resolved in those models. These include dynamic downscaling (Trapp et al. 2007; Gensini and Mote 2014, 2015) and the identification of environmental conditions supportive of severe events (Brooks et al. 2003; Trapp et al. 2009; Diffenbaugh et al. 2013; Allen et al. 2015; Púčik et al. 2017; Viceto et al. 2017). These environmental approaches typically define supportive conditions by formulating thresholds for quantities deemed necessary for hazard development. For severe convective storms, these include instability and wind shear, quantities that may be considered ingredients for organized deep moist convection (Johns and Doswell 1992). In addition to these, the probability of convective initiation must be assessed because the presence of instability does not guarantee storm initiation. The ingredients of instability and wind shear are necessary but not sufficient. For instance, Groenemeijer et al. (2017) found that in the U.S. Great Plains, severe storms occur on only 20% of all days that have sufficient CAPE and wind shear to sustain them. Across Europe, this value is mostly between 60% and 80%. Thus, the probability of convective initiation must be assessed in addition to these ingredients. Therefore, multiple authors (Trapp et al. 2007, 2009; Tippett et al. 2012; Púčik et al. 2017) have used the occurrence of modeled (convective) precipitation as an additional criterion. Van Klooster and Roebber (2009) used a neural network approach, while Taylor et al. (2012) and Diffenbaugh et al. (2013) considered convective inhibition.

These studies count the number of occurrences of situations that exceed the formulated thresholds. In other words, environments are classified in a binary rather than a probabilistic way. Alternatively, a statistical model can be used that assigns any environment an occurrence probability of a convective hazard between 0 and 1.

For instance, Kapsch et al. (2012) developed a Bayesian model for hail-damage days based on objective

weather type (OWT) classifications of the tropospheric flow pattern. OWTs were also used by Mohr et al. (2015), who developed a linear logistic regression based on various meteorological parameters [minimum temperature, surface-based lifted index (LI), and 2-m temperature] and the OWTs (Bissolli and Dittmann 2001) in order to model the frequency of hail potential in Germany in the future, using regional climate model simulations. Although the occurrence of OWTs correlates with the physical ingredients needed for severe convection, identifying the ingredients in simulated weather patterns appears to be only an indirect way of modeling hazard occurrence. Focusing on the physical ingredients rather than on indirect metrics like OWTs allows the model to be used in other regions and seasons, provided that the environmental factors are sufficiently broad brushed to capture these features. For weather forecasting purposes, similar methods have been developed; the difference is that they are applied to numerical weather predictions rather than to climate models or reanalyses. For instance, Schmeits et al. (2005) developed linear logistic regression equations for the probability of mid-April–mid-October thunderstorms in the Netherlands using lightning data and traditional thunderstorm indices. It has not been demonstrated that the linear logistic regression is the most suitable statistical model, which motivated the development of an additive logistic regression. Further, Hart and Cohen (2016) developed a statistical severe convection risk assessment model, which uses the observed frequencies of hazards for particular values of predictor parameters without fitting these data to any particular function. Cintineo et al. (2014) created an empirical Bayesian model for assessing the severe weather potential of developing convection, using predictors based on radar and satellite data.

In the present study, we describe the development of statistical models for convective hazard probability and use them to study hazard trends across Europe during the last 38 years by applying them to a reanalysis dataset. The model is an additive logistic regressive convective hazard model (AR-CHaMo). As we will show, the additive nature of the model fits the observational data better than a linear model would. The full hazard model is a product of two components: 1) the probability of lightning, indicating the occurrence of a convective storm, and 2) the probability of a hazard, given that lightning occurs, that is, $P(\text{hazard}) = P(\text{storm}) \times P(\text{hazard}|\text{storm})$. For each hazard or combination of hazards, different models can be created. Both probabilities are estimated by parameters with a proven physical and empirical association with lightning probability and hazard incidence, respectively. The choice of predictors is based on prior work by Westermayer et al. (2017), who investigated the environmental conditions

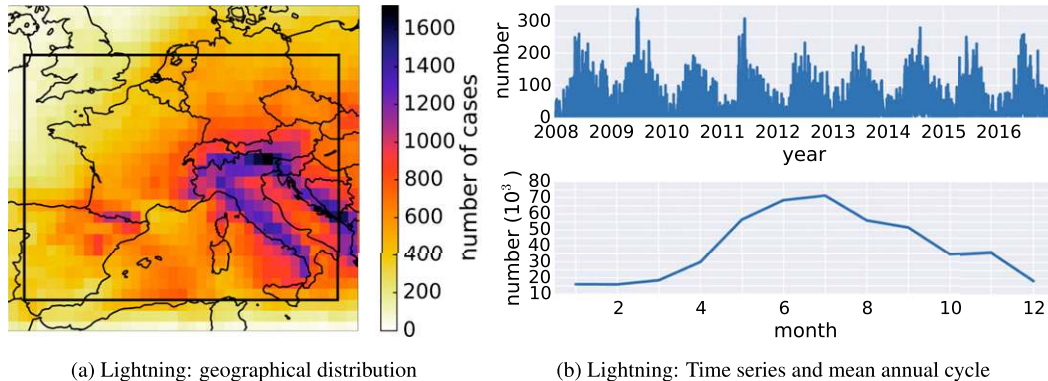


FIG. 1. Sum of EUCLID lightning cases in Europe between 2008 and 2016. A lightning case is defined as a 6-h period with at least two lightning strikes within a $0.75^\circ \times 0.75^\circ$ grid. (a) Geographical distribution of EUCLID lightning cases. (b) Top: time series of lightning cases; bottom: mean annual cycle of lightning cases. The analysis domain in (b) is restricted to data within the black box of (a). This domain is referred to as W&CEurope.

for thunderstorms in reanalysis data, and by Púčík et al. (2015), who studied the environments for large hail and severe wind gusts, given that a storm occurs.

In this article, we model the wind and large hail hazards using a reanalysis dataset across Europe covering data since 1979, with the aim of demonstrating the application of AR-CHaMo. The method can also be applied to other hazards across other areas and can be used with various datasets, including numerical weather prediction models. The primary questions we will answer in this study are as follows:

1. How well can we reproduce the observed climatology of convective hazards and lightning in Europe using AR-CHaMo?
2. Can we distinguish between hazards and account for convective initiation separately?
3. Using AR-CHaMo, are there any detectable trends in hazard occurrence across Europe during the last 38 years?

In section 2, we introduce the reanalysis, lightning, and hazard data used in this study. Section 3 explains the methods and development of the model. Section 4 comprises the results of the model application to reanalysis data. In section 5, the driving factors for the modeled trends are analyzed. Section 6 presents conclusions and a discussion of the results.

2. Data

a. EUCLID lightning data

Cloud-to-ground lightning data from the European Cooperation for Lightning Detection network (EUCLID) serve as indicators of the occurrence of deep moist convection. EUCLID is a cooperation among 23 European countries that provides lightning measurements

from its national networks (Diendorfer et al. 2010; Schulz et al. 2016; Poelman et al. 2016). In this study, a lightning case is defined as a 6-h period (namely, 0000–0600, 0600–1200, 1200–1800, and 1800–0000 UTC) with at least 2 lightning strikes within the same $0.75^\circ \times 0.75^\circ$ grid. Figure 1a depicts the geographical distribution of EUCLID lightning cases between 2008 and 2016. The peak region for lightning case occurrence in Europe is located in north Italy. Most lightning cases occur along the Apennine Mountains and the east coast of the Adriatic Sea. Within Germany, there is a clear south–north gradient toward fewer lightning cases in the north. The time series of lightning cases is illustrated in the top panel of Fig. 1b, which is restricted to the western and central European domain of Fig. 1a, referred to as W&CEurope. The years 2009 and 2011 have a particularly high frequency of lightning cases during the spring and summer months. The mean annual cycle (Fig. 1b, bottom) peaks in July. The second increase in November is caused by the high activity in the Mediterranean region. The EUCLID lightning data are only used to distinguish between cases with and without the occurrence of deep moist convection. To account for the severity of the identified convection, hazard reports are incorporated, as described in the next section.

b. ESWD hazard reports

The European Severe Weather Database (ESWD) is a collection of quality-controlled reports of severe convective events throughout Europe (Dotzek et al. 2009; Groenemeijer and Kühne 2014). For this study, the two hazards of hail ≥ 2 cm and severe wind gusts ≥ 25 m s⁻¹ have been taken into account. Hazard reports from the ESWD were accumulated over 6 h and on a $0.75^\circ \times 0.75^\circ$ horizontal grid. Figure 2 shows the geographical distribution of the number of 6-h periods with hail (Fig. 2b) and wind (Fig. 2c) between 2008 and 2016

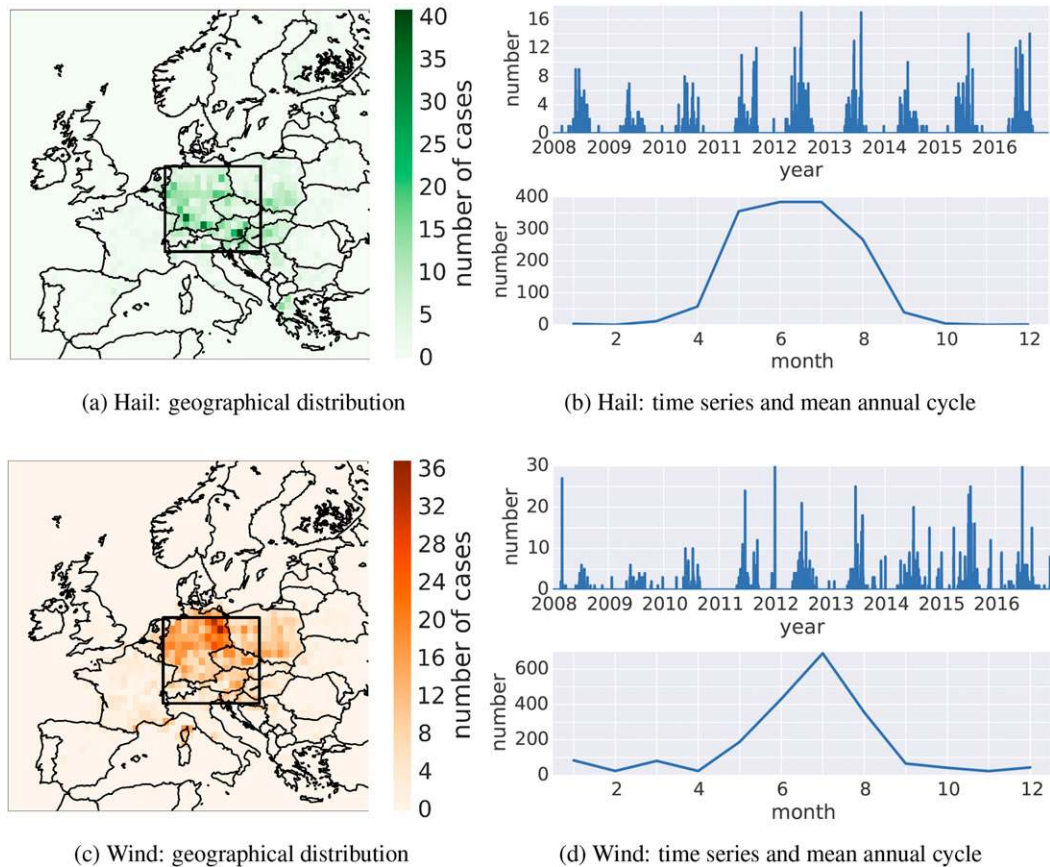


FIG. 2. ESWD hazard cases for time period 2008–16 where lightning occurred. (a) Hail ≥ 2 cm cases. (b) Top: time series of hail cases; bottom: mean annual cycle of hail cases. (c) Wind ≥ 25 m s⁻¹ cases. (d) As in (b), but for wind cases. The analysis domains in (b) and (d) are restricted to data within the black box shown in both (a) and (c). This domain is referred to as G&Alps.

where lightning occurred. The criterion of a hail (wind) case is that at least one report of hail (wind) occurred in conjunction with lightning within a grid box during the 6-h period. Hail and wind reports in a grid box that occurred without lightning are ignored because we want to focus on convective hazards only. Especially for wind, we do not want to include winter storms that occur without embedded convection. The highest density of reports is in Germany and Austria, which is partly because of the inhomogeneous reporting rate (Groenemeijer and Kühne 2014). The reporting rates over France and central Italy have not been very good throughout the period, although they have improved recently. This is primarily due to the number of partners of the ESWD network in each country and how active they are. These partners are usually networks of storm spotters, as well as individuals. The highest reporting rates are in countries in which the European Severe Storms Laboratory has had long-term cooperation, such as Germany, Austria, Poland, Czech

Republic, Switzerland, Slovakia, Croatia, Hungary, and Slovenia. Therefore, only a small domain (referred to as G&Alps; 45.75°–54°N, 6°–18.75°E) covering this area was selected to develop the statistical models (black boxes in Figs. 2a,c). East of the G&Alps domain, the detection efficiency of the EUCLID network decreases rapidly. Figures 2b and 2d represent the time series (top panel) and the mean annual cycles (bottom panel) of hail and wind cases and are restricted to the G&Alps domain. The annual cycle of hail (Fig. 2b, bottom) peaks in June and July, and that of wind cases (Fig. 2d) peaks in July. The slightly elevated activity in December and January is caused by thunderstorms occurring with synoptic-scale storm systems in winter.

c. ERA-Interim data

We have used the ERA-Interim global atmospheric reanalysis, produced at the European Centre for Medium-Range Weather Forecasts (ECMWF; Dee et al. 2011) as a representation of the atmospheric conditions. We

TABLE 1. Atmospheric parameter calculation.

Atmospheric parameter	Calculation of parameter
Instability LI	$\min[\text{LI}_{p(925\text{hPa}) \rightarrow p(500\text{hPa})}, \text{LI}_{p(850\text{hPa}) \rightarrow p(500\text{hPa})}, \text{LI}_{p(700\text{hPa}) \rightarrow p(500\text{hPa})}]$
Midtropospheric moisture RH	$[\text{RH}_{p(850\text{hPa})} + \text{RH}_{p(700\text{hPa})} + \text{RH}_{p(500\text{hPa})}]/3$
Deep-layer shear DLS	$[(u_{10\text{m}} - u_{500\text{hPa}})^2 + (v_{10\text{m}} - v_{500\text{hPa}})^2]^{1/2}$

calculated multiple atmospheric parameters every 6 h on a $0.75^\circ \times 0.75^\circ$ grid between 2008 and 2016 from pressure-level data.

The following parameters were selected for the development of the models. First, the instability was represented by a minimum lifted index, which is calculated by lifting an air parcel adiabatically from three different source layers—925, 850, and 700 hPa—to the 500-hPa layer, as was done by Púčik et al. (2017). The minimum LI value and, hence, the air parcel with the largest measure of instability, is selected and will be referred to as LI. We selected LI instead of CAPE because the method should be applicable to climate models with a reduced vertical resolution precluding the calculation of CAPE. Second, the average relative humidity (RH) from the 850-, 700-, and 500-hPa layers was used because it was identified as relevant for thunderstorm initiation by Westermayer et al. (2017). Third, we calculated the deep-layer (bulk) wind shear (DLS) between 10 m and the 500-hPa level. We decided to choose the 500-hPa layer as the top layer to ensure that we can apply the models to climate simulations for which insufficient vertical pressure levels are available to interpolate to a height level above ground (Púčik et al. 2017). We followed the choice for parameters of Púčik et al. (2017) because we plan to apply AR-CHaMo to the same climate models used in their study. For these models, pressure levels above 500 hPa are not available.

The details of the atmospheric parameter calculations are listed in Table 1.

3. Model development

The model that we developed is multiplicative, in the sense that we treat the probability of hazards as the product of the probability that a thunderstorm occurs and the probability that the storm, given that it occurs, will produce the hazard, that is,

$$P(\text{hazard}) = P(\text{storm}) \times P(\text{hazard}|\text{storm}). \quad (1)$$

This approach implies that any hazards occurring in the absence of lightning are excluded from the model.

a. The dependence of lightning on predictor parameters

Knowledge about the dependence of lightning on ERA-Interim atmospheric parameters is important to

develop the model $P(\text{storm})$. We acknowledge that many predictors can be considered for modeling this, but here, we will focus on the results of Westermayer et al. (2017), which demonstrated a strong relationship among lightning occurrence, instability, and midtropospheric humidity. They analyzed multiple atmospheric parameter distributions and linked them with storm occurrence. A similar approach is done in this study, combining the LI–RH (LI–DLS) parameter space with the occurrence of lightning (convective hazard given lightning) cases.

The joint distribution of all data points in the domain in LI–RH space shows that positive LI values (i.e., stable conditions) are most common and occur within typical RH ranges of 20%–80% (Fig. 3a), while for lightning cases, the maximum values shift toward lower LI and higher RH (Fig. 3b). This means that the relative frequency of lightning $\text{rel}_f(\text{storm})$, defined as the fraction of lightning cases to all cases, is highest for unstable (negative LI) and moist conditions (Fig. 3c).

b. The dependence of hazard occurrence on predictor parameters

To model $P(\text{hazard}|\text{storm})$, the relationship between the relative frequency of hazards, coincident with lightning, and potential predictor parameters has to be explored. Here, we investigate the dependence on the predictor parameters LI and DLS. Figure 4a shows that most lightning cases in the area of Fig. 2 occur with LI of values close to 0 K (neutral stability for a lifted parcel) and DLS within the range of 5–15 m s^{-1} (Fig. 4a). The highest number of such cases that are accompanied by hail occurs for negative LI values in a range from 0 to -5 K, and DLS ranges between 5 and 20 m s^{-1} (Fig. 4b). The relative frequency of hail, given lightning, $\text{rel}_f(\text{hail}|\text{storm})$, increases with decreasing LI (i.e., increasing instability) and with increasing DLS (Fig. 4c).

Increasing instability leads to stronger updrafts that are required for hail formation, and DLS influences the storm types. High values of DLS favor the development of well-organized storms, such as multicells, squall lines, or supercells, which are prone to produce hail.

For wind, the majority of lightning cases occur for higher shear values, between 5 and 20 m s^{-1} (Fig. 4d), and in somewhat more stable conditions. A fair number of wind cases, unlike hail cases, occur with DLS exceeding 30 m s^{-1} and a positive LI. The relative frequency

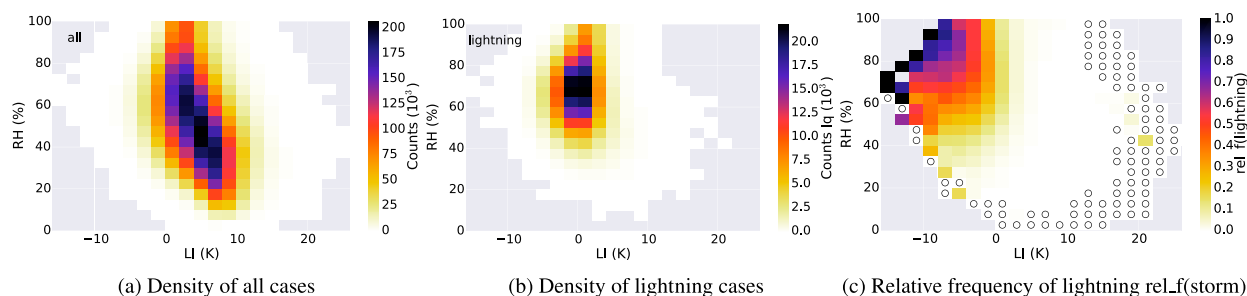


FIG. 3. Histogram of ERA-Interim parameters LI and RH across the W&CEurope domain (2008–16). (a) Histogram of all cases, (b) histogram of EUCLID lightning cases, and (c) relative frequency of lightning rel_f(storm) . Open circles in (c) represent parameter combinations that occurred in (a) but not in (b).

of wind cases, given lightning, rel_f(wind|storm) , exhibits two regions with high values (Fig. 4e): 1) for high instability (and decreasing LI values) and high DLS, for LI between -5 and -10 K and DLS between 20 and 30 m s^{-1} , and 2) under stable conditions (LI of $+4$ K) and even higher DLS of 40 m s^{-1} . This second maximum is mainly associated with winter storms for which the convective layer does not reach up to 500 hPa . It is also possible that the signal in the stable–high DLS parameter space might be reflective of elevated storm activity, leading to evaporative-cooling-driven downdrafts and dry microbursts. However, we already account for elevated

convection because we allow the air parcel to be lifted from three different source layers (925 , 850 , and 700 hPa). Further, the spatial ($0.75^\circ \times 0.75^\circ$) and temporal (6 h) resolutions of ERA-Interim might not be able to capture subgrid-scale instability under all circumstances.

c. Additive logistic regression

In the next step, a statistical model was developed to yield continuous probability functions $P(\text{storm})$ and $P(\text{hazard}|\text{storm})$ across the predictor parameter space. We have chosen to first develop a generalized linear model (GLM), as described by Nelder and Wedderburn

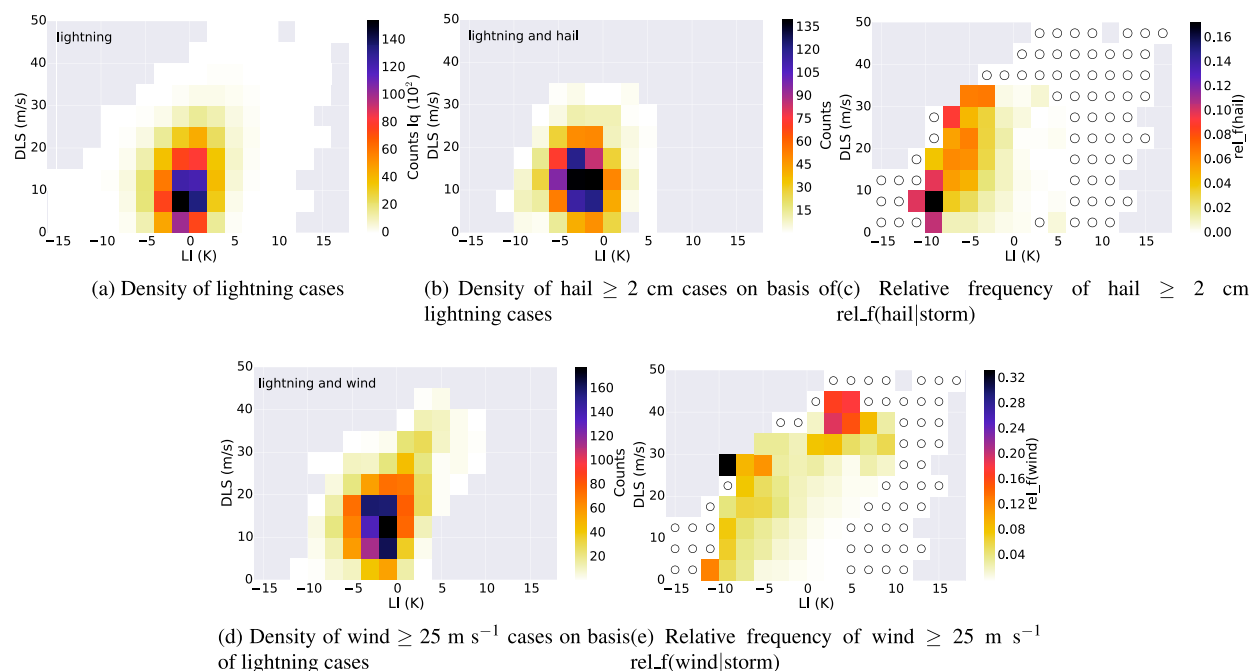


FIG. 4. (a) Histogram of EUCLID lightning cases in LI–DLS parameter space across the G&Alps domain (2008–16). (b) Histogram for hail $\geq 2 \text{ cm}$ cases and (c) relative frequency of hail for lightning cases rel_f(hail|storm) . (d) Histogram for wind $\geq 25 \text{ m s}^{-1}$ cases and (e) relative frequency of wind for lightning cases rel_f(wind|storm) . Open circles in (c) and (e) represent parameter combinations that occurred in (a) but not in (b) or (d).

(1972), and, second, to extend the GLM to a generalized additive model (GAM).

A GLM allows a linear model to be related to the response variable via a “link function” and, therefore, generalizes an ordinary linear regression. Additionally, GLMs give rise to error distribution models other than normal distributions for response variables. Wood (2006) describes the structure of a GLM as

$$g(\mu_i) = \mathbf{X}_i \cdot \boldsymbol{\beta}, \tag{2}$$

where \mathbf{X}_i represents the i th row of the model matrix \mathbf{X} of the explanatory variable or predictor. The regression coefficient $\boldsymbol{\beta}$ is a vector of unknown parameters. The quantity $g()$ is a link function relating the mean μ —in other words, the estimated fitted values $E(y)$ —to the linear predictor $\mathbf{X}_i \cdot \boldsymbol{\beta}$. The quantity μ_i , hence, describes the expected value of the response variable or predictand Y_i , that is, $\mu_i \equiv \mathbb{E}(Y_i)$. The quantity Y_i is distributed according to some exponential family distribution.

In our case, Y_i is distributed binomially because the response variable, or predictand “hazard occurrence” has only two options: yes or no. Logistic regressions are fit to binary predictands and can then be viewed as linear after the application of a certain transformation. In the present example with a binomial family distribution, the necessary transformation is defined by the logarithm of the odds ratio $\mu/(1 - \mu)$, which is also called “logit transformation” (Wilks 2006). Hence, in our case of a linear logistic regression, the logit link function $g(\mu) = \ln[\mu/(1 - \mu)]$ applies. The probabilities for lightning $P_{\text{lin}}(\text{storm})$; for hail, given lightning, $P_{\text{lin}}(\text{hail}|\text{storm})$; and for wind, given lightning, $P_{\text{lin}}(\text{wind}|\text{storm})$, are calculated using the programming language R for statistical computing (R Core Team 2015). We used the package “mgcv” (Wood 2017) for the above-described linear logistic regression for two explanatory variables in each case. The resulting models were exploited to generate lookup tables (Figs. 5a,c,e), which link the probability of a hazard to two atmospheric parameters with predefined ranges. The standard errors of the models (Wood 2006) are depicted as contour lines in Figs. 5a, 5c, and 5e. The probability $P_{\text{lin}}(\text{storm})$ being dependent on RH and LI increases with increasing instability (i.e., decreasing LI values) (Fig. 5a). The dependence of $P(\text{storm})$ on the parameter RH is not reproduced correctly, owing to the inflexibility of the linear logistic regression. This means for 40% RH and -7 K LI, the relative frequency is $\text{rel_f}(\text{storm}) \approx 0.1$, while the modeled probability $P_{\text{lin}}(\text{storm}) \approx 0.8$ (cf. Figs. 3c and 5a). For 80% RH and -7 K LI, $\text{rel_f}(\text{storm}) \approx 0.65$, while $P_{\text{lin}}(\text{storm}) \approx 0.9$. The strong increase of $\text{rel_f}(\text{storm})$ for increasing RH in unstable environments (LI < 0 K)

is also not modeled correctly in $P_{\text{lin}}(\text{storm})$ by the linear model. For $P_{\text{lin}}(\text{hail}|\text{storm})$ and $P_{\text{lin}}(\text{wind}|\text{storm})$, the probability increases with decreasing LI values and increasing DLS values, with a stronger dependence on LI for $P_{\text{lin}}(\text{hail}|\text{storm})$ and on DLS for $P_{\text{lin}}(\text{wind}|\text{storm})$ (Figs. 5c,e).

To avoid the abovementioned deficiencies of the linear approach, and to account for more flexibility in the models, a generalized additive model was investigated. Hastie and Tibshirani (1986) defined a GAM as a generalized linear model with a linear predictor involving a sum of smooth functions of covariates (Wood 2006). The general structure of a GAM is

$$g(\mu_i) = \mathbf{X}_i^* \cdot \boldsymbol{\theta} + f_1(x_{1,i}) + f_2(x_{2,i}) + f_3(x_{3,i}, x_{4,i}) + \dots, \tag{3}$$

where \mathbf{X}_i^* represents the i th row of a model matrix \mathbf{X}^* for any model components, and $\boldsymbol{\theta}$ is the corresponding parameter vector (Wood 2006). The f_i describes the smooth functions of the covariates x_k (Wood 2006). Instead of detailed parametric relationships, this model allows a rather flexible specification of the dependence of the response on the covariates and specifies the model only in terms of “smooth functions” (Wood 2006). This means the smoothing function is estimated from the available data, and no theory or mechanic model assumption is needed. Using this approach, AR-CHaMo is developed and adjusted for each of the predictands’ probabilities of lightning, hail requiring the coincidence of lightning, and wind requiring the coincidence of lightning for two predictors (atmospheric parameters).

If the response variable y is binomially distributed and depends on two explanatory variables x_1 and x_2 , the GAM can be written as

$$\text{logit}[E(y_i)] = f_1(x_{1,i}, x_{2,i}). \tag{4}$$

For example, the lightning model, which depends on two explanatory variables, LI and RH, can be expressed by

$$\text{logit}[E(\text{Lightning}_i)] = f_1(\text{LI}_i, \text{RH}_i). \tag{5}$$

This additive logistic regression yields a probability $P(\text{storm})$ over the parameter space LI and RH after model calculation using thin-plate regression splines (Wood 2006).

The hazard models, which depend on the explanatory variables LI and DLS, are represented by

$$\text{logit}[E(\text{Hail}_j)] = f_2(\text{LI}_j, \text{DLS}_j), \tag{6}$$

yielding $P(\text{hail}|\text{storm})$, and

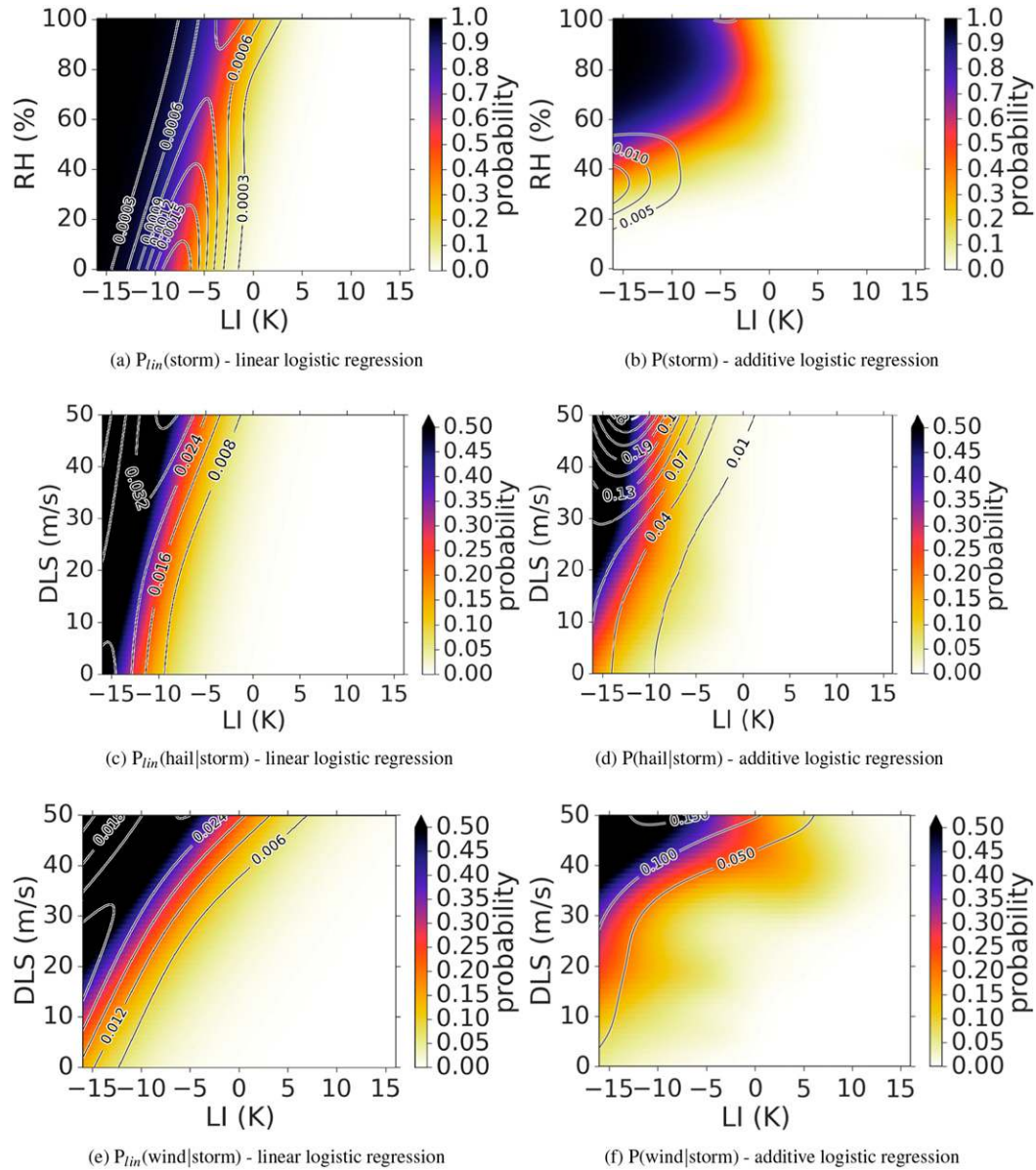


FIG. 5. Probabilities predicted by statistical models depending on LI and DLS for (left) GLMs and (right) GAMs for (a),(b) lightning cases, (c),(d) hail ≥ 2 cm cases, and (e),(f) wind $\geq 25 \text{ m s}^{-1}$ cases. Contour lines are standard error estimates as described by Wood (2017).

$$\text{logit}[E(\text{Wind}_j)] = f_3(\text{LI}_j, \text{DLS}_j), \quad (7)$$

yielding $P(\text{wind}|\text{storm})$, where $i \in [1, n]$ signifies the space of all available n observations and j denotes the subset of all indices i in which a lightning case was observed.

The three AR-CHaMo variants for the probability of lightning $P(\text{storm})$; the probability for hail, given lightning, $P(\text{hail}|\text{storm})$; and the probability for wind, given lightning, $P(\text{wind}|\text{storm})$, are computed using the logit link function for the binomial family distribution

(e.g., lightning “yes” or “no”) and thin-plate regression splines (Wood 2003) to estimate the smooth function. The resulting lookup tables for the hazard probabilities predicted by the different AR-CHaMo variants are shown in Figs. 5b, 5d, and 5f. For $P(\text{storm})$, the highest probability values occur with $\text{LI} < -5 \text{ K}$ and $\text{RH} > 65\%$ (Fig. 5b). The dependence on RH is well pronounced, in contrast to the linear logistic regression model (Fig. 5a).

For $P(\text{hail}|\text{storm})$, the highest values can be found with high DLS and low LI values (Fig. 5d), representing an unstable atmosphere with strong updrafts (balancing

TABLE 2. Comparison between additive logistic regressions and linear logistic regressions. Models are computed based on data from 2008 to 2016. For the lightning models, data from W&CEurope were chosen, and for the hazard models— $P(\text{hazard}|\text{storm})$ — data from G&Alps were taken into account. Parameter n signifies the total number of data points within the selected time and spatial domain.

Model	Predictors	Type	Deviance explained (%)	BIC	n
$P_{\text{lin}}(\text{storm})$	LI and RH	Linear	26.8	2 801 047	10 127 040
$P(\text{storm})$	LI and RH	Additive	28.9	2 718 959	10 127 040
$P_{\text{lin}}(\text{hail} \text{storm})$	LI and DLS	Linear	8.0	15 552.25	150 952
$P(\text{hail} \text{storm})$	LI and DLS	Additive	8.66	15 595.39	150 952
$P_{\text{lin}}(\text{wind} \text{storm})$	LI and DLS	Linear	6.1	20 441.75	150 952
$P(\text{wind} \text{storm})$	LI and DLS	Additive	7.32	20 415.23	150 952

the fall speed of a growing hailstone) and large vertical wind shear, which tends to support storm organization, longevity, and severity (Markowski and Richardson 2010). For $P(\text{wind}|\text{storm})$, two areas emerge: 1) low LI and high DLS—well-organized storms with strong updrafts and large vertical wind shear—and 2) positive LI with very high DLS that is likely to represent atmospheric conditions during winter storms where the convective layer does not reach up to 500 hPa. Another possible explanation might be that narrow and fast-moving instability fields are not resolved by the spatial or temporal resolutions of ERA-Interim (Fig. 5f).

The resulting probabilities of the additive logistic regression models (Figs. 5b,d,f) fit the observational data (relative frequencies of Figs. 3c, 4c,e) better than the resulting probabilities of the linear logistic regression models. The additive approach can overcome the limitations described earlier for the linear logistic regression. The additional flexibility of the additive model can resolve the dependence of $P(\text{storm})$ on RH more accurately. Thunderstorms rarely occur with low values of RH (<40%), even if sufficient instability is available (LI < 0 K). These low relative frequencies are well represented by the additive logistic regression (cf. Fig. 3c and Fig. 5d). For $P(\text{hail}|\text{storm})$ and $P(\text{wind}|\text{storm})$, the differences between the additive and linear logistic regressions are less pronounced.

In addition to the consideration of physical feasibility of the additive and linear models, statistical measures can be taken into account for model comparison. Two common measures for model selection are the residual deviance, which is the deviance of the fitted model (Wood 2006), and the “deviance explained,” which is a standard criterion commonly used in statistical modeling and can be easily derived from the residual deviance. The model with the higher deviance explained can be considered superior to others.

Further, the Bayesian information criterion (BIC) can be exploited to distinguish between different models. Models with lower BIC are preferred, while it has to be assumed that the same data points are considered. To avoid overfitting, a penalty term is introduced for the

number of parameters in the models. Table 2 summarizes the comparison between the statistical measures deviance explained and BIC for the linear logistic regressions and the additive logistic regressions of Fig. 5. Yet again, the additive logistic regressions show a better performance over the linear logistic regressions.

An illustrative graphical means for the verification of statistical models is receiver operating characteristic (ROC) diagrams (Wilks 2006). ROC curves display the ratios between true positive rates and false positive rates for varying thresholds of a binary classifier system, which, in our case, represents the occurrence of a lightning case or a hazard case. The ROC curves for the linear and additive models are shown in Figs. 6a–c, and it becomes evident that all models show a good overall performance, as their ratio of true positive rates exceeds the false positive rates considerably.

To obtain an estimate of the errors in the additive models, we performed a cross validation. The available data were divided into two parts: data from even years (2008, 2010, 2012, and 2014) and data from odd years (2009, 2011, 2013, and 2015). Six additive models were computed for lightning, hail, and wind: $P_{\text{even}}(\text{storm})$, $P_{\text{odd}}(\text{storm})$, $P_{\text{even}}(\text{hail}|\text{storm})$, $P_{\text{odd}}(\text{hail}|\text{storm})$, $P_{\text{even}}(\text{wind}|\text{storm})$, and $P_{\text{odd}}(\text{wind}|\text{storm})$. Models that are developed using data from odd years were applied to data of even years, and vice versa. ROC curves in Figs. 6d–f show that the performance of these models is slightly degraded relative to those that were developed using the full dataset. However, the quality of the even-year models applied to the odd-year data is very similar to that of the odd-year models applied to the even-year data. This indicates that the additive models are robust (Figs. 6d–f).

After examining various statistical measures and comparing additive and linear models, it can be concluded that the additive approach (AR-CHaMo models) can outperform the linear approach and is, therefore, selected for this study. Finally, a cross validation using data from even and odd years is able to underline the robustness of the additive models.

In the next step, we show that AR-CHaMo can reproduce meteorological patterns, such as annual cycles and spatial distributions.

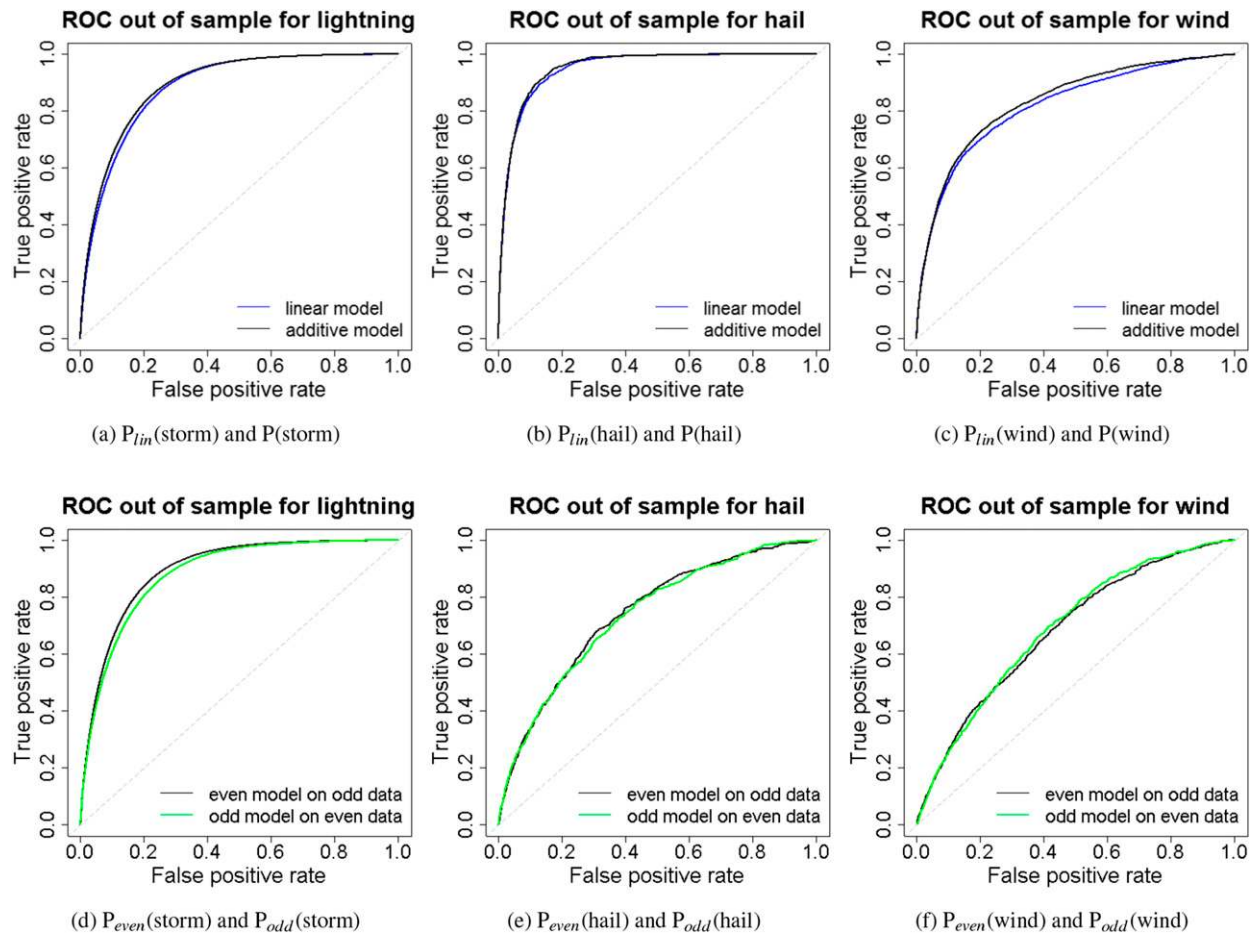


FIG. 6. ROC curves showing (a)–(c) comparison between additive (gray) logistic regressions and linear (blue) logistic regressions for lightning, hail, and wind, respectively; models are developed using data between 2008 and 2016. (d)–(f) As in (a)–(c), but for models based on data from odd years [gray (2009, 2011, 2013, and 2015)] and applied to data from even years [green (2008, 2010, 2012, and 2014)], and vice versa. For the lightning models, data from W&CEurope were taken into account; while for the hazard models, data from G&Alps were chosen.

4. Application to reanalysis data

a. Annual cycles

Another important indicator for the quality of a model is its ability to reproduce the observed annual cycles. We applied the AR-CHaMo lightning models to ERA-Interim data across the W&CEurope domain and the 2008–16 period. The modeled number of cases per month is determined by summing up all individual probabilities for the 6-h periods. The resulting sum is divided by the number of months considered. We find that the modeled monthly number of lightning cases follows the mean annual cycle of observed lightning cases quite well (Fig. 7a). The model does overestimate this number to a small extent from January to April and slightly underestimates it from May to July.

Next, we applied the models for hail and wind to the G&Alps domain. The modeled annual cycle of hail cases follows the mean annual cycle of observed hail cases fairly well (Fig. 7b). Indeed, in May and June, the model underestimates the number of observed cases, while it slightly overestimates it from July onward (Fig. 7b). The modeled annual cycle of wind cases and that of observed wind cases match fairly well throughout the year. A notable difference is that the maximum of the model underestimates the observed maximum in July by approximately 40%. The spikes in the cool season observations may be explained by the fact that the comparison only takes 9 yr of data into account; in some months within this period, rare major synoptic windstorms with electrified convection occurred, whereas by chance they did not in others. They can be identified in Fig. 2d (top) as two winter storms in 2008 and 2012, respectively.

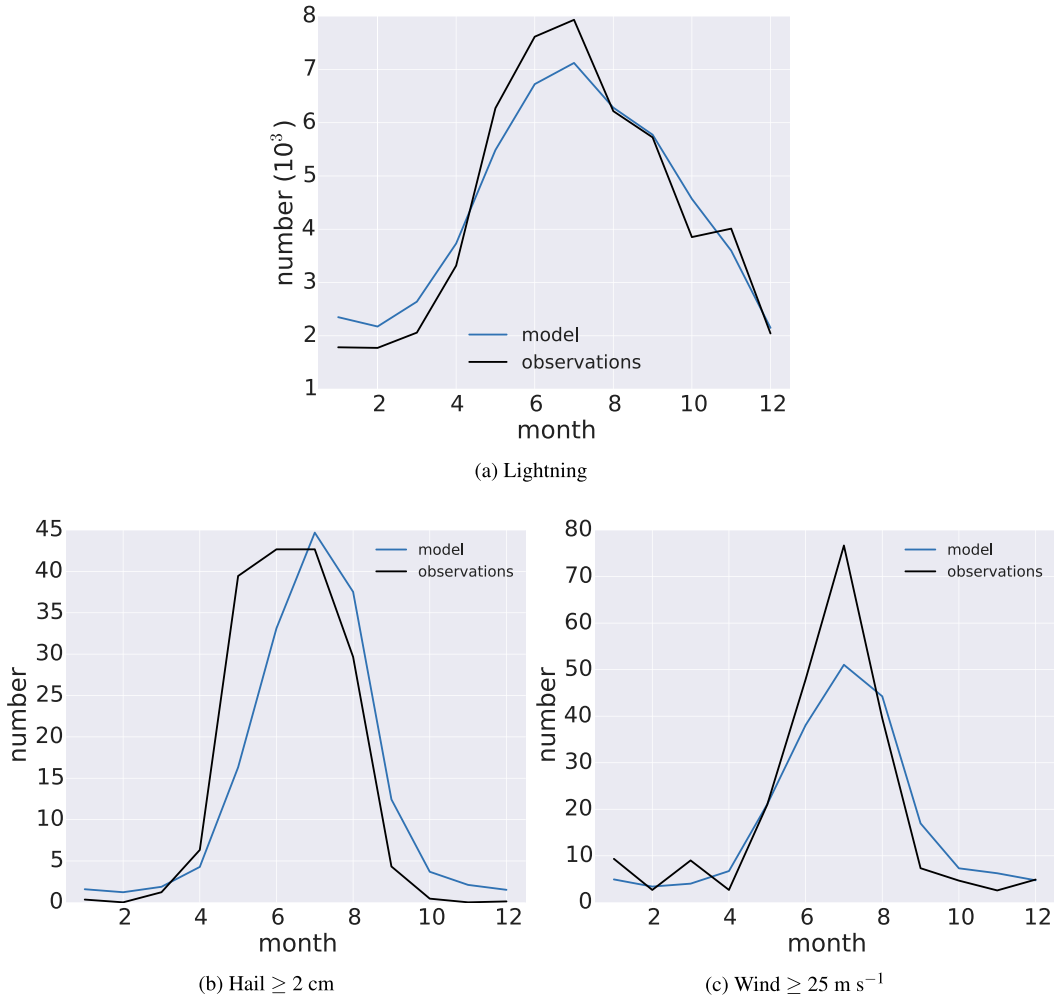


FIG. 7. Mean annual cycles for (a) modeled number of lightning cases (blue) and observed EUCLID lightning cases (black) for W&CEurope; (b) as in (a), but for hail cases (blue) and observed ESWD hail cases where lightning occurred (black) for G&Alps; and (c) as in (b), but for wind cases (blue) and observed ESWD wind cases where lightning occurred (black) for G&Alps. Time period is 2008–16.

One explanation for the hail underestimation in May and June suggests that there are other factors controlling hail occurrence, beyond the three parameters that we have taken into account. Our thinking is that these might be 1) the height of the freezing level, which is lower in those months, and/or 2) the typically steeper midlevel lapse rates in those months, which may favor hail formation in addition to the buoyancy of a theoretical low-level parcel (which we represent by the lifted index). More advanced models might take such effects into account. In this study, we restricted ourselves to two-dimensional models in order to present the method and show its applicabilities.

b. Spatial distribution

Another check of the model’s quality is whether or not it is able to reproduce observed spatial patterns. We,

therefore, applied the models to the reanalysis dataset across all of Europe from 1979 to 2016 and calculated the annual number of hazard cases for each grid point. The modeled spatial distribution of lightning cases (1979–2016) reflects the observed pattern (2008–16) in Fig. 1 quite well (Fig. 8a). The modeled, as well as the observed, thunderstorm activity is maximized in north Italy and is high along the Apennine Mountains and the east coast of the Adriatic Sea. The north–south gradient of thunderstorm activity in Germany is reproduced. The modeled number of hail cases in Fig. 8b presents a different pattern than the observed ESWD reports of Fig. 2a because strong hail activity is no longer limited to the area where the ESWD reporting efficiency is high. North Italy stands out, as well as central Romania and the Pyrenees. The Atlas Mountain Ranges in northern

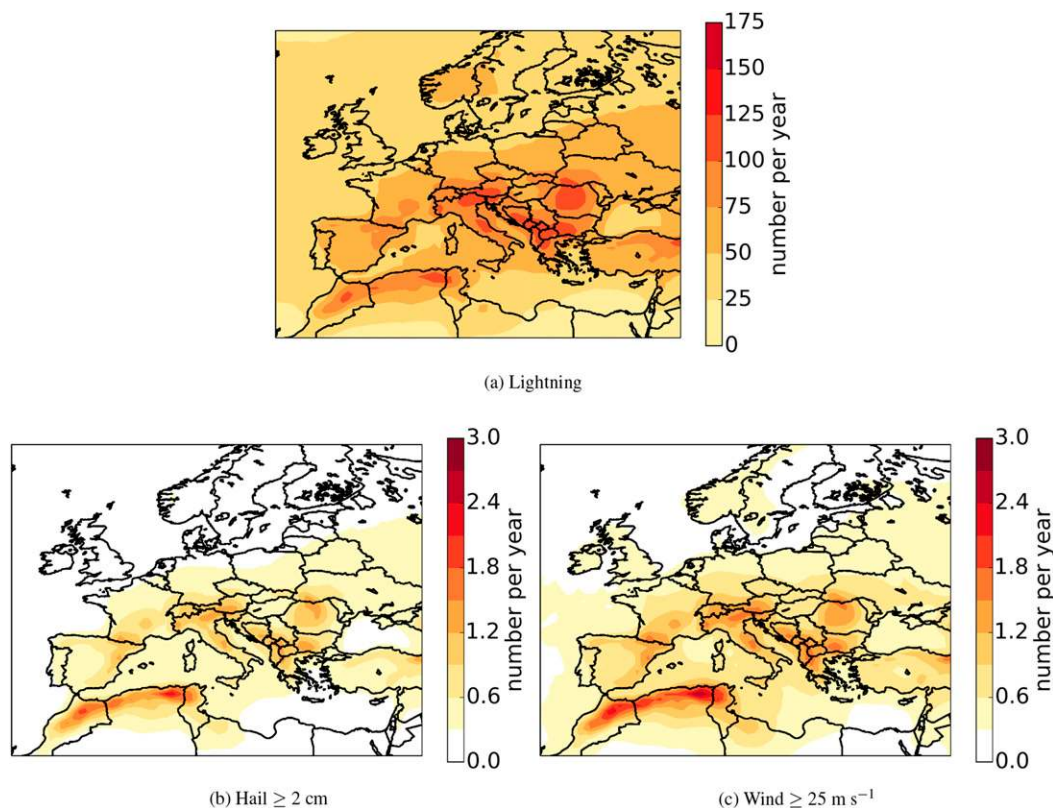


FIG. 8. AR-CHaMo application to ERA-Interim data for the period 1979–2016. The modeled number of cases per year is determined by summing up all individual probabilities for the 6-h periods. The resulting sum is divided by the number of years considered. (a) Modeled number of lightning cases per year, (b) modeled number of hail cases per year (hail ≥ 2 cm), and (c) modeled number of wind cases per year (wind ≥ 25 m s $^{-1}$).

Africa have the highest number of events on the map. Within Germany, most hail cases are modeled in the southwest, with a decrease from south to north. This gradient is more pronounced in the model than in the observations (Fig. 2a). The model result, however, corresponds well to that found in the study by Punge et al. (2014), which is based on the frequency of overshooting tops detected by satellite, or by Puskeiler et al. (2016), who used 3D radar data. Across the G&Alps domain, the absolute number of modeled events of approximately 0.6 per year seems rather low, which may be caused by an underestimation of ERA-Interim instability.

For wind cases (Fig. 8c), the overall pattern may, at first sight, seem similar to the pattern of hail (Fig. 8b), but there are important differences. These differences can be seen when dividing the modeled number of hail (Fig. 9a) and wind (Fig. 9b) cases by the modeled number of lightning cases. In Belarus, for example, between 0.6% and 0.9% of lightning cases are accompanied by hail or wind. However, in Ireland, the percentage of lightning cases with hail is only between

0.3% and 0.6%, while the percentage of lightning cases with wind is between 0.9% and 1.2%.

Across the G&Alps domain, wind events are underestimated in the northern half, relative to the southern half of this domain, that is, in comparison with the observations. We think that this is primarily caused by the difficulty of representing wind gusts that occur with convection under weak instability in strong background flows near synoptic-scale storms. Such storms can produce both lightning and severe wind gusts, but the wind gusts are primarily caused by the strong background wind field. In such cases, the weak instability yields a relatively low $P(\text{wind}|\text{storm})$, resulting in an underestimation of wind gust risks. The AR-CHaMo wind model result compares favorably to a study by Mohr et al. (2017), who have explicitly left out synoptically driven wind gusts from their climatology of convective wind gusts across Germany. By doing so, the maximum number of convective gusts (>18 m s $^{-1}$) observed at standard meteorological stations shifts to the south of Germany and resembles the AR-CHaMo result rather closely.

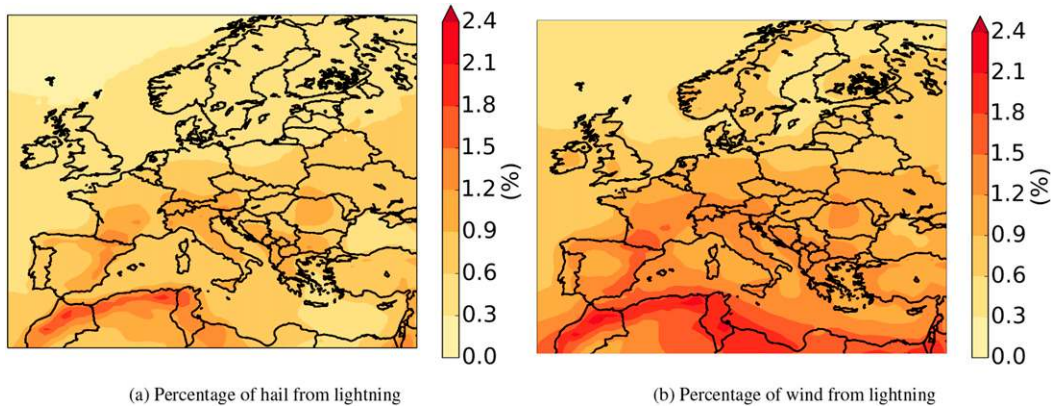


FIG. 9. AR-CHaMo application to ERA-Interim data for the period 1979–2016. The modeled number of cases per year is determined by summing up all individual probabilities for the 6-h periods. The resulting sum is divided by the number of years considered. (a) Modeled number of hail cases per year divided by the modeled number of lightning cases per year. (b) As in (a), but for wind cases. Thus, the fraction of lightning cases occurring with (a) hail or (b) wind is shown.

c. Trends of hazard occurrence

To detect trends of the hail or wind hazards in the last decades, we now consider the modeled annual number of lightning, wind, and hail cases across the W&CEurope and G&Alps domains (Fig. 10). We summed up the individual probabilities for each 6-hourly period within the domains and per year to receive the expected number of cases per year for each domain (Fig. 10). For lightning, the annual number of cases decreases until approximately 1995, followed by a strong increase until 2016, which dominates the trend over the entire time period (Fig. 10a). The increase amounts to 3.77 cases per decade for the G&Alps (blue line) and 2.55 cases per decade for the W&CEurope domain (green line). These numbers result in a relative increase of 23% and 16% within the G&Alps and W&CEurope domains, respectively, between 1979 and 2016. The modeled number of hail and wind cases also increased across both domains (Figs. 10b,c). Since 1979, the respective increases in hail and wind cases were 0.091 and 0.092 per decade across the G&Alps domain (blue line) and 0.047 and 0.051 per decade across the W&CEurope domain (green line). For hail and wind, the relative increases are 86% and 56%, respectively, across the G&Alps domain and 41% and 29%, respectively, across the W&CEurope domain. Table 3 summarizes all linear trends (per decade) along with their significance levels that are shown in Fig. 10.

The time period 1990–2000 stands out with lower lightning and hail activity, relative to the earlier and later periods. It also shows a smaller degree of interannual variability. The low thunderstorm activity between 1990 and 2000 correlates with higher LI values—resulting in less instability—and lower midlevel relative humidity (Fig. 12).

The trends for lightning, hail, and wind are not homogeneous across Europe. The modeled number of annual lightning, wind, and hail cases exhibits a significant positive linear trend in northern Italy and on the southeast coast of the Adriatic Sea (Fig. 11). Modest positive trends are detected across several areas in north-central Europe. At the same time, significant negative trends are identified across parts of northeastern Spain and northern Morocco (Figs. 11a–c).

5. Driving factors for the modeled trends

In the next step, we answer the question of how the detected changes in hazard probability are linked to the underlying atmospheric parameters used in the models, that is, LI, RH, and DLS.

a. Trends and variability of instability parameter LI

The fifth percentile of LI exhibits a significant linear trend of $-0.242 \text{ K decade}^{-1}$ across the G&Alps domain and $-0.126 \text{ K decade}^{-1}$ across the W&CEurope domain (Fig. 12a; Table 4). The negative signs of the trends denote that 5% of most unstable environments have become more unstable. The fifth percentile was chosen, as it corresponds with an amount of instability (near 0 K), which is marginally sufficient to support lightning. For lower (higher) values of LI, the probability of lightning increases (decreases) rapidly (Fig. 3c). The standard deviation of the LI distribution within any year has a positive linear trend of $+0.074 \text{ K}$ ($+0.046 \text{ K}$) across the G&Alps (W&CEurope) domain, which signifies an increase in LI variability during the last three decades (Table 4). The trend of LI differs across Europe (Fig. 13a). Across most of Europe, significant negative trends of LI prevail, while across Spain and

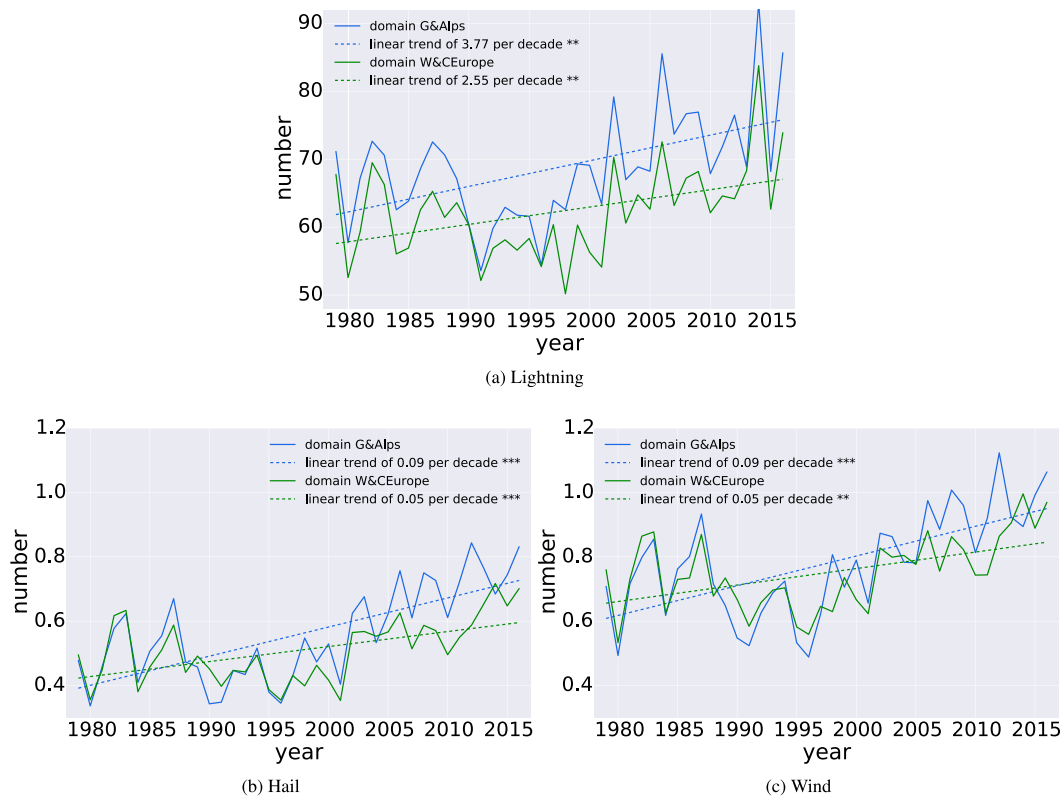


FIG. 10. AR-CHaMo application to ERA-Interim data for the period 1979–2016. The modeled number of cases per year is determined by summing up all individual probabilities for the 6-hourly periods in 1 yr within a chosen domain. The resulting sum is divided by the number of grid boxes considered. Modeled number of cases per year for (a) lightning, (b) hail (≥ 2 cm), and (c) wind (≥ 25 m s $^{-1}$). The trends per decade (dashed lines), along with its significance levels, are summarized in Table 3. The green curves correspond to W&CEurope, and the blue curves correspond to G&Alps. Significance is indicated by the asterisks associated with the linear trend lines with 3 asterisks = 0 and 2 asterisks = 0.001 significance.

Morocco, significant positive trends are detected. The strongest increase in instability occurred in north Italy and the western Balkans (-0.55 K decade $^{-1}$). The spatial pattern of the LI trend (Fig. 13a) is qualitatively similar to the trend of modeled lightning, hail, and wind cases (Fig. 11). This suggests that changes in instability may be the primary driver behind the changes in lightning, hail, and wind cases.

b. Trends and variability of midlevel relative humidity parameter RH

Relative humidity is also a potential driver of the changes, as low levels of relative humidity prevent storm

occurrence even when sufficient instability is present (Fig. 3c). We considered changes of the median relative humidity in cases of negative LI, which is 65.36% (60.52%) across the G&Alps (W&CEurope) domains. These values lie in a region of parameter space with a strong gradient of lightning probability, that is, for higher RH than this value, the probability of lightning increases strongly (Fig. 3c). The median RH exhibits a significant negative linear trend of -1.165% decade $^{-1}$ (Fig. 12b) for the G&Alps domain, but there is no significant trend for the W&CEurope domain (-0.101% decade $^{-1}$). The standard deviation of RH for LI < 0 K

TABLE 3. Trends of modeled number of hazard occurrences from Fig. 10.

Hazard	G&Alps domain		W&CEurope domain	
	Trend (decade $^{-1}$)	Significance level α	Trend (decade $^{-1}$)	Significance level α
Lightning	+3.772	0.001	+2.554	0.001
Hail	+0.091	0	+0.047	0
Wind	+0.092	0	+0.051	0.001

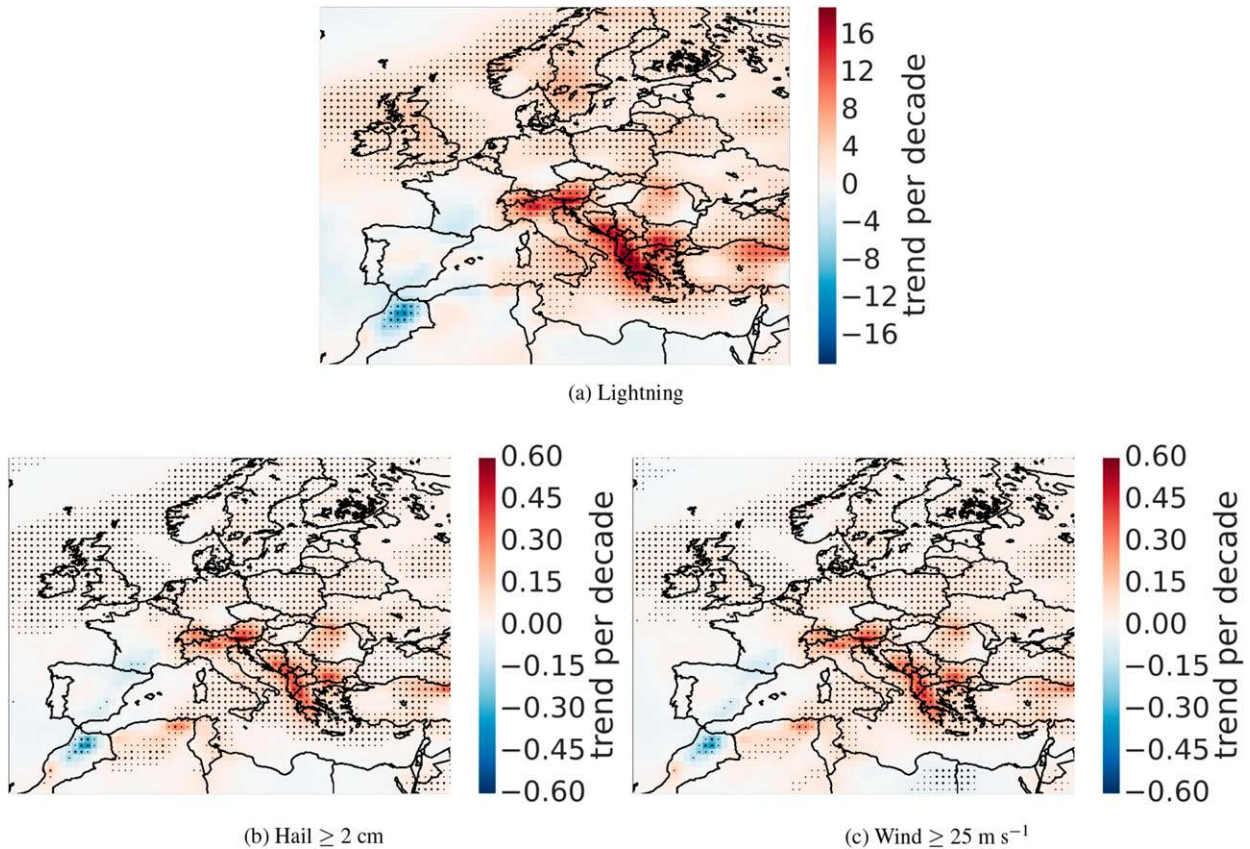


FIG. 11. The modeled number of cases per year is determined by summing up all individual probabilities for the 6-h periods in 1 yr within a grid box. A linear trend is then calculated for each grid box. (a)–(c) Trend of the modeled number of lightning, hail (≥ 2 cm), and wind (≥ 25 m s $^{-1}$) cases per decade, respectively. The small black dots indicate a significance level of 0.01, and the large ones a level of 0.05.

shows no significant trend in either domain (Table 4). The spatial distribution of the trends within Europe is depicted in Fig. 13b. Almost all of eastern and central Europe shows a significant negative linear trend toward lower humidity values. This would have an inhibiting effect on convective storms. The fact that the model predicts an increased number of lightning cases in the area suggests that the effect of increased instability (lower LI) dominates that of decreasing midtropospheric humidity. No significant trend of RH is observed over Italy and France. There is a positive trend over central Spain and eastern Portugal, but its effect on the number of lightning cases is dominated by the decrease of instability (increase of LI).

c. Trends and variability of deep-layer shear parameter DLS

High values of vertical wind shear, that is, DLS, in the presence of instability favor the development of organized convection and attendant hazards. To identify the effect of changes in DLS on hazard probability, we consider the median of DLS in an unstable

atmosphere ($LI < 0$ K). No significant trend of DLS was detected across the G&Alps ($+0.101$ m s $^{-1}$ decade $^{-1}$) and W&CEurope (-0.080 m s $^{-1}$ decade $^{-1}$) domains (Fig. 12c; Table 4). There are no significant trends in the standard deviation for either domain (Table 4). In spatial terms, a significant positive trend is found over northeastern Germany and northern Africa (Fig. 13c), and a significant negative trend is found along the Adriatic east coast (Fig. 13c). All other trends are not significant at the 95% confidence level. Because DLS did not change significantly during the last three decades in most of the domain, it is not responsible for the changes in hail and wind probabilities, with the possible exception of parts of northern Africa.

6. Discussion and conclusions

We have developed a statistical model that yields a probability for convective hazard occurrence on the basis of lightning and hazard observations and reanalysis data. The method was evaluated within the time period 2008–16 for central Europe, and it was shown that it was

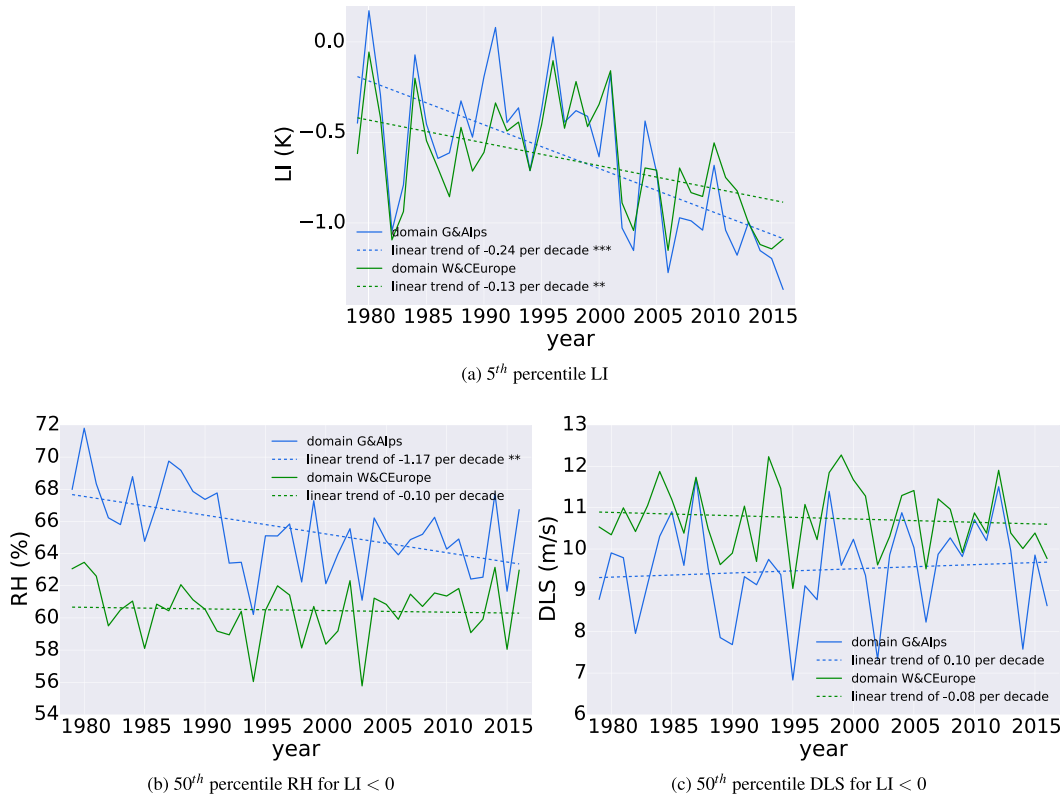


FIG. 12. (a) Change in fifth percentile of LI, (b) change in 50th percentile of RH, and (c) change in the median of DLS. For (b) and (c), only values are taken into account where $LI < 0$ K. Green line corresponds to W&CEurope, and the blue line corresponds to G&Alps. The trends per decade (dashed lines), along with their significance levels, are summarized in Table 4. Significance is indicated by the asterisks associated with the linear trend lines with 3 asterisks = 0, 2 asterisks = 0.001, and no asterisks = 0.05.

able to reproduce the annual cycles and spatial distribution of lightning, hail, and wind hazards fairly well.

The model that was developed has a number of limitations. First, the model was simple and consisted of only two predictor parameters per model component: $P(\text{storm})$ and $P(\text{hazard}|\text{storm})$. The parameter selection and calculation for the AR-CHaMo models is designed to be applicable to climate simulations despite their limited vertical resolution. This model may be improved

further by using additional predictor parameters. Such parameters may include low-level moisture, lapse rates, lifted condensation level, or height of the melting level. In this case, sufficient data must be available to prevent overfitting, which can lead to unphysical relationships between the predictors and the modeled probabilities. This may become a problem when the model has a high number of degrees of freedom. That said, the additional degrees of freedom provided by an additive logistic

TABLE 4. Trends of parameter distribution properties from Fig. 12.

Parameter	G&Alps domain		W&CEurope domain	
	Trend (decade ⁻¹)	Significance level α	Trend (decade ⁻¹)	Significance level α
LI_{q1} (K)	-0.368	0	-0.165	0.001
LI_{q5} (K)	-0.242	0	-0.126	0.001
LI_{std} (K)	+0.074	0	+0.046	0.001
RH_{q50} (%) for $LI < 0$ K	-1.165	0.001	-0.101	0.1
RH_{mean} (%) for $LI < 0$ K	-1.062	0	-0.052	0.1
RH_{std} (%) for $LI < 0$ K	+0.210	0.05	-0.099	0.1
DLS_{q50} (m s ⁻¹) for $LI < 0$ K	-0.101	0.1	-0.080	0.1
DLS_{mean} (m s ⁻¹) for $LI < 0$ K	+0.062	0.1	-0.071	0.1
DLS_{std} (m s ⁻¹) for $LI < 0$ K	-0.014	0.1	-0.022	0.1

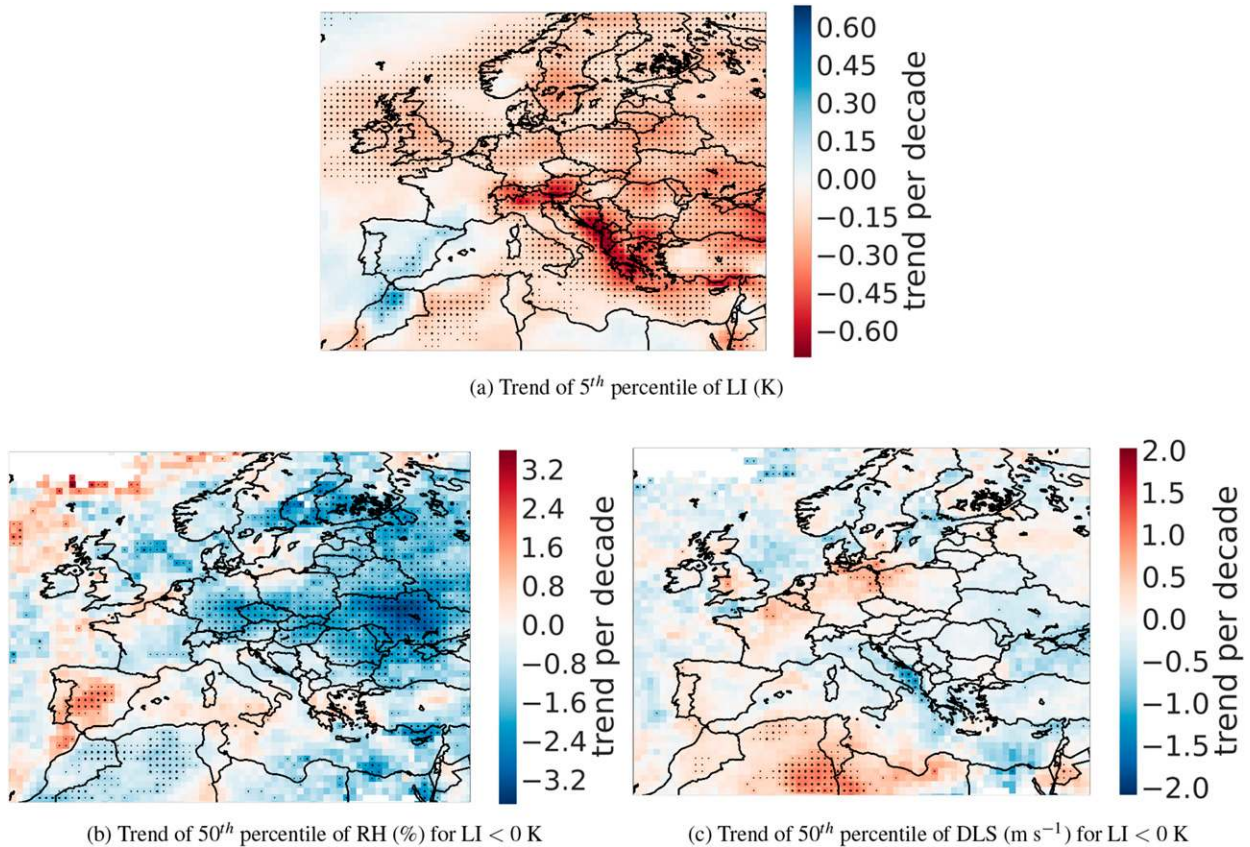


FIG. 13. Trends of ERA-Interim parameter for the period 1979–2016. (a) Trend of fifth percentile of LI (K) per year. (b) Trend of 50th percentile of RH (%) per year. (c) Trend of 50th percentile of DLS (m s⁻¹) per year. For (b) and (c), only those values are taken into account where LI < 0 K, representing an unstable atmosphere. The small black dots indicate a significance level of 0.01, and the large ones a level of 0.05.

regression resulted in a fit to the observations that is much improved relative to the linear logistic regression.

Second, the hazard observations are not complete and feature temporal and spatial inhomogeneities. We tried to minimize any resulting effects on the models by selecting the small domain of G&Alps, across which the collection of data was well organized and homogeneous throughout the 2008–16 period.

Third, the time span of 6 h between subsequent reanalysis data is long relative to the time scale involved in the development of convective storms. For example, an instability value at 1200 UTC might not accurately represent the atmospheric conditions for a storm that formed 5 h later. The spatial resolution of 0.75° implies that the predictor parameter values used are the means over a model grid box of that size. This limited spatio-temporal resolution blurs the relationship between the predictor parameters and the predictands, that is, the lightning, hail, and wind cases. The use of higher-resolution reanalysis data, as they become available in the future, will likely mitigate this problem. Additionally, we have not evaluated the sensitivity of the model

to our choice of the reanalysis data, that is, ERA-Interim. A comparison of results based on other reanalysis datasets can shed light on the robustness of the results. Thorne and Vose (2010) studied reanalysis suitable for characterizing long-term trends. They emphasized that the available reanalyses have obvious and undesirable, unphysical, time-varying biases that lead to discontinuities in long-term trends. These discontinuities may be a result of changes to the assimilated data within the reanalysis (Thorne and Vose 2010).

The analysis of the time series since 1979 indicates a positive trend in the modeled number of lightning cases, as well as a positive trend of hail and wind cases in both domains. The increase of convective hazards can only partly be attributed to the more frequent occurrence of thunderstorms. More important, thunderstorms that occur are more likely to produce severe weather. Both the increase in storms and their efficiency in producing severe weather are driven by an increase in instability, rather than changes in DLS, which were small, or midtropospheric humidity. Although the atmosphere became drier during this time period, the associated inhibiting effect on the occurrence of

storms was overcompensated by the instability increase. The spatial distributions of the different trends and their significance were analyzed in detail. We did not study the reasons for the changes in instability, relative humidity, and wind shear in ERA-Interim. The presented AR-CHaMo modeling framework that we have developed can be applied to other regions of the world, preferably calibrated with observations from that region. An important question that we have not addressed is whether a model that was developed using observations from one region can be applied to another. By applying the models that were calibrated with data from central Europe, we have implicitly assumed their applicability to the rest of Europe and surrounding regions. Further research is needed to test the sensitivity of the models to the region. Besides reanalyses, important potential applications include the use of the framework in the realms of severe weather forecasting and climate projections. The application of the AR-CHaMo method can also be extended to other convective hazards, such as heavy precipitation or tornadoes.

Acknowledgments. The authors thank the reviewers for their detailed comments on the paper that helped to improve the work in many ways. This work was mainly carried out within the Analysis of Changes in the Risk of Severe Convective Storms in Europe (ARCS) project, funded by Munich Re and by the Federal Ministry of Education and Research (BMBF) under Grant 01LP1525A. In addition, Pieter Groenemeijer's work was partly funded by the European Union's Seventh Framework Programme for research, technological development, and demonstration as part of the RAIN project under Grant Agreement 608166. We thank the Statistical Consulting Unit StaBLab, Department of Statistics, LMU Munich, Germany, for their support. We also acknowledge initial work done by Georg Pistotnik, who showed us that generalized additive models may be useful tools to approach the problem. We thank Tomáš Púčik and Lars Tijssen in particular for their scientific and technical support and fruitful discussions. Additionally, we thank Prof. Dr. Peter Höppe, who supported this study. Finally, we acknowledge the European Centre for Medium-Range Weather Forecasts for the ERA-Interim dataset and EUCLID for the lightning detection data.

REFERENCES

- Allen, J. T., and M. K. Tippett, 2015: The characteristics of United States hail reports: 1955–2014. *Electron. J. Severe Storms Meteor.*, **10**, <http://www.ejssm.org/ojs/index.php/ejssm/article/view/149/104>.
- , —, and A. H. Sobel, 2015: An empirical model relating U.S. monthly hail occurrence to large-scale meteorological environment. *J. Adv. Model. Earth Syst.*, **7**, 226–243, <https://doi.org/10.1002/2014MS000397>.
- Bissolli, P., and E. Dittmann, 2001: The objective weather type classification of the German Weather Service and its possibilities of application to environmental and meteorological investigations. *Meteor. Z.*, **10**, 253–260, <https://doi.org/10.1127/0941-2948/2001/0010-0253>.
- Brooks, H. E., J. W. Lee, and J. P. Craven, 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.*, **67–68**, 73–94, [https://doi.org/10.1016/S0169-8095\(03\)00045-0](https://doi.org/10.1016/S0169-8095(03)00045-0).
- , G. W. Carbin, and P. T. Marsh, 2014: Increased variability of tornado occurrence in the United States. *Science*, **346**, 349–352, <https://doi.org/10.1126/science.1257460>.
- Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, and D. T. Lindsey, 2014: An empirical model for assessing the severe weather potential of developing convection. *Wea. Forecasting*, **29**, 639–653, <https://doi.org/10.1175/WAF-D-13-00113.1>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Diendorfer, G., K. L. Cummins, and W. Schulz, 2010: EUCLID—State of the art lightning detection. EUMETNET Lightning Task Force Tech. Doc., 21 pp.
- Diffenbaugh, N. S., M. Scherer, and R. J. Trapp, 2013: Robust increases in severe thunderstorm environments in response to greenhouse forcing. *Proc. Natl. Acad. Sci. USA*, **110**, 16 361–16 366, <https://doi.org/10.1073/pnas.1307758110>.
- Doswell, C. A., 2007: Small sample size and data quality issues illustrated using tornado occurrence data. *Electron. J. Severe Storms Meteor.*, **2**, <http://ejssm.org/ojs/index.php/ejssm/article/view/26/26>.
- Dotzek, N., P. Groenemeijer, B. Feuerstein, and A. M. Holzer, 2009: Overview of ESSL's severe convective storms research using the European Severe Weather Database ESWD. *Atmos. Res.*, **93**, 575–586, <https://doi.org/10.1016/j.atmosres.2008.10.020>.
- Gensini, V. A., and T. L. Mote, 2014: Estimations of hazardous convective weather in the United States using dynamical downscaling. *J. Climate*, **27**, 6581–6589, <https://doi.org/10.1175/JCLI-D-13-00777.1>.
- , and —, 2015: Downscaled estimates of late 21st century severe weather from CCSM3. *Climatic Change*, **129**, 307–321, <https://doi.org/10.1007/s10584-014-1320-z>.
- Groenemeijer, P., and T. Kühne, 2014: A climatology of tornadoes in Europe: Results from the European Severe Weather Database. *Mon. Wea. Rev.*, **142**, 4775–4790, <https://doi.org/10.1175/MWR-D-14-00107.1>.
- , and Coauthors, 2017: Severe convective storms in Europe: Ten years of research at the European Severe Storms Laboratory. *Bull. Amer. Meteor. Soc.*, **98**, 2641–2651, <https://doi.org/10.1175/BAMS-D-16-0067.1>.
- Hart, J. A., and A. E. Cohen, 2016: The Statistical Severe Convective Risk Assessment Model. *Wea. Forecasting*, **31**, 1697–1714, <https://doi.org/10.1175/WAF-D-16-0004.1>.
- Hastie, T., and R. Tibshirani, 1986: Generalized additive models. *Stat. Sci.*, **1**, 297–310, <https://doi.org/10.1214/ss/1177013604>.
- Johns, R. H., and C. A. Doswell III, 1992: Severe local storms forecasting. *Wea. Forecasting*, **7**, 588–612, [https://doi.org/10.1175/1520-0434\(1992\)007%3C0588:SLSF%3E2.0.CO;2](https://doi.org/10.1175/1520-0434(1992)007%3C0588:SLSF%3E2.0.CO;2).
- Kapsch, M.-L., M. Kunz, R. Vitolo, and T. Economou, 2012: Long-term trends of hail-related weather types in an ensemble of

- regional climate models using a Bayesian approach. *J. Geophys. Res.*, **117**, D15107, <https://doi.org/10.1029/2011JD017185>.
- Markowski, P., and Y. Richardson, 2010: *Mesoscale Meteorology in Midlatitudes*. John Wiley & Sons, 430 pp.
- Mohr, S., M. Kunz, and K. Keuler, 2015: Development and application of a logistic model to estimate the past and future hail potential in Germany. *J. Geophys. Res. Atmos.*, **120**, 3939–3956, <https://doi.org/10.1002/2014JD022959>.
- , —, A. Richter, and B. Ruck, 2017: Statistical characteristics of convective wind gusts in Germany. *Nat. Hazards Earth Syst. Sci.*, **17**, 957–969, <https://doi.org/10.5194/nhess-17-957-2017>.
- National Academies of Sciences, Engineering, and Medicine, 2016: *Attribution of Extreme Weather Events in the Context of Climate Change*. National Academies Press, 200 pp., <https://doi.org/10.17226/21852>.
- Nelder, J. A., and R. W. M. Wedderburn, 1972: Generalized linear models. *J. Roy. Stat. Soc.*, **135A**, 370–384, <https://doi.org/10.2307/2344614>.
- Poelman, D. R., W. Schulz, G. Diendorfer, and M. Bernardi, 2016: The European lightning location system EUCLID—Part 2: Observations. *Nat. Hazards Earth Syst. Sci.*, **16**, 607–616, <https://doi.org/10.5194/nhess-16-607-2016>.
- Pučík, T., P. Groenemeijer, D. Rýva, and M. Kolář, 2015: Proximity soundings of severe and nonsevere thunderstorms in central Europe. *Mon. Wea. Rev.*, **143**, 4805–4821, <https://doi.org/10.1175/MWR-D-15-0104.1>.
- , and Coauthors, 2017: Future changes in European severe convection environments in a regional climate model ensemble. *J. Climate*, **30**, 6771–6794, <https://doi.org/10.1175/JCLI-D-16-0777.1>.
- Punge, H. J., K. M. Bedka, M. Kunz, and A. Werner, 2014: A new physically based stochastic event catalog for hail in Europe. *Nat. Hazards*, **73**, 1625–1645, <https://doi.org/10.1007/s11069-014-1161-0>.
- Puskeiler, M., M. Kunz, and M. Schmidberger, 2016: Hail statistics for Germany derived from single-polarization radar data. *Atmos. Res.*, **178–179**, 459–470, <https://doi.org/10.1016/j.atmosres.2016.04.014>.
- R Core Team, 2015: *R: A Language and Environment for Statistical Computing*. The R Foundation for Statistical Computing, <https://www.R-project.org/>.
- Schmeits, M. J., K. J. Kok, and D. H. P. Vogelegang, 2005: Probabilistic forecasting of (severe) thunderstorms in the Netherlands using model output statistics. *Wea. Forecasting*, **20**, 134–148, <https://doi.org/10.1175/WAF840.1>.
- Schulz, W., G. Diendorfer, S. Pedeboy, and D. R. Poelman, 2016: The European lightning location system EUCLID—Part 1: Performance analysis and validation. *Nat. Hazards Earth Syst. Sci.*, **16**, 595–605, <https://doi.org/10.5194/nhess-16-595-2016>.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.
- Thorne, P., and R. Vose, 2010: Reanalyses suitable for characterizing long-term trends. *Bull. Amer. Meteor. Soc.*, **91**, 353–362, <https://doi.org/10.1175/2009BAMS2858.1>.
- Tippett, M. K., A. H. Sobel, and S. J. Camargo, 2012: Association of U.S. tornado occurrence with monthly environmental parameters. *Geophys. Res. Lett.*, **39**, L02801, <https://doi.org/10.1029/2011GL050368>.
- Trapp, R. J., B. A. Halvorson, and N. S. Diffenbaugh, 2007: Telescoping, multimodel approaches to evaluate extreme convective weather under future climates. *J. Geophys. Res.*, **112**, D20109, <https://doi.org/10.1029/2006JD008345>.
- , N. S. Diffenbaugh, and A. Gluhovsky, 2009: Transient response of severe thunderstorm forcing to elevated greenhouse gas concentrations. *Geophys. Res. Lett.*, **36**, L01703, <https://doi.org/10.1029/2008GL036203>.
- Van Klooster, S. L., and P. J. Roebber, 2009: Surface-based convective potential in the contiguous United States in a business-as-usual future climate. *J. Climate*, **22**, 3317–3330, <https://doi.org/10.1175/2009JCLI2697.1>.
- Verbout, S. M., H. E. Brooks, L. M. Leslie, and D. M. Schultz, 2006: Evolution of the U.S. tornado database: 1954–2003. *Wea. Forecasting*, **21**, 86–93, <https://doi.org/10.1175/WAF910.1>.
- Viceto, C., M. Marta-Almeida, and A. Rocha, 2017: Future climate change of stability indices for the Iberian Peninsula. *Int. J. Climatol.*, **37**, 4390–4408, <https://doi.org/10.1002/joc.5094>.
- Westermayer, A. T., P. Groenemeijer, G. Pistotnik, R. Sausen, and E. Faust, 2017: Identification of favorable environments for thunderstorms in reanalysis data. *Meteor. Z.*, **26**, 59–70, <https://doi.org/10.1127/metz/2016/0754>.
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. International Geophysics Series, Vol. 100, Academic Press, 648 pp.
- Wood, S. N., 2003: Thin plate regression splines. *J. Roy. Stat. Soc.*, **65B**, 95–114, <https://doi.org/10.1111/1467-9868.00374>.
- , 2006: *Generalized Additive Models: An Introduction with R*. CRC Texts in Statistical Science Series, CRC Press, 410 pp.
- , 2017: Package ‘mgcv.’ R Project Tech. Rep., 290 pp., <https://cran.r-project.org/web/packages/mgcv/mgcv.pdf>.