

A Poisson regression approach to model monthly hail occurrence in Northern Switzerland using large-scale environmental variables

Erica Madonna^{a,b,*}, David Ginsbourger^{c,d}, Olivia Martius^a

^a*Oeschger Centre for Climate Change Research, Institute of Geography, University of Bern, Switzerland*

^b*Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Norway*

^c*Idiap Research Institute, Martigny, Switzerland*

^d*IMSV, Department of Mathematics and Statistics, University of Bern, Switzerland*

Abstract

In Switzerland hail regularly causes substantial damage to agriculture, cars and infrastructure, however, little is known about its long-term variability. To study the variability, the monthly number of days with hail in northern Switzerland is modeled in a regression framework using large-scale predictors derived from ERA-Interim reanalysis. The model is developed and verified using radar-based hail observations for the extended summer season (April-September) in the period 2002-2014. The seasonality of hail is explicitly modeled with a categorical predictor (*month*) and monthly anomalies of several large-scale predictors are used to capture the year-to-year variability. Several regression models are applied and their performance tested with respect to standard scores and cross-validation.

The chosen model includes four predictors: the monthly anomaly of the two meter temperature, the monthly anomaly of the logarithm of the convective available potential energy (CAPE), the monthly anomaly of the wind shear and the month. This model well captures the intra-annual variability and slightly underestimates its inter-annual variability. The regression model is applied to the reanalysis data back in time to 1980. The resulting hail day time-series shows an increase of the number of hail days per month, which is (in the model) related to an increase in temperature and CAPE. The trend corresponds to approximately 0.5 days per month per decade. The results of the regression model have been compared to two independent data sets. All data sets agree on the sign of the trend, but the trend is weaker in the other data sets.

Keywords: hail, regression models, insurance data, inter-annual variability

1. Introduction

Hail can cause substantial damage to cars, buildings and crops and it can injure livestock or people. It is one of the costliest natural hazards in Switzerland (VKF, 2013; BAFU, 2016). Three main ingredients are needed for the development of thunderstorms and hailstorms: an unstable atmosphere, lower tropospheric moisture and a lifting mechanism, i.e. the thunderstorm trigger (e.g., Johns and Doswell III, 1992). The relative importance of these three ingredients depends on regional factors, such as terrain barriers, thermotopographic wind systems, or warm water surfaces and varies regionally. For example in the U.S. large values of convective available potential energy (CAPE) are essential for severe convection, while in Europe synoptic and mesoscale conditions (as the presence of fronts) seem more important (Brooks et al., 2003, 2007). Several studies (Cacciamani et al., 1995; Costa et al., 2001; Giaiotti et al., 2003; Kunz and Puskeiler, 2010) also pointed out the importance of the interaction between the synoptic flow and complex orography for the development of hail in Europe. Indeed, (Cacciamani et al. 1995), showed that in the Po valley only a small number of thunderstorms develops purely thermodynamically, while the majority of them are associated with synoptic-scale dynamical forcing. Several studies (e.g., Costa et al. 2001, Giaiotti et al. 2003, Kunz and Puskeiler 2010) also pointed out the importance of the interaction between the synoptic flow and complex

*Corresponding author

Email address: erica.madonna@uib.no (Erica Madonna)

orography for the development of hail. In Switzerland, most of the hail events occurs in the pre-alpine regions, north and south of the Alps (Nisi et al., 2016, see also Fig. 1a), and 20-40% of all the hail events are related to surface fronts (Schemm et al., 2016).

Hail occurrence varies not only in space but also in time, exhibiting a large year-to-year variability (see e.g., Mohr et al. 2015a for Europe or Nisi et al. 2016 for Switzerland) but also complex long-term trends. For example, Berthet et al. (2011) analyzed hailpads in France for the period 1989-2009, observing no significant changes in the frequency of hail, but a significant change in the intensity of the hail events. Similar results were found in Italy by Eccel et al. (2012). However, the sign of hail trends can vary considerably in adjacent regions (Hermida et al., 2015) and their strength depends on the considered time period (Sánchez et al., 2017). A detailed overview of hail trends over Europe is provided by Punge and Kunz (2016). Unfortunately, studying hail trends and low frequency variability is limited in most areas of the world by the sparse availability of long-term direct observations (e.g., Martius et al., 2015).

One way to address this limitation and to study the long-term variability is to use large-scale variables as proxies for hailstorm formation. For example stability indices, which can stem from radiosonde data or from reanalyses data sets, can be used as proxies (see e.g., Mohr and Kunz, 2013). The main advantage of using reanalysis data is their spatial and temporal coverage and their availability over longer time periods. The large-scale variables can be validated using direct observation from hailpads, indirect observations from radar reflectivity, satellite cloud top temperatures, or loss data from insurance companies (e.g., Schiesser, 1990; Kunz and Puskeiler, 2010; Cintineo et al., 2012; Cecil and Blankenship, 2012; Merino et al., 2014a; Mohr et al., 2015b; Melcón et al., 2016; Nisi et al., 2016). The ability of various large-scale variables to capture severe (hail producing) thunderstorms have been studied for various regions around the world, such as the U. S. (e.g., Brooks et al., 2003; Rasmusen, 2003; Johnson and Sugden, 2014), Europe (e.g., Groenemeijer and van Delden, 2007; Kunz, 2007; García-Ortega et al., 2012; Manzato, 2012; Mohr and Kunz, 2013; Gascón et al., 2015; Púčik et al., 2015; Tuovinen et al., 2015; Melcón et al., 2017) and Australia (e.g., Allen et al., 2011).

To predict the hail potential of the atmosphere, often a combination of several large-scale variables is needed. An extensive summary of hail relevant large-scale variables, which can be used as environmental proxies, is provided by Mohr and Kunz (2013) for Europe, and by Huntrieser et al. (1997) specifically for Switzerland. Some large-scale variables describe conditional instabilities (e.g., Vertical Total (Miller, 1972)) or latent instabilities (e.g., Lifted Index (Galway, 1956), Showalter Index (Showalter, 1953), CAPE (Moncrieff and Miller, 1976), convective inhibition (CIN)). Others characterize the moisture content in the troposphere (e.g., the vertically integrated liquid water (Greene and Clark, 1972)) or the kinematic conditions (e.g., vertical wind shear (Weisman and Klemp, 1982, 1984)).

The large-scale variables can then be used to develop statistical models, which reproduce or predict hail occurrence. For example, based on daily proxies derived from regional downscaled reanalysis and regional climate model simulations, Mohr et al. (2015b) estimated the potential for hailstorms occurrence in Germany for the period 1971-2000 and 2021-2050 using a logistic regression approach, finding a future increase (but mostly not statistical significant) in the potential of hail. Using monthly means, Allen et al. (2015) developed a Poisson regression model which relates the monthly hail occurrence to the large-scale atmospheric environment in the U.S.. Our motivation to look at the monthly number of hail days is similar, namely to identify very active hail months, to study trends, and to identify potential large-scale drivers of active hail months. There is substantial inter-annual variability in the number of hail days per month in northern Switzerland, e.g., the number of hail days in July varies between 5 and 15 (Fig. 1b). For a number of insurance and agricultural applications it would be important to characterize this variability and to understand the underlying atmospheric drivers.

A first goal of this paper is to determine a set of best reanalysis-based hail predictors for the orographic setting of Northern Switzerland. A second goal is to identify a suitable regression model to study the monthly hail occurrence and its variability in Northern Switzerland. Applying this regression model to a reanalysis data set, we will then analyze the inter-annual hail variability over a 35-year time period.

The paper is structured as follows: Section 2 introduces the data and method, Section 3 frames the development of the statistical model and the choice of the relevant large-scale parameters, Section 4 presents the results from the statistical model, followed by a discussion in Section 5 and some concluding remarks given in Section 6.

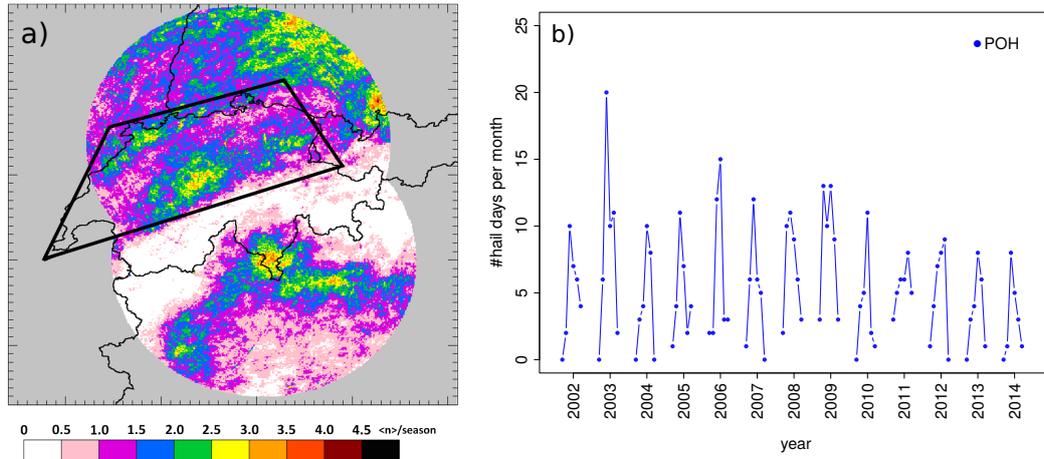


Figure 1: (a) Seasonal average (April-September, 2002-2014) of the number of days with hail per km^2 (i.e. $POH > 80\%$), reproduced from Nisi et al. (2016). The domain in the northern part of Switzerland is marked by the black trapeze. (b) Time series of the number of days with hail per month for April-September 2002-2014.

2. Data and Methods

2.1. Probability of Hail (POH)

Several studies (e.g., Waldvogel et al., 1979; Witt et al., 1998; Foote et al., 2005) used volumetric radar information to estimate the probability of hail (POH) at the ground. The hail data in this study is based on the method of Foote et al. (2005) to estimate POH over Switzerland. The calculation is based on the 45 dBZ EchoTop height from the radar measurement and the freezing level height provided by the numerical weather prediction model COSMO-CH (Baldauf et al., 2011). POH was validated by Nisi et al. (2016) with hail damage from a car insurance company in Switzerland. Nisi et al. (2016) found a POH of 80% as the best threshold to capture hail reaching the ground. Based on these findings, we define days with a POH exceeding 80% over an area of at least $100 km^2$ as days with hail. Different area thresholds (50 - $500 km^2$) have been tested, showing no substantial influence on the results. We do not make any distinction between the size of hailstones. We focus on a domain in the northern part of Switzerland (black trapeze, Fig. 1a). Radar-based hail estimates are indirect observations of hail based on empirical relationships. In complex terrain like in Switzerland a number of issues such as ground clutter or beam shielding can affect the radar observations. A detailed discussion of the effect on these radar issues on the radar hail product as well as a verification of the radar hail product used here is presented in Nisi et al. (2016). They show that shielding effects are negligible due to the high vertical extent of the hail storms, but some ground clutter issues remain. In mitigation we note that here we focus on an area north of the Alps and that we use a threshold of $100 km^2$ of likely hail occurrence to count a day as a hail day. This reduces the effect of clutter. We then sum the number of days with hail for each month of the extended summer season (April to September) for the period 2002 to 2014 and regard these as the “observed” hail occurrences (blue dots, Fig. 1b). Note that with this approach only information about the number of days with hail is evaluated but not the severity of the individual hail days.

2.2. Insurance data

An estimate of the number of hail days per month in the past can be derived using insurance data. Since 1949 agricultural hail damage claims are available from the Swiss Hail Insurance Company, the only insurance company in Switzerland that insures crop damages caused by hail. Schiesser (1990); Willemse (1995); Schiesser et al. (1997) have analyzed and homogenized this data. The damage claims are an indirect measure of hail, because the crop damages depend on the vulnerability of specific cultures, hail prevention measures, and the insurance participation in the area (see Willemse, 1995, for a detailed discussion). Nevertheless they currently constitute the best available estimate of hail occurrence for the time period before radar availability.

name	acronym	unit
Total integrated liquid water	<i>TILW</i>	mm
Surface Lifted Index	<i>SLI</i>	°C
Mixed Layer Convective Available Potential Energy	<i>MLCAPE</i>	J/kg
Natural logarithm of <i>MLCAPE</i>	$\ln(MLCAPE)$	J/kg
Mixed Layer Convective Convective Inhibition	<i>MLCIN</i>	J/kg
Wind shear (10m-500 hPa)	<i>SH</i>	m/s
Two meter temperature	<i>T2m</i>	°C

Table 1: Large-scale variables analyzed in the regression models.

For each day, information about the number of municipalities with damage claims in Northern Switzerland is available. Most municipalities in Switzerland are smaller than 100 km^2 , the area threshold applied to the radar data. The question is therefore how many municipalities need to be affected for a day to correspond to a hail day as defined using the radar data? To answer this question, we compare the radar data and insurance claims on a daily basis while systematically varying the number of affected municipalities used as a threshold to define a hail day. The correlation coefficient is maximized ($r = 0.69$) if we use a threshold of seven affected municipalities to define a hail day. Also for the insurance data we do not distinguish between the severity of the events, but we retain only information about the number of days with hail in a month.

2.3. Large-scale predictors for the statistical model

The ERA-Interim reanalysis data (Dee et al., 2011) from 1980 to 2014 is used, interpolated to a horizontal grid of 1° , with 61 vertical levels, and a temporal resolution of 6 hours. As discussed above the environment favorable for hail development varies regionally and for this reason, the choice of the large-scale variables, listed in Table 1, is based on previous studies performed over Switzerland (Huntrieser et al., 1997), Germany (Kunz, 2007; Mohr and Kunz, 2013; Mohr et al., 2015b) and Europe (Mohr et al., 2015a). The total integrated liquid water (TILW) is an estimate for the moisture content of the atmosphere. Latent instabilities are measured by the Surface Lifted Index (SLI), the Mixed Layer Convective Available Potential Energy (MLCAPE) and the Mixed Layer Convective Inhibition (MLCIN). The vertical wind shear between 10 meters and 500 hPa (SH) is used as an indicator for kinematic conditions. Furthermore, the two meter temperature is included as predictor.

Reanalyses data with a 6-hourly resolution are able to capture the large-scale environment in which hail develops but not the thunderstorm triggers. While most hail events in Northern Switzerland occur in the late afternoon (Nisi et al., 2016), we do not know exactly at which time the vertical profile best captures the hail related instabilities. For this reason we take for each day and each variable the maximum daily value averaged over $47 - 48^\circ N$ and $7 - 9^\circ E$ (i.e. the grid points covering Northern Switzerland). For negative variables as *SLI*, we take the lowest value. We then compute monthly means for each variable, similar to Tippet et al. (2014) and Allen et al. (2015). Moreover, since *MLCAPE* can take values which range from zero to more than 2000 J/kg , a single extreme value can significantly affect the monthly mean and we are interested in months with several hail days. To reduce this effect, following Allen et al. (2015), the monthly mean of the natural logarithm of *MLCAPE* is considered (alternatively instead of the mean value, the median could be used).

Here we work with a limited number of large-scale variables. One reason for this choice is that many variables are highly correlated (e.g., melting layer height and CAPE, Dessens et al., 2015) and highly correlated variables affect the estimation of the regression coefficients (see section 3.1.2 for more details). Another reason is that some variables (e.g., the maximum potential speed of an updraft, Gascón et al., 2015) are meaningful on daily basis, but less on monthly basis (due to the relatively coarse resolution of the data as well as the large day to day variability of the variable within a month). Finally, some variables, for example the melting layer height, are helpful to capture the variations in the hailstone size distributions (e.g., Dessens et al., 2015). Here we do not use this information and are interested only in the number of days with hail of any size.

2.4. The seasonal cycle

Hail is a seasonal phenomenon and it is most frequent in the northern part of Switzerland during the summer months of June and July (Nisi et al., 2016). To represent this seasonal behavior in our model, a categorical predictor

variable (factor), *month*, is introduced. Also most large-scale variables exhibit a strong seasonal cycle, which can as well be represented by the variable *month*. Thus, the seasonal cycle is removed from each variable X and monthly anomalies (X_a) are calculated. In other words, the monthly climatology ($X_{clim,m}$) of each variable X , computed for the period 1980-2014, is subtracted from the monthly mean X_m to obtain the monthly anomaly ($X_a = X_m - X_{clim,m}$).

2.5. Statistical models

Statistical models are widely used to model count data, as for example, the number of tropical storms (e.g., McDonnell and Holbrook, 2004; Villarini et al., 2010; Tippett et al., 2011), the monthly occurrence of tornados (Tippett et al., 2012, 2014) or hail events (Allen et al., 2015). Similarly to the last study, we want to model the number of hail days per month based on one or more predictors (our large-scale variables). For that purpose we first have to choose a statistical model fitting our empirical observations. The Poisson model is a generalized linear model for count data (e.g., Wilks, 2011). Typical issues that have to be tackled when applying a Poisson model include variable selection, model validation, and dealing with potential over-dispersion.

Poisson probability distributions can be considered as a canonical option for modeling count data. Given a parameter $\mu > 0$ called the intensity, let us denote the Poisson distribution with intensity μ by $Po(\mu)$. A random variable Y with non-negative integer values is defined to be $Po(\mu)$ -distributed when its probability mass function is given by:

$$Pr\{Y = y\} = \frac{\mu^y e^{-\mu}}{y!} \quad (y \in \{0, 1, 2, \dots\}) \quad (1)$$

where $Pr\{Y = y\}$ is the probability of occurrence of exactly y events (e.g., Wilks, 2011). Poisson distributions form a so-called exponential family, which comes with a number of useful algebraic properties. In addition, the Poisson family is stable by addition, in the sense that the sum of independent Poisson-distributed random variables is Poisson-distributed. Let us recall also that for a $Po(\mu)$ -distributed Y the intensity μ is at the same time the expectation of Y and its variance. Hence, cases of count data distributions with a variance substantially larger than the expectation are referred to as over-dispersed compared to the Poisson distribution and, as briefly tackled later, several distributions have been considered to deal with over-dispersion issues.

In Poisson regression, which corresponds to predicting under a Generalized Linear Model with a link function naturally associated with the Poisson distribution, the intensity μ is assumed to depend on predictor variables X_1, \dots, X_p in a log-linear way:

$$\mu = \exp(\beta_0 + \beta_1 \cdot X_1 + \dots + \beta_p \cdot X_p) \quad (2)$$

Classically, the regression coefficients β_i are estimated by maximum likelihood (e.g., Wilks, 2011; Hilbe and Robinson, 2013) under the assumption that observations are independently Poisson distributed with intensities given by eq. (2). The probability of the postulated model having generated the observed responses is maximized with respect to β_0, \dots, β_p . The corresponding optimization problem is concave, and efficient maximization algorithms offering convergence guarantees under standard assumptions are available in common softwares.

The Poisson family having only one free parameter μ may be seen as an advantage in terms of simplicity, however, this family often has a limited applicability in practical cases, e.g., it cannot accommodate over-dispersion nor excesses of zero counts. Alternative distributions that account for over-dispersion have been proposed, including notably the Negative Binomial distribution (e.g., Villarini et al., 2010). As for the abundance of zero counts that need to be accounted for in a number of applications, two main variants of the Poisson distribution have been proposed: the zero-inflated Poisson distribution and the hurdle model (see Zeileis et al., 2008, for details). The last distribution will be considered within some of the candidate models presented next.

In the next section we investigate regression models based on subsets of the large-scale predictors X_i and their interactions with the aim to achieve good predictions of the hail occurrence while avoiding both multicollinearity and over-fitting. The considered practical approach consists in first removing redundant, i.e. highly correlated, variables by means of the Variance Inflation Factor (VIF, Mansfield and Helms, 1982), and then by running standard model selection algorithms to find a parsimonious set of predictors explaining the response, including first and second order terms in the considered predictors.

Keeping in mind that model selection procedures such as considered here are prone to false discovery and that conducting statistical tests on final models without accounting for the whole sequence of considered models can be

misleading, we mitigate these issues by conducting predictivity assessments based on k-fold cross-validation. The various model building workflows considered in our comparison are described, applied, and assessed in Section 3 below.

3. Model building

The baseline model that we consider is a standard Poisson regression model with the large-scale variables presented in Table 1 and the categorical variable standing for the current month. The model is constructed using the radar based monthly hail days during the extended summer season (April to September) from 2002 to 2014. Starting from there, several variations of the baseline model were tested that rely on model selection procedures driven by the Akaike Information Criterion (AIC, Akaike, 1974). Also, modifications of the Poisson regression accounting for the excess of zeros are considered, and finally a regression approach consisting of a seasonal trend completed by a remainder linearly depending on the predictors is added for comparison. The models included in the numerical experiments are described in more detail below. Before that, we review the metrics used for assessing the fit of the individual models and for comparing them.

A set of metrics (e.g., Wilks, 2011; Hilbe and Robinson, 2013, see also Appendix A) exists that assess how well a generalized linear model, such as the Poisson regression model, fits the training data. Here we apply simple measures typically used in linear modeling such as the Mean Square Error (MSE - the mean of the squared difference between the observed and predicted values) and the Mean Absolute Error (MAE - the mean of the absolute difference between the observed and predicted values) to simplify comparisons between models of different natures (including notably linear and Poisson regression models). In addition the deviance is a measure often used to discriminate between different regression models. However, criteria such as AIC or the Bayesian Information Criterion (BIC), which complement the log-likelihood with a complexity-penalizing term, are more appropriate to quantify the relative quality of regression models while accounting for the number of parameters. In the considered model selection algorithms, AIC is used to select between Poisson regression models with different numbers of predictors, and also as a metric to compare statistical models of different natures (see Tab. 2).

Beyond fit metrics based primarily on comparing predicted versus actual responses from the training data, an even more crucial question is how regression models will perform when applied to new data. When no external validation data is available, cross-validation approaches can be used to estimate the predictivity of statistical models. Note that some kind of external validation is presented as well in the next section, but it involves data of a different kind. Here we use a k-fold cross-validation (CV), where the regression coefficients are calculated by leaving out in a row each year of the data set and comparing at every iteration the predicted number of hail days with the observations for the left out year. A MSE-based score (coined “CV-MSE” in Tab. 2) is then obtained by computing the MSE values corresponding to the left out years. Note that the cross-validation considered here is to be understood as *post model selection*, in the sense that for each training set with a left out year, coefficients are re-calculated but the set of chosen variables is kept constant to the set that is obtained based on the full data set. Note that the cumbersome alternative of letting the selected subset vary within the cross-validation procedure was found not to substantially change results on a sample of dedicated numerical experiments, so the lighter option was preferred here for simplicity.

3.1. Description of the competing statistical models

3.1.1. Model with all large-scale predictors

The first model (1), reported in Table 2, is a Poisson model where the monthly anomalies (a) of all the large-scale variables listed in Table 1 are used as predictors, together with the categorical variable *month*. We assign the name of *no pretreatment* to this model, as no variable has been removed.

3.1.2. Model with removed multicollinearity and AIC optimization

The second model (2) is the result of a two-steps procedure starting from model (1). In a first step, potential multicollinearity is detected and removed using the VIF criterion. In a second step, the predictor variables are selected based on AIC minimization. The method is illustrated in more detail below and it is referred as *pretreatment*.

The inclusion of predictors with strong correlation is known to potentially lead to poor estimates of the regression parameters (e.g., Wilks, 2011). Indeed, multicollinearity can result in sign and magnitude changes of the estimated

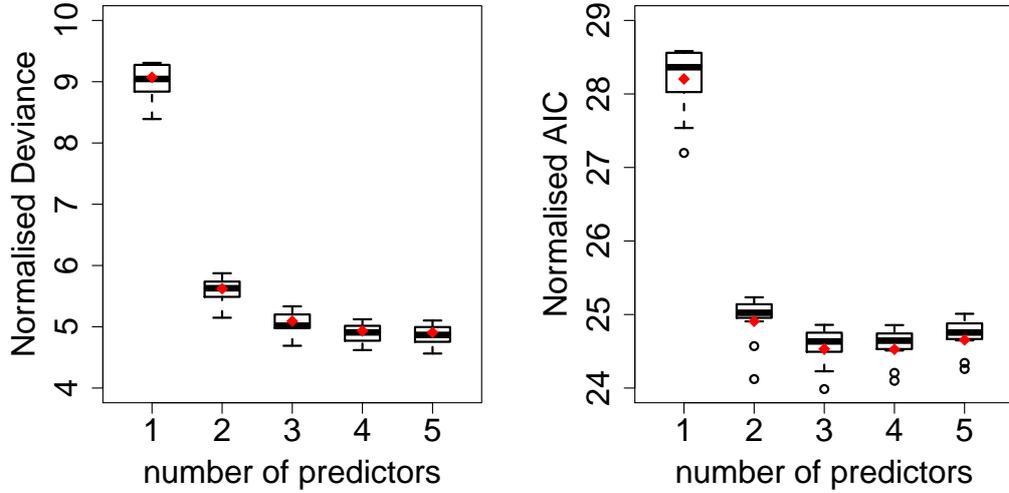


Figure 2: Normalized (i.e. total divided by the number of considered years) a) deviance and b) AIC as function of the number of predictors. The red diamonds show the values for the model using all years (2002-2014), the boxplot (with the median as black solid line) are the results of 13 models with one year left out (cross-validation). The predictors are: 1) *month*, 2) $T2m_a$, 3) $\ln(MLCAPE)_a$, 4) SH_a and 5) $TILW_a$.

regression coefficients. In our case, introducing the categorical variable *month* and using monthly anomalies of each variable (whose absolute values are highly correlated) allows us to remove some redundancy. To furthermore reduce multicollinearity per se, we employ the Variance Inflation Factor (*VIF*, see for example Mansfield and Helms, 1982, and Appendix A). Starting with the anomalies of the large-scale variables reported in Table 1 and the variable *month*, the *VIFs* are first calculated. Then, the predictor with the largest *VIF* is excluded and the *VIFs* are recalculated with the new version of the model, repeating the procedure until all the *VIFs* are small enough (reference values in the literature are 4-10, e.g., O'Brien, 2007). The variable with the largest *VIF* (> 11) is SLI_a , and it is removed together with $MLCIN_a$ and $MLCAPE_a$. After the pretreatment (*VIFs* < 4), the remaining variables are $TILW_a$, $\ln(MLCAPE)_a$, SH_a , $T2m_a$ and *month*.

From there, model (2) is obtained by selecting among models including a combination of the 5 aforementioned predictors, the one with the smallest AIC score. The AIC is minimized with a bidirectional procedure, which combines forward selection and backward elimination and allows adding and removing each predictor at any stage of the selection process.

As an illustration, Figure 2 shows the deviance and AIC values for regression models with an increasing number of predictors (red diamonds). The order considered here for the predictors is based on the p-values from model (1) and is as follows: 1) *month*, 2) $T2m_a$, 3) $\ln(MLCAPE)_a$, 4) SH_a and 5) $TILW_a$. Both deviance and AIC are strongly reduced after the introduction of a second predictor, and smaller reductions are observed up to four predictors. However, with the introduction of the fifth predictor, the deviance almost does not decrease and the AIC increases. The same is observed in the cross validation, where AIC values for each model with a year left out (shown as boxplot in Figure 2) stop decreasing after four predictors. As the AIC considers also the number of included parameters in the model, we continue the analysis with this score.

The final model (2) contains the following four predictors: 1) *month*, 2) $\ln(MLCAPE)_a$, 3) $T2m_a$, 4) SH_a .

3.1.3. Negative binomial model

We use the same procedure described in section 3.1.2 for model (2) using a Negative Binomial probability distribution instead of the Poisson probability distribution. The Negative Binomial probability distribution can account for over-dispersion, which occurs when the variance of the data is larger than its mean. However, we do not find any indication of over-dispersion in our data. Indeed, there is barely any difference in term of predictors, their regression coefficients and validation metrics between the Negative binomial and the Poisson probability distribution. Therefore, the model using the Negative Binomial distribution is not reported in Table 2, as it is close to identical to model (2).

3.1.4. Models with second order interactions

Model (3) and (4) include second order interactions between the predictors. Since the number of models resulting from all possible combinations of all predictors is very large, only the five predictors identified in the pretreatment are considered and a procedure from the R package *glmulti* (Calcagno and De Mazancourt, 2010) is used for model selection. This R package automatically returns all possible unique models involving the specified predictors and their pairwise interactions. Based on the AIC values the best model is selected (model (4)) as well as an average of the best 100 explored models with slightly higher AIC values (here referred to as multi model, (3)).

As model (3) is an average between different models, it is not possible to identify the single predictors and it makes no sense to perform *post model selection* cross validation with it as the regression coefficients are not comparable. Model (4) includes only an interaction term ($\ln(MLCAPE)_a * SH_a$) and three single predictors (*month*, $\ln(MLCAPE)_a$, $T2m_a$).

3.1.5. Two part models, a count and a zero part

Models based on a Poisson distribution such as models (1-4) are often found to underestimate zeros, especially with phenomena like hail where the probability of no hail (as in April, e.g., 2002-2004, Fig. 1) is typically higher in the observations than what a Poisson distribution fitted to whole observation samples predicts. For that reason extensions of the Poisson distribution specifically accounting for zeros have been proposed. One of these extensions is the so-called hurdle model, composed of a part modeling zero counts and another modeling the counting part (Cragg, 1971; Mullahy, 1986; Zeileis et al., 2008). Models (5) and (6) are hurdle models. For simplicity, these are constructed starting from the VIF pretreated predictors and minimizing AIC (as described in section 3.1.2). The main difference between model (5) and (6) consists in the zero part. In (5) all variables are included in the zero part, whereas in (6) only variables with statistical significant zero part coefficients (in this case only $\ln(MLCAPE)_a$) are used.

3.1.6. A linear model

Model (7) is a linear model, modeling the monthly hail occurrence anomalies. First, the seasonal cycle is removed by subtracting from each element of the observed time series the corresponding monthly mean (2002-2014) number of hail days. Second, the obtained residual are fitted with a simple linear model using the variables obtained after the VIF pretreatment. As the seasonal cycle has been removed a priori, the variable *month* is not included in this model.

3.2. Model performance

Next we evaluate the ability of each model to capture the number of hail days for each month. Figure 3a shows in blue dots the observed number of hail days per month for the period 2002-2014, together with the predictions from the seven models (in different colors). The residuals (POH-model(nr)) are shown in Figure 3b. In addition, the metrics used to evaluate the model performance (AIC, MSE, MAE, CV-MSE) are summarized in Table 2. Most of the models with uncorrelated predictors (VIF pretreatment) do not include $TILW_a$, but keep all the other parameters (SH_a , $\ln(MLCAPE)_a$, $T2m_a$, *month*).

From Figure 3 we can see that all models predict a similar number of hail days per month and overall the models tend to predict too few hail days between 2007-2009 and too many after 2011. The simple linear model (7) show the largest errors and its MSE, MAE and AIC values are the highest of all considered models. Its performance after cross-validation is also the worst (see Tab. 2). Apart from model (7), the MAE values vary between 1.36 and 1.44 hail days per month and the MSE mean values (which give more weight to the extreme values compared to the MAE) vary between 2.77 and 3.31. The AICs, which also consider the model complexity, range between 315.9 and 330.0. The CV-MSE (Tab. 2) gives useful information of the model performance when applied to data that were not used for parameter fitting (but partly used in variable selection as explained before). Furthermore, an ideal model should lead to almost no changes of the regression coefficients during the CV procedure.

The model elected as best here is model (4), that minimizes both AIC and CV-MSE scores, and is closely followed by model (2). We choose to work with model (2) for all subsequent analyses. The reason being that model (4) contains a predictor ($\ln(MLCAPE)_a * SH_a$) that is very difficult to interpret meteorologically. The product of the monthly mean SH anomalies and $\ln(MLCAPE)$ anomalies does not contain information if e.g., high SH and high $MLCAPE$ values occurred on the same day or not and this is a crucial information for the meteorological interpretation.

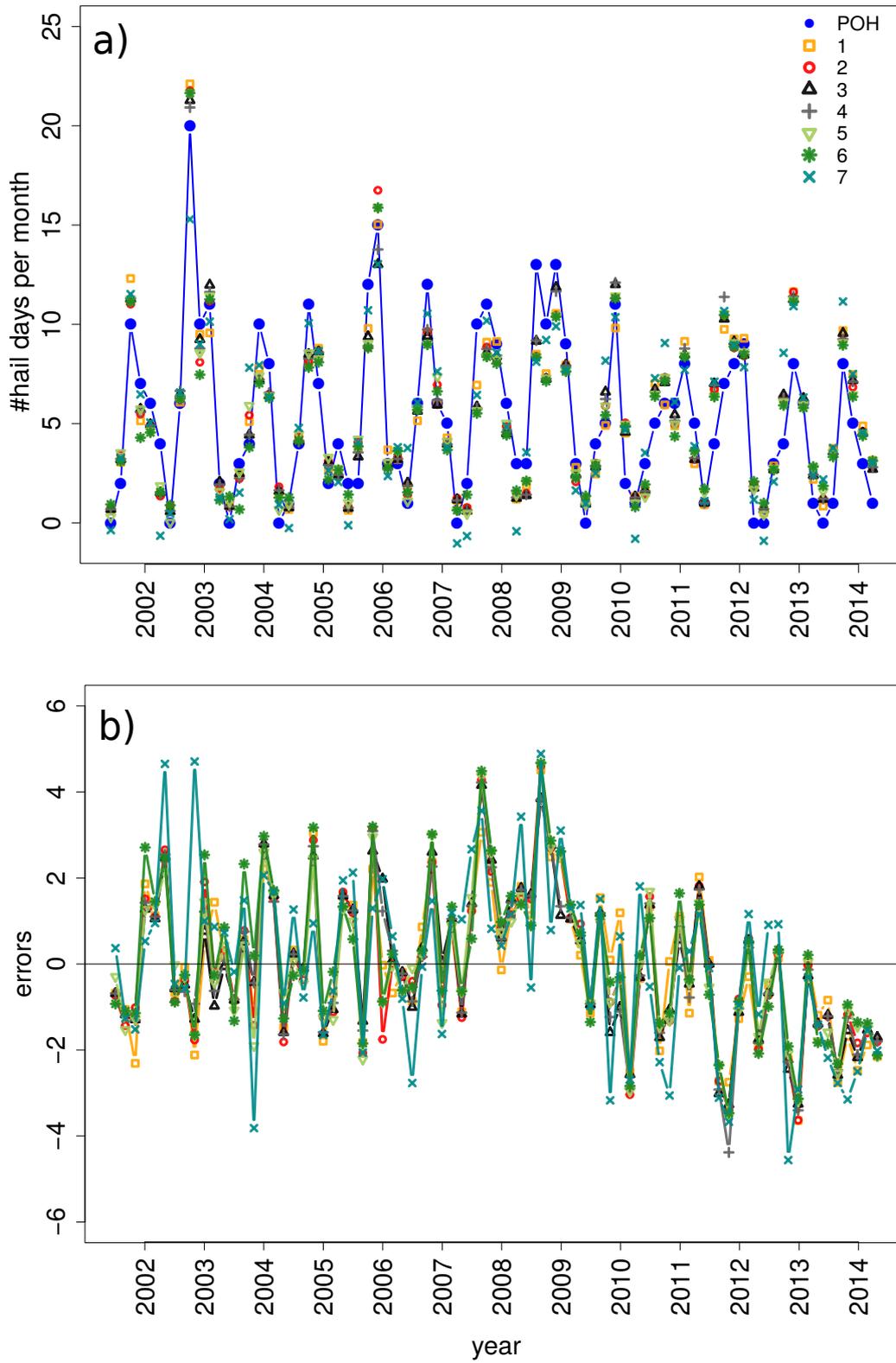


Figure 3: a) Number of hail days per month (April-September) from the POH (blue) and the modeled occurrence using the models described in Table 2. b) Residuals (observed monthly number of days with hail minus estimated) for the models in a).

nr	name	AIC	MSE	MAE	CV-MSE	predictors
1	Poisson model, no pretreatment	322.5	2.77	1.37	4.52	$TILW_a, SH_a, SLL_a, MLCAPe_a, MLCIN_a, \ln(MLCAPe)_a, T2m_a, month$
2	standard: pretreatment VIF and step AIC	318.8	3.01	1.43	4.24	$SH_a, \ln(MLCAPe)_a, T2m_a, month$
3	multi model, VIF pretreatment	320.1	2.74	1.37	-	-
4	best model from multi model (3)	315.9	2.85	1.38	3.96	$\ln(MLCAPe)_a * SH_a, \ln(MLCAPe)_a, T2m_a, month$
5	hurdle, VIF pretreatment	327.1	2.85	1.37	4.20	$SH_a, \ln(MLCAPe)_a, T2m_a, month$
6	hurdle, VIF pretreatment, only $\ln(MLCAPe)_a$ in the zero part	330.0	3.11	1.44	4.33	$SH_a, \ln(MLCAPe)_a, T2m_a, month \ln(MLCAPe)_a$
7	linear model using residual (pretreatment)	344.4	4.15	1.66	5.19	$TILW_a, SH_a, \ln(MLCAPe)_a, T2m_a$

Table 2: Overview of the seven evaluated models, their AIC, MSE and MAE values, CV-MSE and their predictors. In the CV, the MSE is the mean of all MSE calculated for the year not used in the model development. All predictors apart from the categorical variable *month* are monthly anomalies (*a*).

4. Results

4.1. Model verification - seasonal cycle and inter-annual variability

In Section 3 we discussed the steps to select regression model with meaningful predictors. Our chosen model (number (2) from Table 2) to predict the number of hail days per month (μ) is the Poisson model given by:

$$\log(\mu) = \beta_0 + \beta_1 T2m_a + \beta_2 \ln(MLCAPe)_a + \beta_3 SH_a + \sum_{i=4}^8 \beta_i \cdot 1_{month=i+1} \quad (3)$$

where the predictors are $T2m_a$, $\ln(MLCAPe)_a$, SH_a and *month* with the subscript *a* standing for monthly anomalies. The estimated coefficients are $\beta_0 = -0.748$, $\beta_1 = 0.136$, $\beta_2 = 0.248$, $\beta_3 = 0.032$, $\beta_4 = 1.509$ (May), $\beta_5 = 1.965$ (June), $\beta_6 = 2.124$ (July), $\beta_7 = 1.767$ (August), $\beta_8 = 0.837$ (September). We focus only on the summer half year, the variable *month* therefore contains only six months, five of them (May to September) explicitly and one implicitly (April).

The estimated coefficients and associated p-values, as well as standard deviations of the estimates calculated via cross-validation are given in Table 3. The obtained p-values, to be considered with caution given the model selection steps unaccounted for, indicate a significance at the 0.001 level for *month* (apart from September) and $T2m_a$, at a lower significance level for $\ln(MLCAPe)_a$, but not significant at the 5% level for SH_a . The intercept (β_0) is negative but all the other coefficients are positive, indicating an increase in hail days with an increasing value of the quantitative predictors. For the categorical variable *month*, the baseline is April, and so for example the coefficient $\beta_7 = 1.767$ for August means that, under the model, the expected number of days with hail is $\exp(1.767)$ times larger in August than in April, all other variables being constant. The reported standard deviations indicate small to moderate changes in the regression coefficients, which ranges from 4.4% for β_6 to 23.0% for β_3 .

Next we evaluate the performance of the model in reproducing the hail occurrence and its variability. The number of hail days per month in the observations and the number estimated using the Poisson regression (model (2), eq. 3) are shown in Figure 4a (blue dots and red circles respectively). The two time series correlate well (correlation coefficient $r = 0.91$, Fig. 4b). For the 13-year period, 416 days with hail are observed and the same amount is reproduced by the regression model (which is a general property of Poisson regression in the case where parameters are estimated by maximum likelihood and predictors include a constant, as here).

	coefficient	value	p-value	std (%)
intercept	β_0	-0.748	0.031 *	0.108 (14.4%)
$T2m_a$	β_1	0.136	8.01e-05 ***	0.010 (7.3%)
$\ln(MLCAPE_a)$	β_2	0.248	0.011 *	0.029 (11.9%)
SH_a	β_3	0.032	0.150	0.007 (23.0%)
May	β_4	1.509	2.13e-06 ***	0.085 (5.6%)
June	β_5	1.965	1.81e-10 ***	0.090 (4.5%)
July	β_6	2.965	6.48e-12 ***	0.095 (4.4%)
August	β_7	1.767	1.41e-08 ***	0.093 (5.3%)
September	β_8	0.837	0.014 *	0.081 (9.7%)

Table 3: Regression coefficients β_i for each predictor with their p-values (calculated without accounting for the model selection procedure, significance level * (0.05), *** (0.001)), for the final Poisson model (2). The last column shows the standard deviation (std, absolute and relative in %), from the cross validation procedure.

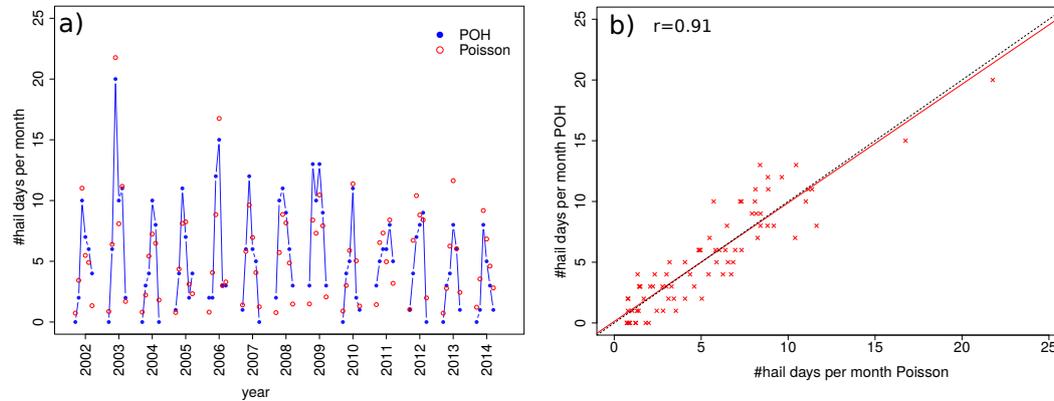


Figure 4: (a) Similar to Figure 3, the observed (derived from POH) number of hail days per month (April-September, blue dots) and the modeled number of hail days using the Poisson regression model (red circles) given in equation 3. (b) Simple linear regression between modeled (x-axis) and observed (y-axis) number of hail days. The black dashed line corresponds to the 1:1 line.

There is a strong seasonal cycle in the number of hail days. The lowest monthly number of predicted hail days occur in April (1 day on average) and in September (2.1 days), followed by May (4.8 days) and August (6.0 hail days) and the highest values are in June (9.2 days) and July (8.8 days). The seasonal cycle is generally well reproduced by the model (see Table 4).

There is large inter- and intra-annual variability in the quality of the statistical model, where some years and some months are better predicted than others. For example, in 8 out of 13 months of April zero hail days were observed. The Poisson model has difficulty to predict zero hail days (zero is at the left extreme of the distribution) and thus the number of hail days in April is often overestimated. However, the number of hail days in April predicted by the models varies and the predicted number of hail days in April is higher if hail was actually observed (e.g., 2009 and 2011). The model is able to capture the two months with the highest number of days with hail, June 2003 (20 days observed, 21.7 days predicted) and July 2006 (15 days observed, 16.7 predicted). The annual maxima are well captured in 2002 and 2010, but they are too high in 2012-2013 and too low in 2007-2009. The variability between the different years is large and the model can capture some, but not all of this variability.

We will next show two example applications of the regression model using the ERA-Interim reanalysis data. First a monthly hail day time series for Northern Switzerland is reconstructed back to 1980. Second the large-scale environmental flow is analyzed for two months of June one with a high number of hail days and one with a low number of hail days.

	POH 2002-2014	regression 2002-2014	regression 1980-2001
April	1.0 (1.2)	1.0 (0.3)	0.7 (0.2)
May	4.8 (3.4)	4.8 (1.9)	3.8 (1.3)
June	9.2 (4.4)	9.2 (4.1)	5.3 (1.1)
July	8.8 (2.9)	8.8 (3.1)	7.6 (3.0)
August	6.0 (2.9)	6.0 (2.4)	5.3 (1.6)
September	2.1 (1.7)	2.1 (0.7)	1.8 (0.7)

Table 4: Mean number of hail days per month and the standard deviation (in brackets) for the period 2002-2014 and 1980-2001. The columns show estimates from POH and from the Poisson regression model (2).

4.2. Application of the model to the ERA-Interim data set

Now that the regression model (2, eq. 3) is selected, it can be used to extend the hail time series back in time based on the large-scale predictors extracted from the ERA-Interim data set (available from 1979).

For the period 1980-2001 a lower number of hail days is estimated for all months compared to the training period 2002-2014 (Fig. 5 and Tab. 4). The differences between the two periods are largest in June (Tab. 4). In the early period (1980-2001) roughly 40% fewer hail days are estimated compared to the recent period and the observations.

The changes in the modeled number of hail days can be related to changes in the anomalies of wind shear (SH_a), instability ($\ln(MLCAPE)_a$) and / or of surface temperatures ($T2m_a$) (see eq. 3). If we consider the changes in the anomalies of the three parameters from the past, we find a small decrease in shear (reduction by approximately 10%) and an increase in both $\ln(MLCAPE)_a$ (+370%) and $T2m_a$ (+27%).

For instance, in the Northern Switzerland the mean April-September surface temperatures for 1980-2001 were on average 0.9°C lower than in 2002-2014. This signal is due mainly to a strong temperature increase in June (+2°C). The increase for the other months (apart from April, which also had an increase of ~ 2°C, but much lower increase in CAPE), is less than 0.5°C. This partly explains why the regression predicts a stronger increase in the number of hail days in June compared to the other months.

The regression predicts lower values for hail in the past, but a similar year to year variability (compare values in the brackets in Table 4), except for June. In June both mean values and standard deviation are lower in the earlier period (5.3 (1.1) days in mean (standard deviation), compared to 9.2 (4.1) in the more recent decade). The months with many days with hail in the regression model in the early period are July 1983 (16.2 days), July 1994 (13.6 days) and July 1995 (11.3 days).

Finally, a first order trend is calculated removing the seasonal cycle, i.e. by subtracting to each month the corresponding long term monthly mean number of hail days. The trend (Fig. 5) indicates a slight increase (approximately 0.5 days per month per decade) in the number of hail days from 1980 to 2014 (t-test significant at 0.001 level). The sign and magnitude of the trend is similar for a set of 1000 subsamples obtained using a bootstrap method (not shown).

4.3. Comparison of a month with few and a month with many hail days

To show an application of the data and to help interpret the large-scale predictors we show a more detailed analysis of the large-scale variables for two example months. We compare the large-scale flow and the large-scale variables considered for the regression model development for a month with very few hail days and one month with many hail days. It is easier to compare the predictors if they have the same climatology and thus we choose the same month (June). The mean number of hail days in June for the 13 analyzed years is 9.2. In June 2004 only 4 hail days occurred (5.4 are predicted), whereas in 2006 12 hail days occurred (8.8 are predicted). We choose these two months for the large-scale variables comparison. We did not choose June 2003 for this comparison, although it is the month with the highest number of hail days (20 days), since this summer was extremely hot in Europe (e.g., Luterbacher et al., 2004).

The climatological June values of four large-scale variables ($TILW$, $T2m$, $MLCAPE$, SH) are shown in the first column of Figure 6, and their anomalies in June 2004 and June 2006 are shown in the second and third column respectively.

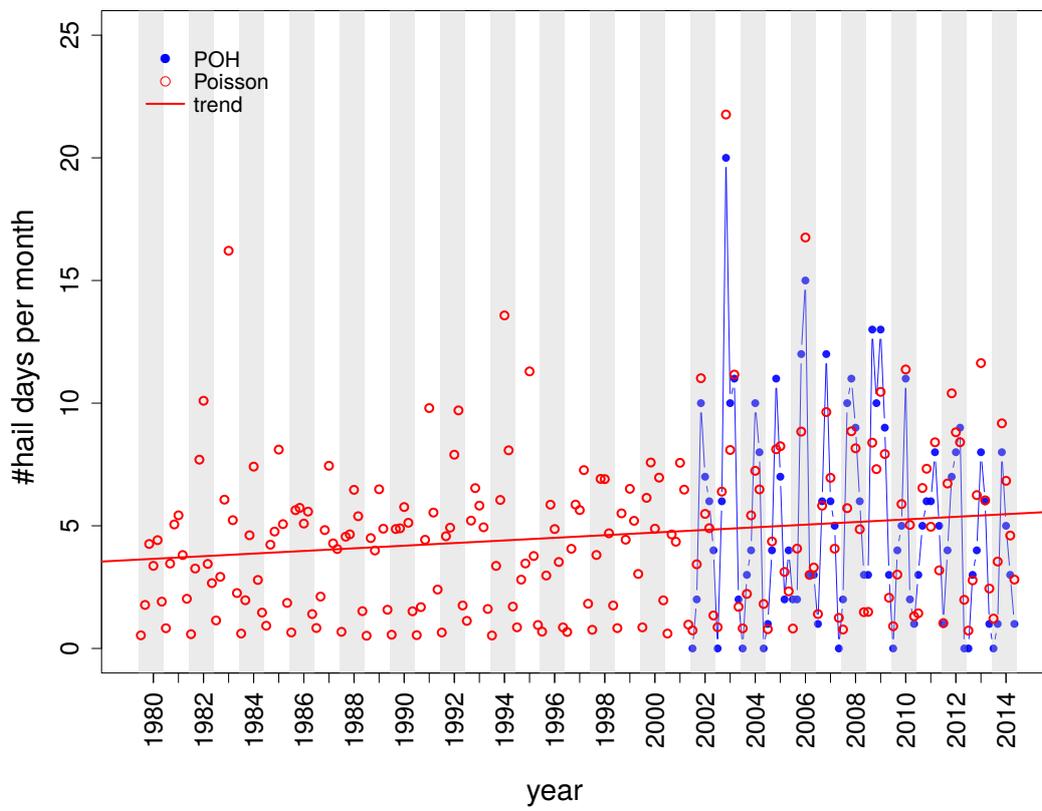


Figure 5: Number of hail days per month (April-September) similar to Fig. 4, but extended back to 1980. The red circles are the results from the regression model, the blue dots show the radar-based observations. A first order trend is added.

The climatological values of the atmospheric moisture content ($TILW$) are higher over the Mediterranean sea (around 25 mm) and lower over the mountain ranges (10-15 mm, Fig. 6, first column, first row). In June 2004 (few hail days) a moderate negative moisture anomaly of -1 to -2 mm was present in the Alpine region, while in June 2006 (many hail days), the atmospheric moisture north of the Alps (black box in Fig. 6) was close to climatology.

In June 2004 temperatures ($T2m$) over Portugal and Spain were warmer than the climatology and the temperatures over the eastern part of Europe were slightly colder (second row, Fig. 6). In June 2006 positive temperature anomalies were present over a major part of the European continent with a positive anomaly of about 2°C over Switzerland. The difference in the sign of the $TILW$ anomalies and the temperature anomalies over Switzerland in June 2006 points reduced evapotranspiration as an increase of $TILW$ with temperature would be expected based on the Clausius-Clapeyron relationship.

The climatology of $MLCAPE$ (third row, Fig. 6) exhibits a SW-NE gradient, with the highest values (around 300 J/kg) found over the Adriatic Sea as well as over Morocco. In June 2004 the maximum values over most of Europe were slightly lower compared to the climatology, whereas in June 2006 the $MLCAPE$ values over the eastern side of the Adriatic Sea were almost twice the climatology and they were also higher over Northern Switzerland (+ 100 J/kg). The picture is similar for $\ln(MLCAPE)$ (not shown). Thus, over Switzerland in 2004 a negative anomaly of $MLCAPE$ and $\ln(MLCAPE)$ was present, whereas the anomaly was positive in 2006.

The last row in Figure 6 shows the wind shear. In June 2004 the wind shear over Northern Switzerland was stronger than in the climatology. A strong zonal jet was located just north of Switzerland (see also Fig. 7). In June 2006 the shear was stronger over the Mediterranean but close to climatology over Switzerland. In 2006 the extratropical jet was located further north at a latitude of about 65°N . These results would suggest that the monthly mean shear is strongly related to the large-scale flow pattern in the upper troposphere (jet). For these two years a higher wind shear over Switzerland was related to fewer hail days and a further discussion in Section 5.3 will address this finding.

Summarizing, we found positive anomalies of the large-scale predictors $T2m$ and $MLCAPE$ over Northern Switzerland in June 2006, the month with more hail days. SH and $TILW$ were close to climatology.

Previous studies showed that specific synoptic configurations (e.g., a trough over the west coast of Europe) were favorable for hailstorm development over France (Merino et al., 2014b; Melcón et al., 2017; Sánchez et al., 2017). Similarly, it is interesting to see whether the aforementioned large-scale variables were related to specific circulation anomalies (e.g., sea surface temperatures, jet stream, blocking). The sea surface temperatures (SSTs) over the Mediterranean were close to climatology during both months (shading in Fig. 7a and Fig. 7b). In the Atlantic positive SST anomalies were located in the East Atlantic in 2004 and over the Labrador Sea in 2006.

In June 2004 Switzerland was located between a high pressure system over the eastern Atlantic and a low pressure system over Scandinavia (contours, Fig. 7a) and hence directly underneath the extratropical jet over Europe. The winds (arrows, Fig. 7a) were blowing from the northwest. In June 2004 an enhanced blocking frequency is found south and west of Greenland (shading, Fig. 7c). With this large-scale flow pattern a frequent passage of low pressure systems and fronts over Northern Switzerland is to be expected. Indeed a comparison of the hail days with the passage of cold fronts over Europe showed that the events in 2004 were mostly pre-frontal (not shown).

In June 2006 Switzerland was located in a high pressure area with its center over Western Europe (Fig. 7b). Two jet streams were well separated; the extratropical jet stream was located over Scandinavia and the subtropical jet stream over North Africa (Fig. 7b). This corresponds to the typical flow situation during a blocking episode over central Europe (see also the enhanced blocking frequency in Fig. 7d). This circulation anomaly is in good agreement with the positive temperature anomalies over most of Europe (cf. second row in Fig. 6). While the low pressure systems crossed Europe north of Switzerland, the cold fronts extended further south and most hail occurred in a pre-frontal environment also in this month (not shown). In contrast to June 2004, the fronts encountered more unstable local conditions over Switzerland.

4.4. Comparison with independent data

Next we compare our model results with insurance data and with hail estimates from another study.

First our estimates are compared to agricultural insurance data from the Swiss hail Insurance company. For this comparison we consider days with damage claims from at least seven municipalities as hail days (see section 2.2).

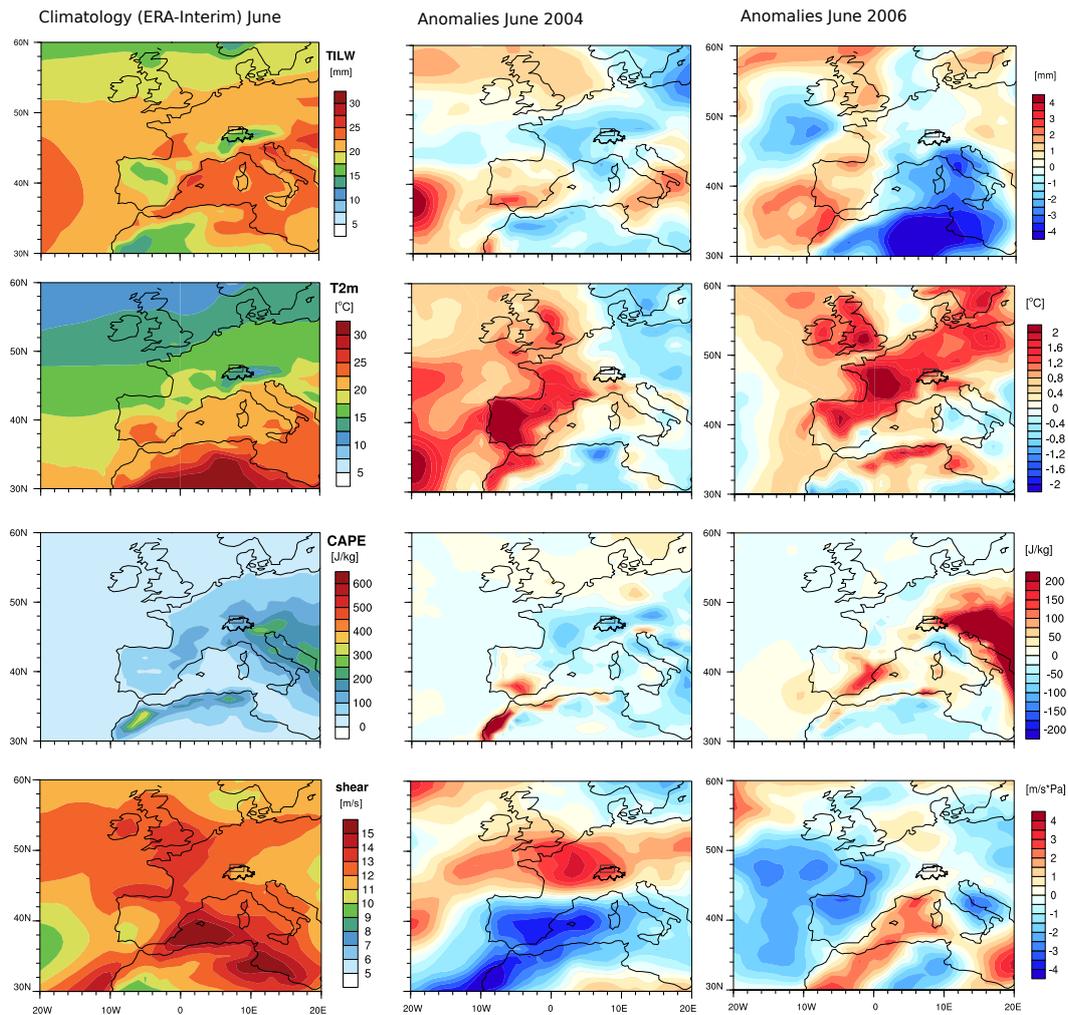


Figure 6: First column: June climatology of *TILW* in mm, *T2m* in °C, *MLCAPE* in J/kg and *SH* in m/s. Second column: monthly mean anomalies of the same quantities for June 2004 (few days with hail). Third column: anomalies for June 2006 (many days with hail). The black box over Switzerland show the considered grid box used for the calculation of the predictors.

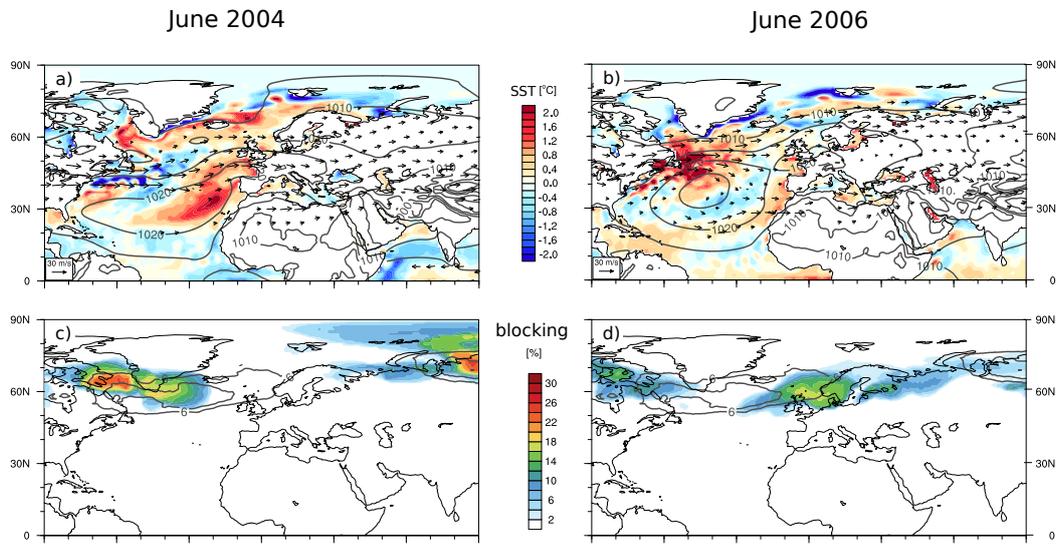


Figure 7: Mean sea level pressure (gray contours, in hPa), upper level (250 hPa) wind (black arrows, in m/s) and sea surface temperature anomalies (SST, shading, in °C) for (a) June 2004 and (b) June 2006. Blocking climatology (ERA-Interim 1980-2014, gray contours, % of time, from 6%, 2% interval) and blocking anomalies (shading, % of time) for (c) June 2004 and (d) June 2006.

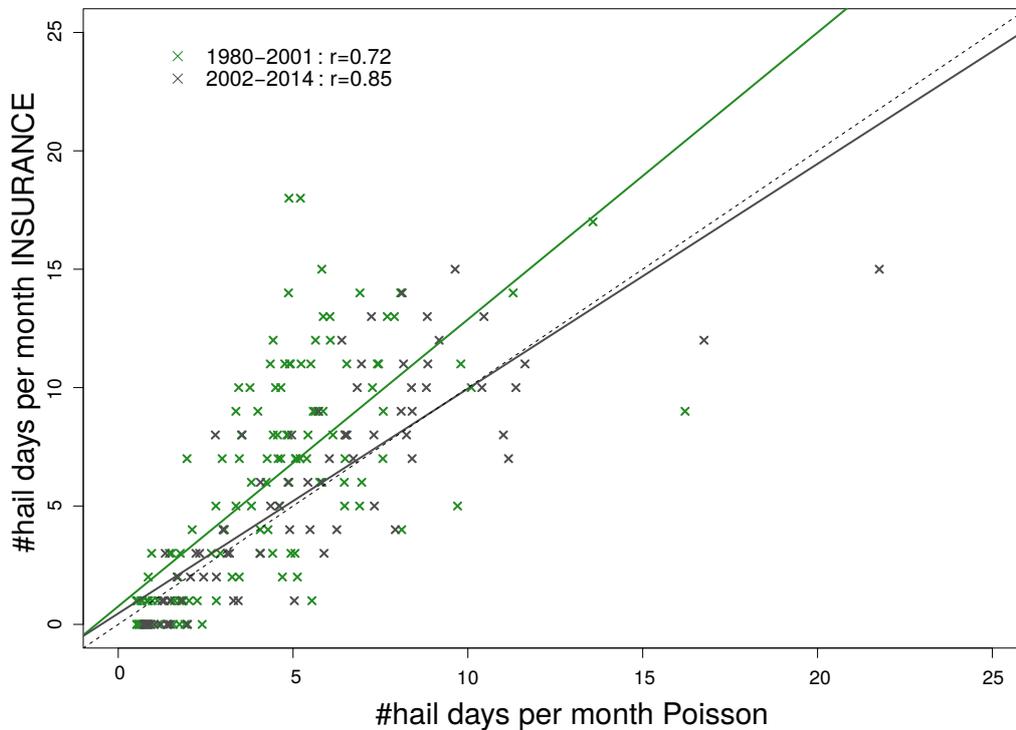


Figure 8: Number of hail days from the Poisson model (x-axis) and from the Swiss hail Insurance company (y-axis) for the 1980-2001 (green) and 2002-2014 (black) period. The solid lines show the regression lines with R^2 added in the legend, the dashed line is the 1:1 line.

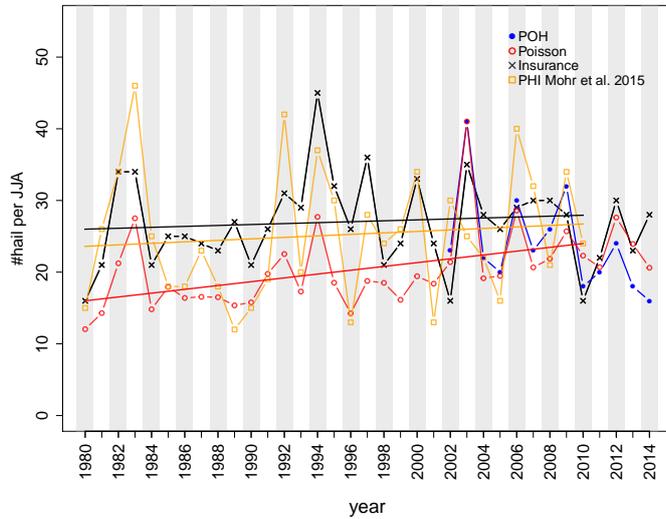


Figure 9: Number of hail days per summer (June, July, August) for the northern part of Switzerland from different sources: from the model study of Mohr et al. (2015a) (orange squares, 1980-2010), from the Swiss Hail Insurance (black crosses, 1980-2014), Poisson model (red circles, 1980-2014) and from POH (blue dots, 2002-2014). Trend lines are calculated only for the period 1980-2010.

For the period 2002-2014, the correlation (r) between the monthly number of hail days from the insurance data and the POH data is slightly weaker ($r = 0.85$) than the correlation between the results from the Poisson model and POH ($r = 0.91$).

Figure 8 shows the scatter plot between the insurance data and the Poisson model results for the period 1980-2001 (green) and for the period 2002-2014 (black). The agreement between the two data sets is better for the later period ($r = 0.85$), compared to the earlier period ($r = 0.72$). In 2002-2014, the number of days with hail per month has similar magnitude for both insurance and Poisson model (black regression line close to the dashed 1:1 line), whereas the insurance data indicates more hail days in the past (1980-2001) compared to the Poisson model (green regression line). As a result the positive trend in the number of hail days for the last 35 years in the insurance data is weaker (increase of about 0.15 vs. 0.5 days with hail per month per decade) and not significant.

Second, we compare our estimates and the insurance data with the Potential Hail Index (PHI) of Mohr et al. (2015a). This index estimates daily hail probabilities based on three large-scale predictors (SLI and T2m at 12 UTC and the minimum near-surface temperature in the morning). Their statistical model is applied to regional retrospective analysis based on NECP/NCAR1 reanalysis at a grid point located approximately in the center of our box (47.0°N, 7.5°E). The number of days with hail is available accumulated over the summer months (June, July and August, JJA) from 1951 to 2010.

Figure 9 shows the summer (JJA) accumulated number of hail days obtained with PHI, our Poisson regression, POH data and the insurance data. The correlation of the Poisson model (red circles) with PHI (orange squares) as well as with the insurance data (black crosses) is approximately $r \approx 0.5$. A similar value ($r = 0.51$) is found between PHI and insurance data. These values are substantially influenced and lowered by the extreme summer of 2003. Moreover, the absolute numbers are generally lower for the Poisson model compared to the two other time series, but the absolute values are also not directly comparable (see section 2.2). These discrepancies result partly from the threshold-based, subjective definition of hail days in all three time series.

The Poisson regression model data shows a significant increase in the number of hail days in the last 35-years (considering the monthly as well as only the JJA time series). The insurance data and PHI (only for summer) show positive but weaker and non-significant trends. Keeping in mind that the trend depends on the observed period, we explain the increase in the number of hail days with positive trends in the temperature and MLCAPE. The influence of temperature on hail was also postulated in other studies. For example Berthet et al. (2011) found for France a correlation between hail intensity and mean minimum surface temperature. They observed an increase in the amount

of hail during spring and related it to the concomitant increase in temperatures.

Let us point out that insurance data are known to have limitations when it comes to study trends, as they depend on the vulnerability of the crops or can be affected by prevention measures taken over time.

5. Discussion

5.1. Nonstationarities

The number of hail days during the last years is overestimated by the Poisson regression model (negative residuals in Fig. 3b). In years 2011/2012 the Swiss radar network was renewed. All the hail data have been homogenized (Nisi et al., 2016). However, the data of the newest radar generation contains less clutter and while the clutter was filtered out for the earlier generation it might still have affected the data slightly. However, one radar had a small negative bias in the number of hail events prior to 2012 and this should result in the opposite signal.

The regression model has its limitations as well. The trend in the model is driven mainly by the trend in T2m and MLCAPE (see section 4.2). However, this strong link between severe convection and the temperature should not be extrapolated into the future as it might break down if the link between moisture availability and temperature changes in a warmer climate, e.g., if the soils dry out in the summer. Moreover, to estimate future change in hail frequency, it is important to consider the interplay of each factor on a daily basis (see e.g., discussion in Diffenbaugh et al., 2013). The change in the hail size distribution, which can depend on additional parameters (e.g., the melting layer height, Dessens et al., 2015), is another interesting question, reserved for future work.

5.2. Comparison with other regression models

We want to briefly compare our model with a similar regression model developed for the U.S. by Allen et al. (2015) based on the North American Regional Reanalysis (Mesinger et al., 2006). Allen et al. (2015) used a Poisson model and monthly means (not anomalies) and identified four significant environmental parameters for the U.S.: the logarithm of the MLCAPE, the logarithm of the storm relative helicity, the logarithm of convective precipitation and the mean specific humidity in the lowest atmosphere. They did not find any direct influence of temperature on hail. The difference in the regression parameters could emphasize regional differences in the type of hail producing storms and their mechanisms of formation. The model for the U.S. underestimated the hail occurrence in summer and overestimated it during fall, similar to the underestimation of the intra-annual variability that we observe in the present study.

Several hail studies have been performed for different parts of Europe using daily data and a logistic regression approach (e.g., López et al., 2014; Gascón et al., 2015; Mohr et al., 2015a,b). However, due to the different temporal resolution of the studies (daily data instead of monthly means), it is difficult to make a direct comparison.

5.3. Wind shear

Wind shear influences the lifetime and structure of a storm. Large shear values are required to form supercells, which are more likely to produce hail than single cells. Many studies (e.g., Brooks et al., 2003; Kaltenböck et al., 2009; Johnson and Sugden, 2014; Púčik et al., 2015; Taszarek et al., 2017) pointed out the importance of wind shear for the development of hail, but most of them also show that hail is found for a wide range of shear values. Moreover, many of these studies consider only larger ($> 2\text{cm}$ or $> 5\text{cm}$) hailstones.

Wind shear is a predictor of our regression model (eq. 3), but it is not statistically significant (p-value of $\beta_3 = 0.150$, Tab. 3). Indeed, wind shear influences mainly the storm type (Markowski and Richardson, 2010) and the size of hailstones (e.g., Dennis and Kumjian, 2017). Furthermore, Groenemeijer and van Delden (2007) found that larger hail events (hail diameter $> 3\text{cm}$) in the Netherlands occurred in lower wind shear environments (below 15m/s) than hail with smaller diameter. They concluded that in this region hail is most often linked to multicell storms rather than supercells. In Switzerland, a substantial fraction of hail storms consist of ordinary storms and intermediate type storms (not super-cells in all aspects, Houze Jr et al., 1993). In Central Europe, in presence of complex topography, strong wind shear seems not to be necessary for the development of hail (see Punge and Kunz, 2016).

With our approach, we consider monthly mean anomalies of wind shear over a complex terrain and we do not distinguish between hail of different size. We have also seen in the example case (Section 4.3, Fig. 6) that stronger monthly mean wind shear is not always related to more hail days. The large-scale flow (e.g., the position of the jet)

can influence the monthly values of wind shear. These could be some reasons why wind shear is included in the regression model without being statistically significant.

Finally, several studies (e.g., Brooks et al., 2003; Craven and Brooks, 2004; Kaltenböck et al., 2009; Eccel et al., 2012; Púčík et al., 2015; Tuovinen et al., 2015) examined the ability of the product of MLCAPE and vertical wind shear to discriminate between severe and non-severe weather, coming to different conclusion depending on the studied region, the data set (temporal and spatial resolution) and the approach (for example if trying to distinguish between large hail or just hail). We tested the combination of the two parameters on daily basis. We found that for the northern part of Switzerland the product is not a better predictor than its single components.

5.4. Further considerations

The main purpose of this study is to offer an alternative framework to study intra-annual and inter-annual variability of hail in absence of long-term direct hail observations in the northern part of Switzerland. It has not the purpose to capture or forecast every single hail event, for which high temporal (beyond the day) resolution data is needed. Moreover, one missing part in the model is the representation of the convection triggering mechanism, which, for example, can be influenced by topography or the passage of frontal systems. The representation of these initiation processes is still challenging for forecasting thunderstorms and thus hailstorms (e.g., Lock and Houston, 2014), and it is even more complex to represent them in a monthly mean and large-scale framework, as they are often mesoscale processes (e.g., Doswell III, 1987).

6. Conclusion

Statistical models which relate the monthly number of hail days in Northern Switzerland to monthly averaged, large-scale environmental variables have been investigated. Model development steps included both the selection of relevant hail predictors and of the most appropriate regression settings. As training data for the model we used radar-derived hail estimates between April and September from 2002 to 2014. The monthly averaged large-scale variables stem from the ERA-Interim reanalysis. Many large-scale variables have a strong seasonal cycle and are therefore highly correlated. To eliminate this correlation the seasonality is explicitly modeled and anomalies from the monthly climatologies are used as predictors.

The appropriateness of the Poisson model has been assessed and compared to other models namely models that include parts to model zero-counts explicitly, as well as a simple linear model. The chosen model is a Poisson model that estimates the number of hail days per month using four predictors: the two meter temperature anomaly ($T2m_a$), the anomaly of the logarithm of the mixed layer convective available potential energy ($\ln(MLCAPE)_a$), the wind shear anomaly (between 10 m-500 hPa) (SH_a) and the *month*. For the years 2002-2014, the model captures the seasonal cycle of hail well but slightly underestimates the year to year variability.

The regression model has been applied to reanalysis data and a time series of the number of hail days in Northern Switzerland has been created (1980-2001), extending the observational 2002-2014 time series. The results show a significant increase in the number of hail days by about 0.5 hail days per month per decade. The positive trend is strongest for the month of June. Weaker and not significant trends have been also found for summer (JJA) in Northern Switzerland using insurance data and results stem from a different study.

The large-scale variables have been analyzed for two Junes, one with few and one with many hail events. The large-scale variables have been compared and related to their climatological values. The results suggest that, despite the coarse data resolution, the large-scale monthly means capture major differences in the Alpine region, which can be related to the large-scale flow pattern over the Atlantic sector.

Acknowledgements

We would like to thank H. Schiesser and H. Lusti for making the insurance data available and L. Nisi for the POH data. We acknowledge S. Mohr for providing the PHI index and are grateful for helpful discussions and feedbacks. MeteoSwiss and the ECMWF are acknowledged for providing access to the ERA-Interim data set. We thank two anonymous reviewers for their useful suggestions and comments that improved the clarity of the paper.

Appendix A. Regression metrics

Given n observations y_i and corresponding predictions \hat{y}_i ($1 \leq i \leq n$) from a regression model (be it linear or other), the Mean Squared Error (MSE) and Mean Absolute Error (MAE) are defined as

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (\text{A.1})$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (\text{A.2})$$

Furthermore, in the linear regression case where p predictors are considered, we denote by R_i^2 the coefficient of determination obtained when regressing the i^{th} predictor on the others. In this context, the Variance Inflation Factor (VIF, Mansfield and Helms, 1982) associated with this predictor is defined by

$$VIF_i = 1/(1 - R_i^2) \quad (\text{A.3})$$

Coming now specifically to the Poisson regression model, the likelihood (given observations y_i corresponding to p -dimensional instances \mathbf{X}_i of the predictor variables X_1, \dots, X_p) is a function of the parameter vector $\boldsymbol{\beta}$ as follows:

$$\mathcal{L}(\boldsymbol{\beta}) = \sum_{i=1}^n \exp(y_i \mathbf{X}_i' \boldsymbol{\beta} - \exp(\mathbf{X}_i' \boldsymbol{\beta})) / y_i! \quad (\text{A.4})$$

From there, it follows that the log-likelihood can be expanded as

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^n y_i \mathbf{X}_i' \boldsymbol{\beta} - \exp(\mathbf{X}_i' \boldsymbol{\beta}) - \log(y_i!), \quad (\text{A.5})$$

where the last term is often neglected when it comes to optimizing ℓ . By the way the maximum of ℓ , say $\ell^* = \ell(\boldsymbol{\beta}^*)$ where ℓ is assumed to admit a global optimum at some $\boldsymbol{\beta}^*$, can be used to quantify the adequacy of a model. For this purpose, it can be compared to another maximum likelihood value obtained by doing less constraining assumptions about the link between the intensity μ and the predictors. Defining instead a vector $\boldsymbol{\lambda}$ of n intensity values, a *saturated* (a.k.a. *full* or *maximal*) model can be defined as a Poisson model with intensity vector $\boldsymbol{\lambda}$, and the associated likelihood is denoted

$$\ell_S(\boldsymbol{\lambda}) = \sum_{i=1}^n y_i \log(\lambda_i) - \lambda_i - \log(y_i!). \quad (\text{A.6})$$

It turns out that the latter is maximized in the current Poisson case for $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*$ with $\lambda_i^* = y_i$ ($1 \leq i \leq n$, with usual convention $0 \log(0) = 0$). Following a general statistical approach, the deviance is defined as twice the difference between the maximum log-likelihood of the saturated model and the maximum log-likelihood of the considered model:

$$D = 2(\ell_S(\boldsymbol{\lambda}^*) - \ell(\boldsymbol{\beta}^*)), \quad (\text{A.7})$$

and in the particular case of the Poisson regression, the deviance can then be calculated in closed form as

$$D = 2 \left(\sum_{i=1}^n y_i (\log(y_i) - \mathbf{X}_i' \boldsymbol{\beta}^*) - \sum_{i=1}^n (y_i - \exp(\mathbf{X}_i' \boldsymbol{\beta}^*)) \right). \quad (\text{A.8})$$

Finally, assuming again for simplicity of exposition a model (stemming from Poisson regression or other) which likelihood admits a global optimum at some $\boldsymbol{\beta}^*$, the Akaike Information Criterion (AIC) is defined as minus twice the maximal log-likelihood penalized by the number of parameters p :

$$AIC = -2\ell(\boldsymbol{\beta}^*) + 2p. \quad (\text{A.9})$$

Appendix B. R-packages

Most of the presented statistical analyses have been performed with the open-source software package R (R Development Core Team, 2016). In particular, the following packages have been used: *car* (Fox and Weisberg, 2011) for the VIF, *MASS* (Venables and Ripley, 2002) for the bidirectional *stepAIC* function, *pscl* (Zeileis et al., 2008) for the hurdle model and *glmulti* (Calcagno and De Mazancourt, 2010) for the multi model selection.

References

- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control* 19 (716–723).
- Allen, J. T., Karoly, D. J., Mills, G. A., 2011. A severe thunderstorm climatology for Australia and associated thunderstorm environments. *Aust. Meteorol. Oceanogr.* J. 61, 143–158.
- Allen, J. T., Tippett, M. K., Sobel, A. H., 2015. An empirical model relating U.S. monthly hail occurrence to large-scale meteorological environment. *J. Adv. Model. Earth Syst.* 7, 226–243.
- BAFU, 2016. Umgang mit Naturgefahren in der Schweiz - Bericht des Bundesrats in Erfuellung des Postulats 12.4271 Darbellay vom 14.12.2012. Tech. rep., Bundesamt fuer Umwelt (BAFU).
- Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., Reinhardt, T., 2011. Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Mon. Weather Rev.* 139, 3887–3905.
- Berthet, C., Dessens, J., Sanchez, J., 2011. Regional and yearly variations of hail frequency and intensity in France. *Atmos. Res.* 100, 391 – 400.
- Brooks, H. E., Anderson, A. R., Riemann, K., Ebberts, I., Flachs, H., 2007. Climatological aspects of convective parameters from the NCAR/NCEP reanalysis. *Atmos. Res.* 83, 294–305.
- Brooks, H. E., Lee, J. W., Craven, J. P., 2003. The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.* 67–68, 73–94.
- Cacciamani, C., Battaglia, F., Patrino, P., Pomi, L., Selvini, A., Tibaldi, S., 1995. A climatological study of thunderstorm activity in the Po valley. *Theor. Appl. Climatol.* 50, 185–203.
- Calcagno, V., De Mazancourt, C., 2010. Glmulti: An R package for easy automated model selection with (generalized) linear models. *J. Stat. Softw.* 34 (12), 1–29.
- Cecil, D. J., Blankenship, C. B., 2012. Toward a global climatology of severe hailstorms as estimated by satellite passive microwave imagers. *J. Clim.* 25, 687–703.
- Cintineo, J. L., Smith, T. M., Lakshmanan, T. M., Brooks, H. E., Ortega, K. L., 2012. An objective high-resolution hail climatology of the contiguous United States. *Weather Forecast.* 27, 1235–1248.
- Costa, S., Mezzasalma, P., Levizzani, V., Alberoni, P., Nanni, S., 2001. Deep convection over northern Italy: synoptic and thermodynamic analysis. *Atmos. Res.* 56, 73–88.
- Cragg, J. G., 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica* 39, 829–844.
- Craven, J. P., Brooks, H., 2004. Baseline climatology of sounding derived parameters associated with deep moist convection. *Natl. Weather Dig.* 28, 13–24.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* 137, 553–597.
- Dennis, E. J., Kumjian, M. R., 2017. The impact of vertical wind shear on hail growth in simulated supercells. *J. Atmos. Sci.* 74, 641–663.
- Dessens, J., Berthet, C., Sánchez, J. L., 2015. Change in hailstone distributions with an increase in the melting level height. *Atmos. Res.* 158–159, 245–253.
- Diffenbaugh, N., Scherer, M., Trapp, R. J., 2013. Robust increases in severe thunderstorm environments in response to greenhouse forcing. *PNAS* 110, 16361–16366.
- Doswell III, C. A., 1987. The distinction between large-scale and mesoscale contribution to severe convection: A case study example. *Weather Forecast.* 2, 3–16.
- Eccel, E., Cau, P., Riemann-Campe, K., Biasioli, F., 2012. Quantitative hail monitoring in an alpine area: 35-year climatology and links with atmospheric variables. *Int. J. Climatol.* 32, 503–517.
- Foot, B. G., Krauss, T., Makitov, V., 10-13 January 2005. Hail metrics using convectional radar. In: *In Proceedings of 16th Conference on Planned and Inadvertent Weather Modification*. Vol. San Diego, California: 1-6. Am. Meteorol. Soc. : Boston, MA.
- Fox, J., Weisberg, S., 2011. *An R Companion to Applied Regression*, 2nd Edition. Sage.
- Galway, J., 1956. The lifted index as a predictor of latent instability. *Bull. Am. Meteorol. Soc.* 37, 528–529.
- García-Ortega, E., Merino, A., López, L., Sánchez, J. L., 2012. Role of mesoscale factors at the onset of deep convection on hailstorm days and their relation to the synoptic patterns. *Atmos. Res.* 114–115, 91–106.
- Gascón, E., Merino, A., Sánchez, J. L., Fernández-González, S., García-Ortega, E., López, L., Hermida, L., 2015. Spatial distribution of thermodynamic conditions of severe storms in southwestern Europe. *Atmos. Res.* 164–165, 194–209.
- Giaiotti, D., Nordio, S., Stel, F., 2003. The climatology of hail in the plain of Friuli Venezia Giulia. *Atmos. Res.* 6768, 247–259, European Conference on Severe Storms 2002.
- Greene, D., Clark, R., 1972. Vertically integrated liquid water – a new analysis tool. *Mon. Weather Rev.* 100, 548–552.

- Groenemeijer, P., van Delden, A., 2007. Sounding-derived parameters associated with large hail and tornadoes in the Netherlands. *Atmos. Res.* 83, 473–487.
- Hermida, L., López, L., Merino, A., Berthet, C., García-Ortega, E., Sánchez, J. L., Dessens, J., 2015. Hailfall in southwest France: Relationship with precipitation, trends and wavelet analysis. *Atmos. Res.* 156, 174–188.
- Hilbe, J. M., Robinson, A. P., 2013. *Methods of Statistical Model Estimation*. No. 255pp. Chapman and Hall/CRC.
- Houze Jr, R., Schmid, W., Fovell, R., Schiesser, H.-H., 1993. Hailstorms in Switzerland: Left movers, right movers, and false hooks. *Mon. Weather Rev.* 121, 3345–3370.
- Huntreser, H., Schiesser, H. H., Schmid, W., Waldvogel, A., 1997. Comparison of traditional and newly developed thunderstorm indices for Switzerland. *Weather Forecast.* 12, 180–125.
- Johns, R. H., Doswell III, C. A., 1992. Severe local storms forecasting. *Weather Forecast.* 7, 588–612.
- Johnson, A. W., Sugden, K. E., 2014. Evaluation of sounding-derived thermodynamic and wind-related parameters associated with large hail events. *Electron. J. Severe Storms Meteorol.* 9, 1–42.
- Kaltenböck, R., Diendorfer, G., Dotzek, N., 2009. Evaluation of thunderstorm indices from ECWMF analyses, lightning data and severe storm reports. *Atmos. Res.* 93, 381–396.
- Kunz, M., 2007. The skill of convective parameters and indices to predict isolated and severe thunderstorms. *Nat. Hazards Earth Syst. Sci.* 7, 327–342.
- Kunz, M., Puskeiler, M., 2010. High-resolution assessment of the hail hazard over complex terrain from radar and insurance data. *Meteorol. Z.* 13, 427–439.
- Lock, N. A., Houston, A. L., 2014. Empirical examination of the factors regulating thunderstorm initiation. *Mon. Weather Rev.* 142 (1), 240–258.
- López, L., García-Ortega, E., Sánchez, J. L., 2014. A short-term forecast model for hail. *Atmos. Res.* 83, 176–184.
- Luterbacher, J., I. Dietrich, D., Xoplaki, E., Grosjean, M., Wanner, H., 2004. European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* 303, 1499–1503.
- Mansfield, E. R., Helms, B. P., 1982. Detecting multicollinearity. *Am. Stat.* 36, 158–160.
- Manzato, A., 2012. Hail in northeast Italy: Climatology and bivariate analysis with the sounding-derived indices. *J. Appl. Meteorol. Climatol.* 51, 449–467.
- Markowski, P., Richardson, Y., 2010. *Mesoscale Meteorology in Midlatitudes*. No. pp 430. John Wiley & Sons, Chichester, UK.
- Martius, O., Kunz, M., Nisi, L., Hering, A., 2015. Conference report 1st european hail workshop. *Meteorol. Z.* 24 (4), 441–442.
- McDonnell, K. A., Holbrook, N. J., 2004. A Poisson regression model approach to predicting tropical cyclogenesis in the Australian/southwest Pacific Ocean region using the SOI and saturated equivalent potential temperature gradient as predictors. *Geophys. Res. Lett.* 31 (20), 120110.
- Melcón, P., Merino, A., Sánchez, J. L., López, L., García-Ortega, E., 2017. Spatial patterns of thermodynamic conditions of hailstorms in southwestern France. *Atmos. Res.* 189, 111–126.
- Melcón, P., Merino, A., Sánchez, J. L., López, L., Hermida, L., 2016. Satellite remote sensing of hailstorms in France. *Atmos. Res.* 182, 221–231.
- Merino, A., López, L., Sánchez, J. L., García-Ortega, E., Cattani, E., Levizzani, V., 2014a. Daytime identification of summer hailstorm cells from MSG data. *Nat. Hazards Earth Syst. Sci.* 14, 1017–1033.
- Merino, A., Wu, X., Gascón, E., García-Ortega, E., Dessens, J., 2014b. Hailstorms in southwestern France: Incidence and atmospheric characterization. *Atmos. Res.* 140–141, 61–75.
- Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., Shi, W., 2006. North American Regional Reanalysis. *Bull. Am. Meteorol. Soc.* 87, 343–360.
- Miller, R., 1972. Notes on analysis and severe storm forecasting procedures of the air force global weather central. *Air Weather Service Tech. Rept.* 200 (Rev.): Air Weather Service, AWS/XTX, Scott Air Force Base, IL., 62225–5118Pp190.
- Mohr, S., Kunz, M., 2013. Recent trends and variabilities of convective parameters relevant for hail events in Germany and Europe. *Atmos. Res.* 123, 211–228.
- Mohr, S., Kunz, M., Geyer, B., 2015a. Hail potential in Europe based on a regional climate model hindcast. *Geophys. Res. Lett.* 42, 904–912.
- Mohr, S., Kunz, M., Keuler, K., 2015b. Development and application of a logistic model to estimate the past and future hail potential in Germany. *J. Geophys. Res. Atmos.*
- Moncrieff, M., Miller, M., 1976. The dynamics and simulation of tropical cumulonimbus and squall lines. *Q. J. R. Meteorol. Soc.* 102, 373–394.
- Mullahy, J., 1986. Specification and testing of some modified count data models. *J. Econom.* 33, 341–365.
- Nisi, L., Martius, O., Hering, A., Kunz, M., Germann, U., 2016. Spatial and temporal distribution of hailstorms in the Alpine region: a long-term, high resolution, radar-based analysis. *Q. J. R. Meteorol. Soc.* 142, 1590–1604.
- O'Brien, R. M., 2007. A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity* 41, 673–690.
- Punge, H. J., Kunz, M., 2016. Hail observations and hailstorm characteristics in Europe: A review. *Atmos. Res.* 176, 159–184.
- Púčik, T., Groenemeijer, P., Rýva, D., Kolář, M., 2015. Proximity soundings of severe and nonsevere thunderstorms in central Europe. *Mon. Weather Rev.* 143, 4805–4821.
- R Development Core Team, 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
URL <http://www.R-project.org/>
- Rasmusen, E. N., 2003. Refined supercell and tornado forecast parameters. *Weather Forecast.* 18, 530–535.
- Sánchez, J. L., Merino, A., Melcón, P., García-Ortega, E., Fernández-González, S., Berthet, C., Dessens, J., 2017. Are meteorological conditions favoring hail precipitation change in Southern Europe? Analysis of the period 1948–2015. *Atmos. Res.* 198, 1–10.
- Schemm, S., Nisi, L., Martinov, A., Leuenberger, D., Martius, O., 2016. On the link between cold fronts and hail in Switzerland. *Atmos. Sci. Lett.* 17, 315–325.
- Schiesser, H., 1990. Hailfall: the relationship between radar measurements and crop damage. *Atmos. Res.* 25, 559 – 582.
- Schiesser, H.-H., Waldvogel, A., Schmid, W., Willemse, S., 1997. *Klimatologie der Stürme und Sturmsysteme anhand von Radar- und Schaden-daten*. No. 132pp. Vdf, Hochschulverlag AG an der ETH Zürich.

- Showalter, A., 1953. A stability index for thunderstorm forecasting. *Bull. Am. Meteorol. Soc.* 34, 250–252.
- Taszarek, M., Brooks, H. E., Czernecki, B., 2017. Sounding-derived parameters associated with convective hazards in Europe. *Mon. Weather Rev.* 145, 1511–1528.
- Tippett, M. K., Camargo, S. J., Sobel, A. H., 2011. A poisson regression index for tropical cyclone genesis and the role of large-scale vorticity in genesis. *J. Clim.* 24, 2335–2357.
- Tippett, M. K., Sobel, A., Camargo, S. J., 2012. Association of U.S. tornado occurrence with monthly environmental parameters. *Geophys. Res. Lett.* 39, L02801.
- Tippett, M. K., Sobel, A. H., Camargo, S. J., Allen, J. T., 2014. An empirical relation between U.S. tornado activity and monthly environmental parameters. *J. Clim.* 27, 2983–2999.
- Tuovinen, J.-P., Rauhala, J., M., S. D., 2015. Significant-hail-producing storms in Finland: Convective-storm environment and mode. *Weather Forecast.* 30, 1064–1076.
- Venables, W. N., Ripley, B. D., 2002. *Modern Applied Statistics with S*, 4th Edition. Springer, New York.
- Villarini, G., Vecchi, G. A., Smith, J. A., 2010. Modeling the dependence of tropical storm counts in the North Atlantic basin on climate indices. *Mon. Weather Rev.* 138, 2681–2705.
- VKF, 2013. *Naturgefahren und Praevention innerhalb der VKF*. Tech. rep., Vereinigung Kantonaler Feuerversicherungen (VKF).
URL [http://vkf.ch/getmedia/ce1cc5de-f5a6-4f67-b5f8-3fd26e8400fa/Broschure ESP 2013 V0-0.pdf.aspx](http://vkf.ch/getmedia/ce1cc5de-f5a6-4f67-b5f8-3fd26e8400fa/Broschure_ESP_2013_V0-0.pdf.aspx)
- Waldvogel, A., Federer, B., Grimm, P., 1979. Criteria for the detection of hail cells. *J. Appl. Meteorol.* 18, 1521–1525.
- Weisman, M. L., Klemp, J. B., 1982. The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Weather Rev.* 110, 504–520.
- Weisman, M. L., Klemp, J. B., 1984. The structure and classification of numerically simulated convective storms in directionally varying wind shear. *Mon. Weather Rev.* 112, 2479–2498.
- Wilks, D. S., 2011. *Statistical Methods in the Atmospheric Sciences*. No. 676pp. Academic Press.
- Willemsse, S., 1995. A statistical analysis and climatological interpretation of hailstorms in Switzerland. Ph.D. thesis, ETH Zürich.
- Witt, A., Eilts, M. D., Stumpf, G. J., Johnson, J. T., De Wayne Mitchell, E., Thomas, K. W., 1998. An enhanced hail detection algorithm for the WSR-88D. *Weather Forecast.* 13, 286–303.
- Zeileis, A., Kleiber, C., Jackman, S., 2008. Regression models for count data in R. *J. Stat. Softw.* 27 (8), 1–25.