https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences
Discussions

Predicting power outages caused by extratropical storms

Roope Tervo[1,*], Ilona Láng[1,*], Alexander Jung[2], and Antti Mäkelä[1]

[1]Finnish Meteorological Institute, B.O. 503, 00101 Helsinki, Finland
[2]Aalto University, Dept of Computer Science, B.O. 11000, 00076 Aalto, Finland
[*]These authors contributed equally to this work.

Correspondence: Roope Tervo (roope.tervo@fmi.fi)

Abstract. Strong winds induced by extratropical storms cause a large number of power outages especially in highly forested countries such as Finland. Thus, predicting the impact of the storms is one of the key challenges for power grid operators. This article introduces a novel method to predict the storm severity for the power grid employing ERA5 reanalysis data combined with forest inventory. We start by identifying storm objects from wind gust and pressure fields by using contour lines of 15 $\mathrm{m\,s^{-1}}$ and 1000 hPa respectively. The storm objects are then tracked and characterized with features derived from surface weather parameters and forest vegetation information. Finally, objects are classified with a supervised machine learning method based on how much damage to the power grid they are expected to cause. Random Forest Classifier, Support Vector Classifier, Naive Bayes, Gaussian Processes, and Multilayer Perceptron were evaluated for the classification task, Support Vector Classifier providing the best results.

1 Introduction

Strong winds, caused by extratropical storms are among the biggest natural hazards in Europe causing massive damage to the forests and society (i.e. Schelhaas et al. (2003); Schelhaas (2008); Ulbrich et al. (2008); Seidl et al. (2014); Valta et al. (2019)); extratropical storms are responsible for 53 percent of all losses related to natural hazards in Europe (Re, 2013). Such storms pose a huge challenge for power distribution companies in highly-forested countries such as Finland (Gardiner et al., 2010) where falling trees cause power outages for hundreds of thousands of customers every year (Niemelä, 2018). Having over 90 000 kilometers overhead line (70 percent of it medium-voltage, 1-35 kV, network) passing through forest (Kufeoglu and Lehtonen, 2015), the windstorms create significant risk for the power supply in Finland. Between the years 2010 and 2018, on average 46 percent of all transmission faults in Finland were caused by extra-tropical storms (Finnish Energy, 2010-2018). During the years of the most damaging storms, 2011 and 2013, the share of windstorm damages of all fault causes was up to 69 percent (Finnish Energy, 2011, 2013) compared to previous years. The need for managing power interruptions is even more urgent since the power suppliers in Finland are obliged to compensate customers of urban areas after 6 hours and rural areas

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020


Natural Hazards and Earth System Sciences
Discussions

after 36 hours of interruption in electricity distribution (Nurmi et al., 2019), thus they require a large amount of manpower to fix caused damages rapidly.

Based on Gregow et al. (2017), the windstorm damages especially in Northern, Central, and Western Europe have increased during the past three decades significantly. Also, other studies are suggesting an increase in wind-related damages in Europe (Csilléry et al. (2017); Haarsma et al. (2013); Gardiner et al. (2010)). Although Ulbrich et al. (Ulbrich et al., 2009) describe the future of extratropical storms to be complex to foresee, it seems that the total number of storms might decrease (i.e. Donat et al. (2011)) but on specific regions, like western Europe and Northeast Atlantic, the number of extreme storms increases (e.g. Pinto et al. (2013)). Besides, the tracks of extratropical storms have already been shifted and are likely to shift also in the future towards the poles (Hoegh-Guldberg et al., 2018), which might affect the storminess in northern Europe. According to (Barredo, 2010) the increased disaster losses to be caused rather by increasing exposure of society than the increased number of windstorms.

Several previous studies respond to the demand for storm impact estimation for power distribution, many of them focusing on the hurricane-induced power blackouts in Northern America (Eskandarpour and Khodaei (2017); Guikema et al. (2014, 2010); Nateghi et al. (2014); Han et al. (2009); Wang et al. (2017); Allen et al. (2014); Chen and Kezunovic (2016); He et al. (2017); Liu et al. (2018)). Convective thunderstorms have been also investigated thoroughly. Li et al. (2015) introduced an area-based outage prediction method further developed to take power grid topology into account (Singhee and Wang, 2017). Shield et al. (2018) studied outage prediction applying a random forest classifier to weather forecast data in a regular grid. Kankanala et al. used data from ground observation stations and experimented regression (Kankanala et al., 2011), a multilayer perceptron neural network (Kankanala et al., 2012), and ensemble learning (Kankanala et al., 2014) to predict outages caused by wind and thunder. Bayesian outage probability (BOP) prediction model (Yue et al., 2018) combines weather radar data and unifies it to a regular grid. Cintineo et al. (2014) create spatial objects from satellite and weather radar data, and track and classify the objects with Naïve Bayesian classifier. Rossi (2015) developed a method to detect and track convective storms. The method was later developed to predict power outages (Tervo et al., 2019).

While much work exists on damage caused by large-scale storms (hurricanes) and small-scale storms (convective thunderstorms), relatively little has been done to be prepared for outages caused by mid-latitude extratropical storms differing from hurricanes and convective storms in available data, time-span, and applicable methods for detecting and tracking. Related forest damage studies have been conducted, though, with random forest classifiers and neural networks. Hart et al. (2019) showed at random forest regression and artificial neural networks can predict a number of falling trees in France caused by the wind. Hanewinkel (2005) conducted a similar study in Germany using artificial neural networks. Artificial neural networks have been used to predict extreme weather in Finland (Ukkonen et al. (2017), Ukkonen and Mäkelä (2019)). To summarise, according to various sources, for example, the framework of IPCC (Masson-Delmotte et al., 2018), the impacts of the extreme weather risks can be analyzed by estimating the hazard, vulnerability, and exposure while machine learning techniques are becoming more popular in the task of connecting the natural hazards with the societal impact forecasts (Chen et al., 2008).

We present a novel method to identify, track, and classify extratropical storm objects based on how much power outages they are expected to induce. We adapt storm object detection (Rossi (2015), Tervo et al. (2019), Cintineo et al. (2014)) to

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020


Natural Hazards and Earth System Sciences Discussions

find potentially harmful areas from extratropical storms by contouring objects from pressure and wind gust fields. We then train a supervised machine learning model to classify storm objects according to their damage potential. To our knowledge, our method is the first that employs the extratropical storm objects as polygons and combines them with meteorological and non-meteorological features to predict power outages. The method can be used as a decision support tool in power distribution companies or as part of elaborating impact forecast by duty forecasters in national hydro-meteorological centers.

This paper is organized as follows: Chapter 2 presents used data and followed by step-by-step method description in Chapter 3. Chapter 3.1 discusses identifying storm objects and present storm tracking algorithm. Chapter 3.2 discusses features. Chapter 3.3 discusses how to define labels of storm objects based on outage data. Chapter 3.4 describes used machine learning methods. In Chapter 4, we discuss the performance of the method followed by conclusion in Chapter 5.

2 Data

We base our method on three main data sources: ERA5 reanalysis data (Hersbach et al., 2019), multi-source national forest inventory (ms-nfi) provided by The Natural Resources Institute Finland (Luke) and occurred power outages obtained from two sources. First, the *local dataset* is gathered from two power distribution companies, Loiste and Järvi-Suomen Energia (JSE), located in Eastern Finland. Second, the *national dataset* is obtained from Finnish Energy (ET), a branch organization for the industrial and labor market policy of the energy sector. All data is gathered from 2010 to 2018. These data are described in the following.

ERA5 is the newest generation reanalysis data provided by ECMWF. ERA5 covers the years from 1979 onward with a one-hour temporal resolution, has a horizontal resolution of 31 km, and covers the atmosphere using 137 levels up to a height of 80 km (Hersbach et al., 2019). Compared to in-situ wind observations, reanalysis data provides a spatiotemporally wider dataset. However, a question may arise about the accuracy of the reanalysis data. Ramon et al. (2019) examined the wind speed characteristics of a total of five state-of-the-art global reanalyses concerning 77 instrumented towers. In their study, ERA5 had the best agreement with in-situ observations on daily time scales; this suggests the ERA5 wind parameters to be decent in windstorm damage examinations as well.

The multi-source forest inventory data is based on field measurements, satellite observations, digital maps, and other geo-referenced data sources (Mäkisara et al., 2016). The data consists of estimates for the forest age, tree species dominance, the mean and total volume, and the biomass (total and tree species-specific).

Each power distribution company in Finland is monitoring and collecting the power interruption data. The raw data consist of many parameters, including the start and end times of the interruption along with a location of a distribution transformer. This data is collected by the Finnish Energy (Finnish Energy, 2010-2018). The national data can be acquired for research purposes but only for areas containing a minimum of five grid companies; this is, for example, to ensure the anonymity of energy users. The same data can be obtained also from some individual power distribution companies with better spatial accuracy.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020


Natural Hazards and Earth System Sciences Discussions

3 Method

We predict power outages by classifying storm objects identified from gridded weather data into three classes based on a number of power outages the storm can typically cause. The overall process contains the following steps: (1) identifying storm objects from weather fields by finding contour lines of some particular threshold, (2) tracking the storm object movement, (3) gathering features of the storm objects, and (4) classifying the objects.

3.1 Identifying and tracking storm objects

Storm objects are identified by finding contour lines of wind gust fields and pressure fields with several thresholds. We define the storm objects using 15 $\mathrm{m\,s^{-1}}$ thresholds from the ERA5 surface level grid with a time step of 1 hour. Thus one storm object represents a solid area (polygon) where hourly maximum wind gust exceeds 15 $\mathrm{m\,s^{-1}}$ during one particular hour. The threshold of 15 $\mathrm{m\,s^{-1}}$ is selected as different sources indicate Finland being vulnerable for windstorms and rather moderate winds (from 15 $\mathrm{m\,s^{-1}}$) causing damages to forests (Valta et al., 2019; Gardiner et al., 2013). To estimate the windstorm impacts on forests, Valta et al. (2019) developed a method by combining the recorded forest damages from the nine most intense storms and their observed maximum inland wind gusts. According to the formula developed in the study, alone the inland wind gusts of 15 $\mathrm{m\,s^{-1}}$ result in forest damages of 1800 $\mathrm{m^3}$.

After identification, storm objects are connected to preceding objects using Algorithm 1. Each object having pressure objects or preceding objects within the threshold, are assigned to the same storm event and gets the same storm ID. Notably, nearby pressure objects are considered along with previous wind objects in assigning ID. This allows several, potentially distant wind objects around the low-pressure center to be assigned to the same storm event. Single wind objects without nearby pressure object or preceding objects are left without ID as they are not assumed to be part of any storm.

We use a 500 km distance threshold for the distance between wind and pressure objects. As the typical diameter of an extratropical storm is approximately 1000 km (Govorushko, 2011), we assume, the damaging wind objects to situate a maximum 500 km from the center of the low pressure. The thresholds for motion speed for wind objects is 200 $\mathrm{km\,h^{-1}}$ and 45 $\mathrm{km\,h^{-1}}$ for pressure objects. Wind objects are i.e. not assumed to move over 200 km and pressure objects over 45 $\mathrm{km\,h^{-1}}$ (Govorushko, 2011). Convective storms may move faster but are outside the focus of this work.

3.2 Extracting storm object features

We characterize the storm objects identified by the methods discussed in Section 3.1 using the features listed in Table 1. The features are structured as four groups. The first group is object characteristics such as size and movement speed and direction are calculated from the contoured storm objects themselves. As a second group, relevant weather conditions, such as wind speed, temperature, etc., are extracted from ERA5 data. To represent each parameter with one number, we aggregate values from the object coverage using functions listed in Table 1. Third, as most of the outages are caused by the trees falling over power grid lines (Campbell and Lowry, 2012), and thus the features of the forest (i.e. tree height, age, or specie) contribute in the damages (Peltola et al., 1999), we support our data with forest information. As with weather parameters, values are

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences
Discussions

Algorithm 1 Storm tracking

Input
Individual storm objects
$pressure\ distance\ threshold$
$wind\ distance\ threshold$
$speed\ threshold$
Output
Connected storm object with storm ID
for all storm objects **do**
if pressure object with ID exists within $pressure\ distance\ threshold$ **then**
Use pressure object ID
else if previous pressure object with ID exists within $speed\ threshold$ **then**
Use previous pressure object ID
else if other object with ID exists within $wind\ distance\ threshold$ **then**
Use other object ID
else if previous object with ID exists within $speed\ threshold$ **then**
Use previous object ID
else if previous object without ID exists within $speed\ threshold$ **then**
Give new ID to the previous and current object
else
Leave object without ID
end if
end for

aggregated from the storm object coverage. The fourth group consists of the number of outages and affected customers used as labels in the model training process discussed more in Chapter 3.4.

We first gather several parameters and further select the most relevant ones. To this end, we plot the difference in fitted Gaussian distribution between all samples and class one and two samples. While many other distributions are known to suit better in modeling particular parameters (such as Gamma in precipitation, Weibull in wind speed, and Lognormal in cloud properties (Wilks, 2011)), Gaussian distribution is a sufficient simplification to help in selecting relevant parameters. Distribution of some selected parameters is shown in Appendices A1 and A2. In total 35 parameters, shown as bolded in Table 1 were chosen for the final classification.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences
Discussions

Table 1. Extracted features. Features used in the final classification marked as bold.

Feature	Aggregation	Explanation
Speed	-	Object movement speed
Angle	-	Object movement angle
Area	-	Object size
Area difference	-	Object area difference to the previous time step
Week	-	Week of the year
Snowdepth	average, minimum, maximum	Snow depth
Total column water vapor	**average, minimum, maximum**	Total amount of water vapour
Temperature	**average, minimum, maximum**	2 meter air temperature
Snowfall	average, minimum, maximum, sum	Snowfall (meter of water equivalent)
Total cloud cover	average, minimum, maximum	Total cloud cover (0-1)
CAPE	**average**, minimum, **maximum**	Convective available potential energy (J/kg)
Precipitation kg/m2	average, minimum, maximum, sum	Precipitation amount (kg/m2)
Wind gust	**average**, minimum, **maximum, standard deviation**	Hourly maximum wind gust ($\mathrm{m\ s^{-1}}$)
Wind Speed	**average**, minimum, **maximum, standard deviation**	10 meter wind speed ($\mathrm{m\ s^{-1}}$)
Wind Direction	**average**, minimum, maximum, **standard deviation**	Wind direction (degrees))
Dewpoint	**average, minimum, maximum**	Dewpoint)
Mixed layer height	**average, minimum, maximum**	Boundary layer height
Pressure	average, **minimum**, maximum	Air pressure
Forest age	**average**, minimum, maximum, standard deviation	The age of the growing stock on a forest stand
Forest site fertility	**average**, minimum, maximum, standard deviation	Group of the forest by vegetation zones
Forest stand mean diameter	**average**, minimum, maximum, standard deviation	Forest stand mean mean diameter
Forest stand mean height	**average**, minimum, maximum, standard deviation	Forest stand mean height
Forest canopy cover	**average**, minimum, maximum, standard deviation	Forest canopy cover fraction (0-100%)
Outages	-	Number of occured outages
Customers	-	Number of affected customers
Tansformers	-	Number of transformers under the object
All customers	-	Number of customers under the object
Class	-	Assigned class

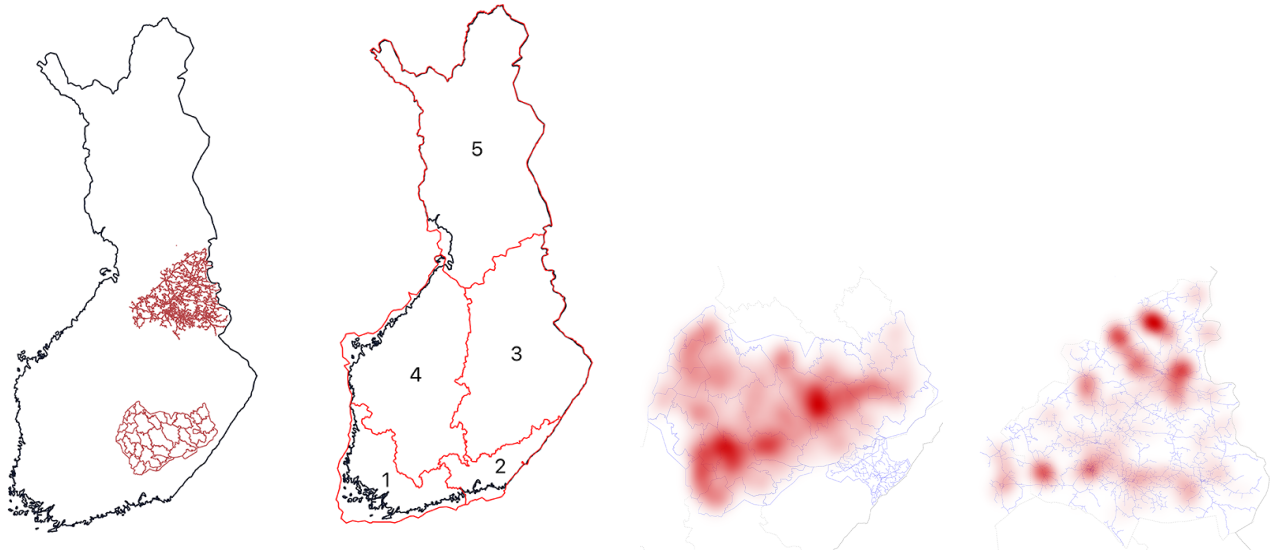
https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences
Discussions

3.3 Defining classes

We use three classes designed together with power grid companies aiming at a simple "at glance" view for power grid operators. Class 0 represents no damage, class 1 low damage, and class 2 high damage. Next, we discuss the power outage data used in this project and then define limits for the classes.

The geographical coverage of the power outage data is illustrated in Figures 1a and 1b. The local dataset contains all outages from Northen Area (Loiste) and outages related to major storms in the Southern area (JSE). The national dataset contains all outages in Finland divided into five regions shown in Figure 1b. As shown in Figures 1c and 1d, the outages are concentrated heavily on 'hot-spots', assumingly, due to forest characteristics and network topology. In total, the local dataset contains 24 542 storm objects and 5 837 outages attributed to a storm object. The national dataset contains 142 873 storm objects and 5 965 324 outages attributed to some storm object.



(a) (b) (c) (d)

Figure 1. (a) Geographical coverage of the outages in local dataset. Red lines represents the power grid where outages are gathered from. (b) Regions in national outage dataset. Outages are gathered from the whole Finland but aggregated to the regions shown in the image. (c) Spatial distribution of the outages in the JSE Network (Southern area), data gathered between 2010 and 2018. (d) Spatial distribution of the outages in the Loiste Network (Northern area), data gathered between 2010 and 2018.

It is notable, that the damage may occur anywhere in the power grid. Outages are, however, always reported as transformers without electricity. Typically one physical damage between the transformers causes several transformers to lose power. Power grid operators can often afterward turn part of the transformers back to operation even before fixing the original damage. This causes an unavoidable noise to the datasets.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences
Discussions

Figure 2 depicts the number of outages and storm objects in both, local and national datasets. We can identify a large amount of 15 $\mathrm{m\ s^{-1}}$ storm objects in both sets indicating that moderate wind objects without other influencing factors, do not cause damage for the transformers. When identifying objects with the contour of 20 and 25 $\mathrm{m\ s^{-1}}$, the number of objects reduces and correlates more with a high number of outages. This supports views of previous studies showing the significance of stronger wind gusts to higher storm damages. The method seems to identify also the most important storm days by capturing several storm objects for those days. For instance, at the end of 2013 when three major storms Eino, Oskari, Seija (Valta et al., 2019) hit Finland, both datasets contain plenty of wind objects with 20 $\mathrm{m\ s^{-1}}$ threshold.

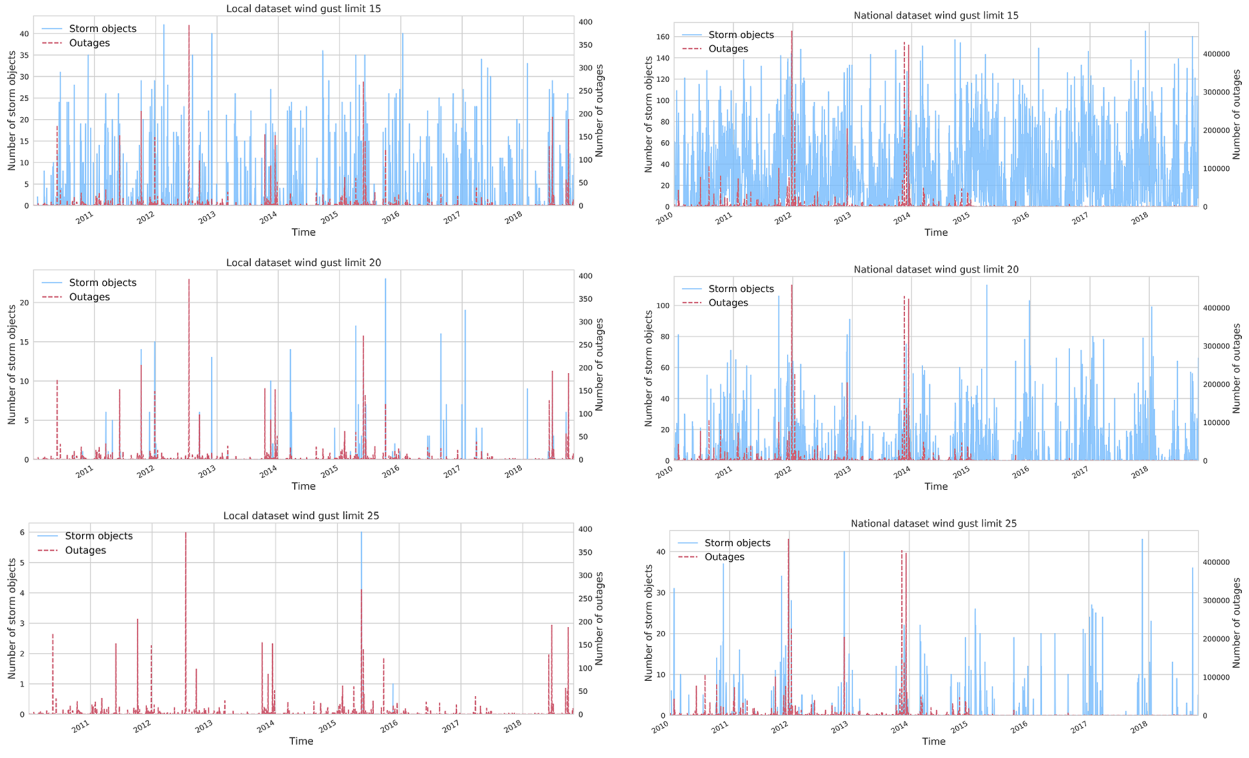


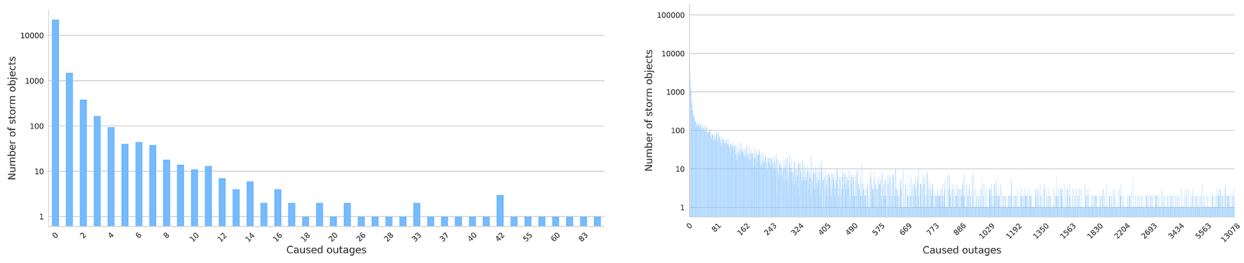
Figure 2. Storm object time series (15, 20 and 25 $\mathrm{m\ s^{-1}}$ contours) with occurred outages for local and national datasets.

Figure 3 illustrates how much outages a single storm object typically produces. In the local dataset, most of the storm objects cause only a few outages. Only 65 storm objects, which are only 0.3 percent of the whole dataset, induced more than 10 outages. On the other hand in the national dataset where one storm object typically affects several different transformers, 17 587 storm objects have caused more than 10 outages which represent 12 percent of the whole dataset. Based on the approximation, shown in Figure 4, one outage typically affects on 200 - 300 customers.

As the number of outages produced by a single storm object varies a lot in the local and national datasets, we end up defining separate limits for the local and the national datasets. The detailed limits are listed in Table 2. Class 1 is defined such that it

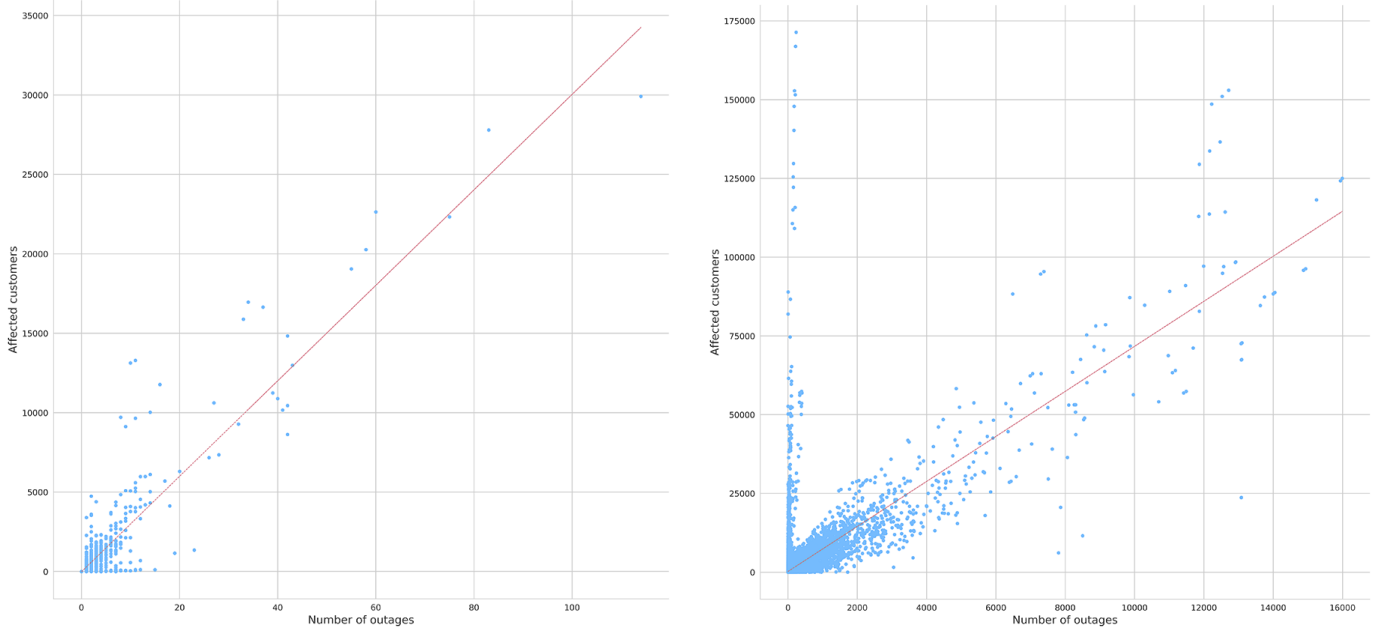
https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU



(a) (b)

Figure 3. Number of storm objects per caused outages in (a) local dataset (b) national dataset.



(a) (b)

Figure 4. Relationship between number of outages and affected customers in (a) local dataset and (b) national dataset.

represents roughly 80 percent of all cases with at least one outage. Class sizes are highly imbalanced as most of the storm objects do not cause any damage.

3.4 Classifying storm objects

We centered and normalized the data points by substracting the empirical mean and then dividing by the empirical standard deviation. The hyperparameters were determined using random search 5-fold cross-validation (Bergstra and Bengio, 2012). To cope with the imbalanced class distribution, we generate artificial training samples using the synthetic minority over-sampling

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020


Natural Hazards and Earth System Sciences
Discussions

Table 2. Classes for local dataset

Class	Outage limit in local tdataset	Local dataset size	Outage limit in national dataset	National dataset size
0	0	5 624	0	76 215
1	1-3	353	1- 140	14 417
2	≥ 4	181	≥ 141	3 085

technique SMOTE (Chawla et al., 2002). Five different models were evaluated to classify storm objects. We omit the more mathematical definitions but shortly discuss the characteristics of different models and describe the implementation details chosen in this work.

Random forest classification (RFC) is based on a random ensemble of decision trees and aggregate results from individual trees to final estimation. Trees in the ensemble are constructed with four steps: 1) use bootstrapping to generate a random sample of the data 2) randomly selected subset of features at each node 3) determine the best split at the node using loss function 4) grow the full tree (Breiman, 2001). RFC is also found to provide adequate performance with imbalanced data (Tervo et al., 2019; Brown and Mues, 2012). We use RFC with the Gini impurity loss function. Hyperparameters listed in Table 3.

Table 3. Hyperparameters for the RFC

Parameter	Value
Number of trees in the forest	500
Max depth	unlimited
Minimum nr. of samples to split	2
Minimun nr of samples to leaf	1
Features to consider for split	$\sqrt{\text{num. of feat.}}$
Max nro of leaf nodes	unlimited

Support Vector Classifiers (SVC) construct a hyper-plane or classification function, in a high dimensional feature space and maximize a distance between training samples and the hyperplane. The hyper-planes may be constructed with non-linear kernels such as gaussian radial basis function (RBF) that often reform a non-linear classification problem to linear. Operating in the high-dimensional feature space without additional computational complexity makes SVC an attractive choice to extract meaningful features from a high dimensional data set. Furthermore, if SVM output is assumed to be log odds of a positive sample, one can fit a parametric model to obtain the posterior probability function and thus get probabilities for samples to belong to the particular class (Platt et al., 1999). For more details, we request the reader to consult for example Chang and Lin (2011) and Platt et al. (1999).

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020


Natural Hazards and Earth System Sciences
Discussions

We implement the SVC in two phases. First, we separate class 0 (no outages) and other samples employing SVC with radial basis function (RBF), defined in Equation 1. Second, we distinguish classes 1 and 2 using SVC with dot-product kernel defined in Equation 2 (Williams and Rasmussen, 2006). The second phase is performed only for the samples predicted to cause outages in the first phase. The approach is similar to often-used one-vs-one classification where a binary classifier is fitted for each pair of classes except that different kernels were used for different pairs.

$$k_{RBF}(\mathbf{x}, \mathbf{x}') = \exp\big(-\gamma||\mathbf{x} - \mathbf{x}'||^2\big) \tag{1}$$

$$k.(\mathbf{x}, \mathbf{x}') = \sigma_0 + x \cdot x' \tag{2}$$

Gaussian Naive Bayes (GNB) (Chan et al., 1979) is a well-known and widely used method based on the Bayesian probability theory. The method assumes that all samples are independent and identically distributed (i.i.d) which does not naturally hold for the weather data. Despite the internal structure of the data, GNB is still sometimes used for weather time series (for example Lindsay and Cox (2005)) and worth investigating also in this context.

Gaussian Processes (GP) (Rasmussen, 2003) is a non-parametric probabilistic method that interprets the observed data points as realizations of a Gaussian random process. GP is widely used for example in weather observation interpolation *kriging* (Holdaway, 1996). They are, however, computationally expensive and they tend to lose power with high-dimensional data. GP models hinge on a kernel function that encodes the covariance between different data points. As a kernel, we use a product of dot-product kernel (Equation 2) and pairwise kernel with laplacian distance, defined in Equation 3. The kernel parameters were optimized on the training data by maximizing the log-marginal-likelihood.

$$k_{pairwise}(\mathbf{x}, \mathbf{x}') = \exp\big(-\gamma||\mathbf{x} - \mathbf{x}'||_1\big) \tag{3}$$

Multilayer perceptrons (MLP) (Goodfellow et al., 2016) are the most basic form of an artificial neural network. Good results achieved by MLP in predicting storms (Ukkonen and Mäkelä, 2019), they are a natural choice to experiment also in this work. The downside of the method is a large number of hyperparameters including the correct network topology. We searched the correct model parameters and network topology for local and national datasets by running multiple iterations of random search 5-fold cross-validation employing Talos library (Autonomio, 2020). Final setup composes of Nadam optimizer (Dozat, 2016), random normal initializer, and relu activation function for hidden layers. Binary cross-entropy was used as a loss function. Optimal network topology varied in different datasets. For the local dataset, the used network contained three hidden layers with 75, 145, and 35 neurons. For the national dataset, the network contained three hidden layers with 75, 195, and 300 neurons.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020


Natural Hazards and Earth System Sciences
Discussions

4 Results

We used two different methods for splitting the data into training and test set. The first method is to use 25 percent of randomly picked samples in the test set. The second method is to construct a test set from a one-year continuous time range (2010-2011). Both approaches have their advantages. Continuous time range ensures that the model has not seen any autocorrelated samples caused by an internal structure of the weather data in the training phase (Roberts et al., 2017). However, having only 9 years of data from a relatively small geographical area, the continuous test set cannot contain many storms as most of the data needs to be reserved for the training process. Thus, the test set may only contain a single type of storms to which the model may work especially well or bad. Picking the test set randomly minimizes this risk and provide more insight to model performance.

We evaluate the models with a weighted average of precision and recall and both weighted and macro average of F1-score. Precision (Equation 4) reports how many samples are correctly predicted to belong to a class. Recall (Equation 5) tells how many samples belonging to a class are found in the prediction. F1-score (Equations 6 and 7) calculates a harmonic mean of precision and recall. Finally, as the datasets are extremely imbalanced we calculate a weighted average of the metrics utilizing a number of samples in each class and a macro average of F1-score using an average of F1-score of each class. A model with a higher macro average of F1-score performs better with small classes.

$$Precision = \frac{1}{\sum_{c \in \mathcal{C}} |\hat{y}_c|} \sum_{c \in \mathcal{C}} \left(|\hat{y}_c| \frac{tp}{tp + fp} \right) \tag{4}$$

where $\mathcal{C}$ represents set of classes, $\hat{y}$ predicted class, tp true positives and fp false positives.

$$Recall = \frac{1}{\sum_{c \in \mathcal{C}} |\hat{y}_c|} \sum_{c \in \mathcal{C}} \left(|\hat{y}_c| \frac{tp}{tp + fn} \right) \tag{5}$$

where $\mathcal{C}$ represents set of classes, $\hat{y}$ predicted class, tp true positives and fn false negatives.

$$F1_{weighted} = \frac{1}{\sum_{c \in \mathcal{C}} |\hat{y}_c|} \sum_{c \in \mathcal{C}} \left(|\hat{y}_c| \frac{Precision_c \times Recall_c}{Precision_c + Recall_c} \right) \tag{6}$$

where $\mathcal{C}$ represents set of classes, $\hat{y}$ predicted class, Precision defined in Equation 4 and Recall defined in Equation 5.

$$F1_{macro} = \frac{1}{|\mathcal{C}|} \sum_{c \in \mathcal{C}} \left(\frac{Precision_c \times Recall_c}{Precision_c + Recall_c} \right) \tag{7}$$

where $\mathcal{C}$ represents set of classes, Precision defined in Equation 4 and Recall defined in Equation 5.

Tables 4 and 5 divulge the results for each models using local and national dataset respectively. Models trained with the local dataset can reach the better-weighted F1-score while the best models trained with the national dataset provide a significantly better macro average of F1-score. The national dataset contains many more samples in classes 1 and 2 which enable models to

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences Discussions

learn the classes better and thus enhance the macro average of the F1-score. Randomly chosen and continuous test set seems not to make a large difference in most cases. The only affected model is the RFC having contradictory better results trained with the continuous test set from the local dataset and the random test set from the national dataset. Assumingly, this squeal more about the unstable performance of RFC than the relevance of the dataset split method.

Confusion matrices are depicted in Figure 5. RFC provides the best results in terms of the selected metrics. Closer exploration revels, however, that this performance is largely due to the best performance in predicting class 0, which is the largest class. SVC results are one of the most balanced ones being the best only in the local dataset with a random test set but yielding good stable results in all cases. The confusion matrix, shown in Figure 5b, displays that it is not the best model to predict class 0 but only a little share of true class 2 cases and the smallest share of true class 1 cases are predicted as class 0. That is to say, SVC misses the smallest number of harmful storms although it confuses in the amount of caused damage.

GP is another strong option that performs an even better job with class 0 while still providing good performance with class 2. A notable connecting aspect between GP and SVC is an almost identical kernel. Based on these experiments, in particular, RBF and pairwise kernels separate harmless and harmful samples from each other while dot-product kernel separates the classes 1 and 2 even better than exponential functions.

Using 15 $\mathrm{m\,s^{-1}}$ threshold for detecting wind objects yields clearly better results than 20 $\mathrm{m\,s^{-1}}$ threshold. For example SVC trained with national dataset using 20 $\mathrm{m\,s^{-1}}$ threshold and randomly chosen test set provide only 0.48 macro average of F1-score being 12 percentage points below corresponding model using 15 $\mathrm{m\,s^{-1}}$ threshold. 15 $\mathrm{m\,s^{-1}}$ threshold have two major advantages compared to 20 $\mathrm{m\,s^{-1}}$. First, it provide significantly larger dataset and second, it is able to catch virtually all extratropical storms causing outages while 20 $\mathrm{m\,s^{-1}}$ can not.

Table 4. Results for each models with local dataset obtained from two local two power grid companies and defined in Chapter 3.3

Model	Split method	Precision	Recall	Weighted F1-score		Macro AVG F1-score	
		test	test	train	test	train	test
Random Forest Classifier (RFC)	Random	0.82	0.76	0.93	0.79	0.93	0.40
	Continuous	0.88	**0.91**	0.93	**0.89**	0.93	**0.48**
Support Vector Classifier (SVC)	Random	0.85	0.73	0.78	0.78	0.78	**0.44**
	Continuous	0.87	0.72	0.77	0.78	0.77	0.42
Gaussian Naive Bayes (GNB)	Random	**0.87**	0.61	0.59	0.70	0.59	0.42
	Continuous	**0.89**	0.59	0.59	0.69	0.59	0.40
Gaussian Processes (GP)	Random	0.84	0.70	1.0	0.76	1.0	0.43
	Continuous	0.85	0.67	0.94	0.74	0.94	0.41
Multilayer perceptor (MLP)	Random	0.82	**0.81**	0.98	**0.80**	0.91	0.41
	Continuous	0.81	0.79	0.97	0.80	0.91	0.41

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences
Discussions

Table 5. Results for each models with national dataset covering whole Finland and defined in Chapter 3.3

Model	test set split method	Precision	Recall	Weighted F1-score		Macro AVG F1-score	
		test	test	train	test	train	test
Random Forest Classifier (RFC)	Random	**0.83**	**0.84**	1.0	**0.83**	1.0	**0.62**
	Continuous	**0.77**	**0.81**	1.0	**0.78**	1.0	0.40
Support Vector Classifier (SVC)	Random	0.81	0.61	0.68	0.68	0.68	0.60
	Continuous	0.62	0.60	0.60	0.60	0.60	0.60
Gaussian Naive Bayes (GNB)	Random	0.75	0.60	0.66	0.66	0.45	0.39
	Continuous	**0.77**	0.60	0.45	0.66	0.45	0.40
Gaussian Processes (GP)	Random	0.57	0.56	0.71	0.55	0.71	0.55
	Continuous	0.67	0.65	0.94	0.65	0.94	**0.61**
Multilayer perceptor (MLP)	Random	0.79	0.75	0.94	0.77	0.90	0.52
	Continuous	0.76	0.78	0.93	0.78	0.85	0.40

4.1 Case Examples

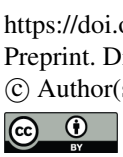
We illustrate the prediction with the three most interesting examples of well-known storms in Figure 6a. We chose the cases among a number of test cases to represent the strengths and weaknesses of the method. The examples are chosen from the randomly picked test set, which was not used to train the model. Because of the random sample, we cannot represent the entire prediction of individual storms, only individually picked time steps. In two of the example cases the model performs well (storms Tapani and Pauliina) and in one (storm Rauli) with less accurate prediction results.

4.1.1 Event 1: Extratropical Storm Tapani (26 December 2011)

The first example is one of the most known extratropical storms in Finland. Storm Tapani, known also as Cyclone Dagmar (Kufeoglu and Lehtonen, 2015), was a rare winter storm, causing wide and long-lasting electricity interruptions. Extreme wind gusts of over 30 $\mathrm{m\ s^{-1}}$ caused widespread damage especially in the southern and western parts of the country. Approximately 570 000 households were left without electricity, causing 30 million euros repair costs and 80 million euros of monetary compensation for electricity distribution companies to their customers. (Hanninen and Naukkarinen, 2012) Exceptionally warm December and the Boxing day being the warmest in 50 years (Finnish Meteorological Institute, 2011) resulted in wet and unfrozen soil, thus, the trees were poorly anchored and exposed to major storm damage.

Figure 6a represents the outage prediction (raster-covered areas) and the actual, true classes (numbers) based on the damage data at 15:00 UTC, 26 December 2011. Wide areas in central and western parts of Finland are predicted to have *high*, class

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020



Natural Hazards and Earth System Sciences
Discussions

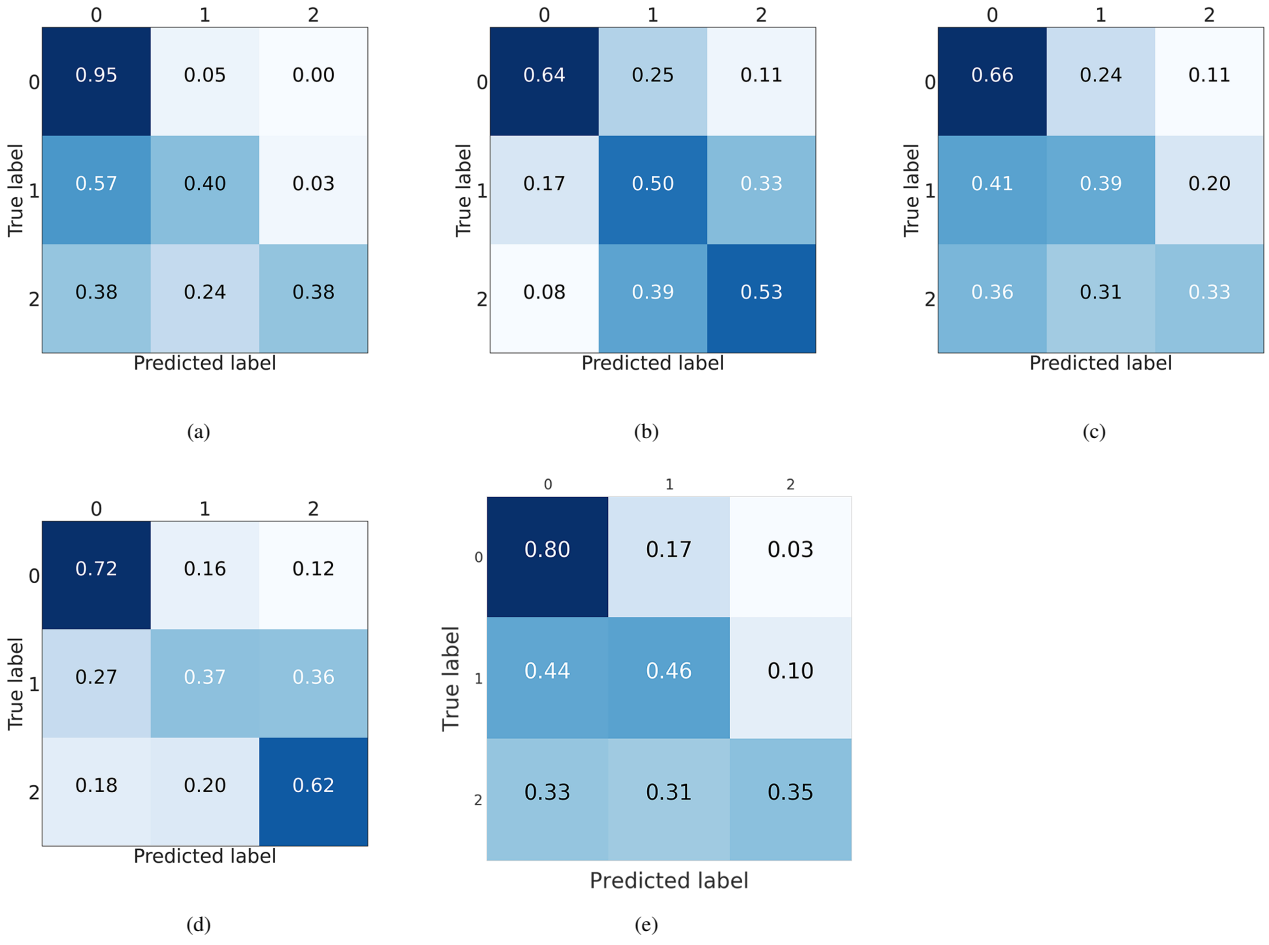


Figure 5. Confusion matrices produced using randomly selected national dataset and (a) RFC (b) SVC (c) GNB (d) GP (e) MLP.

2 damages. The predicted class is in line with the true class. Also, the damage areas of the storm correlate with the wind gust observations of the Finnish Meteorological Institute. The strongest gusts situated in western (15-27 $\mathrm{m\ s^{-1}}$) and southern (18-28 $\mathrm{m\ s^{-1}}$) Finland and north-western part of Lapland (13-31$\mathrm{m\ s^{-1}}$) (Finnish Meteorological Institute, 2020). In the rest of Finland, the maximum wind gusts remained between 10-15 $\mathrm{m\ s^{-1}}$ and therefore the damages were also minor. Overall, the model works in this particular example accurately.

4.1.2 Event 2: Extratropical Storm Rauli (27 August 2016)

Extratropical storm Rauli was exceptionally strong for the summer season, especially regarding the impacts. It caused severe damages for the power grid in the western and middle parts of Finland for various reasons. The trees had leaves on, the soil was wet after a rainy August, the strong wind areas of Rauli were widely spread and the solar radiation was intensifying the wind

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020


Natural Hazards and Earth System Sciences Discussions

gusts during the afternoon (Finnish Meteorological Institute, 2016). Rauli was impacting especially the middle and southern parts of Finland which are also the most densely populated areas. The power outages were increasing rapidly in the middle part of Finland, starting at midday and reaching the highest values, 200 000 households without electricity (Ilta-Sanomat, 2016), around 5 pm. The winds were blowing exceptionally long, nearly 24 hours. The typical duration of summer storms is between 6-12 hours.

Figure 6b shows the predicted outages and true classes at 12:00 UTC, 27 August 2016. In this particular time step, the model is over-predicting the class, however, the predicted outage area seems to correlate with the wind gust maxims during that afternoon. The strongest wind gusts were measured in the southern and middle parts of the country, maximum gusts reaching on land stations up to 24,9 $\mathrm{m\,s^{-1}}$ (Klemettilä, Vaasa and Maaninka, Pohjois-Savo) and on wide areas up to 20 $\mathrm{m\,s^{-1}}$ apart from the northern part of Finland.

4.1.3 Event 3: Extratropical storm Pauliina (22 June 2018)

The last example is a strong extratropical storm, called Pauliina which (Finnish Meteorological Institute, 2018) caused numerous power outages in Finland. The most significant part of the power outages happened in the network of power grid company JSE which is included in the local dataset. The highest peak in the damages was reached between 6 and 8 p.m with over 28 000 households without electricity. The strongest wind gust on land reached 22,7 $\mathrm{m\,s^{-1}}$ on Helsinki, Kumpula observation station and the inland gusts were widely between 15-20 $\mathrm{m\,s^{-1}}$ (Finnish Meteorological Institute, 2020; Finnish Meteorological Institute, Twitter). The strong wind gusts continued until the dawn of the 23rd of June.

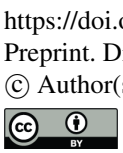
Figure 6c presents the predicted and true damage classes at 01:00, UTC, 22 June 2018. We chose extratropical storm Pauliina as an example storm for two reasons: 1) Pauliina represents a *low damage* class 2) Pauliina represents a rare, summer-season extratropical storm. Figure 6c shows the predicted and true classes correlating. While weather warnings were given to large areas in southern and middle parts of Finland, (Myr) predicted and true damage to the power grid occurred in a relatively small geographical area. This example shows the potential added value of the model compared to weather warnings providing more accurate information to the power grid operators.

5 Conclusions

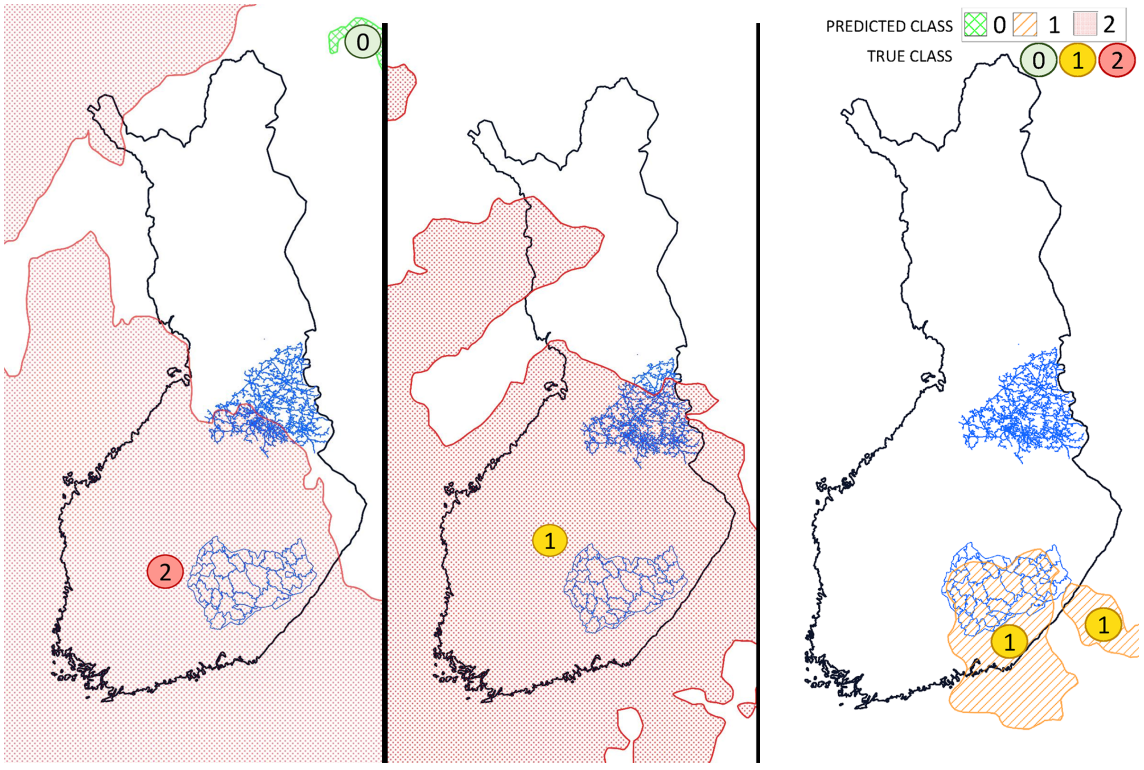
This paper introduced a novel method to predict the damage potential of extratropical storms to power grids. The method consists of identifying wind objects by contouring surface wind gust field with 15 $\mathrm{m\,s^{-1}}$ threshold along with pressure objects with 1000 hPa threshold, tracking the objects, and classifying them in three classes based on their damage potential to the power grid. For the classification task, we evaluated five different machine learning methods all employing in a total of 35 predictive features and trained with 8 years of power outage data from Finland.

The most balanced results were gained with the Support Vector Classifier. The model recognizes harmful storm cells well and can distinguish extremely harmful cells among others adequately. While the results still left a lot to improve, the developed

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020



Natural Hazards and Earth System Sciences
Discussions



(a) Tapani, 26 December 2011 15:00 (b) Rauli, 27 August 2016 12:00 (c) Pauliina, 22 June 2018 01:00

Figure 6. Selected examples (a) Extratropical storm Tapani (b) Extratropical storm Rauli (c) Extratropical storm Pauliina produced employing SVC model trained with national dataset. The storm objects are coloured based on the predicted class while the true class is stated as a coloured number over the object. The time is represented as UTC time.

model can be already used to help decisions made in power grid companies. The model is able to provide a more specific and geospatially accurate prediction of caused damage to the power grid than for example weather warning.

The work opens several possible avenues for further studies. Soil moisture, soil temperature, and leaf index would most probably enhance the results as they would provide critical information about the environmental conditions. Different thresholds could be investigated as well, especially for pressure objects where lower thresholds might yield better results. By design, applying the method on other regions should be possible as well by using available impact and meteorological data. For the classification task, carefully designed Bayesian networks could provide good results as well. Experiments in this study were conducted with ERA5 reanalysis and additional forest data. As the method employs common features existing also in various other datasets, data provided by other vendors could be used as well. In the future, this method could be used as a base for a decision support tool and as a part of an existing early warning system, for both, duty forecasters of national hydro-meteorological centers as well as operators of electricity transmission companies.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences
Discussions

Code and data availability. Source code is available in repositories https://github.com/fmidev/sasse-era5-smartmet-grid and https://github.com/fmidev/sasse-polygon-process. ERA5 data may be downloaded from the Copernicus Climate Data Store: https://cds.climate.copernicus.eu. Forest inventory may be downloaded from LUKE open data service: http://kartta.luke.fi/index-en.html. Power outage data is propriety data which the authors have no property rights to distribute.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences
Discussions

Appendix A: Gaussian distribution fitted to the storm object features

A1 Local dataset

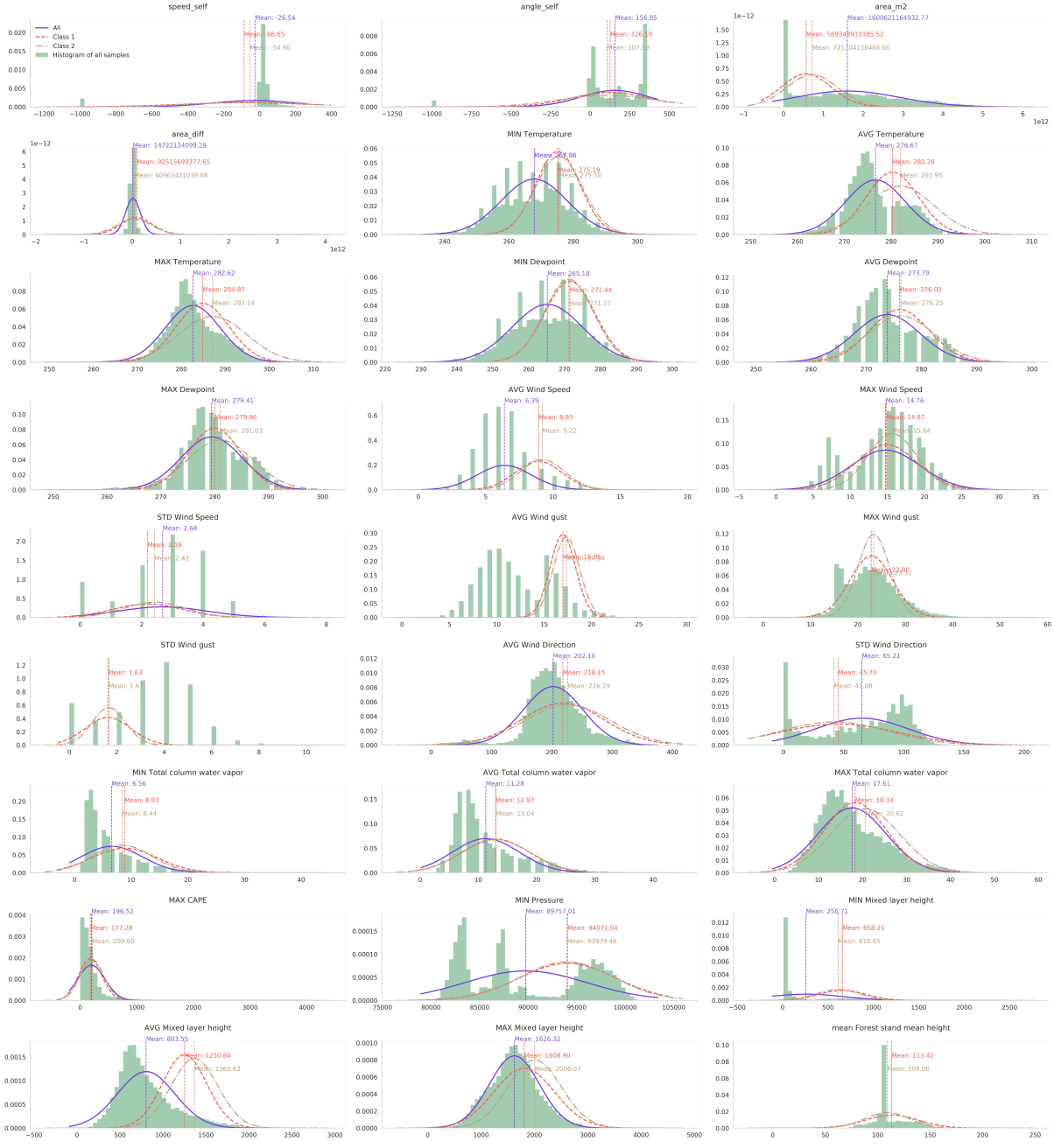


Figure A1. Distribtution of the selected parameters of all storm objects (samples) and objects with class 2. Local dataset.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences
Discussions

A2 National dataset

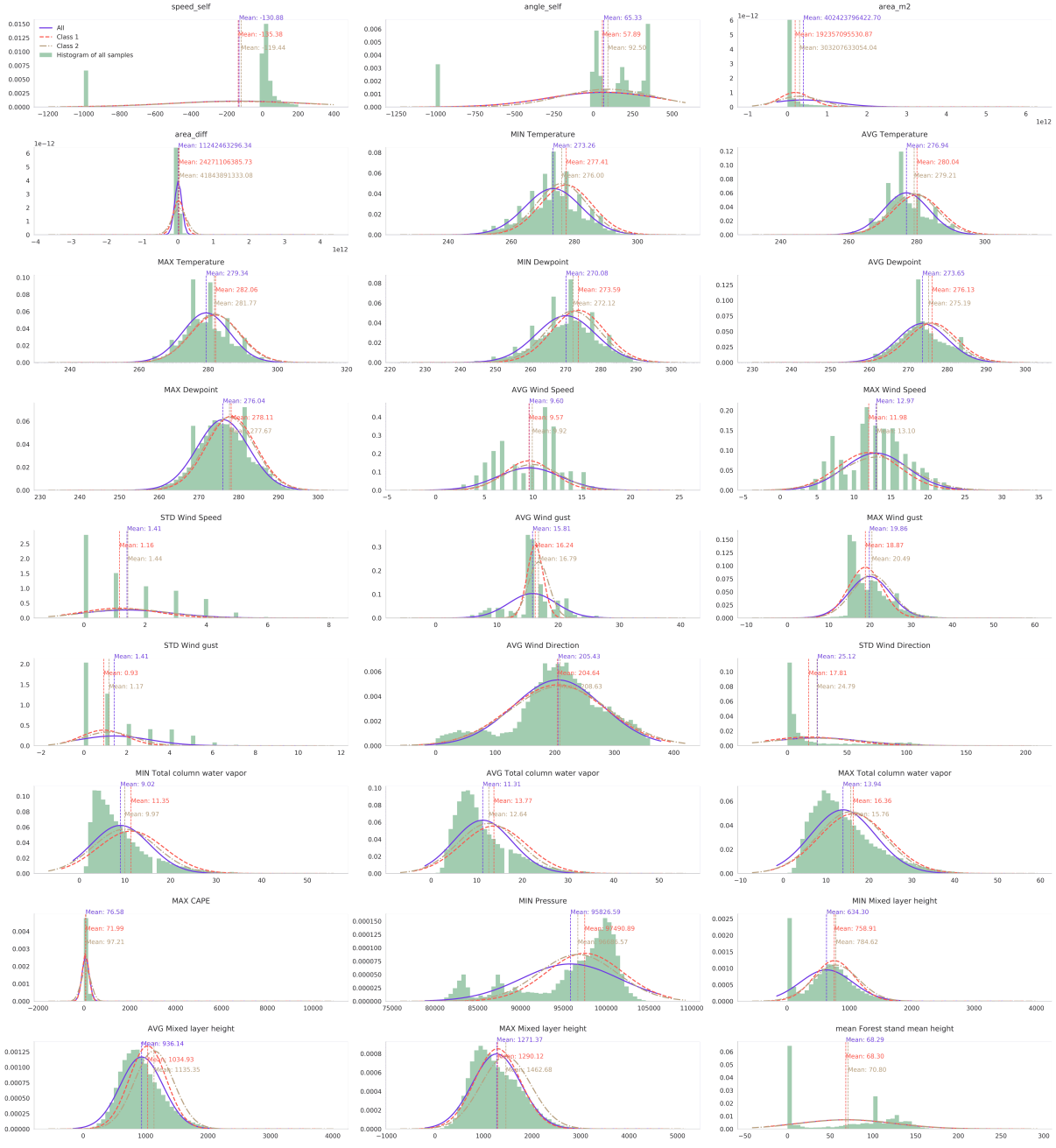


Figure A2. Distribtution of selected parameters of all storm objects (samples) and objects with class 2. National dataset.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences
Discussions

Author contributions. RT conceptualized, designed, and developed the method. IL contributed with meteorological expertise such as selecting used data and meteorological features along with correct thresholds. IL also helped in analyzing the performance. AM provided supervision from a meteorological perspective and AJ from a machine learning perspective. All contributed in presenting the results.

Competing interests. The authors declare that they have no conflict of interests.

Acknowledgements. The authors thanks project partners Järvi-Suomen Energia, Loiste Sähköverkko, and Imatran Seudun Sähkönsiirto for data and experience.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

References

2018 Syystrombeja, https://myrskyvaroitus.com/index.php/myrskytieto/myrskyhistoria/202-2018-syystrombeja.

Allen, M., Fernandez, S., Omitaomu, O., and Walker, K.: Application of hybrid geo-spatially granular fragility curves to improve power outage predictions, Journal of Geography & Natural Disasters, 4, 1–6, https://doi.org/10.4172/2167-0587.1000127, 2014.

Autonomio: Talos (software), http://github.com/autonomio/talos., 2020.

Barredo, J.: No upward trend in normalised windstorm losses in Europe: 1970-2008, Natural Hazards and Earth System Sciences, 10, https://doi.org/10.5194/nhess-10-97-2010, 2010.

Bergstra, J. and Bengio, Y.: Random search for hyper-parameter optimization, Journal of Machine Learning Research, 13, 281–305, 2012.

Breiman, L.: Random forests, Machine learning, 45, 5–32, 2001.

Brown, I. and Mues, C.: An experimental comparison of classification algorithms for imbalanced credit scoring data sets, Expert Systems with Applications, 39, 3446–3453, 2012.

Campbell, R. J. and Lowry, S.: Weather-related power outages and electric system resiliency, Tech. rep., Congressional Research Service, Library of Congress Washington, DC, 2012.

Chan, T., Golub, G., and LeVeque, R.: Updating formulae and a pairwise algorithm for variances computing sample, in: COMPSTAT, p. 30, Springer Science & Business Media, 1979.

Chang, C.-C. and Lin, C.-J.: LIBSVM: A library for support vector machines, ACM transactions on intelligent systems and technology (TIST), 2, 27, 2011.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P.: SMOTE: synthetic minority over-sampling technique, Journal of artificial intelligence research, 16, 321–357, 2002.

Chen, P.-C. and Kezunovic, M.: Fuzzy logic approach to predictive risk analysis in distribution outage management, IEEE Transactions on Smart Grid, 7, 2827–2836, 2016.

Chen, S. H., Jakeman, A. J., and Norton, J. P.: Artificial intelligence techniques: an introduction to their use for modelling environmental systems, Mathematics and computers in simulation, 78, 379–400, 2008.

Cintineo, J. L., Pavolonis, M. J., Sieglaff, J. M., and Lindsey, D. T.: An empirical model for assessing the severe weather potential of developing convection, Weather and Forecasting, 29, 639–653, 2014.

Csilléry, K., Kunstler, G., Courbaud, B., Allard, D., Lassegues, P., Haslinger, K., and Gardiner, B.: Coupled effects of wind-storms and drought on tree mortality across 115 forest stands from the Western Alps and the Jura mountains, Global change biology, 23, https://doi.org/10.1111/gcb.13773, 2017.

Donat, M. G., Leckebusch, G. C., Wild, S., and Ulbrich, U.: Future changes in European winter storm losses and extreme wind speeds inferred from GCM and RCM multi-model simulations, Natural Hazards and Earth System Sciences, 11, 1351–1370, https://doi.org/10.5194/nhess-11-1351-2011, https://www.nat-hazards-earth-syst-sci.net/11/1351/2011/, 2011.

Dozat, T.: Incorporating nesterov momentum into adam, 2016.

Eskandarpour, R. and Khodaei, A.: Machine Learning Based Power Grid Outage Prediction in Response to Extreme Events, IEEE Transactions on Power Systems, 32, 3315–3316, 2017.

Finnish Energy, E. O.: Energy distribution interruptions 2010-2018 - Finnish Energy, https://energia.fi/julkaisut/materiaalipankki/sahkon_keskeytystilastot_2010-2018.html#material-view, 2010-2018.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Finnish Energy, E. O.: Energy distribution interruptions 2011 - Finnish Energy, Tech. rep., Finnish Energy ry, https://energia.fi/files/605/Keskeytystilasto_2011.pdf, 2011.

Finnish Energy, E. O.: Energy distribution interruptions 2013 - Finnish Energy, Tech. rep., Finnish Energy ry, https://energia.fi/files/605/Keskeytystilasto_2013.pdf, 2013.

Finnish Meteorological Institute: Climate Statistics Report (December 2011), pp. 1–24, 2011.

Finnish Meteorological Institute: Rauli nousi myrskyjen raskaaseen sarjaan | Ilmastokatsaus – Ilmatieteen laitos, Ilmastokatsaus, http://www.ilmastokatsaus.fi/2016/08/29/rauli-nousi-myrskyjen-raskaaseen-sarjaan/, 2016.

Finnish Meteorological Institute: Kesäkuun 2018 kuukausikatsaus | Ilmastokatsaus – Ilmatieteen laitos, http://www.ilmastokatsaus.fi/2018/07/04/kesakuun-2018-kuukausikatsaus/, 2018.

Finnish Meteorological Institute: Open data, observation download service, https://en.ilmatieteenlaitos.fi/download-observations, 2020.

Finnish Meteorological Institute, Twitter: Ilmatieteen laitos Twitterissä: "#Juhannus'aaton-#myrsky #Pauliina lukuina klo 15.30: kovin puuska 32,5 m/s ja 10min keskituuli 26,5 m/s (Kaskinen Sälgrund), maa-alueilla Helsinki Kumpula puuska 22,7 m/s. Vaasa Klemettilä 24h sademäärä 41 mm." / Twitter, https://twitter.com/meteorologit/status/1010144460703969280, 2020.

Gardiner, B., Blennow, K., Carnus, J., Fleischner, P., Ingemarson, F., Landmann, G., Lindner, M., Marzano, M., Nicoll, B., Orazio, C., Peyron, J., Reviron, M., Schelhaas, M., Schuck, A., Spielmann, M., and Usbeck, T.: Destructive storms in European Forests: Past and Forthcoming Impacts. Final report to European Commission - DG Environment,, European Forest Institute, 2010.

Gardiner, B., Schuck, A., Schelhaas, M.-J., Orazio, C., Blennow, K., and Nicoll, B.: Living with Storm Damage to Forests What Science Can Tell Us What Science Can Tell Us, European Forest Institute, EFI, Joensuu, Finland, https://doi.org/10.13140/2.1.1730.2400, 2013.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep Learning, in: Deep Learning, pp. 164–223, MIT Press, 2016.

Govorushko, S. M.: Natural processes and human impacts: Interactions between humanity and the environment, Springer Science & Business Media, 2011.

Gregow, H., Laaksonen, A., and Alper, M. E.: Increasing large scale windstorm damage in Western, Central and Northern European forests, 1951-2010, Scientific Reports, 7, https://doi.org/10.1038/srep46397, 2017.

Guikema, S. D., Quiring, S. M., and Han, S. R.: Prestorm Estimation of Hurricane Damage to Electric Power Distribution Systems, Risk Analysis, 30, 1744–1752, https://doi.org/10.1111/j.1539-6924.2010.01510.x, 2010.

Guikema, S. D., Nateghi, R., Quiring, S. M., Staid, A., Reilly, A. C., and Gao, M.: Predicting Hurricane Power Outages to Support Storm Response Planning, IEEE Access, 2, 1364–1373, https://doi.org/10.1109/ACCESS.2014.2365716, 2014.

Haarsma, R. J., Hazeleger, W., Severijns, C., De Vries, H., Sterl, A., Bintanja, R., Van Oldenborgh, G. J., and Van Den Brink, H. W.: More hurricanes to hit western Europe due to global warming, Geophysical Research Letters, 40, 1783–1788, https://doi.org/10.1002/grl.50360, 2013.

Han, S. R., Guikema, S. D., and Quiring, S. M.: Improving the predictive accuracy of hurricane power outage forecasts using generalized additive models, Risk Analysis, 29, 1443–1453, https://doi.org/10.1111/j.1539-6924.2009.01280.x, 2009.

Hanewinkel, M.: Neural networks for assessing the risk of windthrow on the forest division level: a case study in southwest Germany, European Journal of Forest Research, 124, 243–249, 2005.

Hanninen, K. and Naukkarinen, J.: 570 000 customers experienced power losses at the end of the year, (Loppuvuoden sähkökatkoista kärsi 570 000 asiakasta - Energiateollisuus, http://www.mynewsdesk.com/fi/pressreleases/loppuvuoden-saehkoekatkoista-kaersi-570-000-asiakasta-725038, 2012.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Natural Hazards and Earth System Sciences Discussions

Hart, E., Sim, K., Kamimura, K., Meredieu, C., Guyon, D., and Gardiner, B.: Use of machine learning techniques to model wind damage to forests, Agricultural and forest meteorology, 265, 16–29, 2019.

He, J., Wanik, D. W., Hartman, B. M., Anagnostou, E. N., Astitha, M., and Frediani, M. E.: Nonparametric Tree-Based Predictive Modeling of Storm Outages on an Electric Distribution Network, Risk Analysis, 37, 441–458, https://doi.org/10.1111/risa.12652, 2017.

Hersbach, H., Bell, W., Berrisford, P., Horányi, A., J., M.-S., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., and Dee, D.: Global reanalysis: goodbye ERA-Interim, hello ERA5, pp. 17–24, https://doi.org/10.21957/vf291hehd7, https://www.ecmwf.int/node/19027, 2019.

Hoegh-Guldberg, O., Jacob, D., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K., Engelbrecht, F., Guiot, J., et al.: Impacts of 1.5 C global warming on natural and human systems, Global warming of 1.5° C. An IPCC Special Report, 2018.

Holdaway, M. R.: Spatial modeling and interpolation of monthly temperature using kriging, Climate Research, 6, 215–225, 1996.

Ilta-Sanomat: Rauli-myrsky repii Suomea: lähes 200000 taloutta ilman sähköjä - Kotimaa - Ilta-Sanomat, https://www.is.fi/kotimaa/art-2000001248969.html, 2016.

Kankanala, P., Pahwa, A., and Das, S.: Regression models for outages due to wind and lightning on overhead distribution feeders, in: Power and Energy Society General Meeting, 2011 IEEE, pp. 1–4, IEEE, 2011.

Kankanala, P., Pahwa, A., and Das, S.: Estimation of Overhead Distribution System Outages Caused by Wind and Lightning Using an Artificial Neural Network, in: International Conference on Power System Operation & Planning, 2012.

Kankanala, P., Das, S., and Pahwa, A.: AdaBoost^{+}: An Ensemble Learning Approach for Estimating Weather-Related Outages in Distribution Systems, IEEE Transactions on Power Systems, 29, 359–367, 2014.

Kufeoglu, S. and Lehtonen, M.: Cyclone Dagmar of 2011 and its impacts in Finland, in: IEEE PES Innovative Smart Grid Technologies Conference Europe, vol. 2015-January, IEEE Computer Society, https://doi.org/10.1109/ISGTEurope.2014.7028868, 2015.

Li, Z., Singhee, A., Wang, H., Raman, A., Siegel, S., Heng, F.-L., Mueller, R., and Labut, G.: Spatio-temporal forecasting of weather-driven damage in a distribution system, in: 2015 IEEE Power & Energy Society General Meeting, pp. 1–5, IEEE, 2015.

Lindsay, D. and Cox, S.: Effective probability forecasting for time series data using standard machine learning techniques, in: International Conference on Pattern Recognition and Image Analysis, pp. 35–44, Springer, 2005.

Liu, Y., Zhong, Y., and Qin, Q.: Scene Classification Based on Multiscale Convolutional Neural Network, IEEE Transactions on Geoscience and Remote Sensing, 56, 7109 – 7121, https://doi.org/10.1109/TGRS.2018.2848473, https://ieeexplore.ieee.org/abstract/document/8421052, 2018.

Mäkisara, K., Katila, M., Peräsaari, J., and Tomppo, E.: The multi-source national forest inventory of Finland–methods and results 2013, 2016.

Masson-Delmotte, T., Zhai, P., Pörtner, H., Roberts, D., Skea, J., Shukla, P., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., et al.: IPCC, 2018: Summary for Policymakers. In: Global warming of 1.5 C. An IPCC Special Report on the impacts of global warming of 1.5 C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global, Geneva, Switzerland, 2018.

Nateghi, R., Guikema, S., and Quiring, S. M.: Power Outage Estimation for Tropical Cyclones: Improved Accuracy with Simpler Models, Risk Analysis, 34, 1069–1078, https://doi.org/10.1111/risa.12131, 2014.

Niemelä, E.: KESKEYTYSTILASTO 2017 (i), Tech. Rep. 2018-06-14 11:51:52.916, Energiateollisuus Ry, Eteläranta 10, 00130 Helsinki, Finland, https://energia.fi/files/2785/Sahkon_keskeytystilasto_2017.pdf, 2018.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Nurmi, V., Pilli-Sihvola, K., Gregow, H., and Perrels, A.: Overadaptation to Climate Change? The Case of the 2013 Finnish Electricity Market Act, Economics of Disasters and Climate Change, 3, 161–190, https://doi.org/10.1007/s41885-018-0038-1, 2019.

Peltola, H., Kellomäki, S., and Väisänen, H.: Model Computations of the Impact of Climatic Change on the Windthrow Risk of Trees, Climatic Change, 41, 17–36, https://doi.org/10.1023/A:1005399822319, 1999.

Pinto, J. G., Bellenbaum, N., Karremann, M. K., and Della-Marta, P. M.: Serial clustering of extratropical cyclones over the North Atlantic and Europe under recent and future climate conditions, Journal of Geophysical Research: Atmospheres, 118, 12,476–12,485, https://doi.org/10.1002/2013JD020564, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD020564, 2013.

Platt, J. et al.: Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods, Advances in large margin classifiers, 10, 61–74, 1999.

Ramon, J., Lledó, L., Torralba, V., Soret, A., and Doblas-Reyes, F. J.: What global reanalysis best represents near-surface winds?, Natural Disasters and Adaptation to Climate Change, 145, 3236–3251, https://doi.org/10.1002/qj.3616, 2019.

Rasmussen, C. E.: Gaussian processes in machine learning, in: Summer School on Machine Learning, pp. 63–71, Springer, 2003.

Re, M.: After the floods, Natural catastrophes, 2013.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., and Dormann, C. F.: Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure, Ecography, 40, 913–929, https://doi.org/10.1111/ecog.02881, https://onlinelibrary.wiley.com/doi/abs/10.1111/ecog.02881, 2017.

Rossi, P. J.: Object-Oriented Analysis and Nowcasting of Convective Storms in Finland, Ph.D. thesis, Aalto University, http://urn.fi/URN:ISBN:978-952-60-6441-3, 2015.

Schelhaas, M.-J.: Impacts of natural disturbances on the development of European forest resources: application of model approaches from tree and stand levels to large-scale scenarios, Dissertationes Forestales, https://doi.org/10.14214/df.56, 2008.

Schelhaas, M.-J., Nabuurs, G.-J., and Schuck, A.: Natural disturbances in the European forests in the 19th and 20th centuries, Global Change Biology, 9, 1620–1633, https://doi.org/10.1046/j.1365-2486.2003.00684.x, https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2486.2003.00684.x, 2003.

Seidl, R., Schelhaas, M., Rammer, W., and Verkerk, P.: Increasing forest disturbances in Europe and their impact on carbon storage, Nature Climate Change, 4, 806–810, https://doi.org/10.1038/NCLIMATE2318, 2014.

Shield, S. et al.: Predictive Modeling of Thunderstorm-Related Power Outages, Master's thesis, The Ohio State University, 2018.

Singhee, A. and Wang, H.: Probabilistic forecasts of service outage counts from severe weather in a distribution grid, in: 2017 IEEE Power & Energy Society General Meeting, pp. 1–5, IEEE, 2017.

Tervo, R., Karjalainen, J., and Jung, A.: Short-Term Prediction of Electricity Outages Caused by Convective Storms, IEEE Transactions on Geoscience and Remote Sensing, 57, 8618–8626, 2019.

Ukkonen, P. and Mäkelä, A.: Evaluation of machine learning classifiers for predicting deep convection, Journal of Advances in Modeling Earth Systems, 11, 1784–1802, 2019.

Ukkonen, P., Manzato, A., and Mäkelä, A.: Evaluation of thunderstorm predictors for Finland using reanalyses and neural networks, Journal of Applied Meteorology and Climatology, 56, 2335–2352, 2017.

Ulbrich, U., Pinto, J. G., Kupfer, H., Leckebusch, G., Spangehl, T., and Reyers, M.: Changing Northern Hemisphere storm tracks in an ensemble of IPCC climate change simulations, Journal of Climaste, 21, 1669–1679, 2008.

https://doi.org/10.5194/nhess-2020-205
Preprint. Discussion started: 3 August 2020

Ulbrich, U., Leckebusch, G. C., and Pinto, J. G.: Extra-tropical cyclones in the present and future climate: A review, in: Theoretical and Applied Climatology, vol. 96, pp. 117–131, Springer Wien, https://doi.org/10.1007/s00704-008-0083-8, 2009.

Valta, H., Lehtonen, I., Laurila, T., Venäläinen, A., Laapas, M., and Gregow, H.: Communicating the amount of windstorm induced forest damage by the maximum wind gust speed in Finland, Advances in Science and Research, 16, 31–37, https://doi.org/10.5194/asr-16-31-2019, 2019.

Wang, G., Xu, T., Tang, T., Yuan, T., and Wang, H.: A Bayesian network model for prediction of weather-related failures in railway turnout systems, Expert Systems with Applications, 69, 247–256, https://doi.org/10.1016/j.eswa.2016.10.011, 2017.

Wilks, D. S.: Statistical methods in the atmospheric sciences, vol. 100, Academic press, 2011.

Williams, C. K. and Rasmussen, C. E.: Gaussian processes for machine learning, vol. 2, MIT press Cambridge, MA, 2006.

Yue, M., Toto, T., Jensen, M. P., Giangrande, S. E., and Lofaro, R.: A Bayesian approach-based outage prediction in electric utility systems using radar measurement data, IEEE Transactions on Smart Grid, 9, 6149–6159, https://doi.org/10.1109/TSG.2017.2704288, 2018.