Environmental extreme events detection: A survey

Rakesh Matta[1], Rui Wu[1], and Shanyue Guan[2]

[1] Department of Computer Science, East Carolina University, Greenville, North Carolina, U.S.A
mattar18@students.ecu.edu; wur18@ecu.edu

[2] Department of Engineering, East Carolina University, Greenville, North Carolina, U.S.A
guans18@ecu.edu

Abstract

The application of statistics in extreme events detection is quite diverse and leads to diverse formulations, which needs to be designed for the specific problem. Each formula needs to be tailored specifically to work with the available data in the given situation. This diversity is one of the driving forces of this survey towards identifying the most common mixture of components utilized in the analysis of environmental outlier detection. Indeed for some arbitrary applications, it may not always be possible to use off-the-shelf models due to the wide variations in problem formulations. In this paper, we summarize the statistical methods involved in the detection of environmental extremes such as wind ramps, high precipitation and extreme temperatures. Then we organize the discussion along different outlier detection types, present various outlier definitions, and briefly introduce the corresponding techniques. Environmental extreme events detection challenges and possible future work are also discussed.

1   Introduction

Environmental extreme events are unusual and unexpected weather conditions such as heat waves, droughts, heavy downpour, wild winds or hurricanes. These events have increased in frequency and/or magnitude because of anthropogenic climate change. Low precipitation and heat waves leading to forest fires have occurred in different area of the world. It can have deep impacts and cause loss of human and animal lives. These extremes can be devastating events in themselves. For example, floods can cause famines; high temperatures and high speed wind can lead to forest fires or health problems in humans like heat stroke. Therefore, it is very important to study environmental extreme events.

An outlier is an observation that significantly differs from other observations of the same feature. If a time series of data is plotted, outliers are usually the unexpected spikes or dips of observations at given points in time. An outlier may exist due to a rare event, other population, incorrect values as a result of errors or breakdowns or bad recording practices[29]. Outliers can be identified by learning a model based on all the time series sequences in the data base, and then compute an outlier score for each sequence using a scoring function based on the model, which can be either supervised or unsupervised depending on the availability of training data[8]. We can also detect an outlier based on the frequent patterns. A data instance is likely to be an

outlier, if it doesn't contain frequent patterns[32]. We can identify outliers by computing outlier score, which can be the deviation of actual value from the predicted value or the distance to the centroid of the closest data cluster. In [35], they considered the outlier score to be proportional to the density of its k-nearest neighbors over the local density of the data point. In case of sliding window approach, the time series sequence can be broken into multiple overlapping windows of fixed length, an outlier score is computed for each window and aggregated later to identify anomalies. In Local Outlier Factor (LOF) method, anomalous data points are identified by measuring the local deviation of a given data point with respect to its neighbors. A better approach is proposed in [21], where the incremental LOF algorithm computes LOF value for each data record inserted into the data set and instantly determines whether the inserted data record is an outlier. In [30], they proposed a solution for the distributed Local Outlier Factor approach, which is inherently parallel and deployable on virtually any distributed infrastructure. They also designed a multi-step pipeline framework called Distributed Local Outlier Factor which leverages the invariant observation and computes LOF scores in a highly distributed fashion.

2   Taxonomy

Extreme events have generated interest world over, because of their potential for high impacts on ecological, technical and social systems. There arises a need for an organized and detailed study of the work done in extreme events detection using various statistical approaches. Out of all the environmental extreme events, we briefly surveyed and illustrated the statistical methods for detecting climatic and wind ramp events as shown in figure 1.

Climatic events are occasional variations producing extreme values of climate indicators, such as temperature and precipitation. Climate change can potentially change the intensity, frequency, timing and duration of these events. The climatic domain shall cover the statistical approaches used to detect extreme precipitation and temperature related events. Detection of climatic extremes mostly deal with Extreme value theory involving peak over threshold, generalized pareto distribution and maximum likelihood estimation approaches. whereas Wind ramp events are unforeseen increases or decreases in wind speed or detection along a very short period of time. Changing wind turbines design based on wind ramp detection can possibly increase electricity generation rate. wind ramp domain shall specifically cover the detection of wind ramps and increase wind farms efficiency. Sliding window and dynamic programming approaches are often utilized by several researchers to detect wind ramps in an efficient way. Both the climatic and the wind ramps events deal with the temporal data. Modeling temporal data is a challenging task due to the dynamic nature and complex evolutionary patterns in the data.

3   Climatic events

Over the years, changes in the variability of climatic extremes had more impacts compared to the changes in the mean of climate[12]. Simulations of various models are analyzed and compared with each other and those simulations can have statistically significant links. For example, precipitation quantity can vary depending on the frequency or intensity of each precipitation events, or a combination of both factors. The intensity of precipitation refers to the amount of precipitation associated with specific quantities of the precipitation distribution. It is possible to estimate the proportion of any trend in total precipitation that is attributable to changes in frequency versus changes in precipitation intensity. The patterns of overall precipitation, on the
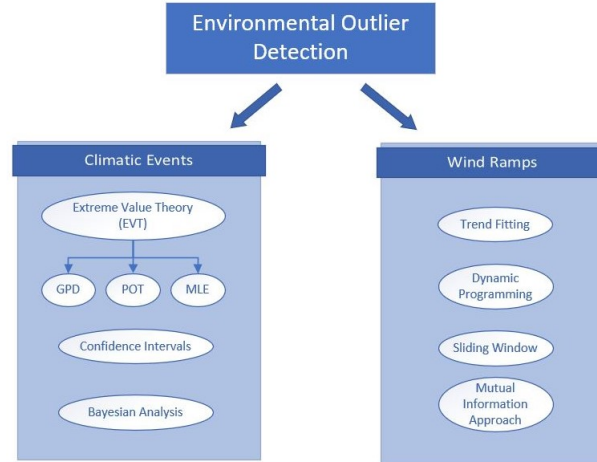
Figure 1: Commonly used statistical methods related to climatic and wind ramp events

other hand, may be conveyed in terms of a percentage of average precipitation over given period. Precipitation patterns may also be derived for particular quantities. Certain statistical data are more frequently used than others, such as the daily maximum quantity of precipitation. This statistic is obtained by identifying the maximum and the median quantities of precipitation per month throughout all the years of recorded data, then by computing the pattern through the obtained values. Alternatively, the ratio of precipitation quantity in a particular area may be matched up to the mean precipitation quantity in the entire area in order to see the variations of precipitation. This statistic also reveals variations in the precipitation distribution that are independent of the variations in the mean quantities[7, 11, 18]. The statistical approaches used for detecting Climatic extremes are:

3.1   Extreme Value Theory

Extreme value theory is a powerful statistical tool developed to study the laws of extreme events i.e. Extreme Value Distributions (EVD). Using these EVDs, we can analyze the input time series using Normal Distribution. When the data doesn't fit a Normal Distribution, we can use Log Normal Distribution, which transforms not-normal distribution into normal[24]. By fitting an EVD to the unknown distribution, it is possible to evaluate the probability of potential extreme events. These Extreme value distributions have the following form:

$$G_\gamma : \ x \ \mapsto \ exp \ \left(-(1+\gamma x)^{-\frac{1}{\gamma}}\right), \ \gamma \in R, \ (1+\gamma x) \geq 0$$

Extremes of common standard distributions follow the above distribution and the extreme value index(γ) depends on it. It may not be reasonable to apply a method directly to detect climatic extreme events if the method assumes data follow a certain distribution. Because climatic data may not follow any data distribution. The extreme value theory, through the POT approach (introduced in Subsection 3.2), gives us a way to estimate threshold without any strong assumption and clear knowledge about the distribution. we have to detect abnormal events in a stream in a blind way without any knowledge about the distribution[3, 27].

3.2   Peak Over Threshold

Climatic data is usually temporal data. Peak Over Threshold (POT) is one of the efficient methods to analyze peak values in the time series. Peak values in the time series can be analyzed using POT method. Initially we set a subjective threshold to obtain all the data points(peaks) which exceed the threshold and then fit a Generalized Pareto Distribution (GPD) to those peaks (introduced in Subsection 3.3). The estimation of GPD parameters i.e., shape and scale parameters can be done using various methods among which maximum likelihood function (introduced in Subsection 3.4), is a preferred method by many other researchers[24].

Many applications rely on high throughput streaming data. In [27], the authors proposed two streaming algorithms: streaming peak over threshold(SPOT) for stationary cases and Drift SPOT (DSPOT) which consider the drift component. Firstly, a POT estimation is performed on the first few observations and a threshold is initialized. Then for all the next observed values, they flagged all the abnormal values and updated the threshold. They built a streaming outlier detector, which uses the next observations to both detect the outliers and refine the threshold value. SPOT assumes that the distribution doesn't change over time, but it might be conservative. They proposed DSPOT which made SPOT run not on the absolute values but on the relative ones considering the variable changes modelling the local behavior at every moment.

3.3   Generalized Pareto Distribution

POT data consist of selecting all the events exceeding a high threshold. When the threshold value increases, the limit distribution of a POT series can be approximated with a Generalized Pareto Distribution. Rather than fitting an EVD to the extreme values of x, the POT approach tries to fit a GPD to the excesses over threshold values (x - t). The excess over a threshold, written (x - t) are likely to follow a GPD with scale(σ) and location(γ) parameters[14]. In fact, the GPD needs a third parameter, the location , but it is null in our case[15, 17, 27, 33].

$$G_{\gamma,\beta}(x) = \begin{cases} 1 - \left(1 + \frac{\gamma x}{\beta}\right), & \text{if } \gamma! = 0 \\ 1 - exp\left(-\frac{x}{\beta}\right), & \text{if } \gamma = 0 \end{cases}$$

valid when,
$$\beta > 0 \text{ and } x \geq 0 \rightarrow \gamma \geq 0$$
$$0 \leq x \leq -\frac{\beta}{\gamma}$$

3.4   Maximum Likelihood Estimation

Some classical methods like method of moments, least squares method and probability weighted moments can be used to estimate σ and γ, but they are less efficient and not robust compared to maximum likelihood estimation. In method of moments, the first 'm' statistical moments of the target distribution are compared to the moments derived from the observation, whereas in least squares method, a flexible method is used for fitting any kind of data sets and distribution functions. Parameters are estimated using linear and non linear regression methods. Linear regression fits the distributions with two parameters and nonlinear regression fits the distributions having three or more parameters[29].

Maximum likelihood estimation helps us to evaluate the parameters using the observations. If $x_1, x_2, ...x_n$ are n independent observations of a random variable x, where density is parameterized by θ, the likelihood function is defined by:[27]

$$\mathcal{L}(x_1,.....x_n;\theta) \;=\; \prod_{i=1}^{n} f_\theta(x_i)$$

This equation represents the joint density of these n observations. The parameter θ is estimated by maximizing the likelihood. Maximum likelihood can be used to estimate the parameters of GPD by maximizing the $log\mathcal{L}$.

$$log\mathcal{L}(\gamma,\sigma) \;-N_t\, log\sigma \left(1+\frac{1}{\gamma}\right) \sum_{i=1}^{N_t} log\left(1+\frac{\gamma}{\sigma}Y_i\right)$$

This method performs well for estimating location and scale parameters, especially when the sample size is large[9, 15]. This is very common for climatic data, which usually contain many features and are collected for decades. In Maximum likelihood Estimation, the parameters of the probability density function are estimated by maximizing the likelihood function by the choice of model parameters for which $\frac{dl}{d\alpha} = 0$. In practice, it is more convenient to work with the logarithm of likelihood.

3.5   Confidence Intervals

In order to identify an area of unpredictability or variability, confidence intervals may be utilized. These intervals may also be used for testing theories or suppositions of significant deviations that fall in the category of the theory of temporal clustering, or cases where there is no apparent pattern in climatic data. The confidence interval clearly identifies the boundaries of expected variability, and all regions that are outside these boundaries are deemed to be statistically significant. Consequently, all regions that are within the boundaries are statistically insignificant i.e, the hypothesis is accepted. Following this standard, the statistical significance of any particular theory can easily be determined. By using the method of parametric bootstrapping, a set of extreme values are placed in distribution. Then random values are created and finally confidence intervals are approximated. Originally, this method was used only for independent values but with the development of block bootstrapping, it is expanded to cover the scrutiny of dependent values. Block bootstrapping groups data in blocks from which the resampling is made and thus, preserving the time within the series. Nonetheless, the data is still considered to be independent since the POT data is taken out only after applying independence measures[19, 29].

3.6   Bayesian Analysis

By using Bayes theorem and applying specific conjectures, values may be neutrally designated to a variety of climatic models that results in a probability distribution of climate change in the future[26]. Bayesian analysis has concerns about the parameters vector as it is viewed as a constant quantity, which must be estimated. Therefore, there has to be a prior distribution that is supposed to reveal earlier awareness of the parameters and can be evaluated without relying solely on observations. Supposing that there is a set of existing data, considering the probability and the prior distribution the parameters of posterior distribution can be obtained through Bayes theorem. This posterior distribution is now considered as the updated version of earlier data. It has a more widely inclusive estimation compared to the traditional point estimation. But as the coverage increases, the posterior distribution can no longer be used in simple applications. Hence more complicated computational approaches are necessary[23].

Researchers have suggested the use of a Bayesian predictive technique to the POT or peaks over threshold method[2]. In this approach, the values in the higher order statistics are treated as separate parameters in the system, along with a suitable prior distribution. A weighted average is also derived from a number of potential threshold values by utilizing the predictive distribution, which eliminates the difficulties introducing from using very small sizes of samples. Given the probability and the prior distributions, the Bayes theorem can thus be used to acquire the posterior distribution, as well as a gamma distribution for values that are underneath the logarithmic threshold.

4   Wind Ramps

Extreme fluctuations in winds can result in damages on wind turbines. Awareness and knowledge about extreme phenomena that can affect safety and the produced power of the wind turbines. Available historical data are used to train various models to classify the future ramp events using various statistical approaches like bayesian classifiers, support vector machines, neural networks, decision trees etc. Features are extracted with the help of these models to categorize the wind ramp events based on a set of predefined thresholds. [31]. In [25], Raffi Sevlian and Ram Rajagopal described an optimal ramp detection technique for identifying wind ramp events of varying lengths under any arbitrary rule set for large time series. A dynamic programming recursion with trend fitting is used to optimally segment wind power data to ramp and non-ramp events. An optimized swinging door algorithm (OpSDA) by utilizing the swinging door algorithm and dynamic programming algorithm to handle the wind power bumps was developed in [4, 6]. The data is segregated through a piecewise linear approximation and the segments are optimized by merging adjacent segments with the same ramp changing direction. This optimized swinging door algorithm (OpSDA) was adopted in [5] and extended to detect wind ramps events, by merging bumps, having a different changing direction, into adjacent ramping segments to improve the performance of the OpSDA method.

4.1   Identifying Wind Ramps

In case of wind ramp events, the time scale of interest is on the order of minutes to hours. First, we must bring all the data to a consistent format by normalizing them to name plate capacity of the system and filling all the missing values . We can also ignore some missing values and neglect them in the statistical analysis. Finally, pre-processed data can be used to model the ramp events using the statistical framework[20]. We can consider an event as a ramp, if a large increase or decrease in wind power within a given time span i.e, the difference between the magnitudes of power generation in a time interval is greater than a specific threshold value$| \Delta p_{(t)} |= p_{(t)} - p_{(t-\Delta t)}$, where $| \Delta p_{(t)} | > \mu$ [10, 25, 31]. There is a chance of missing few ramp events occurring between the end points of an interval. Hence, if the difference between the maximum and minimum values of the wind power within a time span is greater than a specified threshold, i.e, $p_{max_{t_1}} - p_{min_{t_2}} > \mu_0$, then it can be a ramp event[10, 25]. The rate of increase (or decrease) in wind power with in a time span is greater than a specified threshold value, $\frac{p_{(t)} - p_{(t-\Delta t)}}{\Delta t} > \mu_1$, is considered as a ramp[25].

4.2   Ramp Detection Techniques

4.2.1   Trend Fitting

The time series wind power data contain uncertainty and underlying structure at multiple time scales. Investigating phenomenon at appropriate time scales requires eliminating noise or trends outside the time scale of interest. It is important to reduce data size while still preserving appropriate trends in the data set. Trend fitting can be done as a preprocessing step to remove short term fluctuations in wind power. It also eliminates the noises and trends outside the time scale[25].

4.2.2   Dynamic Programming

Optimal ramp intervals simply mean that a ramp event refers to the longest series of points that comply with the predetermine ramp rules. A ramp event has a start and an end point, both of which define the boundaries of the ramp interval. The objective is to find the longest intervals. This can be done by analyzing each subsequence of the original data. The best approach for this is through dynamic programming. If an interval complies with ramp rules, all its subintervals also comply with the ramp rules. Each of these intervals is assigned a score that is a function corresponding to the interval length. Thus, finding a set of intervals that maximizes the score is all needed. Given an appropriate scoring function on intervals of the time series, the optimal ramp start and stop times are recovered by maximizing the objective function J according to the dynamic program.

$$J(i,j) = \max_{i < k \leq j} W(i,k) + J(k+1,j)$$

$J(i,j)$ is the maximum attained score in the interval. (i,j) is computed as the maximum over $j - i$ subproblems. Maximizing over all subproblems yields an optimal solution, since each subproblem is in terms of a maximum score is in the interval $(k+1,j)$, and $W(i,k)$ is the positive weight given to the interval (i,k). The dynamic programming requires a proper cost function that evaluates the cost of each subsequence and the ramp score can be computed as the maximum over sub problems [25].

4.2.3   Sliding Window

To efficiently detect data trends within a large temporal wind dataset, input signals are split into a number of overlapping sections parameterized by window length. Wind Ramps can be identified by processing these overlapping sections individually or in a parallel fashion which greatly reduces the computation time. The advantage of this technique can be maintained by maximizing the ramp duration parameter. i.e, if the window length is greater than the longest ramp duration, all the wind ramps are easily detected. It ensures that no ramp event is missing and the detected ramps are aggregated later. Window-based method can perform better localization of outliers, compared to the methods that output the entire time series as an outlier[25].

4.2.4   Mutual Information

Crucial features are required to build accurate wind ramp models. One of the commonly used method is Mutual Information Approach, where the Mutual independence of two random variables is measured. A feature is considered important, if the Mutual Information value between

the candidate feature and target class is high. Also, the Mutual Information value between the candidate feature and selected features need to be low, making the features independent of each other. Considering only the important features, we can improve the classification accuracy and reduce the computational costs. Hence, for classification of ramp up/down events, Mutual Information is useful for filtering out the less important inputs of the classification engine[1, 31].

5   Challenges and Future Work

Handling incomplete, missing and delayed information is a difficult task. noise, skewed distributions and redundant values also hinders the progress. In traditional offline analysis, data preprocessing is usually done manually by a human expert prior to modeling. In the streaming scenario manual processing is not feasible, as new data arrives continuously. Streaming data needs fully automated preprocessing methods, that can optimize the parameters and operate autonomously. Since new data arrives at every time instant, the scale of the data is very large. This often leads to processing and resource-constraint challenges like energy, memory, computational capacity and communication bandwidth usage.[13, 34].

Association of wind ramps with weather conditions might help us in identifying the days that are likely to have ramps[28]. Most of the existing methods deal with spatial and temporal variables on hard and subjective thresholds and those methods do not have a clear empirical definition accepted universally by every scientist in the world. A neural network-based system might be an alternative approach, which is capable of learning a broad class of patterns from complex multi-variable data and avoiding subjective threshold for extreme events detection. Training a Convolutional Neural Network can help us extract features from the available data sets and classify images to detect the extreme events like tropical cyclones, atmospheric rivers and hurricanes. Convolution neural networks (CNNs) eliminate the need for feature extraction which is one of the most important and time-consuming part of traditional machine learning methods. However, training a deep CNN model with a limited amount of training data is quite challenging as these approaches require more data than other machine learning methods, such as decision trees. Also, their performance is strongly correlated with the amount of available training data. Data are much harder and more expensive to collect than developing and applying the algorithms for execution. Without enough data, neural network may not be able to achieve desired level of accuracy.[16, 22].

6   Conclusion

This paper provides an insight into various existing outlier detection methods depending on the scenario and available data. The survey is classified based on the detection of climatic extremes and wind ramp events. For climatic events, we illustrated the statistical techniques used to detect climatic extremes over the decades, which can be used to identify problem settings in the future. For wind ramp events, we summarized the data pre-processing and feature extraction methods for analyzing the data to obtain useful patterns. We also illustrated various ramp detection techniques to categorize the future ramp events based on the available data. The detection of extreme weather events is significant in assistance of risk management and influencing governmental policy decisions. Extreme events cannot be avoided but their detection can give a cushion to reduce the damage. In order to save human lives and properties, its pertinent that a system is developed that helps in sending out warnings well ahead of the catastrophe.

7   Acknowledgements

References

[1] Roberto Battiti. Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on neural networks*, 5(4):537–550, 1994.

[2] Cibele N Behrens, Hedibert F Lopes, and Dani Gamerman. Bayesian analysis of extreme events with threshold estimation. *Statistical Modelling*, 4(3):227–244, 2004.

[3] Jan Beirlant, Yuri Goegebeur, Johan Segers, and Jozef L Teugels. *Statistics of extremes: theory and applications*. John Wiley & Sons, 2006.

[4] EH Bristol. Swinging door trending: Adaptive trend recording? In *ISA National Conf. Proc., 1990*, pages 749–754, 1990.

[5] Mingjian Cui, Jie Zhang, Cong Feng, Anthony R Florita, Yuanzhang Sun, and Bri-Mathias Hodge. Characterizing and analyzing ramping events in wind power, solar power, load, and netload. *Renewable energy*, 111:227–244, 2017.

[6] Mingjian Cui, Jie Zhang, Anthony R Florita, Bri-Mathias Hodge, Deping Ke, and Yuanzhang Sun. An optimized swinging door algorithm for identifying wind ramping events. *IEEE Transactions on Sustainable Energy*, 7(1):150–162, 2015.

[7] Christoph Frei and Christoph Schär. Detection probability of trends in rare events: Theory and application to heavy precipitation in the alpine region. *Journal of Climate*, 14(7):1568–1584, 2001.

[8] Manish Gupta, Jing Gao, Charu C Aggarwal, and Jiawei Han. Outlier detection for temporal data: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 26(9):2250–2267, 2014.

[9] Jonathan RM Hosking and James R Wallis. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29(3):339–349, 1987.

[10] Chandrika Kamath. Understanding wind ramp events through analysis of historical data. In *IEEE PES T&D 2010*, pages 1–6. IEEE, 2010.

[11] Thomas R Karl and Richard W Knight. Secular trends of precipitation amount, frequency, and intensity in the united states. *Bulletin of the American Meteorological society*, 79(2):231–242, 1998.

[12] Richard W Katz and Barbara G Brown. Extreme events in a changing climate: variability is more important than averages. *Climatic change*, 21(3):289–302, 1992.

[13] Georg Krempl, Indre Žliobaite, Dariusz Brzeziński, Eyke Hüllermeier, Mark Last, Vincent Lemaire, Tino Noack, Ammar Shaker, Sonja Sievi, Myra Spiliopoulou, et al. Open challenges for data stream mining research. *ACM SIGKDD explorations newsletter*, 16(1):1–10, 2014.

[14] John R Lanzante. Resistant, robust and non-parametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 16(11):1197–1226, 1996.

[15] Yun Li, Wenju Cai, and EP Campbell. Statistical modeling of extreme rainfall in southwest western australia. *Journal of climate*, 18(6):852–863, 2005.

[16] Yunjie Liu, Evan Racah, Joaquin Correa, Amir Khosrowshahi, David Lavers, Kenneth Kunkel, Michael Wehner, William Collins, et al. Application of deep convolutional neural networks for detecting extreme weather in climate datasets. *arXiv preprint arXiv:1605.01156*, 2016.

[17] Alexander J McNeil. Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bulletin: The Journal of the IAA*, 27(1):117–137, 1997.

[18] Gerald A Meehl, Julie M Arblaster, and Claudia Tebaldi. Understanding future patterns of increased precipitation intensity in climate model simulations. *Geophysical Research Letters*, 32(18), 2005.

[19] Victor Ntegeka and Patrick Willems. Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at uccle, belgium. *Water Resources Research*, 44(7), 2008.

[20] Tinghui Ouyang, Xiaoming Zha, and Liang Qin. A survey of wind power ramp forecasting. *Energy and Power Engineering*, 5(04):368, 2013.

[21] Dragoljub Pokrajac, Aleksandar Lazarevic, and Longin Jan Latecki. Incremental local outlier detection for data streams. In *2007 IEEE symposium on computational intelligence and data mining*, pages 504–515. IEEE, 2007.

[22] Evan Racah, Christopher Beckham, Tegan Maharaj, Samira Ebrahimi Kahou, Mr Prabhat, and Chris Pal. Extremeweather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. In *Advances in Neural Information Processing Systems*, pages 3402–3413, 2017.

[23] Benjamin Renard, Michel Lang, and Philippe Bois. Statistical analysis of extreme events in a non-stationary context via a bayesian framework: case study with peak-over-threshold data. *Stochastic environmental research and risk assessment*, 21(2):97–112, 2006.

[24] Gianluca Rosso. Extreme value theory for time series using peak-over-threshold method. *arXiv preprint arXiv:1509.01051*, 2015.

[25] Raffi Sevlian and Ram Rajagopal. Detection and statistics of wind power ramps. *IEEE Transactions on Power Systems*, 28(4):3610–3620, 2013.

[26] Mxolisi E Shongwe, Geert Jan van Oldenborgh, Bart van den Hurk, and Maarten van Aalst. Projected changes in mean and extreme precipitation in africa under global warming. part ii: East africa. *Journal of Climate*, 24(14):3718–3733, 2011.

[27] Alban Siffer, Pierre-Alain Fouque, Alexandre Termier, and Christine Largouet. Anomaly detection in streams with extreme value theory. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1067–1075. ACM, 2017.

[28] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[29] Gerbrant Van Vledder, Yoshimi Goda, Peter Hawkes, Etienne Mansard, Maria Jesus Martin, Martin Mathiesen, Eric Peltier, and Edward Thompson. Case studies of extreme wave analysis: a comparative analysis. In *Proceedings of the second international symposium on ocean wave measurement and analysis*, pages 978–992. ASCE New York, 1993.

[30] Yizhou Yan, Lei Cao, Caitlin Kulhman, and Elke Rundensteiner. Distributed local outlier detection in big data. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1225–1234. ACM, 2017.

[31] Hamidreza Zareipour, Dongliang Huang, and William Rosehart. Wind power ramp events classification and forecasting: A data mining approach. In *2011 IEEE Power and Energy Society General Meeting*, pages 1–3. IEEE, 2011.

[32] Weiwei Zhang, Jianhua Wu, and Jie Yu. An improved method of outlier detection based on frequent pattern. In *2010 WASE International Conference on Information Engineering*, volume 2, pages 3–6. IEEE, 2010.

[33] Xuebin Zhang, Francis W Zwiers, and Guilong Li. Monte carlo experiments on the detection of trends in extreme values. *Journal of Climate*, 17(10):1945–1952, 2004.

[34] Yang Zhang, Nirvana Meratnia, and Paul JM Havinga. Outlier detection techniques for wireless sensor networks: A survey. *IEEE Communications Surveys and Tutorials*, 12(2):159–170, 2010.

[35] Zhiruo Zhao, Chilukuri Mohan, and Kishan Mehrotra. Adaptive sampling and learning for unsupervised outlier detection. In *The Twenty-Ninth International Flairs Conference*, 2016.