Technical Report Precursors

E. Ramona Stefanescu

Department of Mechanical Engineering, University of Buffalo,

ers32@buffalo.edu, http://www.acsu.buffalo.edu/~ers32/

Abstract

This report presents an introduction into precursors of extreme increments. It tries to answer the question under which circumstances large events are better predictable than smaller events and how to choose a precursor in order to obtain good predictions.

Key words. Extreme events, time series

1 Introduction

Systems with a complex time evolution, which generate a great impact from time to time, are found everywhere (ubiquitous). Examples include stock market crashes, epileptic seizures, earthquakes etc. Due to the complexity of these systems, a complete modeling is impossible. This is why one applies the framework of prediction with precursory structures. One might expect that the more extreme an event is, the more difficult it is to predict it. However, it has been reported in the literature that extreme events are better predictable than small events.

There have been many attempts to investigate times series of historical data and try to infer knowledge about the future. One possibility to do so, consists in predicting via identifying precursory structures. The focus here is on predictions made via precursory structures, which are identified through the conditional probability that an event follows a given precursory structure.

In all types of phenomena where the event magnitude can take any value inside doe interval, one has to decide beyond which magnitude one calls an event "extreme".

Definition 1 (Extreme Events) *An extreme event needs to fulfill the following properties:*

- ***Quantifiability*** *It is possible to define a variable η which characterizes the magnitude of the event and which is a function $\eta(x)$ of a scalar or vector valued observable x of the system under study. $\eta(x)$ is a large value if the observable x is in a state, referred as extreme event.*

- ***Rareness*** *The event is rare, i.e., its occurrence is reflected by the tails of the probability distribution which describes the occurrence of all possible events (however not every rare event is an extreme event).*

- ***Irregularity*** *The event occurs irregularly, otherwise prediction of the events would be trivial.*

- ***Endogenity*** *The event is generated by the dynamics of the system itself and not as a consequence of external shocks.*

The occurrence and prediction of extreme events can be observed in a scalar time series, i.e., a set of N observations at discrete times t_n, where $t_n = t_0 + n\Delta$, with $n = 0, 1, \ldots, N-1$ with a fixed sampling interval Δ. We are not employing any prior knowledge about the physical process or phenomena under the study, but rely only on recordings of past data.

Due to the different fields in which precursor based prediction are discussed, there are also various interpretations of the term precursor. However, it is assumed that at least a part of the dynamics of the system under study is unknown and there therefore has ti be described by a stochastic terms. The consequence of this is, that we cannot demand from a precursor to precede *every* individual event, however, we can expect the data structure we call precursor to *typically* precede an event.

In order to predict an event occurring at time $(n+1)$ we compare the last τ observations

$$\mathbf{s}_{(n,\tau)} = (x_{n-\tau+1}, x_{n-\tau+2}, \ldots, x_{n-1}, x_n) \tag{1.1}$$

to which we will refer as the *precursory variable*, to a specific precursor, also called *predictor*,

$$\mathbf{u} = (u_{\tau-1}, u_{\tau-2}, \ldots, u_0). \tag{1.2}$$

An *alarm volume* $V(\mathbf{u}, \delta)$ around each precursor as the set of all $\mathbf{s}_n$ for which $\| \mathbf{s}_n - \mathbf{u} \| \leq \delta$, where $\| \cdot \|$ denotes a norm which can be the Euclidean norm or the maximum norm. An alarm is triggered if the precursory structure $\mathbf{s}_n$ is found within the alarm volume. A *decision variable* is defined as:

$$A(\mathbf{s}_{(n,\tau)}, \mathbf{u}, \delta) = \begin{cases} 1 & \text{if } \mathbf{s}_n \in V(\mathbf{u}, \delta), \\ 0 & \text{if otherwise.} \end{cases} \tag{1.3}$$

The precursors can be found using conditional probabilities. The use of conditional probability is linked with Markov-chain. For $t_1 \leq t_2 \leq \cdots \leq t_n$

$$\rho(x_n, t_n | x_1, t_1; \ldots; x_{n-1}, t_{n-1}) = \rho(x_n, t_n | x_{n-1}, t_{n-1}) \tag{1.4}$$

For stationary and homogeneous Markov process the conditional probability $\rho(x_n, t_n | x_{n-1}, t_{n-1})$ does not depend explicitly on time, but just on the time interval r between the successive time steps, such results in $\rho(x_n | x_1, \ldots, x_{n-1}) = \rho(x_n | x_{n-1})$. A chain of CPDFs can be constructed:

$$\rho(x_n, x_{n-1}, \ldots, x_0) = \rho(x_n | x_{n-1}) \rho(x_{n-1} | x_{n-1}) \ldots \rho(x_1 | x_0) \tag{1.5}$$

By considering a vector valued process with $x_j = \{x_n, x_{n-1}, \ldots, x_{n-\tau+1}\}$ we get

$$\rho(x_{n+1}, x_n, \ldots, x_{n-\tau+1} | x_n, x_{n-1}, \ldots, x_0) = \rho(x_{n+1}, x_n, \ldots, x_{n-\tau+1} | x_n, x_{n-1}, \ldots, x_{n-2\tau+1}) \tag{1.6}$$

such that

$$\rho(x_j|x_{j-1}, x_{j-1}, \ldots, x_0) = \rho(x_j|x_{j-1}) \quad (1.7)$$

holds. Another property, which holds for Markov processes and is of importance for forecasting is the fact that they fulfill the Chapman Kolmogrov Equation:

$$\rho(x_{n+1}|x_{n-1}) = \int \rho(x_{n+1}|x_n)\rho(x_n|x_{n-1})dx_n \quad (1.8)$$

The CPDFs that connects observations that are k time steps apart:

$$\rho(x_{n+k}|x_n) = \int \rho(x_{n+k}|x_{n+k-1})\rho(x_{n+k-1}|x_{n+k-2}) \ldots \rho(x_{n+1}|x_n)dx_{n+k-1} \ldots dx_{n+1} \quad (1.9)$$

This allows us to use lead time $k > 1$ to forecast based on $\rho(x_{n+k}|x_n)$.

1.1 Strategies to identify the Precursor

The precursory structure can be chosen using either a maximum of the *a posteriori PDF* or of the maximum of the *likelihood*.

- The a posteriori PDF $\rho(\mathbf{s}_{(n,\tau)}|X)$ takes into account all events of size X and provides the probability density to find a specific precursory structure before an observed event.

  The strategy consists in defining the precursors in a retrospective or *a posteriori* way: once the extreme event X has been identified, one asks for the signals right before it. Formally, this implies that the precursory structure consists of the global maxima in each component $(u^*_{n-\tau+1}, u^*_{n-\tau+2}, \ldots, u^*_n)$ of the a posteriori PDF.

- The likelihood $\rho(X|\mathbf{s}_{(n,\tau)})$ takes into account all possible values of precursory structures, and provides the probability density that an event of size X will follow them. The likelihood is not a density function with respect to the precursory structure, but with respect to the event size X. The precursory structure enters into the likelihood only as a parameter.

  The strategy consist in determining those values of each component x_i of the condition $\mathbf{s}_{(n,\tau)}$ for which the likelihood has a global maximum. The posteriori PDF and the likelihood are linked via Bayes's theorem.

  $$\rho(\mathbf{s}_{(n,\tau)}|X) = \rho(\mathbf{s}_{(n,\tau)})\rho(X|\mathbf{s}_{(n,\tau)}) = \rho(\mathbf{s}_{(n,\tau)}|X)\rho(X), \quad (1.10)$$

  where $\rho(\mathbf{s}_{(n,\tau)})$ represents the marginal PDF to find the precursory structure $\mathbf{s}_{(n,\tau)}$ and $\rho(X)$ represents the marginal PDF to find events of size X.