Risk & Uncertainty Assessment of a city facing rising sea levels

Exercise 1 to the lecture

**Climate Change Uncertainty and Risk:
from Probabilistic Forecasts to Economics of Climate Adaptation**

by
Prof. Reto Knutti and Dr. David Bresch
Spring Term 2016

Tutorials

Anina Gilgen (anina.gilgen@env.ethz.ch)
&
Martin Stolpe (martin.stolpe@env.ethz.ch)
&
Maria Rugenstein (maria.rugenstein@env.ethz.ch)

Introduction

The goal of this exercise is to have a first glimpse into the issue of uncertainty when stochastic processes are involved. This means that we have to take uncertainty - expressed in mathematical term as probability distributions - into account when facing a decision. The numerical implementation of uncertainty will be in MATLAB and we will guide through the programming part of this exercise. There are some voluntary questions denoted with a "*", which give you extra points. Feel free to ask us any question.

The situation is as follows. You are a citizen of Hansetown, which is known to be affected by strong wind gusts. Figure 1 below shows the observations of a **monthly storminess index (SI)** during the years 2001-2011. High winds speed can damage your building when passing a **threshold value** $31\,\mathrm{m\,s^{-1}}$, as indicated in Figure 1.

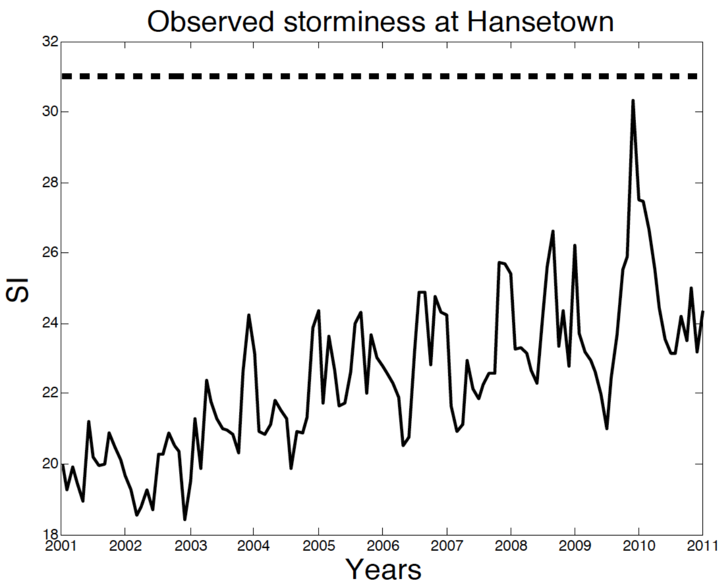

Figure 1: Observations of the storminess index in Hansetown from 2001 to 2011.

There seems to be an upward trend along with some sort of stochastic noise in the maximum wind speed. You have a specific budget which will be given to you later in the exercise. Should you save the money and get interest or should you rather invest in your building buying additional wind speed protection? In the end of the exercise, you will be provided with the real time series for the next couple of months and you will either win or loss.

In the next sections, you will stepwise learn the tools to cope with the nature of this time series and you will be able to make probabilistic forecasts of the storminess and the impact costs. The data can be downloaded from the lecture homepage. The observations are in the `observations_fs2016.mat` file.

1 Statistical modeling of past observations

In this part of the exercise, we construct a statistical model that represents the observations in Figure 1 which will allow to make a prediction of the storminess into the future.

We choose for this purpose a parametric statistical model of four parameters ($\boldsymbol{\beta_{1,2}, \alpha, \sigma}$), which will be described in the course of this section. The model consists of two parts, a deterministic linear trend and a stochastic noise term,

$$x_t = \beta_1 + \beta_2 \cdot t + \epsilon_t, \tag{1}$$

where β_1 and β_2 are the inter-section and the slope of the trend, respectively, and ϵ_t is a stochastic term representing the deviations (residuals) from the regression.

Deterministic part – the trend

One can estimate the trend of the time series by fitting a linear function through the time series. This can be achieved with the function `polyfit`. Which results in an estimate for the model parameters.

- **1a)** What is the corresponding trend estimate $\boldsymbol{\beta_{1,2}}$?
- **1b)** Is the trend significant on a $5\,\%$ level?[1]
- **1c)** What other method can be applied to estimate the trend?

The estimation of the linear trend makes assumptions on the nature of the residuals (ϵ_t), stating that the residuals should be identically normally distributed and independent (iid). Whether these assumptions are fulfilled is evaluated in the next step of the model estimation (ϵ_t).

Stochastic part – the noise

By knowing the linear term of the statistical model, the overlying noise can be analyzed. Subtracting the linear trend from the observations (function `detrend`) gives the residuals (noise) of the linear model. Have a look at the residuals (function `plot`) to get glimpse on its variability.

- **1d)** Are the assumptions for the linear regression concerning the residuals fulfilled? Discuss each assumption looking at your detrended plot. In addition, use the functions `histogram` and `qqplot` and plot the residuals against the fitted values.

[1] The function `regstats.m` gives you a whole lot of statistics for your time series.

Auto-correlation means that the data points of the time series linearly depend on previous values. Auto-correlated noise can be represented by an auto-regressive process (AR). The basic idea is that the value of the time series at time t (ϵ_t) is a linear function of previous values (ϵ_{t-1}, ϵ_{t-2}, ϵ_{t-3}, ...) plus a white noise term ($w_t \sim N(0, \sigma^2)$) which is responsible for the stochastic fluctuations. In the case of an AR(1) process, ϵ_t can be expressed as

$$\epsilon_{t+1} = \epsilon_t \cdot \alpha + w_t \ \ (AR(1)). \tag{2}$$

- **1e)** Examine the properties of an AR(p) process. What is the order of the partial auto-correlation of the noise? Use the function `parcorr` for this purpose.
- **1f)** What is the partial auto-correlation $\boldsymbol{\alpha}$ of the lag-1?
- **1g)** What could be a reason for the observed auto-correlation of the storminess index?
- **1h)** How does the autocorrelation affect the trend estimate and significance of the trend? Can you come up with a more precise estimate of $\boldsymbol{\beta_{1,2}}$ and the significance? You can use a procedure termed *recursive pre-whitening* (write you own little loop for recursive pre-whitening) or you can use *feasible generalized least squares* (function `fgls`) to estimate the regression coefficients and the AR(1) parameters simultaneously. Is the the trend from 2001-2011 still significant?

The remaining quantity we need to estimate is the noise term w_t, which we have assumed to be normally distributed with zero mean for the trend analysis and the AR(1) process.

- **1i)** Evaluate the normal distribution assumption. What changes to our model would you suggest if the normal distribution assumptions would be violated?
- **1j)** What is the standard deviation ($\boldsymbol{\sigma}$) of w_t assuming a normal distribution? Remind that $w_t = \epsilon_{t+1} - \epsilon_t \cdot \alpha$ according to equation 2.

Reproduce observations – Trend+AR(1)

All four parameters of our model are now estimated, the trend parameters $\boldsymbol{\beta_{1,2}}$, the auto-correlation $\boldsymbol{\alpha}$, and the standard deviation of the noise $\boldsymbol{\sigma}$. Using the model we can qualitatively reproduce the observations in Figure 1 by using equation 1 together with your estimates. Use the first observation as initial value for the AR(1) process. Does your model reproduce the characteristics of the observations?

2 Forecast

Using a *Monte-Carlo* computation, generate probabilistic forecasts for the years 2012 to (and including) 2018 based on the parameters you derived in the previous tasks. The principle of a *Monte-Carlo* simulation is to run the forecast a lot of times, in our case 10'000 realizations, to get a statistical distribution of an entity, in our case the storminess index, SI.

Until now, we don't have any knowledge about the likelihood of different wind speeds. A boxplot computes various quantiles of the data and displays the distribution of the data.

- **2a)** In order to have a first glimpse at the future wind speed, plot all the forecasts in one figure to get a feeling of the spread of SI.
- **2b)** Generate boxplots of the 7-year monthly forecasts for each month individually, but all 84 boxplots in one figure.

3 Forecast and "Reality" – a little excursus

Five years have passed, you are now in the year 2016, and you have observations available for the years 2011 until 2015 (Figure 2). How do the observations compare with your forecast? In the next section you will generate evidence supporting or undermining your hypothesis that the increasing trend has been attenuated.

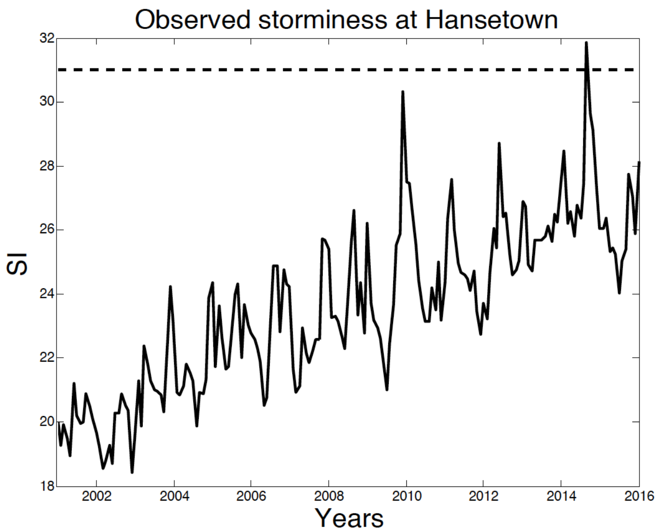

Figure 2: Observations of the storminess index in Hansetown from 2001 to 2016.

- **3a)** Download the extended observational dataset, `observations_fs2016_long.mat` from the course website and produce box plots the same way as you did in **2b**, but this time only for the years 2011-2016. In the same figure, add the observations of the year 2011 to 2016.
- **3b)** Interpret your plot. Would you say your predictions captured the observations during this time period? Are those more recent observations a result of internal vari-

ability of the system? Could it be that the trend from 2011 to 2016 is significantly lower than that of the 2001 to 2011 period?

- **3c*)** In the previous exercise, we take the observations as reference to validate our forecast. Assume that the observed SI is based on wind speed measurements from one anemometer. How could that impact your result?

4 Uncertainty and impact and mitigation costs

For the following tasks use the forecasts for the years 2016 to 2018 you generated in task **2a)**. The boxplot allowed you to have a vague idea how the storminess will change during the next couple of years. But what is the exact probability distribution for the storminess of the next months and how does this translate into the impact and mitigation costs? You will answer this question by means of the storminess forecast for the next 2 years.

- **4a)** Plot a histogram of your forecast of January 2018 and try to fit the data to a certain distribution. Hereby, you can use the functions `histogram`, `normfit`, and `gamfit`. What are the values of the distribution parameters?
- **4b)** Plot the probability of staying below the threshold value of $31\,\mathrm{m\,s^{-1}}$ as function of time for the next twenty-four months.

In the next part, we will have a look at the costs in the case of exceedance of the threshold value and the resulting damage costs on your house which in turn crucially depend on the parameters you've just derived.

4.1 Impact costs

In this part of the exercise we want to know how our SI translates into impact costs due to damages to our house. In risk and impact calculations, one often assumes an exponential relation between the damages on your object and the quantity of interest, in our case the monthly storminess index. Your house, which has a value of 700'000 $, is fully protected until the threshold value of $31\,\mathrm{m\,s^{-1}}$. In the case of a SI above $34\,\mathrm{m\,s^{-1}}$, your house is fully destroyed.

- **4c)** Given the information above, figure out the parameters and expression of the impact costs $\phi(SI)$, including some sort of exponential function of SI. There is no strict right or wrong function. Generate a function that looks meaningful to you and explain why. Plot the impact costs ϕ as a function of SI.

We are now ready to compute the distribution and hence the uncertainty of the impact costs of next months' SI. An elegant way to compute the cost as a function of SI is by defining a function in MATLAB, including the conditions if SI $\leq 31\,\mathrm{m\,s^{-1}}$ and SI $\geq 34\,\mathrm{m\,s^{-1}}$. This way you can proceed and call the function easily.

- **4d)** Considering the distribution of the February 2016 forecasts. How high are the impact costs in terms of the median, 2.5% and 97.5% quantile?
- **4e)** Plot the impact cost distribution for the mean, median and some quantile ranges of your choice for the next twenty-four months. Have also a look at the cumulative distribution of the costs as a function of forecast lead time.

It is assumed that the trend seen in the observations is likely ($\geq$66%) due to climate change (a rigorous detection and attribution framework is omitted here). One might be interested if the forecast of SI and hence its costs would have been different if climate change had not happened.

- **4f*)** Compute and plot the values derived in the previous exercises for the case of no climate change. Are there significant differences?

4.2 Impact and Mitigation Cost Curve

Lastly, we want to know how much mitigation the damage risk costs. Your local store has recognized the demand for wind-protection early and sells incremental protection, e.g. you can shield your house for an additional unit SI. Since the protection depends on the area which you want to shield, and your house doesn't change in size, the price calculation takes the following form

$$\phi^{mitigation}(\delta SI)[\$] = 25'000\$ \cdot \delta SI^2, \tag{3}$$

where δSI denotes the additional wind shield protection ranging from 0 to $5\,\mathrm{m\,s^{-1}}$ with increments of $0.1\,\mathrm{m\,s^{-1}}$. Your budget sums up to 700'000\$. So one option is to invest everything in wind protection, another would be to just invest it otherwise and hope for the best regarding storm damage to your house.

- **4g)** Compare impact and mitigation costs between 31 and $34\,\mathrm{m\,s^{-1}}$. At which level are the total costs (impact and mitigation) lowest, without using any information of your forecast?
- **4h)** Based on the your storminess forecast, how would you protect your house in order to minimize costs? How would an interest rate of $8\,\%$ affect your decision?