Probability, Statistics and Risk, MVE0

Project 11 (improvements of text needed)

In the project description we sketch the analysis of the problem we expect you to do. (Obviously you are welcome to do more.) To pass the project a short report should be written and handed in to the project supervisor. In addition the group should present their results in class. The presentation should take about 15 minutes. Please include a short introduction which will facilitate for other students to understand the results of the project. (Do not assume that the audience knows the subject.)

1 Introduction - Wind energy harvesting.

As a result of the increased concern about the drawbacks of conventional energy sources in terms of emissions of Carbon Dioxide and other Green-House Gases, renewable energy sources such as wind power attract more and more attention from governments, industry, research and the general public. Since after manufacture and installation of the wind mills does not emit any Green-House Gases, wind power is the energy source that is rapidly growing all over the world.

The available power from wind is related to the cube of wind speed w, say, viz

$$P(w) = \frac{1}{2}\rho_{air} A_r\, w^3,$$

where A_r is the area swept by the rotor blades and ρ_{air} is the air density. Obviously the power output from the turbine is smaller than the available power and it is about 20%-30% of that value. In addition wind turbines cannot produce power for all wind speeds. The range in which power is produced is limited below by the cut-in speed and above by the cut-off speed. These take values between 2-4 m/s and 25-30 m/s, respectively, depending on turbine model.

Wind energy is an inexhaustible but variable energy source. The variability is due to the seasonal, daily and regional variations in wind speed. The power output from wind turbines is therefore variable and managing its integration into the power system is a rather complicated issue for Transmission System Operators. This becomes increasingly challenging when larger parts of the power supply are covered by wind power. In order to integrate the power efficiently into the grid, accurate forecasts of the expected wind power output for the coming hours and days are necessary. This problem requires simultaneous modeling of wind speeds variability in the near future, at several locations of wind mills farms.

In this project you will address some aspects of prediction of wind speed two hours ahead at fixed location.

2 Modeling the wind speed variability

In the project wind measurements on Älvsborgsbron year 2008, available at

http://www.gvc.gu.se/Department of earth sciences/climate stations/climate-data/

will be used. The bridge is located at 57,6967 N and 11,9869 E. Average wind speeds are recorded hourly. Write few words about data. Load the data gathered every hour year 2008

```
>>load Alvsborgsbron2008.mat
```

Plot the data on Normal, Weibull probability paper (MATLAB functions normplot, weibplot) to investigate what probability distribution could be used to model variability of the data. It will be quit clear that Gaussian model is less appropriate that the Weibull is.

One way to change the distribution of the data is to transform the data. A possible transformation of wind speed data $w(t)$, say, is the following one

$$x(t) = w(t)^a - m(t).$$

For the winds measure on Älvsborgsbron the following estimates of the parameters have been found; $a^* = 0.575$ and

$$m^*(t) = 2.82 + 0.20\cos(2\pi\, t) + 0.28\sin(2\pi\, t),$$

t has units years. Now the variability of the transformed winds $x(t)$ seems to follow Gaussian distribution pretty well. Motivate this claim using the normal probability plot of $x(t)$. Plot also $x(t)$ to visually inspect whether the data has constant mean and variance over the year. The figures can be plotted using the following script

```
>> wt=Alvsborgsbron2008;
>> N=length(wt);
>> t=(1:N)'/N;
>> mt=2.82 + 0.20*cos(2*pi*t) + 0.28*sin(2*pi*t) ;
>> xt=wt.^0.575-mt;
>> figure, normplot(xt)
>> figure, plot(t,xt)
```

Estimate the mean m_X and variance σ_X^2 of $x(t)$.

Suppose that one wishes to predict the wind speed two hours ahead then the joint variability of pairs $(x(t), x(t+2))$ should be investigated. This can be done by the following plot

```
>> figure, plot(xt(1:8:end-2),xt(3:8:end),'.')
>> hold on
```

Are the pairs $(x(t), x(t+2))$ normally distributed, i.e. have pairs $(x(t), x(t+2))$ a joint pdf $f(x, y)$ similar to the one given in (5.5) in the course book, viz

$$f(x, y) = \frac{1}{2\pi\sigma_X^2} \exp\left(-\frac{1}{2\sigma_X^2(1-\rho^2)}(x^2 + y^2 - 2\rho\, x\, y) \right)?$$

A simple test could be to plot contour lines of the density on the last figure. This can be done using the following script

```
>> S=cov(xt(1:end-2),xt(3:end));
>> s2=S(1,1); ro= S(1,2)/S(1,1);
>> x= -3:0.1:3; y=x; Nx=length(x); Ny=length(y);
>> fxy=zeros(Nx,Ny);
>> for i=1:Nx
>>      for j=1:Ny
>>           fxy(i,j)= exp(-(x(i)^2+y(j)^2-2*ro*x(i)*y(j))/2/(1-ro^2)/s2)/2/pi/s2;
>>      end
>> end
>> contour(x,y,fxy,5,'linewidth',2)
```

Is the fit reasonably good? (Note that negative values of $x(t)$ corresponds to low wind speeds which are of less interest for power calculations.)

3 Predicting future wind speed

In order to predict wind speed two hours ahead the conditional density of $x(t+2)$ given all the information available at time t is needed. The density can be interpreted as the posteriori density of wind speed at time $t+2$ given the knowledge at time t. The prior knowledge is contained in the unconditional density of $x(t+2)$ which is $N(0, \sigma_X^2)$.

Here you consider the simplest case when the only information available is the wind speed at time t, i.e. $w(t)$ and hence also $x(t)$. What could be included into "available information at time t" that could improve the prediction?

Use formula (5.17) in the course book to derive the conditional density. (Hint: the derivations will result in conclusion that the conditional density is normal having the expected value $\rho\, x(t)$ and variance $(1-\rho^2)\sigma_X^2$.)

The conditional expectation $\rho\, x(t)$ is the best predictor of $x(t+2)$ given the available information at time t. Here t has units hours. The prediction error is zero mean normal with standard deviation $\sqrt{(1-\rho^2)\sigma_X^2}$. Write down a formula for the predictor of the wind speed $w(t+2)$. Present the predictor and predictions error graphically. Hint:

```
>> wt_pred=max(0,mt(3:end)+xt(1:end-2)*ro).^(1/0.575);
>> figure
>> subplot(4,1,1)
>> plot(t(3:end),wt(3:end)), hold on
>> subplot(4,1,2)
>>  plot(t(3:end),wt(3:end)-wt_pred,'r')
>> ylabel('m/s')
>> subplot(4,1,3)
>> plot(t(1102:1302),wt(1102:1302)), hold on
>>  plot(t(1102:1302),wt_pred(1100:1300),'r')
>> subplot(4,1,4)
>> plot(t(1102:1302),wt(1102:1302)-wt_pred(1100:1300))
>> ylabel('m/s')
```

Comment the results. Any ideas how to make the prediction error smaller?