Course in Probabilistic Methods in Wind Energy: Assignment 1 (class assignment).

The aim of this exercise is to compute the reliability of a wind turbine blade against fatigue failure in the root section. For that purpose we define a limit state function which characterizes the state of the blade as follows:

$$g(\mathbf{X}) = \begin{cases} < 0 & \text{failure set} \\ = 0 & \text{failure surface} \\ > 0 & \text{safe set} \end{cases}$$

The probability of failure is expressed as the *total* probability that $g(\mathbf{X}) \leq 0$. This is expressed as

$$P_f = \int \mathbb{I}_{g(\mathbf{X})<0}\, f(\mathbf{X}) d\mathbf{X}$$

where the limit state function which we will use is defined as

$$g(\mathbf{X}) = \Delta - \frac{1}{N_{short-term} \cdot k} \sum_{j=1}^{N_{short-term}} \{X_M \cdot M_{x,j}[u(X_W), \sigma_u, \alpha]\}^m \ ;$$

the "indicator function" $\mathbb{I}_{g(\mathbf{X})<0}$ is equal to 1 when $g(\mathbf{X}) < 0$, and is otherwise zero.

Variable	Explanation	Distribution	Mean	CoV
$\boldsymbol{\Delta}$	Uncertain material fatigue capacity	Lognormal	1	0.3
$\boldsymbol{X_M}$	Loads model uncertainty	Normal	1	0.2
$\boldsymbol{X_W}$	Uncertainty in wind (site) conditions	Normal	1	User-defined
$\boldsymbol{\alpha}$	Wind shear exponent	Truncated normal	0.1	$\sigma_\alpha = \min\left[1, \frac{1}{u}\right]$
$\boldsymbol{k}$	Fatigue strength normalization factor	Deterministic	TBD	-
$\boldsymbol{m}$	Fatigue S-N curve slope	Deterministic	3	-
$\boldsymbol{N_{short-term}}$	Number of short-term estimates in the lifetime sum	Deterministic	User-defined	

The vector $\mathbf{X} = (\Delta, X_M, X_W)$ combines the uncertain variables that are part of the reliability analysis. Δ represents the total accumulated fatigue damage that the structure can sustain before experiencing failure. Hence, we need to compute the total fatigue damage that the structure is expected to accumulate over its entire lifetime, and check whether this exceeds the limit (Δ). The total lifetime damage will be dependent on the distribution of the environmental variables (u, σ_u, α). The fatigue strength normalization factor k is simply an arbitrary factor which is needed in order to simplify the material strength part of the problem so we can make the exercise using loads on cross-section level rather than having detailed models of the structural component and computing stresses and strains.

Inputs

You are provided two datasets:

- "HovsoreData_Sonic_100m_2004-2013.csv" – containing about 10 years of simultaneous wind speed and turbulence measurements;
- "ML_ExampleDataSet.xlsx" that contains data of simulated 10-minute realizations of damage-equivalent loads (DEL) for the DTU 10MW reference wind turbine as function of various external conditions, including wind speed, turbulence, wind shear.

Assignment

You are required to set up a procedure to compute the probability against fatigue failure of a wind turbine blade root based on the input information given above. This involves reporting the following steps:

1. Fit a joint probability distribution of (u, σ_u) based on the received wind data set;
2. Determine the probability distribution of the wind climate uncertainty variable X_W by analysing the wind data through bootstrapping, further incorporating a 1% uncertainty for windspeed measurement and 1% for the joint-distribution fit;
3. Set up a load surrogate model M_x that maps the 10-minute DELs as function of u, σ_u and α;
4. Set up a routine to compute the total lifetime fatigue damage as function of (X_M, X_W) by integrating the short-term damage predictions from the load surrogate model over the joint distribution of external conditions;
5. Run structural reliability analysis by applying at least one reliability method and discussing or applying at least one additional method;
6. Estimate the probability of failure, and discuss which factors have the highest influence on the results.

Suggested approach – implementation hints

Fitting the environmental uncertainty variable X_W:

Did you try bootstrapping directly on the data? What did you notice? The uncertainty likely depends on the bootstrap sample size (how many points you use in each bootstrap sample). In order to determine the right bootstrap sample size, we should consider what the wind speed has certain long variation cycles (diurnal, seasonal), hence shuffling the data may not be the most realistic way to determine the uncertainty as it eliminates these long-term variations. The suggested approach is to take each bootstrap sample (or other methodology) based on samples which take a continuous piece of the data (e.g. one continuous period at a time), and the randomness is achieved by randomly shifting the starting point of the sample. What is the minimum period duration you should consider?

On obtaining the total lifetime fatigue damage:

You will need to do a numerical integration (represented in discrete form as a sum) of the short-term damage $M_{x,j}$ over the distribution of environmental conditions – including propagating the environmental uncertainty X_W. We already know a simple integration method – the crude Monte Carlo simulation. However, as we know the outcomes of the Monte Carlo simulation are also uncertain (repeated Monte Carlo simulations under the same conditions will still produce different results). This is acceptable if the reliability analysis (which is another, "outer loop" integration step in our task) is also done with a Monte Carlo simulation. However, there may be challenges if we want to apply gradient-based techniques such as FORM and if the underlying lifetime fatigue damage calculation is done with Monte Carlo. Alternatives include:

- Using Importance Sampling based on a deterministic Quasi-Monte Carlo sequence;
- Using a structured grid design with probability weighting.

Note that in order to generate samples of $M_{x,j}$ you will need to train a regression model based on the available load data, as we don't have a load simulation model available (and even if we had, it would be too computationally expensive to run so many simulations).

Report format guidelines

Please submit one report per study group, in pdf format. The report should normally be written using a word processor. Including the code in the report is in general not required – if for some reason (e.g. explaining a specific bug or a choice of function) you want to include large pieces of code, please do that in an appendix. It is also acceptable to submit printouts of Jupyter notebooks as a report – as long as they maintain the same standard of details, presentation quality and readability as a normal report. This would mean that the code should be accompanied by thorough explanations and formulas described in markdown cells, and that all unnecessary code is removed.

We recommend that the following information is included in the report:

- Short problem description (e.g. could be a reference to the assignment text or a summary).
- Solution methods – how do you approach the problem, what kind of methods and actions have you applied in order to obtain the solution.
- Results (possibly including figures, tables) showing the outcomes of your analysis
- Discussion of your results. Try to explain the observations and provide potential interpretations.

The above is not necessarily a definition of the structure of the report. The structure could be adjusted depending on your preferences, and e.g. if you have several steps which could require separate theory descriptions and separate result/discussion sections for each step.

Normally 10-15 A4 pages will be sufficient for a report with an approximately even split between figures and text and without listing of programming code.