Modellering av reparerbare systemer ved hjelp av Poisson-prosesser og deres utvidelser. En pålitelighetsanalyse av vindturbiner.

Bo Henry Lindqvist and Vaclav Slimacek

Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim

bo@math.ntnu.no

Fremtidens metoder for risikostyring,
Norsk forening for Risiko- og pålitelighetsanalyse,
Universitetet i Oslo, 3. februar 2016

Wind turbines: Smøla vindpark, Møre og Romsdal

Contents of talk

- Presentation of data
  - WMEP - The German Wind Turbine Reliability Database
- Mathematical modeling
  - Homogeneous Poisson process
- Observed covariates
  - Continuous and factor covariates
- Unobserved heterogeneity
  - Frailties
- Concluding remarks

Wind Energy

WIND ENERGY
Wind Energ. (2016)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/we.1964

RESEARCH ARTICLE

Reliability of wind turbines modeled by a Poisson process with covariates, unobserved heterogeneity and seasonality

V. Slimacek and B. H. Lindqvist
Norwegian University of Science and Technology, Trondheim, Norway

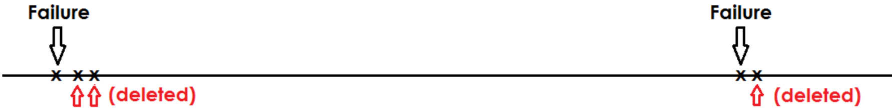
WMEP - The German Wind Turbine Reliability Database (Fraunhofer)

- German onshore turbines
- failure stops between 1989-2009
- 14853 events were considered
- 773 different wind turbines
- 402 different wind farms
- 6 different manufacturers
- 24 different rated powers

Summary statistics

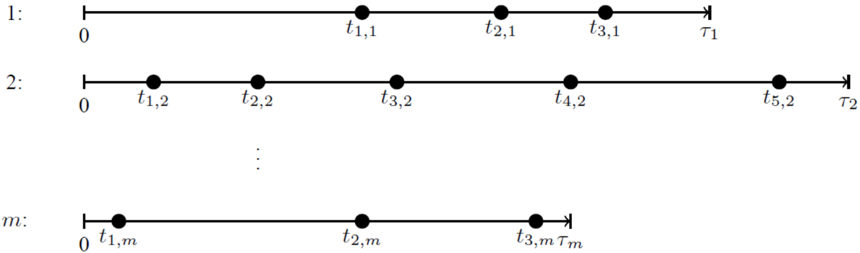
	min	1st Q	median	mean	3rd Q	max
betw fail (days)	15.00	47.00	110.00	234.62	274.00	4011.00
obs length (yrs)	4.26	10.01	10.19	10.75	10.98	17.39
turb fail per yr	0.00	0.60	0.98	1.31	1.80	7.80

- Failures which occurred within 2 weeks from the previous failure were deleted during the preprocessing of the data.
- These deleted failure times are considered to be caused by previous failures which were not properly fixed
- Overall, 3528 such failures (out of 9900), distributed over the whole observation period, were omitted.



Mathematical modeling: notation

Consider m independent repairable systems (wind turbines). The jth system is observed for a length of time τ_j, with n_j failures observed at times t_{ij} ($j = 1, \ldots, m$, $i = 1, \ldots, n_j$) measured from start at time 0.



Homogeneous Poisson Process (HPP):

- ROCOF (rate of occurrence of failures) is constant, $\lambda = a$
- Times between failures are exponentially distributed
- MTBF $= 1/a$
- Estimate of a: *Number of failures divided by total time.*

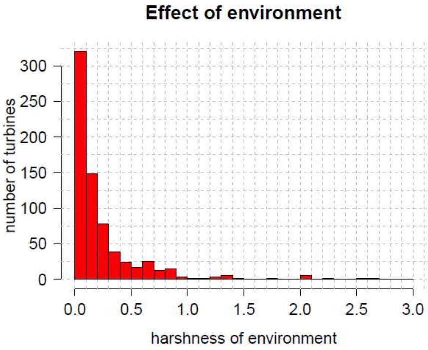
Extension 1: Observable differences through covariates

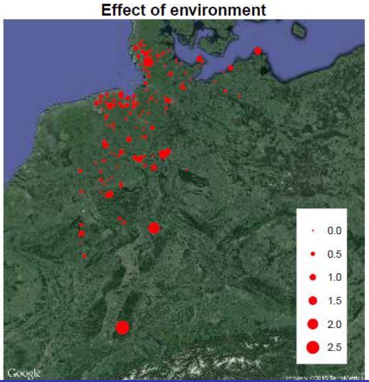
- ROCOF $= a \exp(\boldsymbol{\beta}'\mathbf{x})$, where $\mathbf{x}$ is a vector of covariate measurements (e.g., environmental variables), and $\boldsymbol{\beta}' x = \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_p x_p$.
- Thus, an increase of x_1 by 1 unit means to multiply the ROCOF by e^{β_1}, etc.

Idea: *a is a basic rate common for all systems; the covariate vector* $\mathbf{x}$ *differs from system to system; the β_j describe the effect of each single covariate x_j.*

Observed covariates: Harshness of environment

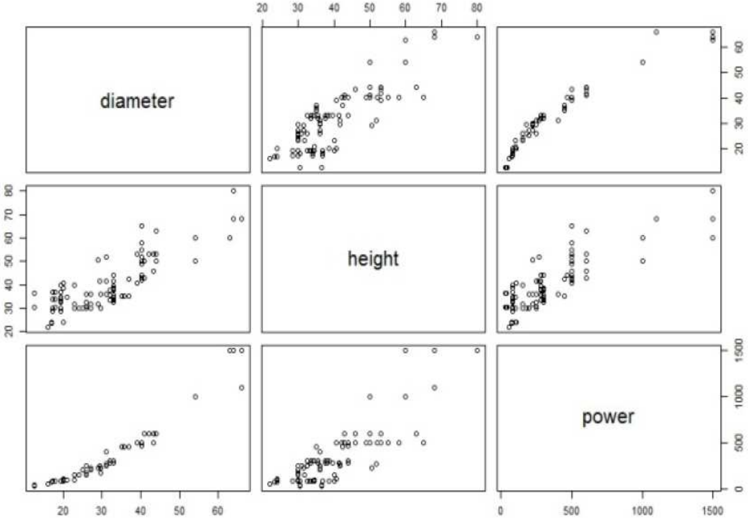
- For each turbine is calculated
  x = average number of stops because of external natural factors (lightning, high wind, icing) per year.
  (Note: those stops were not considered as failures)
- histogram of values on the left
- spatial distribution on the right



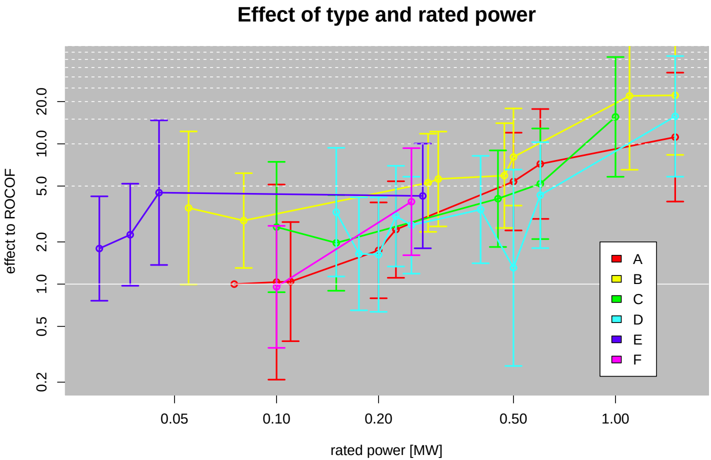


Observed covariates: Size and power

Rated power, rotor diameter and nacelle height are strongly correlated:



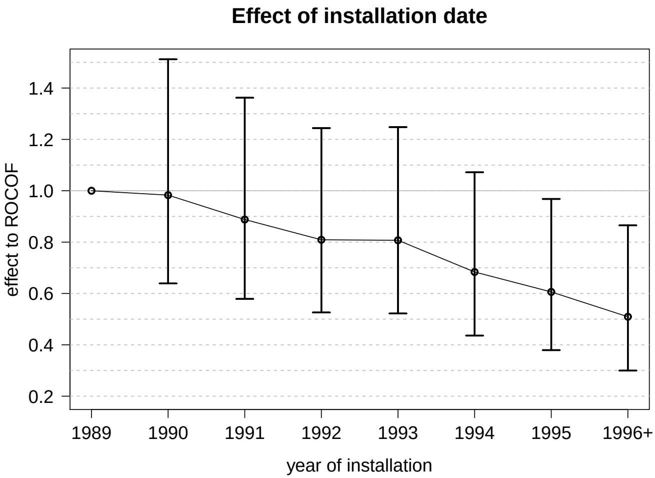
This is a *factor covariate* with 36 combinations, obtained by combining **rated power** and **manufacturer** and hence also covering differences between **technical concepts**. Reference type (value 1) is red circle to the left.



Explanation: Factor covariates

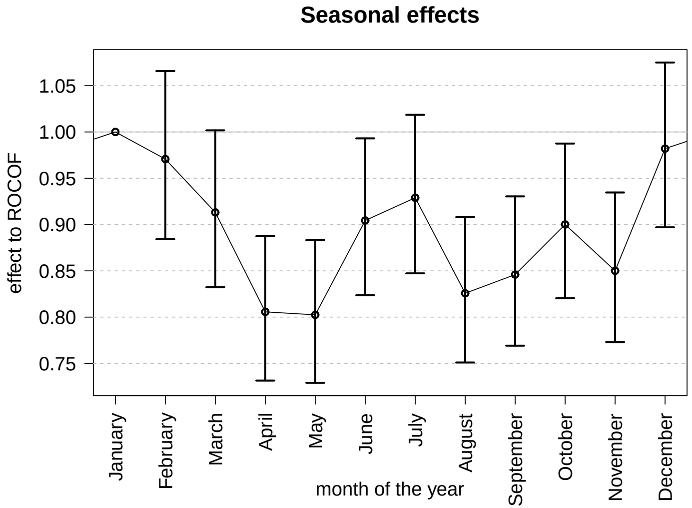
- A **factor covariate** is a categorical variable that can take a number of values or levels, where we estimate the multiplicative effect (e^{β}) of each level with respect to a reference level.
- Covariates are otherwise **continuous measurements** x, where the effect on the hazard is of the multiplicative form $e^{\beta x}$ where x varies.

Factor covariate: installation year



Relative effect on ROCOF of year of installation (together with 95% confidence intervals) with respect to the reference year 1989.

Seasonal effect

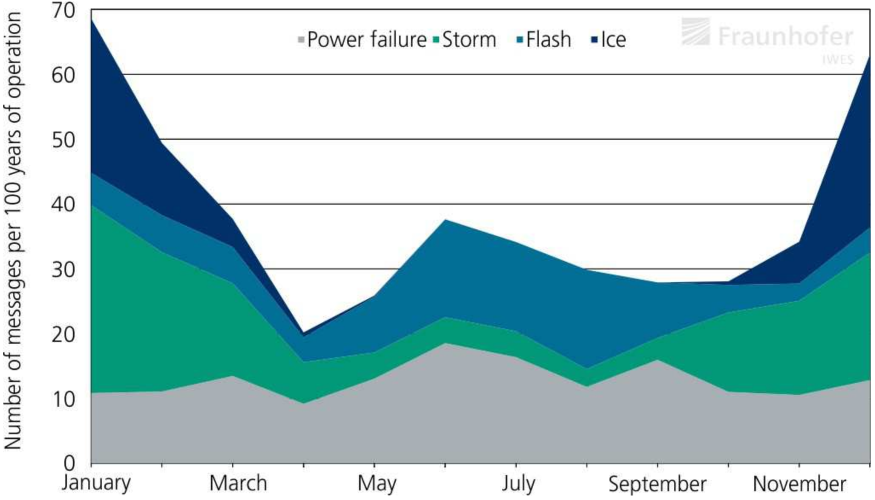


Relative effect of season (together with 95% confidence intervals) with respect to the reference month January.

Extension 2: Modeling of seasonal effect

- Let, e.g., January be reference month and estimate multiplicative change in ROCOF for each month relative to January.
- Techincally, in the model we multiply the ROCOF by a factor $\exp\left(\sum_{s=2}^{12} \delta_s I_s(t)\right)$, where $I_s(t)$ is an indicator function of the sth month, i.e., equals 1 if the calendar time corresponding to t is in the sth month and 0 otherwise.
- Then e^{δ_s}, $s = 2, \ldots, 12$ represent the multiplicative change of the ROCOF for each month.
- Remark that this makes the Poisson process *nonhomogeneous* (NHPP) with a ROCOF cyclically changing over the year.

Confirmation of results on climatic causes for failures...



Frequency of external conditions as a cause of failure. Source: WMEP 2006

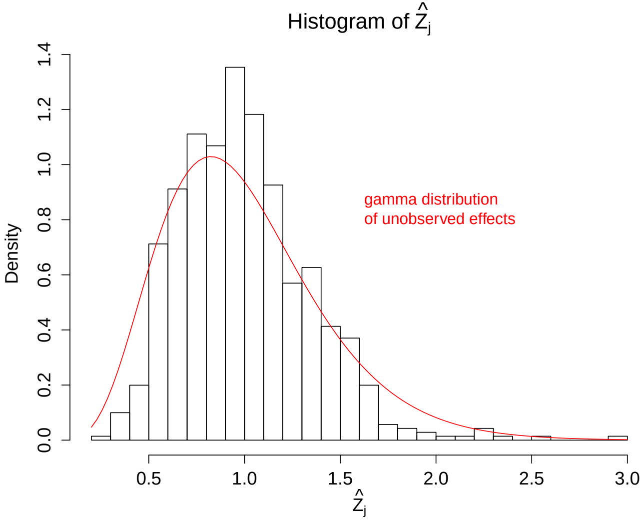
Extension 3: Unobserved heterogeneity

Covariates represent observable differences between systems. *There may also be* **unobservable** *differences* (but why should we care about those...?)

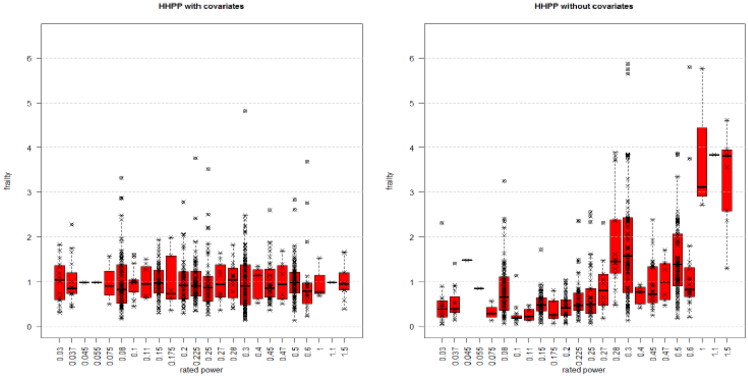
- Multiply ROCOF by a "modifying factor" z, averaging to 1, leading to ROCOF $= z\, a \exp(\beta' \mathbf{x})$
- z represents the so called *frailty* of the corresponding system, which is *unobservable* (while $\mathbf{x}$ as before is observable).
- the frailty is assumed to vary independently from system to system (usually modeled by a gamma-distribution with expected value 1 and a variance α that can be estimated).
- individual frailties z may be estimated from data

Idea: *Systems of the same kind may well exhibit different failure patterns because of external unobserved sources such as differing environmental conditions, differing maintenance philosophies, or differing quality of operators/equipment.*

Individual estimated frailties



Can frailties replace covariates?



Comparison of estimates of individual frailties grouped by rated power and represented by box plots for model with and without covariates

Some mathematics

Final model for ROCOF of jth system:

$$\lambda_j(t) = z_j a \exp\left(\beta' \mathbf{x}_j\right) \exp\left(\sum_{s=1}^{12} \delta_s I_s(t)\right)$$

Likelihood for jth system conditional on the value of z_j:

$$L_j(z_j) = \prod_{i=1}^{n_j} \lambda_j(t_{ij}) \exp\left(-\int_0^{\tau_j} \lambda_j(u) du\right)$$

Likelihood for jth system after averaging over the distribution of z_j using a gamma-distributon with expected value 1 and variance α:

$$L_j = \frac{\Gamma\left(n_j + \frac{1}{\alpha}\right)}{\alpha^{\frac{1}{\alpha}} \Gamma\left(\frac{1}{\alpha}\right)} \frac{a^{n_j} \exp\left(n_j \beta' \mathbf{x}_j\right) \exp\left(\sum_{s=1}^{12} \delta_s u_{s,j}\right)}{\left(a \exp\left(\beta' \mathbf{x}_j\right) \sum_{s=1}^{12} \exp\left(\delta_s\right) v_{s,j} + \frac{1}{\alpha}\right)^{n_j + \frac{1}{\alpha}}}$$

where $u_{s,j}$ is the number of failures experienced by system j in month s and $v_{s,j}$ is the total time spent in month s.

Some mathematics (cont.)

The total likelihood of the data is

$$L = \prod_{j=1}^{m} L_j$$

which after maximization gives the maximum likelihood estimates of the unknown parameters a, α, and the coefficients β_i and δ_s.

Standard errors and confidence intervals of the estimates can be computed using standard maximum likelihood theory.

The unobserved individual frailties z_j can finally be estimated with the use of an empirical Bayes approach, giving

$$\hat{z}_j = \frac{n_j + \frac{1}{\hat{\alpha}}}{\left(\hat{a}\exp\left(\hat{\beta}'\mathbf{x}_j\right)\sum_{s=1}^{12}\exp\left(\hat{\delta}_s\right)v_{s,j} + \frac{1}{\hat{\alpha}}\right)^{n_j+\frac{1}{\hat{\alpha}}}}$$

Concluding remarks

Starting from a homogeneous Poisson process we have introduced

- Effect of covariates
  - Harshness of environment - turns out to have a significant effect on ROCOF
  - Factor covariate for rated power and manufacturer - significant differences found; ROCOF increases with rated power
  - Factor covariate for installation year - strongly decreasing ROCOF is interepreted as improved technology
- Effect of season
  - Modifying model for ROCOF with respect to monthly changes from January
- Unobserved heterogeneity
  - Introduction of frailties improves model fit and interpretative power
  - Individual frailties can be viewed as representing effect of missing covariates

Solnedgang i Smøla vindpark

Thank you!