An Improved Bayesian Update Tool for Components Failure Rates

Ali Ayoub

PhD Student, ETH Zurich
Former Intern in PSA Team at Leibstadt NPP

Valerio Ariu

PSA Analyst, Leibstadt NPP

17.09.2018

PSAM14

Kernkraftwerk Leibstadt AG
CH-5325 Leibstadt | Telefon +41(0)56 267 71 11 | www.kkl.ch

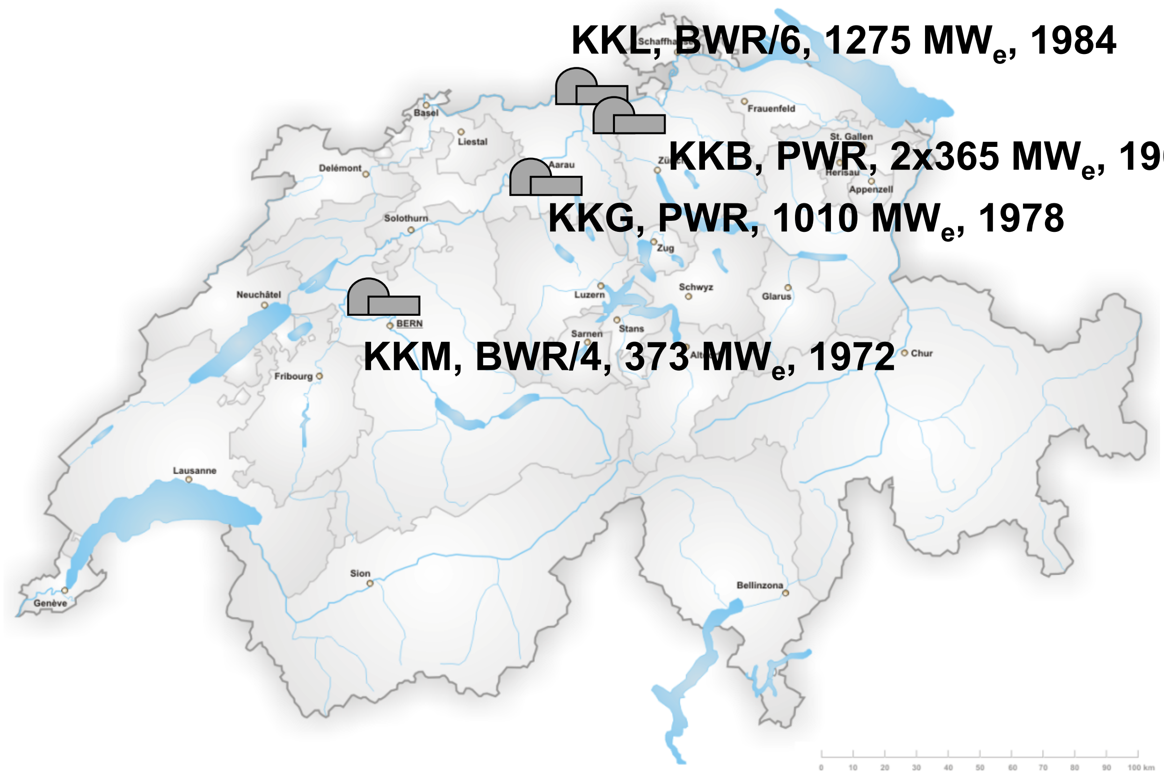
Kernkraftwerk Leibstadt

An Improved Bayesian Update Tool for Components Failure Rates

Content

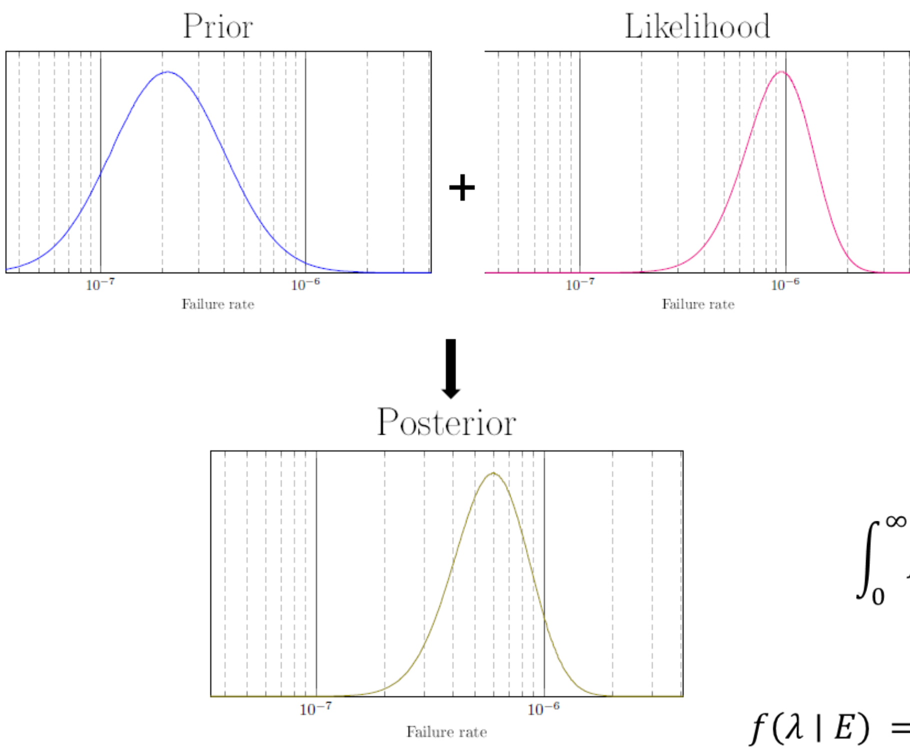
Kernkraftwerk Leibstadt

Introduction and guidelines



Kernkraftwerk Leibstadt

Introduction and formulation



$f(\lambda)$ prior probability density function (international data) (PDF)

$L(E \mid \lambda)$ likelihood function (pdf of the number of failures)

$f(\lambda)\, L(E \mid \lambda)$ describes the shape of the Posterior distribution

$\int_0^{\infty} f(\lambda')L(E \mid \lambda')d\lambda'$ is a normalization factor

$$f(\lambda \mid E) = \frac{f(\lambda)\, L(E \mid \lambda)}{\int_0^{\infty} f(\lambda')L(E \mid \lambda')d\lambda'}$$ Posterior distribution

Introduction and formulation

What's the need for a Bayesian update in nuclear PSA?

Swiss regulatory guideline ENSI-A05 states:

> "The plant-specific reliability parameters shall be derived for each component group by **combining the collected plant-specific raw data with the generic reliability** data through a **Bayesian** update process."
>
> "The mean failure probability and a statistical representation of the **associated uncertainty** (5th, 50th, 95th percentile) shall be provided for each reliability parameter.
>
> The uncertainty distribution resulting from the Bayesian update shall be directly used or **mapped** by an appropriate distribution (e.g., Beta or Gamma distribution)." (*Re-casting*)

Previous Bayesian update tool implemented at KKL

Main features:

$$f(\lambda \mid E) = \frac{f(\lambda)\, L(E \mid \lambda)}{\int_0^\infty f(\lambda')L(E \mid \lambda')d\lambda'}$$

- Programmed in script language Ruby
- The discretization (to calculate the integral) used was solely based on the prior information
- Integral was always resolved numerically
- Used very large number of integration points

→ Not always efficient
→ In some cases was unable to correctly capture the mean value and percentiles
→ In some cases lead to curtailment of distributions (therefore to optimistic failure rates)

Previous Bayesian update tool implemented at KKL

Main drawback

- discretization was solely based on prior distribution

Example for curtailment due to discretization based on prior distribution

Prior distribution type: Lognormal

Prior distribution Mean: 6.19E-7 [1/h]

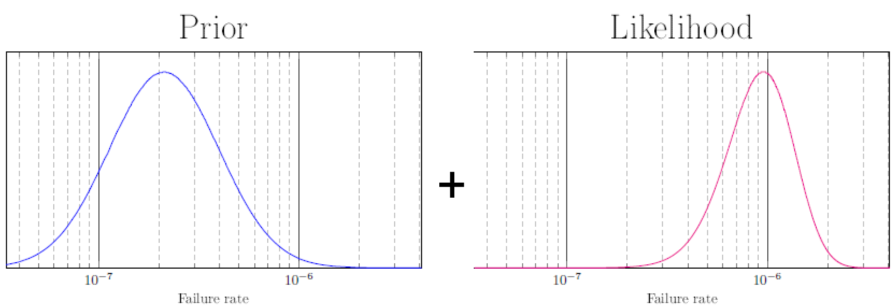
Prior distribution Error Factor: 3.57

Evidence Number of failures: 8

Evidence Exposure time: 2.52E5 [h]

Obtained Posterior Mean = 7.04E-6 [1/h]

Expected result ~ 8/2.52E5 = 3.17E-5 [1/h]



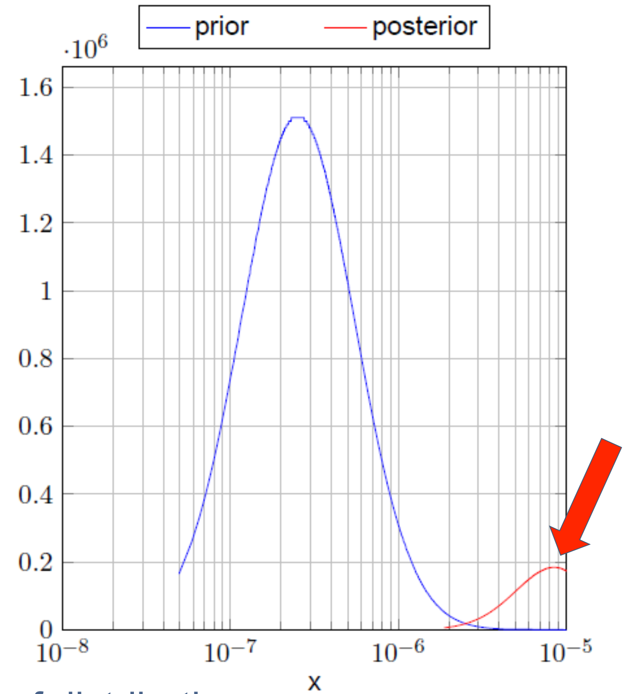
Prior and Likelihood distributions are “not in-phase

Previous Bayesian update tool implemented at KKL

Main drawback

- discretization was solely based on prior distribution

Example for curtailment due to discretization based on prior distribution



Prior distribution type: Lognormal

Prior distribution Mean: 6.19E-7 [1/h]

Prior distribution Error Factor: 3.57

Evidence Number of failures: 8

Evidence Exposure time: 2.52E5 [h]

Obtained Posterior Mean = 7.04E-6 [1/h]

Expected result ~ 8/2.52E5 = 3.17E-5 [1/h]

Curtailment of distribution:

Loss of information of the posterior distribution

→ Underestimation of the failure rates

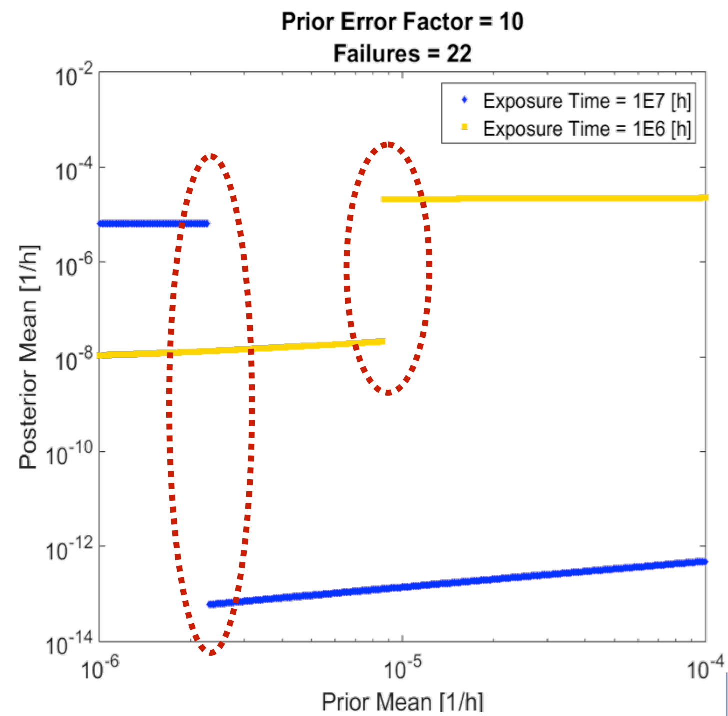
Previous Bayesian update tool implemented at KKL

Is there an issue with the integration algorithm? Test against MATLAB

Torture test (Mean vs Mean):

Each point represents the Bayesian updated posterior mean as function of the prior mean, for different exposure times and constant number of failures.

MATLAB built-in integration function shows **instabilities and divergences**



Improved Bayesian update algorithm

Investigated combinations of prior distributions (available in RiskSpectrum) and likelihood functions commonly used in PSA

	Case	Prior	Likelihood	Update Method
Non-Conjugables	1	Lognormal	Binomial	Numerical Integration (modal method)
	2	Lognormal	Poisson	Numerical Integration (modal method)
	3	Normal	Binomial	Numerical Integration (modal method)
	4	Normal	Poisson	Numerical Integration (modal method)
	5	Uniform	Binomial	Analytical Derivation
	6	Uniform	Poisson	Analytical Derivation
Conjugables	7	Gamma	Binomial	Transformation + Conjugation
	8	Gamma	Poisson	Conjugation
	9	Beta	Binomial	Conjugation
	10	Beta	Poisson	Transformation + Conjugation
Discrete	11	Discrete*	Binomial	Analytical Derivation
	12	Discrete*	Poisson	Analytical Derivation

* Adopts the meaning of RiskSpectrum Discrete distribution, i.e. a piecewise constant (uniform) distribution

Improved Bayesian update algorithm
Non-conjugable distributions

The Modal Method

Based on 4 variables: prior mean, prior variance, # failures, exposure time (or number of trials).

Discretization after foreseeing the posterior distribution

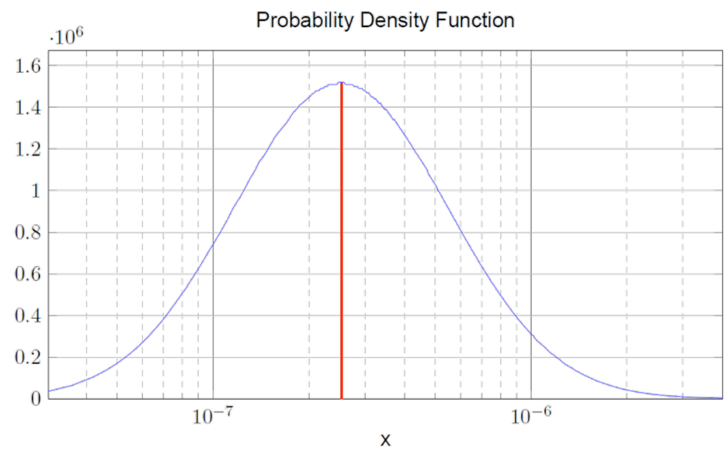
1. **Solve for the mode of the posterior:**

$$\frac{d}{d\lambda}\left(f(\lambda \mid k)\right) = 0$$

$$\frac{d}{d\lambda}\left(\frac{f(\lambda)L(k \mid \lambda)}{\int_0^\infty f(\lambda')L(k \mid \lambda')d\lambda'}\right) = 0$$

$$\frac{d}{d\lambda}\left(f(\lambda)L(k \mid \lambda)\right) = 0$$

$$L(k \mid \lambda)\frac{df(\lambda)}{d\lambda} + f(\lambda)\frac{dL(k \mid \lambda)}{d\lambda} = 0$$

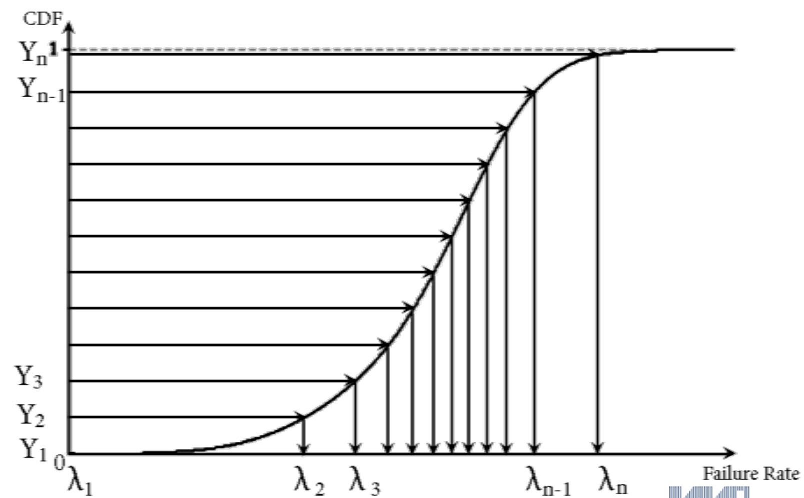


Improved Bayesian update algorithm
Non-conjugable distributions

2. Construct a *pseudo* posterior assuming
- the posterior will have the shape of the prior, i.e. same distribution type
- the variance of the posterior is equal to the variance of the prior (conservative)

Example:
A Lognormal distribution
can be defined by
its Mode (Step 1)
its Variance (Asm. 2)

3. Define the «smart» discretization points
- discretization points based on the inverse CDF of the pseudo posterior distribution



4. Obtain the *real* posterior distribution
- using the «smart» discretization points (integration)

5. Obtain the posterior distribution properties
- mean, p05, p50, p95 (integration)

Improved Bayesian update algorithm

Conjugable distributions (analytical solution)

The posterior distribution can be obtained analytically from the prior data distribution and the plant specific data, without the need to perform a numerical integration.

Prior Distribution (λ)	Likelihood Distribution $(k \mid \lambda)$	Exact posterior Distribution $(\lambda \mid k)$	Posterior mean $E[\lambda \mid k]$	Posterior variance $Var[\lambda \mid k]$	Posterior percentiles p_n
$Gamma(\alpha, \beta)$	$Binomial(n, \lambda)$	$Gamma(\alpha + k, \beta + n)$	$\frac{\alpha + k}{\beta + n}$	$\frac{\alpha + k}{(\beta + n)^2}$	$Gamma$ CDF
$Gamma(\alpha, \beta)$	$Poisson(\lambda T)$	$Gamma(\alpha + k, \beta + T)$	$\frac{\alpha + k}{\beta + T}$	$\frac{\alpha + k}{(\beta + T)^2}$	$Gamma$ CDF Inverse
$Beta(\alpha, \beta)$	$Binomial(n, \lambda)$	$Beta(\alpha + k, \beta + n - k)$	$\frac{\alpha + k}{\alpha + \beta + n}$	$\frac{(\alpha + k)(\beta + n - k)}{(\alpha + \beta + n)^2(\alpha + \beta + n + 1)}$	$Beta$ CDF Inverse
$Beta(\alpha, \beta)$	$Poisson(\lambda T)$	$Beta(\alpha + k, \beta + T - k)$	$\frac{\alpha + k}{\alpha + \beta + T}$	$\frac{(\alpha + k)(\beta + T - k)}{(\alpha + \beta + T)^2(\alpha + \beta + T + 1)}$	$Beta$ CDF Inverse

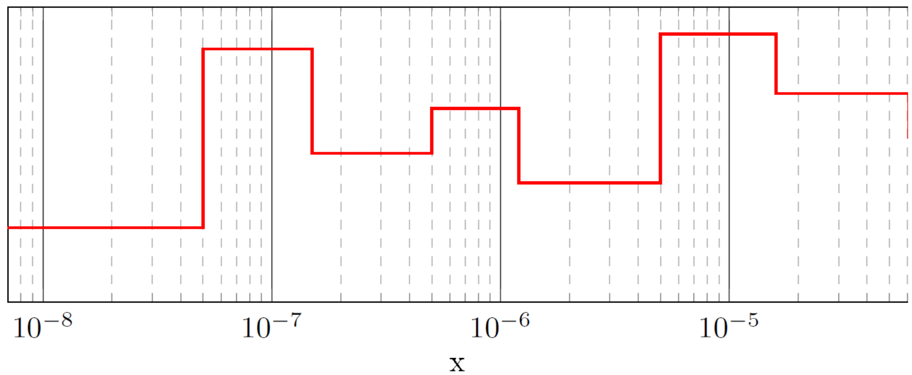
α, β: parameters of Gamma and Beta distribution
k: # of observed failures

Kernkraftwerk Leibstadt

Improved Bayesian update algorithm
Discrete distribution

The Discrete distribution is defined (in RiskSpectrum) as a *piecewise constant distribution*

- Used to model fragilites
- Can mimic any distribution
- Bayesian update performed analytically



PDF of a discrete distribution

→ Can be used to

- Approximate any distribution
- Benchmark the results of the Bayesian update obtained with the non-conjugables/conjugables methods

Results: Improved Bayesian update algorithm
Non-conjugable distributions

Example (same as before)

Prior distribution type: Lognormal

Prior distribution Mean: 6.19E-7 [1/h]

Prior distribution Error Factor: 3.57

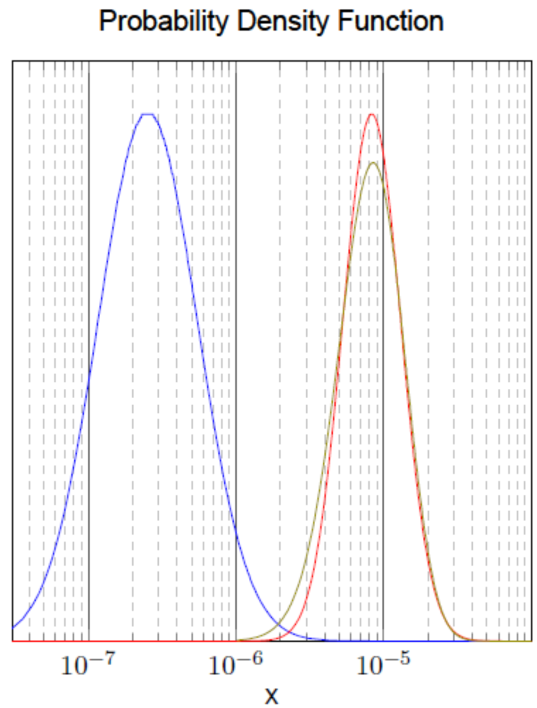
Evidence Number of failures: 8

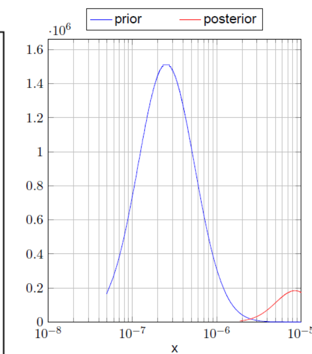
Evidence Exposure time: 2.52E5 [h]

Expected result ~ 8/2.52E5 = 3.17E-5 [1/h]

Old algorithm Posterior Mean: **7.04E-6 [1/h]**

New algorithm Posterior Mean: ***1.13E-5 [1/h]***





→ **Robust, fast, and efficient Bayesian update algorithm**

Results: Improved Bayesian update algorithm
Conjugable distributions (analytical solution)

Example

Prior distribution type: Gamma

Prior distribution Mean: 3.10E-6 [1/h]

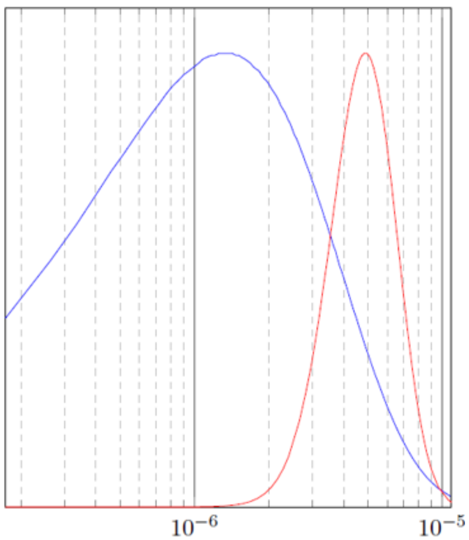
Prior distribution α = 3.1E-6

Prior distribution β = 1.75

Evidence Number of failures: 10

Evidence Exposure time: 1.63E6 [h]

Expected result ~ 10/1.63E6 = 6.1E-6 [1/h]



	Prior	Posterior
Mean	3.10E-06	5.35E-06
Alpha	1.75E+00	1.18E+01
Beta	5.65E+05	2.19E+06
5 %-percentile	4.60E-07	3.07E-06
50 %-percentile	2.53E-06	5.20E-06
95 %-percentile	7.67E-06	8.16E-06

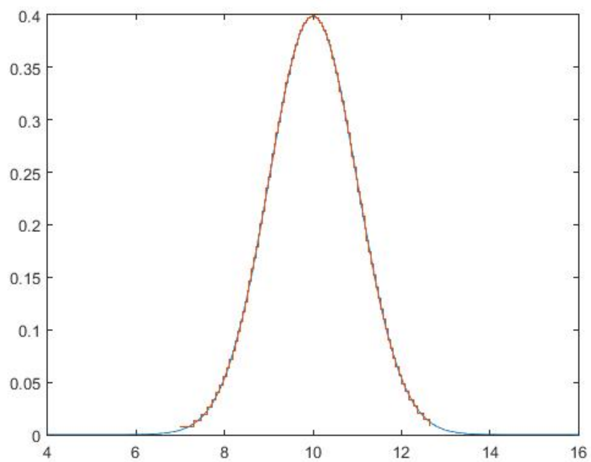
New algorithm Posterior mean: 5.35E-6 [1/h]

→ **Robust, fast, and efficient Bayesian update algorithm**

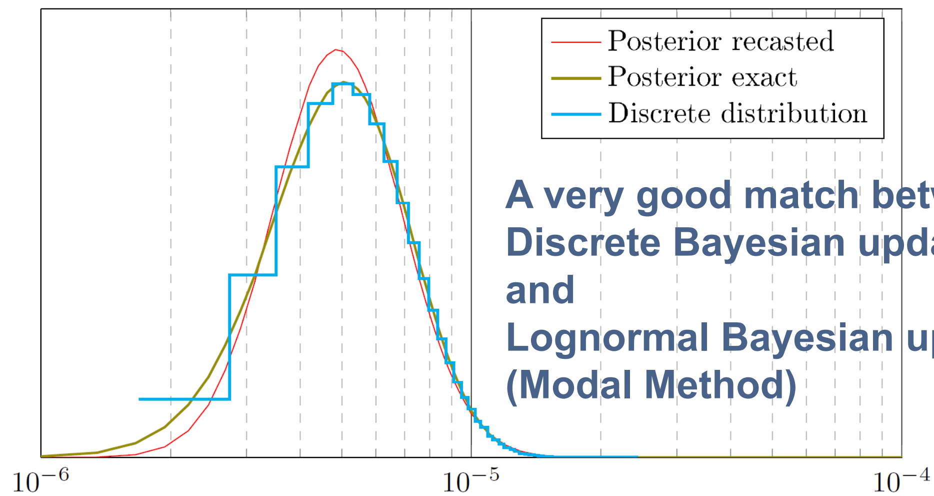
Results: Improved Bayesian update algorithm
Discrete distribution

The Discrete distribution is defined (in RiskSpectrum) as a *piecewise constant distribution*

- Can mimic any distribution (by defining a reasonable number of discretization points)
- Comparison of a discrete distribution Bayesian Update with a Lognormal distribution



Prior distribution

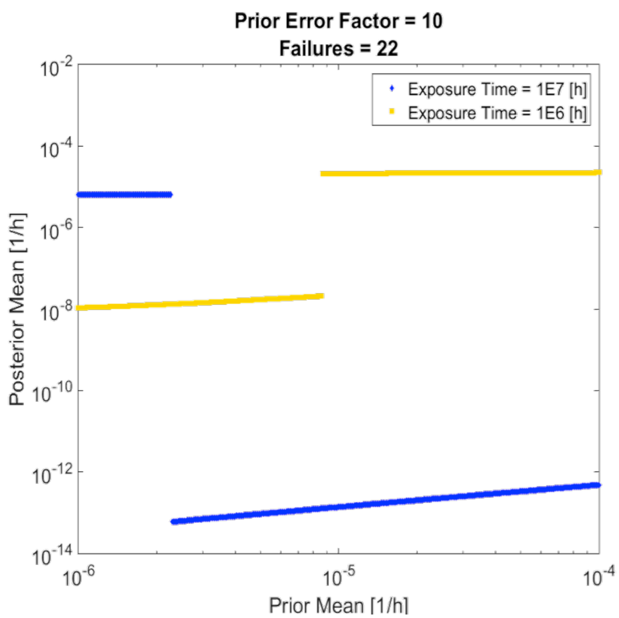


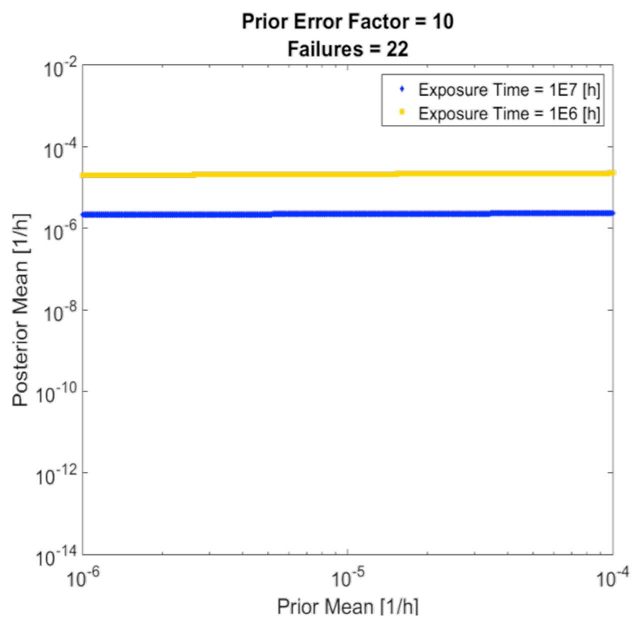
A very good match between Discrete Bayesian update and Lognormal Bayesian update (Modal Method)

Posterior distribution

Results: Improved Bayesian update algorithm
Stability check (against established numerical codes)

Stability check of the implemented algorithm (right) compared to MATLAB (left)





The developed tool shows a **very robust and stable behavior** thanks to the **smart adaptive discretization** algorithm

Kernkraftwerk Leibstadt

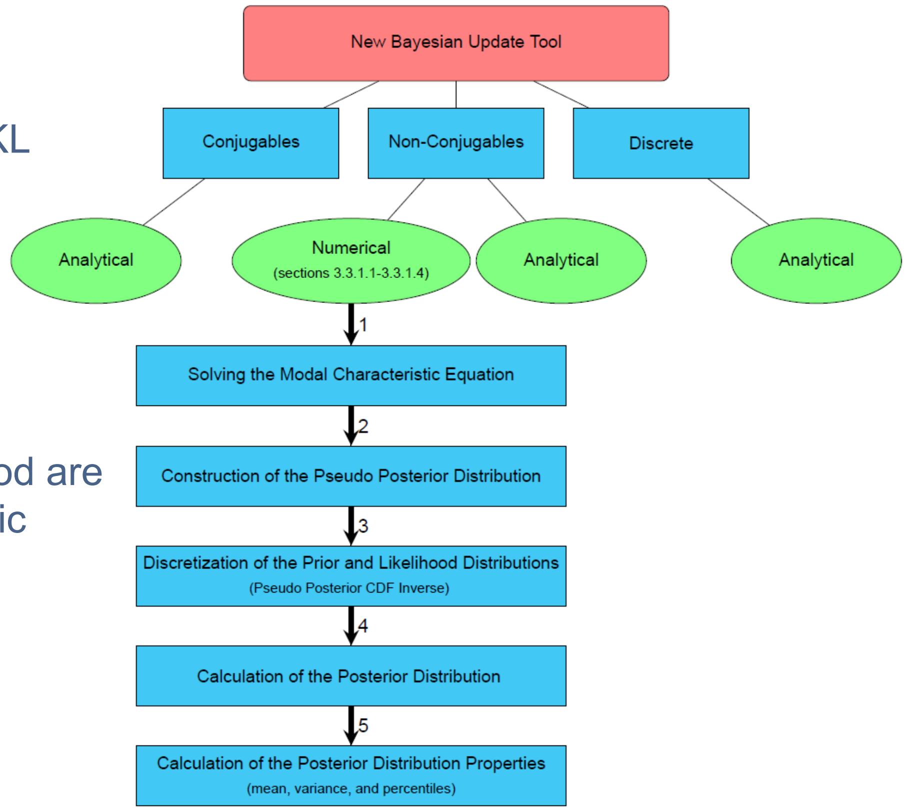
Conclusion

A new Bayesian update tool was developed at KKL

12 combinations of distributions and likelihood functions were investigated and implemented in script language Ruby

12 combinations of prior distributions and likelihood are split into 3 categories. For each category a specific Bayesian update method was developed.

The method is **fast, reliable and robust**, even against well established numerical codes and torture testing



Kernkraftwerk Leibstadt

Thank you for your attention!

Ali Ayoub
PhD Student, ETH Zurich

Valerio Ariu
PSA Analyst, Leibstadt NPP

valerio.ariu@kkl.ch
+41 56 268 4076

17.09.2018
PSAM14

Backup 1
Improved Bayesian update algorithm
Non-conjugable distributions
Re-casting (mapping)

- The Bayesian update results in a posterior distribution having some irregular shape (distribution-free)
- Fit the posterior distribution into a known parametric distribution so that it becomes **easy to handle** in RiskSpectrum (requested by the regulator)

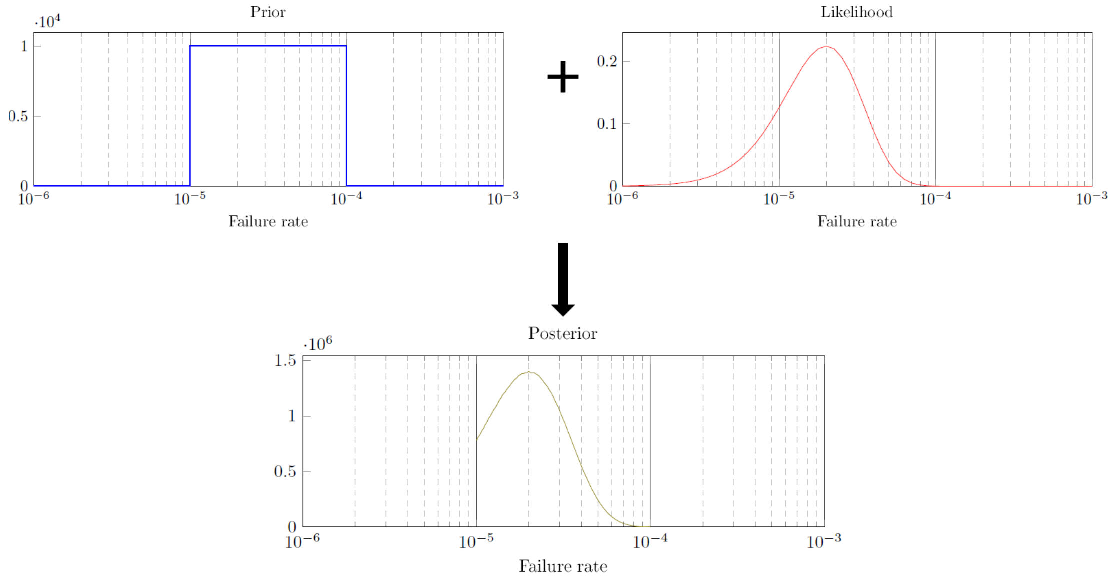
e.g Re-casting preserving the mean and median (or p95)

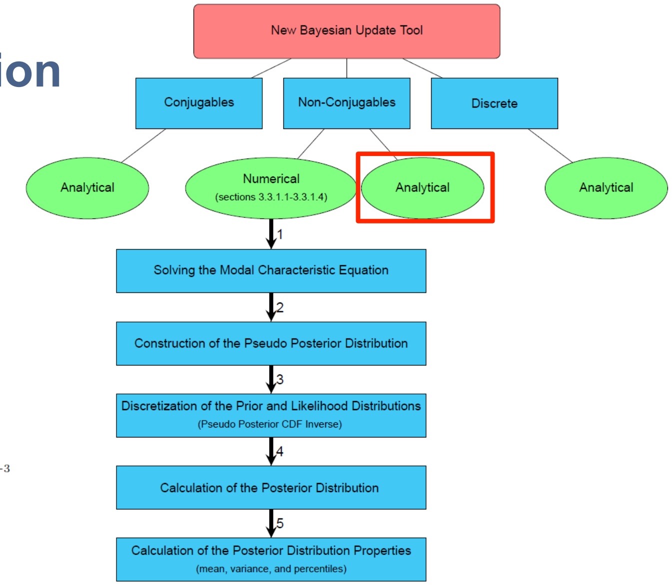
$$\tilde{\mu}_{posterior} = \ln\left(p_{50,\,\text{exact}}\right)$$
$$\tilde{\sigma}_{posterior} = \sqrt{2\ln\left(mean_{posterior}\right) - 2\tilde{\mu}_{posterior}}$$
$$\widetilde{EF}_{posterior} = e^{1.64485\cdot\tilde{\sigma}_{posterior}}$$

Kernkraftwerk Leibstadt

Backup 2
Improved Bayesian update algorithm
Non-conjugables distributions, analytical solution

Uniform prior distribution, Poission or Binomial likelihood





Kernkraftwerk Leibstadt

Backup 3
Improved Bayesian update algorithm
Non-conjugables distributions, analytical solution

Uniform prior distribution, Poission or Binomial likelihood

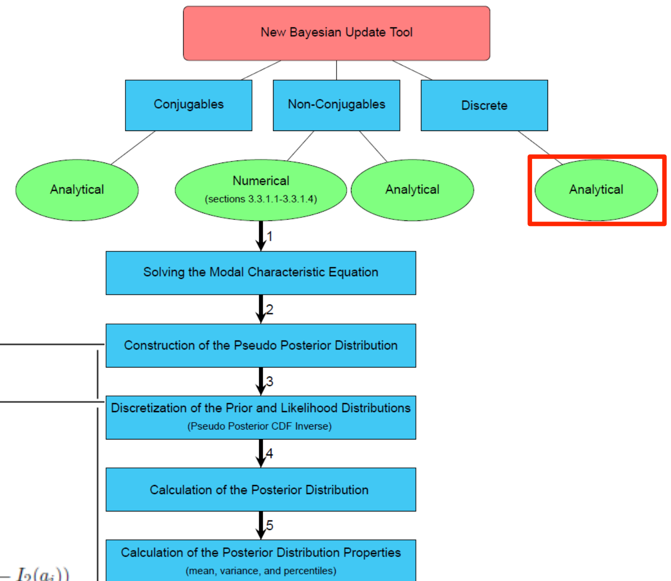
Fully analytical Bayesian update

Tab. 5-2: Summary table of the analytical non-conjugable combinations Bayesian updates

Prior Distribution (λ)	Likelihood Distribution $(k \mid \lambda)$	Exact posterior Distribution $(\lambda \mid k)$	Posterior mean $E[\lambda \mid k]$	Posterior variance $Var[\lambda \mid k]$	Posterior percentiles p_n
$Uniform(a,b)$	$Binomial(n,\lambda)$	$\frac{\binom{n}{k} \cdot \lambda^k (1-\lambda)^{n-k}}{\frac{1}{n+1} \cdot (I_1(b) - I_1(a))}^{*}$	$\frac{k+1}{n+2} \cdot \frac{(I_2(b) - I_2(a))}{(I_1(b) - I_1(a))}^{*}$	$\frac{(k+2)(k+1)}{(n+3)(n+2)} \cdot \frac{(I_3(b) - I_3(a))}{(I_1(b) - I_1(a))} - (E[\lambda \mid k])^{2*}$	$I_1^{-1}(n\% \cdot I_1(b) + (1-n\%) \cdot I_1(a))^{*}$
$Uniform(a,b)$	$Poisson(\lambda T)$	$\frac{\frac{(\lambda T)^k}{k!} \cdot e^{-\lambda T}}{\frac{1}{T} \cdot (\gamma_1(bT) - \gamma_1(aT))}^{**}$	$\frac{(k+1)}{T} \cdot \frac{(\gamma_2(bT) - \gamma_2(aT))}{(\gamma_1(bT) - \gamma_1(aT))}^{**}$	$\frac{(k+2)(k+1)}{T^2} \cdot \frac{(\gamma_3(bT) - \gamma_3(aT))}{(\gamma_1(bT) - \gamma_1(aT))} - (E[\lambda \mid k])^{2**}$	$\frac{\gamma_1^{-1}(n\% \cdot \gamma_1(bT) + (1-n\%) \cdot \gamma_1(aT))}{T}^{**}$

Kernkraftwerk Leibstadt

Backup 4
Improved Bayesian update algorithm
Discrete distribution, analytical solution



Discrete prior distribution, Poission or Binomial likelihood

Fully analytical Bayesian update

Prior Distribution (λ)	Likelihood Distribution $(k \mid \lambda)$	Exact posterior Distribution $(\lambda \mid k)$	Posterior mean $E[\lambda \mid k]$
Discrete	$Binomial(n, \lambda)$	$\begin{cases} \dfrac{pdf_0 \cdot \binom{n}{k}\lambda^k(1-\lambda)^{n-k}}{\sum_{i=0}^{m-1} \frac{pdf_i}{n+1}(I_1(a_{i+1})-I_1(a_i))} & \text{if } a_0 \leq \lambda \leq a_1 \\ \dfrac{pdf_1 \cdot \binom{n}{k}\lambda^k(1-\lambda)^{n-k}}{\sum_{i=0}^{m-1} \frac{pdf_i}{n+1}(I_1(a_{i+1})-I_1(a_i))} & \text{if } a_1 \leq \lambda \leq a_2 \\ \vdots & \vdots \\ \dfrac{pdf_{m-1} \cdot \binom{n}{k}\lambda^k(1-\lambda)^{n-k}}{\sum_{i=0}^{m-1} \frac{pdf_i}{n+1}(I_1(a_{i+1})-I_1(a_i))} & \text{if } a_{m-1} \leq \lambda \leq a_m \\ 0 & \text{otherwise} \end{cases}$	$\dfrac{(k+1)}{(n+2)} \cdot \dfrac{\sum_{i=0}^{m-1} pdf_i \cdot (I_2(a_{i+1})-I_2(a_i))}{\sum_{i=0}^{m-1} pdf_i \cdot (I_1(a_{i+1})-I_1(a_i))}$