



Bayesian inference to study low-level radioactivity in the environment: Application to the detection of xenon isotopes of interest for the CTBTO

I. Rivals¹ , C. Fabbri¹, G. Euvrard¹ & X. Blanchard²

¹ Équipe de Statistique Appliquée, ESPCI ParisTech, France

² CEA, DAM, DIF, France

S&T2011, Vienna - Austria,
June 8-10, 2011

Introduction

- The **decision** that a given detection level corresponds to the effective presence of a radionuclide, or to its absence, is **widely made using a classic hypothesis test (Currie)**.
- **Shortcomings** of the classic framework / **promising perspectives of Bayesian statistics** (part of recent standards [ISO 11929-7, 2005])

⇒ we propose a **novel Bayesian approach** providing estimates of:

- the **probability of zero radioactivity**, together with **physically meaningful point and interval estimates**,
- the **prior density of the radioactivity**, obtained by **fitting previously recorded radioactivity data**.

The classic framework [Currie, 1968]

True net radioactivity level μ ?

Observed net count : $x = x_g - x_b$

with blank count $x_b \rightarrow \text{Poisson}(\mu_b)$ and gross count $x_g \rightarrow \text{Poisson}(\mu + \mu_b)$

Independence $\Rightarrow E(x) = \mu$, $\text{var}(x) = \sigma^2 = \mu + 2\mu_b$

Large enough counting time $\Rightarrow \approx$ Gaussian CDF: $F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$

Test of $H_0: \mu = 0$ against $H_1: \mu > 0$ with a type I error (false alarm) risk α

Currie's critical level L_C :

$$L_C = \Phi^{-1}(1 - \alpha)\sqrt{2x_b}$$

• if $x > L_C$, H_0 is rejected with error risk α . $1 - \alpha$ confidence interval:

$$\mu \in x \pm \Phi^{-1}(1 - \alpha/2)\sqrt{x_b + x_g}$$

\Rightarrow **the confidence interval may include negative values**

• if $x \leq L_C$, H_0 is accepted. Currie's detection limit L_D :

$$L_D = 2 \times L_C + \left(\Phi^{-1}(1 - \alpha)\right)^2$$

\Rightarrow **unknown probability of missing a real event** (type II error)

No use of past observed values of x and x_b .

The Bayesian framework (1)

True radioactivity $\mu \equiv$ random variable

\Rightarrow “ H_0 is true” ($\mu = 0$) and “ H_1 is true” ($\mu > 0$) form a complete set of events, with *a priori* and *a posteriori* (given the observed net count x) probabilities:

$$\begin{array}{ccc} \text{posterior} & \longrightarrow & \text{prior} \\ \text{probability} & & \text{probabilities} \\ & \longrightarrow & \longleftarrow \\ & \text{P}(H_i | x) = \frac{f(x | H_i)P(H_i)}{\sum_{k=0,1} f(x | H_k)P(H_k)} & \end{array}$$

Medical diagnosis

- $H_0 \equiv$ “healthy”, $H_1 \equiv$ “sick”, $x \equiv$ result of a medical test (“+” or “-”)
- the ***a priori* probability $P(H_0)$** and the ***conditional* probabilities $f(x|H_k) \equiv P(x|H_k)$** are estimated with the frequencies observed on a large representative sample of patients \Rightarrow **no theoretical problem**

Radioactivity detection

Real-valued x depending on H_k , i.e. on $\mu \Rightarrow$ **conditional densities $f(x|H_k)$**

$$f(x | H_k) = \int_{\mu \in \Theta_k} f(x | \mu) \pi(\mu | H_k) d\mu$$

- $\Theta_0 = \{0\}$ (under H_0) and $\Theta_1 =] 0 ; +\infty [$ (under H_1)
- $f(x|\mu)$: the Gaussian density ($= \varphi((x-\mu)/\sigma))$
- $\pi(\mu|H_k)$: ***a priori* density of the activity μ under H_k**

The Bayesian framework (2)

Denominator of the a *posteriori* probabilities = **prior density of the true radioactivity μ** :

$$\pi(\mu) = \sum_{k=0,1} \pi(\mu | H_k) P(H_k)$$

If $\pi(\mu)$ can be estimated:

- **marginal density of the net count x** : $f(x) = \int f(x | \mu) \pi(\mu) d\mu$

- **posterior probability of H_0 (no radioactivity)**: $P(H_0 | x) = \frac{f(x | \mu = 0) P(H_0)}{f(x)}$

- posterior density μ : $f(\mu | x) = \frac{f(x | \mu) \pi(\mu)}{f(x)}$

- **point estimate of the true radioactivity**:

$$\mu^* = E(\mu | x) = \int \mu f(\mu | x) d\mu$$

- **$1 - \gamma$ credibility interval $[\mu^- ; \mu^+]$** :

$$\frac{\gamma}{2} = \int_{\mu^+}^{+\infty} f(\mu | x) d\mu = \int_{-\infty}^{\mu^-} f(\mu | x) d\mu$$

positive values only

Bayesian priors (1)

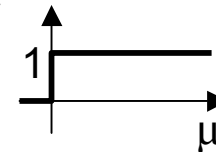
The prior density of the true radioactivity μ should be given by:

$$\pi(\mu) = P(H_0) \delta_0(\mu) + P(H_1) \pi(\mu | H_1)$$

probability of no radioactivity Dirac peak at zero radioactivity density under H_1

Improper prior approach [Zähringer & Kirchner, 2008]:

$$\pi(\mu) = \mathbb{I}_{[0;+\infty[}(\mu)$$



- no Dirac peak at zero $\Rightarrow P(H_0|x) = 0$ whatever the observed x
- not integrable \Rightarrow **the marginal density $f(x)$ cannot be estimated**

Implicit prior approach [Vivier et al., 2009]:

- $P(H_0|x) = 1 - \Phi(x/\sigma) \Rightarrow P(H_0|x=0) = 0.5$ whatever σ (depends on μ_b)
- no explicit prior \Rightarrow **the marginal density $f(x)$ cannot be estimated**

Bayesian priors (2)

Dirac peak
at zero

Proposed prior (proper):

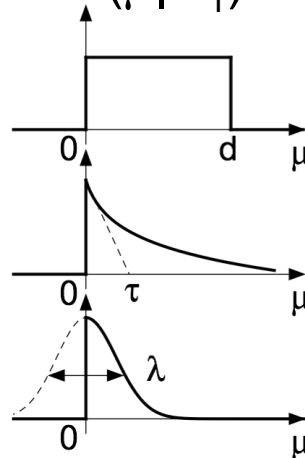
$$\pi(\mu) = P(H_0) \delta_0(\mu) + P(H_1) \pi(\mu | H_1)$$

probability of no radioactivity

radioactivity density under H_1

Possible functional forms for $\pi(\mu|H_1)$:

- uniform ($d > 0$):
- exponential ($\tau > 0$):
- half-Gaussian ($\lambda > 0$):



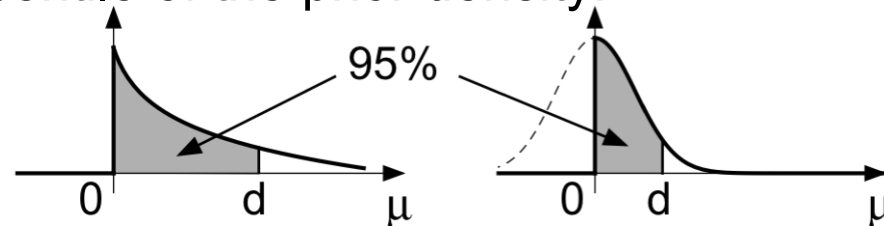
We calculated all the Bayesian estimates for these three priors.

Estimate $P(H_0)$ and $\pi(\mu|H_1)$, i.e. the parameter d , τ or λ by fitting the marginal density $f(\mathbf{x}) = \int f(\mathbf{x}|\mu) \pi(\mu) d\mu$ to past records of observed \mathbf{x} .

Empirical fit of the prior parameters

- d , τ , and λ are merged into **a single parameter, d** , by taking τ and λ so that d is the 95th centile of the prior density:

- **$P(H_0) \equiv p_0$**



Expression of the marginal density of x (uniform prior):

$$f(x) = \frac{p_0}{\sigma} \varphi\left(\frac{x}{\sigma}\right) + \frac{1-p_0}{d} \left(\Phi\left(\frac{x}{\sigma}\right) - \Phi\left(\frac{x-d}{\sigma}\right) \right)$$

In fact, $\sigma^2 = \mu + 2\mu_b$. Numerical simulations show that all the results hold when σ^2 varies by simply replacing σ^2 by $x + 2\mu_b$.

Fit $f(x)$ to past records of x with maximum likelihood, two options:

- as a **function of p_0 , d and μ_b** ,
- first estimate μ_b using measured values of $x_b \Rightarrow$ **check whether $\mu_b \approx \text{cte}$** then fit $f(x)$ as a **function of p_0 and d only (chosen option)**.

Experimental results: empirical fit of the priors

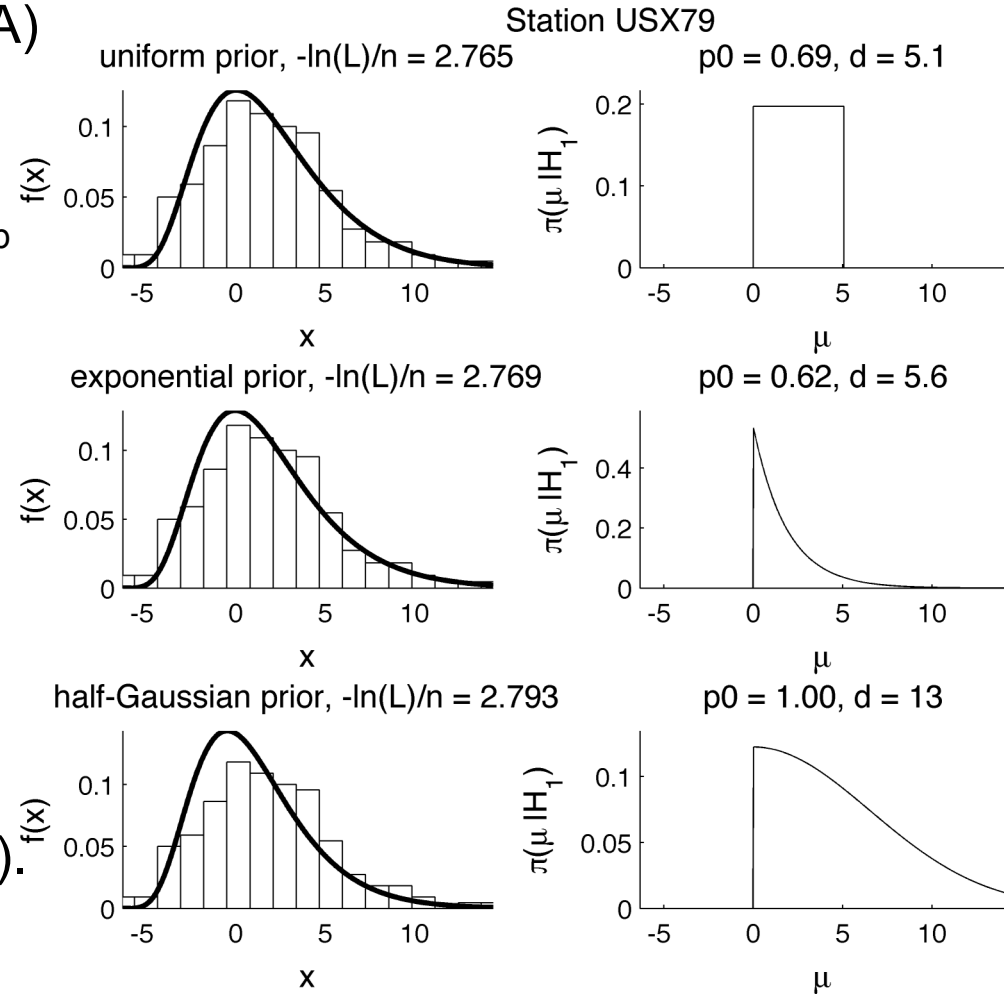
Six month of 2009 daily measurements for a dozen of CTBTO stations:

- Hawaiï station USX79 (SAUNA)
- Xe131m data

- Poisson distribution of blank x_b with $\mu_b \approx 4$

- ML fit of the marginal density $f(x)$ of the observed net counts \Rightarrow **good results with the uniform and exponential priors:**

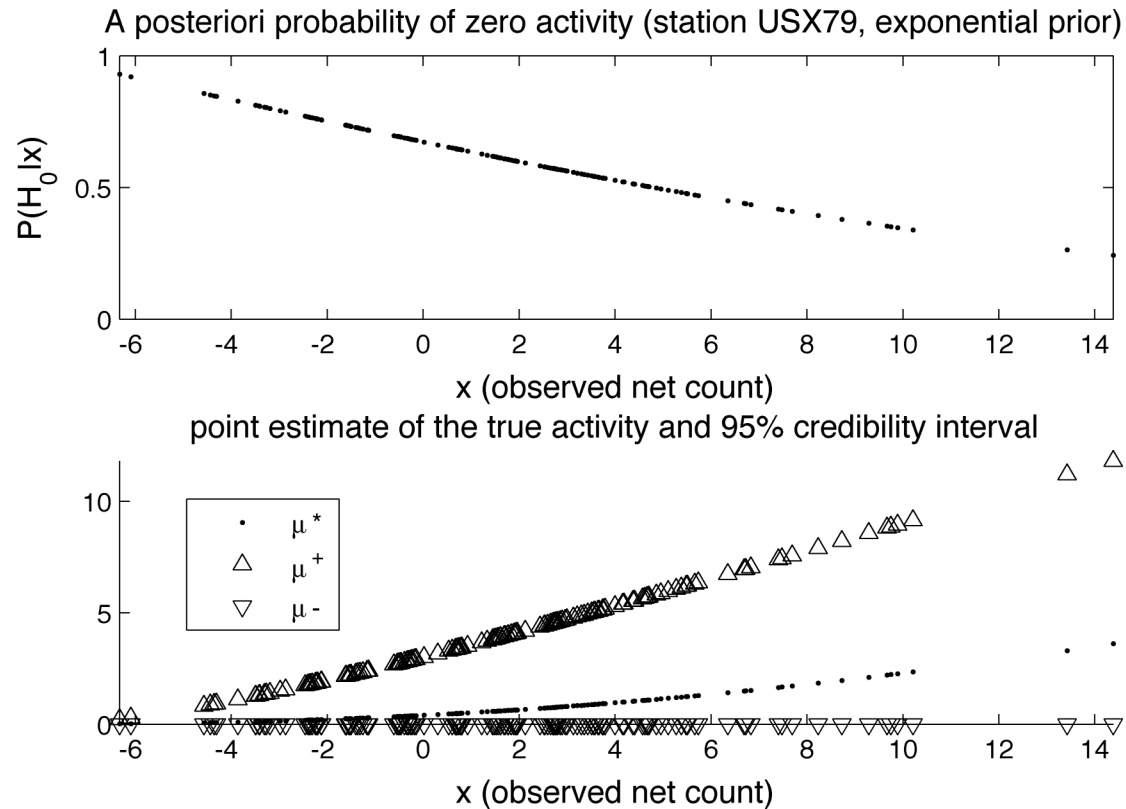
there is an *a priori* probability $p_0 \approx 2/3$ of a small net radioactivity ($d \approx 5-6$).



Experimental results: Bayesian estimates

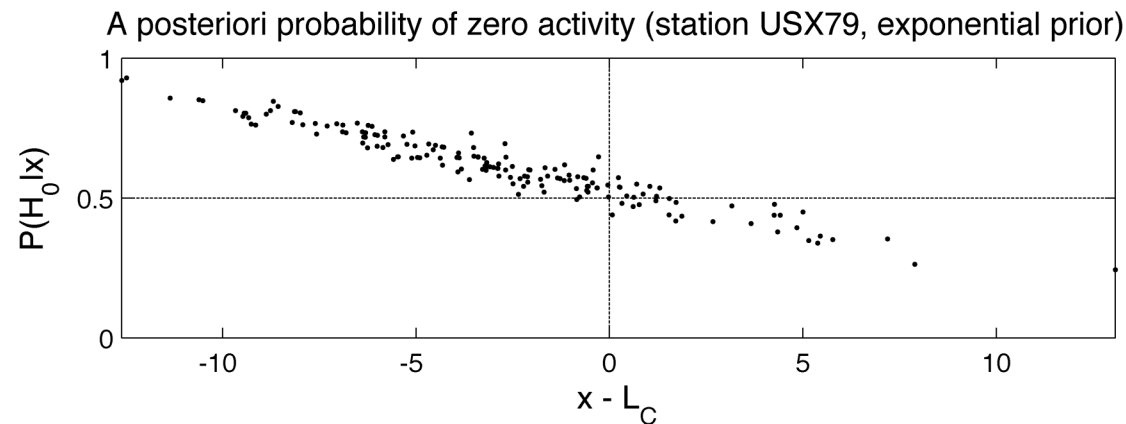
For each observed net count x :

- **estimate of the posterior probability of zero radioactivity** $P(H_0|x)$
- **point estimate** μ^* , **credibility interval** $[\mu^+ ; \mu^-]$ **for the true activity** μ



Experimental results: comparison to the classic framework

The posterior probability of zero activity $P(H_0|x)$ is now shown as a function of $x - L_C$ (L_C is Currie's critical level): $x - L_C > 0$: reject H_0 , else accept H_0



⇒ $x - L_C > 0$ (reject H_0) coincides with $P(H_0|x) < 0.5$

In the Bayesian framework, H_0 should be rejected when:

$$P(H_0|x) < C_{II}/(C_I+C_{II})$$

where C_I and C_{II} are the costs associated to type I (**false alarm**) and type II (**missing a real event**) errors respectively ⇒ **quantifying these costs would define the decision threshold for CTBTO.**

Conclusion

Realizations

Novel approach based on rigorous Bayesian principles (proper priors):

- probability estimate of a truly radioactive sample
- physically meaningful estimates of its radioactivity
- *a priori* knowledge is taken into account via the fit of the prior to data observed in the past

Outlooks

Extend the method to:

- time-varying blank level
- analysis of several isotopes jointly (Xe131m, Xe133m, Xe133, Xe135)
- analysis of other fission and activation products (aerosols) of interest for CTBTO

Supplementary slide 1: references

- **L. A. Currie** (1968)

Limits for qualitative detection and quantitative determination; application to radiochemistry

Anal. Chem. **335**, 586-593.

- **ISO 11929-7** (2005)

Determination of the Detection Limit and Decision Threshold for Ionizing Radiation Measurements, Part 7: Fundamentals and General Applications.

- **M. Zähringer & G. Kirchner** (2008)

Nuclide ratios and source identification from high-resolution gamma-ray spectra with Bayesian decision methods.

Nuclear Instruments and Methods in Physics Research A **594**, 400-406.

- **A. Vivier, Gilbert Le Petit, B. Pigeon & X. Blanchard** (2009)

Probabilistic assessment for a sample to be radioactive or not: application to radioxenon analysis

Journal of Radioanalytical and Nuclear Chemistry **282**, 743–748.

- **I. Rivals, C. Fabbri, G. Euvrard & X. Blanchard**

A Bayesian method with empirically fitted priors for the evaluation of environmental radioactivity: application to low-level radioxenon measurements

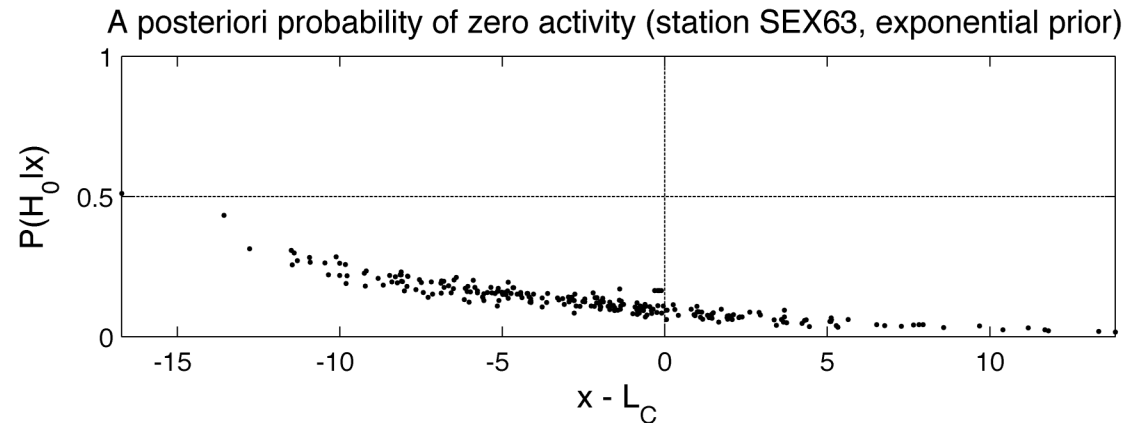
Submitted to *Nuclear Instruments and Methods in Physics Research A*. 13

Supplementary slide 2: results for Stockholm station SEX63

Constant low blank level $\mu_b \approx 6$

Good fit with the exponential prior: $p_0 = 0.13$ (low *a priori* probability of no radioactivity), $d = 7.1$ (small activity when present)

⇒ consistent with the station location in northern Europe (vicinity of nuclear power plants)



⇒ with a threshold of 0.5 for $P(H_0|x)$, **the proposed Bayesian approach detects more radioactive events than the classic approach.**

Supplementary slide 3: validity of the Gaussian approximation **even for very small counts**

Let $X \rightarrow \text{Poisson}(a)$ and $Y \rightarrow \text{Poisson}(b)$, and $Z = X - Y$. Then:

$$P(Z = z) = e^{-(a+b)} \left(\frac{a}{b}\right)^{z/2} I_{|z|} \left(2\sqrt{ab}\right) \quad \text{with} \quad I_z(u) = (u/2)^k \sum_{j=0}^{+\infty} \frac{(u/2)^{2j}}{j!(j+z)!}$$

Application to:

- gross count: $X_g \rightarrow \text{Poisson}(\mu_g = 8)$
- blank count: $X_b \rightarrow \text{Poisson}(\mu_b = 5)$
- net count: $X = X_g - X_b$ ($\mu = 3$)

