Modeling and estimation of Accident Rate and Trend in Air Transport

Henk A.P. Blom & Edwin A. Bloem
National Aerospace Laboratory NLR
Amsterdam, The Netherlands
blom@nlr.nl & bloem@nlr.nl

Abstract - An established approach in the evaluation of aviation accident statistics is to determine point estimates of the accident rate by dividing number of accidents by number of flights and to determine an uncertainty interval through evaluation of the underlying binomial distribution. The trend, however, is not estimated. Another established approach is to perform a regression analysis to estimate rate and trend, but then uncertainty is not estimated. In this paper we overcome these limitations of established approaches by studying the problem as one of Bayesian estimation of the joint conditional density function of accident rate and trend given accident and flight statistical data. Subsequently, a particle filter is used in order to perform numerical evaluations. The novel approach is shown to work well on commercial aviation accident data.

Keywords: Bayesian estimation; accident statistics; particle filtering; uncertainty estimation.

1 Introduction

Aviation accident data provides essential information to monitor aviation safety. If collected at large scale then this data provides insight into the progress made by aviation industry and they may indicate possible safety bottlenecks. Based on these insights, the aviation industry can set their strategy and priorities right.

A long standing problem in the evaluation of aviation accident statistics is the joint estimation of accident rate, trend and uncertainty. An established approach is to divide the number of accidents by the number of flights, and to determine a 95% uncertainty area by using the underlying binomial distribution [1]. However, this approach does not estimate trend. Another established approach in estimating rate and trend is to perform a regression analysis by which the relationship between a dependent variable and one or more independent variables is analyzed. Now the uncertainty is not estimated. Thus with established approaches, either rate and uncertainty or rate and trend are jointly estimated, but not all three.

The aim of this in this paper is to overcome the limitation of the established approaches by developing a Bayesian approach [2] towards the estimation of the joint probability density function of the accident rate and the trend. The numerical evaluation of such a Bayesian approach has become possible due to the development of powerful sequential Monte Carlo simulation techniques [3].

For the problem of estimating the joint conditional probability density function of accident rate and trend given large scale aviation accident and flight statistics, an exact Bayesian characterization is being developed first. Subsequently, a particle filter approximation is introduced in order to perform numerical evaluations. This particle filter is then used to perform joint estimation of accident rate, trend and uncertainty from commercial aviation accident data. In [5] we studied this problem under the assumption that the trend was constant. The current study extends these results by assuming that the trend may slowly evolve over time.

The paper is organized as follows. Section 2 provides a mathematical formulation of the problem. Section 3 derives a recursive Bayesian characterization of the joint conditional probability density function for the rate and trend. Section 4 presents a particle filter approach towards the evaluation of this joint conditional probability density function. Sections 5 and 6 present the results for the particle filter applied to worldwide aviation accident data, and compares the results obtained with classical estimation results. Section 5 considers the case of constant trend, and Section 6 considers the case of slowly evolving trend. Section 7 draws conclusions.

2 Mathematical formulation of the problem

This section proposes a specific mathematical formulation of the problem addressed in this paper. In order to characterize a model for the number of accidents per flight in a year, we assume that the accident rates are piecewise constant. We also assume that observed data is available on the number of flights and on the number of accidents per year.

Let $\lambda_k \in [0,1]$ denote the accident rate per flight in year k and $a_k \in \mathbb{R}$ the accident trend in year k, $k \in [0,...,F]$ and assume that these two evolve according to the following model:

$$a_k = a_{k-1} + \sigma w_k$$
$$\lambda_k = a_k \lambda_{k-1} \quad (1)$$

with $\lambda_0 \in [0,1]$, $a_0 \in [1-\varepsilon, 1+\varepsilon]$, $\sigma \geq 0$ and $\{w_k\}$ a sequence of i.i.d. $N\{0,\sigma^2\}$ random variables. Furthermore we assume that the initial joint probability distribution p_{λ_0,a_0} is a Uniform distribution on $[0,1]\times[1-\varepsilon,1+\varepsilon]$.

Let h_k denote the number of flights in year k. The accident rate Λ_k in year k is then given by

$$\Lambda_k = h_k \lambda_k \quad (2)$$

Let κ_k denote the number of accidents in year k. We assume that κ_k given Λ_k has a Poisson distribution:

$$p_{\kappa_k|\Lambda_k}(\kappa \mid \Lambda) = \begin{cases} \dfrac{\Lambda^\kappa}{\kappa!}\exp(-\Lambda) & \kappa = 0,1,2,... \\ 0 & \text{else} \end{cases} \quad (3)$$

Bayesian filtering problem

Let $\mathcal{H}_k = \{\underline{h}_0,...,\underline{h}_k\}$ and $\mathcal{K}_k = \{\underline{\kappa}_0,...,\underline{\kappa}_k\}$ where $\underline{h}_k$ and $\underline{\kappa}_k$ are the observed realizations of h_k and κ_k. Given the flight statistics $\mathcal{H}_k$ and the accident statistics $\mathcal{K}_k$, the problem is to characterize the joint conditional density of λ_k and a_k:

$$p_{\lambda_k,a_k|\mathcal{H}_k,\mathcal{K}_k}(\lambda,a)$$

Estimates $\hat{\lambda}_k \triangleq E\{\lambda_k \mid \mathcal{H}_k,\mathcal{K}_k\}$ and $\hat{a}_k \triangleq E\{a_k \mid \mathcal{H}_k,\mathcal{K}_k\}$ are characterized from this joint conditional density:

$$\hat{\lambda}_k = \int\int \lambda\, p_{\lambda_k,a_k|\mathcal{H}_k,\mathcal{K}_k}(\lambda,a)\, da\, d\lambda \quad (4a)$$

$$\hat{a}_k = \int a\left(\int p_{\lambda_k,a_k|\mathcal{H}_k,\mathcal{K}_k}(\lambda,a)\, d\lambda\right) da \quad (4b)$$

From the joint conditional density, 95% uncertainty intervals for λ_k can also be determined by the values $\hat{b}^{\text{lower}}_{\lambda_k}$ and $\hat{b}^{\text{upper}}_{\lambda_k}$ such that

$$\int_0^{\hat{b}^{\text{lower}}_{\lambda_k}}\left(\int_{1-\varepsilon}^{1+\varepsilon} p_{\lambda_k,a_k|\mathcal{H}_k,\mathcal{K}_k}(\lambda,a)\, da\right) d\lambda = 0.025$$
$$\int_{\hat{b}^{\text{upper}}_{\lambda_k}}^{1}\left(\int_{1-\varepsilon}^{1+\varepsilon} p_{\lambda_k,a_k|\mathcal{H}_k,\mathcal{K}_k}(\lambda,a)\, da\right) d\lambda = 0.025$$

Similarly, 95% uncertainty intervals for a_k can be determined by the values $\hat{b}^{\text{lower}}_{a_0}$ and $\hat{b}^{\text{upper}}_{a_0}$ such that

$$\int_{1-\varepsilon}^{\hat{b}^{\text{lower}}_{a_k}}\left(\int_0^1 p_{\lambda_k,a_k|\mathcal{H}_k,\mathcal{K}_k}(\lambda,a)\, d\lambda\right) da = 0.025$$
$$\int_{\hat{b}^{\text{upper}}_{a_k}}^{1+\varepsilon}\left(\int_0^1 p_{\lambda_k,a_k|\mathcal{H}_k,\mathcal{K}_k}(\lambda,a)\, d\lambda\right) da = 0.025$$

3 Characterization of joint conditional density

In this section a recursive Bayesian characterization of the joint conditional density $p_{\lambda_k,a_k|\mathcal{H}_k,\mathcal{K}_k}(\lambda,a)$ is derived.

Applying Bayes' rule yields:

$$p_{\lambda_k,a_k|\mathcal{K}_k,\mathcal{H}_k}(\lambda,a) = p_{\lambda_k,a_k|\kappa_k,\mathcal{K}_{k-1},\mathcal{H}_k}(\lambda,a \mid \underline{\kappa}_k)$$
$$= \frac{1}{c_k} p_{\kappa_k|\lambda_k,a_k,\mathcal{K}_{k-1},\mathcal{H}_k}(\underline{\kappa}_k \mid \lambda,a)\, p_{\lambda_k,a_k|\mathcal{K}_{k-1},\mathcal{H}_k}(\lambda,a)$$

where c_k denotes a normalising constant, $\underline{\kappa}_k$ is the observed realization of κ_k.

Since $p_{\lambda_k,a_k|\mathcal{K}_{k-1},\mathcal{H}_k}(\lambda,a)$ is independent of the number of flights in year k, we have

$$p_{\lambda_k,a_k|\mathcal{K}_{k-1},\mathcal{H}_{k-1},h_k}(\lambda,a) = p_{\lambda_k,a_k|\mathcal{K}_{k-1},\mathcal{H}_{k-1}}(\lambda,a).$$

Substitution of the latter equation in the former yields

$$p_{\lambda_k,a_k|\mathcal{K}_k,\mathcal{H}_k}(\lambda,a) = \frac{1}{c_k} p_{\kappa_k|\lambda_k,a_k,\mathcal{K}_{k-1},\mathcal{H}_k}(\underline{\kappa}_k \mid \lambda,a)\, p_{\lambda_k,a_k|\mathcal{K}_{k-1},\mathcal{H}_{k-1}}(\lambda,a) \quad (5)$$

Using equations (1), (2) and (3) to evaluate (5) yields

$$p_{\lambda_k,a_k|\mathcal{K}_k,\mathcal{H}_k}(\lambda,a) = \frac{1}{c_k}\frac{(\underline{h}_k\lambda)^{\underline{\kappa}_k}}{\underline{\kappa}_k!}\exp(-\underline{h}_k\lambda)\cdot$$
$$\cdot p_{\lambda_k,a_k|\mathcal{K}_{k-1},\mathcal{H}_{k-1}}(\lambda,a) \quad (6)$$

Next we characterize $p_{\lambda_{k+1},a_{k+1}|\mathcal{K}_k,\mathcal{H}_k}(\lambda,a)$ in terms of $p_{\lambda_k,a_k|\mathcal{K}_k,\mathcal{H}_k}(\lambda,a)$:

$$p_{\lambda_{k+1},a_{k+1}|\mathcal{K}_k,\mathcal{H}_k}(\lambda,a) = \int p_{\lambda_{k+1},a_{k+1}|\lambda_k,a_k}(\lambda,a \mid \lambda',a')\cdot$$
$$\cdot\, p_{\lambda_k,a_k|\mathcal{K}_k,\mathcal{H}_k}(\lambda',a')\, d\lambda' da' \quad (7)$$

The transition density in (7) satisfies:

$$p_{\lambda_{k+1},a_{k+1}|\lambda_k,a_k}(\lambda,a \mid \lambda',a') = p_{\lambda_{k+1}|a_{k+1},\lambda_k}(\lambda \mid a,\lambda')\, p_{a_{k+1}|a_k}(a \mid a')$$

Due to (1) this yields:

$$p_{\lambda_{k+1},a_{k+1}|\lambda_k,a_k}(\lambda,a\,|\,\lambda',a') = \delta_{[a\lambda']}(\lambda)\,N\{a;a',\sigma^2\} \qquad (8)$$

Substituting this in (7) yields:

$$p_{\lambda_{k+1},a_{k+1}|\mathcal{K}_k,\mathcal{H}_k}(\lambda,a) = \int \delta_{[a\lambda']}(\lambda) N\{a;a',\sigma^2\}\cdot \cdot p_{\lambda_k,a_k|\mathcal{K}_k,\mathcal{H}_k}(\lambda',a')\,d\lambda' da' \qquad (9)$$

4 Particle filtering towards estimation of rate, trend and uncertainty

In this section a Sampling Importance Resampling (SIR) type of particle filter approach [6]-[7] towards the numerical evaluation of $p_{\lambda_k,a_k|\mathcal{H}_k,\mathcal{K}_k}(\lambda,a)$ is presented. This particle filter is able to provide an arbitrarily close approximation of the true Bayesian solution by increasing the number of particles. The main idea behind this particle filter is to approximate the joint conditional density of λ_k and a_k given $\mathcal{H}_k$ and $\mathcal{K}_k$ by an empirical density that is defined by a set of particles, i.e. samples from the joint conditional density with corresponding weights.

Particles are randomly drawn from the initial distribution p_{λ_0,a_0}, each with same weights. Next the particles evolve and are updated according to the underlying stochastic model and the new measurements, where for each particle its weight is adapted based on the likelihood of the measurements for that particle. For the problem at hand, the underlying stochastic model is given by equations (1), (2) and (3), the measurements are given by $\mathcal{H}_k$ and $\mathcal{K}_k$, and the weights are adapted based on equation (6). With this particle filter, estimates $\hat{\lambda}_k$ and $\hat{a}_k$ of λ_k and a_k can be obtained by simply taking the weighted average over all particles. A formal description of this particle filter reads as follows:

A *particle* is defined as a triplet $(\mu_0^j,\lambda_0^j,a_0^j)$, $\mu_0^j\in[0,1]$, $\lambda_0^j\in[0,1]$, $a_0^j\in[1-\varepsilon,1+\varepsilon]$, $j\in[1,...,N]$, where N denotes the number of particles, j refers to the j^{th} particle, μ^j denotes the weight of particle j, λ_0^j denotes the expected number of flights in year 0 of particle j, and a_0^j denotes the trend parameter of particle j. With these particles the joint conditional density of λ_0 and a_0 can be approximated by

$$p_{\lambda_0,a_0}(\lambda,a) \approx \sum_{j=1}^{N} \mu_0^j \delta_{\left[\lambda_0^j,a_0^j\right]}(\lambda,a)$$

For $k=0,...,F$, the particle filter cycles through steps 1 through 5 below:

Step 1: Resampling or initiation of particles.
For $k>0$, Step 1 performs a resampling of the particles when the effective sample size drops below a certain threshold [8]. For $k=0$, Step 1 initializes a set of N particles in $[0,1]\times[0,1]\times\mathbb{R}$ i.e.

$$\left\{\left(\mu_0^j,\lambda_0^j,a_0^j\right);\, j\in[1,N]\right\}$$

With $\mu_0^j = 1/N$, λ_0^j independently drawn from $p_{\lambda_0}(\lambda)$ and a_0^j independently drawn from $p_{a_0}(a)$ for each $j\in[1,...,N]$, e.g. both λ_0^j and a_0^j independently drawn from Uniform distributions on $[0,1]$ and $[1-\varepsilon,1+\varepsilon]$ respectively.

Step 2: Measurement processing
Using measurement (h_k,κ_k), determine new weights per particle,

$$\left\{\left(\hat{\mu}_k^j,\lambda_k^j,a_k^j\right);\, j\in[1,N]\right\} \qquad (10)$$

with for the new weights, using equation (6) for $k=0,1,2,...$:

$$\hat{\mu}_k^j = \frac{1}{c_k}\frac{\left(h_k\lambda_k^j\right)^{\kappa_k}}{\kappa_k!}\exp\left(-h_k\lambda_k^j\right)\mu_k^j \qquad (11)$$

with c_k a normalising constant such that $\sum_{j=1}^{N}\hat{\mu}_k^j = 1$.

Step 3. Joint conditional density at year k:
The particle filter outputs the joint conditional density of λ_k and a_k given $\mathcal{H}_k$ and $\mathcal{K}_k$ in the form of empirical density

$$p_{\lambda_k,a_k|\mathcal{K}_k,\mathcal{H}_k}(\lambda,a) \approx \sum_{j=1}^{N}\hat{\mu}_k^j\delta_{\left[\lambda_k^j,a_k^j\right]}(\lambda,a) \qquad (12)$$

Estimates $\hat{\lambda}_k$ and $\hat{a}_k$ are obtained by calculating the weighted average over all particles:

$$\hat{\lambda}_k \approx \sum_{j=1}^{N}\hat{\mu}_k^j\lambda_k^j \qquad (13)$$

$$\hat{a}_k \approx \sum_{j=1}^{N}\hat{\mu}_F^j a_k^j \qquad (14)$$

Step 4. Prediction
Based on equation (1) we generate new particles

$$a_{k+1}^j = a_k^j + \sigma w_k^j$$
$$\lambda_{k+1}^j = a_{k+1}^j \lambda_k^j$$
$$\mu_{k+1}^j = \hat{\mu}_k^j$$

$$p_{\lambda_{k+1},a_{k+1}|\mathcal{K}_k,\mathcal{H}_k}(\lambda,a) \approx \sum_{j=1}^{N} \mu_{k+1}^j \delta_{\left[\lambda_{k+1}^j, a_{k+1}^j\right]}(\lambda,a) \qquad (15)$$

Step 5. In this step we determine the probability density of κ_{k+1} given $\mathcal{H}_{k+1}$ and $\mathcal{K}_k$.
From the law of total probability we have:

$$p_{\kappa_{k+1}|\mathcal{K}_k,\mathcal{H}_{k+1}}(\kappa) = \int p_{\kappa_{k+1}|\lambda_{k+1},\mathcal{K}_k,\mathcal{H}_{k+1}}(\kappa\,|\,\lambda)\, p_{\lambda_{k+1}|\mathcal{K}_k,\mathcal{H}_{k+1}}(\lambda)\, d\lambda$$

Together with (15) this yields

$$p_{\kappa_{k+1}|\mathcal{K}_k,\mathcal{H}_{k+1}}(\kappa) \approx \int p_{\kappa_{k+1}|\lambda_{k+1},h_{k+1}}(\kappa\,|\,\lambda,h_{k+1}) \sum_{j=1}^{N} \mu_{k+1}^j \delta_{\left[\lambda_{k+1}^j\right]}(\lambda)\, d\lambda$$

Thanks to (3), this becomes for $\kappa = 0,1,2,\ldots$:

$$p_{\kappa_{k+1}|\mathcal{K}_k,\mathcal{H}_{k+1}}(\kappa) \approx \int \frac{(\lambda h_{k+1})^\kappa}{\kappa!} \exp(-\lambda h_{k+1}) \sum_{j=1}^{N} \mu_{k+1}^j \delta_{\left[\lambda_{k+1}^j\right]}(\lambda)\, d\lambda$$

Subsequent evaluation yields for $\kappa = 0,1,2,\ldots$

$$p_{\kappa_{k+1}|\mathcal{K}_k,\mathcal{H}_{k+1}}(\kappa) \approx \sum_{j=1}^{N} \mu_{k+1}^j \frac{\left(\lambda_{k+1}^j h_{k+1}\right)^\kappa}{\kappa!} \exp\left(-\lambda_{k+1}^j h_{k+1}\right) \qquad (16)$$

which is the equation to be used in step 5.

5 Application of the particle filter to worldwide aviation accident data

The particle filter of section 4 is now applied to worldwide commercial aviation accident and flight data. First we explain which input data is used. Subsequently the particle filtering results are presented.

5.1 Input data

The aviation accident and flight data used in this paper are from [4]. The table below specifies the criteria that have been used for the selection from this database.

Table 1. Data selection criteria

Selection Criteria of accidents and flights	
Time period	1/1/1990 – 31/12/2008
Occurrence Class	Accident
Aircraft Category	Fixed Wing
Aircraft Mass group	> 5700 kg
Location of occurrence	Worldwide (not filtered)
Operation Type	Scheduled Commercial Air Transport

In figures 1 and 2 the resulting $\mathcal{H}_F$ and $\mathcal{K}_F$ data is visualised. Figures 1 and 2 provide the number of flights $\underline{h}_k$ and the number of accidents $\underline{\kappa}_k$ for $k = 0,1,\ldots,18$, where $k = 0$ corresponds with year 1990. Figure 3 provides for each year the ratio between the number of accidents $\underline{\kappa}_k$ and the number of flights $\underline{h}_k$.

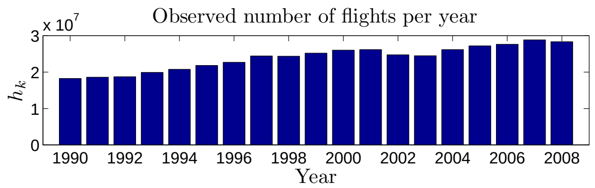

Figure 2. Observed number of flights $\underline{h}_k$ for $k = 0,1,\ldots,18$, corresponding with years 1990 – 2008.

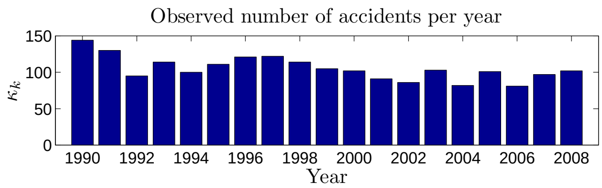

Figure 2. Observed number of accidents $\underline{\kappa}_k$ for $k = 0,1,\ldots,18$, corresponding with years 1990 – 2008.

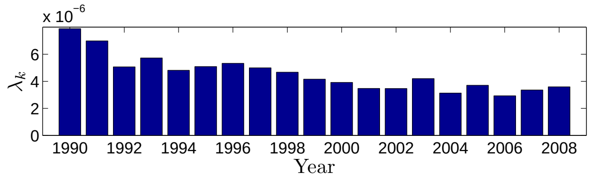

Figure 3. Ratio $\underline{\kappa}_k / \underline{h}_k$ for $k = 0,1,\ldots,18$, corresponding with years 1990–2008.

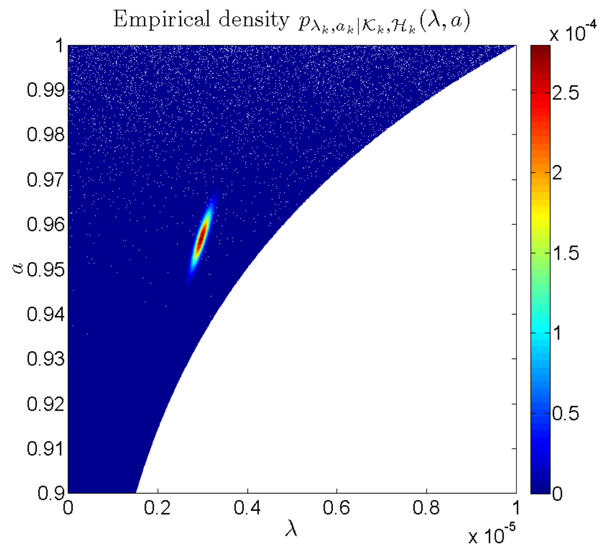

Figure 4a. Joint conditional density of $p_{\lambda_k,a_k|\mathcal{K}_k,\mathcal{H}_k}$ for $k = F$ when $\sigma = 0$.

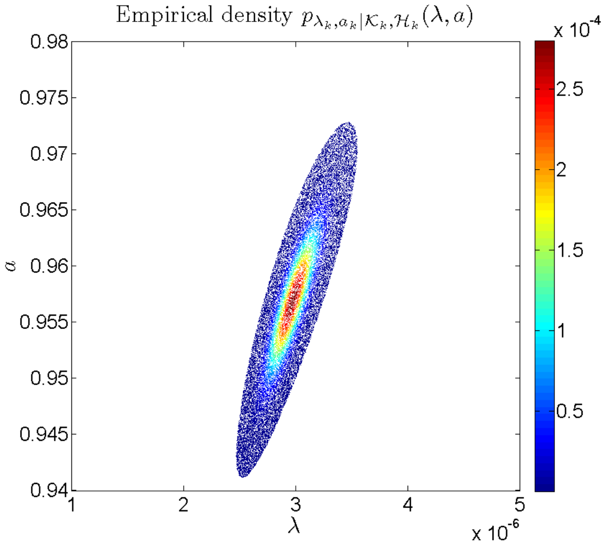

Figure 4b. Joint conditional density of $p_{\lambda_k, a_k | \mathcal{K}_k, \mathcal{H}_k}$ for $k = F$ when $\sigma = 0$ and with all particles with weights below 10^{-7} ignored.

5.2 Particle filtering results when $\sigma = 0$

We use the particle filter equations of section 4 with one million particles (i.e. $N = 10^6$), and without applying resampling in Step 1. First we consider the case $\sigma = 0$, i.e. fixed trend $a_k = a_{k-1}$. Particle filter based numerical evaluation of $p_{\lambda_F, a_F | K_F, H_F}(\lambda, a)$ is depicted in the form of the empirical joint conditional densities in figure 4a,b for $\sigma = 0$. Figure 4a shows all particles whereas Figure 4b ignores all particles with weights below value 10^{-7}. From the orientation of the empirical joint conditional densities in Figure 4b it can be observed that estimation of accident rate λ_F and accident trend a_F involves strong correlation. Figures 5 and 6 provide marginal empirical densities for $p_{\lambda_F | \mathcal{K}_F, \mathcal{H}_F}(\lambda)$ and $p_{a_F | \mathcal{K}_F, \mathcal{H}_F}(a)$ as these resulted from the empirical joint conditional density $p_{\lambda_F, a_F | \mathcal{H}_F, \mathcal{K}_F}(\lambda, a)$.
Corresponding 95% uncertainty intervals are also given (dotted vertical lines). From the shapes of the empirical densities in these figures it can be observed that they closely resemble Gaussian densities.
Figure 7 provides the marginal empirical density for $p_{\kappa_F | \mathcal{K}_{F-1}, \mathcal{H}_F}(\kappa)$ which follows from particle filter step 4.
For year 2008, the estimates of the accident rate $\hat{\lambda}_F$ and accident trend $\hat{a}_F$ are given by

$$\hat{\lambda}_F = 2.97 \times 10^{-6}$$
$$\hat{a}_F = 0.957$$

The corresponding standard deviations $\hat{\sigma}_{\lambda_F}$ and $\hat{\sigma}_{a_F}$, and correlation coefficient $\hat{\rho}_{\lambda_F, a_F}$ are given by

$$\hat{\sigma}_{\lambda_F} = 0.14 \times 10^{-6}$$
$$\hat{\sigma}_{a_F} = 0.004$$
$$\hat{\rho}_{\lambda_F, a_F} = 0.87$$

Hence, standard deviations amount some 5% of the estimated means, and there is strong correlation.

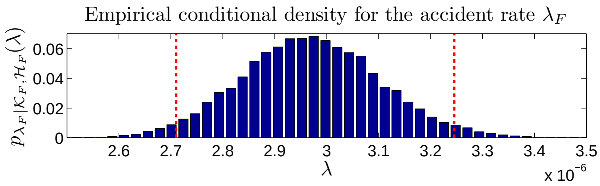

Figure 5. Empirical conditional density for the accident rate at year 2008, $\lambda_F : p_{\lambda_F | \mathcal{K}_F, \mathcal{H}_F}(\lambda)$ with 95% uncertainty interval, when $\sigma = 0$.

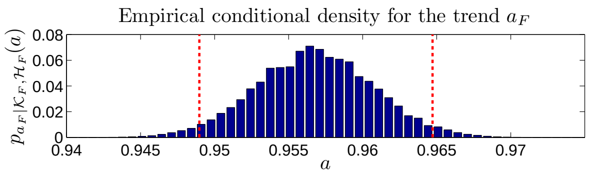

Figure 6. Empirical conditional density for the trend at year 2008, $a_F : p_{a_F | \mathcal{K}_F, \mathcal{H}_F}(a)$ with 95% uncertainty interval, when $\sigma = 0$

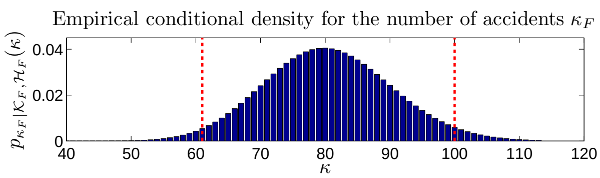

Figure 7. Empirical conditional density $p_{\kappa_F | \mathcal{K}_{F-1}, \mathcal{H}_F}(\kappa)$ for κ_F at year 2008, with 95% uncertainty interval, when $\sigma = 0$

5.3 Comparison with classical estimation results when $\sigma = 0$

An established, approach, e.g. [1], is to determine for each year the ratio between the number $\underline{\kappa}_k$ of accidents and the number $\underline{h}_k$ of flights as an indication of estimated accident rate for each year, and to use the underlying binomial distribution to determine a 95% uncertainty area around this point estimate.
Now we compare the classical estimated 95% uncertainty intervals with our new 95% uncertainty intervals that apply to $\kappa_k / \underline{h}_k$ with $p_{\kappa_k | \mathcal{K}_{k-1}, \mathcal{H}_k}(\kappa)$ as is illustrated in Figure 7 for $k = F$. Figure 8 shows for each year the particle filtering based estimation of mean rate and 95%

uncertainty interval of $\kappa_k / \underline{h}_k$ versus the classical point estimates and 95% uncertainty interval of the accident probability. The dashed line represents the mean, and the dotted lines represent the 95% uncertainty interval of our new estimation results. The classical point estimates of the accident probability with 95% uncertainty intervals are depicted as circles with corresponding error bars.

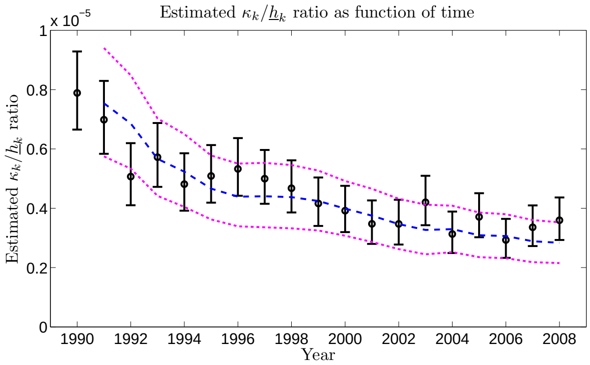

Figure 8. Newly estimated $\kappa_k / \underline{h}_k$ ratio (▬ ▬ ▬ = mean) with 95% uncertainty interval (▪▪▪▪▪) versus classical point estimates (● = mean) and 95% uncertainty interval (I) of $p_{\kappa_k / \underline{h}_k | \mathcal{K}_{k-1}, \mathcal{H}_k}(\cdot)$.

Figure 8 shows that the sizes of the 95% areas of the $\kappa_k / \underline{h}_k$ ratios are quite similar for both approximations. However the new approach yields a much smoother estimate of the evolution of the mean over time.

6 Filtering results when $\sigma \neq 0$

In the previous section, we assumed that accident trend $\{a_k\}$ does not change. However it is more realistic to assume that $\{a_k\}$ itself is changing over time. In this section we show the particle filtering results for $\sigma = 0.01$.

6.1 Particle filtering results when $\sigma = 0.01$

Again we use the particle filter equations of section 4 with one million particles (i.e. $N = 10^6$) and without applying resampling in Step 1. We consider the case $\sigma = 0.01$, which allows for small random trend variations $a_k = a_{k-1} + \sigma w_k$. Particle filter based numerical evaluation of $p_{\lambda_F, a_F | K_F, H_F}(\lambda, a)$ is depicted in the form of the empirical joint conditional densities in figure 9a,b for $\sigma = 0.01$. From the orientation of the empirical joint conditional densities in Figure 9b it can observed that estimation of accident rate λ_F and accident trend a_F involves strong correlation.

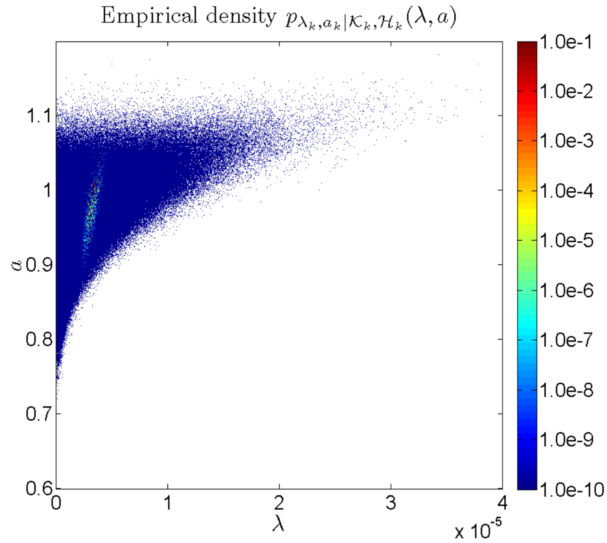

Figure 9a. Joint conditional density of $p_{\lambda_k, a_k | \mathcal{K}_k, \mathcal{H}_k}$ for $k = F$ when $\sigma = 0.01$ and with all particles with weights below 10^{-1000} ignored.

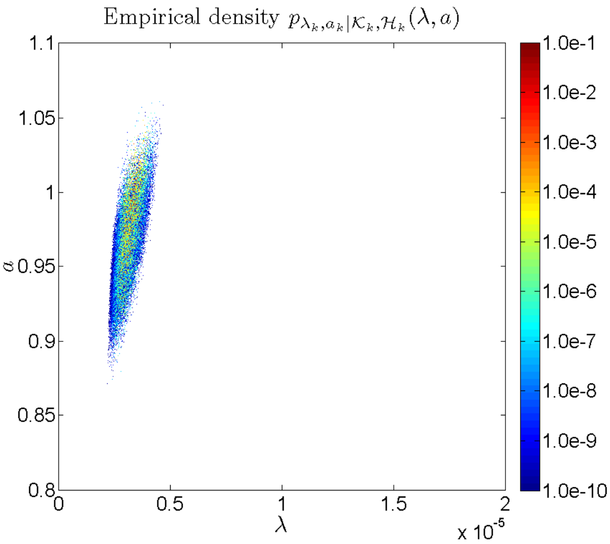

Figure 9b. Joint conditional density of $p_{\lambda_k, a_k | \mathcal{K}_k, \mathcal{H}_k}$ for $k = F$ when $\sigma = 0.01$ and with all particles with weights below 10^{-10} ignored.

Figures 10 and 11 provide marginal empirical densities for $p_{\lambda_F | \mathcal{K}_F, \mathcal{H}_F}(\lambda)$ and $p_{a_F | \mathcal{K}_F, \mathcal{H}_F}(a)$ as these resulted from the empirical joint conditional density $p_{\lambda_F, a_F | \mathcal{H}_F, \mathcal{K}_F}(\lambda, a)$. Corresponding 95% uncertainty intervals are also given (dotted vertical lines). From the shapes of the empirical densities in these figures it can be observed that they still resemble Gaussian densities.
Figure 12 provides the marginal empirical density for $p_{\kappa_F | \mathcal{K}_{F-1}, \mathcal{H}_F}(\kappa)$ which follows from particle filter step 4.

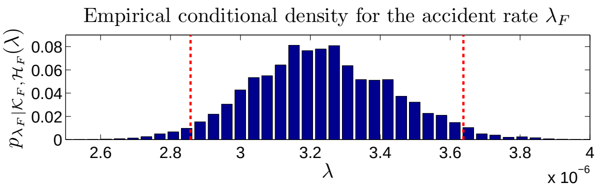

Figure 10. Empirical conditional density for the accident rate at year 2008, $\lambda_F : p_{\lambda_F|\mathcal{K}_F,\mathcal{H}_F}(\lambda)$ with 95% uncertainty interval, when $\sigma = 0.01$.

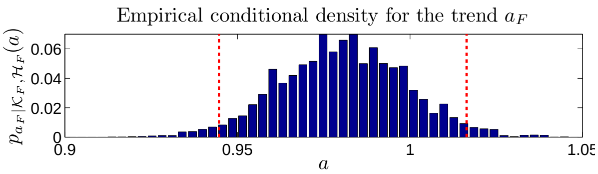

Figure 11. Empirical conditional density for the trend at year 2008, $a_F : p_{a_F|\mathcal{K}_F,\mathcal{H}_F}(a)$ with 95% uncertainty interval, when $\sigma = 0.01$

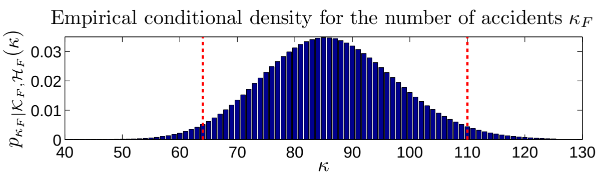

Figure 12. Empirical conditional density $p_{\kappa_F|\mathcal{K}_{F-1},\mathcal{H}_F}(\kappa)$ for κ_F at year 2008, with 95% uncertainty interval, when $\sigma = 0.01$

For year 2008, the estimates of the accident rate $\hat{\lambda}_F$ and accident trend $\hat{a}_F$ are given by

$$\hat{\lambda}_F = 3.24\times10^{-6}$$
$$\hat{a}_F = 0.981$$

The corresponding standard deviations $\hat{\sigma}_{\lambda_F}$ and $\hat{\sigma}_{a_F}$, and correlation coefficient $\hat{\rho}_{\lambda_F,a_F}$ are given by

$$\hat{\sigma}_{\lambda_F} = 0.20\times10^{-6}$$
$$\hat{\sigma}_{a_F} = 0.018$$
$$\hat{\rho}_{\lambda_F,a_F} = 0.70$$

Comparison with the results for $\sigma = 0$, in Section 5, shows that the estimated trend value is no longer significant better than the value 1.

6.2 Comparison with classical estimation results when $\sigma = 0.01$

Now we compare the classical estimated 95% uncertainty intervals with our new 95% uncertainty intervals that apply to $\kappa_k / \underline{h}_k$ with $p_{\kappa_k|\mathcal{K}_{k-1},\mathcal{H}_k}(\kappa)$ as is illustrated in Figure 10 for $k = F$. Figure 13 shows for each year the particle filtering based estimation of mean rate and 95% uncertainty interval of $\kappa_k / \underline{h}_k$ versus the classical point estimates and 95% uncertainty interval of the accident probability. The dashed line represents the mean, and the dotted lines represent the 95% uncertainty interval of our new estimation results. The classical point estimates of the accident probability with 95% uncertainty intervals are depicted as circles with corresponding error bars.

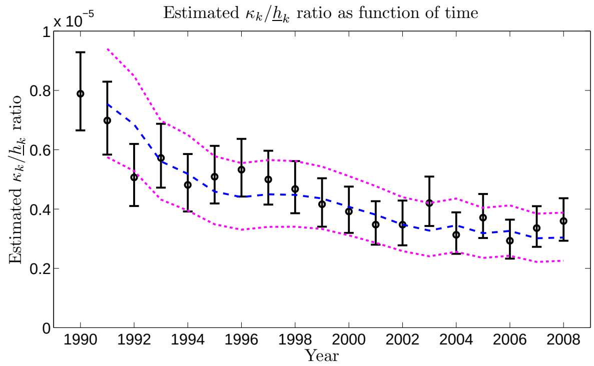

Figure 13. Newly estimated $\kappa_k / \underline{h}_k$ ratio (— — — = mean) with 95% uncertainty interval (⋯⋯) versus classical point estimates (● = mean) and 95% uncertainty interval (I) of $p_{\kappa_k/\underline{h}_k|\mathcal{K}_{k-1},\mathcal{H}_k}(\cdot)$.

Figure 13 shows that the sizes of the 95% areas of the $\kappa_k / \underline{h}_k$ ratios are now slightly larger than the uncertainty intervals of the classical approach. However the new approach still yields a much smoother estimate of the evolution of the mean over time.

7 Concluding remarks

In this paper a novel approach towards joint estimation of accident rate, trend and uncertainty in aviation accident data has been developed. This novel approach is based on the exact Bayesian estimation of the joint conditional density of the accident rate and trend given observed accident and flight data. For numerical evaluation a particle filter approximation has been developed.
Numerical evaluations and comparison to classical approach shows the validity of the novel approach for joint estimation of accident rate, trend and uncertainty. A relevant finding is the following. If the trend is assumed to be fixed over the period 1990 – 2008, then the estimated trend for 2008 is statistically significant lower than the

value 1. However, if the trend over the period 1990 – 2008 is assumed to be slightly variable, then the trend estimated for 2008 does not have a statistically significant deviation from value 1. Because the value of σ has significant impact on the estimated trend, it seems to be relevant to consider the joint estimation of rate λ, trend a and trend evolution parameter σ in a follow-up study.

References

[1] G. van Es. "A Review of Civil Aviation Accidents Air Traffic Management Related Accidents:1980-1999," Proc. USA/Europe R&D Seminar on Air Traffic Management, Santa Fe, USA, 2001.

[2] J.M. Bernardo and A.F.M. Smith, Bayesian Theory, Wiley, 1994.

[3] A. Doucet, N. de Freitas, and N. Gordon, Sequential Monte Carlo Methods in Practice, Springer-Verlag, 2001.

[4] NLR Air Safety Data Base, version 2009.

[5] E.A. Bloem and H.A.P. Blom, Bayesian estimation of accident rate, trend and uncertainty, Proc. ATOS conference, Delft, Netherlands, 14-15 April 2010.

[6] D.B. Rubin, "Using the SIR algorithm to simulate posterior distributions," Eds. J.M. Bernardo et al., Bayesian Statistics 3, Oxford Univ. Press, 1988, pp. 395-402.

[7] N.J. Gordon, D. J. Salmond, A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," IEE Proceedings-F, Vol. 140, pp. 107-113, 1993.

[8] J.S. Liu, R. Chen, Sequential Monte Carlo methods for dynamical systems, J. American Statistical Association, Vol. 93 (1998), pp. 1032-1044.