Safety Science 104 (2018) 216–230

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

Prediction of aircraft safety incidents using Bayesian inference and hierarchical structures

Rosa María Arnaldo Valdés*, V. Fernando Gómez Comendador, Luis Perez Sanz, Alvaro Rodriguez Sanz

Universidad Politecnica de Madrid, School of Aerospace Engineering, Plaza de Cardenal Cisneros, n3, 28040 Madrid, Spain

ARTICLE INFO

Keywords:
Risk
Incident rate
Predictive models
Bayesian inference
Hierarchical models

ABSTRACT

Today, aviation is immersed in a shift from old-fashioned reactive and compliance-based safety approaches towards proactive and performance-based methods and tools. Stakeholders have to monitor, gather and analyse safety-related data and information in order to anticipate and predict actual and emerging safety risks. In this context safety analytics and statistics need to evolve to forecast future safety performances and risks.

This research adopts an innovative statistical approach involving the use of Bayesian inference and Hierarchical structures to develop statistical estimation and prediction models with different complexities and objectives. The study develops and analyses five Bayesian models of increasing difficulty, two basic and three Hierarchical models, which allows us to explore safety incident data, efficiently identify anomalies, assess the level of risk, define an objective framework for comparing air carriers, and finally predict and anticipate incidents.

1. Introduction

Aviation product and service providers and safety oversight authorities around the world are moving from traditional reactive and compliance-based safety and operational approaches towards new proactive and performance-based tools and methods (ICAO, Third Edition, 2013).

As such, nowadays, aviation stakeholders not only have to monitor, gather and analyse safety-related data and information, but also must identify, anticipate and predict safety-related trends and proactively address actual and emerging safety risks. Such a shift introduces a parallel need for new safety indicators and new statistical methods suitable for modelling accidents and incidents, and their evolution.

The reporting and evaluation of less serious safety events is of prime importance in safety analysis. Statistics on less serious incidents can potentially provide a great deal of information as they are more frequent and reporting is simpler. ICAO, EASA, EUROCONTROL and the FAA (EUROCONTROL and FAA, 2012), and other safety authorities, have put a lot of effort into identifying common information and performance indicators for use in safety monitoring. Specifically, ICAO Annex 13 (ICAO, 2010) requires States to establish accident and incident reporting systems to gather information on real or potential safety shortcomings. European Regulation (EU) No 376/2014 (European Union, 2014) mandates the reporting, analysis and follow-up of incidents in civil aviation. Furthermore, the European Union developed the ECCAIRS database (European Co-ordination Centre for Accident and Incident Reporting Systems), which offers standard and flexible accident and incident data collection, representation, exchange and analysis tools.

A combination of reactive, proactive and predictive analytics is required to exploit the potential of safety data to give feedback about operational hazards and risks. In this context safety analytics and statistics need to evolve from the analysis of reactive indicators towards a more predictive perspective, to forecast future safety performance and risk.

Over recent decades, there has been substantial research on the use of modern statistical techniques to predict safety incident and accident precursors in other modes of transport, specifically road transport (Lord, et al., 2005). The statistical models commonly used include Binomial, Poisson and Poisson-gamma (or negative binomial). The tools used to develop predictive models include random-effect models (Miaou & Lord, 2003), generalised estimating equations (GEE) (Lord & Persaud, 2000), and Markov Chain Monte Carlo (MCMC) methods (Qin, et al., 2004).

Within the context of aviation, there is limited available research that compares different predictive modelling approaches, identifies the most appropriate statistical models for forecasting safety events, or helps to predict safety performance (Drees, 2014; Di Gravio, et al.,

* Corresponding author at: School of Aerospace Engineering, Plaza de Cardenal Cisneros, n3, Building B, room 224, 28040 Madrid, Spain.
E-mail address: rosamaria.arnaldo@upm.es (R.M. Arnaldo Valdés).

https://doi.org/10.1016/j.ssci.2018.01.008
Received 29 January 2017; Received in revised form 2 December 2017; Accepted 15 January 2018
Available online 03 February 2018
0925-7535/

2015; Janic, 2000; Kunimitsu Iwadare, 2015; Knecht, 2015; Panagiotakopoulos, et al., 2014).

Two modern statistical tools that are increasingly being used for predictive applications in areas such as microbiology, sociology, psychology, econometrics, structural engineering, nuclear physics, and so on, are Hierarchical structure and Bayesian models.

Relevant academic literature gives interesting examples of the types of aviation issues to which Bayesian techniques can be applied. Examples include: Spatial Analysis of Pilot Fatality Rates in General Aviation Crashes (Curriero et al., 2002); Probability of Midair Collision During Ultra Closely Spaced Parallel Approaches (Houck & Powell, 2003); Analysis of Failure Risk in Engine Rotor Disks (Enrigth and Huyse, 2003); Probabilistic Forecasts for Aviation Traffic (Bhadra & Shaufele, 2007); and A Probabilistic Influence Diagram for Landing Runway Overrun Excursion Risk Analysis (Sanchez Ayra, 2015). Object-oriented Bayesian network (OOBN) has also been used to integrate the safety risks contributing to an in-flight loss-of-control aviation accident, allowing to quantitatively drawing inferences about changes to the states of the accident shaping or causal factors (Ancel et al. (2015)).

In this research, we have adopted an innovative statistical approach as opposed to the reactive methodologies used up to now. This distinctive approach involves the use of Bayesian inference to develop statistical estimation and prediction models with different complexities and objectives. One of the most noteworthy features of Bayesian inference is that it allows us to easily develop Hierarchical Models, with differing orders of complexity, that have different objectives. It also enables us to evaluate the predictive efficacy of the models and to compare them with one another. In a discipline such as air safety, a hierarchical structure is extremely useful since it is conceptually consistent with techniques already in use in the sector, such as Bayesian networks or fault trees.

The aim of this paper is to demonstrate that such techniques are applicable to the study of aviation safety data, and that they enable us to extract information about the magnitude of risk involved and to make predictions. As such, they may be used in preventive and predictive methodologies to improve the safety conditions under which airlines operate, thereby minimising the level of risk.

In this study, we developed and analysed five Bayesian models of increasing complexity. Specifically, there were two basic and three Hierarchical Models which allowed us to:

- Explore safety incident data and efficiently identify anomalies;
- Assess the level of risk;
- Define an objective framework for comparing air carriers, and finally;
- Predict and anticipate incidents.

The models are demonstrated and applied in a specific aviation industry “Commercial Air Transport and its safety” by the use and prediction of airline's safety occurrences; although the models could be applied to any safety industry. More specifically, the results of this study are illustrated using data of monthly incidents from four fleets of different models of aircraft belonging to Company A. The incidents correspond to: *Pilots Reports* in the technical log book about breakdowns or malfunctions of any aircraft system; subsystems or components *Faults or Failures Deferred*; *In-Flight Shutdowns*, *In-Flight Turn-Backs*, *Delays and Cancellations* for Technical Reasons; *Rejected Take-Offs* for technical reasons; Non-Stabilised Approaches; and *Flight Time Limitations* exceeded.

The models can be of great utility in aviation industry for safety oversight and safety improvement and they can be applied at three different levels:

(a) By the Safety Oversight Authority to identify sources of risk in Commercial Air Transport at national level, quantify the specific risks, compare different sources of risk and, most importantly, quantify the level of uncertainty and predict the future distributions of incidents.
(b) To benchmark the performance of different air carriers as it allows identifying the parameters that are characteristic of the safe operation of each operator and of the entire system. These parameters allow identifying and quantifying trends, to establish benchmarks to compare the current year's performance with that of previous years, and to compare the performance of different companies.
(c) By the own airline to improve the safety conditions under which it operate. The airline may better evaluate the performance of the entire set of fleets of the Company regarding each category of events as well as analyse the spread of the incident rate of a fleet for a particular incident category compared to the overall rate for the entire Company. It could also identify those fleets that require most attention within a company and quantitatively compare the stress effect in one fleet compared to that in others.

The following sections explain the theory behind each of the models and the process used to design and evaluate them.

2. Model based Bayesian inference

Let us consider two possible outcomes, A and B, and assume that $A = A_1 \cup \cdots \cup_n^A$ where $A_i \cap_j^A = \varnothing$ for every $i \neq j$. Bayes' theorem gives the following expression for the conditional probability of A_i given B:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\int_{i=1}^{n} P(B|A_i)P(A_i)}$$

The above theorem can be used for inverse inference. B is the observed result, or outcome, and A_i represents the possible factors that cause B. Then, $P(B|A_i)$is the probability of B occurring when cause A_i is present, and $P(A_i|B)$ is the probability that A_i is responsible for the incident of previously observed outcome, B.

Bayes' theorem is a very useful statistical tool for extracting information from data (Bernardo, 1994). Therefore, using observed data $(y_1,y_2,\ldots,y_n)$ and its posterior distribution $f(\theta|y_1,\ldots,y_n)$, which combines both prior and observed data, we can calculate certain information. This posterior distribution is central to the idea of Bayesian inference. This differs from frequentist statistics as, in prior distribution, all unknown parameters are considered to be random variables. As such, prior distribution must be defined at the outset. Prior distribution represents the state of knowledge before collecting any data. The objective is to determine the posterior distribution, $f(\theta|y)$, of the parameters θ given a set of observed outcomes.

Posterior distribution integrates both prior and observed data and information, and is a combination of prior distribution, $f(\theta)$, and probability, $f(y|\theta) = \prod_{i=1}^{n} f(y_i|\theta)$. Both distributions must then be completely specified to complete the Bayesian model.

The whole modelling procedure may be divided into four stages:

i. Construction of model;
ii. Calculation of posterior distribution;
iii. Analysis of posterior distribution, and
iv. Inference.

The final step is to use the model to predict probable outcomes. Bayesian inference offers an accurate and straightforward means of predicting future outcomes via calculated predictive distribution (Gelman, et al., 2013).

3. Safety incidents data and information

This study is based on the analysis of safety incidents that occurred to air operators and were registered in a national Mandatory Incident Reporting (MOR) system. The events or incidents that occur during aircraft operations are usually precursors to more serious accidents that may occur in the future (Brooker, 2005; Margaryan & Littlejohn, 2016; Gnoni & Homer, 2017).

By studying these incidents, we are able to obtain information about the operational risk associated with different areas of air transport, identify dangerous trends in the performance of operators, and establish preventive measures. When we analyse the information gathered we can identify current risks to safety and detect emerging risks. We are also able to identify changes in trends and predict how these will evolve. More importantly, this information will guide us when making improvements to safety. It is, therefore, essential to develop analytical methodologies that have predictive capacity, such as those proposed in this study. These provide an anticipatory and well-documented approach, which will help us to identify and adopt adequate safety measures.

The results of this study are illustrated using data of monthly incidents from four fleets of different models of aircraft belonging to generic Company A. The data came from the following sources:

1. Pilot Reports: Notes made by pilots in the aircraft technical log book to inform maintenance personnel about breakdowns or malfunctions of any aircraft system.
2. Deferred Items: Faults or failures of subsystems or components whose repair has been deferred for a period not greater than that defined in the minimum equipment list (MEL) of the aircraft.
3. In-Flight Shutdown (IFSD): when an engine ceases to function (when the airplane is airborne) and is shutdown, whether self-induced, flight crew initiated or caused by an external influence, for example, flameout, foreign object ingestion, or inability to obtain or control desired thrust or power.
4. In-Flight Turn-Back (IFTB): Any operation in which the aircraft has been forced to land at a destination airport other than the one originally scheduled, due to technical reasons.
5. Delays and Cancellations for Technical Reasons: Delays exceeding 15 min with respect to the expected time of departure or flight cancellations, both for technical reasons.
6. Rejected Take-Offs (RTO) for Technical Reasons: Rejected take-off performed after the aircraft has started the take-off roll and its speed has exceeded 90 knots.
7. Non-Stabilised Approaches: Approaches that do not meet all the criteria of a stabilised approach, as defined by the Flight Safety Foundation, with regard to minimum height, rate of descent, location on flight path, landing configuration, and so on.
8. Flight Time Limitations (FTL) exceeded: Number of occasions on which the Flight Time Limitations (FTL) of the technical crews were exceeded.

The five Bayesian Models developed in this study are computationally feasible using Markov Chain Monte Carlo (MCMC) simulation based on Gibbs sampling with BUGS (Bayesia*n* inference Using Gibbs Sampling).

4. Development of the Gamma-Poisson model: the basic model for estimating the rate of incidents

When modelling safety events, the most commonly used approach is to approximate the probability of a specific event occurring using a Poisson distribution, defined by a parameter λ. This parameter represents the risk expressed as average rate of incidents, a ratio such as *No. of safety incidents/No. of operations* $\left(\frac{No.\ of\ safety\ incidents}{No.\ of\ operations}\right)$ or, failing that, *No. of safety incidents/No. of flight hours* $\left(\frac{No.\ of\ safety\ incidents}{No.\ of\ flight\ hours}\right)$.

Given a series of n data points where $D = y_1, y_2, y_3, \ldots, y_n$, and a parameter to be estimated λ, the density function of the Poisson distribution $Po(\lambda)$ may be expressed as $Po(\lambda) = p(y/\lambda) = \frac{e^{-\lambda}\lambda^{y}}{y!}$.

The model consists of the following elements:

- **Likelihood function:** This function is the product of the likelihood functions for each separate event, and is equivalent to a single event modelled by a Poisson distribution function, $Po(n\lambda)$:
- $L(\lambda) = p(y_1, y_2, \ldots, y_n/\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{y_i}}{y_i!}$, *where* $y_i = incident_i/operation_i$
- **Function representing prior knowledge about the parameter being studied:** In the specific case of a process that generates events which are modelled using a Poisson distribution, the conjugate prior function (Lynch, 2007; Kruschke, 2010) must be a gamma distribution of the generic parameters $\alpha$and β:
- $p(\lambda) = f(\lambda, \alpha, \beta) = \frac{\beta^{\alpha}\lambda^{\alpha-1}e^{-\beta\lambda}}{\Gamma(\alpha)} = C\lambda^{\alpha-1}e^{-\beta\lambda} \propto \lambda^{\alpha-1}e^{-\beta\lambda}$
- **Posterior function:** The posterior function is proportional to the product of the likelihood function and the prior density function.
- $p(\lambda/y_1, y_2, \ldots, y_n) \propto e^{-n\lambda}\lambda^{\sum_i^y}\lambda^{\alpha-1}e^{-\beta\lambda} \propto e^{-(\beta+n)\lambda}\lambda^{(\alpha+\sum_i^y)-1}$
- This expression can be rearranged in the form of a gamma density function with parameters β^* y α^*, such that: $\alpha^* = \alpha + \sum_{i=1}^{n} y_i$; $\beta^* = \beta + n$; where $\alpha$and β are the values representing the prior belief and n is the total number of data points to which the model has been applied.
- **Posterior predictive function:** The posterior predictive function represents the probability of a certain number of incidents occurring in the future, assuming that the number of operations that there will be that month is known. This may be expressed as:

$$p(y_{new}|y_1, y_2, \ldots, y_n) = \int_{\lambda} p(y_{new}|\lambda n_{new}) p(\lambda|y_1, y_2, \ldots, y_n) d\lambda$$

where $p(\lambda|y_1, y_2, \ldots, y_n)$ is the posterior distribution; $p(y_{new}|\lambda n_{new})$ the Poisson distribution of parameter λn_{new}; and y_i the number of incidents in a given month.

4.1. Choice of prior function

The information supplied to the model, before applying it to the data being studied, is represented by the values α and β of the Gamma distribution. There are several alternatives when selecting these parameters.

- **Expert opinion:** Expert opinion on the mean and the variance of the parameter being studied can be modelled using the parameters α and β using the following expressions: $\boldsymbol{E}[\lambda] = \frac{\alpha}{\beta}$; $\boldsymbol{Var}[\lambda] = \frac{\alpha}{\beta^2}$.
- In this case the information contained in the resulting gamma distribution is similar to that which could be extracted from a sample of $neq = \beta$ data of the Poisson distribution (Lynch, 2007).
- **Non-informative functions:** If there is no prior belief about the value of λ, then a non-informative function must be chosen which represents the lack of knowledge about the value of parameter λ. In this study, we tested two non-informative functions:

I. **Non-informative uniform prior:** This distribution is the simplest possible and accurately represents the lack of information regarding the value of the parameter λ. It gives the same weight to all values of $\lambda \geqslant 0$: $f(\lambda) = 1; \lambda \geqslant 0$

The resulting posterior distribution is given by the following expression:

$$p(\lambda/y_1, y_2, \ldots, y_n) \propto p(y_1, y_2, \ldots, y_n|\lambda)p(\lambda) \propto e^{-n\lambda}\lambda^{\sum y_i}$$

Identifying the terms with those of the gamma probability density function we get the following changes in the hyperparameters: $\alpha^* = 1 + \sum_{i=1}^{n} y_i$; $\beta^* = n$;

The uniform prior function may also be defined as a gamma probability density function with parameters $\alpha = 1$; $\beta = 0$.

II. **Jeffreys' prior:** This distribution has been extensively studied in the field of Bayesian inference as the expression depends on the likelihood function only (Jeffreys, 1946). This is given mathematically as:

$p(\theta) \propto \sqrt{\det\,(I(\theta))}$ where $I(\theta)$ is the Fisher information.

Fisher information is a measure of the information that can be extracted from the likelihood, and gives a function of its curvature. For a single parameter and a Poisson likelihood function, the Jeffreys' prior would be: $p(\lambda) = 1\sqrt{\lambda}, \lambda > 0$.

The posterior distribution is given by the following expression:

$$p(\lambda/y_1,y_2,...,y_n) \propto p(y_1,y_2,...,y_n|\lambda)p(\lambda) \propto e^{-n\lambda}\lambda^{-\frac{1}{2}+\sum y_i}$$

By identifying the exponents in the above expression with the parameters of a gamma distribution, we get: $\alpha^* = \frac{1}{2} + \sum_{i=1}^{n} y_i$; $\beta^* = n$;

Jeffreys' prior can also be defined as a gamma distribution of parameters $\alpha = \frac{1}{2}$ and $\beta = 0$;

Jeffreys' prior is not strictly non-informative in this case, since it gives slightly more weight to smaller values of λ. However, this is not a valid reason for ruling out its use since the values of λ obtained by calculating the ratio of *No. of safety incidents/No. of operations* $\left(\frac{No.\ of\ safety\ incidents}{No.\ of\ operations}\right)$ are, in almost all cases, less than one.

The uniform function provides a simple, intuitive solution. On the other hand, Jeffreys' prior is popular in the academic field because it is invariant under reparameterisation of the parameter vector θ (Canfield and Teed (1977)). We have decided to use both in the model, and compare the results obtained. In so doing, we can verify the robustness of the posterior function to changes in the prior function.

An initial analysis of the posterior gamma distributions, resulting from applying both prior functions (Fig. 1), shows how sensitive the model is to the choice of prior function. The posterior distribution function given by the Jeffreys' prior is slightly more focused on lower values of λ, since it gives more weight to small values as it is inversely proportional to square root of λ. In this example, the Jeffreys' prior gives a median value of 0.1055 while the Uniform prior gives a value of 0.1206. When the number of data points is small, the functions give widely differing values. However, once there are more than approximately ten data points, both posterior functions start to give similar values, and when there are about thirty data points the results are very similar.

4.2. Results of the Gamma-Poisson model

Fig. 2 gives a graph of estimated values of parameter λ. The posterior function resulting from a Jeffreys' prior is Gamma (4.3128, 35), while the posterior function resulting from a Uniform prior is Gamma (4.8128, 35).

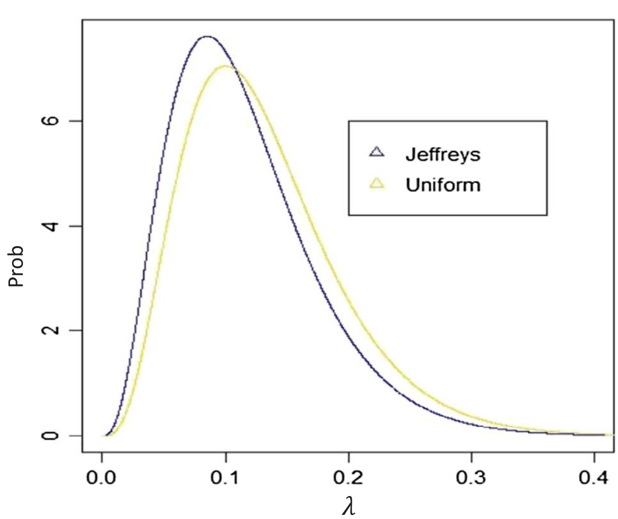

Fig. 1. Comparison of the posterior functions resulting from Jeffreys' and Uniform priors using data from Pilot Reports for Fleet 1 of Company A between January 2009 and January 2012.

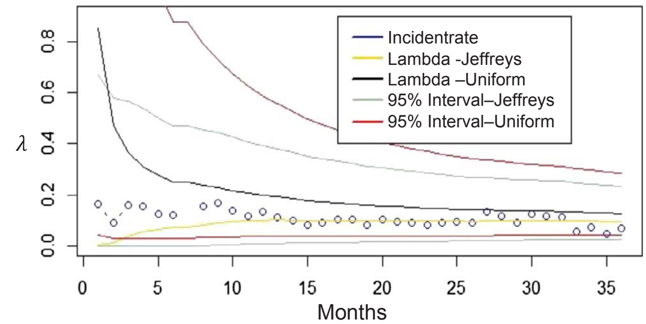

Fig. 2. Estimate of parameter λ for Fleet 3 of Company A, using the Gamma-Poisson model.

The above graph gives the values of the incident rate as per the Pilot Reports. It also shows how the median and 95% confidence intervals of parameter λ converge towards the estimated value as more data points are added.

The Jeffreys' function appears to give a better fit to the data than the Uniform function. This is because the Jeffreys' prior, although it is non-informative, gives more weight to the possibility that the values of λ will be small which, in this case, turns out to be true. The 95% confidence limits are not symmetrical, nor is the posterior function. Furthermore, the lower limit is zero as values of the incident ratio tend to be relatively small.

4.3. Results predicted by the model

In this section, we compare actual data with predictions made based on the number of incidents. We use a posterior predictive $Poisson(n_i\lambda)$ distribution, according to the expression:

$$p(y_{new}|y_1,y_2,...,y_n) = \int_\lambda p(y_{new}|\lambda n_{new})p(\lambda|y_1,y_2,...,y_n)d\lambda$$

Fig. 3, below, shows the situation in Month 20 using Jeffreys' prior. The red line represents the actual value for Month 20 based on the Pilot Reports. The vertical blue lines represent the median (in the middle) and the limits of the 90% confidence interval. In this case, the model predicted 61 Pilot Reports while the real value was 55. The approximation is quite good, however, the 90% safety interval ranges from 22 to 128, which is very broad.

Fig. 4 gives the results for Month 20 using the Uniform prior. Once again, we see that the predicted values of rate of incidents per operation tends to be higher when using the Uniform prior. In this case, the median is 69, and the 90% confidence interval ranges from 27 to 138.

Fig. 5 shows the results predicted by the model and allows us to make an initial qualitative evaluation of the model. The predicted values, as represented by the median of the predictive functions obtained, are in line with the actual data. However, the 90% confidence interval would appear to be too big. There is no value outside the confidence interval when, according to the definition of the 90% interval, approximately 3–4 values should be outside.

The explanation is that for the Poisson model to represent these

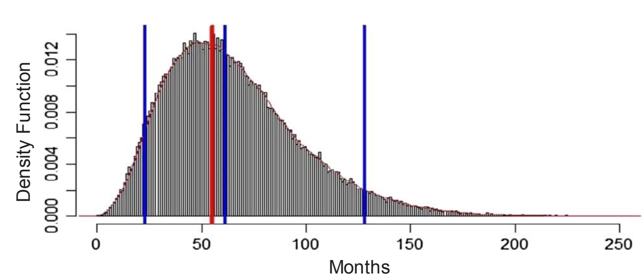

Fig. 3. Prediction for Month 20 using Jeffreys' prior.

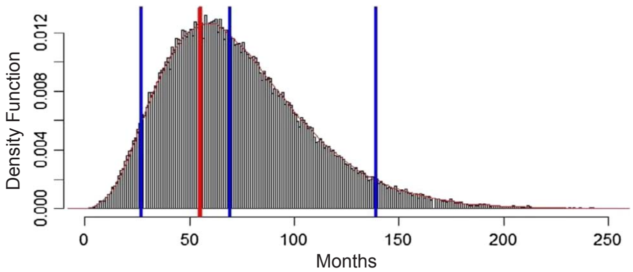

Fig. 4. Prediction for Month 20 using the Uniform prior.

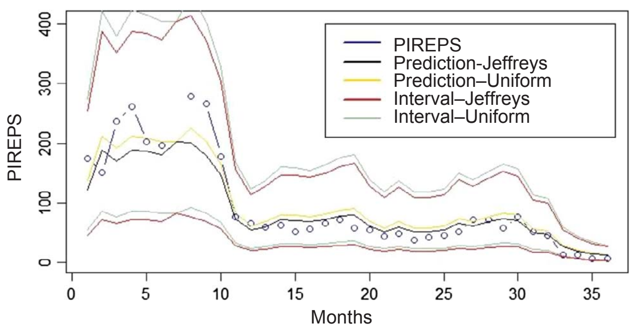

Fig. 5. Estimate of the number of PIREPS.

effects, the implicit assumption is that the variance of the process must be approximately of the same order as the expected value. Therefore, the confidence interval narrows as the number of incidents decreases. As we will see later on, the confidence interval gives a good fit when the data points are low in value and there is little variation between them.

Another significant conclusion, from this example, is that the actual value of the incident rate has varied throughout the study. This is evidenced by the fact that a model based on a Uniform prior distribution gives a better fit when the number of incidents is high. Conversely, a model based on a Jeffreys' prior gives a better fit when the number of incidents is low.

On analysing the data, we see that the fleet decreased in size from nine aircraft to one aircraft from January 2009 to January 2012. Also, if we look at the rate of Pilot Reports per operation over the same period, we see that the rate decreased progressively from 0.16 Pilot Reports per operation to 0.07.

A possible explanation for the lack of seasonality in the incident rate, is the effect of stress. Due to the stress effect, the company could experience a higher number of incidents per operation when the number of aircraft and operations increases. As a result, the value of parameter λ would not be constant.

To safeguard against this effect, we studied a data sample in which the number of aircraft remained constant. This was the data from the Pilot Reports for Fleet 1 over the same period of time. The results are shown in Fig. 6.

In Fig. 6, Jeffreys' prior gives a more accurate prediction over the

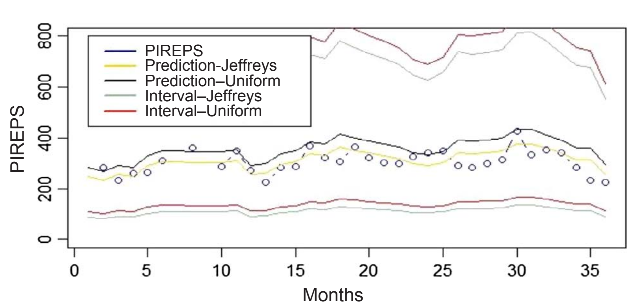

Fig. 6. Estimate of the number of Pilot Reports for Fleet 1.

time period studied. However, the number of incidents is quite large for Poisson distribution (200–400). This is reflected in the wide range of the 90% confidence interval. An increase in the variance when the number of events is high is a characteristic of the Poisson model. Therefore, we can see that the Gamma-Poisson model is useful for predicting the number of incidents using the median of posterior predictive distribution. However, the confidence intervals are overstated. As such, it seems reasonable to combine this model with others, or to extend the model so as to optimise the amount of information that can be extracted from the available data.

5. Development of the Beta-Binomial model

An obvious alternative to the Poisson model is the Beta-Binomial model. This compares the number of safety incidents that an airline experiences in a month with the positive results of n experiments carried out that month. Each of these n experiments corresponds to an operation performed that month.

The problem may then be restated by establishing that the number of incidents corresponds to a $Binomial(n_i,\theta)$ distribution, where parameter θ represents the probability of each of the n trials ending in a positive result, that is, a safety incident.

The model consists of the following elements:

- **Likelihood function:** The Binomial distribution is given by the expression: $y_i \sim Binomial(n_i,\theta) = \binom{n_i}{y_i}\theta^{y_i}(1-\theta)^{n_i-y_i}$

Therefore, the likelihood function, for the total number of months, is: $\mathscr{L}(\theta) \propto C\theta^{S}(1-\theta)^{N-S}$ donde $S = \sum_{i=1}^{m} y_i$, $N = \sum_{i=1}^{m} n_i$

- **Prior function** The corresponding conjugate distribution of the Binomial function is the Beta function (Glickman & Dyk, 2007): $p(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1} \propto C\theta^{a-1}(1-\theta)^{b-1}$

The following Beta functions are used as priors: Beta (0.5,0.5), Beta (1,1), and Beta (1,3) (Dyer & Chiou, 1982).

- **Posterior function:** The posterior function also reduces to a Beta function.

$$p(\theta|D) \propto C\theta^{S}(1-\theta)^{N-S}\theta^{a-1}(1-\theta)^{b-1} = \frac{\theta^{S+a-1}(1-\theta)^{N-S+b-1}\Gamma(a+b+N)}{\Gamma(a+S)\Gamma(b+N-S)}$$

The posterior function is equivalent to the prior Beta function with its parameters a and b modified as follows: $a^* = a + S$; $b^* = b + (N-S)$. Parameter a is the number of positive results of the "theoretical experiment" (in this case, incidents). Every time there is a new incident the value of this parameter increases. Parameter b is the number of failed experiments each month ("theoretical failure" or number of operations that occurred without incident).

5.1. Estimation of parameter θ

Fig. 7 shows that the posterior function of parameter θ is the same regardless of the prior function chosen. Therefore, it is true to say that this model is much more robust than the previous model as regards the choice of prior function. Furthermore, this model also limits the range of plausible values of θ, which will lead to more accurate predictions for the posterior predictive function.

The characteristic percentiles of the posterior functions, shown in Table 1 below, vary considerably. In the case of the Binomial posterior distribution the difference in values is approximately 0.004, whereas in the Gamma-Poisson models the range is 0.22 and 0.25.

Looking at the process for estimating parameter θ, shown in Fig. 8, we can see that there is a greater degree of convergence and that the

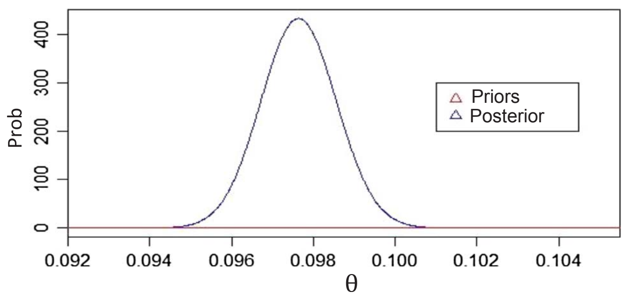

Fig. 7. Beta-Binomial posterior distribution of parameter θ for the Pilot Reports of Fleet 1.

Table 1
Comparison on the confidence intervals between the Beta-Binomial and the Gamma-Poisson models.

	P2.5	P50	P97.5
Poisson Jeffreys'	0.0301	0.1055	0.2569
Poisson Uniform	0.0380	0.1206	0.2797
Binomial	0.0958	0.0976	0.0995

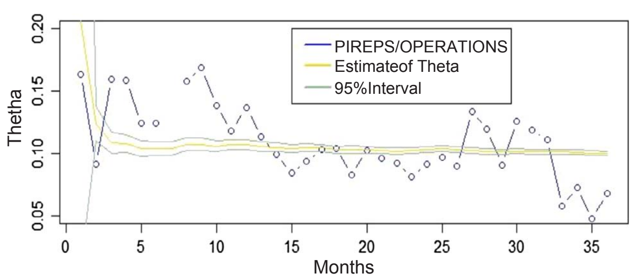

Fig. 8. Estimation of θ.

convergence is much faster than with the previous model. It is also obvious that, in this case, the 95% interval does not contain 95% of the Pilot Reports.

5.2. Predictions of the model

Fig. 9 shows the prediction for Month 20 calculated using the Beta-Binomial model. This can be compared with the results for the same month predicted using the Poisson function: shown in Figs. 3 and 4.

Comparing both models, we see that the graph of the Gamma-Poisson model with Uniform prior distribution has a wider spread. The Beta-Binomial model, on the other hand, gives a more accurate result. Nevertheless, it is true that the medians of both distributions are practically identical. Therefore, both models make good predictions. Fig. 10 shows that the Beta-Binomial model gives a very good fit for the entire data series, and not just for Month 20. In this case, the confidence interval is much narrower and gives a better fit to the data.

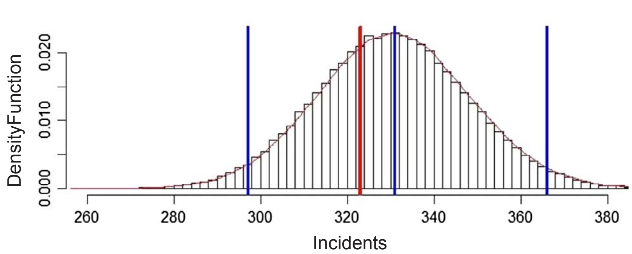

Fig. 9. Prediction for Month 20 using the Beta-Binomial model.

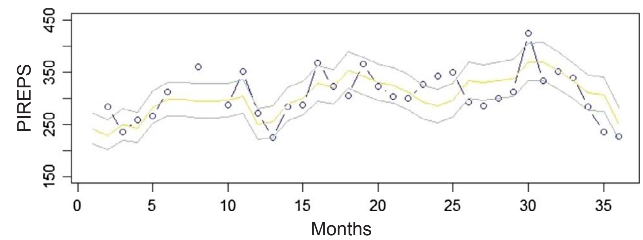

Fig. 10. Estimation of the number of Pilot Reports using the Beta-Binomial model.

6. Development of hierarchical models

The data indicates that the incident rates of a parameter may vary depending on the type of fleet. Basic Gamma-Poisson and Beta-Binomial models do not allow us to evaluate this relation; to do so we require Hierarchical Models. A Hierarchical Model entails two parts:

i. A hierarchy of levels;
ii. Bayesian models to estimate the parameters.

A Hierarchical Model consists of a number of sub-models which are integrated, into one overall model, using the Bayes theorem. The Hierarchical Model enables us to calculate the values of the different parameters and the uncertainty of these estimates, using MCMC (Markov Chain Monte Carlo) methods (Allenby et al., 2005). In this study, we developed three Hierarchical Models.

6.1. Hierarchical Model 1

Model 1 has three levels, as shown in Fig. 11:

- Number of monthly incidents, expressed as a *Binomialdistribution*(n_i,θ_j) where subscript i represents the month and subscript j identifies the fleet.
- Distribution of θ_j, expressed as a *Beta distribution*(a_j,b_j).
- Parameter μ, which defines the incident rate of an event for all the fleets of a company (Gelman, et al., 2013).

θ_j is defined by the *Beta distribution*(a_j,b_j). Analysis of the first and second moments of a *Beta distribution*(a,b) gives us:

$$a = \mu K, b = (1-\mu)K \text{ with } k = \left[\frac{\mu(1-\mu)}{Var} - 1\right]$$

where parameter K is a measure of the dispersion of parameter μ for a specific fleet, and parameter μ is the rate of a certain incident for the entire company, in other words, all fleets. Parameter K is inversely proportional to the variance (Var) and, therefore, is a good indicator of the dispersion of μ characteristic of each fleet. The higher the value of K, the lower the variance and, therefore, the nearer the value of parameter θ_j for a specific fleet to the mean value of the parameters (μ) for the company. Conversely, the lower the value of K, the greater the dispersion of θ_j.

As previously stated, there is only one parameter μ for Company A, in other words, the value of the parameter is the same for all fleets. On

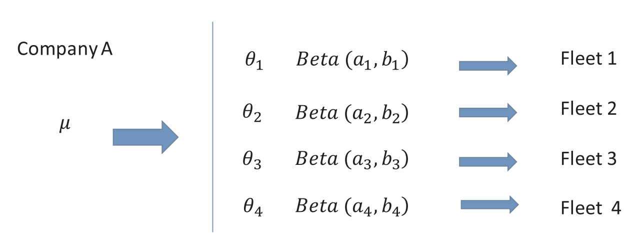

Fig. 11. Hierarchical Model 1.

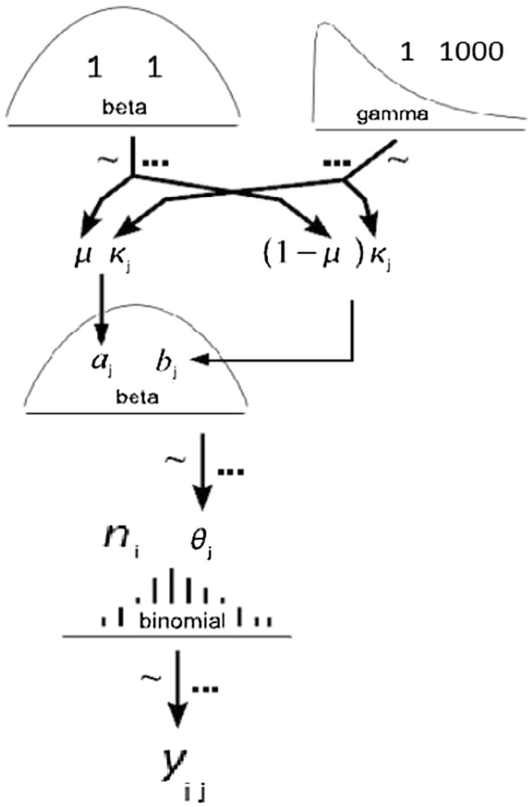

Fig. 12. Structure of Hierarchical Model 1.

the other hand, each fleet has its own value of the dispersion parameter, K, which is identified as K_j.

According to the literature the prior distribution of μ is characterised by a Beta (1,1) distribution. The prior distribution of K_j is expressed as Gamma (1,0.001) functions, (Kruschke, 2010).

Fig. 12 shows the Hierarchical Model and the interrelations between the different levels. The upper level consists of, so-called, hyper-prior distributions, which model the hyper-parameters μ and K_j. The second level comprises the prior Beta distributions that characterise each of the parameters θ_j. Finally, the third level is made up of Binomial likelihoods that estimate the monthly events for a specific fleet, y_{ij}, based on the parameters θ_j and the number of operations, n_i, in the month in question.

6.1.1. Results given by Hierarchical Model 1

Fig. 13 shows the distributions of parameter μ, for the four fleets of Company A, under the following headings: *Pilot Reports, Deferred Items, Delays and Cancellations for Technical Reasons, Non-Stabilised Approaches and Rejected Take-Offs for Technical Reasons.*

As can be seen in the above figure, the model identifies the characteristic orders of magnitude for each category of incidents. As such, we can compare these with the orders of magnitude of incidents of other companies. The model also allows us to identify any changes in the value of the incident rate for the company as a whole. This change is evident in the distribution functions for "Deferred Items" and "Delays and Cancellations", each of which has two clearly defined maxima.

Fig. 14, below, shows the values of parameter K_j. K is specific to each fleet and measures how much the behaviour of an individual fleet deviates from the characteristic parameter of the category of incidents (μ) in question. The greater the value of K for a fleet, the closer the incident rate per operation for that fleet, parameter θ, to the characteristic parameter, μ, of the company.

The figures show that the median values of K_j are generally high. This agrees with the hypothesis that incident rates for the different fleets (θ_j) are of the same order and similar to μ. Exceptions to this are found in the category Pilot Reports for Fleets 3 and 4, in which the values of parameter K are much lower than elsewhere. Extreme values

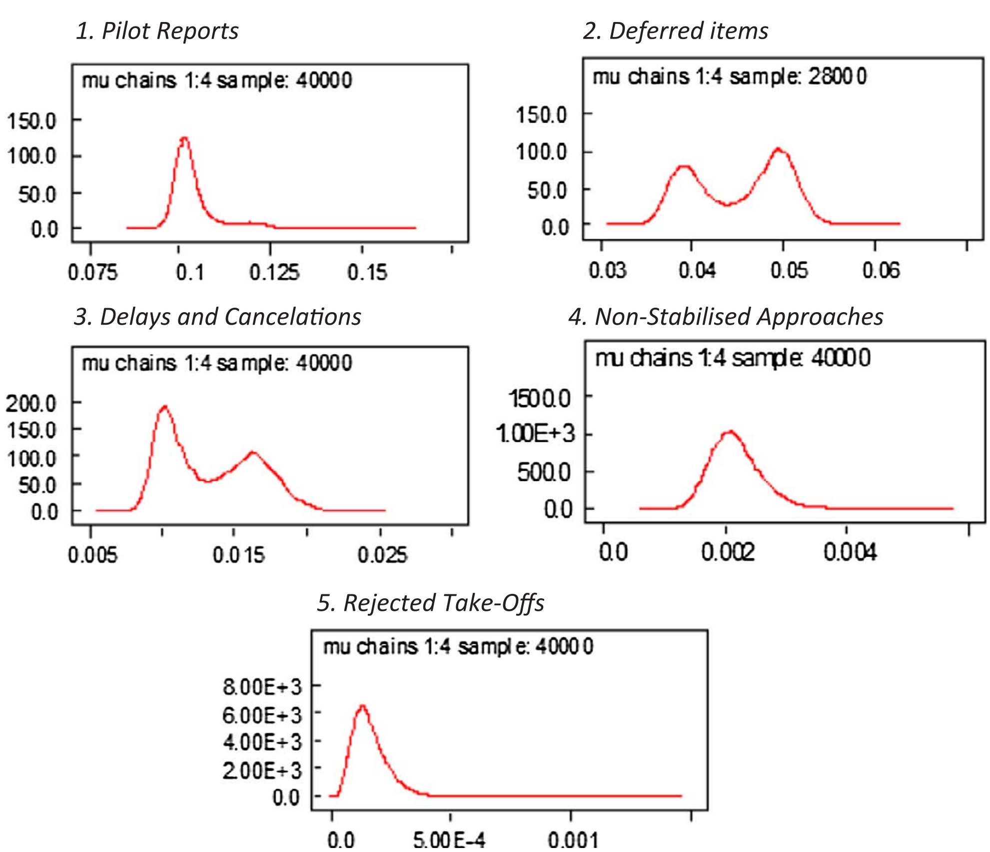

Fig. 13. Estimation of parameter μ using Hierarchical Model 1.

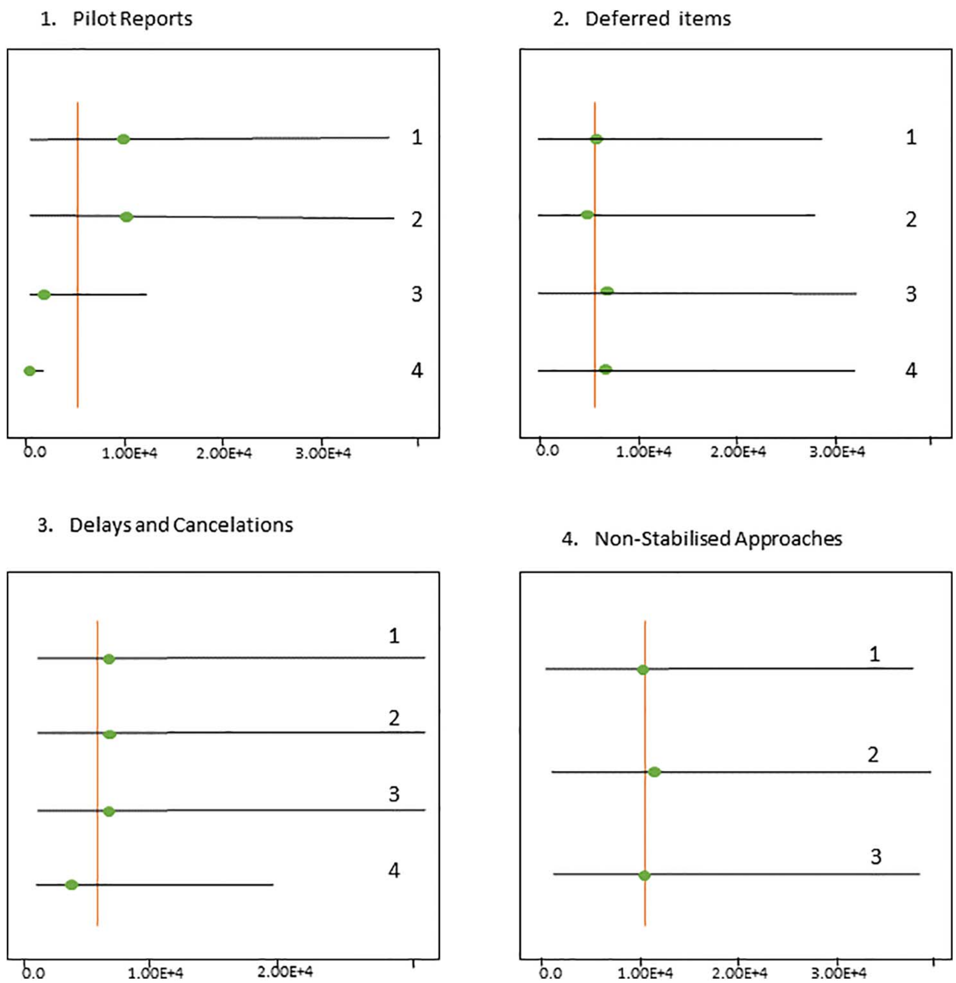

Fig. 14. Estimation of parameter K_j using Hierarchical Model 1.

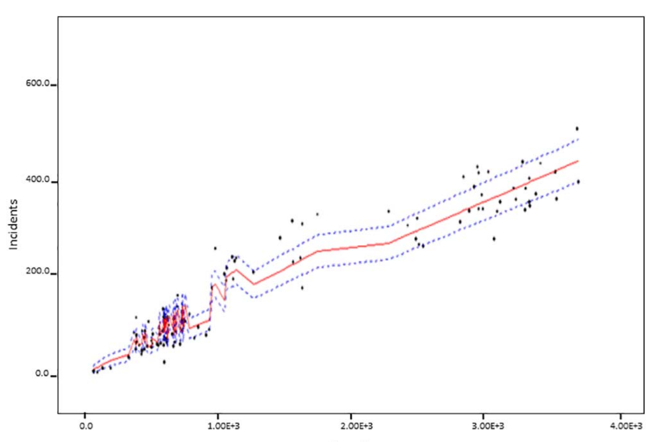

Fig. 15. Predicted incident of Deferred Items using Hierarchical Model 1.

of K_j, that is, values that are unusually high or low, compared to the values normally found within the Company A, can be useful in identifying atypical incident rates in fleets.

Fig. 15 gives the predicted Deferred Items. The initial data is shown in black; the 95% confidence interval is shown in blue; and the mean of the predicted posterior distribution is in red.

Hierarchical Model 1 gives a good fit to the data and the predictions it makes are similar to those of the simple Gamma-Poisson and Beta-Binomial models. It is a very useful tool for evaluating and comparing fleets, types of incidents and different companies.

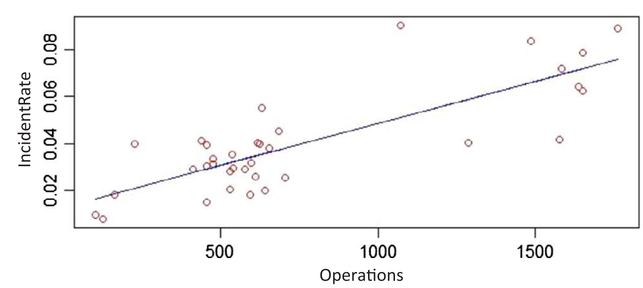

Fig. 16. Evolution of the rate of Differed Items as a function of the total number of operations.

6.2. Hierarchical Model 2

This aim of this model is to identify the stress effect, in other words, to determine if the incident rate is influenced by the number of operations and the number of aircraft. Recent studies have identified the stress effect and indicate that the relationship between accident frequency/rate and Total Flights Hours for General Aviation can be represented by a log-linear model (Bazargan & Guzhva, 2007; Knecht & Smith, 2012).

The data used in this study provides evidence of the stress effect. In Fig. 16 we see that, in the case of *Deferred Items* for Fleet 3 of Company A, the incident rate per operation increases with the number of operations.

The model assumes a linear relationship between incidents and operations via parameter β, which is given by the expression: $\beta = \alpha_1 * n_i + \alpha_2$.

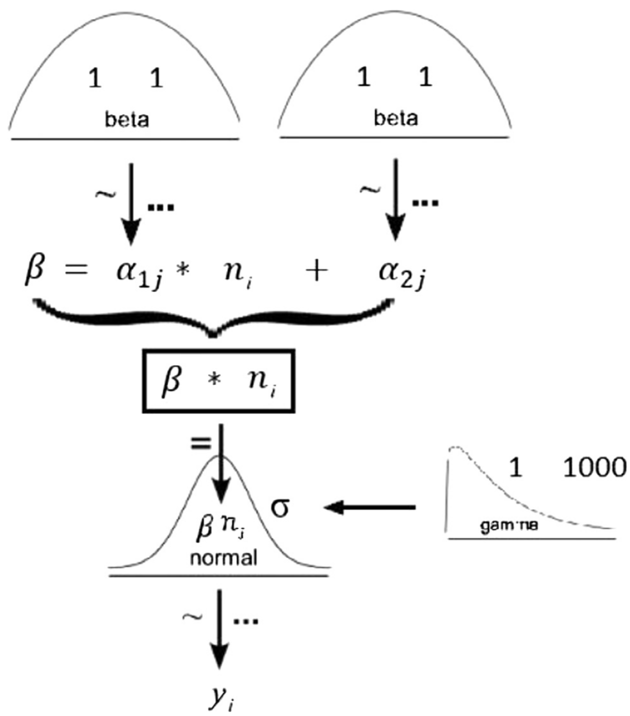

Fig. 17. Structure of Hierarchical Model 2.

Therefore, each fleet may be characterised by two parameters, α_1 and α_2, to which we assign prior Beta (1,1) distribution. The incident data was also modelled as a $Normal(\beta n_i,\sigma)$ distribution, as proposed by Breslow (Breslow & Clayton, 1993). The proposed structure for Hierarchical Model 2 is indicated in Fig. 17.

6.2.1. Results of Hierarchical Model 2

Fig. 18 shows the quartiles and medians of α_1 for every fleet, in four incident categories. As we can see, the values are very small, typically $< 10^{-5}$. The results of the model are quite accurate because for Fleet 1, in which the number of aircraft is constant, the value of α_1 is very close to zero.

The morphology of the posterior distributions of this parameter varies and depends both on the characteristics of the fleet and on the category of event being analysed. The graphs show that the less likely the event, the greater the similarity between the distributions. This is because it is difficult to quantify the stress for events that are highly unlikely. Fig. 19 gives the posterior distribution of α_1, for 4 incident categories.

The other new parameter, α_2, may be considered to be analogous to the incident rate, as defined for the other models. As such, the less the incident rate is influenced by the number of incidents, the greater the similarity between $\alpha_2$and the incident rates calculated using the other methods.

The posterior distributions also have different shapes, although in all cases a non-informative Beta prior was used. A representative selection of posterior distributions is given in Fig. 20, below.

Fig. 21 compares the initial data and the results predicted by the model for Deferred Items.

The model clearly gives a better fit to the data for this category of event. It is useful when there is a change in the incident rate for all fleets.

7. Hierarchical Model 3

This model is a modification of Hierarchical Model 2. The aim is to produce a model that combines:

– The smooth functioning of a normal distribution (by using a quadratic regression to establish confidence intervals, especially when there are many incidents per month), and
– The capacity of Poisson and Binomial distributions to make the variance proportional to the number of incidents.

The model makes the variance proportional to the number of operations by estimating a proportionality parameter, M_j, for each fleet. When the variance of a fleet is constant for the entire range of incidents, a Gaussian noise component is added to the regression, so that parameter M_j can be set to 0, if required. The expression would then be as follows: $\beta = \alpha_1 * n_i^2 + \alpha_2 * n_i + \varepsilon_j$

This is combined with a structure similar to that in Hierarchical Model 1. The difference is that the lower-level function of the model uses normal distribution instead of discrete gamma or binomial functions.

The structure of this model is shown in Fig. 22. It is a combination of the two previous models. The lower-level function uses a normal distribution that is dependent on the number of operations in order to simulate the variable variance of binomial and Poisson distributions. However, the model caters for situations where, according to the data set in question, it is possible that there is no such dependence.

The function of parameters μ and K_j is very like Hierarchical Model 1 and, therefore, both parameters can be useful in drawing conclusions. It is also worthwhile studying the values of ε_j since this parameter indicates if a fleet has too much spread, in other words, whether there is greater uncertainty with respect to the fleet's incident rate compared to that of other fleets, for a particular incident category.

7.1. Results of hierarchical model 3

Fig. 23, below, gives the fit and predicted values for the category Deferred Items.

The main advantage of this model, compared to the previous ones, is that it the confidence levels show a very good fit to the data and the predictions are, in all cases, in line with the initial data.

The "Deferred Items" of Company A is specific case where there is a drastic change in the incident rate, despite which the model still gives a good fit.

Hierarchical Model 3 has many parameters. This means that it is flexible and capable of quantifying the data spread. However, having that many parameters means that some of them do not have that same significance that they had in Hierarchical Model 1 and are more difficult to interpret.

Fig. 24 shows the posterior distributions of the estimates for parameter μ.

The distributions are of a similar order of magnitude to those of parameter μ in Hierarchical Model 1. However, the distributions are less precise than those in Hierarchical Model 1 with a greater degree of error. They also do no provide as much useful information. For example, the distributions of μ no longer capture the changes in the parameter for Categories 2 and 3.

Moreover, this model gives lower values of parameter K than those obtained using Hierarchical Model 1. The probable explanation is that μ, K and ε are, in some way, related.

8. Comparison of the models

In this study, we carried out two comparisons of the models: one graphical or qualitative, and the other quantitative.

8.1. Qualitative comparison of the models

The aim of comparative graphical analysis is twofold: the first is to compare the fit of the distributions to the data, and the second is to

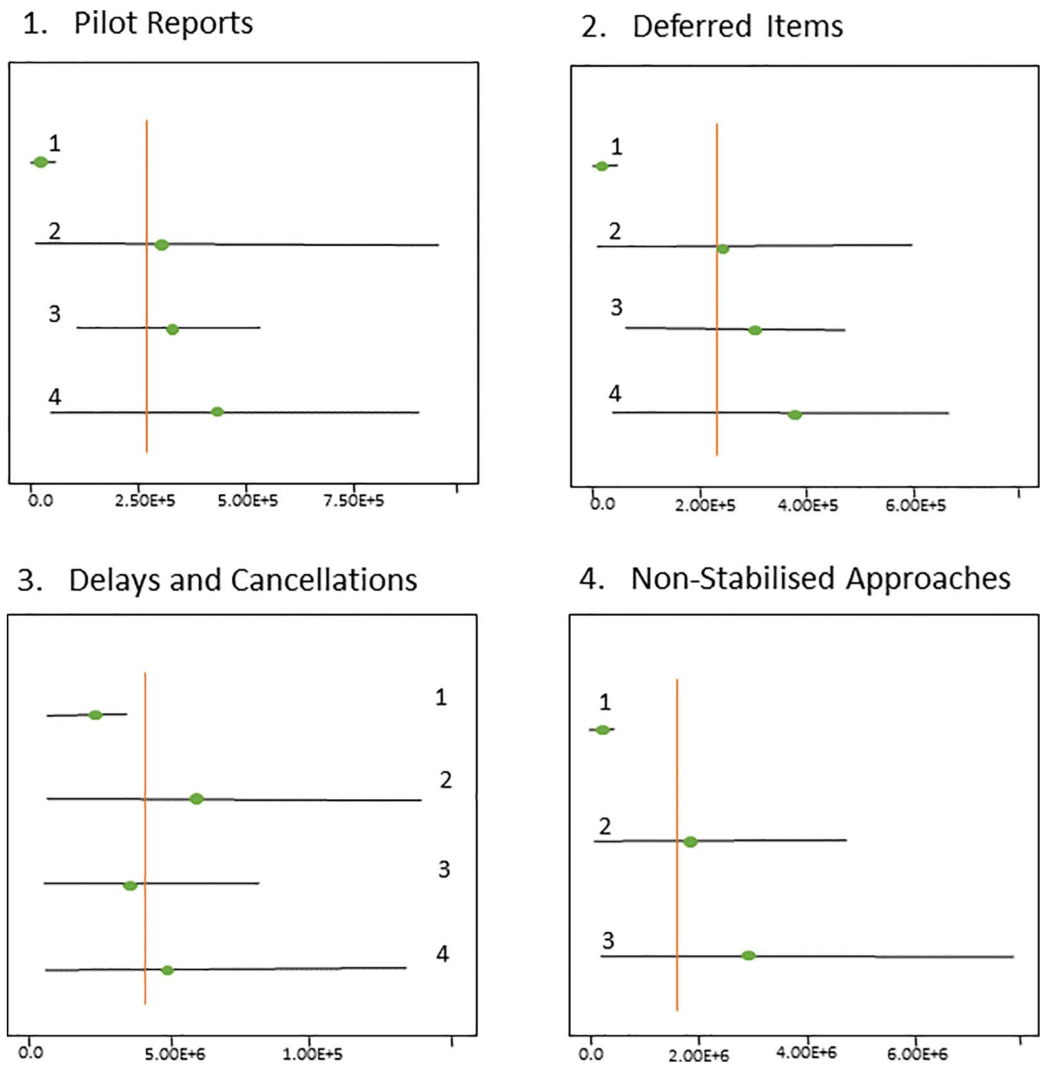

Fig. 18. Estimation of $\alpha_1$using Hierarchical Model 2.

compare the estimated values of the parameters.

The most significant parameter, as regards the level of safety of a fleet with respect to a certain category of incidents, is the incident rate per operation. This rate is equivalent to the parameter λ_j of the Gamma–Poisson model and the parameter θ_j of the Beta-Binomial model and Hierarchical Model 1. With slight variations, the incident rate per operation, parameter α_{2j} of Hierarchical Model 2, and parameter δ_j of Hierarchical Model 3 may also be considered to be equivalent. These parameters can be compared using a box diagram as shown in Fig. 25, for *Pilot Reports*.

The discrete models: Gamma-Poisson, Beta-Binomial and Hierarchical Model 1, yield similar percentiles. The Beta-Binomial models, that is, the simple Beta-Binomial model and Hierarchical Model 1, give a narrower band of values for the incident rate. Parameter α_{2j} of Hierarchical Model 2 follows the same trend as the incident rates of the discrete models, however, it gives a wider spread and the values are two one-hundredths lower than those given by the discrete models. These two one-hundredths are added back, later, via a parameter that measures the quadratic component of the fit.

On the other hand, parameter δ_j of Hierarchical Model 3 has a much wider spread, since it is affected by the noise component, ε_j, of each fleet. This model gives a significantly different trend to those of the other models, however, it has a similar order of magnitude.

In Figs. 6, 10, 15, 21 and 23 we see that the predicted fit of the non-hierarchical models and Hierarchical Model 1 are worse than those of the hierarchical models with normal distribution (Hierarchical Models 2 and 3). The latter take the stress effect, that may be experienced by some fleets, into consideration.

The hierarchical models with normal distribution (Hierarchical Models 2 and 3) work best when the data series contain sufficiently frequent incidents. On the other hand, the Gamma-Poisson and Beta-Binomial models have the advantage that the variance is of the same order as the expected value. In such cases, the higher the value of the incident rate, the wider the confidence interval.

8.2. Quantitative comparison of the models

We will use the DIC (Deviance Information Criterion) to compare the models using Bayesian statistical tools (Spiegelhalter, et al., 2002). This parameter, which is easily calculated using Monte Carlo simulations with Markov chains, computes the deviation as the mean of the posterior predictive function:

$$DIC = -2\ln(p(y|\bar{\theta})) + 2pD,\ where\ D = -2\int^{\ln}(p(y|\bar{\theta}))d\theta$$

This criterion is a measure of the goodness-of-fit of the data while, at the same time, introducing a term to evaluate the complexity of the model. This criterion is asymptotically the same as performing cross-validations using part of the data to estimate the data and the rest to calculate the goodness of the predictions (Stone, 1997).

Specifically, for the same data, the lower the value of the DIC calculated, the more accurate the predictions in the short term, which is

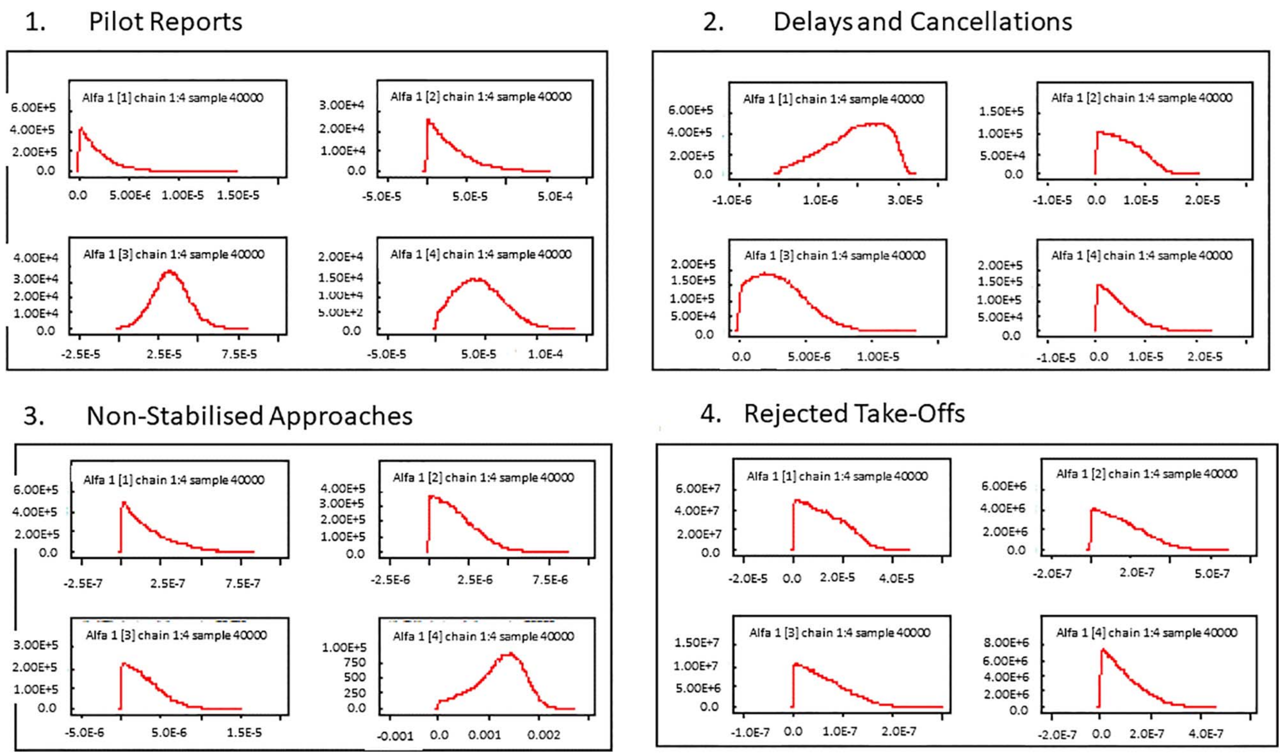

Fig. 19. Posterior distributions of α_1 for Categories 1, 3, 4 and 5.

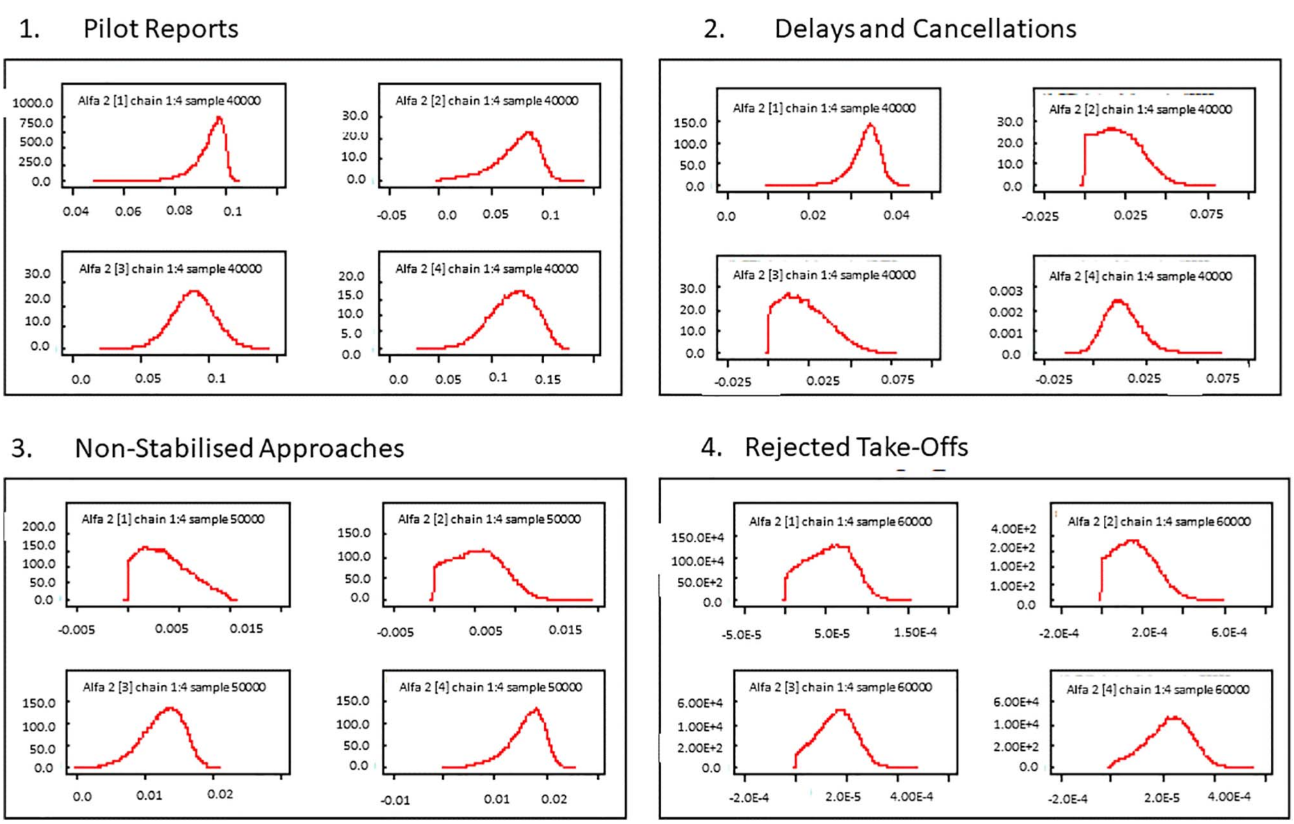

Fig. 20. Posterior distributions of α_2 for Categories 1, 3, 4 and 5.

what interests us. A model may be considered to be superior to another, when its DIC is more than 5 points below that of its competitor (Celeus, et al., 2006).

Table 2 presents the DIC of the five incident categories for all fleets, and each of the models. The values in bold are the lowest values of DIC for each category. However, it should be borne in mind that differences of less than 5 points are not significant.

The table confirms that the Hierarchical Models, especially Models 2 and 3, are better at predicting incidents when there are many monthly events, such as *Pilot Reports, Deferred Items* and *Delays and Cancellations*.

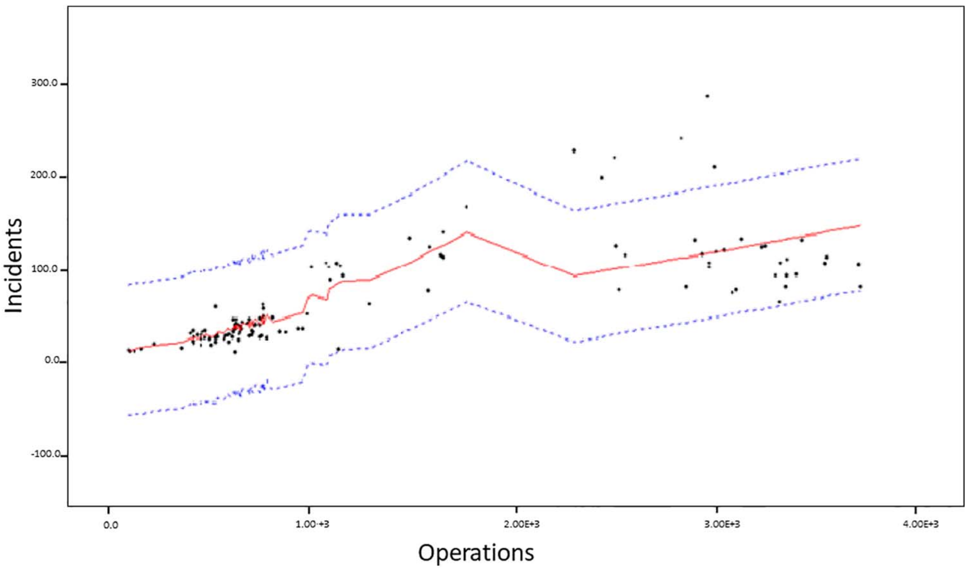

Fig. 21. Predicted incident of Deferred Items using Hierarchical Model 2.

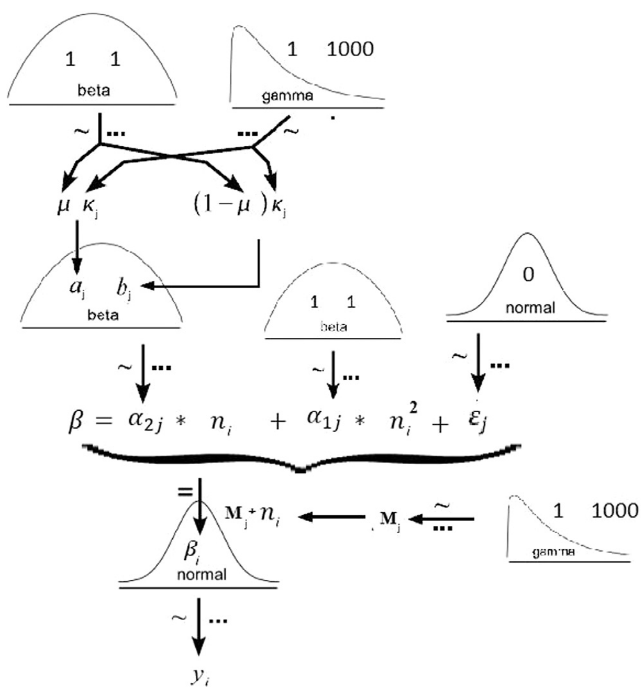

Fig. 22. Structure of Hierarchical Model 3.

For less likely categories of incidents, the simple Gamma-Poisson and Beta-Binomial models are more efficient. This is especially true of extremely infrequent events in which normal approach does not apply, for example, *Rejected Take-Offs* and so on. Finally, in the case of categories with intermediate frequency, of the order of ten incidents per month, both types of model behave similarly.

9. Conclusions

The set of models developed in this study can be used to boost and improve current risk estimation and forecasting methodologies.

Like the Gamma-Poisson and Beta-Binomial models, Hierarchical Model 1 is capable of giving an approximate value of the incident rate per operation, indicate trends and orders of magnitude of the incidents. It estimates parameter μ, for each category of event, representing the performance of the entire set of fleets of the Company and is, therefore, a good tool for comparing the performance of different companies. Parameter μ, also, allows us to detect changes in trends as the posterior distributions are bimodal.

The K parameter is an effective measure of the spread of the incident rate of a fleet, for a particular incident category, compared to the overall rate for the entire Company. Therefore, a rapid, visual analysis of K values will alert us to those fleets that require more in-depth analysis, due to their behaviour deviating from the ideal.

If the values of K are high and similar to one another for all fleets, this means that there is a central tendency around the mean for the Company, since K is inversely proportional to the variance of μ. On the contrary, small values of K indicate that a specific fleet has an unusually low or high number of incidents in a particular category, compared to the Company average. In general, if the value of K has a significantly different order of magnitude compared to that of the other fleets then in-depth analysis is required to identify the cause.

When using Hierarchical Model 2 it is implicitly assumed that incident data has a normal distribution. Therefore, for series in which the data points have widely varying orders of magnitude Hierarchical Model 2 gives remarkably good results. At first sight, the model is very simple and provides sufficiently good results when making predictions and estimates. The confidence intervals give a good fit the data. However, when the monthly values are small the predictions are overestimated.

Hierarchical Model 2 performs a quadratic regression on the data series for the two variables: incidents and operations. The model estimates the regression coefficients. It also quantifies what contribution the linear rate of incidents per operation makes to the generation of data. Furthermore, it calculates how much this rate varies with the number of operations due to the so-called stress effect.

The value of the quadratic coefficient in the regression, parameter α_1, is a measure of the importance of the stress effect. The greater the value of $\alpha 1$, the greater the positive correlation between the incident rate and the number of operations and, consequently, the greater the stress effect.

However, if the value of $\alpha 1$ is very close to zero then the linear parameter of the regression, $\alpha 2$, is the dominant parameter. $\alpha 2$ is the incident rate per operation that was estimated using the previous models. As such, parameter $\alpha 1$ enables us to quantitatively compare the stress effect in one fleet compared to that in others. It also allows us to quantify the importance of the stress effect in the different incident categories.

Finally, Hierarchical Model 3 is a combination of the most attractive

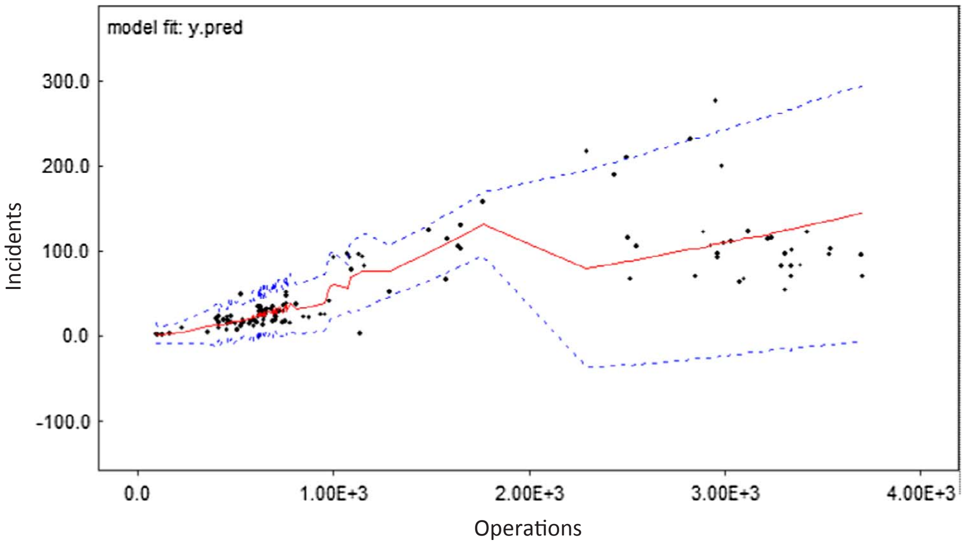

Fig. 23. Predicted incident of Deferred Items using Hierarchical Model 3.

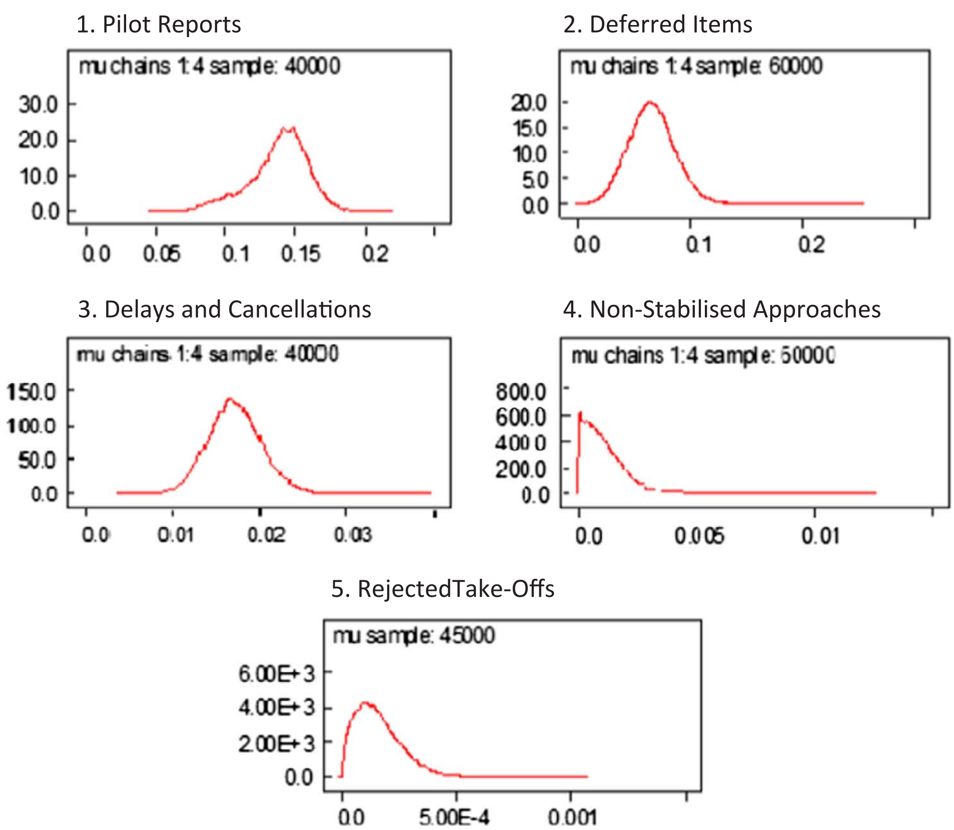

Fig. 24. Posterior distributions of the estimates for parameter μ using Hierarchical Model 3.

features of the two previous models. In general, it performs better than the other two although, as it is a compromise solution, it lacks some of their respective advantages.

Hierarchical Model 3 reparametrises the linear coefficient of the regression of Model 2 as a function of μ and the kJ. Therefore, it inherits the precision of Model 2 as regards detecting the stress effect. It also gives better predictions by adjusting the confidence intervals, as it assumes that the data has a greater spread when the order of magnitude of the number of monthly incidents increases. In other words, it adds a parameter M that acts as a proportionality coefficient between the variance of the normal function and the number of operations.

Hierarchical Model 3 not only gives a very good fit to the data but also has very logical confidence intervals with respect to the spread of the data series for any order of magnitude. It is, therefore, the best model in predicting distributions of incidents. Like Hierarchical Model 2, Model 3 quantifies the importance of the stress effect in a fleet and, via parameter α1, enables comparisons with other fleets.

However, as Hierarchical Model 3 is a compromise between the two previous models, it also has some of the advantages of Hierarchical Model 1. For example, it estimates μ parameters that can be used to compare the performance of different companies and to compare categories of events. Also, the K parameters allow us to extract information about which fleets require most attention within a company.

However, due to its having more parameters, Hierarchical Model 3 does not have the same ability as Hierarchical Model 1 to detect abnormal parameter changes in cases where parameter μ is bimodal.

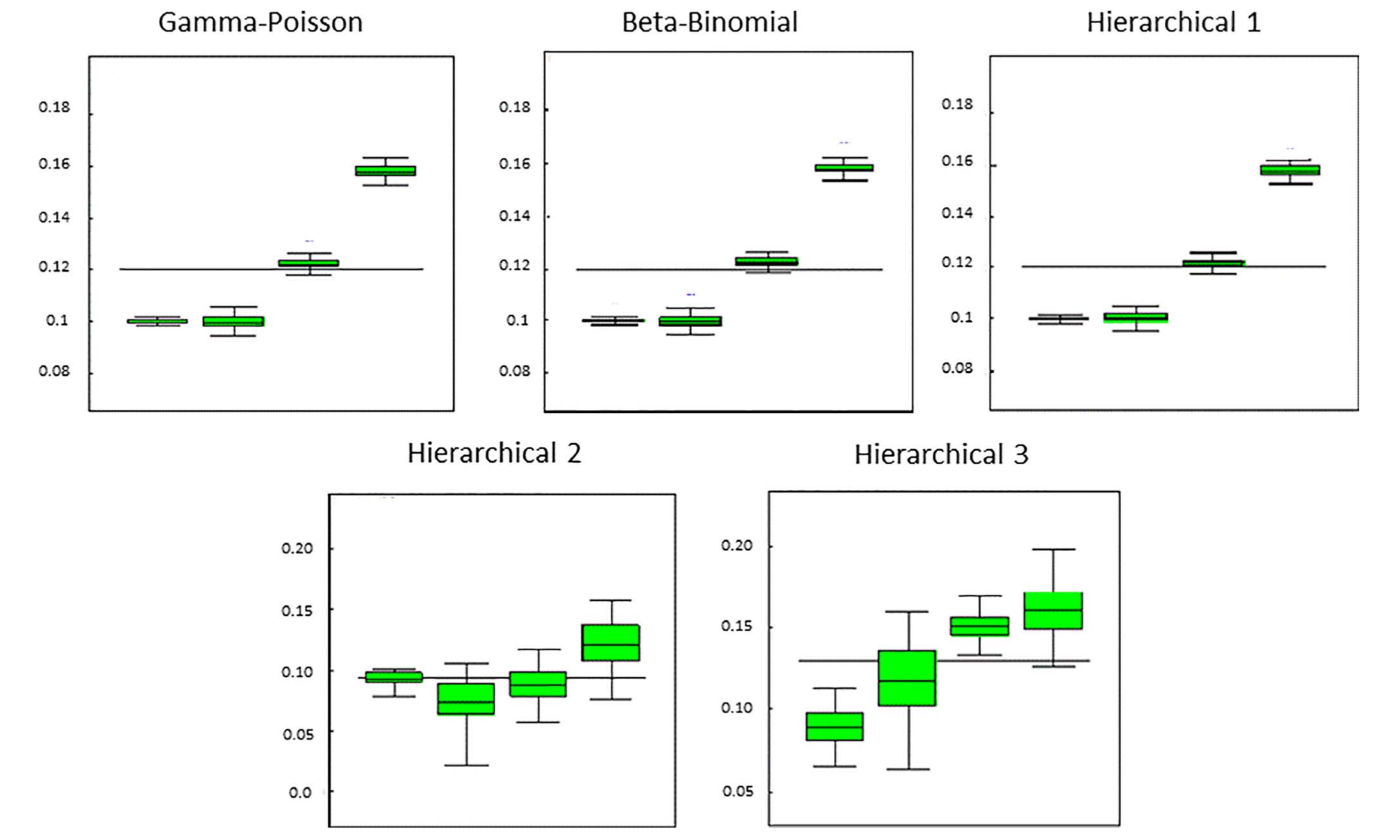

Fig. 25. Comparison of the incident rates estimated by each model.

Table 2
Comparison of DIC for the five models.

DIC	Pilot reports	Deferred items	Delays and cancellations	Non-stabilised approaches	Rejected take-Offs
Poisson-Gamma	1382.29	2383.73	799.679	190.646	84.004
Beta-Binomial	1447.44	2460.90	801.768	190.646	86.605
Hierarchical 1	1447.37	2461.11	802.34	**189.726**	**83.683**
Hierarchical 2	1138.81	1185.73	802.052	200.489	101.147
Hierarchical 3	**1101.46**	**1110.97**	**764.667**	194.215	95.621

Despite this disadvantage, Hierarchical Model 3 has the advantage of uniting many positive characteristics in a single model. Not only does it give the most accurate predictions but, at the same time, it combines most of the advantages and analysis tools found in the other two hierarchical models. As such, Hierarchical Model 3 provides a simple means of performing many different types of analysis that typically form part of the risk assessment carried out by aircraft operators.

This study demonstrates the usefulness of hierarchical structures when it comes to obtaining parameters that can be used to effectively quantify risk. We can identify the parameters that are characteristic of the safe operation of each operator and of the entire system. These parameters allow us to identify and quantify trends, and to establish benchmarks to compare the current year's performance with that of previous years.

In summary, by developing methodologies based on hierarchical models we can extract information in a hierarchical way to identify sources of risk and quantify the specific risks. We can also compare different sources of risk and, most importantly, quantify the level of uncertainty.

References

Allenby, G.M., Rossi, P.E., McCulloch, R., 2005. Hierarchical Bayes Models: A Practitioners Guide, s.l.:s.n.
Ancel, E., Shih, A.T., Jones, S.M., Reveley, M.S., Luxhøj, J.T., Evans, J.K., 2015. Predictive safety analytics: inferring aviation accident shaping factors and causation. J. Risk Res. 18 (4).
Bazargan, M., Guzhva, V., 2007. Factors contributing to fatalities in general aviation. World Rev. Int. Transp. Res. 170–181.
Bernardo, J.S.A., 1994. Bayesian Theory. Wiley & Sons, Chichester.
Bhadra, D., Shaufele, R., 2007. Probabilistic forecasts for aviation traffic at FAA's commercial terminals. Transp. Res. Record 37–46.
Breslow, N., Clayton, D., 1993. Approximate inference in generalized linear mixed models. J. Am. Statist. Assoc.
Brooker, P., 2005. Reducing mid-air collision risk in controlled airspace: lessons from hazardous incidents. Safety Sci. 715–738.
Canfield, R., Teed, J., 1977. Electing the prior distribution in Bayesian estimation. IEEE Reliab.
Celeus, G., Forbes, F., Robert, C., Titterington, D., 2006. Deviance information criteria for missing data models. Bayesian Anal.
Curriero, F., Baker, S., Li, G., Grabowski, J., 2002. Exploratory spatial analysis of pilot fatality rates in general aviation crashes using geographic information systems. Am. J. Epidemiol. 398–405.
Di Gravio, G., Mancini, M., Patriarca, R., Costantino, F., 2015. Overall safety performance of air traffic management system: forecasting and monitoring. Safety Sci. 351–362.
Drees, L., 2014. Model based Prediction of Incidents and Accidents Probability. s.l., s.n.
Dyer, D., Chiou, P., 1982. Most Conservative Beta Prior Distribution for Binomial-Sampling. Technical Report ii187, s.l.: Department of Mathematics. The University of Texas at Arlington. Arlington, Texas.
Enrigth, M., Huyse, L., 2003. Probabilistic Surface Damage Tolerance Assessment of Aircraft Rotors. Atlanta s.n.
EUROCONTROL & FAA, 2012. U.S./Europe comparison of ATM-related Operational Performance. Bruxells: s.n.
European Union, 2014. Regulation (EU) No 376/2014 of the European Parliament and of

the Council of 3 April 2014 on the reporting, analysis and follow-up of occurrences in civil aviation, amending Regulation (EU) No 996/2010 of the European Parliament and of the Council and rep. s.l.:s.n.

Gelman, A., Calin, J., Stern, H., Dunson, D., 2013. Bayesian Data Analysis. Chapman & Hall/CRC Texts in Statistical Science, s.l.

Glickman, M., Dyk, D., 2007. Basic Bayesian methods. Methods Mol. Biol.

Gnoni, M.G., Homer, J., 2017. Near-miss management systems and observability-in-depth: handling safety incidents and accident precursors in light of safety principles. Safety Sci. 154–167.

Houck, S., Powell, W., 2003. Probability of mid air collision during ultra closely spaced parallel approaches. J. Guidance Cont. Dynam. 702–710.

ICAO, 2010. International Standards And Recommended Practices Annex 13 To the Convention on International Civil Aviation Aircraft Accident and Incident Investigation. s.l.: s.n.

ICAO, Third Edition – 2013. Safety Management Manual, (Doc 9859-AN/474). s.l.:s.n.

Janic, M., 2000. An assessment of risk and safety in civil aviation. J. Air Transp. Manage.

Jeffreys, H., 1946. An Invariant Form for the Prior Probability in Estimation Problems. s. l., s.n., pp. 453–461.

Knecht, W.R., 2015. Predicting Accident Rates from General Aviation Pilot Total FLigth Hours. FAA, Oklahoma City.

Knecht, W., Smith, J., 2012. Predicting general aviation accident frequency from pilot total flight hours. (Technical Report No. DOT/FAA/AAM-12/13),Federal Aviation Administration, Office of Aerospace Medicine. Administration Office of Aerospace Medicine, Washington, DC.

Kruschke, K., 2010. Doing Bayesian Data Analysis: A Tutorial with R and BUGS. Academic Press s.l.

Kunimitsu Iwadare, T.O., 2015. Statistical data analyses on aircraft accidents in Japan: occurences, causes and countermeasures. Am. J. Operat. Res. 222–245.

Lord, D., Persaud, B., 2000. Accident prediction models with and without trend: application of the generalized estimating equations procedure. Transp. Res. Record 102–108.

Lord, D., Washington, S.P., Ivan, J.N., 2005. Poisson, Poisson-gamma and zero-inflated regression models of motor. Safety Sci. 35–46.

Lynch, M., 2007. Introduction to Applied Bayesian Statistics and Estimation for Social Scientists. Springer-Verlag, New York.

Margaryan, A., Littlejohn, A., 2016. Research and development agenda for Learning from Incidents. Safety Science.

Miaou, S.P., Lord, D., 2003. Modeling traffic crash-flow relationships for Intersections: dispersion parameter, functional form, and bayes versus empirical bayes. Transp. Res. Record 31–40.

Panagiotakopoulos, D., Majumdar, A., Ochieng, W., 2014. Characterizing the distribution of safety ocurrences in aviation : an approach using extreme value theory. J. Transport. Res. Board.

Qin, X., Ivan, J., Ravishankar, N., 2004. Selecting exposure measures in crash rate prediction for two lana highway segments. Accid. Anal. Prev. 183–191.

Sanchez Ayra, E., 2015. Risk analysis for unintentional slide deployment during airline operations: risk analysis for slide deployment. Risk Anal.

Spiegelhalter, D., Best, N., Carlin, B., 2002. Bayesian measures of model complexity and fit. J. Roy. Statistic. Soc.: Series B (Statistic. Method.).

Stone, M., 1997. An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. J. Roy. Statistic. Soc. Ser. B (Methodolog.).