Received: 19 April 2021 | Revised: 2 December 2021 | Accepted: 6 February 2022

DOI: 10.1002/asmb.2675

RESEARCH ARTICLE

WILEY

Forecasting aviation safety occurrences

Bruno Flores[1] | **David Rios Insua**[1] | **Cesar Alfaro**[2] | **Javier Gomez**[2]

[1]Institute of Mathematical Sciences, ICMAT-CSIC, Madrid, Spain

[2]Department of Computer Science, Rey Juan Carlos University, Mostoles, Spain

Correspondence
Bruno Flores, Institute of Mathematical Sciences, ICMAT-CSIC, Madrid, Spain.
Email: bruno.flores@icmat.es

Funding information
AXA Research Fund; Comunidad de Madrid; Ministerio de Economía, Industria y Competitividad, Gobierno de España

Abstract

We present a general framework for aviation safety occurrence forecasting. This is a major component of a methodology for aviation safety risk management at national level. It covers novel models as well as novel combinations of earlier models. Having good quality occurrence and severity forecasting models is paramount to properly manage risks, maintain the confidence of its users and preserve the status of aviation as a safe transportation mode. The problem is involved due to the presence of complex effects like seasonality, trends, or stress that impact the rates of various occurrences and the uncertainty about future number of operations.

KEYWORDS

aviation safety, Bayesian forecasting, dynamic models, nonhomogeneous Poisson models

1 | INTRODUCTION

Despite the high safety levels of the aviation industry, aviation safety (AS) occurrences continue to take place.[1] These may entail undesirable consequences like deaths, injured people, delays, aircraft destruction or reputation loss, among others. Countries develop national AS plans, implemented through regulations and/or resource allocation, to try to make safety occurrences in their airspace less frequent and/or less severe should they happen. As air freight and passenger traffic is expected to increase in the forthcoming years,[2] notwithstanding the current pandemic, the implementation of effective safety plans in air transportation is of major importance for governments, not only for the safety of its citizens, but also because economic prosperity and employment critically depends on a robust flow of goods and people.

In earlier work,[3] a framework to support AS risk management at country level was presented. Such framework, which has been successfully applied in Spain,[4] supports a government in deciding how to allocate resources to improve AS levels. It is based on Bayesian decision analysis[5] and includes as ingredients models to: (a) forecast the numbers of safety occurrences; (b) forecast occurrence severities; (c) forecast the consequences of safety occurrences; and (d) assess such consequences through a multi-attribute utility function. Such models are integrated to monitor safety, screen occurrences and, more importantly, allocate AS resources. In Rios Insua et al.[6] we detailed ingredients (c) and (d). Here we introduce a novel and comprehensive forecasting methodology to support parts (a) and (b).

Most of the literature concerning AS occurrence forecasting has focused on predicting the circumstances that cause them, for example, aircraft icing[7] or turbulence.[8] While this can be useful for short horizon and operational planning, AS strategic planning at country or airline level requires focusing more on medium to long-term forecasts of occurrences. Moreover, due to the many different types of AS occurrences, there is a need for flexible models capable of forecasting occurrences of very different nature, instead of models that focus on one particular occurrence, as for example, those for runway excursions,[9] flight delays,[10] or go-around/missed approaches.[11] Thus, our aim is at providing a comprehensive modeling approach that can accurately forecast diverse AS occurrences with long or short horizons, that is, the interest lies in modeling time series of nonnegative counts, taking into account the particularities of our application domain.

Using the terminology in Cox,[12] time series models for count data can be divided between *observation-driven* models, with examples in Alzaid and Al-osh,[13] Benjamin et al.,[14] Ferland et al.[15] and Heinen;[16] and *parameter-driven* models, with recent works in West,[17] Aktekin et al.,[18] Chen et al.,[19] Aktekin and Soyer,[20] and Gamerman et al.[21] As discussed in Snyder et al.[22] and Yelland,[23] *parameter-driven state-space* approaches tend to be more flexible and provide a closer match to the empirical properties of the series than *observation-driven* ones. The general approach adopted here is based on dynamic Bayesian analysis of nonhomogeneous Poisson models. Several models are proposed given the effects present in the AS time series to forecast, starting with a standard Poisson-Gamma model, we morph it into novel models (or novel combinations of standard models) to cover the identified issues including stress, trends, seasonal, clustering, and dependence effects, with novel computational algorithms in Data S2.

Section 2 introduces the problem and provides exploratory analysis illustrating key effects in the proposed forecasting domain. They are incorporated gradually in Section 3, stemming from a basic model to which we add one feature in turn to deal with such effects. These models are then expanded to account for the uncertainty in the number of operations, an especially important theme when interested in long-horizon AS forecasts. Then, Section 5 covers models that allow us to predict occurrence severity, and deal with the possible under-reporting in some of those severity groups. The models are illustrated with three cases in Section 6. We use probabilistic influence diagrams throughout[24] to graphically support the presentation of models. An appendix in Data S2 compiles various algorithms introduced for the novel models.

2 | EXPLORATORY DATA ANALYSIS OF AVIATION SAFETY OCCURRENCES

In our motivating scenario, airlines and national AS agencies periodically register occurrences along with the number of operations with the aim of monitoring and improving their occurrence rate as well as reducing their severity. As an example, in our case, referring to AS risk management at country level, 86 occurrence types are considered, ranging from *runway incursions* to *ground handling events* going through *low altitude operations* or *ground collisions*. Each of the occurrences is classified into one of five severity groups as proposed by International Civil Aviation Organization (ICAO[25]): (1) *Accident* (entailing fatalities and/or aircraft destruction); (2) *Serious Incident*; (3) *Major Incident*; (4) *Significant Incident*; and (5) *Occurrence with no safety effect*. Thus, we may talk, for example, about a *severity 2 ground handling* occurrence. For our analysis, we use daily recorded occurrences from 2010 to 2018. For each occurrence, the following information is available: date and airport code; its type and severity; the corresponding number of fatalities, and serious and minor injuries, if any; and, finally, information concerning the aircraft such as its model, and maximum certificated take-off mass. Furthermore, for each airport, the recorded data include information such as its ICAO code, latitude and longitude, and number of daily flight operations. In this paper, as well as in the strategic risk management aspects of our project, we focus on monthly forecasts aggregating the data as required.

For a given occurrence type, for example, *runway incursions*, the data available at the beginning of the kth forecasting period is denoted with D_k and consist of the operations-occurrences pairs $\{(n_1, x_1), \ldots, (n_{k-1}, x_{k-1})\}$, where n_i represents the number of operations and x_i the number of occurrences during the ith period. Hereafter, we introduce the typical effects that can be observed in AS occurrence data, with the aid of exploratory tools based on the primary data $\{(n_i, \hat{\lambda}_i)\}_{i=1}^{k-1}$, with $\hat{\lambda}_i = x_i/n_i$ designating the observed occurrence rate at month ith period.

Figure 1A represents occurrence rate versus number of operations for the *Communication, navigation, and surveillance failures* type, suggesting a relatively stable occurrence rate. However, several effects may appear altering such stability. The first one is showcased in Figure 1B referring to the *TCAS warning* occurrence. In it, we observe that higher numbers of operations induce higher occurrence rates, perhaps due to the increasing pressure over the involved agents (pilots, controllers). We refer to this as a *stress effect*.

Some occurrences can show a *seasonal effect*, typically when affected by regional weather patterns. This is the case of the *bird strike* occurrence, Figure 2A, whose time series of rates shows a pronounced seasonal behavior produced by natural causes related to the migratory movements of birds and their passage near airports. The auto-correlation function (ACF) of occurrence rates can be used to explore and showcase the presence of this effect, Figure 2B.

Sometimes, we can discern a, possibly piecewise, linear variation in the rate, expressed through a *trend*, evolving from one period to another, as Figure 3 displaying the annual *wind shear* occurrence rate shows. Other times certain *grouping effect* is appreciated over an occurrence rate, as shown by the two clusters of airports in Figure 4, referring to the *bird strike* occurrence. As a possible explanation, airports with akin location face analogous weather patterns that, in turn, may induce similarities in the corresponding occurrence rates. It could also be the case that a group of airports

TABLE 1 Models suggested for the 86 different types of occurrences

Model	Types of occurrences
No effect	25
Stress	1
Seasonal	1
Linear	44
Stress+Seasonal	3
Seasonal+Linear	12

is operated by a same company with specific operational procedures which, consequently, induces specificities in their rates.

Finally, several occurrences show relevant *correlation* due to common causes, for technical or physical reasons, or because one of them is a precursor of another. Figure 5 portrays correlations between eleven of them. Note, for example, the high correlation between the *wind* and *wind shear* occurrence rates, most likely due to a common cause.

Several cases shall include more than one effect, as illustrated in Section 6. Indeed, Table 1 summarizes the effects detected for the 86 types of occurrences relevant in our case. As an example, twelve of the occurrence types suggested incorporating both seasonal and linear effects.

3 | FORECASTING AS OCCURRENCES WITH DYNAMIC POISSON MODELS

Let us describe the class of models used to predict the monthly number of occurrences for the 86 relevant types. Our emphasis is on monthly forecasts, although other time granularity forecasts will be required (mainly, annual, for strategic planning, and weekly, for monitoring purposes). We start with a standard Poisson-Gamma model (Section 3.1) and gradually adapt it producing novel models to incorporate procedures to deal with the effects described in Section 2 (stress, seasonal, trend, clustering, dependence). The corresponding algorithms are presented in Data S2.

3.1 | Basic model

Let us begin with the basic version of our model. In it (as in Figure 1A) the occurrence rate remains relatively stable throughout the year (and over the years), not appreciating the effects suggested in Section 2. We deal with it with a standard Poisson-Gamma model, taking into account the approximation of the Poisson distribution to the binomial, for example, Rios Insua et al.,[26] and that the Gamma is conjugate for the Poisson. Thus, our basic model is

$$x_k|\lambda, n_k \sim \mathrm{Po}\left(\lambda n_k\right), \quad \lambda \sim \mathrm{Ga}(a, p),$$

where x_k is the number of occurrences during the kth period; n_k is the number of operations during such period; and, finally, λ is the occurrence rate. It is well known that the posterior distribution for the occurrence rate at the kth period, having observed D_k, is $\lambda|D_k \sim \mathrm{Ga}\left(a_k, p_k\right)$, with $a_k = a_{k-1} + x_{k-1}$ and $p_k = p_{k-1} + n_{k-1}$; moreover, the posterior predictive distribution for the number of occurrences is $x_k|D_k \sim \mathrm{Neg-Bin}\left(a_k, p_k/(n_k + p_k)\right)$, from which the predictive mean, variance and intervals follow easily.[27]

3.2 | Variants over the basic model

Several nontrivial variants of the basic model need to be considered to take into account the effects in Section 2 leading to novel models. They are dealt with one at a time. Section 6.1 provides an example with an effect combination (linear trend, seasonal and group), by aggregating the proposed modeling ideas.

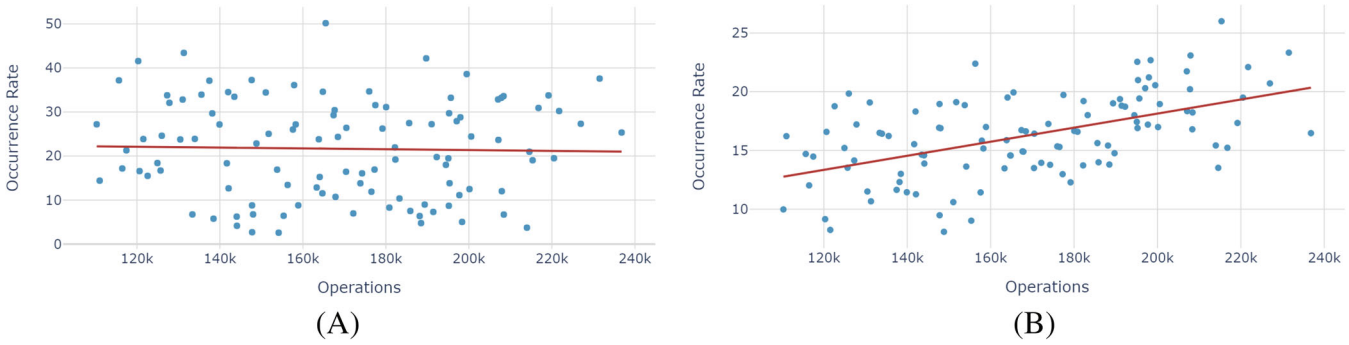

FIGURE 1 Stable relation (A) and stress effect (B) in occurrence rates

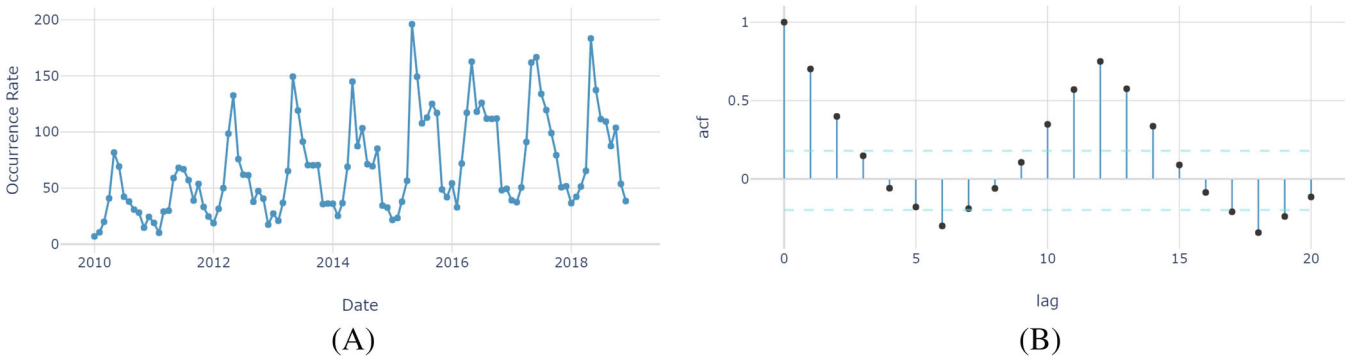

FIGURE 2 Seasonal effect of the *bird strike* occurrence

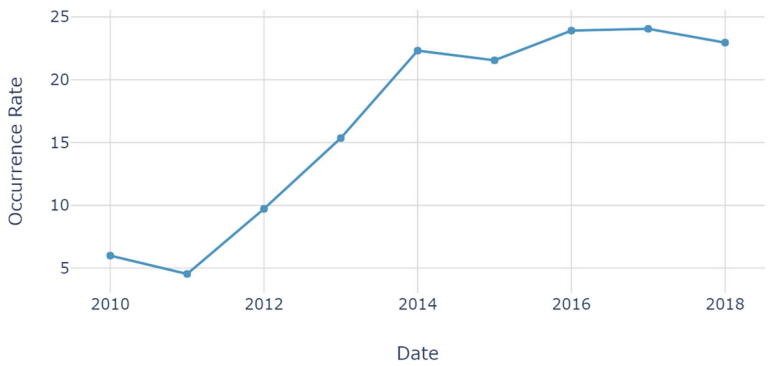

FIGURE 3 Trend in *wind shear* occurrence

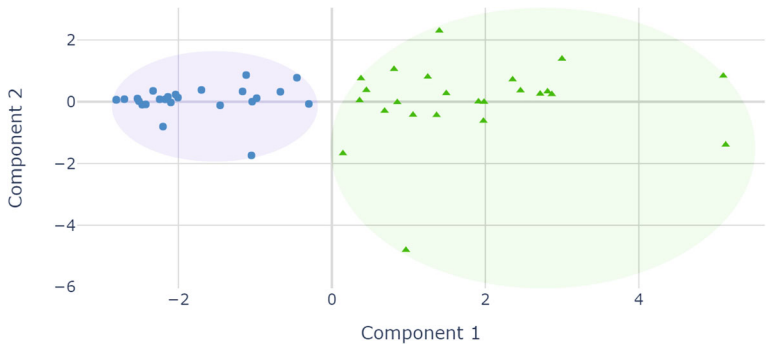

FIGURE 4 Group effect of *bird strike* occurrence

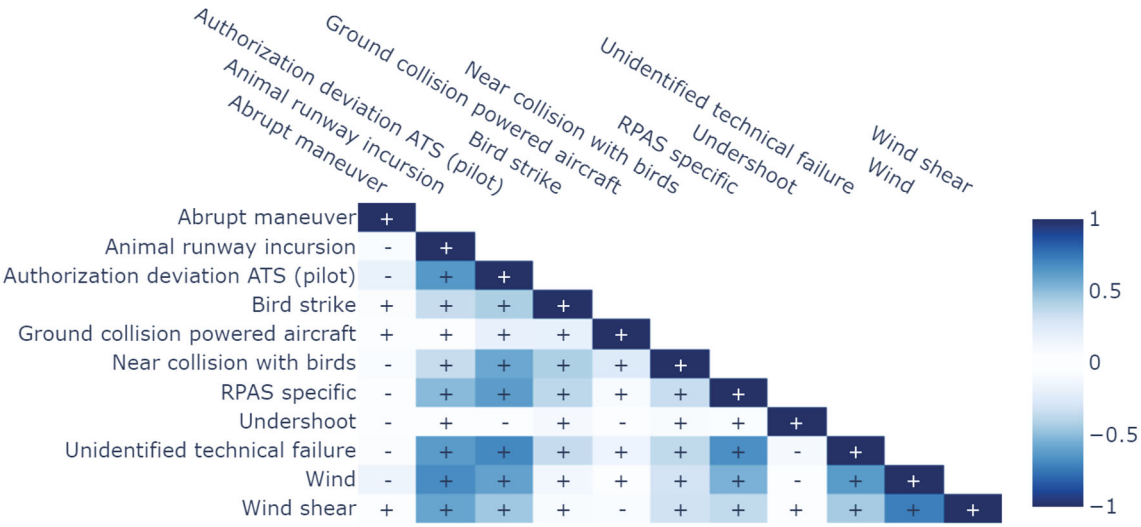

FIGURE 5 Correlation matrix for 11 occurrence types

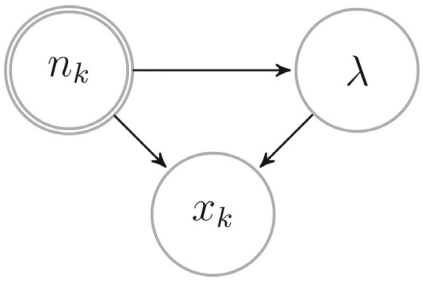

FIGURE 6 Influence diagram for stress effect (kth period)

3.2.1 | Stress effect

Figure 6 reflects how we deal with stress effects (Figure 1B). Its simplest expression would be based on a linear relationship between the rate and the number of operations as follows:

$$\begin{aligned} x_k|\lambda, n_k &\sim \text{Po}\left(\lambda n_k\right) \\ \lambda &= a n_k + b + \varepsilon_k, \quad \varepsilon_k \sim N\left(0, \sigma^2\right), \\ a \sim N(\mu_a, \sigma_a^2), \quad b &\sim N(\mu_b, \sigma_b^2), \quad \sigma^2 \sim \text{Inv} - \text{Gamma}(\alpha, \beta). \end{aligned} \tag{1}$$

Given data D_k, the posterior $p(\lambda, a, b, \sigma^2|D_k)$ is easily seen to be proportional to

$$\lambda^{\sum_{i=1}^{k-1} x_i} \sigma^{-3-2\alpha} \exp\left(-\lambda \sum_{i=1}^{k-1} n_i - \frac{(\lambda - a n_k - b)^2}{2\sigma^2} - \frac{(a-\mu_a)^2}{2\sigma_a^2} - \frac{(b-\mu_b)^2}{2\sigma_b^2} - \frac{\beta}{\sigma^2}\right).$$

The conditional posterior distributions are then

$$p(a|\lambda, b, \sigma^2, D_k) \sim N\left(a \,\middle|\, \frac{\sigma^2 \mu_a + n_k \sigma_a^2 (\lambda - b)}{\sigma^2 + n_k^2 \sigma_a^2}, \frac{\sigma^2}{n_k^2 + \sigma^2/\sigma_a^2}\right),$$

$$p(b|\lambda, a, \sigma^2, D_k) \sim N\left(b \,\middle|\, \frac{\sigma^2 \mu_b + \sigma_b^2 (\lambda - a n_k)}{\sigma^2 + \sigma_b^2}, \frac{1}{1/\sigma_b^2 + 1/\sigma^2}\right),$$

$$p(\sigma^2|\lambda, a, b, D_k) \sim \text{Inv} - \text{Gamma}\left(\sigma^2 \;\middle|\; \alpha + \frac{1}{2}, \beta + \frac{1}{2}(b + an_k - \lambda)^2\right),$$
$$p(\lambda|a, b, \sigma^2, D_k) \propto \lambda^{\sum x_i} \exp\left(-\frac{1}{2\sigma^2}\left(\lambda^2 - 2\lambda\left(an_k + b - \sigma^2 \sum n_i\right)\right)\right).$$

Based on these, Algorithm 1 in Data S2 provides a hybrid scheme to sample from the posterior, using Gibbs steps to sample from the conditional posteriors of a, b, and σ^2 and a Metropolis–Hastings step to sample from the conditional posterior of λ. In particular, note that we may check whether the posterior distribution of a concentrates around 0 to eventually discard the presence of a stress effect. We do this by computing the posterior probability of an interval around 0, as illustrated in Section 6.2. Note that the nonnegativity of λ is controlled through the support of the proposal distribution of our MCMC sampler (Algorithm 1).

Regarding the predictive distribution for the number of occurrences, we use

$$\begin{aligned} Pr(x_k = z|D_k) &= \iint \iint Pr(x_k = z|\lambda, n_k)\; p(\lambda|a, b, \sigma^2, D_k)\; p(a, b, \sigma^2|D_k)\; d\lambda\; da\; db\; d\sigma^2 \\ &\approx \frac{1}{N}\sum_{j=1}^{N} Pr(x_k = z|\lambda_j, n_k) = \frac{n_k^z}{Nz!}\sum_{j=1}^{N} \exp(-\lambda_j n_k)\left(\lambda_j\right)^z, \end{aligned}$$

based on the sample $\{\lambda_j\}_{j=1}^N$ from Algorithm 1. From it, predictive means, variances and intervals follow easily.

3.2.2 | Effects dealt with through dynamic linear models

In many cases, one cannot consider the occurrence rate as constant, recall Figures 2 and 3. Trend and seasonal effects are modeled with Dynamic Linear Models (DLM).[28] As observations are considered to come from a Poisson distribution, they cannot be directly dealt with a DLM. Then, the model is described through

$$\begin{aligned} &x_k|\lambda_k, n_k \sim \text{Po}\left(\lambda_k n_k\right), \qquad \lambda_k = \exp(u_k), \\ &u_k = \boldsymbol{F}_k\boldsymbol{\theta}_k + v_k, \qquad v_k \sim N(0, V_k), \\ &\boldsymbol{\theta}_k = \boldsymbol{G}_k\boldsymbol{\theta}_{k-1} + \boldsymbol{w}_k, \quad \boldsymbol{w}_k \sim N(\boldsymbol{0}, \boldsymbol{W}_k), \\ &\boldsymbol{\theta}_0 \sim N(\boldsymbol{m}_0, \boldsymbol{C}_0), \end{aligned} \tag{2}$$

where $\boldsymbol{F}_k$ and $\boldsymbol{G}_k$ are known matrices; and v_k and $\boldsymbol{w}_k$ are independent sequences of normal variables with zero mean and variances V_k and $\boldsymbol{W}_k$, respectively. The exponential transformation in the second equation guarantees positivity of the occurrence rate. The model is represented in Figure 7, in which we obviate the deterministic relationship between λ and u.

Note that another option would be to use a Dynamic Generalized Linear Model (DGLM[28]) by eliminating the noise v_k above, and not assuming any particular distribution for $\boldsymbol{w}_k$, just its mean and variance. However, we prefer to use (2) in our AS context since the dual source of error offers additional flexibility and adapts better to the motivating case. Models making use of DGLMs and its sequential updating might be beneficial in applications dealing with very large number of time series and/or time series with frequent zeros, which is not our case.

Let us briefly discuss modeling possibilities that are convenient in our AS domain. The DLMs considered can be used as building blocks combined through the superposition principle[29] to form a model when both effects are deemed relevant.

Trend effect. The basic models to deal with a dynamic occurrence rate are the first order polynomial model, characterized by $\boldsymbol{F} = \boldsymbol{G} = 1$, and the second order polynomial model, or linear growth, based on

$$\boldsymbol{F}_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad \boldsymbol{G}_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

We choose the later as it is more general and allows us to deal more adequately with our datasets, both for modeling and forecasting purposes.

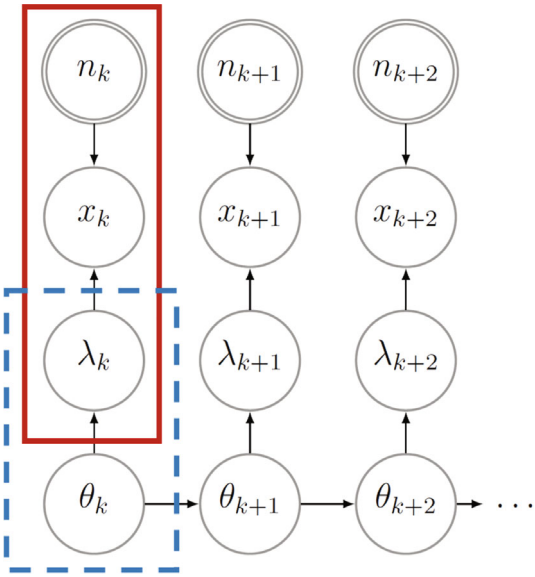

FIGURE 7 Influence diagram with log-rate as dynamic linear models (DLM). Poisson component, solid; DLM, dash

Seasonal effect. With monthly data, when considering the presence of a seasonal effect of period 12, as in Figure 2, we use a DLM with $\boldsymbol{F}_2 = (1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)$,

$$\boldsymbol{G}_2 = \begin{pmatrix} -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Inference and prediction procedures are common in both cases (and their combination) through this scheme: *Step 0. Forecast at period k*. At the beginning of the kth period, before observing x_k, we have the distributions $\pi(x_k|\lambda_k, n_k)$, $\pi(\lambda_k|u_k)$, $\pi(u_k|\boldsymbol{\theta}_k)$ and $\pi(\boldsymbol{\theta}_k)$, (typically, this last one will be characterized by a sample $\{\boldsymbol{\theta}_k^i\}_{i=1}^N$, with weights $\pi_k^i \geq 0$, and $\sum_{i=1}^N \pi_k^i = 1$). To make predictions about the number of occurrences x_k, the distribution is

$$\begin{aligned} \pi(x_k|n_k) &= \iiint \pi(x_k|\lambda_k, n_k)\ \pi(\lambda_k|u_k)\ \pi(u_k|\boldsymbol{\theta}_k)\ \pi(\boldsymbol{\theta}_k)\ d\lambda_k\ du_k\ d\boldsymbol{\theta}_k \\ &= \iint \pi(x_k|\exp(u_k), n_k)\ \pi(u_k|\boldsymbol{\theta}_k)\ \pi(\boldsymbol{\theta}_k)\ du_k\ d\boldsymbol{\theta}_k, \end{aligned}$$

estimated by simulation through

Sample $\{\boldsymbol{\theta}_k^i\}_{i=1}^N \sim \pi(\boldsymbol{\theta}_k)$ (possibly already available);

Do $\lambda_k^i = \exp(\boldsymbol{F}_k \boldsymbol{\theta}_k^i)$, for $i = 1, \ldots, N$;

Approximate the predictive $\pi(x_k|n_k)$ with $\quad \pi(x_k|n_k) \approx \frac{n_k^{x_k}}{N x_k!} \sum_{i=1}^N \exp(-\lambda_k^i n_k)(\lambda_k^i)^{x_k}$.

The predictive mean and second moment are approximated through

$$E(X_k|n_k) = \sum x_k \pi(x_k|n_k) \approx \frac{n_k}{N}\sum_{i=1}^{N} \exp\left(\frac{1-2\boldsymbol{F}_k\boldsymbol{\theta}_k^i}{2V_k}\right) =: \eta_k,$$
$$E(X_k^2|n_k) = \sum x_k^2\, \pi(x_k|n_k) \approx \eta_k + \frac{n_k^2}{N}\sum_{i=1}^{N} \exp\left(\frac{2-2\boldsymbol{F}_k\boldsymbol{\theta}_k^i}{V_k}\right) =: \overline{\eta_k}.$$

Therefore, the predictive variance is approximated through $V(X_k|n_k) \approx \overline{\eta_k} - \eta_k^2 =: \kappa_k^2$.

Step 1. Observation of x_k and update. Once x_k is observed, it is propagated to obtain $\pi(\theta_k|x_k)$. From step 0, preserve samples $\{\boldsymbol{\theta}_k^i\}_{i=1}^N$ and $\{\lambda_k^i\}_{i=1}^N$; for each $\boldsymbol{\theta}_k^i$, draw $\{u_k^{ih}\}_{h=1}^N \sim N\left(\boldsymbol{F}_k\boldsymbol{\theta}_k^i, V_k\right)$, do $\lambda_k^{ih} = \exp(u_k^{ih})$, $h = 1, \dots, N$, and approximate

$$\pi(x_k|\boldsymbol{\theta}_k^i) = \iint \pi(x_k|\lambda_k, n_k)\pi(\lambda_k|u_k)\pi(u_k|\boldsymbol{\theta}_k^i)\, d\lambda_k\, du_k \approx \frac{n_k^{x_k}}{Nx_k!}\sum_{h=1}^{N} \exp(-\lambda_k^{ih}n_k)(\lambda_k^{ih})^{x_k}.$$

Suppressing dependence on u_k, which is fixed, we get

$$\pi(\boldsymbol{\theta}_k^i|x_k) = \frac{\pi(x_k|\boldsymbol{\theta}_k^i)\pi(\boldsymbol{\theta}_k^i)}{\pi(x_k)} \approx \frac{\left(\frac{1}{N}\sum_{h=1}^N \pi(x_k|\lambda_k^{ih})\right)\pi(\boldsymbol{\theta}_k^i)}{\frac{1}{N}\sum_{i=1}^N \pi(x_k|\lambda_k^i)}$$
$$\propto \pi(\boldsymbol{\theta}_k^i)\sum_{h=1}^{N} \exp(-\lambda_k^{ih}n_k)(\lambda_k^{ih})^{x_k}. \tag{3}$$

Step 2. Propagation to period $(k+1)$. We have $\pi(\boldsymbol{\theta}_{k+1}|D_k) = \int \pi(\boldsymbol{\theta}_{k+1}|\boldsymbol{\theta}_k)\pi(\boldsymbol{\theta}_k|x_k)d\boldsymbol{\theta}_k$. Thus, the distribution for the d-dimensional state vector $\boldsymbol{\theta}_k$ is approximated by

$$\pi(\boldsymbol{\theta}_{k+1}|D_k) \approx \frac{1}{N}\sum_{i=1}^{N}\pi(\boldsymbol{\theta}_{k+1}|\boldsymbol{\theta}_k^i)$$
$$= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sqrt{(2\pi)^d|\boldsymbol{W}_k|}}\exp\left(-\frac{1}{2}(\boldsymbol{\theta}_{k+1} - \boldsymbol{G}_k\boldsymbol{\theta}_k^i)'\boldsymbol{W}_k^{-1}(\boldsymbol{\theta}_{k+1} - \boldsymbol{G}_k\boldsymbol{\theta}_k^i)\right),$$

with $\{\boldsymbol{\theta}_k^i\}_{i=1}^N$ a sample of $\boldsymbol{\theta}_k|x_k$, from step 1.

After this, we would be again at step 0 and re-initiate the process. Algorithm 2 in Data S2 summarizes the above. An MCMC approach could also be adopted for this model but, as pointed out by Storvik[30] and Aktekin et al.,[18] this presents disadvantages over the particle filter approach.

3.2.3 | Group effect

A simple approach to dealing with *group effects* (Figure 4) would be to model each cluster of observations separately. A more elaborate version uses hierarchical modeling relating the cluster components through parameters coming from a same hyperdistribution. For L groups, using the basic model in Section 3.1, we have

$$x_k^i|\lambda^i, n_k^i \sim \text{Po}\left(\lambda^i n_k^i\right), \quad \lambda^i \sim \text{Ga}\left(a^i, p^i\right)\ i = 1, \dots, L$$
$$a^i \sim \text{Ga}(\alpha, \beta) \quad p^i \sim \text{Ga}(\gamma, \delta).$$

The posterior distribution is

$$\pi(\lambda^1, \dots, \lambda^L, a^1, \dots, a^L, p^1, \dots, p^L|D_k) \propto$$
$$\prod_{i=1}^{L}\left(\left[\prod_{j=1}^{k-1}\pi(x_j^i|\lambda^i, n_j^i)\right]\pi(\lambda^i|a^i, p^i)\pi(a^i|\alpha, \beta)\pi(p^i|\gamma, \delta)\right) \propto$$

$$\prod_{i=1}^{L}\left[\exp\left(-p^i\lambda^i - a^i\beta - p^i\delta - \lambda^i\sum n_j^i\right)(\lambda^i)^{a^i-1+\sum x_j^i}(p^i)^{a_i+\gamma-1}\frac{\beta^\alpha\delta^\gamma}{\Gamma(a^i)\Gamma(\alpha)\Gamma(\gamma)}(a^i)^{\alpha-1}\right].$$

The conditional posterior distributions for each parameter $i = 1, \dots, L$ are:

$$\pi(\lambda^i|a^i,p^i,D_k) \propto \exp\left(-\lambda^i\left(p^i+\sum n_j^i\right)\right)(\lambda^i)^{a^i+\sum x_j^i-1} \sim \mathrm{Ga}\left(a^i+\sum x_j^i, p^i+\sum n_j^i\right),$$
$$\pi(p^i|\lambda^i,a^i,D_k) \propto \exp(-p^i(\lambda^i+\delta))p^{i(a^i+\gamma-1)} \sim \mathrm{Ga}\left(a^i+\gamma,\lambda^i+\delta\right),$$
$$\pi(a^i|\lambda^i,p^i,D_k) \propto \frac{(\lambda^i p^i\exp(-\beta))^{a^i-1}(a^i)^{\alpha-1}}{\Gamma(a^i)}.$$

The last ones lack a standard form, but can be treated through a Metropolis–Hastings step as shown in Algorithm 3 in Data S2.

3.2.4 | Dependence of AS occurrence types

So far, we have assumed that occurrence types were independent. However, as discussed in Section 2, it is reasonable to assume that some of them might be related (Figure 5). Although there are other variants, a relevant representation for dependent occurrences would be as in Figure 8.

If the relation between λ_1 and λ_2 is assumed to be linear, a relevant model would be

$$x_{1,k}|\lambda_1,n_k \sim \mathrm{Po}\left(\lambda_1 n_k\right), \quad x_{2,k}|\lambda_2,n_k \sim \mathrm{Po}\left(\lambda_2 n_k\right),$$
$$\lambda_1 \sim \mathrm{Ga}(r,p), \quad \lambda_2 = a\lambda_1 + b + \varepsilon, \varepsilon \sim N\left(0,\sigma^2\right),$$
$$a \sim N(\mu_a,\sigma_a^2), \quad b \sim N(\mu_b,\sigma_b^2), \quad \sigma^2 \sim \mathrm{Inv-Gamma}(\alpha,\beta).$$

Given data $D_k = \{(x_{1,i}, x_{2,i}, n_i)\}_{i=1}^{k-1}$, the joint posterior would be

$$\pi(\lambda_1,\lambda_2,a,b,\sigma^2|D_k) \propto$$
$$\lambda_1^{r-1+\sum_{i=1}^{k-1}x_{1,i}}\lambda_2^{\sum_{i=1}^{k-1}x_{2,i}}\sigma^{-2\alpha-3}\exp\left(-(\lambda_1+\lambda_2)\sum_{i=1}^{k-1}n_i - \frac{(\lambda_2-a\lambda_1-b)^2}{2\sigma^2} - p\lambda_1 - \frac{(a-\mu_a)^2}{2\sigma_a^2} - \frac{(b-\mu_b)^2}{2\sigma_b^2} - \frac{\beta}{\alpha^2}\right).$$

The conditional posterior distributions are

$$\pi(\lambda_1|\lambda_2,a,b,\sigma^2,D_k) \propto \lambda_1^{r-1+\sum x_{1,i}}\exp\left(-\lambda_1\left(p+\sum n_i + \frac{(\lambda_2-a\lambda_1-b)^2}{2\sigma^2}\right)\right),$$
$$\pi(\lambda_2|\lambda_1,a,b,\sigma^2,D_k) \propto \lambda_2^{\sum x_{2,i}}\exp\left(-\lambda_2\left(\sum n_i + \frac{(\lambda_2-a\lambda_1-b)^2}{2\sigma^2}\right)\right),$$
$$\pi(a|\lambda_1,\lambda_2,b,\sigma^2,D_k) \sim N\left(a \,\middle|\, \frac{\sigma^2\mu_a+\lambda_1\sigma_a^2(\lambda_2-b)}{\sigma^2+\lambda_1^2\sigma_a^2}, \frac{\sigma^2}{\lambda_1^2+\sigma^2/\sigma_a^2}\right),$$
$$\pi(b|\lambda_1,\lambda_2,a,\sigma^2,D_k) \sim N\left(b \,\middle|\, \frac{\sigma^2\mu_b+\sigma_b^2(\lambda_2-a\lambda_1)}{\sigma^2+\sigma_b^2}, \frac{1}{1/\sigma^2+1/\sigma_b^2}\right),$$
$$\pi(\sigma^2|\lambda_1,\lambda_2,a,b,D_k) \sim \mathrm{Inv-Gamma}\left(\sigma^2 \,\middle|\, \alpha+\frac{1}{2}, \beta+\frac{(\lambda_2-a\lambda_1-b)^2}{2}\right).$$

Based on them, we build a hybrid MCMC sampler to generate from the posterior as in Algorithm 4 in Data S2. As in 3.2.1, nonnegativity of λ_2 is guaranteed by using the recommended jumping distribution in Data S2.

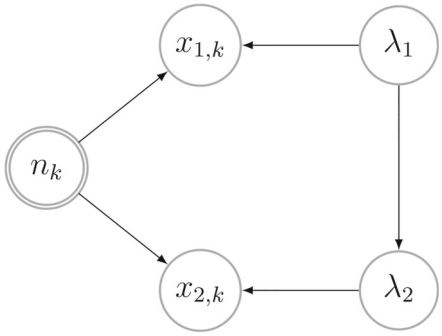

FIGURE 8 Influence diagram for the dependence effect between occurrences

As a final comment, note that should the rate of one of the occurrences be higher than that of the other, which does not hold in our case, we could alternatively consider the use of McKay's bivariate gamma model.[31]

4 | FORECASTING AS OCCURRENCES WITH AN UNCERTAIN NUMBER OF OPERATIONS

The number n_k of aviation operations in the k-th period was assumed known in previous models. This may be realistic for short-term forecasts in which there is little uncertainty about the number of operations to be held; on the other hand, for long horizons, for example, in annual operational planning, there is uncertainty about such quantities, which should be taken into account to improve occurrence forecasting. Consider thus the case in which the number of operations is uncertain and both the occurrence rate and such number evolve according to DLMs. The corresponding influence diagram is reflected in Figure 9.

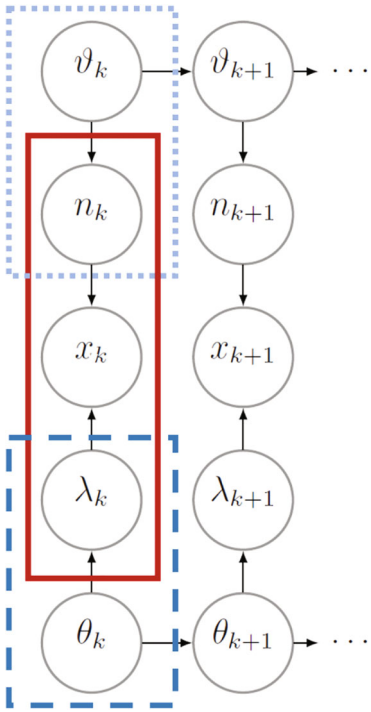

FIGURE 9 Predicting occurrences with uncertain number of operations. Occurrence rate dynamic linear models (DLM), dashed; operations DLM, dotted; Poisson, solid

The resulting model would be

$$\begin{cases} \begin{cases} n_k = \boldsymbol{H}_k \boldsymbol{\vartheta}_k + z_k, z_k \sim N(0, \Sigma_k) \\ \boldsymbol{\vartheta}_k = \boldsymbol{J}_k \boldsymbol{\vartheta}_{i-1} + \boldsymbol{\xi}_k, \boldsymbol{\xi}_k \sim N(\boldsymbol{0}, \boldsymbol{S}_k) \\ \boldsymbol{\vartheta}_0 \sim N\left(\boldsymbol{\eta}_0, \boldsymbol{S}_0\right) \end{cases} \\ x_k | \lambda_k, n_k \sim \mathrm{Po}\left(\lambda_k n_k\right), \quad \lambda_k = \exp(u_k) \\ \begin{cases} u_k = \boldsymbol{F}_k \boldsymbol{\theta}_k + v_k, v_k \sim N(0, V_k) \\ \boldsymbol{\theta}_k = \boldsymbol{G}_k \boldsymbol{\theta}_{k-1} + \boldsymbol{w}_k, \boldsymbol{w}_k \sim N(\boldsymbol{0}, \boldsymbol{W}_k) \\ \boldsymbol{\theta}_0 \sim N(\boldsymbol{m}_0, \boldsymbol{C}_0), \end{cases} \end{cases}$$

where, in addition to the features in model (2), $\boldsymbol{\vartheta}_k$ are the state variables for the number of operations; $\boldsymbol{H}_k$ and $\boldsymbol{J}_k$ are the regression vector and evolution matrix of the operations DLM; and, finally, z_k, $\xi_k$would be independent sequences of normal variables (independent of v_k and w_k) with zero mean and variances Σ_k and $\boldsymbol{S}_k$, respectively. Contrary to λ_k, the number of occurrences n_k is modeled directly with a DLM, and therefore some probability is assigned to nonpositive values; however, since the minimum aggregation level we are interested in is the number of operations in an airport during a month, which is consistently in the tens of thousands, this is not significant and allows to use the straightforward sequential updating of DLMs for the number of operations. The prediction procedure at the kth step would thus be:

Step 0. Prediction of x_k and n_k *at period* k. We have distributions $\pi(\boldsymbol{\theta}_k)$, $\pi(u_k|\boldsymbol{\theta}_k)$, $\pi(x_k|u_k, n_k)$, $\pi(n_k|\boldsymbol{\vartheta}_k)$, $\pi(\boldsymbol{\vartheta}_k)$, and the relation $\lambda_k = \exp(u_k)$ (again, it could be the case that some of these distributions are expressed through samples; in particular, that of $\pi(\boldsymbol{\theta}_k)$ would be given by a sample $\{\boldsymbol{\theta}_k^i\}_{i=1}^N$, with weights $\pi_k^i \geq 0$, $\sum \pi_k^i = 1$). To predict n_k, use the predictive distribution $\pi(n_k) = \int \pi(n_k|\boldsymbol{\vartheta}_k)\pi(\boldsymbol{\vartheta}_k)\, d\boldsymbol{\vartheta}_k$, based on the DLM predictive formulae, Appendix A, having a normal distribution $n_k \sim N(f_k, Q_k)$. To predict x_k, use

$$\pi(x_k) = \iiint \pi(x_k|\exp(u_k), n_k)\pi(n_k)\pi(u_k|\boldsymbol{\theta}_k)\pi(\boldsymbol{\theta}_k)\, dn_k\, du_k\, d\boldsymbol{\theta}_k.$$

We simulate it as follows:

Sample $\{\boldsymbol{\theta}_k^i\}_{i=1}^N \sim \pi(\boldsymbol{\theta}_k)$ (possibly already available);
Do $\lambda_k^i = \exp(\boldsymbol{F}_k \boldsymbol{\theta}_k^i)$, for $i = 1, \ldots, N$;
Sample $\{n_k^i\}_{i=1}^N \sim N(f_k, Q_k)$;
Approximate $\pi(x_k) \approx \frac{1}{N x_k!} \sum_{i=1}^N \exp(-\lambda_k^i n_k^i)(\lambda_k^i n_k^i)^{x_k}$.

The approximate predictive mean and second moment are

$$E(X_k) \approx \frac{f_k}{N} \sum_{i=1}^N \exp\left(\frac{1 - 2\boldsymbol{F}_k \boldsymbol{\theta}_k^i}{2V_k}\right) =: \eta_k,$$
$$E(X_k^2) \approx \eta_k + \frac{f_k^2 + Q_k^2}{N} \sum_{i=1}^N \exp\left(\frac{2 - 2\boldsymbol{F}_k \boldsymbol{\theta}_k^i}{V_k}\right) =: \overline{\eta_k}.$$

Then, the predictive variance would be approximated by $\overline{\eta_k} - \eta_k^2$.

Step 1. Observation of (x_k, n_k) *and update*. At the end of the kth period, observe x_k, n_k and propagate this information to obtain $\pi(\boldsymbol{\theta}_k|x_k, n_k)$ and $\pi(\boldsymbol{\vartheta}_k|n_k)$. First, invert the relation $x \to \lambda$. The new distribution at node x is $\pi(x_k|n_k, \boldsymbol{\theta}_k) = \int \pi(x_k|n_k, \lambda_k)\pi(\lambda_k|\boldsymbol{\theta}_k)\, d\lambda_k$. The posterior for λ_k is

$$\pi(\lambda_k|x_k, n_k, \boldsymbol{\theta}_k) = \frac{\pi(\lambda_k|\boldsymbol{\theta}_k)\pi(x_k|\lambda_k, n_k)}{\pi(x_k|n_k, \boldsymbol{\theta}_k)} \propto \pi(\lambda_k|\boldsymbol{\theta}_k)\pi(x_k|\lambda_k, n_k).$$

Propagate now the evidence of x_k and n_k resulting in

$$\pi(\boldsymbol{\theta}_k|x_k, n_k) = \int \pi(\lambda_k|x_k, n_k, \boldsymbol{\theta}_k)\pi(\boldsymbol{\theta}_k)\, d\lambda_k.$$

It can be approximated by

$$\pi(\boldsymbol{\theta}_k|x_k, n_k) \approx \frac{1}{N}\sum_{i=1}^{N}\pi(\lambda_k|x_k, n_k, \boldsymbol{\theta}_k^i) = \frac{1}{N}\sum_{i=1}^{N}\frac{\pi(\lambda_k|\boldsymbol{\theta}_k^i)\pi(x_k|\lambda_k, n_k)}{\int \pi(x_k|n_k, \lambda_k)\pi(\lambda_k|\boldsymbol{\theta}_k^i)\, d\lambda}.$$

The propagation of evidence n_k to $\boldsymbol{\vartheta}_k$ is done through

$$\pi(\boldsymbol{\vartheta}_k|n_k) = \frac{\pi(n_k|\boldsymbol{\vartheta}_k)\pi(\boldsymbol{\vartheta}_k)}{\pi(n_k)},$$

with $\pi(n_k) = \int \pi(n_k|\boldsymbol{\vartheta}_k)\pi(\boldsymbol{\vartheta}_k)\, d\boldsymbol{\vartheta}_k$. Thereupon, the DLM equations for sequential updating in Appendix A are used.

Step 2. Propagation to period $k+1$. It is $\pi(\boldsymbol{\theta}_{k+1}|D_k) = \int \pi(\boldsymbol{\theta}_{k+1}|\boldsymbol{\theta}_k)\pi(\boldsymbol{\theta}_k|x_k, n_k)\, d\boldsymbol{\theta}_k$. Thus, the distribution of the d-dimensional state vector $\boldsymbol{\theta}_k$ gets approximated through

$$\pi(\boldsymbol{\theta}_{k+1}|D_k) \approx \frac{1}{N}\sum_{i=1}^{N}\pi(\boldsymbol{\theta}_{k+1}|\boldsymbol{\theta}_k^i) = \frac{1}{N}\sum_{i=1}^{N}\frac{\exp\left(-\frac{1}{2}(\boldsymbol{\theta}_{k+1} - \boldsymbol{G}_k\boldsymbol{\theta}_k^i)'\boldsymbol{W}_k^{-1}(\boldsymbol{\theta}_{k+1} - \boldsymbol{G}\boldsymbol{\theta}_k^i)\right)}{\sqrt{(2\pi)^d|\boldsymbol{W}_k|}},$$

where $\{\boldsymbol{\theta}_k^i\}_{i=1}^N$ is a sample of $\boldsymbol{\theta}_k|x_k, n_k$ from step 1. The one step ahead predictive distribution for state $\boldsymbol{\vartheta}_k$ is obtained with the DLM equations in Appendix A.

After that, we would be back at step 0, restarting the process. The above can be grouped into a scheme similar to Algorithm 2 in Data S2.

5 | FORECASTING SEVERITIES

We also need to predict how many of the x_k occurrences in the kth period correspond to the five severity classes. Let $\boldsymbol{p} = (p_1, p_2, p_3, p_4, p_5)$ be a vector representing their proportions, with $p_i \geq 0$, $\sum_{i=1}^5 p_i = 1$; $\boldsymbol{s} = (s_1, s_2, s_3, s_4, s_5)$ be the vector with the number of occurrences of each severity, with $s_i \geq 0$ and $\sum_{i=1}^5 s_i = x_k$; and $D_k = \{(s_1^j, s_2^j, s_3^j, s_4^j, s_5^j)\}_{j=1}^{k-1}$ the data at the beginning of the kth period, where s_i^j is the number of occurrences of severity i in period j.

In our problem, the number x_k of occurrences in the kth period is unknown, and predicted as in Sections 3 and 4. For example, if we consider the initial basic model for x_k and a Multinomial-Dirichlet model for the severity, Figure 10, we have

$$\begin{aligned} x_k &\sim \mathrm{Po}\,(\lambda n_k)\,, \qquad \lambda \sim \mathrm{Ga}(a, b), \\ \boldsymbol{s}|\boldsymbol{p}, x_k &\sim \mathcal{M}(x_k; p_1, p_2, p_3, p_4, p_5), \\ \boldsymbol{p} &\sim \mathcal{D}ir(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5). \end{aligned}$$

The predictions would be

$$\begin{aligned} Pr(s_i|D_k) &= \sum_{r=s_i}^{\infty}\binom{r}{s_i}\frac{\mathrm{B}(s_i + \alpha_i', A + r - s_i - \alpha_i')}{\mathrm{B}(\alpha_i', A - \alpha_i')}\frac{b_k^{a_k}}{(b_k+1)^{a_k+r}}\frac{\Gamma(a_k + r)}{r!\,\Gamma(a_k)} \\ &\approx \sum_{r=s_i}^{h}\binom{r}{s_i}\frac{\mathrm{B}(s_i + \alpha_i', A + r - s_i - \alpha_i')}{\mathrm{B}(\alpha_i', A - \alpha_i')}\frac{b_k^{a_k}}{(b_k+1)^{a_k+r}}\frac{\Gamma(a_k + r)}{r!\,\Gamma(a_k)}, \end{aligned}$$

for big enough h, with $\alpha_i' = \alpha_i + \sum_{j=1}^{k-1} s_i^j$, $A = \sum_{i=1}^5 \alpha_i'$, $a_k = a + \sum_{j=1}^{k-1} x_j$, and $b_k = b + \sum_{j=1}^{k-1} n_j$. The predictive expected number of occurrences is $E(s_i|D_k) = E(x_k)E(p_i)$.

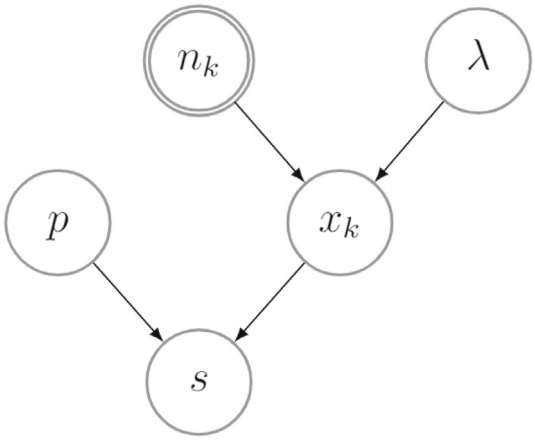

FIGURE 10 Influence diagram to forecast aviation occurrences severity

As an example, with the basic model, Section 3.1, $E(s_i|D_k) = (x_k a_k \alpha_i')/(b_k A)$, and the predictive variance is $E(p_i(1-p_i))E(x_k^2) + (E(p_i))^2 Var(x_k)$. Notice that although shown here in combination with the most basic model, extensions to other cases follow the same path, for example, as illustrated in Section 6.1.

5.1 | The problem of underreporting

A major obstacle to forecast occurrences, and AS management in general, would be the unavailability of accurate data, which could hide latent conditions that end up causing more severe ones. Indeed, in the absence of a strong reporting culture among the agents involved (pilots, air controllers,...) it would be common not to report low severity occurrences.[32]

In a case with suspected underreporting, we can apply the logic of the influence diagram in Figure 11, based on our basic model in Section 3.1. Again, extensions to the other models can be developed. Introduce vector of reported occurrences for each severity class, $\boldsymbol{z} = (z_1, z_2, z_3, z_4, z_5)$, and vector with the proportion of reported occurrences $\boldsymbol{\varrho} = (\varrho_1, \varrho_2, \varrho_3, \varrho_4, \varrho_5)$. Partial information about ϱ_i is typically available, in particular, $\varrho_1 = 1$, $\varrho_2 \approx 1$ and $\varrho_2 > \varrho_3 > \varrho_4 > \varrho_5$, that is, as occurrences get less severe they are less likely to be reported. Thus, additional features to model in Figure 10, are $\boldsymbol{\varrho}$, with $\varrho_i \sim \text{Be}(\gamma_i, \beta_i)$, $i = 1, \ldots, 5$; and $\boldsymbol{z}$, with $z_i|s_i, \varrho_i \sim \text{Bin}(s_i, \varrho_i)$, $i = 1, \ldots, 5$.

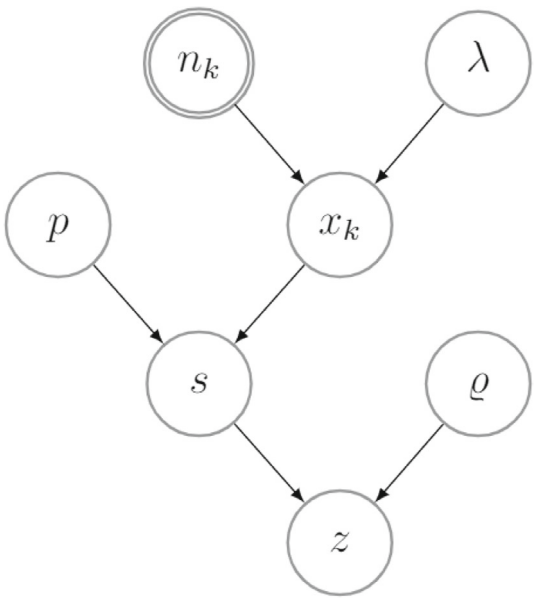

FIGURE 11 Model for the underreporting problem

At the beginning of each period, our goal is to predict $\boldsymbol{s}$, x and $\boldsymbol{z}$; then, after observing $\boldsymbol{z}$, to propagate that information to the different levels of the graph in Figure 11. To simplify the problem, assume $\boldsymbol{p}$ known[1]. Due to the Poisson process partition property, $s_i|\lambda \sim \text{Po}(p_i\lambda)$.[26] Furthermore,

$$\Pr(Z_i = z_i|\lambda, \varrho_i) = \sum_{s_i=z_i}^{\infty} \Pr(Z_i = z_i|s_i, \varrho_i)\Pr(S_i = s_i|\lambda) = \frac{(p_i\lambda\varrho_i)^{z_i}}{z_i!}\exp(-p_i\lambda\varrho_i).$$

Hence, $z_i|\lambda, \varrho_i \sim \text{Po}(p_i\lambda\varrho_i)$, and the likelihood of the observed data $\boldsymbol{z}$ is

$$\pi(\boldsymbol{z}|\lambda, \boldsymbol{\varrho}) \propto \lambda^{\sum z_i}\exp\left(-\lambda\sum_{i=1}^{5}p_i\varrho_i\right)\prod_{i=1}^{5}\varrho_i^{z_i}.$$

Assuming all parameters independent, and under the noninformative prior $\pi(\lambda) \propto \lambda^{-1}$, the posterior distribution would be

$$\pi(\lambda, \boldsymbol{\varrho}|\boldsymbol{z}) \propto \lambda^{\sum z_i - 1}\exp\left(-\lambda\sum_{i=1}^{5}p_i\varrho_i\right)\prod_{i=1}^{5}\varrho_i^{z_i+\gamma_i-1}(1-\varrho_i)^{\beta_i-1},$$

and the posterior conditionals are

$$\pi(\lambda|\boldsymbol{\varrho}, \boldsymbol{z}) \propto \lambda^{\sum z_i - 1}\exp\left(-\lambda\sum_{i=1}^{5}p_i\varrho_i\right) \sim \text{Ga}\left(\sum_{i=1}^{5}p_i\varrho_i, \sum_{i=1}^{5}z_i\right)$$
$$\pi(\varrho_i|\lambda, \boldsymbol{z}) \propto \varrho_i^{z_i+\gamma_i-1}(1-\varrho_i)^{\beta_i-1}\exp(-\lambda p_i\varrho_i), \quad i = 1, \dots, 5.$$

The last distributions are not standard, but are unimodal and log-concave, so it is easy to sample from them as in Algorithm 5 in Data S2, obtaining samples $\{(\lambda^j, \varrho_1^j, \dots, \varrho_5^j)\}_{j=1}^N$. Our interest lies in sampling from the distributions $s_i|\boldsymbol{z}$. Observe that

$$Pr(s_i|\boldsymbol{z}) = \iint Pr(s_i|\lambda, \boldsymbol{\varrho}, \boldsymbol{z})\pi(\lambda, \boldsymbol{\varrho}|z)\, d\lambda\, d\boldsymbol{\varrho}.$$

Then $s_i|\lambda, \boldsymbol{\varrho}, \boldsymbol{z} \sim z_i + Po(p_i\lambda(1-\varrho_i))$, and $s_i^j \sim z_i + Po(p_i\lambda^j(1-\varrho_i^j))$ constitutes a sample from $s_i|\boldsymbol{z}$. We summarize it with $\frac{1}{N}\sum_{j=1}^N s_i^j$ which approximates $E(s_i|\boldsymbol{z})$.

6 | CASES

As application examples, we present the models used for the *wind shear* and *TCAS warning* occurrences, and to a simulated occurrence type showing dependence. For space reasons, more emphasis is placed on the *wind shear* model, because it is the most versatile and used in our domain. Core ideas are given for the other two.

6.1 | Wind Shear

Wind shear consists of a change in wind speed and/or direction over a short distance.[33] It can occur either horizontally or vertically, at high or low altitude, most often associated with strong temperature inversions or density gradients. It may significantly affect the airspeed and trajectory of a plane, being more dangerous the closer to the ground and the slower the aircraft is. Therefore, AS occurrences reported in relation to wind shear usually happen during take-off or landing.

[1]Extensions to unknown p_i's follow in a straightforward manner.

6.1.1 | Exploratory analysis

Table 2 displays the evolution of the number of occurrences from 2010 to 2018, the number of operations (in blocks of 100,000), the occurrence rate (number of occurrences per 100,000 operations) as well as the evolution for the five severities. As we see, the occurrence rate has been growing annually, especially during the first 5 years, then stabilizing. Note that in 2011 there were 23% less occurrences compared to the previous year, while the number of operations increased slightly.

Regarding occurrence severity, Figure 12, observe that every year, severity 4 occurrences were the most reported, followed by those of severity 5. Finally, note that there has been only two severity 2 occurrences, and one severity 1 during the considered period, suggesting that this event is not very severe.

6.1.2 | Effects

Graphical and numerical analyses used to identify the relevant effects follow.

Stress effect. Figure 13A shows the scatter plot for the number of operations versus occurrence rates, as well as the regression line relating both variables. The correlation coefficient is -0.23 and no *stress effect* is included.

Seasonal effect. The monthly ACF is in Figure 13B suggests a seasonal effect, through the relevance of the autocorrelation of lag 12, due to weather relevance over this phenomenon. In addition, the first ones, although not strong, are relevant, suggesting a relationship between rates at consecutive months.

TABLE 2 Number of occurrences and operations for *wind shear*, 2010–2018

	Sev. 1	Sev. 2	Sev. 3	Sev. 4	Sev. 5	Total occ.	Ops.	Occ. rate
2010	0	0	5	113	9	127	21.20	5.99
2011	1	0	0	91	5	97	21.40	4.53
2012	0	0	4	160	23	187	19.25	9.71
2013	0	1	2	265	7	275	17.91	15.35
2014	0	1	5	357	46	409	18.33	22.31
2015	0	0	1	385	24	410	19.03	21.55
2016	0	0	2	474	13	489	20.45	23.91
2017	0	0	2	511	10	523	21.74	24.05
2018	0	0	2	518	8	528	23.00	22.95

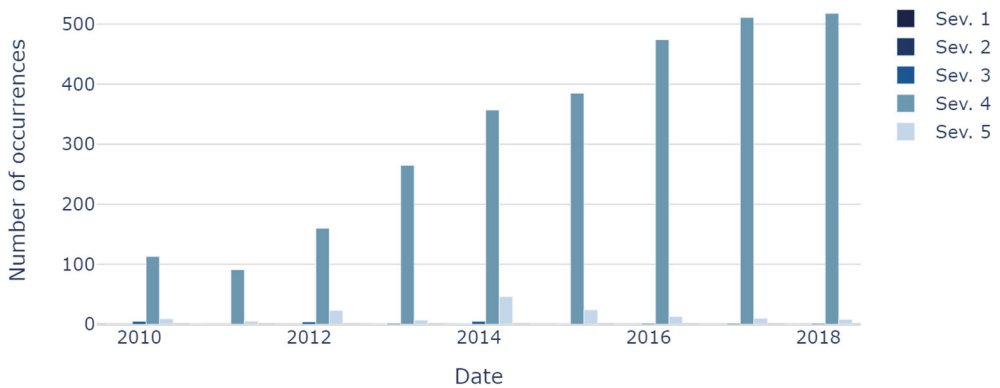

FIGURE 12 *Wind shear* occurrences, period 2010–2018, by severity

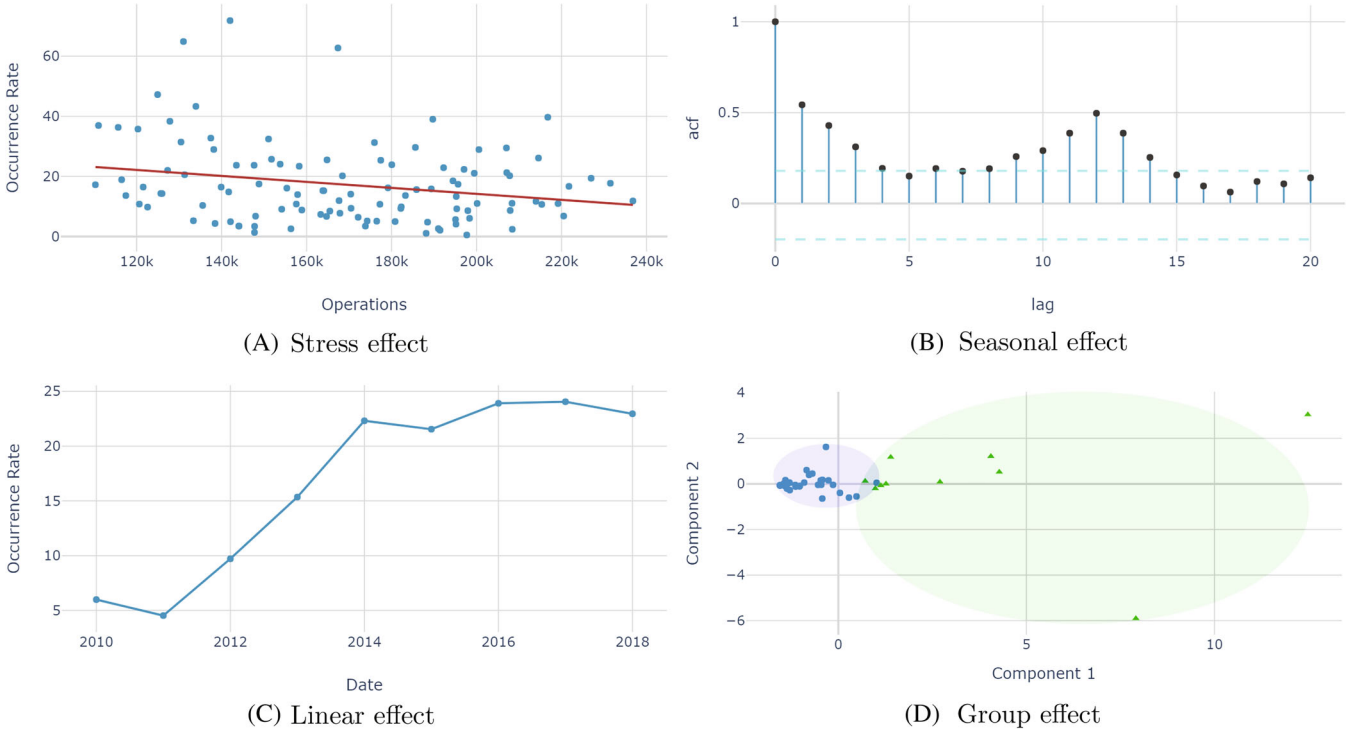

FIGURE 13 Effect analysis for *wind shear*. (A) Stress effect; (B) Seasonal effect; (C) Linear effect; (D) Group effect

Linear effect. Figure 13C represents the annual evolution of occurrence rates. The annual time series suggests a linear increase of wind shear occurrence rates during the first 5 years. The effect is considerable because, except for year 2011, in which it was slightly less than the previous one, the rate has grown annually.

Group effect. A cluster analysis allows us to identify two airport groups with similar *wind shear* occurrence rate, Figure 13D. The first one (triangles) includes the 10 airports in temperate coastal areas. The second group (circles) would be formed by the remaining airports. Because of the climate differences, we deal with both groups hierarchically and aggregate the forecasts. Note that this would also allow for a certain *lag* between the seasonalities in groups, which for example would be relevant when dealing with occurrences related to migratory birds that arrive at airports at different times of the year.

6.1.3 | Model

We thus have detected a seasonal effect of period 12, a (possible) linear growth effect and two groups of airports. Hence, we consider a hierarchical model for the occurrences x_k^i at group i of airports based on the Poisson DLM of Section 3.2.2, with a linear growth component, a seasonal component of period 12, and a common prior,

$$\begin{aligned}
&x_k^i \,|\, \lambda_k^i, n_k^i \sim Po(\lambda_k^i n_k^i), \quad \lambda_k^i = \exp(u_k^i), \qquad i = 1, 2\\
&u_k^i = \boldsymbol{F}\boldsymbol{\theta}_k^i + v^i, \quad v^i \sim N\left(0, V^i\right),\\
&\boldsymbol{\theta}_k^i = \boldsymbol{G}\boldsymbol{\theta}_{k-1}^i + \boldsymbol{w}^i, \quad \boldsymbol{w}^i \sim N\left(\boldsymbol{0}, \boldsymbol{W}^i\right), y\\
&\boldsymbol{\theta}_0^i \sim N(\boldsymbol{m}_0, \boldsymbol{C}_0),
\end{aligned} \tag{4}$$

where $\boldsymbol{F} = (\boldsymbol{F}_1 \quad \boldsymbol{F}_2)$ and $\boldsymbol{G} = \text{blockdiag}\,(\boldsymbol{G}_1 \quad \boldsymbol{G}_2)$. Matrices $\boldsymbol{W}^i$ and V^i are initialized based on the observations, using maximum likelihood.[34] To complete model specification, we need the prior moments $\boldsymbol{m}_0$ and $\boldsymbol{C}_0$. These priors aim to be flexible enough and somewhat informative. The prior mean vector, $\boldsymbol{m}_0$, is initialized[2]using the the first year of data

[2] m_0^{Lj} and m_0^{Sj} indicate the jth parameters of the linear growth and seasonal blocks, respectively.

(12 observations). Parameter m_0^{L2} describes the expected growth and it is initialized with $m_0^{L2} = (y_{12} - y_1)/11$, where $y_k = \log(\sum x_k^i / \sum n_k^i)$; m_0^{L1} describes the expected level and is initialised with $m_0^{L1} = (\sum_{k=1}^{12} y_k - 78\ m_0^{L2})/12$; m_0^{Sj} describes the jth seasonal component and, to assess it, we use $m_0^{Sj} = y_{13-j} - m_0^{L1} + (j-13)\ m_0^{L2}$. Hence, for both groups of airports, we have

$$\boldsymbol{m}_0 = (\overbrace{1.6}^{m_0^{L1}}, \overbrace{0.0}^{m_0^{L2}}, \overbrace{-0.3}^{m_0^{S1}}, \overbrace{-0.7}^{m_0^{S2}}, \overbrace{0.0}^{m_0^{S3}}, \overbrace{-0.9}^{m_0^{S4}}, \overbrace{0.6}^{m_0^{S5}}, \overbrace{0.8}^{m_0^{S6}}, \overbrace{0.1}^{m_0^{S7}}, \overbrace{-0.6}^{m_0^{S8}}, \overbrace{-0.4}^{m_0^{S9}}, \overbrace{0.6}^{m_0^{S10}}, \overbrace{1.1}^{m_0^{S11}})'$$

For more accuracy, we could repeat the calculations as many times as years of data are available, and take the average value for each parameter. Since we are confident about the suitability of our $\boldsymbol{m}_0$, we use a relatively small prior variance $\boldsymbol{C}_0 = \boldsymbol{I}_{13}/10$. Also, for both groups of airports, we include the prior parameters α_j ($j = 1, \ldots, 5$) for the different severities and, based on expert judgement, set at 1, 2, 3, 7, and 5, respectively, not very high to facilitate learning.

We then adjust the previous models using the approximations described in Section 3.2.2 (implemented via a hierarchical version of Algorithm 2 in Data S2) and obtain two samples, one for the predictive distribution of each group of airports. The aggregation of both samples facilitates a predictive sample for the total number of wind shear occurrences. Figure 14 shows 1-month ahead predictions for 2010–2018 observations (black dots), the predictive mean (solid line) and the 90% probability band (dashed lines). We also show the 12-step ahead forecast for year 2019 (without actual observations, since they were not available), for which the uncertainty in the future number of operations is modeled as in Section 4.

Routine forecasting and monitoring will be responsible for checking the stability of the model and suggesting anomalies, sudden instabilities, and deterioration in forecast performance that have not been anticipated through expert intervention. We would therefore raise alarms whenever observed values lay outside predictive intervals, like the observations marked with a circumference in Figure 14. As Figure 15 shows, the credible intervals for one-step ahead predictions (solid blue) adequately capture the observations, showing only slight under-coverage for credible intervals between 50% and 95%.

Compared to other popular models used with non-negative integer time series, like generalized linear ARMA (GLARMA[14]), and integer-valued GARCH (INGARCH[15]), Table 3 shows that our proposed model (4) offers significantly better point forecasts than any of them, using either a Poisson or Negative Binomial distribution for the observations, or a standard DLM. Additionally, as seen in Figure 15, the credible intervals (solid blue line) of our model closely match the nominal coverage probability (45 degree dashed line) and perform nicely when compared with the analyzed competitors.

The analysis is completed with the prediction of the number of occurrences of each severity class. Table 4 shows a summary of the predictions, with μ' and σ'designating the predictive mean and standard deviation of the number of occurrences for the next period to forecast (in this case, the next month January 2019), and α_j', the parameter of the posterior distribution of the jth severity class of such event. Thus, the expected number of occurrences for each severity would be $\mu'\alpha_j' / \sum_{h=1}^{5} \alpha_h'$, for example, 33.27 severity 4 occurrences.

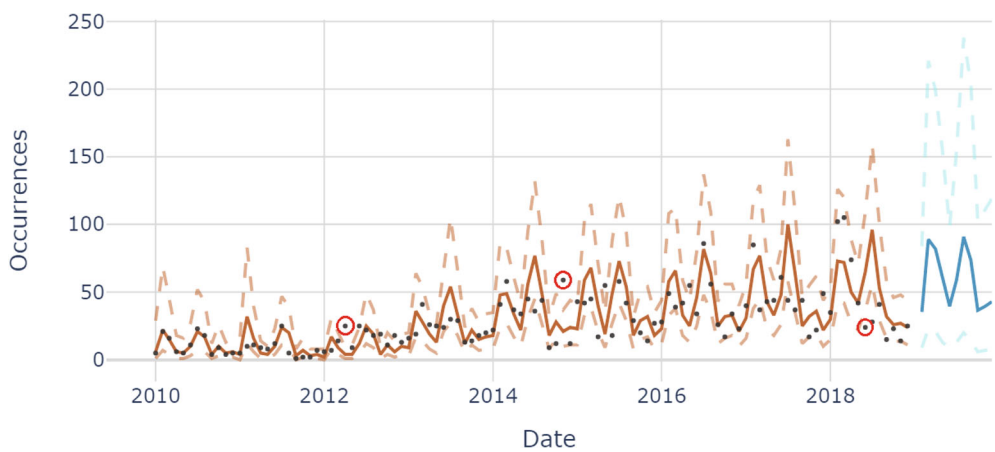

FIGURE 14 Prediction of number of occurrences

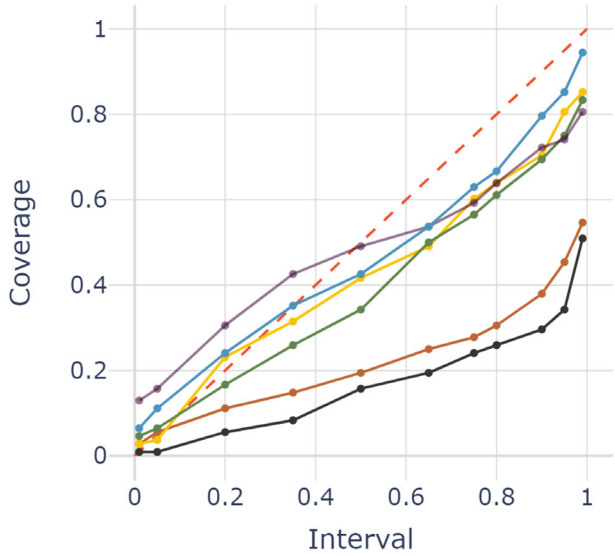

FIGURE 15 Empirical (solid) versus nominal (dashed) coverage of credible intervals for *wind shear*. Besides our proposed model in blue, also shows coverage for Poisson generalized linear ARMA (GLARMA) (brown), NB GLARMA (yellow), Poisson integer-valued GARCH (INGARCH) (black), NB INGARCH (green) and dynamic linear models (purple)

TABLE 3 Error metrics for the predictive median of various models as point forecast for *wind shear*

	Our model	GLARMA (Poi)	GLARMA (NB)	INGARCH (Poi)	INGARCH (NB)	DLM
MSE	296.74	350.33	392.35	425.51	471.22	379.7
MAE	11.96	13.35	14.03	14.81	15.37	14.53
MAPE	0.53	0.63	0.65	0.61	0.59	0.8
Theil's U	0.91	0.94	0.99	1.03	1.09	1.03

Abbreviations: GLARMA, generalized linear ARMA; INGARCH, integer-valued GARCH.

TABLE 4 Prediction summary

μ'	σ'	α'_1	α'_2	α'_3	α'_4	α'_5
35.38	26.55	2	4	26	2881	150

Regarding sensitivity to the hyper-parameters $\boldsymbol{m}_0$, $\boldsymbol{C}_0$ for the prior of the initial state; the election of a different $\boldsymbol{m}_0$, for example, the usual vector of zeros which does not use prior information, has little effect on forecast performance beyond the first observations (algorithm particles still arrive relatively quickly to a zone with high probability) unless we deviate a lot from these values for the states (with prior values for all states outside the interval $[-2, 2]$). Moving from the proposed variance $\boldsymbol{C}_0 = \boldsymbol{I}_{13}/10$ up to $\boldsymbol{I}_{13}$ results in too much dispersion and very high predictive intervals during the first observations, more resamples and overall worse forecasting performance; using lower values for the diagonal down to $\boldsymbol{I}_{13}/100$ also worsens the point forecast metrics and coverage of the predictive distributions, although less dramatically.

6.2 | TCAS warnings

Traffic Collision Avoidance Systems (TCAS) warn pilots of the presence of other aircraft which may present a threat of mid-air collisions. This type of occurrences, unlike *wind shear*, presents a *stress effect*, Figure 1B. Since it does not show any other of the effects mentioned in Section 2, we model it through *stress effect* model in Section 3.2.1.

Using Algorithm 1 we obtain the one-month ahead forecasts in Figure 16, with predictive mean (solid line) and 99% probability band (dashed lines), and we can check that it adequately predicts time series with this AS specific effect. Figure 17 shows that the predictive distributions adequately cover the observations at different credible intervals (blue line) and that there is an improvement in coverage over the *basic* model (yellow) from Section 3.1; this is also the case for

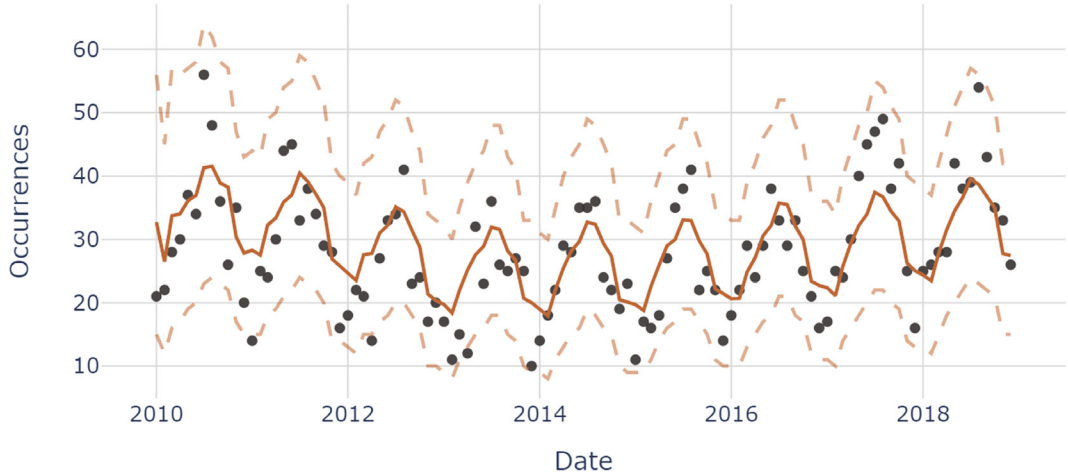

FIGURE 16 Forecast of *Traffic Collision Avoidance Systems warnings*

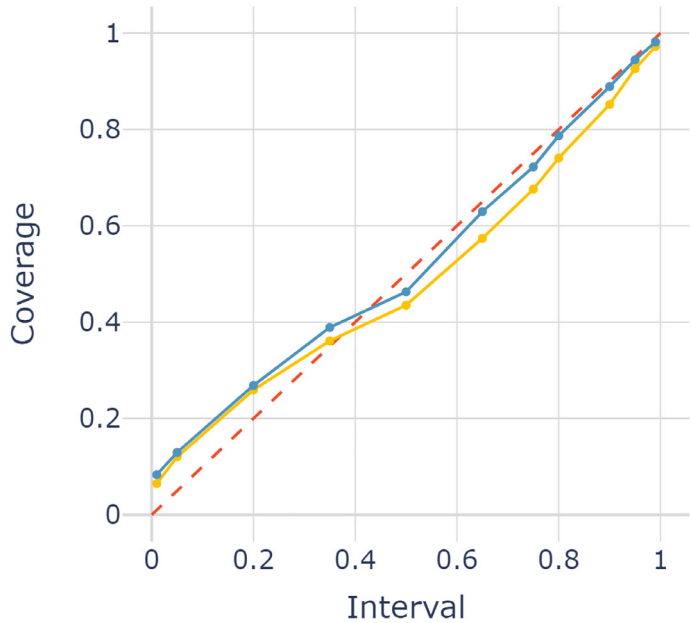

FIGURE 17 Covplot of *Traffic Collision Avoidance Systems* occurrence

the error metrics of the point forecasts (Table 5). Furthermore, as Figure 18 shows, the posterior distribution of parameter a in the model concentrates around 12, away from 0, which is consistent with the relevance of a *stress effect*.

A sensitivity analysis concerning the hyper-parameters of the priors over a, b, and σ^2 suggests that, as long as plausible values are chosen, the performance of the *stress effect* model (1) is robust. For example, for *TCAS warnings*, even if we would to select negative values for the means of a and $b(\mu_a = \mu_b = -5)$, indicating a negative stress effect, with variances $\sigma_a^2 = \sigma_b^2 = 10$ we arrive to similar posterior distributions. Hence we recommend the election of means consistent with the available data and relatively high variances.

6.3 | Dependence of occurrence types

The relevance of the *dependence* model in Section 3.2.4 can be readily exemplified with simulated data of a new occurrence type (Figure 19) that shows dependence with *TCAS warnings*. The use of the *dependence* model and Algorithm 4 to sample from it, improves the forecast performance over the basic model from Section 3.1 that assumes independence as shown in Table 6. Similarly to Algorithm 1, the proposed MCMC sampler is quite robust for reasonable choices of the values in the hyper-parameters of the priors.

TABLE 5 Error metrics *traffic collision avoidance systems* occurrence

	Stress model	Basic model
MSE	36.57	45.67
MAE	4.78	5.28
MAPE	0.2	0.22
Theil's U	0.77	0.82

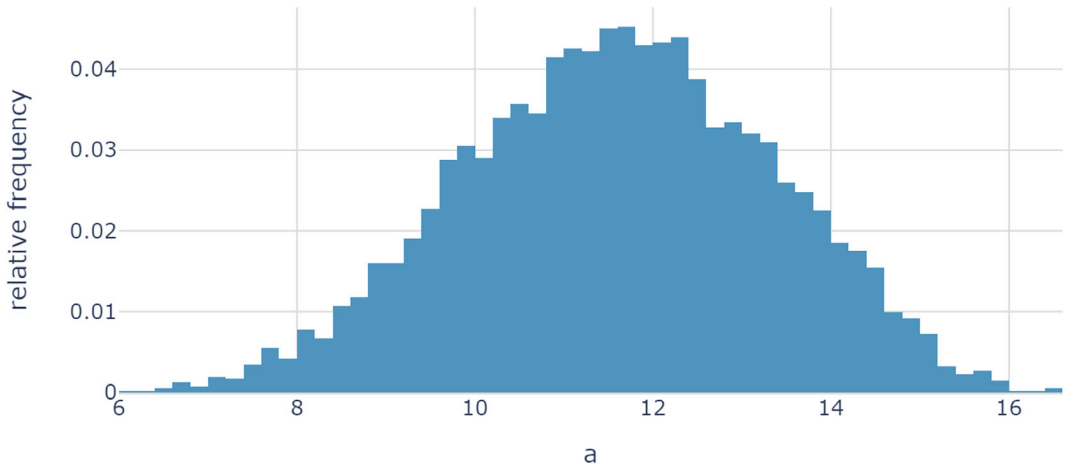

FIGURE 18 Posterior distribution of *a*

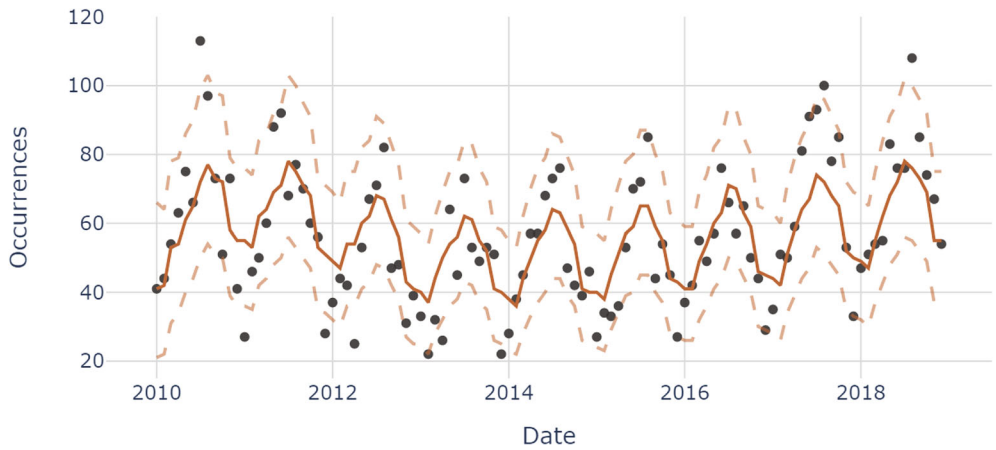

FIGURE 19 Forecast of *dependent* occurrence

TABLE 6 Error metrics *dependent* occurrence

	Dependence model	Basic model
MSE	159.54	245.25
MAE	9.89	11.51
MAPE	0.21	0.22
Theil's U	0.78	0.94

7 | DISCUSSION

We have provided a methodology to forecast aviation safety occurrences and their severities, based on an initial standard model, suitable for situations in which the occurrence rate remains relatively stable over the period of interest. In most cases, various effects impact the rate evolution. Thus, we adapted that model by adding specific components (stress effect, seasonal and trend effect, group effect and dependence) and proposed algorithms to forecast with these new models. Several of them need to be combined for certain occurrences, as shown with the case in Section 6.1. In addition, we have described a model to predict the severity of AS occurrences according to the classification proposed by the ICAO. The above models are suitable when all the information about different types of occurrence is available. However, in some cases underreporting needs to be addressed, possibly as suggested.

The proposed models are fundamental in our risk management methodology[3] feeding our AS resource allocation models. They are also important in predicting and monitoring events that allow identifying anomalies related to an unexpected increase (or decrease) in the number of occurrences. In particular, the methodology emphasizes a *management by exception principle*[28] with our models used for routine inference, prediction (and decision support) under standard circumstances until exceptional ones arise in which case an intervention is requested.

The forecasting performance of our model was compared to other popular ones like DLM, generalized linear ARMA (GLARMA), and integer-valued GARCH (INGARCH) models, showing better forecasting performance with the AS time series studied. However, some of these models assuming negative binomially distributed observations might be more relevant when exploring approaches at smaller time (weeks) and spatial (airport) frames, which might present more over-dispersion. Also, given the high safety levels in the aviation system we should expect numerous zero counts and, in particular, models such as those in Berry and West[35] would be relevant. Future work includes exploring these and other approaches needed for modeling time series with frequent zeros and different levels of dispersion in the observations.

Finally, although we have applied these models to forecast aviation safety occurrences, they can also be used to predict safety and reliability occurrences in other areas, like maritime transport, industry, or supply chain networks.

ACKNOWLEDGMENTS

Research supported by the Spanish Ministry of Economy program MTM2017-86875-C3-1-R AEI/FEDER, UE; Community of Madrid program IND2018/TIC-9901 FEDER, UE; the AXA-ICMAT Chair on Adversarial Risk Analysis and the RAC-AESA Agreement. The authors would like to thank Fran Bernal, Veronica Elvira and Pablo Hernandez-Coronado, from AESA, for the insights provided during the development of this work. Referees' suggestions are also deeply appreciated.

DATA AVAILABILITY STATEMENT

Actual data from the Aviation Safety Agency cannot be freely distributed, but an anonymized version of the original data is provided to illustrate the use of the proposed models.

ORCID

Bruno Flores https://orcid.org/0000-0002-7870-9103
David Rios Insua https://orcid.org/0000-0002-5748-9658
Cesar Alfaro https://orcid.org/0000-0002-9146-4067
Javier Gomez https://orcid.org/0000-0001-6434-7263

REFERENCES

1. ICAO. State of global aviation safety. Technical report; 2019.
2. IATA. Economic performance of the airline industry. Technical report; 2019.
3. Rios Insua D, Alfaro C, Gomez J, Hernandez-Coronado P, Bernal F. A framework for risk management decisions in aviation safety at state level. *Reliab Eng Syst Saf*. 2018;179:74-82.
4. Elvira V, Bernal F, Hernandez-Coronado P, et al. Safer skies over spain. *INFORMS J Appl Anal*. 2020;50(1):21-36.
5. Clemen R, Reilly T. *Making Hard Decisions with DecisionTools*. Cengage Learning; 2013.
6. Rios Insua D, Alfaro C, Gomez J, Hernandez-Coronado P, Bernal F. Forecasting and assessing consequences of aviation safety occurrences. *Saf Sci*. 2019;111:243-252.
7. McCann DW. NNICE – A neural network aircraft icing algorithm. *Environ Model Softw*. 2005;20:1335-1342.
8. Gill PG, Buchanan P. An ensemble based turbulence forecasting system. *Royal Meteorol Soc Meteorol Appl*. 2013;21:12-19.
9. Ayra ES, Rios Insua D, Cano J. Bayesian network for managing runway overruns in aviation safety. *J Aerosp Inf Syst*. 2019;16(12):546-558.

10. Khanmohammadi S, Tutun S, Kucuk Y. A new multilevel input layer artificial neural network for predicting flight delays at JFK airport. *Proc Comput Sci*. 2016;95:237-244.
11. Subramanian SV, Rao AH. Deep-learning based time series forecasting of go-around incidents in the national airspace system. Proceedings of the AIAA Modeling and Simulation Technologies Conference; 2018.
12. Cox DR. Statistical analysis of time series: some recent developments. *Scand J Stat*. 1981;8(2):93-115.
13. Alzaid AA, Al-osh M. An integer-valued pth-order autoregressive structure (INAR(p)) process. *J Appl Probab*. 1990;27(2):314-324.
14. Benjamin MA, Rigby RA, Stasinopoulos DM. Generalized autoregressive moving average models. *J Am Stat Assoc*. 2003;98(461):214-223.
15. Ferland R, Latour A, Oraichi D. Integer-valued GARCH process. *J Time Ser Anal*. 2006;27(6):923-942.
16. Heinen A. Modelling Time Series Count Data: An Autoregressive Conditional Poisson Model. *SSRN Electronic Journal*. 2003. doi:10.2139/ssrn.1117187
17. West M. Bayesian forecasting of multivariate time series: scalability, structure uncertainty and decisions (with discussion). *Ann Inst Stat Math*. 2020;72:1-44.
18. Aktekin T, Polson NG, Soyer R. Sequential Bayesian analysis of multivariate count data. *Bayesian Anal*. 2018;13(2):385-409.
19. Chen X, Irie K, Banks D, Haslinger R, Thomas J, West M. Scalable Bayesian modeling, monitoring and analysis of dynamic network flow data. *J Am Stat Assoc*. 2018;113:519-533.
20. Aktekin T, Soyer R. Call center arrival modeling: a Bayesian state-space approach. *Naval Res Logist (NRL)*. 2011;58(1):28-42.
21. Gamerman D, dos Santos TR, Franco GC. A non-gaussian family of state-space models with exact marginal likelihood. *J Time Ser Anal*. 2013;34(6):625-645.
22. Snyder RD, Martin GM, Gould P, Feigin PD. An assessment of alternative state space models for count time series. Monash econometrics and business statistics. Working papers 4/07, Monash University, Department of Econometrics and Business Statistics; 2008.
23. Yelland PM. Bayesian forecasting for low-count time series using state-space models: an empirical evaluation for inventory management. *Int J Prod Econ*. 2009;118:95-103.
24. Shachter R. Probabilistic inference and influence diagrams. *Oper Res*. 1988;36(4):589-604.
25. ICAO. *Safety Management Manual*. 4th ed. International Civil Aviation Organization (ICAO); 2018.
26. Rios Insua D, Ruggeri F, Wiper M. *Bayesian Analysis of Stochastic Process Models*. Vol 978. John Wiley & Sons; 2012.
27. Rios Insua S, Martin J, Rios Insua D, Ruggeri F. Bayesian forecasting for accident proneness evaluation. *Scand Actuar J*. 1999;1999(2):134-156.
28. West M, Harrison J. *Bayesian Forecasting and Dynamic Models*. Springer; 1997.
29. Prado R, West M. *Time Series: Modelling, Computation & Inference*. Chapman & Hall/CRC Press; 2010.
30. Storvik G. Particle filters for state-space models with the presence of unknown static parameters. *IEEE Trans Signal Process*. 2002;50(2):281-289.
31. McKay AT. Sampling from batches. *Suppl J R Stat Soc*. 1934;1(2):207-216.
32. Haslbeck A, Schmidt-Moll C, Schubert E. Pilot's willingness to report aviation incidents. Proceedings of the International Symposium on Aviation Psychology; 2015.
33. FAA. Wind shear. Technical report; 2008.
34. Petris G, Petrone S, Campagnoli P. *Dynamic Linear Models with R*. Springer; 2009.
35. Berry LR, West M. Bayesian forecasting of many count-valued time series. *J Bus Econ Stat*. 2020;38:872-887.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Flores B, Rios Insua D, Alfaro C, Gomez J. Forecasting aviation safety occurrences. *Appl Stochastic Models Bus Ind*. 2022;38(3):545-567. doi: 10.1002/asmb.2675

APPENDIX A . DYNAMIC LINEAR MODELS

A normal dynamic linear model (DLM) for univariate observations X_t, specified by the quadruple $\{\boldsymbol{F}_t, \boldsymbol{G}_t, V_t, \boldsymbol{W}_t\}$, is defined through

$$
\begin{aligned}
x_t &= \boldsymbol{F}_t \boldsymbol{\theta}_t + v_t, & v_t &\sim N(0, V_t),\\
\boldsymbol{\theta}_t &= \boldsymbol{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{w}_t, & \boldsymbol{w}_t &\sim N(\boldsymbol{0}, \boldsymbol{W}_t),\\
\boldsymbol{\theta}_0 &\sim N(\boldsymbol{m}_0, \boldsymbol{C}_0),
\end{aligned}
$$

with v_t and w_t internally and mutually independent.[28]

For the univariate DLM, if we denote the available information at the beginning of period t as $D_t = \{D_{t-1}, x_{t-1}\}$, the sequential update and forecast procedure is given by the recursion:

- One-step ahead predictive distribution of $\boldsymbol{\theta}_t$, given D_t. It is $N(\boldsymbol{a}_t, \boldsymbol{R}_t)$, with $\boldsymbol{a}_t = \boldsymbol{G}_t \boldsymbol{m}_{t-1}$ and $\boldsymbol{R}_t = \boldsymbol{G}_t \boldsymbol{C}_{t-1} \boldsymbol{G}_t' + \boldsymbol{W}_t$.
- One-step ahead predictive distribution of x_t, given D_t. It is $N(f_t, Q_t)$, with $f_t = \boldsymbol{F}_t \boldsymbol{a}_t$ and $Q_t = \boldsymbol{F}_t \boldsymbol{R}_t \boldsymbol{F}_t' + V_t$.
- Filtering or posterior distribution of $\boldsymbol{\theta}_t$, given D_t and x_t. It is $N(\boldsymbol{m}_t, \boldsymbol{C}_t)$, with $\boldsymbol{m}_t = \boldsymbol{a}_t + \boldsymbol{R}_t \boldsymbol{F}_t' Q_t^{-1}(x_t - f_t)$ and $\boldsymbol{C}_t = \boldsymbol{R}_t - \boldsymbol{R}_t \boldsymbol{F}_t' Q_t^{-1} \boldsymbol{F}_t \boldsymbol{R}_t$.