Moving Integrated Product Development to Service Clouds in the Global Economy
J. Cha et al. (Eds.)

doi:10.3233/978-1-61499-440-4-706

Stochastic Forecasting of Lumpy-distributed Aircraft Spare Parts Demand

Wim J.C. VERHAGEN[a,1] and Richard CURRAN[a]
[a] *Delft University of Technology, Kluyverweg 1, Delft, The Netherlands, 2629 HS*

Abstract. An approach is proposed to forecast lumpy spare parts demand associated with non-routine aircraft maintenance based on a stochastic model. The model assumes demand arrival in the form of a Homogeneous Poisson Process (HPP). The developed model is applied to a dataset from an aircraft maintenance operator consisting of lumpy spare parts demand for nine Line Replacable Units (LRUs) from a shared customer pool of parts, which are associated with significant repair turnaround time delays. To meet the operator's service level requirements and associated stocking strategy, the model is evaluated for its capacity to forecast peak demand, rather than its capacity to minimize forecasting error. Under these terms, the proposed model outperforms the existing company approach.

Keywords. Forecasting, Spare Parts Demand, Homogeneous Poisson Process

Introduction

Maintenance is an integral part of the aircraft lifecycle and ensures the continued safe operation of aircraft. The maintenance function is subject to diverse and contrary high-level regulatory and business requirements: to ensure product safety (as embodied in aircraft airworthiness), at minimal maintenance effort in terms of time and cost, while optimizing aircraft availability for revenue-generating purposes.

Maintenance operations can be subdivided into routine and non-routine maintenance. Routine maintenance tasks have specified resource requirements, including spare parts and labor. This form of maintenance can be scheduled a long time in advance, allowing for a near-perfect forecast and supply of spare parts at actual maintenance execution. In contrast, non-routine maintenance is essentially of a stochastic nature. Nevertheless, accurate prediction of associated requirements such as spare parts demand is possible in the case of regular patterns in quantity and/or intervals. Forecasting accuracy decreases sharply when demand displays irregular patterns in terms of quantity and intervals, known as 'intermittent' and 'lumpy' demand [1]. Various efforts have been made to forecast lumpy demand (see Section 2) using time series techniques, but accuracy often lacks, especially when the span of time series data is limited.

Given the generally short time series spans associated with aircraft non-routine maintenance, a stochastic forecasting model is proposed to forecast spare parts demand. To properly position this contribution, a brief overview of the theoretical context is given, followed by introduction of the forecasting model in Section 3. In the Results

[1] Corresponding Author: W.J.C.Verhagen@tudelft.nl

section, forecasts are produced for a specific use case, where the forecasting model is applied to a sample of maintenance data for pooled LRU demand. This dataset has been scoped to include units having lumpy demand patterns, currently causing significant repair turn-around time delays for the maintenance operator that cooperated in this research. The consequence is that the forecast model generates monthly forecasts for peak demand related to LRUs. Forecasting performance is evaluated by considering the accuracy of the new model with respect to the existing company approach.

1. Theoretical Context

Ghobbar and Friend [1] and Syntetos [2] characterize demand types by considering variations in the frequency and quantity of demand. The associated metrics are the Average Inter Demand (ADI) and Coefficient of Variation (CV^2) – see Equations (1), (2) and (3), where t is time, n is the number of periods, ε_a is the average demand and ε_i is the actual demand.

$$ADI = \frac{\sum_{i=1}^{n} t_i}{n} \tag{1}$$

$$CV^2 = \left(\frac{\sqrt{\frac{\sum_{i=1}^{n} (\varepsilon_i - \varepsilon_a)^2}{n}}}{e_a} \right)^2 \tag{2}$$

$$\varepsilon_a = \frac{\sum_{i=1}^{n} \varepsilon_i}{n} \tag{3}$$

Throughout literature, ADI and CV^2 values are used to identify different demand patterns. Though differences in nomenclature exist, the following terms are frequently used [1]:

- **Smooth demand**: regular demand with a limited variation in quantity
- **Intermittent demand**: extremely sporadic demand, with no accentuated variability in the quantity of the single demand
- **Erratic demand**: large variation in quantity, but constant distribution over time
- **Lumpy demand**: great number of zero-demand periods and large variation in quantity.

These demand patterns (or types) are shown in Figure 1, together with suggested cut-off values by Ghobbar and Friend [1]. More involved classification criteria for stock keeping units are available [3], as well as further insights on insensitivity of cut-off values [4], but the categorization given below is used in the remainder of the research.

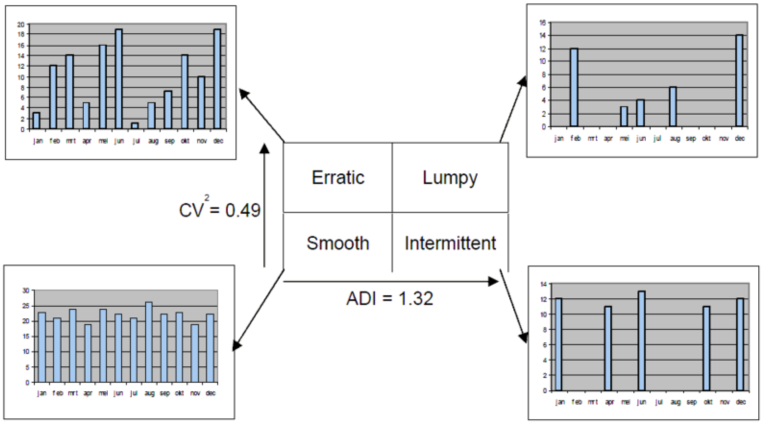

Figure 1. Demand classification [1]

Non-routine maintenance (and associated material demand) is characterized by many instances of zero-demand and significant variation in quantity. As such, the demand classification of non-routine material is typically intermittent or lumpy. Various studies have reviewed the development of endogenous forecasting methods for intermittent and lumpy demand [1, 5], typically in the form of time series techniques which consider historical patterns of demand to generate forecasts. Examples of time series techniques include Single Exponential Smoothing (SES), Croston's method (CR), and variants of Croston's method such as the Syntetos-Boylan Approximation (SBA) and the Teunter-Syntetos-Babai (TSB) variation. In recent years, Neural Networks (NN) have been applied to forecast lumpy demand, for instance by Gutierrez *et al.* [6] and Kourentzes [7]. However, these forecasting methods (both time series techniques and NN) generally require a substantial set of demand observations to enable proper initialisation, calibration or training of the applied methods. However, [8] note that non-routine aircraft maintenance is characterized by small time series datasets, as the number of demand observations associated with aircraft non-routine maintenance is typically small. As a direct result, prediction of lumpy spare part demand for aircraft non-routine maintenance frequently lacks accuracy.

2. Stochastic Modelling of Spare Parts Demand

Given the general inability of time series-based forecasting methods to accurately predict lumpy spare parts demand associated with non-routine aircraft maintenance, an alternative stochastic approach is proposed. The following general characteristics are assumed to apply for the arrival process of spare parts demand:

1) **Demand can be represented as a random variable:** demand can be represented by a function assigning a number to each sample point in sample S.
2) **The random variable describing demand is discrete:** a discrete random variable is 'a random variable X that can assume only a particular finite or countably infinite set of values' [9]. Demand occurrence is a variable that can only assume one of two possible values (i.e. a Boolean), namely demand occurs or does not occur at any given point of time. Demand volume can also be considered discrete in the current context, as the possible values are always integer and are part of a countable and finite set, being bounded by the total population of components in operation.
3) **Demand arrival is a continuous-time process:** demand can occur at any point in a given time period.

Given these characteristics, demand arrival can be modelled using a Poisson process, which is a stochastic point process on the basis of the Poisson distribution describing the number of occurrences of an event in a given time period. When the demand is random and constant and can therefore be described by an exponential time to failure distribution, the Homogenous Poisson Process (HPP) can be used. The latter condition is often considered to be the case when predicting spares demand for Line Replaceable Units (LRUs) in steady state conditions. The homogenous variant HPP is a counting process N(t) that satisfies the following three conditions [10]:

1) $N(0) = 0$
2) The process has stationary and independent increments
3) The number of demands in any interval of length t is Poisson distributed with mean demand λt, meaning that for all h, t ≥0,

$$P[N(t+h) - N(h) = n] = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, n = 0,1,2 \ldots \tag{4}$$

where $N(t+h) - N(h) = n$ is the number of events in time interval [*t, t*+*h*].

As mentioned, using the HPP to estimate demand arrival assumes that the arrival has a constant distribution per time unit analyzed. Alternatively, a Non-Homogenous Poisson Process (NHPP) or variant thereof can be used to predict demand arrival in the case that the arrival rate λ is a function of time t, i.e. the arrival rate λ is not constant. In this case, condition 2) is adjusted to indicate that the process has independent increments, but not stationary.

Estimating unit demand for a given time period requires the following steps:

1) **Determination of time period *t*:** the length of the time period *t*, representing the increment or interval of the process, has a direct effect on the expected number of demand arrivals (see also Equation (5) for the HPP process). However, it has a more indirect and subtle effect in that the determination of a time period may influence the determination of the arrival rate *λ(t)*, given that data for a time period may exhibit seasonality or a trend. Shifting external conditions during a time period *t* may also influence the counting process. As

such, data that is subjected to Poisson process modelling and analysis must be checked for seasonality and trends, and must be used at the right level of granularity for forecasting use.

2) **Determination of arrival rate *λ(t)*:** if the arrival rate is a function of time, the generalized rate function is given as *λ(t)*. In that case, the expected number of events between time t_1 and time t_2 is given by $\lambda_{t1,t2} = \int_{t1}^{t2} \lambda(t)dt$. This relation can be used to determine the arrival rate *λ(t)* for a given period of time, where *λ(t)* itself may be deterministic or stochastic, and where the latter variant is described by a Cox process.

The arrival rate *λ(t)* can be used to generate 'flat' forecasts of demand for a given time period in the case of HPP, when *λ(t)* is constant with respect to time. The expected number of demands during the duration of length *t* is then given by:

$$E[N(t)] = \lambda t \quad (5)$$

For a non-homogenous Poisson process (NHPP), demand arrival will be a function of time.

3. Results

To test the suitability of the HPP model for predicting lumpy spare parts demand for non-routine aircraft maintenance, a dataset from a maintenance operator has been collected and analyzed. This dataset concerns Line Replaceable Units (LRUs) that are maintained for a customer *pool*, which requires meeting specified service levels. This necessitates timely performance above all, which in practice requires immediate replacement of failed LRUs, which is followed by repair of LRUs at the maintenance operator premises. The repaired LRUs can be reintroduced into the stock at a later stage. The most important implication of this is that from a strategic point of view, the maintenance operator requires an ability to meet *peak demand* at any time, rather than average demand. Consequently, the forecasting approach should take this into account.

In this section, the scope and characteristics of the dataset are discussed, followed by application of the HPP model to generate forecasts for peak demand. The assumption that HPP is the appropriate and most suitable probabilistic distribution will be checked by evaluating a range of probability density functions. The accuracy of these stochastic distributions are checked using a goodness-of-fit test based on the root mean square deviation (RMSD). Finally, the approach is validated by comparing performance to the current method used by the maintenance operator.

3.1. Dataset characteristics

For the use case, a dataset consisting of 58202 repairs (assumed equivalent to spare part demand, given the spare parts pool context) covering a time span of 730 days (January 2012- December 2013) has been assembled. This dataset has subsequently been reduced in two ways to uncover the items of interest, i.e. LRUs that cause a significant cost impact due to turn-around time (TAT) delays and that exhibit lumpy or intermittent demand.

The first step in scoping the dataset has been to uncover the repair occurrences which are associated with a TAT delay. Of the total number of repairs, the number of TAT delays is given in Table 1. This number is subsequently reduced to incorporate TAT delays that were caused by a shortage of parts (categorized as TA by the maintenance operator). In total, about 8% of the units exceeded TAT because of a shortage of parts.

Table 1. Component repair occurrences

Categorization	Occurrences	Percentage of total (%)
Total repairs	58202	0
Total TAT	20816	36
TA	4727	8

Next, from this reduced dataset a top nine rank has been established. Ranking has been achieved through calculation of an impact figure, which consists of a multiplication of the amount of occurrences that the unit exceeded the TAT over an specific amount of time with the price value of the unit (at latest list price). The top 9 units are shown in Table 2.

Table 2. Selected component impact ranking

Rank	Unit code number	# TAT delays	Impact
1	332000	44	7,711,132
2	471126	24	4,679,280
3	495337	18	2,941,578
4	433016	16	2,631,040
5	481090	15	2,006,400
6	497540	7	1,116,899
7	854140	19	993,567
8	495285	17	733,057
9	495680	17	733,057

Subsequently, the demand characteristics of these top 9 units have been evaluated for the time span of the dataset. Demand characteristics have been analyzed on multiple levels of granularity (days, months, half-yearly semesters). The characteristics on a semester basis are given in Table 3. 55.6% of the units showed intermittent demand behavior, whereas the rest (44.4%) showed lumpy behavior. None of the units showed smooth (0%) or erratic (0%) demand characteristics. Both by extrapolation and by analysis, demand on a daily or monthly basis likewise shows intermittent or lumpy behavior.

Table 3. Unit demand characteristics

	Demand type			
	Smooth	**Erratic**	**Intermittent**	**Lumpy**
Semester				
Year 1, Semester 1	0	0	5	4
Year 1, Semester 2	0	0	4	5
Year 2, Semester 1	0	0	5	4
Year 2, Semester 2	0	0	6	3
Percentage total	0.0	0.0	55.6	44.4

3.2. Application of stochastic forecasting model

Data for the selected top 9 of units has been used to generate demand forecasts (see also Section 4.3). This consisted of two elements:

1) **Determination of time period t:** the maintenance operator uses forecasts on a monthly basis and aims to stock towards peak demand during that period. To minimize organizational impact transition when moving from one forecast method to another, the time period t is determined to concern monthly intervals.
2) **Determination of arrival rate λ(t):** The main approach has been the HPP process (implying determination of a constant arrival rate), but for completeness a number of other probability density functions have been analysed in order to validate whether the Poisson distribution shows the best forecasting accuracy (see also Section 3.3). The program Easyfit has been used to determine the coefficients of the probability density functions shown in Table 4, based on the historical demand data of the top 9 units and a first selection of density functions based on statistical goodness-of-fit. The data in Table 4 are average coefficients, but these can show significant variation from the given values when considering monthly intervals. Due to data confidentiality, it has been chosen to represent the average values instead of a monthly example.

Table 4. Probability density function coefficients for top-9 LRUs (averaged over total time period of 2 years)

Unit	General Extreme Value (k,σ,μ)	Weibull (α,β,γ)	Poisson (λ)	Normal (σ,μ)
332000	-0.28378	3.2778	5	1.0954
	1.0594	5.4111		5
	4.6269	0		
433016	0.22372	1.145	4.8333	3.3714
	2.2006	4.7785		4.8333
	2.9451	0		
471126	0.5148	4.9606	9.6667	4.2269
	1.4029	8.6842		9.6667
	7.4165			
481090	0.16246	1.2977	5.1667	3.4881
	2.4644	5.0773		5.1667
	3.2768	0		
495285	-0.08673	1.0605	5	2.9665
	2.6928	5.3671		5
	3.6599	0		
495337	-0.28378	3.1876	8	2.2804
	2.4719	8.3017		8
	7.1295			
495680	-0.1418	2.6246	5.1667	3.5449
	3.2913	6.9996		5.1667
	3.6765	0		
497540	-0.3541	3.5981	9.3333	2.8752
	3.166	9.191		9.3333
	8.3565	1.0732		
854140	0.547	0.79972	3.1667	3.9707
	1.234	1.906		3.1667
	1.0122	0		

3.3. Validation of forecasts

Using the determined coefficients (on a monthly basis), forecasts have been generated per month. Again, an average representation is given (see Table 5) which can be compared to the average actual monthly removals.

Table 5: Forecasts versus actual monthly average demand

Unit	GEV	Weibull	Poisson	Normal	Actual monthly average
332000	8.06	8.64	8.00	8.22	7.83
433016	16.24	12.37	9.00	11.16	5.00
471126	18.73	16.07	15.00	17.06	6.00
481090	24.64	15.56	11.00	14.72	6.83
495285	12.22	12.21	12.00	13.02	5.42
495337	19.21	18.50	17.00	17.09	9.58
495680	12.43	12.20	12.00	13.61	6.58
497540	13.73	24.32	11.00	14.06	6.00
854140	64.23	51.63	15.00	33.20	13.08

It is apparent that the Poisson process is closest to the actual values. However, this averaged representation foregoes the variation on a monthly basis. To check the accuracy of the various methods, the Root Mean Square Deviation (RMSD) as given in equation (6) has been used for various time window resolutions (i.e. intervals of t). The results are given in Table 6. It can be seen that the Poisson distribution consistently performs best. This compares favorably to the method currently used by the company in question, which uses the normal distribution as a basis for generating forecasts.

$$RMSD = \sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n}} \tag{6}$$

Table 6. RMSD of selected probability function forecasts versus actual demand for different time windows

Time window	GEV	Weibull	Poisson	Normal
1	55	43	29	41
2	34	31	24	26
3	76	65	33	48
4	39	31	21	26
5	48	37	25	35

For more in-depth validation, units can be analyzed separately. An example is given in Figure 2.

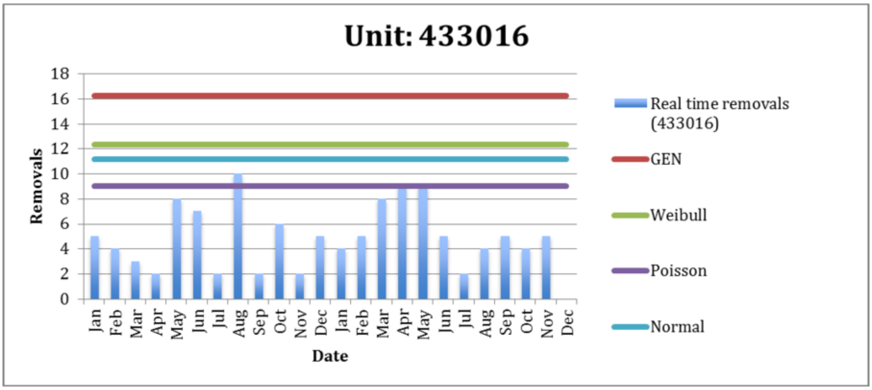

Figure 2. Unit forecasts versus monthly removals

This example concerns unit number 433016. As can be observed, the HPP model has been evaluated on basis of the average arrival rate over the dataset timespan, as the monthly estimates are considered confidential. As such, a constant forecast is generated for the full timespan. It can be seen that the HPP model is closest to the actual monthly demand values, though the forecasts on monthly basis are in reality more accurate. What is striking is that the HPP model is able to forecast peak demand for all instances except one (August 2012). The same is true when considering forecast based on monthly arrival rates, though these results cannot be represented here.

4. Conclusions & Future Research

A stochastic method to generate demand forecasts for aircraft non-routine demand has been presented. Demand arrival is assumed to be modelled most accurately by a Poisson Process. It has been shown that application of the Homogeneous Poisson Process (HPP) outperforms other probability distributions.

This work represents an initial step in using stochastic point processes to forecast spare part demand. Future extensions are considered, including 1) comparison of accuracy with time series techniques as discussed in [8]; 2) attenuation of the HPP coefficients using explanatory variables; 3) incorporation of time-dependent arrival rates (using a NHPP variant such as the Cox Process).

References

[1] A.A. Ghobbar and C.H. Friend, Evaluation of forecasting methods for intermittent parts demand in the field of aviation: a predictive model, *Computers & Operations Research* **30** (2003), 2097-2114.

[2] A.A. Syntetos, A note on managing lumpy demand for aircraft spare parts, *Journal of Air Transport Management* **13** (2007), 166-167.

[3] A. Bacchetti, et al., Empirically-driven hierarchical classification of stock keeping units, *International Journal of Production Economics* **143** (2013), 263-274.

[4] J.E. Boylan, A. Syntetos, and G.C. Karakostas, Classification for forecasting and stock control: a case study, *Journal of the Operational Research Society* **59** (2008), 473-481.

[5] R.H. Teunter and L. Duncan, Forecasting intermittent demand: a comparative study, *Journal of the Operational Research Society* **60** (2009), 321-329.

[6] R.S. Gutierrez, A.O. Solis, and S. Mukhopadhyay, Lumpy demand forecasting using neural networks, *International Journal of Production Economics* **111** (2008), 409-420.

[7] N. Kourentzes, Intermittent demand forecasts with neural networks, *International Journal of Production Economics* **143** (2013), 198-206.

[8] M. Zorgdrager, W.J.C. Verhagen, and R. Curran, An Evaluation of Forecasting Methods for Aircraft Non-Routine Maintenance Material Demand, *International Journal of Agile Systems and Management* **7** (in press).

[9] J. Knezevic, *Reliability, Maintainability and Supportability: A Probabilistic Approach*, McGraw-Hill, London, 1993.

[10] U.D. Kumar, et al., *Reliability, Maintainability and Logistic Support - A Lifecycle Approach*, Kluwer Academic Publishers, Boston, 2000.