Sortie-based aircraft component demand rate to predict requirements

Thomas R. O’Neal, John M. Dickens, Lance E. Champagne, Aaron V. Glassburner, Jason R. Anderson and Timothy W. Breitbach

Department of Operational Sciences, Graduate School of Engineering and Management, Air Force Institute of Technology, Wright-Patterson AFB, Ohio, USA

Abstract

Purpose – Forecasting techniques improve supply chain resilience by ensuring that the correct parts are available when required. In addition, accurate forecasts conserve precious resources and money by avoiding new start contracts to produce unforeseen part requests, reducing labor intensive cannibalization actions and ensuring consistent transportation modality streams where changes incur cost. This study explores the effectiveness of the United States Air Force’s current flying hour-based demand forecast by comparing it with a sortie-based demand forecast to predict future spare part needs.

Design/methodology/approach – This study employs a correlation analysis to show that demand for reparable parts on certain aircraft has a stronger correlation to the number of sorties flown than the number of flying hours. The effect of using the number of sorties flown instead of flying hours is analyzed by employing sorties in the United States Air Force (USAF)’s current reparable parts forecasting model. A comparative analysis on D200 forecasting error is conducted across F-16 and B-52 fleets.

Findings – This study finds that the USAF could improve its reparable parts forecast, and subsequently part availability, by employing a sortie-based demand rate for particular aircraft such as the F-16. Additionally, our findings indicate that forecasts for reparable parts on aircraft with low sortie count flying profiles, such as the B-52 fleet, perform better modeling demand as a function of flying hours. Thus, evidence is provided that the Air Force should employ multiple forecasting techniques across its possessed, organically supported aircraft fleets. The improvement of the forecast and subsequent decrease in forecast error will be presented in the Results and Discussion section.

Research limitations/implications – This study is limited by the data-collection environment, which is only reported on an annual basis and is limited to 14 years of historical data. Furthermore, some observations were not included because significant data entry errors resulted in unusable observations.

Originality/value – There are few studies addressing the time measure of USAF reparable component failures. To the best of the authors’ knowledge, there are no studies that analyze spare component demand as a function of sortie numbers and compare the results of forecasts made on a sortie-based demand signal to the current flying hour-based approach to spare parts forecasting. The sortie-based forecast is a novel methodology and is shown to outperform the current flying hour-based method for some aircraft fleets.

Keywords Demand, Forecast, D200, United States air force, Inventory, Spare parts

Paper type Research paper

Introduction

Predicting future needs for aircraft spare parts is a critical issue within the United States Air Force (USAF). In the USAF’s complex multi-echelon, multi-indenture supply repair cycle, an inaccurate demand forecast may result in improper work schedules at the repair depots, incorrect operating stock levels at base supply warehouses and incorrect stock levels in aircraft deployment readiness kits. The consequences of such inaccuracy include a spare part
or multiple spare parts not being available for an aircraft that the USAF needs to fly. The impact of an unavailable aircraft could result in missed training opportunity for a pilot or, more severely, degradation of core missions such as personnel recovery or air superiority. If demand for any given component is overestimated, too many spare parts are stocked, and other needed parts are not purchased or repaired due to sustainment budget constraints. Consequently, demand forecasting is a critical component to supply chain resilience (Fu and Chien, 2019). Accurate aircraft spare part forecasts empower the Air Force with the robustness and agility (Wieland and Wallenburg, 2013) required to meet its core competency in support of national defense.

The current USAF aircraft spare parts planning process employs a combination of reliability theory and forecasting techniques to determine future demand. Critics of USAF forecasting methods suggest that the USAF’s methodologies should be updated because it continues to underperform (Eckbreth et al., 2011). This study is a parsimonious effort to improve forecasts and strays from the recommendation of Eckbreth et al. (2011) to employ “more sophisticated data analysis.” For most spare parts, the USAF currently calculates reliability on the number of yearly flying hours associated with the aircraft platform. The following section further clarifies this problem and the cost of inaccurate demand predictions.

Sustainment costs for USAF weapon systems, especially legacy systems, are untenable. Eckbreth et al. (2011) tied the USAF challenges of future weapon system sustainment to the service’s inability to accurately predict part needs. In 2011, Eckbreth et al. (2011) claim the demand forecast for spare parts was only 19% accurate. Furthermore, the USAF expenditures to operate and maintain the active fleet ballooned to $63.7 billion in 2019 dollars. Part of the growth in sustainment expenses were due to the age of the fleet (Gunzinger et al., 2019). To illustrate, Figure 1 shows the upward trend of operation and maintenance costs per aircraft across all of its fleets, i.e. fighter, bomber and cargo aircraft.

The 35 years prior to 1997, the USAF funding for operations and maintenance increased by $3.4 million per aircraft. However, in the 20 years since, this portion of funding increased by $5.1 million per aircraft. Figure 2 highlights the downward trend in new aircraft procurement spending.

Figures 1 and 2 indicate that the cost of sustaining and operating an aging fleet is increasing and cause for concern. Congressional Budget Office (2018) found that since 2001 operations and sustainment cost growth rates have exceeded 5% above inflation.

![Figure 1. Trends in O&M funding per aircraft](source(s): Gunzinger et al. (2019))
Considering prior to 2001 these growth rates were between 1 and 2%, the office finds it alarming that the cost of maintaining and operating an aging fleet remain to grow at a faster rate. With plans to fly legacy systems like the B-52 until 2050 and increasing part sustainment costs, it is imperative that USAF supply chain planners continue to adapt and find more accurate calculations for future spare part needs beyond the inadequacies of legacy forecasting methods currently employed.

The USAF employs a centrally planned flying hour program to determine future yearly flying requirements and associated spare parts consumption rates. Ironically, in most cases the rate at which aircraft spare parts fail and correspondingly place a demand on the supply system does not show a strong correlation with actual hours flown. With a forecasting program that heavily relies on flying hours, this weakness subsequently translates to a demand rate calculation that produces an exceedingly inaccurate future year forecast. To address this problem and how it is claimed to affect sustainment costs, this study aims to address the main research question: (1) Can sortie data be employed to reduce USAF forecast error? To appropriately address this research question, this study employs a correlation analysis to illustrate that demand for reparable aircraft parts has a stronger relationship to the number of sorties flown rather than the number of flying hours. Then the effects of using the number of sorties flown instead of flying hours are analyzed by inputting the number of sorties in the USAF’s current forecasting model. To appropriately assess the complexity of the USAF, both the F-16 and the B-52 fleets were chosen to analyze the difference a sortie-based demand rate will have fleets of differing size and flying profiles. For both fleets, errors from the current system and the proposed model are compared to determine which produces the least amount of error.

Related studies
The purpose of this section is to explore the vast knowledge of forecasting as it relates to our study. It starts with the structure of forecasting. Then, the literature review follows the lineage of time series forecasting techniques as they have grown in complexity and accuracy.
Then, the literature review differentiates times series forecasting from the techniques that count data like inventory demand require.

**Forecasting**

The review of the literature pertaining to this study begins with the concept of predicting future outcomes or forecasting. Forecasting techniques can be placed into two main categories: quantitative and qualitative. Quantitative forecasting can be employed when there is enough empirical information regarding the past, and it can be assumed that the past demand will continue into the future. If these conditions are not met, qualitative techniques can be employed. If neither condition is met, the topic of interest is unpredictable. Moreover, quantitative forecasting methods can be characterized as a continuum with two extremes. At one extreme, there are intuitive or ad hoc methods, and at the other extreme there are formal statistical methods. Another dimension for classifying quantitative forecasting methods distinguishes this method by the model used. The two main forecasting models are time series and explanatory models. Explanatory models, such as linear regression, assume that the variable to be predicted has some relationship with one or more independent variables. However, time series models make no effort to explain the factors that may affect the variable that is being predicted. Time series models attempt to find a pattern in the historical data and generalize that pattern into the future (Makridakis et al., 1998).

**Time series forecasting**

Bowerman et al. (2005) define time series as a chronological sequence of observations on a particular variable that is quantifiable over some time measure. The authors explain that the components of a time series are trend, cycle, seasonal variation and irregular fluctuation. The authors argue that due to the irregular fluctuation, no single best forecasting model exists. So, the biggest problem with forecasting is fitting an appropriate model to the pattern in the available time series data. The fluctuations are modeled as part of the error in forecasting. So, according to the authors, large forecasting errors can indicate that the irregularity is too great for forecasting, or another model or technique could be more appropriate. Next, it is important to review the different time series forecasting methods.

Beyond using averages or naïve one month moving averages to predict future occurrences, researchers like Wiener (1949) began using the statistical concepts from communication engineering and cybernetics to make predictions based on the smoothing of stationary time series. Much like early communication devices depended on probability distributions to predict the most likely intended message and provide it to the receiver, Wiener proposed that time series data behave this way and can be used to make predictions. Brown (1959) based his work on much of Wiener’s ideas by using statistical forecasting for inventory control. His work was an early application of smoothing and other advanced tools like Monte-Carlo simulation to advanced demand predictions. It may have been unknown to Brown, but Holt (2004) documented the idea of smoothing variation or random fluctuations and also derived equations to model trend and seasonal fluctuations. Brown’s work in 1959 attempted to make the abstract concept and the mathematics more user friendly for an inventory control specialist or manager. Winters’ (1960) work added to the time series forecasting body of knowledge by comparing weighted exponential smoothing to traditional methods of the time to show that it can model trend and seasonality if present and provide a more accurate forecast. Additive and multiplicative forms of exponential smoothing were theorized in much of the early works. However, Pegels (1969) formally
presented the nine possible models in graphical form and summarizes them into one formula that readers can comprehend.

Before Pegels’ work, Muth (1960) was the first to apply statistical concepts like linear regression to time series and showed that this method of simple exponential smoothing (SES) provided an optimal forecast for what he called a “random walk with noise.” Later, Box and Jenkins (1970) examined time series that are nonstationary. Nonstationary time series are exceedingly difficult to forecast using moving average methods because the data have multiple windows of time that have different means. However, nonstationary time series does display homogeneity in the sense that at least one part of the series behaves much like others. To model this behavior, the authors proposed a technique called autoregressive integrated moving average (ARIMA) by summing the stationary processes by the number of differences in the time series. With the expansive work and the multiple approaches to time series forecasting, Box and Jenkins work provided a clear and robust method for time series identification, parameter estimation and verification known as the Box–Jenkins approach.

With research, scholars articulated other methods like state-space models (Ord et al., 1995). These advanced methods are outside the scope of this study. This section of the literature focused on time series with linear relationships between the variable and time. Time series can also show a nonlinear relationship between time and the variable of interest, which are much more complex than the aforementioned linear methods. Furthermore, they are difficult to perform and are outside the scope of this study.

Method
Sherbrooke (1997) noted that F-16 sortie data were preferable to flying hours because demand was more related to sorties than flying hours. However, current demand forecasts are made via D200 using forecast flying hours to predict demand for parts. The USAF primarily uses an eight-quarter moving average factor method to calculate the next years demand for each spare parts (Defrank, 2017). Using the following equation, the USAF begins by calculating the average demand per flying hour:

\[
F_j = \frac{\sum_{i=T-8}^{T} \text{Demand of } j\text{th part in the } i\text{th quarter}}{\sum_{i=T-8}^{T} \text{Flying hours in the } i\text{th quarter}}
\]

(1)

where \(F_j\) is average demand of part \(j\), and \(T\) is the most recent three-month period (Defrank, 2017). The D200 A multiplies this average number of demands per actual flying hour by the USAF’s flying hour forecast for the next time period(s) to forecast the future demand. This research proposes reformulating equation (1) by modeling the denominator as a function of sorties versus the flying hours in the denominator of equation (1). The resulting demand factor is determined by equation (2) below:

\[
F_j = \frac{\sum_{i=T-8}^{T} \text{Demand of } j\text{th part in the } i\text{th quarter}}{\sum_{i=T-8}^{T} \text{Sorties flown by the } j\text{th part in the } i\text{th quarter}}
\]

(2)

where \(F_j\) is average demand of part \(j\), and \(T\) is the most recent three-month period. In calculating the annual demand, the factor computed via equation (2) is multiplied by the forecasted sorties for the airframe. Currently, the USAF does not forecast sorties, so this research employed the Holt–Winters triple exponential smoothing to forecast future sorties.

Data
This research explored two data sets to validate the hypothesized forecasting method, F-16 and B-52 spares. The two data sets were selected to highlight performance of the proposed
method in relation to airframes with differing flying hour to sortie ratios. That is, fighter aircraft tend to fly sorties of shorter duration than bomber aircraft, so the average flying hours per sortie will be larger for bomber fleets than for fighter fleets. Additionally, the F-16 fleet is significantly larger than the B-52 fleet. As a result, forecast accuracy may be influenced by fleet size.

The Air Force Sustainment Center (AFSC) provided two initial sets of data from the USAF D200 A system. The first data set contained flying hour and demand data for 539 repairable items on the F-16 weapon system from fiscal year 2004 through fiscal year 2018. The second data set consisted of 561 repairable items with flying hour and demand data on the B-52 weapon system over the same time frame.

The number of sorties was taken from the USAF’s Logistics Installation Mission Support-Enterprise View (LIMS-EV) for both fleets. LIMS-EV brings a large amount of logistics data to include supply information from across at least 60 different data repositories to provide one true picture of all USAF resources. Finally, the AFSC provided the forecast that the D200 produced for each item from fiscal year 2011–2018. The D200 demand forecast for each aircraft’s items in those years allows this study to compare the proposed model’s predictions with the D200’s predictions.

**Accuracy metric**
This study uses the mean absolute percent error (MAPE) to measure forecast error. The MAPE of the D200 demand forecast and the MAPE of the proposed demand forecast are calculated and compared to determine the method that will provide the USAF with the best estimate of demand. The MAPE is calculated with the following equation:

\[
\text{MAPE} = \left( \frac{1}{n} \sum_{j=1}^{n} \frac{|D_j - F_j|}{|D_j|} \right)
\]

where \(n\) is the number of spare parts in the fleet inventory, \(D_j\) is the actual demand of part \(j\), and \(F_j\) is the forecast demand of part \(j\). A lower MAPE indicates a more accurate forecast of demand.

**Correlation of fleet demand to flying hours and sorties**
Following the data collection, a correlation analysis was performed to compare each item’s demand with both flying hours and number of sorties. The Pearson’s \(\rho\) value was calculated for each item via equation (4):

\[
\rho_{x,y} = \frac{\sum (x_j - \bar{x}_j)(y - \bar{y})}{\sqrt{\sum (x_j - \bar{x}_j)^2} \sqrt{\sum (y - \bar{y})^2}}
\]

where \(x_j\) is the forecast demand for part \(j\), \(\bar{x}_j\) is the sample mean of the actual demand for part \(j\) from 2004 to 2018, and \(\bar{y}\) is the sample mean of the aircraft’s actual flying hours or number of sorties, respectively, from the same quarter. Correlation values for each part range between \(-1\) and \(1\), where correlation values between 0 and 1 indicate positive correlation and values between 0 and \(-1\) indicate negative correlation.

**Forecasting sorties**
The USAF does not provide the logistics community a forecast or plan for number of sorties. Therefore, this research forecasts future sorties via Holt–Winter’s triple exponential
smoothing. The method accounts for seasonality and trends in the demand. The forecasted number of sorties is multiplied by the new sortie-based factor from equation (2) to produce a demand forecast.

Results and discussion

The results of the correlation analysis and the demand forecast are designed to give decision makers insight for when to use a sortie-based demand forecast rather than a flying hour-based forecast. Analysis of correlations within the fleets showed mixed results. Comparison resulted in 200 of 539 parts (37.1%) for the F-16 fleet and 295 of 561 parts (52.6%) for the B-52 fleet with demand with stronger correlation to number of sorties than to flying hours. The remaining items, 339 of 539 and 266 of 561 parts, respectively, have annual demand that is more strongly correlated to flying hours. While our correlation analysis does not identify the most effective forecasting technique, the results suggest that there is value in exploring the link between sorties and demand more closely. If nothing else, this underscores the complexity in forecasting part demands for military aircraft.

The number of parts used in the forecasts for each airframe differs from the correlation analysis as not all parts reflected a full 14 years of data. Those parts, while useful for the correlation analysis, were subsequently removed from the forecasts. Table 1 summarizes the MAPE resulting from the forecasts for part demands for the F-16 and B-52 fleets. In the case of the F-16 fleet, flying-hour based forecasts within D200 produced a MAPE of 52.5%, and the proposed sortie-based method produced a MAPE of 39.9%. This represents an overall decrease of 12.6% in the MAPE or a 24% reduction in error. In contrast, the D200 flying-hour-based forecast for the B-52 produced a MAPE of 40.9% while the proposed sortie-based forecast produced a demand MAPE of 44.5%. This an overall increase in error of 3.6% or an 8.1% increase in error. To provide clarity, we believe that the flying profile of the F-16 and aircraft whose missions are of short duration lend themselves to a forecast based on sorties counts in comparison to aircraft flying fewer sorties but for longer hours (e.g. B-52, C-5, etc.). This underscores the complexity of predicting spare part demands across multiple and diverse (e.g. age, mission type, flying duration variability, extreme operating conditions etc.) airframes. Consequently, we believe the Air Force “one-size fits all” approach to predicting spares requirements requires further investigation and perhaps should be tailored to each airframe. More clearly, forecasts for some airframes should be sortie-based while others should be flying hours based and still others a combination of both techniques or some other higher performing technique not considered.

The analysis of this study supports extant research, specifically Sherbrooke’s (1997), that demonstrates a lack of statistical significance between flying hours (e.g. sortie length) and demand for various aircraft types. The calculated error for F-16 parts was smaller for sorties-based demand forecast counts, while B-52 parts forecasts were improved under a flying hour-based forecasting model. Our analysis backs Sherbrooke’s lack of support to the Air Force’s assumption that demand linearly increases with flying hours. The resulting correlation and forecasting error analyses of this study supports further investigation into adjusting D200’s parts forecasting algorithm to a sorties-based forecast for some aircraft types.

<table>
<thead>
<tr>
<th></th>
<th>F-16</th>
<th>B-52</th>
</tr>
</thead>
<tbody>
<tr>
<td>D200 (Flying hour-based) forecast MAPE</td>
<td>52.5%</td>
<td>40.9%</td>
</tr>
<tr>
<td>Sortie-based forecast MAPE</td>
<td>39.9%</td>
<td>44.5%</td>
</tr>
<tr>
<td>Number of parts (n)</td>
<td>418</td>
<td>555</td>
</tr>
</tbody>
</table>

Table 1. 2018 forecast error comparison (D200 versus sortie-based)
Conclusions and recommendations
This study showed that aircraft spare part demand is not always strongly correlated to the number of hours that are flown. In fact, 37.1% of the F-16 items from 2004 to 2018 had demand that was more correlated to the number of sorties flown. Due to the historic error of USAF forecasts and the finding that demand for many F-16 parts have a relationship with the number of sorties flown, this study adjusted the USAF’s historic demand rate forecast system to use sorties as a measure of time or demand interval. This demand rate is applied to the aircrafts predicted usage to calculate demand for the next periods (Berger and Murphy, 2014). However, generally the USAF predicts usage in terms of flying hours not sorties. B-52 spare part demand data were obtained to explore the robustness of potential findings.

The results of this study empirically show that there is a possibility for decreased error in USAF spare part forecasting. Although forecasting count data that are intermittent and nonstationary like spare part demand is difficult, this study employs a parsimonious method to decrease error and improve spare parts forecasting. This study gives evidence for USAF supply chain analysts and program managers to consider in comparing their current flying hour-based forecasts to a sortie-based approach. A quick comparison could result in better buying decisions and budget proposals for spare parts. With weapon system sustainment costs growing at an alarming rate, better decisions based on this study could decrease funds being erroneously allocated among spare part requirements.

Future research
The simplicity of this study lends itself to a number of areas for future research. First, future studies or improvements to the model proposed in this study should apply quantity per application (QPA) and percent application to the number of sorties attributed to each item. Applying these factors to the number of sorties USAF aircraft fly allows for a more precise allocation of sorties to each item installed on the aircraft. Future research can employ the model of this study with the more precise sortie allocation and could improve the forecast.

Furthermore, future research should explore the appropriateness of applying the Poisson process to every item. The literature suggests that some aircraft parts may have a failure distribution that differs from the Poisson process (Ebeling, 2004). Future research should investigate the failure distribution of a sample of parts. If the distributions are significantly different, parameter estimates can be explored and possibly implemented into this study’s model. If a future study of this nature could show decreased forecast error further, it could help drive sustainment costs down.

Finally, research should be done to identify possible trends regarding the time measure that predicts demand more accurately. Research should be focused on finding the most appropriate variable for each item, class of items or repair cycle group. Having greater confidence in the time measure selected, forecasters can more effectively employ a possible hybrid system.

References


**Further reading**


O'Malley, T.J. and Bachman, T. (1990), *Depot Maintenance in the Air Force: How Requirements Are Determined and How They Relate to Aircraft Readiness and Sustainability*, Report No. AF901R1, Logistics Management Institute, Bathesda, MD.


**Corresponding author**

Aaron V. Glassburner can be contacted at: aaron.glassburner@afit.edu