NLR-TP-2021-023 | October 2021

Maintenance Add-on to Logistics

Data fusion for a reduced logistics footprint

CUSTOMER: TKI Dinalog

Royal NLR – Netherlands Aerospace Centre

Dedicated to innovation in aerospace

EXECUTIVE SUMMARY

UNCLASSIFIED

Maintenance Add-on to Logistics

Data fusion for a reduced logistics footprint

REPORT NUMBER
NLR-TP-2021-023

AUTHOR(S)
D. Spirtovic
K. Oosterwaal
J.W.E. Wiegman
R. Brink
J. Donders
H. Corstens

REPORT CLASSIFICATION
UNCLASSIFIED

DATE
October 2021

KNOWLEDGE AREA(S)
Health Monitoring and Maintenance of Aircraft

DESCRIPTOR(S)
Artificial intelligence
Logistics
Maintenance
Operational Research
Prognostics & Health Management

Project partners

Gordian Logistic Experts

ILIAS smart. reliable. ready.

OneLogistics

Koninklijke Luchtmacht

Problem area

Aviation companies use prognostics for logistics and maintenance, but data fusion between these domains is rare. Supply chains have short planning horizons and vendors store extra spare parts at strategic locations for security of delivery. However, this approach results in additional cost and large logistics footprints. Accurate and timely predictions of failures and failure modes enable just in time spare parts delivery and efficient parts repair.

Description of work

The project Maintenance Add-on to Logistics (MATLOG) combines predictive maintenance and predictive logistics to reduce cost and logistics footprints. Diagnostics models and prognostics models for maintenance, as well as multi-echelon models for logistics are developed as add-ons to a logistics management system. These items are combined to prove the concept of enriching logistics planning with information about forthcoming component failures to increase the planning horizon and reduce the troubleshooting workload.

Results and conclusions

Component histories, flight data and maintenance records are combined to enrich data. Two line replaceable components are selected for case studies. Failures are predicted through binary classification. Supervised and unsupervised learning approaches are evaluated to obtain diagnostics models. Multi-echelon models are developed for network stock optimisation. All models are used as plug-ins to a logistics management system.

Prognostics models provide triggers for components that are likely to fail within 3 days with a confidence level above a preset threshold. Multi-echelon models calculate significantly reduced network stock and repair lead time with 3 days warning time, while maintaining an aggregate service level of 95% and taking up to 10% false negatives into account.

Conclusions are that: 1) material condition data is required for accurate diagnostics and prognostics; 2) accuracy and timeliness of failure predictions are key drivers of network stock; and 3) data-driven failure diagnostics, combined with multi-echelon models and a logistics management system, reduce lead times and cost of repairs.

Applicability

The concept of data fusion between logistics and maintenance is applicable to asset management of aircraft, infrastructure, installations, vehicles and vessels. A constraint to the concept is the presence of a geographically distributed supply chain for repairables.

GENERAL NOTE

This report is based on a presentation held at the EURO 2021 conference, Athens, 11-14 July 2021.

This publication has been refereed by the “Advisory Committee Aerospace Systems”.

Royal NLR
Anthony Fokkerweg 2
1059 CM Amsterdam, The Netherlands
p) +31 88 511 3113
e) info@nlr.nl i) www.nlr.nl

Dedicated to innovation in aerospace

NLR-TP-2021-023 | October 2021

Maintenance Add-on to Logistics

Data fusion for a reduced logistics footprint

CUSTOMER: TKI Dinalog

AUTHOR(S):

D. Spirtovic	NLR
K. Oosterwaal	Gordian Logistic Experts
J.W.E. Wiegman	NLR
R. Brink	NLR
J. Donders	Gordian Logistic Experts
H. Corstens	Ilias Solutions

Royal NLR – Netherlands Aerospace Centre

This report is based on a presentation held at the EURO 2021 conference, Athens, 11-14 July 2021.

The contents of this report may be cited on condition that full credit is given to NLR and the author(s).
This publication has been refereed by the "Advisory Committee Aerospace Systems".

CUSTOMER	**TKI Dinalog**
CONTRACT NUMBER	**2018-1-148TKI**
OWNER	**NLR + partner(s)**
DIVISION NLR	**Aerospace Vehicles**
DISTRIBUTION	**Unlimited**
CLASSIFICATION OF TITLE	**UNCLASSIFIED**

APPROVED BY:		
AUTHOR	**REVIEWER**	**MANAGING DEPARTMENT**
D. Spirtovic	Arjan de Jong 2021.10.08 15:36:21 +02'00'	Marcel J. Bos 2021.10.10 12:58:40 +02'00'
DATE 0 7 1 0 2 1	DATE 0 8 1 0 2 1	DATE

Abstract

Aviation companies use prognostics for logistics and maintenance, but data fusion between these domains is rare. Supply chains have short planning horizons and vendors store extra spare parts at strategic locations for security of delivery. However, this approach results in additional cost and large logistics footprints. Accurate and timely predictions of failures and failure modes enable just in time spare parts delivery and efficient parts repair.

Component histories, flight data and maintenance records are combined to enrich data. Two line replaceable components are selected for case studies. Failures are predicted through binary classification. Supervised and unsupervised learning approaches are evaluated to obtain diagnostics models. Multi-echelon models are developed for network stock optimisation. All models are used as plug-ins to a logistics management system.

Failure of a shop replaceable component in a main landing gear and a valve is predicted. Average accuracy and recall for the former is 75%, whereas the latter scores 60%. Available data does not generate reliable diagnostics models. Multi-echelon models calculate significantly reduced network stock and repair lead time with 3 days warning time, while maintaining an aggregate service level of 95% and taking up to 10% false negatives into account.

Conclusions are that: 1) material condition data is required for accurate diagnostics and prognostics; 2) accuracy and timeliness of failure predictions are key drivers of network stock; and 3) data-driven failure diagnostics, combined with multi-echelon models and a logistics management system, reduce lead times and cost of repairs.

Contents

Abbreviations

ACRONYM	DESCRIPTION
ADI	Advance Demand Information
AI	Artificial Intelligence
ALIS	Autonomic Logistics Information System
AOG	Aircraft On Ground
CMMS	Computerised Maintenance Management System
DT	Decision Tree
LRC	Line Replaceable Component
MATLOG	Maintenance Add-on To Logistics
METRIC	Multi-Echelon Technique for Recoverable Item Control
ML	Machine Learning
MLG	Main Landing Gear
NN	Neural Network
OLS	OneLogistics System
PdM	Predictive Maintenance
PHM	Prognostics and Health Management
PRC	Percentage Cost Reduction
PSI	Product Support Integrator
RF	Random Forest
SRC	Shop Replaceable Component
SVM	Support Vector Machine
TCO	Total Cost of Ownership
XGBoost	Extreme Gradient Boosting

1 Introduction

The Netherlands is a hotspot for logistics in Europe when measured with the aggregated logistics performance index (Arvis et al., 2018). Service providers add value to supply chains with their networks and regional warehouses, while some want to become a regional control tower. This concept implies connecting streams of goods and information to facilitate efficient collaboration between service providers. The regional control tower could be enhanced through the application of prognostics.

Prognostics are increasingly used in logistics and maintenance, but data fusion between these domains is rare. One example is the F-35 Air System, which comprises of the F-35 Air Vehicle and the Autonomic Logistics Information System (ALIS). Its Prognostics and Health Management (PHM) system is a key enabler for autonomic logistics, which features on-board enhanced diagnostics and off-board prognostics. The objective of autonomic logistics is *"to generate the required number of sorties at the lowest cost"* (Hess, Calvello and Dabney, 2004).

Supply chains sometimes have relatively short planning horizons and store extra spare parts in warehouses at strategic locations to fulfill customer demand with acceptable delivery times. However, this approach results in additional logistics cost and a larger logistics footprint. Reliable predictions of Line Replaceable Component (LRC) failures enable the supply chain to deliver spares just-in-time. The logistics process becomes more efficient due to a longer planning horizon. Even a one-day warning time can substantially reduce the network stock of spares. The hypothesis in this paper is:

"The application of prognostics helps to reduce network stock, logistics cost and repair cost."

Models predict failures and failure modes of LRCs with sufficient accuracy. Failure mode in this context means the failed Shop Replaceable Component (SRC) within the LRC. Information about the moment of failure is used to reduce network stock and logistics cost, whereas information about the failure mode is used to reduce maintenance cost in terms of troubleshooting and repair. The aforementioned concept could be applied to asset management of aircraft, infrastructure, installations, vehicles and vessels. A constraint to the concept is the presence of a geographically distributed supply chain for repairables. This paper aims to prove the concept of a maintenance add-on to logistics for the European supply chain of a military aircraft program.

2 Literature survey

2.1 Predictive maintenance

Inside as well as outside the (aircraft) maintenance domain, business analytics play an essential role in Industry 4.0 applications. To improve and support practical applications of predictive maintenance (PdM) solutions, Tiddens (2018) proposed a framework for optimally approaching the main problems of PdM applications based on the available data type, solution technique and goal of resulting solution. The approach of our study can be classified as a data driven model-based approach for diagnostic and prognostic purposes with access to asset history and usage data. As Tiddens (2018) states the data model-based approach relies on data analytics to extract and fuse features from various data sources and apply Machine Learning (ML) algorithms to derive patterns.

A systematic literature review of machine learning methods applied to predictive maintenance is conducted by Carvalho *et al.* (2019). The literature explores common and state-of-the-art ML techniques applied to PdM. Notable information from their review is that Artificial Intelligence (AI) or ML approaches are increasingly applied in PdM applications and are proven to outperform classic statistical approaches. The authors conclude that the top 5 most used ML methods for PdM are: Random Forests (RF), Neural Networks (NN), Support Vector Machines (SVM) and k-means clustering. Moreover, the most common data type used in the reviewed studies was vibration signals.

Finally, to tackle aircraft domain specific PdM difficulties, Adhikari, Rao and Buderath (2018) propose a data driven diagnostics & prognostics framework. Similar to Tiddens (2018) the authors stress the importance of identifying features with sufficient prediction value with regards to component degradation levels. The proposed framework follows a common ML approach to PdM application with additional domain specific steps. Namely, Data pre-processing, feature engineering, anomaly detection, fault isolation and identification.

2.2 Predictive logistics

The literature survey focuses on existing methods and models for inventory control optimization in a multi-echelon spare parts distribution system, subject to an availability constraint. Furthermore, the literature survey covers existing methods to include advance demand information (i.e. predictions of component failures) and lateral transhipments in these inventory control models.

The typical problem of repairable inventory systems is concerned with the optimal stocking of repairable components at local and central warehouses. The central warehouse manages repairs of failed components returned from local warehouses while providing some predetermined service level. The objective of such a system is typically minimizing the backorders and hence the number of grounded aircraft, subject to a budget constraint.

The Multi-Echelon Technique for Recoverable Item Control (METRIC) model represents a fundamental development in repairable inventory theory. Many repairable inventory theory models are based on Sherbrooke’s METRIC model for setting inventory levels and allocating components to achieve some desired level of expected backorders at the local warehouse level (Sherbrooke, 1968, 1986 & 2004). METRIC takes a system view of the repairable inventory problem, since it is concerned with setting the initial levels of inventory and the distribution of the inventory among the warehouses so as to support a system-wide objective of minimizing backorders. A continuous one-for-one inventory replenishment policy is used. The repairable inventory problem faced by the METRIC model is identical to the problem at hand. Therefore, the METRIC model is selected as the baseline model.

Demand lead times are the opposite of supply lead times for including Advance Demand Information (ADI) (Hariharan & Zipkin, 1995). The effect of a demand lead time on overall system performance is precisely the same as a corresponding reduction in the supply lead time. Predictions of LRC failures result in early warnings, which indicate that an SRC (e.g. a circuit board) within an LRC (e.g. a radio) is going to fail within a few days with a certain probability. The warning time is the time from the moment a warning is received until the moment the LRC fails. Therefore, by using the ADI of failures the inventory planner knows a few days in advance where an LRC will be needed with a certain probability. Based on this, the decision is made to ship a part from the central warehouse. Besides, the early warnings also specify the failure mode, so it is also known in advance which SRC will cause the LRC to fail. The repair shop can use this information to order the SRC that is needed for the repair earlier. Therefore, the warning time resulting from the ADI of failures, has to be included in the model as a reduction in the supply lead times.

The imperfectness of the ADI has to be included in the model as well. This is done by making a distinction between false negatives and false positives. False negatives are LRCs that fail without advance warning and false positives are LRCs that fail later than predicted. For false negatives the result is no reduction in supply lead times, whereas for false positives there is a reduction in supply lead times, but the LRC is actually not needed yet. Topan et al. (2018) revealed that using imperfect ADI yields substantial savings, but only when the demand lead time is larger than the supply lead time. Furthermore, having fewer false negatives is more desirable than having fewer false positives and returning excess inventory is quite effective in coping with the consequences of false ADI, particularly for slow-moving items. The one-for-one inventory replenishment strategy can still be used, but the best way to incorporate ADI when using real-time condition-based sensor information is adopting a condition-based inventory policy (Lin et al., 2017; Li & Ryan, 2011).

Finally, for including lateral transshipments the METRIC model must be extended with a new variable, namely the fraction of demand satisfied by lateral transshipments. Alfredsson & Verrijdt (1999) compared their model with lateral transshipments to the VARI-METRIC model and they found a maximum cost reduction of 43.9% and a minimum cost reduction of 13.2%. The results also show that in many cases the stock levels are lower, especially the central stock level. In addition, Kranenburg (2006) discussed a semi-conductor company and showed that using reactive lateral transshipments can save the company up to 50% of annual inventory related costs for spare parts. These results indicate that using lateral transshipments in a distribution network for service parts can be very beneficial. Therefore, the use of lateral transshipments is investigated for the spare parts supply chain of a military aircraft program. This is done by including lateral transshipments in the models and by comparing results to output from models without lateral transshipments.

3 Methods

3.1 Predictive maintenance

3.1.1 Approach

Component histories, flight data and maintenance records are combined to generate a training and test set for this study. Namely, a historic time series for each unique component is constructed by summarizing the parameters of each flight (or cycle) into relevant features. This leaves one with a time series of damage/wear and maintenance features for each component from new state until failure. As mentioned previously, damage and wear are of cumulative nature, i.e. the component accumulates a certain amount of damage/wear until it fails. Therefore, the extracted usage and maintenance features are summed up accumulatively for each cycle to generate a cumulative feature set. To accomplish the latter, the selected features must be available for the majority of flights. Thus, the completeness of each components historical usage time series is of major importance to the quality of the training samples.

Failure prognostics

Note that for prognostic purposes, each flight from the historic time series is treated as a training sample for the binary classification model. We assuming that decisions should take place X cycles before the actual failure occurs to enable just-in-time spare parts delivery. Samples are labeled true if failure occurred X cycles later. Part of the effort includes determining an acceptable tradeoff between X and model performance (see Figure 1).

Figure 1: Illustrative example for binary class labels where X = 3

Failure diagnostics

For diagnostics on the other hand, failure mode classification is done at a time when some failure is confirmed. Specifically, the samples used to obtain diagnostic models are the samples with parameters and features recorded at confirmed moments of failure.

3.1.2 Data

The dataset used for this study is taken from a record system which measures and records a variety of on-board digital and analog data. In addition to the operational data, maintenance related features as well as failure labels are extracted from the maintenance logs. Both failure moment and reason are extracted to obtain labels respectively for failure prognostics and diagnostics.

Usage features

For the usage related features, the specific interest goes to the moment of touch-down during the landing, as this provides the relevant information for damage/wear done to the Main Landing Gear (MLG). Therefore, only processes have been selected that contain this moment at a sufficiently high sample rate. To determine the damage/wear done to the MLG during a landing, the most relevant data would be the strains in the MLG itself. Such direct measurement, however, is not available. Therefore, derived parameters have been selected. The applicable parameters are believed to be parameters related to forces on the MLG: aircraft weight and accelerations, both translational and rotational. Additionally, the flight angles during touch-down are related to the way the forces act on the MLG. One step further away from the strains are the aircraft velocities directly before touch-down, providing some measure for the landing severity. More general features like number of touchdowns are included as well.

Maintenance features

Key maintenance features are derived from the maintenance logs. For example, number of cycles since last inspection and number of time component is swapped from one aircraft to another. Additionally, features related to failure of other components related to the studied components are extracted as well. Finally, historic maintenance logs are merged with the usage data set to synchronize data points and enable generating cumulative maintenance features (see Figure 2).

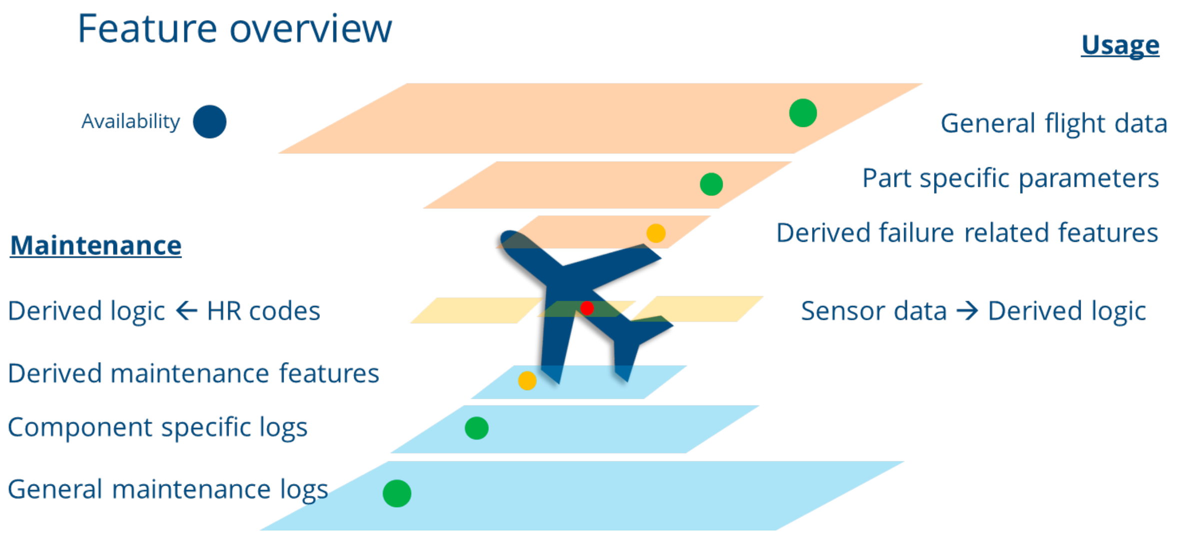

Figure 2: Overview of features available for this study (Green, yellow and red respectively correspond to available, partially available and not available)

3.1.3 Models

Explainability of the machine learning models is key for practical acceptance and certification of the proposed solution. Therefore, we mainly focus on obtaining so called white-box models such that (partial) decision logic can be extracted and validated by domain experts.

Supervised learning

For this study we explore a supervised learning approach for components where labels can be determined from the available data. Established models such as Decision Trees (DT), Random Forests (RF) and eXtreme Gradient Boosting (XGBoost) are trained for prognostic and diagnostics purposes. Binary classification for prognostics is single label while the failure mode classification for diagnostics is multilabel. The latter means a component can fail due to multiple causes. DTs and RFs are inherently capable of multilabel classification. To enable multilabel classification for XGBoost we employ the One-vs-Rest method.

Unsupervised learning

Certain components might have a missing description of failure. As part of this study we attempt to determine classes of these missing labels in an unsupervised fashion. Dimensionality reduction techniques, such as principal component analysis, are employed to visually validate clustering results.

3.1.4 Experiments

Common Machine Learning (ML) practices are employed for data preparation, model training, (cross)validation, testing and parameter tuning. In this section we give a brief overview of the experimental setup.

3.1.4.1 Data split

For prognostic models: training, validation and test sample is split by unique component ID such that the trained model is evaluated on historic samples from components it has not seen during training.

3.1.4.2 Performance measures

Considered inventory models require a high rate of true positive predictions therefore in this paper we focus present performance results in overall accuracy and recall. For multilabel diagnostic classifiers, we expose performance based on sample accuracy and hamming loss. For more details on these performance metrics we refer the reader to the literature.

3.2 Predictive logistics

The multi-echelon models are based on the METRIC model that is selected from literature. First, the multi-echelon model without ADI is designed. This model determines the optimal stock levels for the supply chain when ADI of component failures is not used. This model serves as the baseline. Next the multi-echelon model with ADI is designed. Finally, the model with ADI and lateral transhipments is designed.

3.2.1 Model description

There is a non-empty set J^{loc} of local warehouses, numbered $j = 1, \ldots, |J^{\text{loc}}|$. Each local warehouse serves a number of aircraft, which is stated in parentheses in Figure 3. Each aircraft consists of a non-empty set of repairable items, which are called LRCs, numbered $i = 1, \ldots, |I|$. The total stream of failures of item $i \in I$ as observed by local warehouse $j \in J^{\text{loc}}$ constitutes a Poisson process with a constant rate $\lambda_{i,j} \geq 0$, as indicated in Figure 3. For at least one item i and local warehouse j, it holds that $\lambda_{i,j} > 0$. Apart from the local warehouses, there exists a central warehouse, denoted by index 0. Define $\lambda_{i,0} = \sum_{j \in J^{\text{loc}}} \lambda_{i,j}$ as the total demand rate for item i at the central warehouse. The demand at the central warehouse is also a Poisson process, since it is the superposition of the Poisson demand processes at the local warehouses. Let J denote the set of all stock points, i.e., $J = \{0\} \cup J^{\text{loc}}$.

If an item i fails at a local warehouse j, a spare part stocked at the local warehouse j is used to replace the defective item, if local warehouse j has a part on stock. Otherwise, a backorder arises, until a spare part becomes available from the central warehouse and this results in downtime for the aircraft. Upon failure, also immediately a replenishment order is placed at the central warehouse. The replenishment order arrives after a deterministic lead time $L_{i,j}$ (supply lead time), if stock for item i is available at the central warehouse. Otherwise, the order is backordered until a spare part becomes available from a repair shop. It is assumed that, for each item $i \in I$, backordered replenishment orders from the local warehouses at the central warehouse are fulfilled according to first-come first-served policy.

The defective part at the local warehouse is immediately sent to the assigned repair shop to be repaired there. It takes a certain random lead time with mean $L_{i,0}$ (i.e. repair lead time) before the defective item is repaired and back in stock at the central warehouse. Repaired parts are considered as new parts. Equivalently, from a modeling point of view, the defective part can be scrapped and after a certain random lead time, a newly purchased part is back in stock at the central warehouse.

Notice that each item i at each stock point j is controlled according to a one-for-one inventory replenishment strategy, with base stock level $S_{i,j}$. The policy in the total network is defined by the $|I| \times |J|$ matrix $\mathbf{S}$, consisting of elements $S_{i,j}$. Each column in this matrix, denoted by a vector $\mathbf{S}_j$, consists of all base stock levels at stock point $j \in J$.

3.2.2 Assumptions

These assumptions are made in the model:

- Demands for the different items occur according to a stationary Poisson process.
- All items are repaired successfully and there is no scrapping of items.
- There are no lateral transhipments in the distribution network (relaxed later).

- Repair lead times for different items are independent and random.
- For each item, the supply lead times are assumed to be deterministic.
- A one-for-one inventory replenishment policy is applied for all items.
- Replenishment orders at the central warehouse are fulfilled in first-come first-served order.
- Ample servers are available at the repair facility, hence no waiting queue is present before repair is started.
- The probability of failure of one item is independent of failures occurring for other items.
- Each item failure is caused by a failure of at most one single subcomponent.

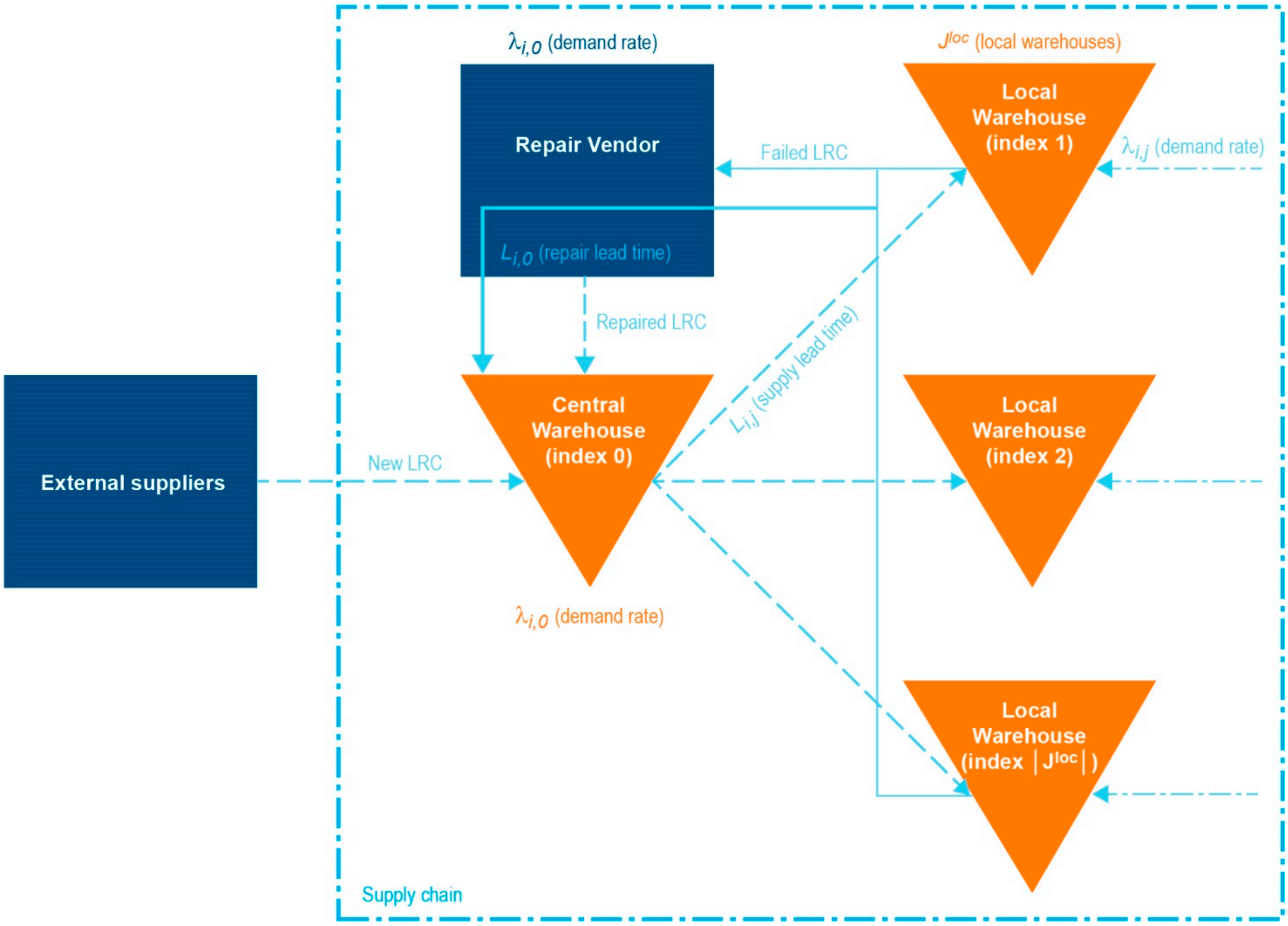

Figure 3: Supply chain including notations for a military aircraft program

3.2.3 Optimization problem

Inventory investment cost c_i^{inv} is counted for every item i put on stock, and the aggregate inventory investment costs are given by:

$$C(\mathbf{S}) = \sum_{i \in I} \sum_{j \in J} c_i^{\text{inv}} S_{i,j}$$

The expected number of backorders for an item $i \in I$ at local warehouse $j \in J^{\text{loc}}$ at an arbitrary point in time at the long run is given by $EBO_{i,j}(S_{i,0}, S_{i,j})$,. Therefore, the aggregate expected number of backorders is:

$$EBO_j(\mathbf{S}_0, \mathbf{S}_j) = \sum_{i \in I} EBO_{i,j}(S_{i,0}, S_{i,j})$$

At local warehouse j, there is a maximum level EBO_j^{obj} given for the aggregate expected number of backorders. The goal of the METRIC model is to determine optimal stock levels that minimize the total inventory investment costs subject to a target for the expected number of backorders per local warehouse. The optimization problem is formulated as follows:

$$\begin{gathered}\text{Min } C(\mathbf{S}) \\ \text{Subject to } EBO_j(\mathbf{S}_0, \mathbf{S}_j) \leq EBO_j^{\mathrm{obj}}, \qquad \forall j \in J^{loc} \\ S_{i,j} \in \mathrm{N}_0, \qquad \forall i \in I, \qquad \forall j \in J,\end{gathered}$$

where N_0 denotes all positive integers starting from 0.

3.2.4 Greedy algorithm

Feasible solutions can be obtained in an efficient way via a Greedy procedure similar to the procedures described in Wong et al. (2007). The basic idea of the Greedy algorithm is the following. For each level of central stock and each local warehouse the expected backorders must be computed as a function of the local stock. For each level of central stock, the optimal allocation of the units of stock to the several local warehouses must be determined, so as to minimize the sum of expected backorders at all local warehouses. This is accomplished by a marginal allocation. At each step, the next unit of stock is added to that local warehouse where it will cause the largest decrease in expected backorders.

3.2.5 Advance demand information

The warning time resulting from the ADI of a component failure has to be included in the model as a reduction in the supply lead time (Hariharan & Zipkin, 1995). The variable that will be used for the warning time is $L_{i,j}^{\mathrm{war}}$. Therefore, the new supply lead times for satisfying a demand are $\max\{L_{i,j} - L_{i,j}^{\mathrm{war}}, 0\}$ for the central warehouse. The imperfectness of the ADI has to be included in the model as well. This is done by making a distinction between false negatives and false positives. False negatives result in backorders for Aircraft On Ground (AOG) situations when the local warehouse has no stock and do not reduce supply lead times. False positives lead to reduced supply lead times, but components are actually not needed yet. If, for example, the false negative percentage is 10%, the supply lead times for the central warehouse become $\left(0.9 \times \max\{L_{i,j} - L_{i,j}^{\mathrm{war}}, 0\}\right) + \left(0.1 \times L_{i,j}\right)$.

Until this point only the effect of warning time on the required stock levels is accounted for. However, early warnings also specify the failure mode of an LRC (i.e. the Shop Replaceable Component (SRC) causing the failure). The repair shop can use this information to order the SRC needed for the repair earlier in time. The effect of warning time on the repair lead time is added. Simulation models are developed for the situations with and without warning time where the repair lead time is simulated. From these simulations, the average repair lead time can be determined in both situations. At this point the effect of warning time on the repair lead time is known. The average repair lead time with and without warning time is then used in the designed multi-echelon models as input variable. The multi-echelon models will give the required stock levels as output for both situations. Therefore, the effect of ADI on the repair lead time and in turn on the required stock levels is determined.

3.2.6 Lateral transhipments

The same two-echelon model is considered as before, but now the following alternative option is considered to satisfy a demand for an item $i \in I$ at local warehouse $j \in J^{\mathrm{loc}}$ if local warehouse j is out of stock; lateral transshipment from a local warehouse. The variables used are $L_{i,j,k}^{\mathrm{lat}}$ for the lateral transshipment lead time, and $c_{i,j,k}^{\mathrm{lat}}$ for the costs of the lateral transhipment. The procedure is as follows. First, the stock is checked at the central warehouse. If the central warehouse has an LRC on stock then the demand is satisfied by the central warehouse. If the central warehouse is out of stock, then the stocks are checked at one or more other local warehouses $k \in J^{\mathrm{loc}}$, $k \neq j$, that have an LRC on stock. If one of these local warehouses has an LRC on stock, then the demand is immediately coupled to that item and the LRC is delivered at the required place. If two or more local warehouses have an LRC on stock, then the location with the lowest demand rate is chosen as the sending source because this location has the lowest backorder probability and the least impact on the aggregate availability level. If none of these local warehouses have an LRC on stock, then the demand is backordered and this results in an AOG situation. The demand is satisfied as soon as a part becomes available from a repair shop at the central warehouse.

4 Results

4.1 Predictive maintenance

The purpose of the failure prediction model we aim to obtain is to serve as a trigger mechanism for a logistic model, which in turn enables just-in-time delivery and inventory optimization. Therefore, the objective of our experiments was to determine if a reliable model can be obtained and for which prediction time window, i.e. how many cycles before a failure occurs can we accurately predict. Moreover, to enable further optimization of the logistic chain this study aims to obtain diagnostic models for failure classification as well. Two components are considered, namely a landing gear component and a pressure valve for respectively prognostics and diagnostics.

4.1.1 Prognostics

Prognostic model results are presented below for the landing gear component. Prediction window of X cycles corresponds to the binary classification if a component is predicted to fail within X cycles.

Landing gear

As mentioned previously, model training and parameter tuning is performed using 5-fold group split cross-validation on a time series corresponding to 160 unique landing gear components. To deal with the class imbalance only samples recorded no longer that 50 cycles prior to failure are considered. The best results are presented in Tables 1 and 2 for the landing gear and valve respectively.

Table 1: Summary of binary classification model results after model parameter tuning for the landing gear component (component failure within next X cycles)

Model	Prediction window	Accuracy	Recall
Decision Tree	5 cycles	63 %	58 %
	10 cycles	78 %	87 %
Random Forest	5 cycles	70 %	71 %
	10 cycles	79 %	80 %
XGBoost	5 cycles	73 %	67 %
	10 cycles	80 %	75 %
Average	5 cycles	68 %	65 %
	10 cycles	79 %	80 %

One can observe that better results were obtained with a prediction window of 10 cycles compared to 5 cycles prior to failure. However, individual (10 cycle window) models do not show significant deviation from the average performance of 79% accuracy and 80% recall. Thus, the differences may be due to a lucky test set split.

Pressure valve

Similar to the landing gear, data is split using 5-fold cross-validation by (50) unique valve components. However, for the valve we were only able to obtain random forest and XGBoost models that do not overfit to the training data entirely. However, the performance of the best obtained model is poor. This might be due to the low prediction value of the available features with regards to the component degradation.

Table 2: Summary of binary classification model results after model parameter tuning for the valve component (component failure within next X cycles)

Model	Prediction window	Accuracy	Recall
Random Forest	5 cycles	65 %	38 %
	10 cycles	63 %	46 %
XGBoost	5 cycles	67 %	33 %
	10 cycles	61 %	60 %

4.1.2 Diagnostics

The performance results of the obtained diagnostic models for the valve component are given below for both the unsupervised and supervised learning approach. Recall that only (50) samples at moment of failure are used for the diagnostic models.

Unsupervised models

Dimensionality reduction of the available features combined with the available multilabel failure modes of the valve indicate there are no clear clusters.

Supervised models

The imbalance of labels for the considered failure modes has led to poorly performing diagnostic models which overfit to the imbalance of the individual labels. This can be observed from the results in Table 3 where the overall sample accuracy is relatively low while the hamming loss is low as well. Furthermore, the poor performance of Random Forest and XGBoost is likely due to high correlation of the available features. These models profit from a large number of features.

Table 3: Summary of multilabel failure classification model results, after model parameter tuning, for the valve component

Model	Sample Accuracy	Hamming Loss
Decision Tree	20 %	27 %
Random Forest	33 %	22 %
XGBoost	20 %	22 %
R Neural Net	53 %	13 %

4.2 Predictive logistics

The numerical results are obtained by implementing the optimization problem and corresponding Greedy solution algorithm. Furthermore, simulations are developed to validate the results of the different multi-echelon models.

4.2.1 Base model

The model results in Table 4 show that the base model requires a network stock of 17 LRCs and achieves an availability level of 95.9%. 9 LRCs are kept in stock at the central warehouse and every local warehouse keeps 1 LRC in stock, i.e., $S_{i,0} = 9$ and $S_{i,j} = 1$. These results are the baseline results.

Table 4: Base model output for aggregate availability

Stock point J	Stock levels $S_{i,j}, S_{i,0}$	Expected availability	Expected inventory	Expected backorders	Expected waiting time
Local warehouse 1	1	97.3%	0.97	0.0004	0.0580
Local warehouse 2	1	98.6%	0.99	0.0001	0.0262
Local warehouse 3	1	98.5%	0.98	0.0001	0.0257
Local warehouse 4	1	95.6%	0.96	0.0010	0.0906
Local warehouse 5	1	93.6%	0.94	0.0022	0.1247
Local warehouse 6	1	98.6%	0.99	0.0001	0.0237
Local warehouse 7	1	95.7%	0.96	0.0010	0.0774
Local warehouse 8	1	98.9%	0.99	0.0001	0.0266
Central warehouse	9	95.9%	2.95	0.2028	3.2442

4.2.2 Model with advance demand information

The model results in Table 5 show that including ADI of component failures in the base METRIC model results in a significant reduction of the network stock and therefore, in a significant reduction of the inventory investment costs. When facing perfect ADI, i.e., 3 days warning time without false negatives, the model with ADI is the best model to use. This model results in a network stock of 10 LRCs, a 96.7% availability level, and an inventory investment reduction of 41.2% compared to the base model. All 10 LRCs should be kept in stock at the central warehouse (centralized allocation), i.e., $S_{i,0} = 10$ and $S_{i,j} = 0$.

Table 5: Output model with warning time for aggregate availability

	Warning time $L_{i,j}^{\text{war}}$ (days)					
	0	1	2	3	4	5
Local warehouse 1	1	1	0	0	0	0
Local warehouse 2	1	0	0	0	0	0
Local warehouse 3	1	0	0	0	0	0
Local warehouse 4	1	0	0	0	0	0
Local warehouse 5	1	0	0	0	0	0
Local warehouse 6	1	0	0	0	0	0
Local warehouse 7	1	0	0	0	0	0

	Warning time $L_{i,j}^{\text{war}}$ (days)					
	0	1	2	3	4	5
Local warehouse 8	1	0	0	0	0	0
Central warehouse	9	13	11	10	10	9
Network stock	**17**	**14**	**11**	**10**	**10**	**9**

4.2.3 Model with lateral transhipments

When facing imperfect ADI, i.e., 3 days warning time with 10% false negatives, the model with ADI and lateral transhipments is the best model to use (see Table 6). This model results in a network stock of 11 LRCs, a 95.8% availability level, and an inventory investment reduction of 35.3% compared to the base model. 3 LRCs should be kept in stock at the central warehouse and every local warehouse should keep 1 LRC in stock (decentralized allocation), i.e., $S_{i,0} = 3$ and $S_{i,j} = 1$. It can be concluded that using reactive lateral transshipments between the local warehouses almost neutralize the negative effect of the false negatives.

Table 6: Simulation results using min BO, 3 days warning time with 10% false negatives, and $L_{i,0}$=85

	Network stock = 11	Network stock = 12	Network stock = 13
Availability Level simulation (without lateral transhipments)	91.6%	94.5%	95.7% *
Availability Level simulation (with lateral transhipments)	95.8% *	97.8%	98.5%

* indicates optimal solution

4.3 Proof of concept

A demonstrator is built to show how generated information leads to decisions of a supply chain manager. This demonstrator is based on the OneLogistics System (OLS), which uses ILIAS as its Enterprise Resource Planning (ERP) backbone, a predictive logistics add-on by Gordian and a predictive maintenance add-on by NLR. The demonstrator shows the data flow from an operator's Computerised Maintenance Management System (CMMS) to the OLS and its add-ons.

The scenario presented in Figure 4 is used to clarify the data requirement and to visualise interpretation and management of data in the supply chain. Figure 5 shows the information presented on the OneLogistics Console to the supply chain managers. The proof of concept visualises the actions and interventions that supply chain managers initiate to minimise downtime for the operator, whilst reducing the required number of spares to minimise Total Cost of Ownership (TCO) for the military aircraft program. Pro-active failure information ensures that LRCs are shipped on time from the Central Warehouse or the Repair Vendor to the required Local Warehouse. Supply chain managers at OneLogistics use periodic data analysis to advice the Product Support Integrator (PSI) to reduce or expand the available spare parts in Europe.

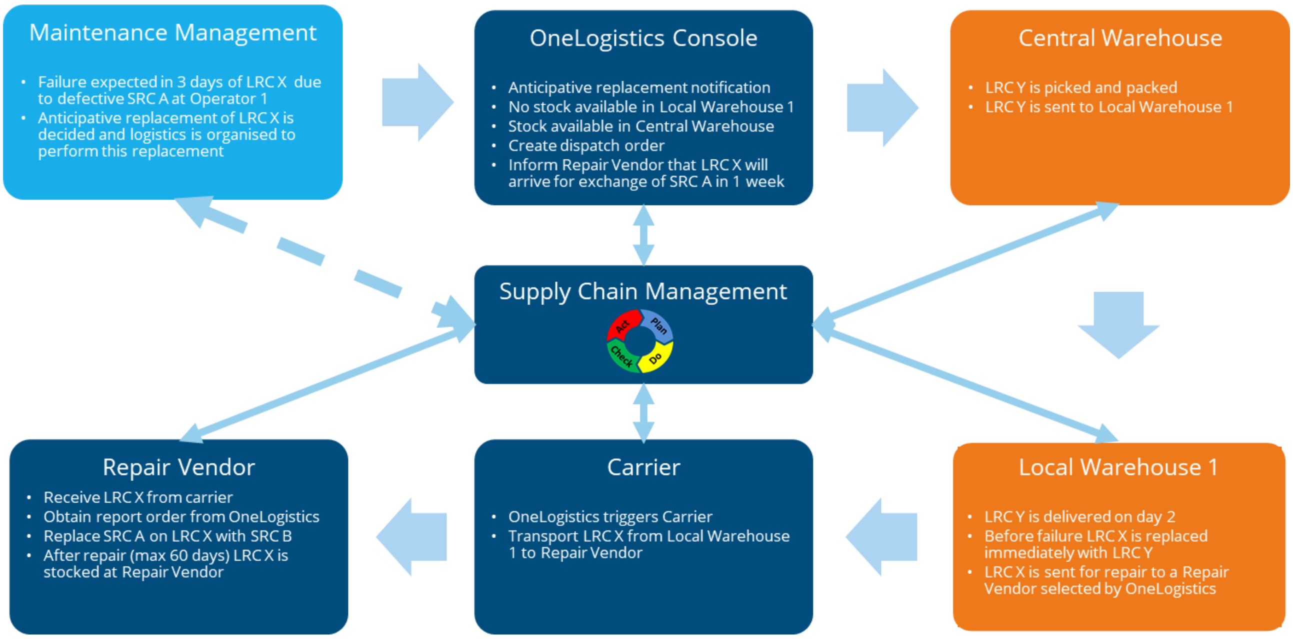

Figure 4: Scenario for proof of concept

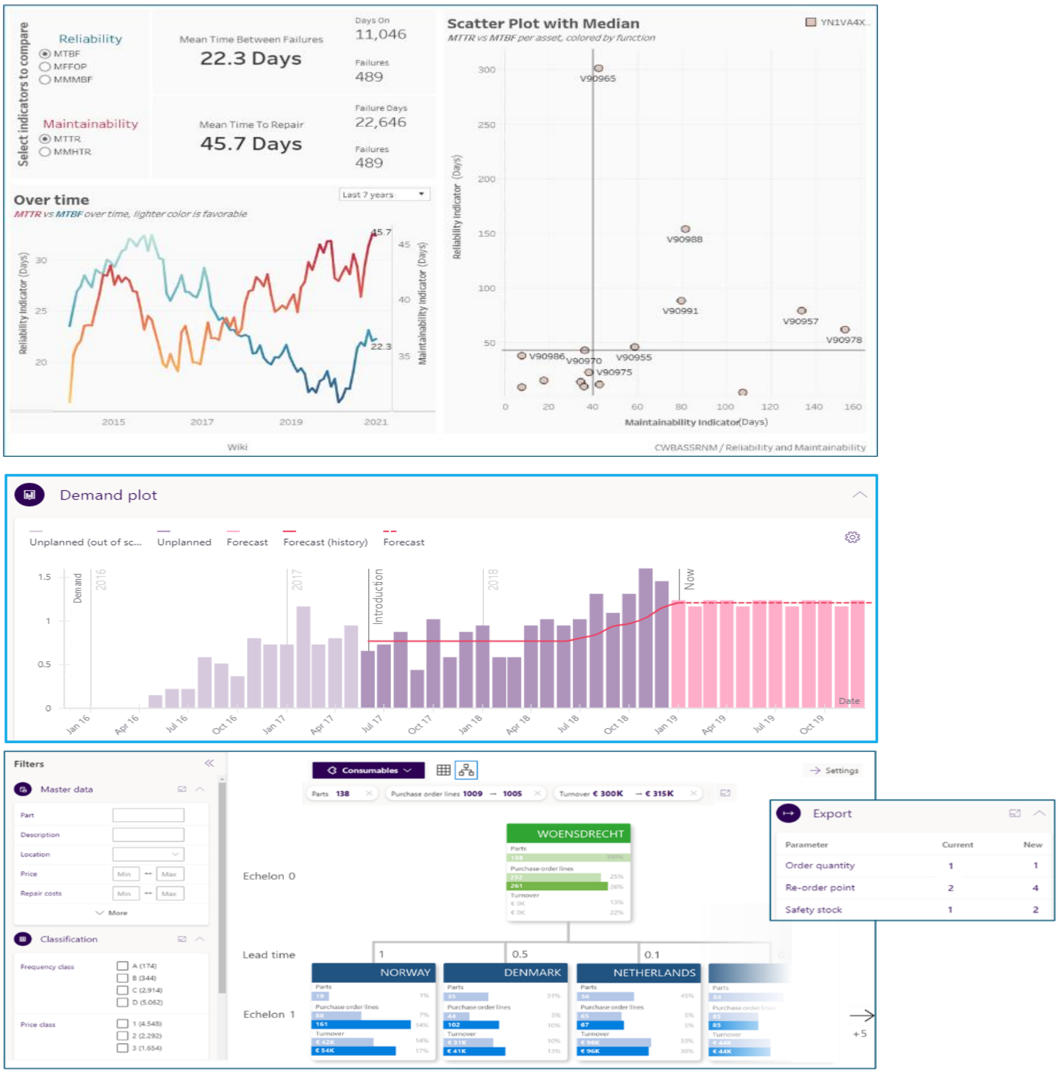

Figure 5: OneLogistics Console

5 Conclusions and recommendations

The following conclusions are drawn from the work presented in this paper:

- Material condition data is required for accurate diagnostics and prognostics. Usage data at aircraft level or system level is often not representative for failure behaviour of components. Furthermore, maintenance data in the form of unscheduled removals does not provide information about the degradation process of components. Material condition data provides insight in the degradation process of components that eventually leads to failure.
- Accuracy and timeliness of failure predictions are key drivers of network stock. False positives result in unnecessary shipments of spare parts to local warehouses, which implies additional logistics cost. False negatives result in aircraft on ground events, if no spares are available in local warehouses. Given the availability level, more spare parts in local warehouses and lateral transhipments are required, since defect rectification is less expensive than downtime. The sensitivity analysis indicates that increased warning time and mean time between failures results in a lower demand for network stock. Besides, shorter repair lead times facilitated by accurate failure diagnoses help to reduce the network stock requirement.
- Data-driven failure diagnostics, combined with multi-echelon models and a logistics management system, reduce lead times and cost of repairs. Development of a reliable data-driven approach requires access to the failure mode, effect and criticality analyses, as well as engineering analyses of components. Knowledge obtained through the design process is used to select parameters, which are used to predict failure of components in-service on the basis of sensor data.

From the conclusions above follow these recommendations for future work:

- Every failure mode needs a dedicated model that is fed with data tailored to its nature. Data available in-service must be sufficient to classify and characterise the failure mode to be detected. The added value of this data is increased by using a predefined syntax and recording meta data of events (e.g. reason for removal or actions taken for defect rectification).
- A Greedy heuristic could be developed that is able to solve the METRIC model with lateral transhipments. The heuristic should be based on an efficient, but still accurate, approximate evaluation method.
- The developed models could be extended with stochastic supply lead times, taking disturbances in the supply chain and disturbances in flight operations into account. In this paper deterministic supply lead times are assumed. Furthermore, these models could be applied to an inventory problem with more than two echelons to investigate differences in performance.
- Coupling predictive maintenance and predictive logistics is a consideration during the design phase of an aircraft program. The ability to maximise affordability and sustainability depends on the capabilities of the PHM system, as well as the flexibility and responsiveness of the supply chain. Besides, the implementation at the aircraft operator is critical to success. Maintainers and logisticians must be aware of the fundamentals to make the system function properly.
- Legacy aircraft may be equipped with additional sensors to obtain data for diagnostics or prognostics at sub-system or component level. However, integrated vehicle management and autonomic logistics are fully achieved, only when considered during the design phase.

6 References

Adhikari, P., Rao, H.G. and Buderath, M. (2018), “Machine learning based data driven diagnostics & prognostics framework for aircraft predictive maintenance”, in: *10th International Symposium on NDT in Aerospace*, 24-26 October 2018, Dresden, DGZfP, Berlin.

Alfredsson, P. and Verrijdt, J. (1999), “Modeling emergency supply flexibility in a two-echelon inventory system”, *Management Science,* Vol. 45, No. 10, p. 1416-1431.

Arvis, J.F., et al. (2018), *Connecting to Compete: Trade Logistics in the Global Economy*, The World Bank, Washington DC.

Carvalho, T. P., et al. (2019), “A systematic literature review of machine learning methods applied to predictive maintenance”, *Computers & Industrial Engineering*, Vol. 137, Art. 106024.

Hariharan, R. and Zipkin, P. H. (1995), Customer-order information, lead times, and inventories, *Management Science,* Vol. 41, No. 10, p. 1599-1607.

Hess, A., Calvello, G. and Dabney, T. (2004), “PHM a key enabler for the JSF autonomic logistics support concept”, in: *IEEE Aerospace Conference Proceedings*, 6-13 March 2004, Big Sky, MT, IEEE, New York, NY.

Kranenburg, B. (2006), *Spare Parts Inventory Control under System Availability Constraints* (PhD thesis), BETA research school, Eindhoven.

Li, R. and Ryan, J. K. (2011), A Bayesian inventory model using real-time condition monitoring information, *Production and Operations Management,* Vol. 20, No. 5, p. 754-771.

Lin, X., Basten, R. J. I., Kranenburg, A. A. and van Houtum, G. J. (2017), Condition based spare parts supply. *Reliability Engineering and System Safety,* Vol. 168, p. 240-248.

Sherbrooke, C. C. (1968), Metric: A multi-echelon technique for recoverable item control, *Operations Research,* Vol. 16, No. 1, p. 122-141.

Sherbrooke, C. C. (1986), Vari-metric: Improved approximations for multi-indenture, multi-echelon availability models, *Operations Research,* Vol. 34, No. 2, p. 311-319.

Sherbrooke, C. C. (2004), *Optimal inventory modeling of systems: Multi-echelon techniques*, Wiley, Hoboken NJ.

Tiddens, W. (2018), *Setting sail towards predictive maintenance : developing tools to conquer difficulties in the implementation of maintenance analytics* (PhD thesis), University of Twente, Enschede.

Topan, E., Tan, T., van Houtum, G. J. and Dekker, R. (2018), Using imperfect advance demand information in lost-sales inventory systems with the option of returning inventory, *IISE Transactions,* Vol. 50, No. 3, p. 246-264.

Wong, H., Kranenburg, B., Van Houtum, G. J. and Cattrysse, D. (2007), Efficient heuristics for two-echelon spare parts inventory systems with an aggregate mean waiting time constraint per local warehouse, *OR Spectrum,* Vol. 29, p. 699-722.

Acknowledgements

The authors would like to express their gratitude to TKI Dinalog for co-funding the project MATLOG on behalf of the Ministry of Economic Affairs. Furthermore, we express a special thanks to the project team, René de Koning at OneLogistics and Lieutenant Colonel Roberto Joannes at the Royal Netherlands Air Force for the hard work, good ideas and constructive feedback.

Dedicated to innovation in aerospace

Royal NLR - Netherlands Aerospace Centre

NLR operates as an objective and independent research centre, working with its partners towards a better world tomorrow. As part of that, NLR offers innovative solutions and technical expertise, creating a strong competitive position for the commercial sector.

NLR has been a centre of expertise for over a century now, with a deep-seated desire to keep innovating. It is an organisation that works to achieve sustainable, safe, efficient and effective aerospace operations.
The combination of in-depth insights into customers' needs, multidisciplinary expertise and state-of-the-art research facilities makes rapid innovation possible. Both domestically and abroad, NLR plays a pivotal role between science, the commercial sector and governmental authorities, bridging the gap between fundamental research and practical applications. Additionally, NLR is one of the large technological institutes (GTIs) that have been collaborating over a decade in the Netherlands on applied research united in the TO2 federation.

From its main offices in Amsterdam and Marknesse plus two satellite offices, NLR helps to create a safe and sustainable society. It works with partners on numerous programmes in both civil aviation and defence, including work on complex composite structures for commercial aircraft and on goal-oriented use of the F-35 fighter. Additionally, NLR helps to achieve both Dutch and European goals and climate objectives in line with the Luchtvaartnota (Aviation Policy Document), the European Green Deal and Flightpath 2050, and by participating in programs such as Clean Sky and SESAR.

For more information visit: **www.nlr.org**

Postal address
PO Box 90502
1006 BM Amsterdam, The Netherlands
e) info@nlr.nl i) www.nlr.org

Royal NLR
Anthony Fokkerweg 2
1059 CM Amsterdam, The Netherlands
p) +31 88 511 3113

Voorsterweg 31
8316 PR Marknesse, The Netherlands
p) +31 88 511 4444