Dynamic Modeling of Maintenance Prices in the Aerospace Industry

George S. Eskander Ekladious, Xiaojun Zhao, Hala Mostafa, Ramona Georgescu

United Technologies Research Center, East Hartford, U.S.A

ekladigs@utrc.utc.com
zhaox@utrc.utc.com
mostafh@utrc.utc.com
georgera@utrc.utc.com

Abstract

In the aerospace industry, pricing of the maintenance, repair and operations involve complex business rules that are being formulated by domain experts. Automation of such a process becomes a challenge, especially when pricing rules significantly differ based on contract conditions, operators, maintenance shops, etc. This paper presents a pricing prediction approach, where the predictors are dynamically built to fit the different pricing rules. To this end, a clustering mechanism efficiently splits the space to dissimilar clusters that are likely to follow different pricing rules. Then, candidate models are designed and ranked for the different clusters. At the exploitation phase, a testing data sample is assigned to a cluster, and processed using the best model for that cluster. Results show significant accuracy improvement compared to the static modeling approach.

1. Introduction

Prediction of maintenance, repair and operations (MRO) price has a central role in planning and negotiation of future contracts, especially for industries involving expensive MRO processes, such as the aerospace industry (Steven R. Erickson & Summerour., 1997). More specifically, accurate prediction of such prices is needed to optimize the manufacturer warranty reserves (Thomas & Rao, 1999) and maintenance reserve funds (Ackert., 2012).

Estimation of the maintenance price of an aircraft, requires prediction of the price and intervals for all systems and subsystems subject to several major events such as, air-frame heavy maintenance visit, engine overhaul, landing gear overhaul, etc. Such events involve different levels of uncertainty that makes forecasting the maintenance price over the aircraft life cycle, a challenging problem.

Maintenance activities are also influenced by non-engineering factors such as business decisions and environmental conditions which introduce additional analytical challenges.

The traditional methodology to building financial models for MRO activities is to having domain experts formulating the pricing models. Some examples can be found in (Ackert., 2011), (Clifford J. Landreth, 2015), (Kenneth E. Marks, 1981) for estimating maintenance prices for the engine, APU, and air-frame, respectively. Such knowledge-based models need to be manually updated by the domain expert to capture any changes in pricing conditions, such as policy changes and new catalog prices, which is an expensive and a hard to maintain process.

For fully automating the maintenance price estimation processes, a data driven approach is recommended (M. & R., 1979), (Kang M. & C., 2008), (Kennet., 1994), (EZIK., 2003), (Hanumanthan., 2009). This approach is only feasible when enough historical data are available to train the models. In the aerospace industry, shortage of historical data that cover a wide range of contract types, operating conditions, fleets, etc., (due to historical paper records) hinders the analytical development.

In this work, the main challenge we observed is that it is not easy to consider estimating maintenance prices as a single problem that applies across operators. For instance, for individual operators, significant differences might exist in the contract and business rules, environmental conditions, etc., that result in differences in pricing rules for the maintenance performed for each operator. Alternatively, it might be more promising to design different models for different operators. However, this approach might not be possible when a few historical maintenance activities are available for the specific operator. Compounding the problem of having a small dataset to build models on, is having a mixture of conditions including the type of the maintenance activity, the maintenance shop, etc., which causes the problem to be divided into many sub-problems. Coming back to the first approach of generating a single model, that is generic for all of the maintenance activities, is the subject of the current work because of data shortage. Accuracy of the model might differ significantly for the different sub-problems according to the extent to which each sub-problem is captured in the generic model.

In this paper, a new ensemble-type approach for price estimation is proposed, where both training data and regression models are dynamically selected to best fit a specific sample. A previous approach has been applied for designing classification systems (Alceu Britto & Oliveira, 2014), while here we study the applicability of it on regression problems focusing on the price estimation. To this end, our algorithm selects the best model that fits a specific region in the feature space and trains it with a subset of the training data that lies in that region.

Accuracy of the proposed method is compared with both the static data approach, where all training data are used to train the model, and the static models approach where a single model is employed to fit the whole feature space. Results of the dynamic approach in which both data and model selection is performed dynamically show a promising improvement in prediction accuracy compared to both the static data and static models approaches.

Next section illustrates the baseline of the static modeling approach to designing price estimation models. Section 3 describes the proposed dynamic modeling approach. Results of the dynamic approach are reported and discussed in Section 4, followed by a conclusions section.

2. STATIC MODELING

The traditional approach to modeling maintenance prices is to use a single model to represent the whole problem space. In other words, a specific machine learning algorithm and a specific model structure are used to generate a formula that is used to predict the price for all possible combinations of maintenance activities types, and for all operators, maintenance shops, contracts, etc. This approach, even if it succeeds to accurately model specific regions in the feature space has no guarantee that it will perform similarly over the whole space.

For price estimation problems, there are multiple factors that might split the problems into different sub-problems that should be modeled differently. For simple visualization, we consider here only two, out of many, of the factors that we observed to have a significant impact on the pricing models: maintenance type and operator (for simplicity, in Figure 1 both of them are chosen to be binary factors where operator is modeled as a certain operator vs the rest of the operators). With only two binary factors, the feature space is split into

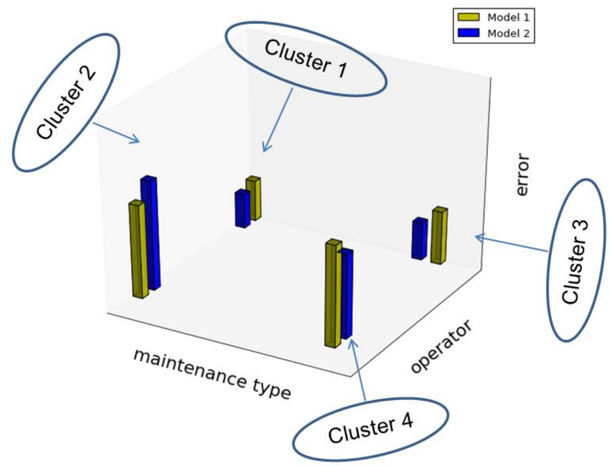


Figure 1. Illustration of regions of competence for the different models.

four different clusters. When considering all possible factors that impact the pricing models, a high dimensional space is generated and a high number of clusters exists.

For the static modeling approach, it is assumed that it is possible for a single model (actually a single formula, single machine learning approach, same training data, etc.) to generalize over the whole feature space. Such assumption was observed to be invalid for the maintenance price prediction problem at hand.

Figure 1 illustrates the disadvantage of employing the static modeling approach. We consider two possible model structures ($Model1$ and $Model2$). It can be seen that cluster 2 is the region of competence for $Model1$ (since $Model1$ provides a lower error than $Model2$), while $Model2$ is more competent everywhere else. The intuition behind our proposed dynamic modeling approach is to learn the regions of competence of all models, using the training data. During the exploitation phase, one can use a model for predicting a sample, only if the sample belongs to its region of competence. The task increases in complexity with a higher dimensional feature space and many models having different regions of competence.

Another assumption behind designing a static model, using the whole training data, is that the training data are equally similar to all future testing samples. This assumption is also questionable, since a testing sample might belong to a region in the feature space that significantly differs from the other training samples. Accordingly, it is logical to not only use region-specific models, but also to train models with region-specific data. Simply speaking, maintenance prices should be predicted using models that work for similar maintenance work and that were trained using similar historical data.

The aforementioned proposals (dynamic selection of models

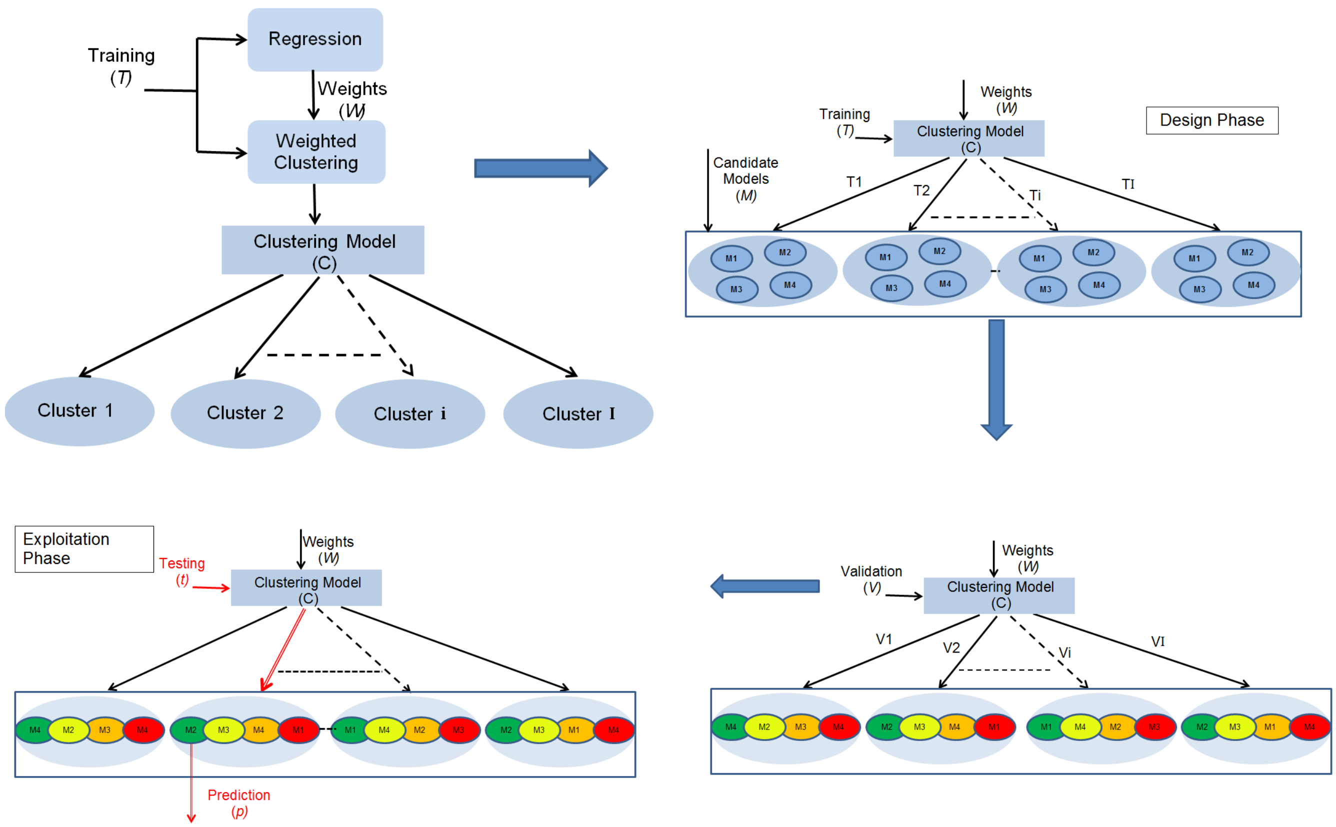


Figure 2. Proposed Dynamic Modeling Approach.

and dynamic selection of data) are combined in the proposed dynamic modeling algorithm, described in the next section. Performance of the proposed algorithm, using the same above illustrative example, is discussed in Section 4.

3. DYNAMIC MODELING

This section describes the proposed dynamic modeling approach. In the traditional static modeling approach, a predictor is trained using the whole training data, and all testing samples are processed using the same predictor. In our proposed approach, each testing sample uses the best fitting predictor that was designed using similar training samples.

Figure 2 illustrates our proposed approach. The target of the design phase is to learn a dictionary of cluster-model pairs, where a cluster represents a subspace that is likely to follow a different pricing model. Each cluster should be assigned a model that differs in its structure, method and training dataset, from the other clusters in the feature space. To this end, we propose a method to cluster the feature space in a way that is guided by the problem at hand (top-left of the diagram). Firstly, a linear regression model is designed using the training data, so that all features in the space are weighted according to their importance. Then, the weights learned are used to manipulate the feature space, by simply weighting the features, so that important features will have more of a say in partitioning the feature space. A pair of samples that have high similarity in the original (unweighted) feature space might become dissimilar in the new (weighted) feature space. For instance, consider the case of providing similar maintenance services to operators that differ in their pricing criteria according to business conditions. Using the original feature space, the similar maintenance services are close and the representation is similar for the different operators. However, they are better separated in the new manipulated space due to the linear regression placing higher weight on the operator feature. Applying this weighted clustering method leads to efficiently splitting the pricing problem into dissimilar subproblems, so that a specific model is then designed for each problem.

Once the feature space is clustered, all models in the candidate model pool are then used to design region-specific versions, where models are trained using the region-specific data (top-right part of the diagram in Figure 2). In order to select the best models that fit each region in the feature space, an independent validation data set is clustered and used to rank the models for each region. Ranking is done based on the accuracy of models for the specific region (bottom-right part of the diagram).

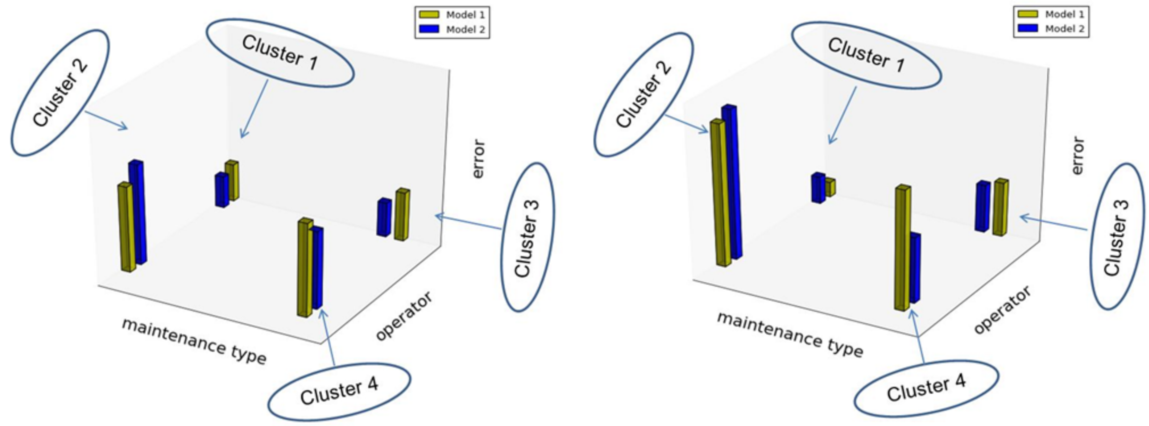


Figure 3. Dynamic selection of models algorithm learns the regions of competence: left) actual ranking, right) learned ranking.

Algorithm 1 Dynamic Modeling- Design Phase

1: **Input:** Training Set T, Validation Set V, Candidate Models M.
2: **Clustering:** T is used to design a clustering model C with I clusters.
3: **Dynamic Data Selection:**
4: **for** each cluster i: **do**
5:     subset T_i of T is extracted from C_i.
6:     T_i is used to train models M.
7: **end for**
8: **Dynamic Model Selection:**
9: **for** each cluster i: **do**
10:     subset V_i of V is extracted from C_i.
11:     samples of V_i are predicted using trained models M, and models are ranked according to accuracy.
12:     best model M_i is assigned to the cluster C_i.
13: **end for**
14: **Output:** Output Cluster-Model dictionary: $\{C_i, M_i\}_{i=1}^{I}$

Algorithm 2 Dynamic Modeling- Exploitation Phase

1: **Input:** Test sample t, Cluster-Model dictionary: $\{C_i, M_i\}_{i=1}^{I}$
2: t is assigned to a cluster C_i using C.
3: best model M_i is located from $\{C_i, M_i\}_{i=1}^{I}$ using C_i.
4: **Output:** produce prediction p for t using M_i

Algorithms 1 and 2 list the steps for the design and exploitation phases of the proposed dynamic modeling approach, respectively. There are three key processes that are executed sequentially: 1) clustering (step 2 in Algorithm 1) where the weights extracted from the regression model are used to cluster the feature space for dissimilar sub-problems, 2) Dynamic data selection (steps 3-7 in Algorithm 1) where the clustered training data are used to design different version of the models for each region, and 3) Dynamic model selection (steps 8-13 in Algorithm 1) where models are ranked, using an independent validation set, and the most accurate model is assigned to each region. The impact of each of these key processes is investigated through the experimental analysis reported in the next section. For all the different experimental setups, we executed Algorithm 2 to compute the predicted price.

4. SIMULATION RESULTS

In order to evaluate the proposed approach, 1000s, 100s and 100s of data points were used for training, validation and testing, respectively. More than 20 inputs, including operator, maintenance type, etc. are used for both clustering the feature space and for designing the pricing models. A variety of candidate models listed in Table 1 that differ in structure, machine learning algorithm, etc. were used in our dynamic modeling framework.

We use the same illustrative example, as discussed in Section 2, to visualize the advantages of the proposed approach.

Figure 3 shows how the proposed Algorithm 1 learned the region of competence of the different models. As in Section 2, this example shows only two inputs being used to split the problem space, and only two model structures being available to the model selection process. In this reduced example, the region of competence of $Model1$ structure is correctly captured, which is cluster 2. This is clear when matching the left plot (actual ranking for the testing set) and the right plot (learned ranked using the validation data set).

The other model structure ($Model2$) is better performing in the other three clusters (clusters 1, 3 and 4). The proposed algorithm correctly learned two out of these three clusters (3 and 4), and incorrectly ranked models in cluster 1. The incorrect ranking results from the dissimilarity between the test-

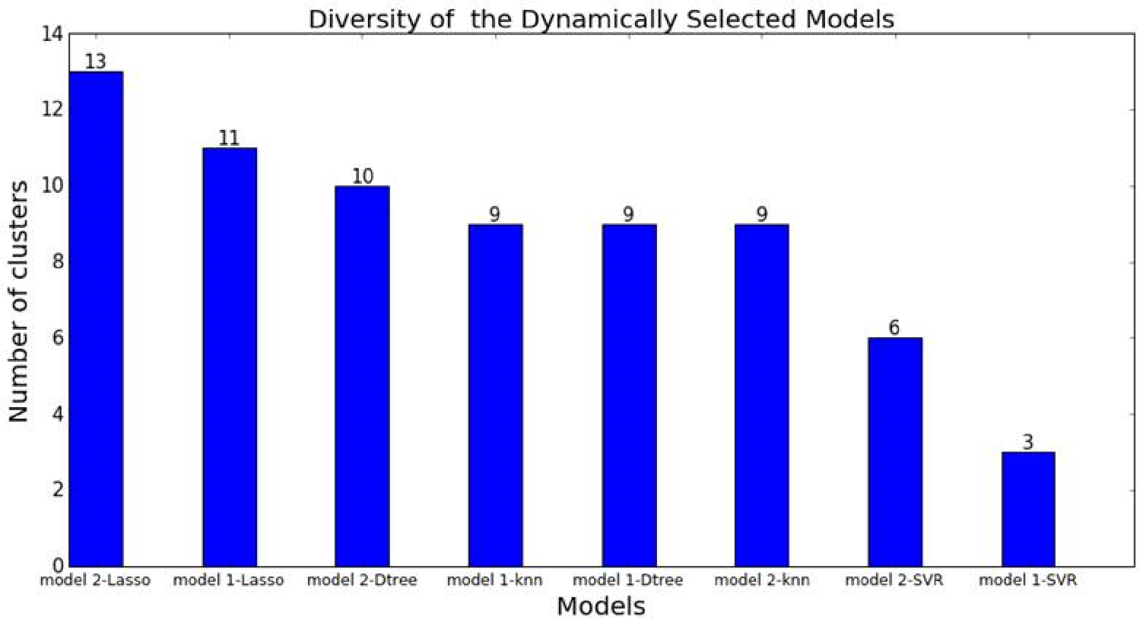


Figure 4. Diversity of the Dynamically Selected Models.

Table 1. Comparison of Using Static Data AND Dynamically Selected Data With Static Models

Metric	Static Data							
	Model1				Model2			
	Lasso	SVR	KNN	Dtree	Lasso	SVR	KNN	Dtree
Err%	2.13	2.01	1.88	2.85	1.25	2.03	1.52	3.12
Metric	Dynamic Data							
	Model1				Model2			
	Lasso	SVR	KNN	Dtree	Lasso	SVR	KNN	Dtree
Err%	2.63	1.95	1.70	1.93	0.41	1.95	2.30	1.56

ing and validation data that lie in the specific region (cluster 1). For improved performance, the clustering method needs more consideration, where some similarity metric learning methods might be investigated for better splitting the problem space.

The above example is used only for illustrative purposes. However, in order to globally evaluate the significance of the proposed approach, we report the prediction results over the whole testing set, and when all splitting factors and models are used by our algorithm.

Figure 4 shows how frequent the different models are found to be most competent in a cluster. For instance, model structure 2 with the Lasso algorithm (Model2-Lasso) is assigned to 13 regions, while structure 1 with the SVR (Support Vector Regression) algorithm (Model1-SVR) is selected only by 3 regions. The distribution of selected models over the problem space provides some insights on the importance of having a diverse of model structures and algorithms in the selection pool.

In order to test the impact of the two main contributions of our algorithm (dynamic data selection and dynamic model selection), we compare the results of applying the individual steps against the static modeling approach (where all training data are used to train the models, and the individual ”off-the-shelf” models are used for prediction).

Table 1 shows the experimental results, where the individual candidate models are used, and no dynamic model selection step is applied (steps 8-13 of Algorithm 1 are not executed, and static models are used instead). Here we only test the significance of employing the dynamic data selection step, by comparing two cases: static data, where all training data are used for training, and dynamic data selection, where data are dynamically selected to fit the specific testing sample. For all experiments, the metric for prediction accuracy reported is Error percentage ($Err\%$).

Note that almost all models are doing better when the training data are dynamically selected for the specific testing sample (in 6 out of 8 cases, the results reported in the bottom cells outperform their corresponding results in the top cells). These

Table 2. Comparison of Using Static Models AND Dynamically Selected Models

Metric	Static Models		Dynamic Models	
	Static Data	Dynamic Data	Static Data	Dynamic Data
Err%	1.25	0.41	0.82	0.2

results supports the claim that it is advantageous to dynamically select training data.

It is also worth noting that, although model structure $Model2$ with the Lasso regression algorithm has best overall prediction accuracy, we observed that the best model changes for the different testing regions, and there is no generic model that performs best all the time. Accordingly, the second main step of our algorithm (dynamic selection of models) is needed, not only to further boost the performance, but also to adaptively learn the best model over time and so sustainability of the pricing modeling is addressed.

Table 2 shows the results when we investigate the impact of the dynamic selection of models step. The best results from Table 1 (for static model selection with both static and dynamic data) are compared to the dynamic selection of model results. Two scenarios of dynamic model selection are tested: with static data (where steps 3-7 of Algorithm 1 are not executed, and all training data are used instead), and with dynamic data (where all the algorithm steps are performed).

It is evident that dynamic selection of models boosted the prediction accuracy for both scenarios. The best performance is achieved when both dynamic data and dynamic model selection steps are executed.

5. CONCLUSIONS

In this paper, we introduced an approach to boosting regression accuracy through dynamic selection of data and models. The method is motivated by the need for designing reliable and sustainable price estimation models for industries that involve expensive maintenance activities, such as the aerospace industry. A practical challenge arises from the shortage of available historical data for the different types of maintenance activities that is needed to design specific models for the different regions in the problem space. Our method benefits from the whole available training set, yet generates specific models for the relevant sub-spaces. A weighted clustering method is proposed to define different regions for which the pricing model is likely to differ. Results have shown significant improvement in pricing prediction accuracy as compared to the traditional static models approach.

There is however room for performance improvement which requires further experimentation over a wider range of price estimation problems, and more generally, on different regression problems. For instance, the weighted clustering step needs more investigation through the exploration of other similarity learning and clustering strategies. Moreover, in this work, we only employed the most competent model for prediction. However, dynamic selection of the most competent ensemble method could lead to improved predictions (Albert H. R. Ko, 2008).

Finally, success of ranking the models properly, through finding the regions of competence of each model, relies not only on the clustering method we propose, but also on the extent to which the validation set is similar to the future testing set. Here, we chose the most recent portion of the available historical data as a validation set, with the promise that the pricing rules are more likely to be smooth in time. However, future work can focus on how to better quantify the impact of selecting the validation data on determining the models' regions of competence.

REFERENCES

Ackert., S. (2011). Engine maintenance concepts for financiers. In *Aircraft monitor.*

Ackert., S. (2012). Basics of aircraft maintenance reserve development and management. In *Aircraft monitor.*

Albert H. R. Ko, A. S. B., Robert Sabourin. (2008). From dynamic classifier selection to dynamic ensemble selection. *Pattern Recognition*, *41*, 1718-1731.

Alceu Britto, R. S., & Oliveira, L. (2014). Dynamic selection of classifiers- a comprehensive review. *Pattern Recognition*, *47*, 3665-3680.

Clifford J. Landreth, L. L. C., Richard H. Wilhelm. (2015). Performance-based logistics. In *Naval postgraduate school.*

EZIK., O. (2003). Calculation of the actual cost of engine maintenance. In *Phd. thesis, air force institute of technology.*

Hanumanthan., H. (2009). Severity estimation and shop visit prediction for civil aircraft engines. In *Phd. thesis, cranfield university.*

Kang M., P. P., Ogaji S., & C., K. (2008). An approach to maintenance cost estimation for aircraft engines. In *Asme turbo expo 2008: Power for land, sea, and air.*

Kennet., M. (1994). A structural model of aircraft engine maintenance. *Journal of Applied Econometrics*, *9*, 351-369.

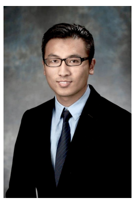
Kenneth E. Marks, R. W. H. (1981). Estimating aircraft depot maintenance costs. In *Rand reports.*

M., D., & R., S. (1979). A technique for engine maintenance cost forecasting. In *4th international symposium on air breathing engines.*

Steven R. Erickson, R. J. M., & Summerour., R. (1997). Aviation maintenance contract management: A survey of defense and commercial practices. In *Logistics management institute.*

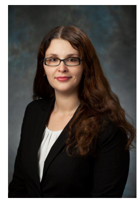
Thomas, M. U., & Rao, S. S. (1999). Warranty economic decision models: A summary and some suggested directions for future research. *Operations Research*, *47*(6), 807-820.

George Ekladious is a Research Scientist at United Technologies Research Center, where he is working on a broad array of business challenges include analytics of jet engine data and diagnostics of building and transporation systems. His main research interests are in the areas of machine learning, biometrics, telecommunication networks and control systems. George received his PhD degree in Automatic Production Engineering from the École de Technologie Supérieure, Montreal, in 2014.



Xiaojun Zhao received his M.S. in mechanical Engineering from the University of Connecticut in 2015. He was a data scientist fellow in the Data Incubator program after graduation. He is currently a Research Engineer at the United Technologies Research Center, East Hartford, CT, USA. His current research interests include data analytics, machine learning, and deep learning.

Hala Mostafa is a Research Scientist at United Technologies Research Center (UTRC) with expertise in reasoning under uncertainty, automated decision making and machine learning. She works on business-oriented data analytics projects as well as more future-looking research projects. Before UTRC, Hala worked at Raytheon BBN Technologies where she was the Technical Lead on government-funded projects. Hala obtained her PhD in 2011 from University of Massachusetts, Amherst.



Ramona Georgescu received her PhD in Electrical Engineering from the University of Connecticut in 2012. She joined United Technologies Research Center in 2013 where she leads the digital project portfolio supporting the needs of United Technologies Aerospace Systems. Her interest is in the development of custom and enterprise analytics addressing problems across product life cycle which brings together machine learning, optimization, big data and user experience. Ramona has one patent, five journal publications, multiple conference publications and is the Associate Editor for Fusion Applications for the ISIF Journal of Advances in Information Fusion.