ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

A Kalman Filter and Neural Network Hybrid Approach for Health Monitoring of Aircraft Engines

Solene Thepaut[1], Sebastien Razakarivony[1], Dong Quan Vu[1] and Alfred Bauny[2]

1- Safran Tech, Digital Sciences & Technologies Department
Rue des Jeunes Bois, Chateaufort, 78114 Magny-Les-Hameaux, France

2- Performance/ Operability, Safran Aircraft Engine
77550 Moissy-Cramayel, France

Abstract. In aircraft engine monitoring, estimating performance indicators from observed measurement data has been an important and long-standing subject, as these indicators provide highly beneficial information to assist maintenance activities. The two main resolution approaches in tackling this problem are Bayesian inferences and machine learning methods, each having its own limitations: inferences are not robust against model-reality gap and non-linearity, while current implementations of machine learning algorithms do not take into account temporal information.

In this work, we focus on a use case in estimating engine performance indicators from snapshot data. We explore several hybrid approaches, aiming to simultaneously leverage the advantages of Bayesian inferences and machine learning methods. We demonstrate that the estimation precision provided by one of our hybrid methods significantly improves upon that of state-of-the-art methods in the tested cases.

1 Introduction

In aeronautics, monitoring engine performance plays a crucial role in the development of predictive maintenance and prognostic activities [1]. In particular, estimating performance indicators from operational data is amongst the most popular subjects of study.

In this work, we focus on a use case of aircraft engines' performance estimation, which revolves around two well-used types of performance indicators: modular *efficiencies* and modular corrected *air mass flow rates* [2]. Instead of directly calculating these indicators , the dominant perspective in the literature is to leverage existing physical forward models. These models are typically thermodynamic simulators that map performance state (i.e., values of efficiencies and flow rates) to selected measurements at certain stable operating conditions. As a consequence, given a real dataset of sensor measurements, estimating performance indicator values is then considered as an inverse problem of the chosen forward model. In the sequel, we refer to this as the **engine performance inverse problem**. A real-life example is provided in Table 1.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Table 1: A performance inverse problem of a turbofan engine (from left to right) and its associated forward model (from right to left). Notation: η = efficiency, Γ = air mass flow rate, Turb = Turbine.

6 measurable quantities		10 performance indicators
Core rotation speed (N2)		η, Γ Fan
Compressor inlet temperature (T25)		η, Γ Booster
Compressor outlet temperature (T3)	$\rightarrow$	η, Γ Compressor
Exhaust gas temperature (EGT)	$\leftarrow$	η, Γ Low-pressure Turb
Compressor outlet pressure (PS3)		η, Γ High-pressure Turb
Combustion chamber fuel flow (WF)		

The two main approaches dominating the state-of-the-art of engine performance inverse problem on in-service data are: *filtering-based methods* [3] and *machine learning (ML) oriented methods* [4, 5]. While filters take into account temporal information and are theoretically optimal in linear cases, they fall short in highly non-linear systems and are non-robust in cases with high model-reality gaps [6, 7]. On the other hand, machine learning techniques, notably neural networks, can be fit to (simulated) dataset generated by thermodynamic simulators to learn the mapping between measurements and performances. Once fully trained, ML-models can quickly estimate performance indicators; however, current implementations only provide point-by-point estimations and only work well if noises in real data are mild enough with respect to the training simulated data. Note importantly that none of the two state-of-the-art approaches have completely resolved the so-called under-determined scenario often encountered in practice, where the number of to-be-estimated indicators is higher than the dimension of observed measurements.

Challenges. In this work, we focus on a use case of monitoring the performance of a turbofan engine, as described in Table 1. This problem is challenging due to its under-determination nature. Even by following recent developments of the state-of-the-art, we find that both Kalman filters and neural networks have not achieved desired performance in this use case. From this starting point, we pose the following research question: *"Can a hybrid usage of Kalman filters and neural networks improve the precision in estimating performance indicators?"*

Contributions. In this work, we investigate several hybrid approaches of Kalman filters (KF) and neural networks (NN) in the engines' performance inverse problem. First, we establish two simulated datasets corresponding to two degradation scenarios: a linear degradation and a non-linear degradation with maintenance recoveries. We then establish three hybrid approaches: *(i)* `NN_UKF` method: using KF as a downstream component to filter out NN's estimations; *(ii)* `UKF_NN` method: using KF as an upstream component to filter out noises in sensors measurements before fitting, as input, to NN models and *(iii)* `UKF_with_NN` method: additionally concatenating NN's estimations as artificial measurements of a KF. We conducted a series of experiments using these

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

three hybrid approaches and compared their performances with the baseline algorithms (`non-hybrid KF` and `non-hybrid NN`). We demonstrate the improvements given by these hybrid methods; in particular, the `UKF_with_NN` method shows a superiority in performance.

Related Works. Besides the two approaches introduced above, the literature of the engine performance inverse problem also includes a traditional method called gas path analysis (GPA) [8, 9] which relies on direct optimization. However, GPA is theoretically impossible in under-determined cases. Moreover, there exist several hybrid approaches. For example, Volponi et al. [10] use NNs to predict residuals between outputs of a given inference and observed measurements (at a pre-determined set of flight conditions); this prediction is used to calibrate the a priori models (that can be used in another inference). Another hybrid usage of NN is proposed by Kobayashi et al. [11] but it is combined with genetic algorithms instead of filters.

2 Problem Statement and Methods

2.1 Problem Statement

We consider an engine system and represent the value of its performance indicators by θ (we also call θ the health state of the engine). Corresponding with θ and an operational condition u (i.e., either Cruise, Take-off or Climbs), the engine gives a measurement y (possibly influenced by unknown noises). We focus on the following problem: given a time series of data points $y_t, t = 1, \ldots, T$ that are measured at some known conditions u, we aim to retrieve all $\theta_t, t = 1, \ldots, T$ corresponding to y_t and u. Importantly, to solve this problem, we assume a thermodynamic simulator S that gives simulated measurements $\tilde{y} = S(\theta, u)$.

In this work, we focus on a use case having data from 4 pre-determined operating conditions. For this reason, in the sequel, we simplify the notation and omit u in all formulas; for example, we write $\tilde{y} = S(\theta)$ and implicitly understand that y being the concatenated outputs at these 4 operating conditions.

2.2 Non-hybrid versus Hybrid Methods

Kalman Filter *(KF).* Intuitively, KF intends to solve inverse problems by combining an a priori model of propagation and the observed measurements. The a priori model predicts measurements with the help of a measurement model (e.g., a simulator), and the difference between such predictions and real measurements allows the filter to correct the current estimation of θ. In this work, we focus on a non-linear version, the Unscented Kalman Filter (UKF) and use it as our first baseline non-hybrid method. The main parameters to be tuned in our usage of UKF are Q—the covariance of the a priori propagation model, and R—the covariance of the measurement noises.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

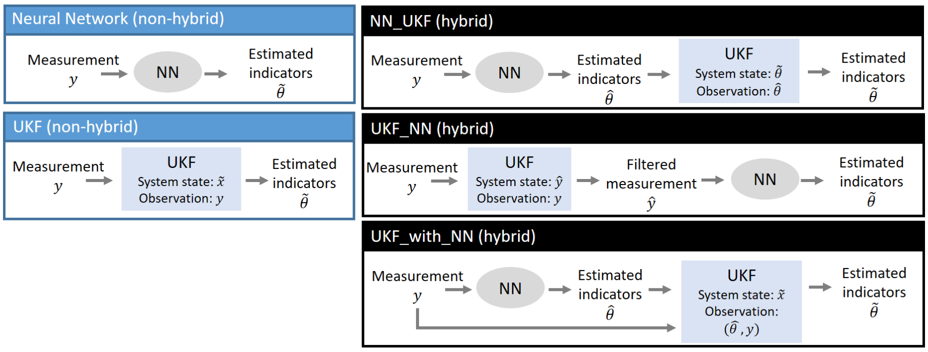

Figure 1: The considered hybrid models in comparison with standard non-hybrid methods.

Neural network *(NN).* Following the typical pipeline of applying NN in the literature of engine performance inverse problems, in this work, we construct the experiments involving NN as follows: we generate a dataset of two millions different pairs (performance indicators, measurements) with a simulator, then train a dense network containing 3 layers of size 100 with mean squared error loss on this dataset to learn the inverse function.

Hybrid methods. In this work, building upon the two aforementioned approaches, we construct three different hybrid methods aiming to obtain a best-of-both-world results. We investigated three approaches: `NN_UKF`, `UKF_NN` and `UKF_with_NN`, as described in Section 1 and Figure 1. Note that technically, one can replace NN by other regression methods (such as random forest or gradient boosting); however, in our tests, we observe that NN outperforms other regressors by a large margin, hence, we only focus on hybrid methods with NN.

3 Experiments and Results

Data generation. We present 2 setups of engine degradation: *(i)* the "linear scenario" where θ_t decreases linearly through time and *(ii)* the "with-maintenance scenario" where maintenance processes cause sudden recoveries in θ_t. In particular, in the first setup, we choose a degradation trajectory, where each performance indicators has a different degradation speed, representing 10000 flights (as a standard life for an aircraft engine). After each 100 flights, we simulate the measurements associated with the corresponding values of performance indicators (and 4 pre-determined operating conditions). We then add a zero-mean Gaussian noise to each measurement representing sensors' noises. Secondly, in the with-maintenance scenario, every 1000 flights, we simulate a maintenance recovery, randomly drawn from a Gaussian distribution around average recoveries observed in practice, while the degradation between these maintenance times

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Metric	(Trajectory Error, Slope Error)									
Setup	Linear					With Maintenance				
Rank	1	2	3	4	5	1	2	3	4	5
UKF_with_NN	(9,4)	(6,8)	(0,3)	(0,0)	(0,0)	(8,7)	(6,6)	(0,2)	(0,0)	(0,0)
UKF	(6,6)	(9,4)	(0,5)	(0,0)	(0,0)	(6,5)	(8,7)	(0,2)	(0,1)	(0,0)
NN_UKF	(0,5)	(0,3)	(15,7)	(0,0)	(0,0)	(1,1)	(0,1)	(14,10)	(0,3)	(0,0)
NN	(0,0)	(0,0)	(0,0)	(15,0)	(0,15)	(0,2)	(1,0)	(0,1)	(14,0)	(0,12)
UKF_NN	(0,0)	(0,0)	(0,0)	(0,15)	(15,0)	(0,0)	(0,0)	(0,1)	(0,11)	(15,3)

Table 2: Couplings showing ranks, by trajectory and slope errors, of each method in the considered setups. E.g., in linear scenarios, UKF_with_NN ranks first (having smallest errors) 9 times with trajectory errors and 4 times with slope errors.

is also linear. For each of these setups, we generate 5 datasets at random.

Evaluation methodology. Dealing with time series estimations, we choose to work with two metrics : *(i)* the *trajectory error* which accumulates the absolute error between true state and estimation, and *(ii)* the *slope error* which accumulates the differences between estimated slopes and the true slopes with a sliding window of fixed sizes. We fix a specific value of Q (an identity matrix multiplied by a coefficient of $10^{-7}, 10^{-8}$ or 10^{-9}) for UKF as a benchmark then compare with other hybrid methods with the same Q.[1] We then report each algorithm's performance in the 15 cases (5 datasets and 3 values of Q for each dataset).

Moreover, NN_UKF and UKF_with_NN require an assessment on uncertainty for NN's estimations. This uncertainty is dictated by a covariance matrix called R_{NN}. We observe that while the performance of hybrid methods varies with R_{NN}, there is a large interval where it works well (hence, we do not specify the value of R_{NN} here).

Results. In Table 2, we summarize the results of our experiments by ranking the algorithms' performances in the considered 15 tested cases. Overall, we observe that UKF_with_NN outperforms non-hybrid approaches, with interesting gains in both considered setups (with and without maintenance). This is also demonstrated in Figure 2. Second, NN_UKF is not competitive enough to surpass non-hybrid UKF, but achieves an improvement in comparison to the tested non-hybrid NN model. Third, UKF_NN fails to surpass non-hybrid UKF but outperforms the non-hybrid NN method in the slope error metric.

Despite its inferior performance, NN_UKF should not be dismissed right away since its performance might be boosted by a more robust predictor and more scenarios should be investigated where UKF is less favored (such as severely non-linear cases). The under-performance of UKF_NN might correspond to the phenomenon observed in [5]: adding noises might improve performances of the neural networks; however, the level of added noises should match that of measurements.

[1] While the assumption of identity matrix in building Q is rather limited, it is enough for the purpose of comparing between different methods in our tests without enforcing too much (unrealistic) a priori knowledge.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

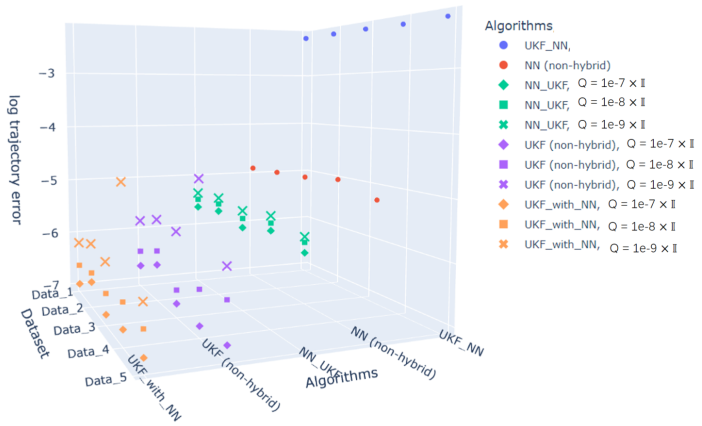

Figure 2: The trajectory error (in logarithmic scale) of considered algorithms in the with-maintenance scenario showing `UKF_with_NN` is the best algorithm most of the time (comparing between algorithms using the same value of matrices Q).

References

[1] A.D. Fentaye, A.T. Baheta, S.I. Gilani, and K.G. Kyprianidis. A review on gas turbine gas-path diagnostics: State-of-the-art methods, challenges and opportunities. *Aerospace*, 6(7):83, 2019.

[2] J.P. Holman. *Thermodynamics*. McGraw-Hill, 1988.

[3] A.J. Volponi, H. DePold, R. Ganguli, and C. Daguang. The use of Kalman filter and neural network methodologies in gas turbine performance diagnostics: a comparative study. *J. Eng. Gas Turbines Power*, 125(4):917–924, 2003.

[4] I.G. Castillo, I. Loboda, and J.L. Pérez Ruiz. Data-driven models for gas turbine online diagnosis. *Machines*, 9(12):372, 2021.

[5] D.Q. Vu, S. Razakarivony, S. Thepaut, G. Doquet, Y. Marnissi, and M. Nocture. Aircraft engines performances estimation from multi-point and multi-time operational data via neural networks. In *2024 IEEE Conference on Artificial Intelligence (CAI)*. IEEE, 2024.

[6] F. Lu, J. Huang, C. Ji, D. Zhang, and H. Jiao. Gas path on-line fault diagnostics using a nonlinear integrated model for gas turbine engines. *International Journal of Turbo & Jet-Engines*, 31(3):261–275, 2014.

[7] F. Lu, T. Gao, J. Huang, and X. Qiu. Nonlinear Kalman filters for aircraft engine gas path health estimation with measurement uncertainty. *Aerospace Science and Technology*, 76:126–140, 2018.

[8] L.A. Urban. *Gas turbine engine parameter interrelationships*. Hamilton Standard Division of United Aircraft Corporation, 1969.

[9] P.C. Escher. Pythia: An object-orientated gas path analysis computer program for general applications. 1995.

[10] A.J. Volponi and R. Rajamani. Hybrid models for engine health management. *Machine Learning and Knowledge Discovery for Engineering Systems Health Management*, pages 395–422, 2012.

[11] T. Kobayashi and D.L. Simon. Hybrid neural-network genetic-algorithm technique for aircraft engine performance diagnostics. *Journal of Propulsion and Power*, 21(4):751–758, 2005.