CEAS-GNC-2022-049

CEAS EuroGNC 2022
"Conference on Guidance, Navigation and Control"
3-5 May 2022 @ Technische Universität Berlin, Germany

A new hybrid model approach coupling a physical model and an artificial neural network through joint estimation

Franz A. R. Enkelmann Research Associate, Technical University of Darmstadt, Flight Systems and Automatic Control, 64287, Darmstadt, Germany. enkelmann@fsr.tu-darmstadt.de

Saleh H. Krüger Research Associate, Technical University of Darmstadt, Flight Systems and Automatic Control, 64287, Darmstadt, Germany. krueger@fsr.tu-darmstadt.de

ABSTRACT

The hybrid model approach presented in this paper is characterized by coupling a physical model and an artificial neural network, which are identified through joint estimation. By using the physical parameters as the interface between the physical model and the artificial neural network, the hybrid model structure combines both approaches directly. Joint estimation using a modified iterated Unscented Kalman Filter (UKF) ensures parallel updating of the dynamic, parameter and artificial neural network weight states. This approach represents an innovation in the context of the literature, in which hybrid model approaches are often cascaded, with separate identification.

When using the hybrid model approach for complex modeling problems, especially for aerospace applications, high nonlinearity, demanding requirements for a robust filter and stability issues can occur. To handle nonlinearity the UKF is chosen. To achieve the required robustness and stability, a modification is derived that separates noisy state and covariance estimation. Testing the hybrid model approach with the modified UKF in a simulation environment on a simplified oscillating problem with time-variant parameters shows convincing results. Both the time-variant and constant parameters can be estimated and predicted with sufficient accuracy. The stability is confirmed by the use of the newly introduced Unscented Wiener Filter.

Keywords: hybrid model, coupling, physical model, artificial neural network, joint estimation, unscented kalman filter

Nomenclature

n	=	number of states
m	=	number of measurements
$\boldsymbol{u}(t)$	=	system input, $[u \times 1]$
$\boldsymbol{m}(t)$	=	sensor measurement, $[m \times 1]$
$\boldsymbol{x}$	=	state vector, $[n \times 1]$
$\boldsymbol{y}$	=	system output, $[m \times 1]$
$\boldsymbol{e}$	=	error between measurement and system output, $[m \times 1]$
$f(\boldsymbol{x}, \boldsymbol{u})$	=	state function model for filter update, $[n \times 1]$
$h(\boldsymbol{x})$	=	measurement function model for residual generation, $[m \times 1]$

$\widehat{\boldsymbol{P}}_{yy}$	=	estimated filter output covariance, $[m \times m]$
$\widehat{\boldsymbol{P}}_{ee}$	=	residual covariance, $[m \times m]$
$\widehat{\boldsymbol{P}}_{xx}$	=	estimated state covariance, $[n \times n]$
$\widehat{\boldsymbol{P}}_{xy}$	=	estimated model covariance, $[n \times m]$
$\widehat{\boldsymbol{Q}}_{xx} \equiv \boldsymbol{S}$	=	system noise covariance, $[n \times n]$
$\widehat{\boldsymbol{R}}_{yy}$	=	measurement noise covariance, $[m \times m]$
$\boldsymbol{V}_{k\|k-1}$	=	minimum variance and time variant filter gain, $[n \times m]$
$\boldsymbol{V}_{c}$	=	constant filter gain, $[n \times m]$

Indices/Accents:

$\widehat{\ }$	=	estimated parameter
$*$	=	corrected parameter
$_{init\ or\ 0}$	=	initial parameter value
$_{c}$	=	constant value
$_{v}$	=	adaptive value
$_{j}$	=	iteration step, zero means no iteration
$_{k\|k-1}$	=	value at time step k based on value at time step k-1.
$_{k\|k}$	=	updated value at time step k
$_{k+1\|k}$	=	propagated value for time step k+1 at time step k
$_{d}$	=	dynamic
$_{p}$	=	parameter
$_{nn}$	=	(artificial) neural network

1 Introduction

Hybrid models describe the beneficial combination of physical models and machine learning approaches. Different structures and categories of hybrid models are already discussed in literature [1,2]. Their use increases the need of domain knowledge and data, but also the accuracy and the scope of the model [3]. On this way machine learned models may gain access to the reliability of physical models. Examining the existing approaches of hybrid modeling this is the standard way, extending machine learning approaches with the inclusion of physical based math or replacing the deterministic standard with structural variants, such as Artificial Neural Networks (ANN) [4,5]. Several attempts in combination of ANN and physics are given by A. Karpatne with the Physical Guided Neural Network (PGNN), which is discussed in several publications [6–8]. Again, the consideration of existing knowledge in the training process rules the data-only training procedure in terms of accuracy and reliability. Another branch of research is being explored with the explainability of Artificial Intelligence (XAI), which is often built with deep neural networks [9–11]. In this case an increase of the model's reliability and understanding is aimed, which may automatically be granted by the consideration of a physical model.

But why not thinking upside down? Consequently, this paper proposes the idea of using the suitable a priori knowledge, the physical model, and its extension via data-based approaches, using ANN. The goal stays the same: model a complex technical system, such as a manned or unmanned aerial vehicle, as accurately as possible in a reliable and robust manner, but by exploiting the complete a priori knowledge at acceptable costs. Grey areas, which are not or not sufficiently described within the physical model, are to be filled by using an ANN, whereby any additional data mapping environmental disturbances, operational conditions, degradation states and maintenance measures may be considered.

A major challenge is seen in the identification of the proposed hybrid model, which is solved by joint estimation using an iterated Unscented Kalman Filter (iUKF). The iUKF extends the Unscented Kalman

Filter (UKF) with an additional iteration[1] that can compensate for the loss of accuracy due to minor modifications discussed later. Both filter exaggerate dissimilar filter approaches like the Extended Kalman Filter (EKF) in cases of high non-linearities [12,13]. In addition, the UKF has been sufficiently established in various estimation challenges, such as state and parameter estimation as well as for training ANN [14–17].

At first the new hybrid model approach and the hybrid model´s structure is introduced in section 2. By using joint estimation to identify the prior introduced hybrid model the UKF and iUKF are employed, whose algorithm is described in the appendix. Subsequently, some flaws of the UKF are worked out and minor modifications separate state and covariance calculation in section 3. Under theoretical analysis the modifications can potentially increase robustness and stability making them candidates for implementation in the preliminary study in section 4. Using a simple simulation of a mass-spring-damper model with time-variant parameters, the performance and robustness of the chosen filter methods and the hybrid model approach are discussed. Later, the results thus found are examined in terms of stability and reliability, introducing the Unscented Wiener Filter (UWF)[2], which is also part of the algorithm given in the appendix. This procedure also presents a novel approach for stability testing of filter results according to the literature. Finally, section 5 concludes a good performance and convincing stability of the hybrid model approach and provides an outlook for further research.

2 A New Hybrid Model Approach

This paper proposes a new hybrid model approach and structure, Fig. 1. The focus lies on the extension of a physical model using an ANN. The physical parameters form the interface between the two modeling approaches. It is assumed, that these parameters are not constant, as typically specified in the physical model, but exhibit dependencies. Just as the dependencies are unknown to the user the unknown parameter behavior should be described by an ANN. In order to characterize the physical parameters, incomplete physical models can be extended and additional effects can be considered using sensor, input and additional data to train the ANN. While the accuracy of the model should be increased, the reliability of the model must be maintained. In addition, the ANN provides knowledge about the characteristic behavior of the physical parameters as a function of the input variables. Constraints[3] for the physical parameters can guarantee physical meaningful behavior of the model.

In context to the literature, where hybrid models describe any combination of experienced-based, data-driven and physics-based models, the proposed hybrid model combines the data-driven and physics-based approach [2]. Generally covered by prognostics technology, these models are often used for estimating the remaining useful lifetime [2]. In contrast to this, the described approach remains a state estimator, which can include dynamic states, physical parameters, and neural network weights, depending on the user´s specifications. Gained knowledge about the system behavior can potentially be used for further examination.

[1] Within this paper only one additional iteration is suggested.

[2] Nota bene, the iterated UWF (iUWF) is also a meaningful derivative, but is not pointed out here.

[3] Constraints for the physical parameters imply the use of Kalman Filtering with state or equality constraints, which are discussed in [18–20] but not considered further in this paper.

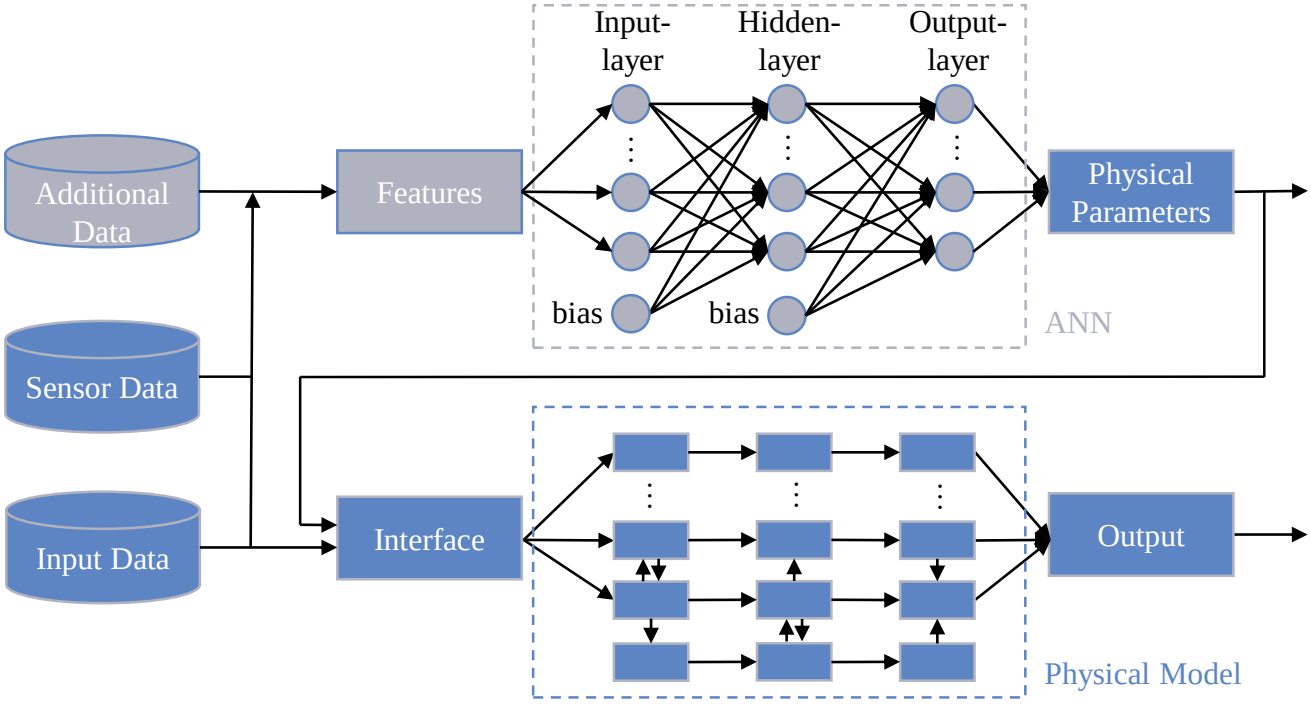

Fig. 1: Qualitative illustration of the hybrid model structure

To address the different states in the state propagation the system equation is extended:

$$\hat{\boldsymbol{x}}_{k+1|k} = f_k\left(\hat{\boldsymbol{x}}_{k|k}\right) = \begin{bmatrix} \hat{\boldsymbol{x}}_{d_{k+1|k}} \\ \hat{\boldsymbol{x}}_{p_{k+1|k}} \\ \hat{\boldsymbol{x}}_{nn_{k+1|k}} \end{bmatrix} = \begin{bmatrix} f_d\left(\hat{\boldsymbol{x}}_{d_{k|k}}, \hat{\boldsymbol{x}}_{p_{k|k}}, \boldsymbol{u}(k)\right) \\ f_{nn}\left(\hat{\boldsymbol{x}}_{p_{k|k}}, \hat{\boldsymbol{x}}_{nn_{k|k}}, \boldsymbol{u}_{nn}(k)\right) \\ \hat{\boldsymbol{x}}_{nn_{k|k}} \end{bmatrix}. \tag{1}$$

The extended system equation f_k involves the known system equation f_d from the physical model to propagate the dynamic states $\hat{\boldsymbol{x}}_{d_{k+1|k}}$, which now depend on the prior dynamic states $\hat{\boldsymbol{x}}_{d_{k|k}}$, the estimated parameters $\hat{\boldsymbol{x}}_{p_{k|k}}$ and the input , $\boldsymbol{u}(k)$ at timestep k. The ANN, given with f_{nn}, propagates the parameters $\hat{\boldsymbol{x}}_{p_{k+1|k}}$ depending on the weights of the ANN $\hat{\boldsymbol{x}}_{nn_{k|k}}$ and the input $\boldsymbol{u}_{nn}(k)$, which can be specified and extended by additional input data to train the ANN. Previously estimated parameters $\hat{\boldsymbol{x}}_{p_{k|k}}$ can be used by the ANN in addition. The weights of the ANN are hold constant with $\hat{\boldsymbol{x}}_{nn_{k+1|k}} = \hat{\boldsymbol{x}}_{nn_{k|k}}$. Note that physical parameters may also be constants, additional states can be defined to cover sensor errors and the input $\boldsymbol{u}_{nn}(k)$ can include $\boldsymbol{u}(k)$ as well as the sensor measurements $\boldsymbol{m}(k)$.

A good introduction and overview about structures of ANN are presented in [21,22] with applications in the aviation sector [4,5,23] and under the use of EKF and UKF [14–17,24]. Possible settings about the number of hidden layers, nodes and activation functions are found in the literature and are not described further here. The setup used to test the approach and implementation is explained in section 4.

Considering the extended system equation **(1)**, the joint estimation can be initiated. Joint estimation has a decisive advantage over the methods commonly used in the literature. The physical model and the ANN are jointly updated, at each time step. This ensures a close coupling of the two modeling approaches. Often discussed under the topic of dual estimation, separating the state calculation, joint and dual estimation are described in [25–28].

The new hybrid model structure with identification through joint estimation provides the basis of physical reliability with the potential to increase accuracy and provide insight into system behavior by

examining the parameter dependencies thus found. To maintain these promising attributes the iUKF is considered including modifications discussed in section 3.

3 A modification of the Unscented Kalman Filter to identify the proposed hybrid model

The UKF shelters a different perspective on the calculation of the system covariance $\widehat{\boldsymbol{P}}_{xx}$ and performing its propagation than the more widely used EKF. Once introduced as "A New Approach for Filtering Nonlinear Systems", the UKF established as one state of the art Kalman Filter (KF) for problems of high nonlinearities today [29]. Exploiting the idea of the Unscented Transformation (UT) any functional dependency can be used to map a random variable on a new set. The representation of these variables, typically Gaussian Random Variables (GRV) in dynamic systems, is done with formally determined Sigma Points (SP).

> These sample points completely capture the true mean and covariance of the GRV, and when propagated through the true nonlinear system, capture the posterior mean and covariance accurately to the second order (Taylor series expansion) for any nonlinearity. [24]

But choosing the spread of those SP and their direct dependency on the noisy state value covers issues in forms of reliability, robustness and stability. Indeed, the robustness has already become an object of discussion in literature, where quite complex solutions have been worked out – increasing the robustness at the expanse of the computational effort [30–32]. As shown in the following, implementing a robust solution, that raises expectations for reliability and stability, does not necessarily require complication; Quite the opposite, simplification leads to the goal.

The gain in reliability and stability is of great interest to the hybrid model approach, as the coupling of a physical model and an ANN promises high non-linearities, as well as filter tuning and initialization challenges to configure a stable filter. The implemented algorithm including the UKF, iUKF and UWF is given in the appendix. In contrast to the literature [33,34], the algorithm applied in this paper handles the measurement update first and the state prediction afterwards. The modification concerns about the covariance calculation $\widehat{\boldsymbol{P}}^{j}{}_{xy_{k|k-1}}$, $\widehat{\boldsymbol{P}}^{j}{}_{yy_{k|k-1}}$ and $\widehat{\boldsymbol{P}}_{xx_{k|k}}$, which include the noisy state $\widehat{\boldsymbol{x}}^{j}{}_{k|k-1}$ as the mean value, see equations **(19)**, **(20)** and **(31)**. This means a direct dependence between the state and covariance calculation, which can lead to instability.

As already discovered in [33] the state calculation does not need to be done with the weighted sum necessarily. Contradictorily, an optimal estimate is expected at optimal gain, but one's own estimate is then no longer trusted. We now trust our estimation and assume that the transformed center SP, the mean value of the state, equals the UT prediction of the measurement. Equations **(17)** and **(30)** can be written as:

$$\widehat{\boldsymbol{y}}^{j}{}_{k|k-1} = h_k\left(\widehat{\boldsymbol{x}}^{j}{}_{k|k-1}\right) \tag{2}$$

and

$$\widehat{\boldsymbol{x}}_{k+1|k} = f_k\left(\widehat{\boldsymbol{x}}_{k|k}\right). \tag{3}$$

This also simplifies the definition of the SP as they are only and separately used for the covariance calculation later on. In Equations **(15)** and **(28)**, the first entries are eliminated:

$$\widehat{\boldsymbol{X}}^{j}{}_{k|k-1} = \left[\widehat{\boldsymbol{x}}^{j}{}_{k|k-1} + \sqrt{(L+\lambda)\widehat{\boldsymbol{P}}_{xx\,k|k-1}} \qquad \widehat{\boldsymbol{x}}^{j}{}_{k|k-1} - \sqrt{(L+\lambda)\widehat{\boldsymbol{P}}_{xx\,k|k-1}} \right], \tag{4}$$

$$\widehat{\boldsymbol{X}}_{k|k} = \left[\widehat{\boldsymbol{x}}_{k|k} + \sqrt{(L+\lambda)\widehat{\boldsymbol{P}}_{xx\,k|k}} \qquad \widehat{\boldsymbol{x}}_{k|k} - \sqrt{(L+\lambda)\widehat{\boldsymbol{P}}_{xx\,k|k}} \right]. \tag{5}$$

The elimination is legitimate as the covariance calculation uses delta values where the mean value is subtracted. Starting from $i = 1$ now, equations **(19)**, **(20)** and **(31)** are independent from the weight $w_0^{(c)}$, defined in equation **(10)**. While the definition of $w_0^{(m)}$ and $w_i^{(m)}$ as well as the parameter β are obsolete, equation **(11)** remains as the only necessary weight equation to define $w_i^{(c)}$. A closer look on the $\widehat{\boldsymbol{P}}^{*}{}_{xx\,k+1|k}$ calculation in equation **(31)** shows the effect of the simplification on the covariance calculation in general. The delta SP for

$$\Delta\widehat{\boldsymbol{X}}_{k+1|k} = \left(\widehat{\boldsymbol{X}}_{k+1|k} - \widehat{\boldsymbol{x}}_{k+1|k}\right) = \left[f_k\left(\sqrt{(L+\lambda)\widehat{\boldsymbol{P}}_{xx\,k|k}}\right) \qquad - f_k\left(\sqrt{(L+\lambda)\widehat{\boldsymbol{P}}_{xx\,k|k}}\right) \right] \tag{6}$$

means a direct dependency of the propagated system covariance to the measurement-updated covariance, system-equations, weights, number of states and the spread of the SP, excluding the system state:

$$\widehat{\boldsymbol{P}}^{*}{}_{xx\,k+1|k} \sim \widehat{\boldsymbol{P}}_{xx\,k|k}, f_k, w_i^{(c)}, L, \lambda. \tag{7}$$

This separates the covariance and state calculation sufficiently how it is known from different filters, such as the EKF. To discover the gained stability of the modification and to examine its behavior in terms of accuracy an application is necessary. In the event that the theoretical based assumptions are confirmed, the modifications are of course also suitable for any other nonlinear application that require a high degree of filter stability. In this paper, the modifications found are used in the context of the identification process of the hybrid model approach and its application in the next section; further research discussing the applicability of the modification using other filter problems is needed.

4 Preliminary study in a simulation environment and results

This section follows a preliminary study to prove the prototypical implementation of the concept using a simple and reproduceable example. The example comprises an oscillating technical system describing a mass-spring-damper model whose equations are given in the appendix. Notably about the design used in this paper, is the introduction of time-variant parameters. Therefore, a degradation of the spring linear over time is proposed, which is replaced in definite intervals. The effect is implemented by characterizing a sawtooth trend of the Eigenfrequency, which is set as the fifth state of the model.

The parameter values are given with $K = 1$, $D = 0.3$ and the Eigenfrequency, which is defined as

$$\omega_0 = \omega_{0_{init}} - \dot{\omega} \cdot t + \dot{\omega} \cdot t_{recovery} \cdot u_{maintenance}, \qquad \dot{\omega} = \frac{\omega_{0_{init}} - \omega_{0_{deg}}}{t_{max}}. \tag{8}$$

The Eigenfrequency $\omega_{0_{init}} = 2\frac{rad}{s}$ degrades over time and a value of $\omega_{0_{deg}} = 1\frac{rad}{s}$ might be reached after the observed maximum time $t_{max} = 120\ s$ defined in the simulation. A full recovery in fixed intervals is added, which takes place every $t_{recovery} = 25\ s$. The maintenance procedure is recorded with $u_{maintenance}$ being incremented by 1 every 25 seconds. Finally, the input of the ANN includes the time t and $u_{maintenance}$, which are both normalized by t_{max} and $u_{maintenance_{max}} = 4$. With a defined sampling frequency of 100 Hz the time step is $0.01\ s$. A white noise with the standard deviation of $r = 0.1$ is added on the output to create the measurement m.

Initialized with $\hat{\boldsymbol{x}}_{1_{init}} = 0$ and $\hat{\boldsymbol{x}}_{2_{init}} = 0$ the starting values of the dynamic states are defined. The parameter states are assumed with a deviation of 20 %: $\hat{\boldsymbol{x}}_{3_{init}} = K' = 1.2$, $\hat{\boldsymbol{x}}_{4_{init}} = D' = 0.36$ and $\hat{\boldsymbol{x}}_{5_{init}} = {\omega_0}' = 2.4\frac{rad}{s}$. These values are also used for the last three weights of the ANN weighting the bias node from the hidden layer. The other weights are said to be small with randomly generated numbers[4] multiplied with 10^{-3}. This means that the ANN is initialized with an estimate of parameter values that are initially assumed to be constant over time. The ANN is kept simple with one hidden layer keeping three nodes and additional bias nodes in the input and hidden layer. A linear relation is set to be the activation function for both the hidden layer and output layer nodes.

The data set is split in two halves, the training ($t_{training} = 60\ s$) and the validation data ($t_{validation} = 60\ s$). The validation data helps to assess the quality of the derived model as well as for tuning initial filter parameters[5]. These filter parameters are given in the appendix. The training data is used to identify the model including the dynamic states, parameters and ANN weights.

Table 1: Training, comparison of unscented filter variants using training data and the RMSE

	Standard UKF	Modified UKF	Modified iUKF	UWF
$RMSE(\hat{y}, y_{ideal})$	0.0204	0.0180	0.0228	0.0099
$RMSE(\hat{x}_1, x_{1_{ideal}})$	0.0042	0.0082	0.0263	0.0039
$RMSE(\hat{x}_2, x_{2_{ideal}})$	0.0107	0.0156	0.0367	0.0081
$RMSE(\hat{x}_3, x_{3_{ideal}})$	0.0190	0.0265	0.0629	0.0038
$RMSE(\hat{x}_4, x_{4_{ideal}})$	0.0266	0.0361	0.0646	0.0197
$RMSE(\hat{x}_5, x_{5_{ideal}})$	0.0469	0.0553	0.1352	0.0183

Table 2: Validation, comparison of unscented filter variants using validation data and the RMSE

	Standard UKF	Modified UKF	Modified iUKF	UWF
$RMSE(\hat{y}, y_{ideal})$	0.0038	0.0088	0.0039	0.0040
$RMSE(\hat{x}_1, x_{1_{ideal}})$	0.0010	0.0036	0.0012	0.0016
$RMSE(\hat{x}_2, x_{2_{ideal}})$	0.0012	0.0046	0.0028	0.0031
$RMSE(\hat{x}_3, x_{3_{ideal}})$	0.0055	0.0046	0.0033	0.0016
$RMSE(\hat{x}_4, x_{4_{ideal}})$	0.0029	0.0096	0.0077	0.0081
$RMSE(\hat{x}_5, x_{5_{ideal}})$	0.0056	0.0219	0.0056	0.0084

The implemented filters are feasible to identify the hybrid model structure. The chosen error metric, the Root Mean Squared Error (RMSE), records low values for the implemented filters. This indicates a

[4] In this work, we used small randomly generated numbers to initialize the ANN weights. Since this initialization can affect the deterministic behavior of the filter, constant values might be a better choice for initialization, which will be considered in future work.

[5] Solving a linear problem with the presented unscented filters lead to equal results. The implementation of the ANN triggers a necessary fine-tuning to handle convergence issues. Further research is needed and will be part of future publications.

good performance of the hybrid model approach compared to RMSE values of $RMSE(\hat{y}, y_{ideal})_{\omega_0=2} = 0.1470$ and $RMSE\left(\hat{x}_5, x_{5_{ideal}}\right)_{\omega_0=2} = 0.1166 \frac{rad}{s}$ for the output and the fifth state of the test data set, which are obtained at constant parameters $\hat{K} = 1$, $\hat{D} = 0.3$ and $\hat{\omega}_0 = 2\frac{rad}{s}$. The values of the error metrics are reduced in the order of 10^2. The hybrid model approach is also convincing in the identification of the parameter progression since the found solution of the ANN does not manipulate K and D to minimize the variance, but finds the correct behavior of ω_0 as a function of the input of the ANN.

Contrary to expectations the standard UKF has no stability issues to deal with the hybrid model approach. But the requirements for stability and robustness may increase with the complexity of the model, which argues for a rasion d'être for the given modifications of the UKF in future. In the following, the focus is on the evaluation of the filter performance and the final investigation of the stability of the iUKF solution found for the hybrid model. The results of the standard UKF now help to rank the results of the modified UKF variants. The RMSE is used to compare the filter performances. The error values concerning the training data are listed in Table 1 and results about the validation data is given in

Table 2. The Standard UKF performs best and achieves the lowest $RMSE_y$ value when using the validation data. The modified UKF loses accuracy to a tolerable extent. This loss can be compensated using a modified iUKF with just one additional iteration. The modified iUKF takes the most time to converge as the training data records the highest error value and with respect to the iUKF behavior shown in Fig. 3.

The reliability and accuracy can also be assessed with the given RMSE of the dynamic and parameter states. The modified iUKF stands out due to the highest error values in the training process and the lowest values in predicting the validation data under use of the identified model. Comparable results are delivered by the standard UKF, which finds the best solutions in this test for the states $\hat{x}_1$, $\hat{x}_2$ and $\hat{x}_4$ in the validation, but is defeated in other cases.

As the modified iUKF fulfills reliability in theory and performs with a convincing accuracy especially in finding the parameter courses, the following investigations are done, keeping the eye on the results of the modified iterated filter. Fig. 2 visualizes the convergence of the 21 ANN weights using the Frobenius Norm (FN). It becomes clear, that the weights significantly change in the first 10 seconds of training time. Smaller adjustments follow for the input layer weights with a small peak after 25 seconds. During the training progresses, negligible changes in the weights of the hidden layer are observed, but there is a drift in the weights of the input layer, which comes to a halt just before the end of the training. From this point a stable finding of the ANN weights are assumed and an increase of the training data provided could be considered.

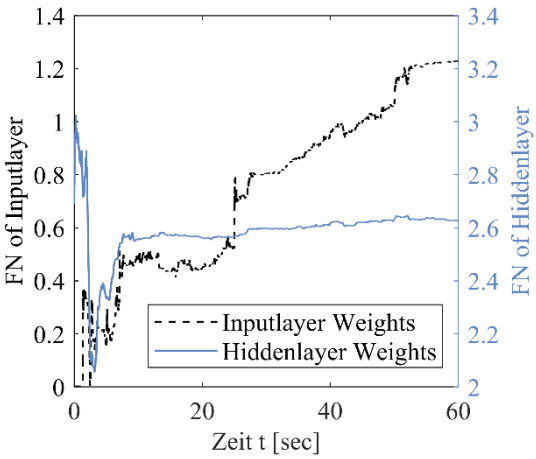

Fig. 2: Frobenius Norm of the ANN weights

The training behavior can be confirmed with the visualization of the output y and the first five states of the studied problem including the ideal and estimated values as well as the residuals, Fig. 3. In the first 10 seconds, great disturbances of both output and states can be seen in the residuals, which afterwards

converge close to zero. Comparison of the estimated curves with the ideal courses provides a better understanding of the numerically interpreted states. The offset in the beginning of the estimation of K, D and ω_0 hardly reduces and the parameters are finally identified after about 30 seconds of training.

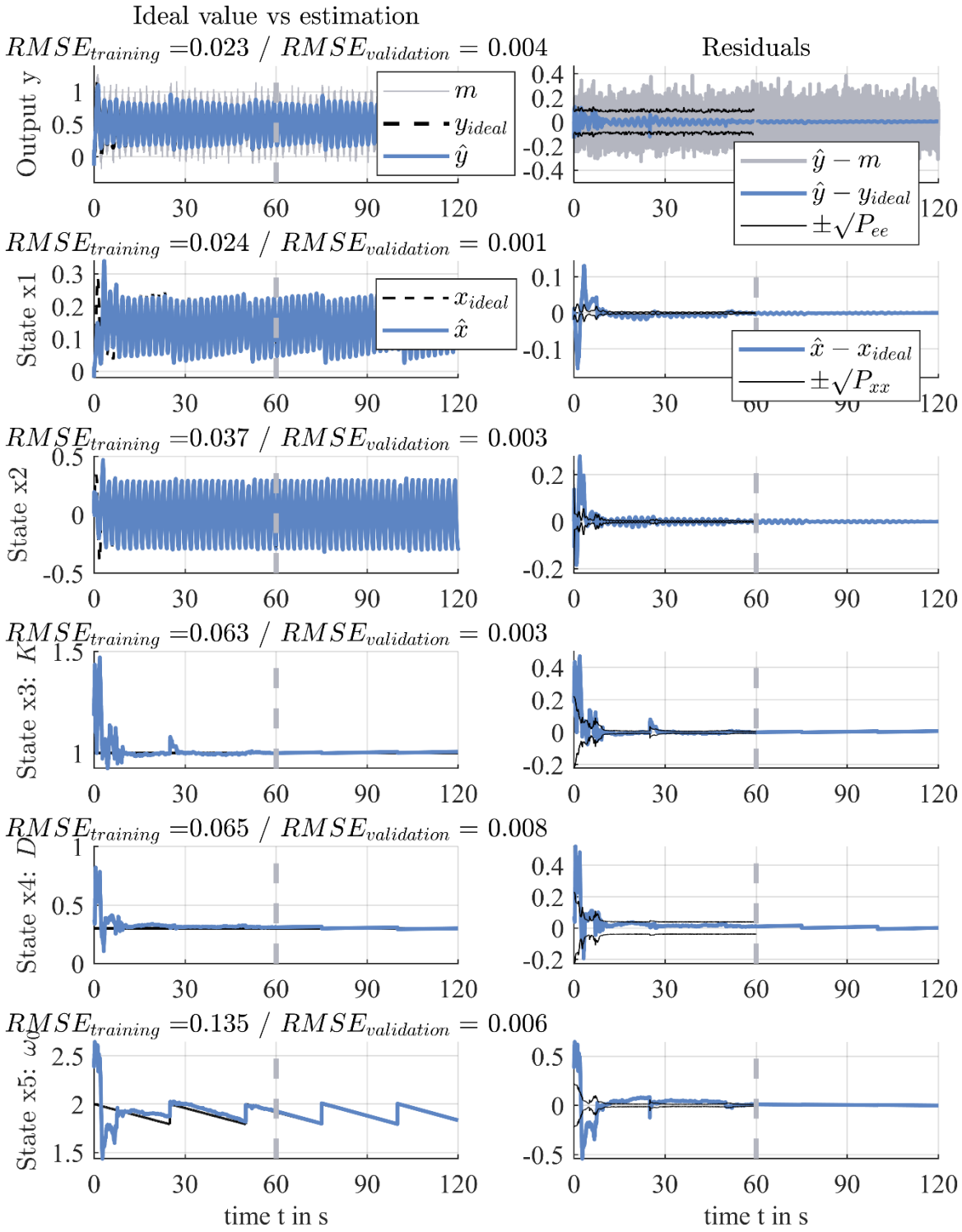

Fig. 3: Comparison of output and states using the certifiable modified iUKF, Training from 0-60s and Validation from 60-120s

Even the sawtooth trend of ω_0 is fully captured within the validation dataset, where the derived model predicts the future parameter behavior. Having a closer look on the presented curves of the square root covariance matrices, some potential in the initiation of the filter algorithm can be concluded. The matrices have to adopt higher values in the beginning in order to capture the associated courses, what is controllable

via the initial $\widehat{\boldsymbol{P}}_{\boldsymbol{xx}}$, $\widehat{\boldsymbol{Q}}^{v}_{xx}$ and $\widehat{\boldsymbol{R}}^{v}_{yy}$ and via the hyper parameters to adapt Q and R with CCC[6]. An additional perspective is granted by a zoom on the first 30 and 10 seconds of the of the training, which is available in the appendix.

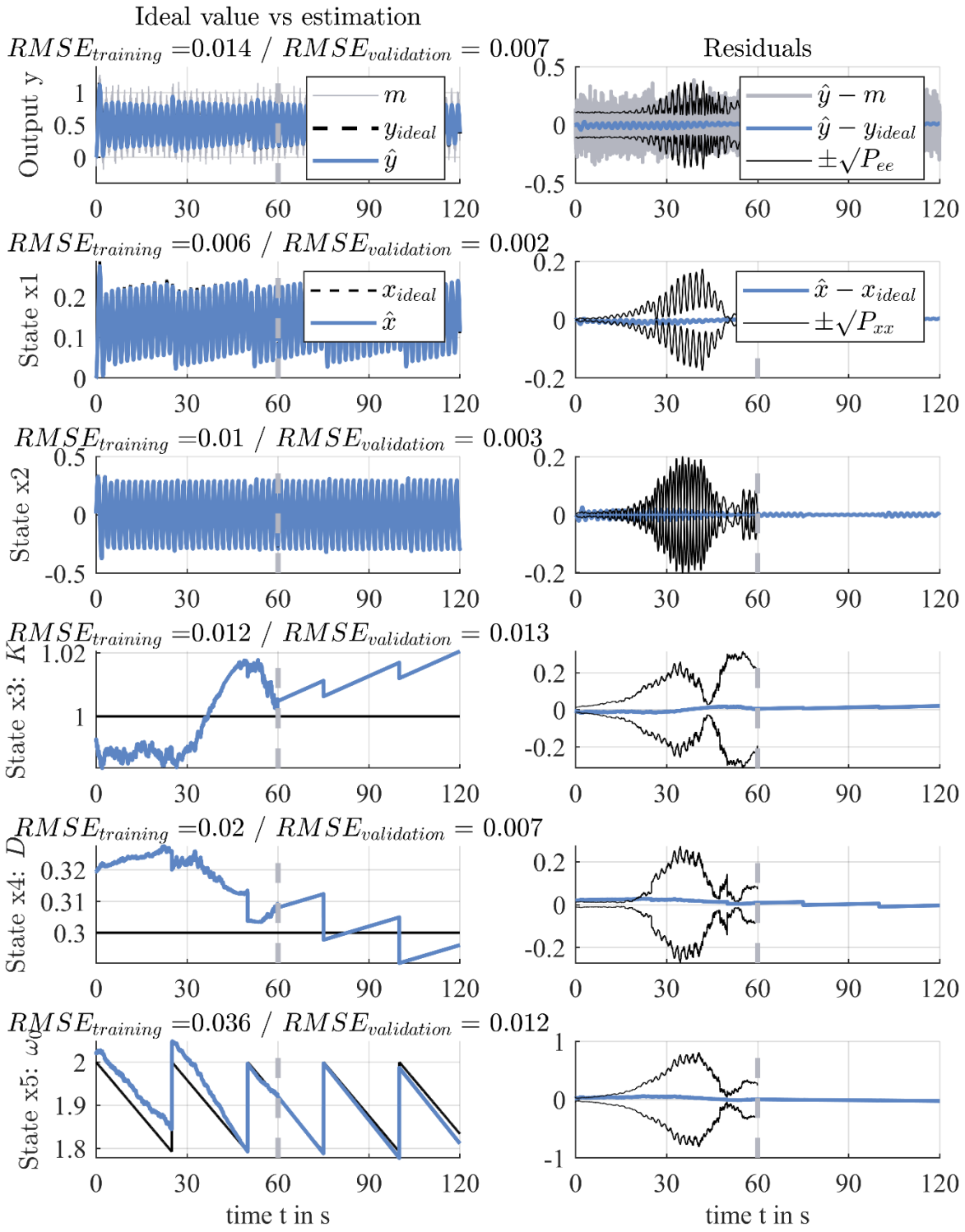

Fig. 4: Stability proof of output and states using the UWF

[6] The adaptive adjustment of the covariances has turned out to be particularly import for the hybrid model approach. Since we control both the state covariance Q and measurement covariance R over the initial value and the values of the last time step, we call it Controlled Combined Covariance (CCC). We do this with a key difference from the literature: The adaptive adjustments are bounded by the initial values of Q and R and cannot diverge as can occur with the techniques reported in the literature, see equations in the appendix.

Finally, a proof of stability is required, where the UWF is utilized and the found solution by the iUKF is investigated, see Fig. 4. The Kalman gain achieved in the last time step using the modified iUKF[7] is used to equip the UWF. As the UWF is forced to handle high non-linearities, which may overcome the filter's abilities, the stable state vector and the state covariance matrix from the last time step of the modified iUKF are also used for the UWF initialization in order to provide an adequate reference condition. With the given initialization of the UWF, the output and states in the training and validation data are accurately estimated.

As the residuals are initially small, the scale of the diagrams decreases significantly, which allows a closer look on the behavior of the states. While the sawtooth trend of ω_0 is still preserved, some deviations are visible for K and D. Somehow, the maintenance data contained in u_{nn} have a negative effect on the constants, which are also trained by the ANN. The covariances show a sharp increase after about 30 seconds, which then decays. Again, the filter tuning parameters can be the trigger. Initializing with the stable state vector, the state covariance matrix and the constant Kalman gain set to the Kalman gain from the last timestep of the iUKF, the UWF provides excellent results from the beginning and confirms the stability of the previously found solution by the iUKF.

5 Conclusion and Outlook

The developed hybrid modeling approach using a modified UKF in a joint estimation represents an innovation in context to the literature. Modifications of the UKF show the potential to increase the stability and robustness with a little loss of accuracy and less complexity. The developed approach has the potential to identify complex model structures and their extension with an ANN. Additional data may be concerned and unknown effects can be mapped. A first application on a simple example has shown a high functionality identifying an oscillator including a time-variant parameter in the presence of noise. The time-variant parameter course is unknown to the user and not covered by the physical model. The behavior of the parameters, both constant and variant can be learned and predicted successfully by the ANN.

This makes the derived approach a perfect complement to accelerate research on hybrid unmanned aerial vehicles, their identification and energy efficiency assessment under consideration of degradation and maintenance procedures. Furthermore, an extension of physical models used in commercial aviation and the application of the hybrid modeling approach is planned to allow a more accurate fuel efficiency assessment, as already postulated in [35].

The development and application of the presented method has raised further questions, naturally. Recursive estimation methods such as the Kalman Filter have compelling properties in dealing with noisy sensor measurements in combination with ANN. The ability to adjust the confidence level with respect to the sensor measurements allows a direct response to overfitting problems. It is obvious, that the ANN of the hybrid approach can be built on these noisy sensor measurements in future, which can lead to biased estimates. An alternative approach for dealing with noisy sensor measurements and bias has already been developed, which includes recursive minimum variance (RMV) [36]. RMV algorithms replace noisy sensor measurements with model estimates and provide unbiased parameter estimates. The Implementation of this approach can potentially increase the accuracy and reliability of the new hybrid model and ANNs considering filter methods, which will be the content of future research.

[7] The Kalman gain had to be adjusted by shifting the gain of the 24th state towards zero, an absolute decrease of 70 % was necessary.

Acknowledgments

Prof. Dr.-Ing. Jürgen Beyer for sharing his experience in working with filters and his guiding influence in the development of this paper, Prof. Dr.-Ing. Uwe Klingauf for the fruitful conversations and Dr.-Ing. Martin Nowara for ideas of filter testing on simplified problems.

References

[1] Kim, N.-H., An, D., and Choi, J.-H., *Prognostics and Health Management of Engineering Systems*, Springer International Publishing, Cham, 1 Jan. 2017.

[2] Liao, L., and Kottig, F., Review of Hybrid Prognostics Approaches for Remaining Useful Life Prediction of Engineered Systems, and an Application to Battery Life Prediction. *IEEE Transactions on Reliability*, Vol. 63, No. 1, 1 Jan. 2014, pp. 191–207. doi: 10.1109/TR.2014.2299152.

[3] Guo, J., Li, Z., and Li, M., A Review on Prognostics Methods for Engineering Systems. *IEEE Transactions on Reliability*, Vol. 69, No. 3, 1 Jan. 2020, pp. 1110–1129. doi: 10.1109/TR.2019.2957965.

[4] Chao, M. A., Kulkarni, C., Goebel, K., and Fink, O., Fusing Physics-based and Deep Learning Models for Prognostics., 2 Mar. 2020.

[5] Chao, M. A., Kulkarni, C., Goebel, K., and Fink, O., Hybrid deep fault detection and isolation: Combining deep neural networks and system performance models., 1 Jan. 2019.

[6] Muralidhar, N., Bu, J., Cao, Z., He, L., Ramakrishnan, N., et al., PhyNet: Physics Guided Neural Networks for Particle Drag Force Prediction in Assembly. *Proceedings of the 2020 SIAM International Conference on Data Mining*, edited by C. Demeniconi and N. Chawla, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1 Jan. 2020, pp. 559–567.

[7] Jia, X., Willard, J., Karpatne, A., Read, J., Zwart, J., et al., Physics Guided RNNs for Modeling Dynamical Systems: A Case Study in Simulating Lake Temperature Profiles. *Proceedings of the 2019 SIAM International Conference on Data Mining*, edited by T. Berger-Wolf and N. Chawla, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1 Jan. 2019, pp. 558–566.

[8] Daw, A., Karpatne, A., Watkins, W., Read, J., and Kumar, V., Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling., 31 Oct. 2017.

[9] Carvalho, D. V., Pereira, E. M., and Cardoso, J. S., Machine Learning Interpretability: A Survey on Methods and Metrics. *Electronics*, Vol. 8, No. 8, 1 Jan. 2019, p. 832. doi: 10.3390/electronics8080832.

[10] Das, A., and Rad, P., Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey., 16 Jun. 2020.

[11] Hong, C. W., Lee, C., Lee, K., Ko, M.-S., Kim, D. E., et al., Remaining Useful Life Prognosis for Turbofan Engine Using Explainable Deep Neural Networks with Dimensionality Reduction. *Sensors (Basel, Switzerland)*, Vol. 20, No. 22, 1 Jan. 2020. doi: 10.3390/s20226626.

[12] Rudolph Van Der Merwe, and Eric Wan, Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models. *In Proceedings of the Workshop on Advances in Machine Learning*, 1 Jan. 2003.

[13] Simon, D., *Optimal State Estimation. Kalman, H [infinity] and nonlinear approaches*, John Wiley & Sons, Inc, Hoboken, NJ, USA, 1 Jan. 2006, 526.

[14] Lima, D. P. de, Neural Network Training Using Unscented and Extended Kalman Filter. *Robotics & Automation Engineering Journal*, Vol. 1, No. 4, 1 Jan. 2017. doi: 10.19080/RAEJ.2017.01.555568.

[15] Vellayikot, S., and Vaidyan, M. V., ANN Approach for State Estimation of Hybrid Systems and Its Experimental Validation. *Mathematical Problems in Engineering*, Vol. 2015, 1 Jan. 2015, pp. 1–13. doi: 10.1155/2015/382324.
[16] Wu, X., and Wang, Y., Extended and Unscented Kalman filtering based feedforward neural networks for time series prediction. *Applied Mathematical Modelling*, Vol. 36, No. 3, 1 Jan. 2012, pp. 1123–1131. doi: 10.1016/j.apm.2011.07.052.
[17] Mauri Aparecido de Oliveira, and Escola Paulista de Política, An Application of Neural Networks Trained with Kalman Filter Variants (EKF and UKF) to Heteroscedastic Time Series Forecasting., 1 Jan. 2012.
[18] Simon, D., Kalman filtering with state constraints: a survey of linear and nonlinear algorithms. *IET Control Theory & Applications*, Vol. 4, No. 8, 1 Jan. 2010, pp. 1303–1318. doi: 10.1049/iet-cta.2009.0032.
[19] Julier, S. J., and LaViola, J. J., On Kalman Filtering With Nonlinear Equality Constraints. *IEEE Transactions on Signal Processing*, Vol. 55, No. 6, 1 Jan. 2007, pp. 2774–2784. doi: 10.1109/TSP.2007.893949.
[20] Simon, D., and Simon, D. L., Aircraft Turbofan Engine Health Estimation Using Constrained Kalman Filtering. *Journal of Engineering for Gas Turbines and Power*, Vol. 127, No. 2, 1 Jan. 2005, pp. 323–328. doi: 10.1115/1.1789153.
[21] Rashid, T., *Make Your own neural network*, CreateSpace Independent Publishing Platform, s.l., 1 Jan. 2016, 222.
[22] Rojas, R., *Neural Networks. A Systematic Introduction*, 1st ed., Springer Berlin Heidelberg, Berlin/Heidelberg, 1 Jan. 1996, 510.
[23] Baumann, S., and Klingauf, U., Modeling of aircraft fuel consumption using machine learning algorithms. *CEAS Aeronautical Journal*, Vol. 11, No. 1, 1 Jan. 2020, pp. 277–287. doi: 10.1007/s13272-019-00422-0.
[24] Haykin, S., *Kalman Filtering and Neural Networks*, John Wiley & Sons Inc, Hoboken, 1 Jan. 2004, 302.
[25] Nelson, A., Nonlinear estimation and modeling of noisy time-series by Dual Kalman filtering methods., 1 Jan. 2000.
[26] Wan, E., Merwe, R., and Nelson, A., Dual Estimation and the Unscented Transformation. *Advances in Neural Information Processing Systems*, Vol. 12, 1 Jan. 2000.
[27] Wan, E., and Nelson, A., Dual Kalman Filtering Methods for Nonlinear Prediction, Smoothing, and Estimation. *Advances in Neural Information Processing Systems*, 1 Jan. 2000.
[28] Zhuang, H., Lu, J., and Li, J., Joint estimation of state and parameter with maximum likelihood method. *2017 36th Chinese Control Conference (CCC)*, IEEE, 1 Jan. 2017, pp. 5276–5281.
[29] Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F., A new approach for filtering nonlinear systems. *Proceedings of 1995 American Control Conference - ACC'95*, 1 Jan. 1995, 1628-1632 vol.3.
[30] Zheng, B., Fu, P., Li, B., and Yuan, X., A Robust Adaptive Unscented Kalman Filter for Nonlinear Estimation with Uncertain Noise Covariance. *Sensors (Basel, Switzerland)*, Vol. 18, No. 3, 1 Jan. 2018. doi: 10.3390/s18030808.
[31] Milschewski, T., and Bariant, J.-F., A numerically stable formulation of the square root unscented Kalman filter for state estimation. *2017 20th International Conference on Information Fusion (Fusion)*, 1 Jan. 2017, pp. 1–7.
[32] Qi, J., Sun, K., Wang, J., and Liu, H., Dynamic State Estimation for Multi-Machine Power System by Unscented Kalman Filter With Enhanced Numerical Stability. *IEEE Transactions on Smart Grid*, Vol. 9, No. 2, 1 Jan. 2018, pp. 1184–1196. doi: 10.1109/TSG.2016.2580584.
[33] Skoglund, M., Gustafsson, F., and Hendeby, G., On Iterative Unscented Kalman Filter using Optimization. *22th International Conference on Information Fusion (FUSION)*, 1 Jan. 2019.
[34] Zhan, R., and Wan, J., Iterated Unscented Kalman Filter for Passive Target Tracking. *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 43, No. 3, 1 Jan. 2007, pp. 1155–1163. doi: 10.1109/TAES.2007.4383605.

[35] Enkelmann, F., Heigl, R., and Pfingsten, K. C., Comparison of a physical model and a machine learning approach for a more accurate assessment of fuel efficiency measures. *AIAA Scitech 2020 Forum*, American Institute of Aeronautics and Astronautics, Reston, Virginia, 1 Jan. 2020.
[36] Beyer, J., and Klingauf, U., Rekursives Verfahren zur biasfreien Parameterschätzung mit minimaler Schätzfehlerkovarianz (RMV). *at - Automatisierungstechnik*, Vol. 42, 1-12, 1 Jan. 1994, pp. 346–355. doi: 10.1524/auto.1994.42.112.346.
[37] Akhlaghi, S., Zhou, N., and Huang, Z., Adaptive adjustment of noise covariance in Kalman filter for dynamic state estimation. *2017 IEEE Power & Energy Society General Meeting*, IEEE, 1 Jan. 2017, pp. 1–5.
[38] Herbert Robbins, and Sutton Monro, A Stochastic Approximation Method. *The Annals of Mathematical Statistics*, Vol. 22, No. 3, 1 Jan. 1951, pp. 400–407. doi: 10.1214/aoms/1177729586.
[39] Belanger, P. R., Estimation of Noise Covariance Matrices for a Linear Time-Varying Stochastic Process. *IFAC Proceedings Volumes*, Vol. 5, No. 1, 1 Jan. 1972, pp. 265–271. doi: 10.1016/S1474-6670(17)68341-1.
[40] Mehra, R., Approaches to adaptive filtering. *IEEE Transactions on Automatic Control*, Vol. 17, No. 5, 1 Jan. 1972, pp. 693–698. doi: 10.1109/TAC.1972.1100100.
[41] Akhlaghi, S., Zhou, N., and Huang, Z., Adaptive adjustment of noise covariance in Kalman filter for dynamic state estimation. *2017 IEEE Power & Energy Society General Meeting*, IEEE, 72017, pp. 1–5.
[42] Haykin, S. S. (ed.), *Kalman filtering and neural networks*, Wiley, New York, NY, Weinheim, 1 Jan. 2001.

Appendix

UKF, iUKF and UWF algorithm

The definition of the weighting and the scaling parameters to calculate the sigma points and their spread around the system state $\widehat{\boldsymbol{x}}_{k|k-1}$ is initially given in the first step of the algorithm. The spread is defined by $\lambda = \alpha^2(L + \kappa) - L$, with the number of states L, a small positive value of $\alpha \in (0, 1]$ (usually set to $\alpha = 10^{-3}$) and the secondary parameters κ and β. These are set to $\kappa = 0$ or for parameter estimation to $\kappa = 3 - L$ [24] and to $\beta = 2$ for Gaussian distributions [12]. The weights used for the weighted approximation of the sample mean index (m) or the covariance (c) are given with:

$$w_0^{(m)} = \frac{\lambda}{L + \lambda}\,, \tag{9}$$

$$w_0^{(c)} = \frac{\lambda}{L + \lambda} + 1 - \alpha^2 + \beta, \tag{10}$$

$$\boldsymbol{w}_i^{(m)} = \boldsymbol{w}_i^{(c)} = \frac{\lambda}{2(L + \lambda)}, i = 1, \dots, 2L. \tag{11}$$

The input $\mathbf{u}_k$ is considered in the beginning to define the system and the measurement function (f_k, h_k) at time step k. The UKF (no iteration means $j = 0$) including the iUKF and the UWF is presented in the following scheme:

0) Input from last time step or initialization: $\widehat{\boldsymbol{x}}_{k|k-1}, \widehat{\boldsymbol{P}}_{xx\,k|k-1}$ and optional: $\widehat{\boldsymbol{Q}}^{v}_{xx_{k-1}}, \widehat{\boldsymbol{R}}^{v}_{yy_{k-1}}$ in case of adaptive covariance adjustment

$$\widehat{\boldsymbol{x}}^{0}{}_{k|k-1} = \widehat{\boldsymbol{x}}_{k|k-1}, \tag{12}$$

$$f_k(\widehat{\boldsymbol{x}}_k) = f(\widehat{\boldsymbol{x}}_k, \boldsymbol{u}_k) \tag{13}$$

$$h_k(\hat{\boldsymbol{x}}_k) = h(\hat{\boldsymbol{x}}_k, \boldsymbol{u}_k) \tag{14}$$

1) Residual and covariance calculation at time k:

Calculation of sigma points:

$$\hat{\boldsymbol{X}}^j{}_{k|k-1} = \left[\hat{\boldsymbol{x}}^j{}_{k|k-1} \qquad \hat{\boldsymbol{x}}^j{}_{k|k-1} + \sqrt{(L+\lambda)\hat{\boldsymbol{P}}_{xx_{k|k-1}}} \qquad \hat{\boldsymbol{x}}^j{}_{k|k-1} - \sqrt{(L+\lambda)\hat{\boldsymbol{P}}_{xx_{k|k-1}}}\right], \tag{15}$$

$$\hat{\boldsymbol{Y}}^j{}_{k|k-1} = h_k\left(\hat{\boldsymbol{X}}^j{}_{k|k-1}\right), \tag{16}$$

$$\hat{\boldsymbol{y}}^j{}_{k|k-1} = \sum_{i=0}^{2L} \boldsymbol{w}_i^{(m)} \hat{\boldsymbol{Y}}^j{}_{k|k-1}, \tag{17}$$

$$\boldsymbol{e}^j_{k|k-1} = \boldsymbol{m}(k) - \hat{\boldsymbol{y}}^j{}_{k|k-1}. \tag{18}$$

2) Matrix $\hat{\boldsymbol{P}}_{xx_{k|k-1}}$ and output equation $h_k\left(\hat{\boldsymbol{x}}^j{}_{\mathrm{k|k-1}}\right)$ provide:

$$\hat{\boldsymbol{P}}^j{}_{xy_{k|k-1}} = \sum_{i=0}^{2L} w_i^{(c)} \left(\hat{\boldsymbol{X}}^j{}_{k|k-1} - \hat{\boldsymbol{x}}^j{}_{k|k-1}\right)\left(\hat{\boldsymbol{Y}}^j{}_{k|k-1} - \hat{\boldsymbol{y}}^j{}_{k|k-1}\right)^T, \tag{19}$$

$$\hat{\boldsymbol{P}}^j{}_{yy_{k|k-1}} = \sum_{i=0}^{2L} w_i^{(c)} \left(\hat{\boldsymbol{Y}}^j{}_{k|k-1} - \hat{\boldsymbol{y}}^j{}_{k|k-1}\right)\left(\hat{\boldsymbol{Y}}^j{}_{k|k-1} - \hat{\boldsymbol{y}}^j{}_{k|k-1}\right)^T. \tag{20}$$

3) Optimal filter gain calculation:

$$\hat{\boldsymbol{P}}^j{}_{ee_{k|k-1}} = \hat{\boldsymbol{P}}^j{}_{yy_{k|k-1}} + \hat{\boldsymbol{R}}^v{}_{yy_{k-1}}, \tag{21}$$

$$\boldsymbol{V}^j{}_{k|k-1} = \hat{\boldsymbol{P}}^j{}_{xy_{k|k-1}} \cdot \hat{\boldsymbol{P}}^j{}^{-1}_{ee_{k|k-1}}. \tag{22}$$

4) Apply the optimal correction:

$$\hat{\boldsymbol{H}}^j{}_{k|k-1} = \hat{\boldsymbol{P}}^j{}^{T}_{xy_{k|k-1}} \cdot \hat{\boldsymbol{P}}^j{}^{-1}_{xx_{k|k-1}}, \tag{23}$$

$$\hat{\boldsymbol{x}}^{j+1}{}_{k|k-1} = \hat{\boldsymbol{x}}^0{}_{k|k-1} + \boldsymbol{V}^j{}_{k|k-1} \cdot \left(\boldsymbol{e}^j_{k|k-1} - \hat{\boldsymbol{H}}^j{}_{k|k-1} \cdot \left(\hat{\boldsymbol{x}}^0{}_{k|k-1} - \hat{\boldsymbol{x}}^j{}_{k|k-1}\right)\right), \tag{24}$$

Note: the standard covariance correction is calculated in the last iteration step:

$$\hat{\boldsymbol{P}}_{xx_{k|k}} = \hat{\boldsymbol{P}}_{xx_{k|k-1}} - \boldsymbol{V}^j{}_{k|k-1} \cdot \hat{\boldsymbol{P}}^j{}^{T}_{xy_{k|k-1}} \tag{25}$$

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}^{j+1}{}_{k|k-1} \tag{26}$$

Note: the standard covariance correction using any fixed filter gain $\boldsymbol{V}_c$ is:

$$\hat{\boldsymbol{P}}_{xx_{k|k}} = \hat{\boldsymbol{P}}_{\boldsymbol{xx}_{k|k-1}} - \boldsymbol{V}_c \cdot \hat{\boldsymbol{P}}^{T}_{xy_{k|k-1}} - \hat{\boldsymbol{P}}_{xy_{k|k-1}} \boldsymbol{V}_c^T + \boldsymbol{V}_c \cdot \hat{\boldsymbol{P}}_{ee_{k|k-1}} \cdot \boldsymbol{V}_c^T \tag{27}$$

5) In case of an iterative approach:

Go to step 1)

6) Execute propagation for time step k+1

Calculate the state vector covariance transformation $\hat{\boldsymbol{P}}^*{}_{xx_{k+1|k}}$ from corrected values $\hat{\boldsymbol{x}}_{k|k}$, $\hat{\boldsymbol{P}}_{xx_{k|k}}$ and state function $f_k\left(\hat{\boldsymbol{x}}_{k|k}\right)$ via new SP:

$$\hat{\boldsymbol{X}}_{k|k} = \left[\hat{\boldsymbol{x}}_{k|k} \qquad \hat{\boldsymbol{x}}_{k|k} + \sqrt{(L+\lambda)\hat{\boldsymbol{P}}_{xx_{k|k}}} \qquad \hat{\boldsymbol{x}}_{k|k-1} - \sqrt{(L+\lambda)\hat{\boldsymbol{P}}_{xx_{k|k}}}\right] \tag{28}$$

$$\widehat{\boldsymbol{X}}_{k+1|k} = f_k\left(\widehat{\boldsymbol{X}}_{k|k}\right) \tag{29}$$

$$\widehat{\boldsymbol{x}}_{k+1|k} = \sum_{i=0}^{2L} \boldsymbol{w}_i^{(m)} \widehat{\boldsymbol{X}}_{k+1|k} \tag{30}$$

$$\widehat{\boldsymbol{P}}^*{}_{xx_{k+1|k}} = \sum_{i=0}^{2L} w_i^{(c)} \left(\widehat{\boldsymbol{X}}_{k+1|k} - \widehat{\boldsymbol{x}}_{k+1|k}\right)\left(\widehat{\boldsymbol{X}}_{k+1|k} - \widehat{\boldsymbol{x}}_{k+1|k}\right)^T \tag{31}$$

$$\widehat{\boldsymbol{P}}_{xx_{k+1|k}} = \widehat{\boldsymbol{P}}^*{}_{xx_{k+1|k}} + \widehat{\boldsymbol{Q}}^v_{xx_{k-1}} \tag{32}$$

7) In case of adaptive $\widehat{\boldsymbol{Q}}^v_{xx_{k-1}}, \widehat{\boldsymbol{R}}^v_{yy_{k-1}}$ calculate the update $\widehat{\boldsymbol{Q}}^v_{xx_k}, \widehat{\boldsymbol{R}}^v_{yy_k}$
8) Wait for next sampling step k = k + 1

In case of the iUKF: Step 5) is performed and a new set of SP, based on the measurement-updated state $\widehat{\boldsymbol{x}}^{j+1}{}_{k|k-1}$, is built in the beginning cumulating the iteration variable j by one with every iteration. The matrix square root $\sqrt{(L+\lambda)\widehat{\boldsymbol{P}}_{xx_{k|k-1}}}$ is calculated by the lower-triangular Cholesky factorization in this paper [33]. Equation **(23)** has been derived to implement the additional term in the optimal state correction in equation **(24)** known from the iEKF [33]. The UWF skips the optimal correction in 4) and uses the overall covariance correction given in equation **(27)** based on a constant gain V_c set initially.

Adaptive covariance adjustment using Combined Covariance Control (CCC)

In case of using adaptive measurement and system covariance, $\widehat{\boldsymbol{R}}^v{}_{yy_{k-1}}$ and $\widehat{\boldsymbol{Q}}^v{}_{xx_{k-1}}$, step 7) can be executed using different adaptive adjustment procedures. These procedures are described in the literature [37]. Most relevant publications are seen in the years 1951 [38] and 1972 [39,40]. Since divergence is observed in [41] and [42] when using the adaptive approach, instability may occur. Therefore, we use an adaptive approach which is controlled by the initial values $\left(\widehat{\boldsymbol{Q}}_{xx_{init}}, \widehat{\boldsymbol{R}}_{yy_{init}}\right)$:

$$\begin{aligned} \widehat{\boldsymbol{Q}}^v_{xx_{k|k-1}} &= v_q \cdot \boldsymbol{Q}_{xx_{k|k-1}} + \left(1 - v_q\right) \cdot \widehat{\boldsymbol{Q}}_{xx_{init}} \\ \widehat{\boldsymbol{R}}^v_{xx_{k|k-1}} &= v_r \cdot \boldsymbol{R}_{xx_{k|k-1}} + (1 - v_r) \cdot \widehat{\boldsymbol{R}}_{xx_{init}} \end{aligned} \tag{33}$$

The updated covariances $\boldsymbol{Q}_{\boldsymbol{xx_{k|k}}}, \boldsymbol{R}_{\boldsymbol{xx_{k|k}}}$ are mixed with the initial values using the blending factors v_q and $v_r \in \mathbb{R}, [0, 1]$. This ensures a lower bound on the covariances and thus stability according to the initially defined covariances. During initial testing the blending factors of $v_q = [0.8, 0.9]$ and $v_r = [0.1, 0.5]$ lead to good results. The innovations are calculated as:

$$\boldsymbol{Q}_{xx_{k|k-1}} = \left(\boldsymbol{V}_{k|k-1}\hat{\boldsymbol{e}}_{k|k-1}\right) \cdot \left(\hat{\boldsymbol{e}}^T_{k|k-1}\boldsymbol{V}^T_{k|k-1}\right), \tag{34}$$

$$\begin{aligned} \boldsymbol{R}_{xx_{k|k-1}} &= diag\left\{\boldsymbol{P}_{ee_{k|k-1}} - \widehat{\boldsymbol{P}}_{yy_{k|k-1}}\right\} \\ \text{with: } \boldsymbol{P}_{ee_{k|k-1}} &= \hat{\boldsymbol{e}}_{k|k-1} \cdot \hat{\boldsymbol{e}}^T_{k|k-1}, \end{aligned} \tag{35}$$

as described in the literature. The filter gain $\boldsymbol{V}_{k|k-1}$ and the residual $\hat{\boldsymbol{e}}_{k|k-1}$ of the current time step are used to calculate the adapted state covariance. Notice that in case of the adapted $\boldsymbol{R}_{xx_{k|k}}$, which describes the positive definite pseudo measurement noise covariance, the filter model output $\widehat{\boldsymbol{P}}_{yy_{k|k-1}}$ is subtracted from the total residual covariance using the diagonal entries only. This suppresses a supposed cross-

correlation of measurement noise from different sensors. To ensure stability using the presented adaptive covariance adjustments we recommend small values of v_q and v_r as well as the implementation of lowpass filtered covariances:

$$\begin{aligned} \widehat{\boldsymbol{Q}}^{v}_{xx_k} &= \widehat{\boldsymbol{Q}}^{v}_{xx_{k-1}} + w_q \cdot \left(\widehat{\boldsymbol{Q}}^{v}_{xx_{k|k-1}} - \widehat{\boldsymbol{Q}}^{v}_{xx_{k-1}} \right), \\ \widehat{\boldsymbol{R}}^{v}_{xx_k} &= \widehat{\boldsymbol{R}}^{v}_{xx_{k-1}} + w_r \cdot \left(\widehat{\boldsymbol{R}}^{v}_{xx_{k|k-1}} - \widehat{\boldsymbol{R}}^{v}_{xx_{k-1}} \right). \end{aligned} \tag{36}$$

The covariance adjustment is advanced in increments using the constants w_q and $w_r \in \mathbb{R}, [0, 1]$. We recommend values of $w_q = [0.5, 0.9]$ and $w_r = [0, 0.1]$. As we use adaptive covariance methods described in the literature controlled by the initial covariance values and the implementation of a lowpass filter[8] we name the procedure Combined Covariance Control (CCC).

Mass-Spring-Damper model

State equation:

$$\dot{\boldsymbol{x}}(t) = \begin{bmatrix} 0 & 1 \\ -\omega_0 & -2D\omega_0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t). \tag{37}$$

Measurement equation:

$$y(t) = [K\omega_0^2 \quad 0]\, \boldsymbol{x}(t). \tag{38}$$

The measurement noise is added to the ideal measurement $y_{id}(t)$ as gaussian distributed normal numbers with variance $\sigma^2_{noise} = 0.01$:

$$m(t) = y_{id}(t) + \sigma_{noise}. \tag{39}$$

The input signal is a square wave signal with amplitude $A_u = 0.5$, period time $T_u = 2\ s$ ·and constant offset of $K_u = 0.5$:

$$u(t) = A_u \cdot sgn\left(\cos\frac{2\pi}{T_u} \cdot t + \frac{\pi}{2}\right) + K_u. \tag{40}$$

For discretization the forward Euler approximation, i.e. $x_{1_{k+1}} = x_{1_k} + x_{2_k} \cdot T_0$, with a sample time of $T_0 = 0.01\ s$ is used.

[8] Note: For specific high noise signals and systems, this lowpass filter can also be applied separately to improve estimation and reduce noise levels.

Table 3: Initialization values in section 4

	Standard UKF	BS-UKF	BS-iUKF and UWF
$\alpha = 10^{x}$	-2.4151	-3.9263	-2.0825
κ	3.0001	-24.3460	-22.0000
β	2.0036	-	-
v_q	0.8398	0.2875	0.1
w_q	0.4247	0.7524	0.2163
v_r	0.5248	0.3382	0.5246
w_r	0.0314	0.0658	0.0658
$\hat{R}^{v}_{yy_{init}} = 10^{x}$	-0.9590	-1.0776	-1.0776
$\hat{Q}^{v}_{xx_{init}}(1) = 10^{2*x} \cdot 10^{-2}$	-4.0000	-5.8704	-3.7720
$\hat{Q}^{v}_{xx_{init}}(2) = 10^{2*x} \cdot 10^{-2}$	-3.9931	-2.3837	-8.8495
$\hat{Q}^{v}_{xx_{init}}(3) = 10^{2*x} \cdot 10^{-2}$	-7.2280	-7.2307	-6.7098
$\hat{Q}^{v}_{xx_{init}}(4) = 10^{2*x} \cdot 10^{-2}$	-7.4135	-6.7325	-0.3928
$\hat{Q}^{v}_{xx_{init}}(5) = 10^{2*x} \cdot 10^{-2}$	-0.3362	-8.9793	-6.4787
$\hat{Q}^{v}_{xx_{init}}(6-26) = 10^{*x} \cdot 10^{-2}$	-4.7443	-7.1263	-9.4375
$\widehat{\boldsymbol{P}}_{xx_{init}}(1) = 10^{x}$	-1.2950	-6.7504	-0.3160
$\widehat{\boldsymbol{P}}_{xx_{init}}(2) = 10^{x}$	-5.6771	-4.7985	-1.3010
$\widehat{\boldsymbol{P}}_{xx_{init}}(3) = 10^{x}$	-1.7935	-2.2553	-0.5358
$\widehat{\boldsymbol{P}}_{xx_{init}}(4) = 10^{x}$	-7.1054	-7.1025	-3.4683
$\widehat{\boldsymbol{P}}_{xx_{init}}(5) = 10^{x}$	-0.5933	-1.3010	-0.3539
$\widehat{\boldsymbol{P}}_{xx_{init}}(5-26) = 10^{x}$	-0.8928	-0.9006	-1.3010

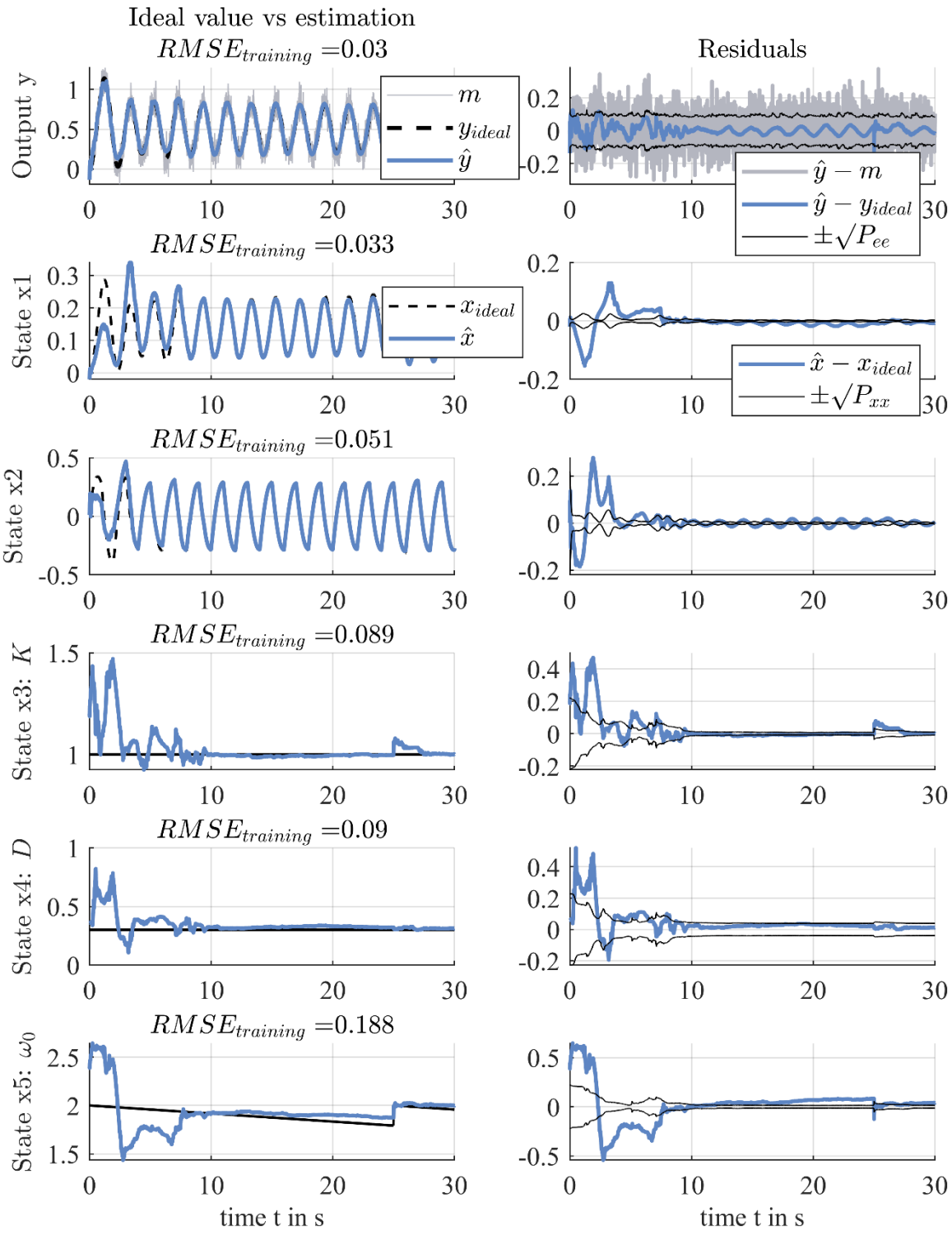

Fig. 5: Comparison of output and states using the certifiable modified iUKF with zoom on the first 30 seconds

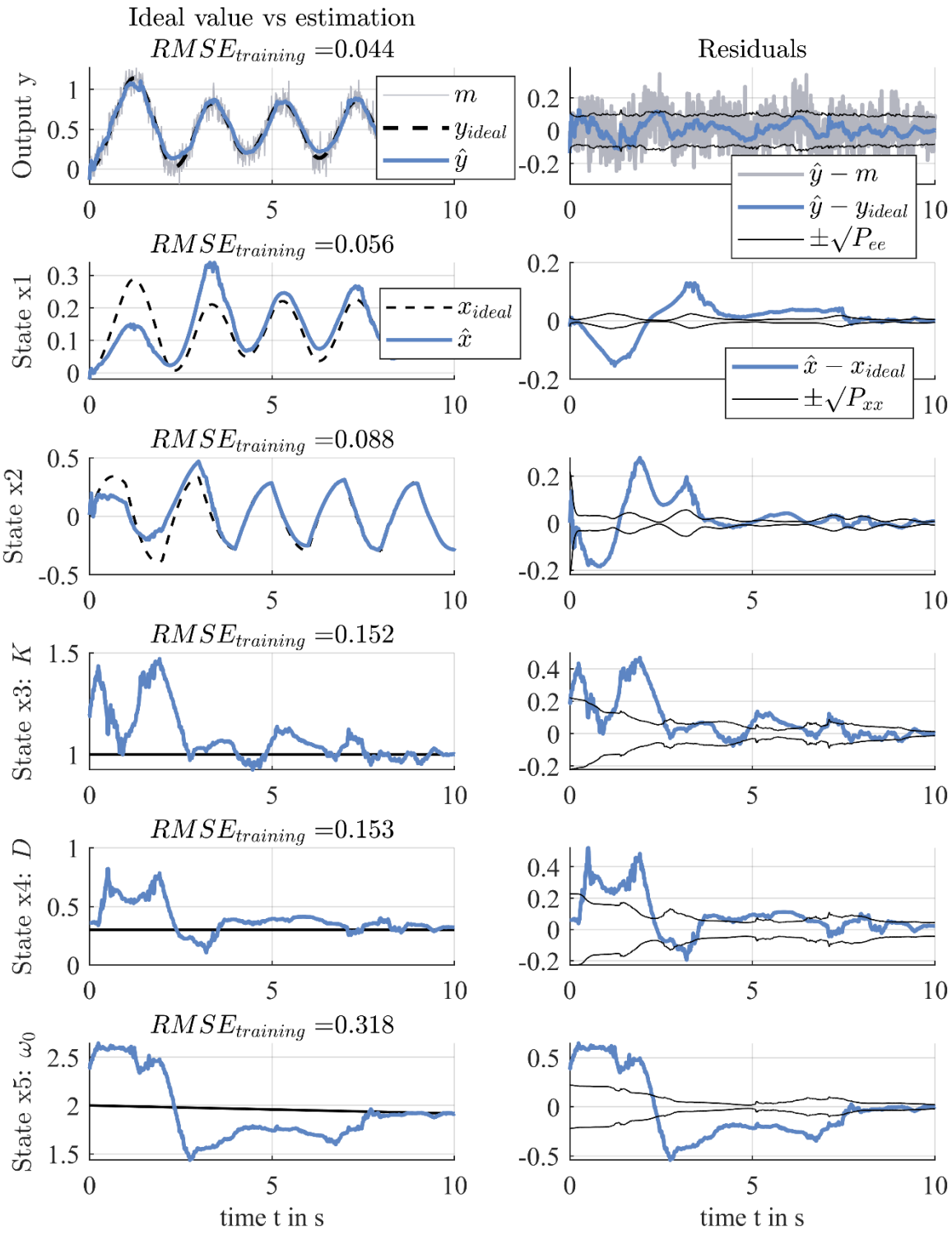

Fig. 6: Comparison of output and states using the certifiable modified iUKF with zoom on the first 10 seconds