

A Hybrid Extended Kalman Filter Based on a Parametrized FeedForward Neural Network for the Improvement of the Results of Numerical Wave Prediction Models [†]

Athanasios Donas ¹, George Galanis ² and Ioannis Th. Famelis ^{2,*} 

¹ microSENSSES Laboratory, Department of Electrical and Electronics Engineering, Ancient Olive Grove Campus, University of West Attica, 12241 Athens, Greece; adonas@uniwa.gr

² Mathematical Modeling and Applications Laboratory, Hellenic Naval Academy, Hatzikiriakion, 18539 Piraeus, Greece; ggalanis@hna.gr

* Correspondence: ifamelis@uniwa.gr

[†] Presented at the 16th International Conference on Meteorology, Climatology and Atmospheric Physics—COMCAP 2023, Athens, Greece, 25–29 September 2023.

Abstract: Aiming to develop a novel optimization model for numerical weather and wave predictions, this study proposes a hybrid approach based on the combination of Artificial Neural Networks (ANNs) and Kalman Filters (KFs). The KF technique uses fixed covariance matrices, which is inappropriate for dynamic conditions where the measurement error varies due to outside factors. This work proposes an alternative method for updating the covariance matrices, depending on a selected parameter that was obtained from a parametrized FeedForward Neural Network. The FeedForward Neural Network is trained for various values of the parameter for a set of historical data to provide the best choice based on a predetermined objective. Following this procedure, the parameter is added to the hybrid Extended Kalman Filter, which improves the direct outputs of numerical forecasts by decoding the systematic error between the model's direct output and the corresponding observations. The suggested approach was used in a number of time periods in various locations with very promising results.

Keywords: optimization of numerical forecasts; extended Kalman filter; feedforward neural networks; WAM-cycle 4



Citation: Donas, A.; Galanis, G.; Famelis, I.T. A Hybrid Extended Kalman Filter Based on a Parametrized FeedForward Neural Network for the Improvement of the Results of Numerical Wave Prediction Models. *Environ. Sci. Proc.* **2023**, *26*, 199. <https://doi.org/10.3390/environsciproc2023026199>

Academic Editors: Konstantinos Moustris and Panagiotis Nastos

Published: 20 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Due to numerous important applications, including search and rescue, shipping and renewable energy sources, the necessity for accurate wave predictions has substantially increased. Numerical wave prediction models, which are employed by many operational and research centers, are a trustworthy and effective instrument to reach these goals, as they are fairly successful in simulating global environmental conditions. Nevertheless, when the forecasting concerns a local area of interest, numerical models fail to provide precise forecasts. The main reason for this is the multiparametric nature of the various factors involved, such as the strong dependence on initial and lateral conditions or the parametrization of some atmospheric or wave procedures [1].

The enhancement of the model's resolution could be a potential solution to these problems, but the success of this strategy is uncertain and at the same time extremely costly, as the computational time will rise dramatically. An alternative approach could be to use post-processing algorithms [1–3] to improve the starting conditions based on assimilation systems. These instruments, like ANNs and Kalman Filters, can be used to improve the output generated by statistical models by reducing the systematic error of the simulation. In order to achieve this, a cost function is defined as the “distance” (bias) between the direct output of the model and the actual observations.

Under this framework, we introduce a hybrid Extended Kalman Filter based on a parametrized FeedForward Neural Network (pFFNN) for the prediction of the Significant Wave Height (SWH). In particular, for a set of historical data, the pFFNN is trained for various values of the parameter and the optimal value is added to the hybrid Extended Kalman Filter, which reduces the systematic error between the direct output of the WAM-cycle 4 [4] model and the corresponding observation.

2. Extended Kalman Filter Algorithm

Kalman Filtering (KF) is a recursive prediction algorithm which provides an optimal estimation for an unknown state vector x_t given information about an observable vector y_t based on the assumption that there is a linear correlation between these variables. However, when the dynamics of the model are nonlinear, a different filtering approach is necessary. One such technique is the Extended Kalman Filter, which is an extension of the Kalman Filter through a linearization process based on Taylor's approximation [5].

To lay the groundwork for the construction of the EKF, we consider a nonlinear system which is described by the following state measurement model:

$$x_t = h_{t-1}(t-1, x_{t-1}) + w_{t-1} \quad (1)$$

$$y_t = g_t(t, x_t) + v_t \quad (2)$$

where w_t and v_t are independent Gaussian noise processes with zero mean and covariance matrices Q_{t-1} and R_t , respectively. The quantities $h_{t-1}(\bullet)$ and $g_t(\bullet)$ are the nonlinear transition matrices functions, which are considered to be C^1 and in some cases are time varying.

The fundamental idea behind the Extended Kalman Filter is to linearize the state space model of Equations (1) and (2) each time around the previous or most recent state estimate, which is taken to be either $\hat{x}_{t-1}$ or $\hat{x}_{t/t-1}$ depending on which specific functional is being taken into consideration [5]. The variables $\hat{x}_{t-1}$ and $\hat{x}_{t/t-1}$ express the priori and the posteriori estimate of the state, respectively.

3. Parametrized FeedForward Neural Network

Artificial Neural Networks (ANNs) are adaptable artificial systems that draw inspiration from the processes of the human brain, and are able to rebuild the fuzzy rules that determine the unknown correlation between various variables. The simplest type of ANN is the FeedForward Neural Network, which is a multi-layer network where every neuron in the previous layer connects to each neuron of the next layer [6].

In this study, a parametrized three-level FeedForward Neural Network has been applied in order to determine the optimal parameter a , with one hidden layer, twenty neurons, the sigmoid function, $f(x) = \frac{1}{1+e^{-x}}$ as the activation function and one linear output layer. In particular, the network's direct output is given by:

$$y_{out} = V * f(W\tilde{x}^T + b_1) + b_2$$

where W is a $d \times m$ matrix, with m being the number of neurons in the hidden layer and d being the size of the input data $\tilde{x} = (x_1, x_2, \dots, x_{d-1}, \alpha)$. V and b_1 are $1 \times m$ matrices, while b_2 is the (scalar) bias of the output layer.

The pFFNN can be trained for a range of values of the parameter α (here $0 \leq \alpha \leq 1$), enhancing network's efficiency. The idea behind the application of a parameterized ANN is based on the work of Famelis and Kaloutsas, in which an NN was trained for different values of a parameter in order to determine to what extent this affects the stiffness of the problem [7]. Furthermore, in order to determine the network's parameters, the Levenberg–Marquardt method was applied as learning a rule.

4. The Proposed Methodology

The Kalman Filtering procedure uses fixed covariance matrices; thus, choosing the quantities Q_{t-1} and R_t is one of the most challenging issues for the filter's successful implementation.

To avoid this problem, Shahrokh Akhlaghi and Zhenyu Huang [8] proposed an estimating method to adaptably adjust the quantities Q and R at each step of the EKF algorithm, based on:

$$R_t = \alpha R_{t-1} + (1 - \alpha) (\varepsilon_t \varepsilon_t^T + G_{t/t-1} P_{t/t-1} G_{t/t-1}^T)$$

and

$$Q_t = \alpha Q_{t-1} + (1 - \alpha) (K_t d_t d_t^T K_t^T)$$

Their choice was $\alpha = 0.3$ for all their tests. Here, the quantity ε_t is the residual and can be determined by the difference between the measurement's actual value and its projected value based on the data provided at step t . On the other hand, d_t is the innovation and expresses the difference between the actual measurement and its predicted value at time t .

However, the choice of the parameter α cannot be independent of the available data, as this loses essential information necessary for the study of dynamical systems. For this reason, aiming to improve the forecasting capabilities of WAM-cycle 4, the proposed methodology implements the parametrized FeedForward Neural Network presented in Section 2 to determine parameter α and therefore to modify the covariance matrices.

More specifically, the available historical data are divided into three sections:

- Data Set1. Train the pFFNN and obtain the parameter a .
- Data Set2. Train the hybrid Extended Kalman Filter with the parameter a .

$$Data\ Set2 = Data\ Set1 + 1 : DataSet1 + range$$

- Data Set3. Forecast using the hybrid Extended Kalman Filter.

$$Data\ Set3 = Data\ Set2 + 1 : DataSet2 + forecast.$$

For all case studies, $DataSet1 = 1000$, $range = 250$ and $forecast = 40$.

Hybrid Extended Kalman Filter

In order to reduce the systematic error of the simulation, we define y_t as the difference between the actual Significant Wave Height and the FeedForward Neural Network's direct output, which uses as input the previous forecast of the model, with the following form:

$$y_t = m_t(t, N_t, SWH_{t-1}) + v_t. \quad (3)$$

where $m(\bullet)$ is the nonlinear transition matrix function, while the quantity N_t express a vectorial parameter that needs to be estimated by the filter. Note here that N_t includes the parameters of each layer of the FFNN, i.e., $N_t = [W, V, b_1, b_2]$ and $m_t(t, N_t, SWH_{t-1}) = V * f(W * SWH_{t-1} + b_1) + b_2$, where $f(\bullet)$ is the sigmoid function.

Furthermore, as we lack accurate knowledge about the evolution of the state vector N_t in time, we presume that its change would be random and thus we set the system's nonlinear transition matrix function $h_{t-1} = 1$. Therefore, the state equation for the EKF can be described as follows:

$$N_t = N_{t-1} + w_{t-1} \quad (4)$$

Combining (3) and (4), we construct the state space model:

$$N_t = N_{t-1} + w_{t-1}$$

$$y_t = m_t(t, N_t, SWH_{t-1}) + v_t$$

When the dynamic system is defined, the hybrid Extended Kalman Filter follows the following steps:

1. Step 1. Initialization. For $t = 0$, set

$$\hat{N}_0 = E[N_0], P_0 = E[(N_0 - E[N_0])(N_0 - E[N_0])^T], R_0, Q_0$$

2. Step 2. Iteration. For $t = 1, 2, \dots$, compute:

- The Jacobian matrix of the nonlinear transition function $m(\bullet)$.

$$F_{t/t-1} = \nabla m_{t/t-1} | \hat{N}_{t/t-1}$$

- The Prediction Step.

$$\hat{N}_{t/t-1} = \hat{N}_{t-1}$$

$$P_{t/t-1} = P_{t-1} + Q_{t-1}$$

$$d_t = y_t - m_t(\hat{N}_{t/t-1})$$

$$K_t = P_{t/t-1} F_{t/t-1}^T [F_{t/t-1} P_{t/t-1} F_{t/t-1}^T + R_t]^{-1}$$

$$\hat{N}_{t/t} = \hat{N}_{t/t-1} + K_t [y_t - m_t(\hat{N}_{t/t-1})]$$

$$\varepsilon_t = y_t - m_t(\hat{N}_{t/t})$$

$$R_t = \alpha R_{t-1} + (1 - \alpha) (\varepsilon_t \varepsilon_t^T + F_{t/t-1} P_{t/t-1} F_{t/t-1}^T)$$

$$K_t = P_{t/t-1} F_{t/t-1}^T [F_{t/t-1} P_{t/t-1} F_{t/t-1}^T + R_t]^{-1}$$

$$P_{t/t} = [I - K_t F_{t/t-1}] P_{t/t-1}$$

$$Q_t = \alpha Q_{t-1} + (1 - \alpha) (K_t d_t d_t^T K_t^T)$$

$$\hat{N}_{t/t} = \hat{N}_{t/t-1} + K_t [y_t - m_t(\hat{N}_{t/t-1})]$$

5. Case Study

In this section, we compare the results derived from the hybrid Extended Kalman Filter with the tuned parameter a to those obtained by the same filter but with a stable parameter ($a = 0.3$). The presented case study concerns data from the area of Lesvos in the Aegean Sea for the period 2010–2012. As training data for the pFFNN, one thousand WAM forecasts were used (DataSet1), while as training data for the hybrid EKF, the next two hundred and fifty model estimations were applied (DataSet2). Finally, in order to validate the performance of the proposed method, the following forty WAM predictions were used (DataSet3).

The values of Bias, Absolute Bias, Root Mean Square Error (Rmse) and Nash–Sutcliffe model efficiency coefficient (Ns) are reported in Table 1. More specifically, the indications “HybridEKFwithStableParameter” and “HybridEKFwithTuning Parameter” express the corresponding results from the proposed methodology, while the indication “Model” presents the direct output of the numerical wave model prior to the use of the method.

Additionally, time series diagrams per year are included. In particular, the indicator “True Values” expresses the actual observations of the Significant Wave Height, while “Model”, “HybridEKFwithStablePar” and “HybridEKFwithTuningPar” present the results produced from the numerical model before and after the implementation of the method, respectively.

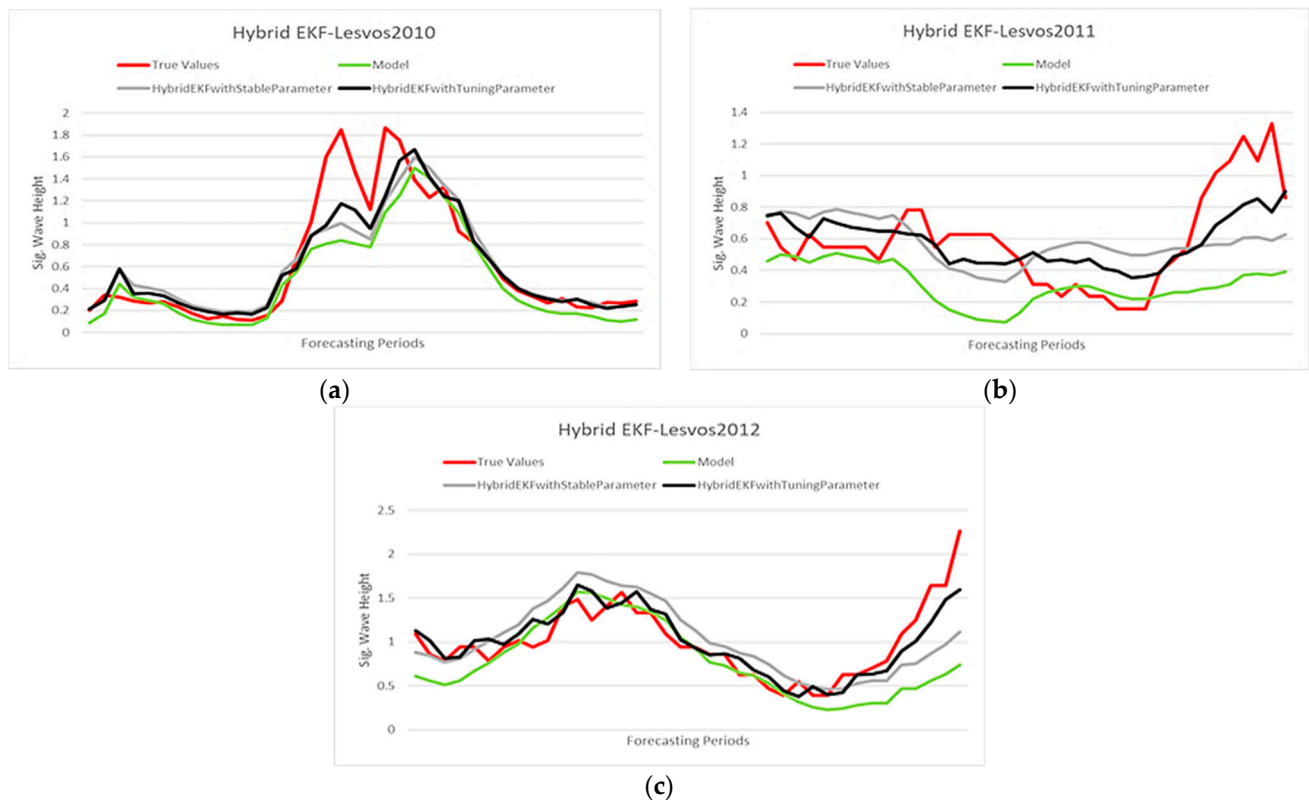
Table 1. Aggregate results for the different periods. The historic database contains 1250 observations and 40 correct predictions.

2010	Bias	AbsoluteBias	Rmse	Ns
Model	0.148526	0.189	0.3007057	0.6992427
HybridEKFwithStablePar a = 0.3	0.027886	0.1643306	0.259394	0.7762037
HybridEKFwithTuningPar a = 0.7	0.030033	0.1359092	0.2185181	0.941179
2011	Bias	AbsoluteBias	Rmse	Ns
Model	0.275632	0.2956842	0.4009535	−0.9108604
HybridEKFwithStablePar a = 0.3	0.117658	0.117658	0.3021639	−0.0852406
HybridEKFwithTuningPar a = 0.6	0.023235	0.1586907	0.1962634	0.5421536
2012	Bias	AbsoluteBias	Rmse	Ns
Model	0.208316	0.2834211	0.4355139	−0.172398
HybridEKFwithStablePar a = 0.3	−0.02932	0.2656988	0.3486826	0.2484957
HybridEKFwithTuningPar a = 0.1	−0.00571	0.1414679	0.1940013	0.7673621

Improving the Forecasting Capabilities-Lesvos Region

As new data for the Significant Wave Height were provided every three hours, our goal was to employ the suggested strategy to improve the wave model predictions for a period of five days.

Analyzing the time series (Figure 1) and Table 1, it is observed that the proposed method enhances the forecasts of the simulation model for all time periods. In particular, the various indicators improved greatly, most of all the Nash–Sutcliffe model efficiency coefficient index (Ns) and the Root Mean Square Error (Rmse).

**Figure 1.** Forecasts from Lesvos in 2010–2012. The historic database contains 1250 observations and 40 correct predictions. (a) Results from 2010; (b) results from 2011; (c) results from 2012.

According to the N_s , the greatest improvement was observed in the period of 2012, which prior to the usage of the method had a negative value ($N_s = -0.172398$), while after the implementation of the hybrid filter with a tuning parameter, the corresponding value increased to $N_s = 0.7673621$.

Furthermore, in the same period, the $Rmse$ indicator significantly reduced, as its prior value was $Rmse = 0.4355139$, while after the application of the proposed methodology with the adaptive parameter, the corresponding value was $Rmse = 0.1940013$. Additionally, it should be noted that in all cases, the hybrid filter with a fixed parameter improves the model's performance, but to a lesser extent compared to the adapted parameter filter and therefore it cannot be considered as the optimal choice.

6. Conclusions

In this study, we introduce a hybrid Extended Kalman Filter based on a parametrized FeedForward Neural Network (pFFNN) for the prediction of the Significant Wave Height (SWH). In particular, for a set of historical data, the pFFNN is trained for various values of the parameter and the optimal value is selected. Then, the chosen parameter is used in the Extended Kalman Filter in order to modify the covariance matrices.

The efficiency of the above methodology was evaluated through various time periods, which concerned (among others) the area of Lesvos in the Aegean Sea in 2010–2012. For every case, we compared the hybrid Extended Kalman Filter with a tuned parameter with the corresponding constant parameter filter to determine to what extent the value of the parameter influenced the performance of the proposed methodology.

The hybrid filter significantly improved the forecasting results of the simulation system. More specifically, in all cases, we observed that the hybrid filter based on a tuned parameter provided better results in comparison to those derived from the fixed parameter filter and therefore it is considered the better option to improve the forecasts of the WAM-cycle 4 model.

Finally, the proposed method was also implemented for the Pacific and other Aegean regions with similar success, but the results are not included in this paper. Furthermore, it should be noted that the hybrid approach of ANNs and EKF is not limited to WAM-cycle 4 applications, but can be extended to other numerical prediction models as well.

Author Contributions: Conceptualization, A.D., G.G. and I.T.F.; methodology, A.D., G.G. and I.T.F.; software, A.D. and I.T.F.; validation, A.D., G.G. and I.T.F.; formal analysis, A.D., G.G. and I.T.F.; investigation, A.D., G.G. and I.T.F.; resources, G.G.; data curation, A.D. and G.G.; writing—original draft preparation, A.D. and I.T.F.; writing—review and editing, A.D., G.G. and I.T.F.; visualization, A.D., G.G. and I.T.F.; supervision, G.G. and I.T.F.; project administration, G.G. and I.T.F.; funding acquisition, I.T.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been funded by the Unit of Economic and Administrative Support (E.L.K.E) of University of West Attica.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is unavailable due to privacy restrictions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Famelis, I. Runge-Kutta solutions for an environmental parameter prediction boundary value problem. *J. Coupled Syst. Mul-tiscale Dyn. Exp.* **2014**, *2*, 62–69. [\[CrossRef\]](#)
2. Famelis, I.; Galanis, G.; Ehrhardt, M.; Triantafyllou, D. Classical and Quasi-Newton Methods for a Meteorological Parameters Prediction Boundary Value Problem. *Appl. Math. Inf. Sci.* **2014**, *8*, 2683–2693. [\[CrossRef\]](#)
3. Famelis, I.; Tsitouras, C. Quadric Shooting Solution for an Environmental parameter prediction model. *Far East J. Appl. Math.* **2015**, *91*, 81–98. [\[CrossRef\]](#)

4. Galanis, G.; Chu, P.C.; Kallos, G. Statistical post processes for the improvement of the results of numerical wave prediction models. A combination of Kolmogorov-Zurbenko and Kalman filters. *J. Oper. Oceanogr.* **2011**, *4*, 23–31. [\[CrossRef\]](#)
5. Haykin, S. *Kalman Filtering And Neural Networks*; John Wiley & Sons: Hamilton, ON, Canada, 2001; p. 304.
6. Famelis, I.; Donas, A.; Galanis, G. Comparative study of FeedForward and Radial Basis Function Neural Networks for solving an Environmental Boundary Value Problem. *Results Appl. Math.* **2022**, *16*, 100344. [\[CrossRef\]](#)
7. Famelis, I.; Kaloutsas, V. Parameterized neural network training for the solution of a class of stiff initial value systems. *Neural Comput. Appl.* **2021**, *33*, 3363–3370. [\[CrossRef\]](#)
8. Shahrokh, A.; Huang, Z. Adaptive Adjustment of Noise Covariance in Kalman Filter for Dynamic State Estimation. In Proceedings of the IEEE General Meeting Power & Energy Society, Chicago, IL, USA, 16–20 July 2017.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.