The parameter employed for the tests is the wind speed at 10 m as simulated by the atmospheric prediction system Skiron. This choice has been made mainly for two reasons:

- Wind speed is one of the most critical parameters of numerical weather forecasting having direct impacts in renewable energy site assessment, maintenance and forecasting.
- On the other hand, due to its variable and discontinuous nature and the increased correlation with local area characteristics, it is one of the most commonly biased parameters by meteorological forecasting models.

For the present study, two different and independent testing approaches have been adopted:

- Validation against artificially created data randomly developed from specific probability density functions, testing the performance of the proposed methodology in an ideal framework.
- Evaluation in real world conditions against observations of meteorological stations in Greece and the Aegean Sea.

The first approach is based on idealized data randomly created which, however, are coming from the real distributions that describe wind speed values over specific sites with the associated “error” considered as Gaussian. More precisely, the following procedure has been followed:

- The distribution that optimally describes the real records from a meteorological station is defined.
- Idealized observations are randomly created by the obtained probability density function (pdf).
- The corresponding “model” values are considered to emerge from the idealized observations by a randomly created noise from the normal distribution.

In the second – real-world – approach, observations and corresponding forecasts from the Skiron model are employed over a number of sites covering all the latitude range of Greece as presented in Figure 1. The available wind speed time series cover a period of 3 years during which independent test periods have been selected corresponding to different seasonal characteristics.

**The statistical indexes**

The evaluation of the results has been based on the following statistical indexes:

\[
\text{Bias} = \frac{1}{n} \sum_{i=1}^{n} (\text{obs}_i - \text{mod}_i)
\]

where obs and mod stand for the observations and modeled values respectively. It gives the size of the sample under study. This is a basic but particularly important measure of success for any optimization system and especially Kalman type filters.

**Root Mean Square Error:**

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\text{obs}_i - \text{mod}_i)^2}
\]

a measure of the variation of the model discrepancies.

The Nash-Sutcliffe model efficiency coefficient (NS):

\[
\text{NS} = 1 - \frac{\sum_{i=1}^{n} (\text{obs}_i - \text{mod}_i)^2}{\sum_{i=1}^{n} (\text{obs}_i - \text{mean}_i)^2}
\]

where \(\text{mean}_i\) is observed values and \(\text{mean}_m\) is modeled values at time \(i\).

This is a commonly used index to assess the models predictive accuracy. Nash-Sutcliffe efficiencies can range from -\infty to 1, where an efficiency of 1 corresponds to a perfect match between model observations and an efficiency of 0 indicates that the model predictions are as accurate as the mean of the observed data as prediction value.

**Finally,** the Percentile, that is the values under which specific percentiles of the data set under study are recorded, of the modeled and observed data are utilized for comparing the full corresponding distributions and revealing discrepancies both in the most common appearing cases as well as in non-frequent/extreme values that could be of great importance for applications.

**Beginning with the analysis of the results,** in Figure 2, the above statistical measures are presented for the site of Skyros Island in the Aegean Sea for the application of a classical standard Kalman filter and the new proposed IG-Kalman filter over idealized data. The latter have been randomly produced based on the Weibull distribution with scale parameter 1.67 and shape 7.10, which have been estimated by all the available real observations in the area.

The obtained results for the proposed new IG-Kalman filter are more satisfactory leading to the significant reduction of the Bias, both of the direct model outputs and the standard Kalman filter, and to a noticeable improvement of RMSE and NS indexes, facts that prove the general qualitative improvement of the final filtered forecasts. This is further underlined by the correspondence of the various values in all their range for observations and IG-Kalman filtered predictions.

Similarly good results have been obtained over all the stations under study in the ideal environment of artificially produced data based on the probability distribution functions that optimally fit the history data of each station. A second indicative example is presented in Figure 3 for the case of Soula in Crete island. Artificially idealized data have been tested also in this case randomly obtained by the Weibull distribution with scale parameter 1.67 and shape 1.10.

**The analysis of idealized wind speed data indicates an overall improvement of model predictions using the proposed Kalman filter.** In the same direction, the study of real data results confirms the level of improvement presented in idealized cases. In cases where the original direct model outputs are of a very good quality leaving limited space for improvements both the classical Kalman and IG-Kalman filters provide similar level of improvements. More over, the proposed filter practically eliminates systematic error in the vast majority of idealized cases. In real terms the error is reduced by at least 50% and reaching a 90-100% rate of improvement in a significant number of cases tested.

Except BIAS statistical index, the level of improvement concerning RMSE and Nash-Sutcliffe coefficient was calculated. In idealized tests, both RMSE and NS indices were improved, while NS displayed a significant rate of correction of more than 50%. The improvement to RMSE reached 20-40% in all cases. The aforementioned behavior was repeated in real data where in the 45% of the test cases (95% confidence coefficient) were improved at an average rate of 15% and 80% respectively.

The results concerning significant wind height correction using classic Kalman filter and Kalman-IQ present a slightly different behavior than wind speed data due to their particularly low values of variance. Nevertheless, the use of a constant variance value of 0.0010 offers the same level and quality of corrections to wave data as those achieved by classic Kalman.

As a conclusion, the main target of the proposed Kalman filter and Kalman filters in general is the elimination of systematic errors. In the present analysis, systematic errors are quantified by the statistical index BIAS, which is significantly improved by Kalman filter correction. The correction achieved in other statistical parameters such as RMSE and NS confirms the ability of the proposed filter to reduce the variability of the error.

Finally, as for another improvement feature, the amount of CPU time needed to complete the execution of both filters, Kalman IQ and classical Kalman, were recorded during the run of all test cases, both idealized and real. There is a reduce in average CPU time of 22.9 % in the case of the run of new Kalman IQ filter compared to runs of the classical Kalman filter, indicating the potential of the proposed filter to correct model predictions accuracy and reduce the needed CPU resources.

**Conclusions**

The analysis of idealized wind speed data indicates an overall improvement of model predictions using the proposed Kalman filter. In the same direction, the study of real data results confirms the level of improvement presented in idealized cases. In cases where the original direct model outputs are of a very good quality leaving limited space for improvements both the classical Kalman and IG-Kalman filters provide similar level of improvements. More over, the proposed filter practically eliminates systematic error in the vast majority of idealized cases. In real terms the error is reduced by at least 50% and reaching a 90-100% rate of improvement in a significant number of cases tested. Except BIAS statistical index, the level of improvement concerning RMSE and Nash-Sutcliffe coefficient was calculated. In idealized tests, both RMSE and NS indices were improved, while NS displayed a significant rate of correction of more than 50%. The improvement to RMSE reached 20-40% in all cases. The aforementioned behavior was repeated in real data where in the 45% of the test cases (95% confidence coefficient) were improved at an average rate of 15% and 80% respectively.

The results concerning significant wind height correction using classic Kalman filter and Kalman-IQ present a slightly different behavior than wind speed data due to their particularly low values of variance. Nevertheless, the use of a constant variance value of 0.0010 offers the same level and quality of corrections to wave data as those achieved by classic Kalman.

As a conclusion, the main target of the proposed Kalman filter and Kalman filters in general is the elimination of systematic errors. In the present analysis, systematic errors are quantified by the statistical index BIAS, which is significantly improved by Kalman filter correction. The correction achieved in other statistical parameters such as RMSE and NS confirms the ability of the proposed filter to reduce the variability of the error.

Finally, as for another improvement feature, the amount of CPU time needed to complete the execution of both filters, Kalman IQ and classical Kalman, were recorded during the run of all test cases, both idealized and real. There is a reduce in average CPU time of 22.9 % in the case of the run of new Kalman IQ filter compared to runs of the classical Kalman filter, indicating the potential of the proposed filter to correct model predictions accuracy and reduce the needed CPU resources.