

Data-Driven Modeling and Optimization of Wind Turbine Performance

(Evidence from a Wind Farm in Kosovo)

Naim Baftiu¹, Ana Atanasova², Enis Baftiu³, Tatjana A. Pachemska⁴,
Egzon Baftiu⁵, Vladimir Zdraveski⁶, Petre Lameski⁷

Faculty of Computer Science, Goce Delchev University, Stip, North Macedonia^{1,4}

Faculty of Computer Science and Engineering, "Ss. Cyril and Methodius" University in Skopje, North Macedonia^{2,6,7}

IT Director, Northern Technologies International Corporation (NTIC), Circle Pines, Minnesota, United States of America³

Software Developer, Adaptech LLC, Austrabe 21, 65527 Niedernhausen, Germany⁵

Abstract—This study proposes an integrated data-driven framework that couples wind power forecasting with the operational optimization of wind turbines under real meteorological conditions. Unlike conventional approaches that focus primarily on improving the predictive accuracy of energy production models, the proposed framework extends forecasting toward proactive operational decision-making, establishing a direct link between prediction, turbine operation, and energy output. Machine learning techniques are used to model the complex nonlinear relationships between meteorological variables and turbine power output, after which a grid-based optimization procedure identifies the operating conditions that maximize generation. The framework is evaluated on ten years of meteorological data for Kosovo covering the period 2015–2024, providing an empirical basis for assessing its applicability. Particular attention is given to identifying the most influential operational and environmental parameters and to determining optimal operating conditions under varying wind regimes. Feature-importance analysis indicates that wind speed and wind direction dominate the response of the model, while rotor speed and blade pitch contribute almost as strongly as the external conditions. Applying the optimization layer to the predicted operating points yields an average increase in power output of 73.70 kW, corresponding to an estimated annual energy gain of approximately 645.60 MWh. A comparative evaluation further shows that a recurrent architecture based on Long Short-Term Memory achieves substantially lower forecasting error than the Random Forest model employed for optimization. The main contribution of the study is a unified forecasting–optimization framework that moves beyond predictive accuracy toward measurable performance improvement, and that can serve as a practical decision-support tool for wind farm operators.

Keywords—Wind energy; wind power forecasting; machine learning; operational optimization; wind turbine performance; random forest; long short-term memory

I. INTRODUCTION

The integration of wind energy into modern power systems is constrained less by the availability of the resource than by its variability. Because the power delivered by a turbine responds nonlinearly to a set of interacting meteorological variables, system operators require reliable estimates of production in order to schedule generation, balance the network, and limit curtailment. A substantial body of work has therefore been

devoted to wind power forecasting, ranging from classical statistical models to modern machine learning and deep learning architectures [1]–[4].

The dominant evaluation paradigm in this literature is statistical: competing models are ranked according to how closely their predictions match observed production, using error measures such as the mean absolute error, the root mean squared error and the coefficient of determination. This paradigm has produced considerable methodological progress, but it leaves an important question unanswered. A forecast, however accurate, does not by itself increase the energy delivered by a turbine; it acquires value only when it is translated into an operating decision. Comparatively little attention has been paid to the step that connects the two, namely the use of a trained predictive model as a surrogate over which turbine control parameters can be searched.

The present study addresses this gap by developing a unified forecasting–optimization framework and evaluating it on real meteorological data for Kosovo covering the period 2015–2024. A Random Forest regressor is first trained to predict power output from meteorological and operational variables. The trained model is then used in a second phase as an objective function: for every instance in the test set, combinations of rotor speed and blade pitch are evaluated and the combination that maximizes predicted output is retained. The difference between the power predicted under the current control setting and the power predicted under the optimized setting is reported both instantaneously, in kW, and as an estimated annual energy gain, in MWh.

A. Research Objectives

To achieve this aim, the following objectives were defined:

- To assemble and prepare real meteorological data for Kosovo for the period 2015–2024;
- To develop an accurate model for forecasting power output;
- To optimize the key turbine control parameters, specifically rotor speed and blade pitch;
- To quantify the resulting energy gains in kW and in MWh per year;

- To compare the results obtained with reference studies on wind power forecasting.

B. Contribution of this Study

The contribution of the work is threefold. First, it formulates operational optimization as a search performed directly over a trained forecasting model, so that prediction and control are treated within a single computational pipeline rather than as separate exercises. Second, it quantifies the outcome of that search in physical rather than purely statistical units, expressing the benefit as an average power increment and an estimated annual energy gain. Third, it evaluates the framework on a ten-year meteorological record for Kosovo, a setting characterized by light to moderate wind conditions in which precise turbine control is particularly consequential.

C. Organization of the Study

The remainder of the study is organized as follows. Section II reviews related work on statistical, machine learning, and deep learning forecasting models and on optimization in wind energy systems. Section III describes the data, the construction of the model variables and the proposed forecasting-optimization methodology. Section IV reports the experimental results. Section V discusses their interpretation, implications, and limitations, and Section VI concludes.

II. RELATED WORK

A. Statistical Forecasting Models

Classical time-series approaches remain a standard benchmark in wind power forecasting. The Box-Jenkins family of autoregressive integrated moving average models, including its seasonal variant, provides a well-established framework for representing linear temporal dependence [2], while general methodological guidance for the construction and evaluation of such models is given by Hyndman and Athanasopoulos [3]. Decomposition-based procedures such as Prophet, which describe a series through additive trend, seasonality, and holiday components, have been proposed as robust and easily configurable alternatives [4]. Because these formulations assume approximately linear relationships, their accuracy tends to degrade when the underlying atmospheric processes are strongly nonlinear. The importance of rigorous out-of-sample evaluation in this setting has long been emphasised [16], and broader reviews situate deterministic forecasting within the wider move toward probabilistic energy forecasting [1].

B. Machine Learning Approaches

Ensemble methods based on decision trees are widely used for wind power prediction because they capture nonlinear interactions without requiring an explicit functional form. The Random Forest algorithm aggregates a large number of decorrelated trees and provides, as a by-product, a measure of the relative importance of each input variable [5]. Wei et al. applied this approach to SCADA records for short-term wind power forecasting and reported competitive accuracy relative to conventional baselines [6]. Kernel-based regression constitutes a second established family: support vector regression [7], [8] performs well on moderately sized datasets.

Although its accuracy depends strongly on the selection of the regularisation and kernel parameters, practical guidance has been developed [9].

C. Deep Learning Approaches

Recurrent neural networks with gated memory cells were introduced to address the vanishing-gradient problem that limits the modelling of long-range temporal dependence [10], and a systematic evaluation of the architectural variants of this family is provided by Greff et al. [11]. Applications to wind power are numerous: Liu et al. demonstrated that a Long Short-Term Memory network outperforms conventional statistical baselines on short-horizon forecasting [12], while hybrid formulations that combine a linear model with a recurrent network have been proposed to capture linear and nonlinear structure simultaneously [13]. Convolutional recurrent architectures extend the same principle to spatio-temporal fields [14].

D. Optimization in Wind Energy Systems

Optimization has traditionally been applied in wind energy at the design stage, for instance in the placement of turbines so as to maximize the energy captured across a site [15]. Operational optimization at the level of an individual turbine, in which control parameters are adjusted in response to predicted conditions, has received considerably less attention in the data-driven literature. The present study is positioned at this intersection: it retains the forecasting model as the analytical core, but treats it as a surrogate over which an operating decision is searched, thereby linking the two strands of research described above.

III. MATERIALS AND METHODS

A. Study Area and Data Source

The meteorological data used in this study were obtained from the renewables.ninja platform for the territory of the Republic of Kosovo and cover the ten years from 2015 to 2024. The retrieved variables comprise wind speed, temperature, humidity, and air density. Fig. 1 presents the statistical distribution of wind speed over the observation period. The distribution follows the Weibull form that is characteristic of wind energy studies, with a high concentration of observations at low and moderate speeds. The modal wind speed lies between approximately 0.8 and 1.2 m/s, and although the majority of measurements fall below 4 m/s, the distribution exhibits a long right tail extending to between 6 and 8 m/s. It is this tail that accounts for the greater part of the energy that can be extracted, since the available power scales with the cube of wind speed. The high absolute frequencies, reaching approximately 16,000 observations in the modal bin, indicate that the models were trained on a sufficiently large sample.

The wind regime shown in Fig. 1 has a direct bearing on the design of the study. Under light to moderate conditions, the operating point of a turbine lies for much of the time on the steep part of the power curve, where small changes in blade pitch and rotor speed translate into comparatively large changes in output. This is precisely the regime in which an optimization layer is expected to be most valuable.

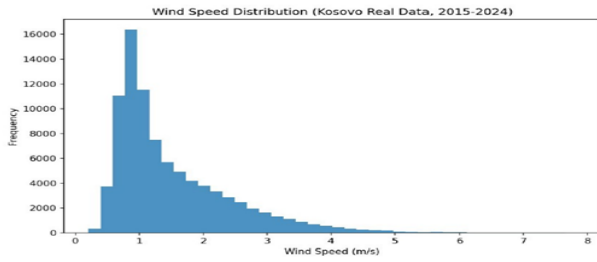


Fig. 1. Statistical distribution of wind speed in Kosovo over the period 2015–2024.

B. Data Preparation

The raw records were harmonised on a common timestamp, filtered to the period 2015–2024, and arranged in time-series format. This procedure ensures the quality and internal consistency of the dataset and provides a uniform temporal resolution for the subsequent modelling stages.

C. Construction of the Model Variables

The variables of the final dataset are summarised in Table I. The pressure variable was not measured directly but derived from the ideal gas law, which links temperature and air density. Because complete SCADA records were not available for the site, the power output, rotor speed and blade pitch variables were modelled so as to reproduce the characteristic operating behaviour of a wind turbine. This is an acknowledged constraint of the present work and is discussed in Section V-D.

TABLE I. VARIABLES OF THE FINAL MODELLING DATASET

Variable	Description	Role
Timestamp	Time index of the observation	Index
WindSpeed	Horizontal wind speed	Predictor
WindDirection	Direction of the incident flow	Predictor
Temperature	Ambient air temperature	Predictor
Humidity	Relative air humidity	Predictor
Pressure	Derived from the ideal gas law	Predictor
RotorSpeed	Rotational speed of the rotor	Control
BladePitch	Pitch angle of the blades	Control
TurbineStatus	Operating state of the turbine	Auxiliary
PowerOutput	Electrical power delivered	Target

D. Train–Test Partition

The dataset was partitioned chronologically, with random shuffling explicitly disabled. Preserving the temporal ordering of the observations is essential in a time-series setting: a random partition would allow information from later periods to influence the fitted model and would therefore produce optimistically biased estimates of out-of-sample accuracy through data leakage.

E. Phase One: the Forecasting Model

The first phase of the framework trains a Random Forest regressor to predict power output from the meteorological and operational variables listed in Table I. The algorithm constructs an ensemble of decision trees and aggregates their individual predictions, which yields three properties that are relevant here:

it captures nonlinear relationships without an explicit functional specification; it is robust to noise in the input data; and it provides an interpretable ranking of variable importance [5].

F. Phase Two: Grid-Based Operational Optimization

The second phase reuses the trained model as an objective function. Let $\hat{f}$ denote the fitted regressor, x_t the vector of meteorological conditions observed at time t , and ω and β the rotor speed and blade pitch, respectively. For each instance of the test set, a systematic search is performed over a discrete grid of admissible control values, and the combination that maximises the predicted output is retained:

$$(\omega^*, \beta^*) = \arg \max \hat{f}(x_t, \omega, \beta) \quad (1)$$

The instantaneous benefit of the optimization at time t is the difference between the power predicted under the optimized setting and the power predicted under the control setting actually applied:

$$\Delta P_t = \hat{f}(x_t, \omega^*, \beta^*) - \hat{f}(x_t, \omega_t, \beta_t) \quad (2)$$

The corresponding annual energy benefit is obtained by extrapolating the mean gain over the hours of a full year:

$$E = \Delta \bar{P} \times 8760 / 1000 \quad (3)$$

where, E is expressed in MWh per year and $\Delta \bar{P}$ in kW. Grid search was preferred over gradient-based alternatives because the control space is low-dimensional and bounded, which makes an exhaustive evaluation computationally inexpensive while guaranteeing that the returned solution is globally optimal on the grid. This step transforms the model from a forecasting instrument into a decision-support instrument.

G. Evaluation Metrics

The predictive accuracy of the models was assessed using three standard measures. The mean absolute error reports the average magnitude of the deviations:

$$\text{MAE} = (1/n) \sum |y_i - \hat{y}_i| \quad (4)$$

The root mean squared error penalises large deviations more heavily, since the errors are squared before averaging:

$$\text{RMSE} = \sqrt{(1/n) \sum (y_i - \hat{y}_i)^2} \quad (5)$$

The coefficient of determination expresses the proportion of the variance of the observed series that is explained by the model:

$$R^2 = 1 - \sum (y_i - \hat{y}_i)^2 / \sum (y_i - \bar{y})^2 \quad (6)$$

In addition to these metrics, the effect of the optimization layer was quantified through the average power gain in kW and the estimated annual energy gain in MWh per year, as defined in Eq. (2) and (3).

IV. RESULTS

A. Relative Importance of the Input Variables

Fig. 2 reports the relative contribution of each input variable to the predictions of the Random Forest model. Wind speed emerges as the dominant predictor, with an importance above 0.175, which confirms that the model correctly identifies the primary physical driver of power output. Wind direction ranks

second, indicating that the alignment of the rotor with the incident flow is also material to performance.

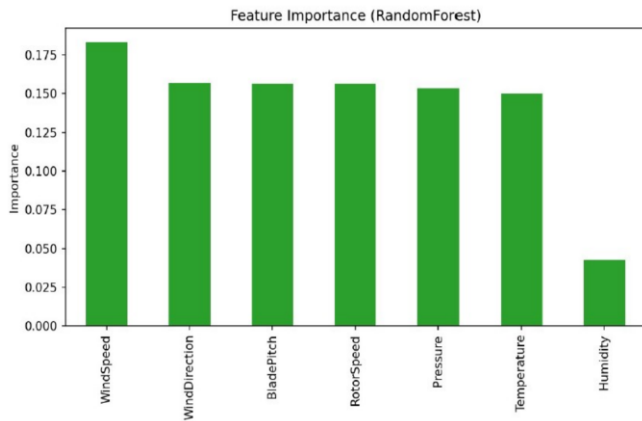


Fig. 2. Distribution of feature importance obtained from the Random Forest model.

The two control variables, blade pitch and rotor speed, exhibit almost equal importance at approximately 0.15. This result is central to the argument of the study: the internal control of the turbine carries nearly as much weight in determining the output as the external atmospheric conditions, which provides an empirical justification for treating these two parameters as decision variables. Pressure and temperature have a comparable influence, consistent with their role in determining air density, whereas humidity is the least influential input, with an importance below 0.05. The ranking as a whole is physically coherent, which supports the view that the fitted model is well behaved rather than merely well fitted.

B. Effect of Operational Optimization on Power Output

Fig. 3 compares the power predicted under the control setting currently applied with the power predicted under the optimized setting, over the final segment of the test period. Under the current setting, the output fluctuates between approximately 20 kW and 40 kW, with isolated excursions to about 65 kW. Under the optimized setting, the series operates in a band between roughly 100 kW and 120 kW.

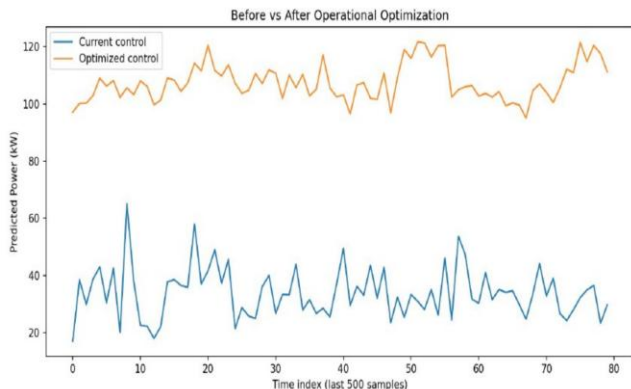


Fig. 3. Predicted power output before and after operational optimization.

Two features of Fig. 3 deserve comment. The first is the magnitude of the shift: the mean output rises from approximately 35 kW to above 105 kW over the segment shown.

The second is the change in the character of the series. The optimized trajectory not only lies higher but is also more stable, in the sense that the extended intervals of depressed output visible in the baseline are absent. The optimization therefore acts less as a uniform additive gain than as the removal of a set of low-performance operating states.

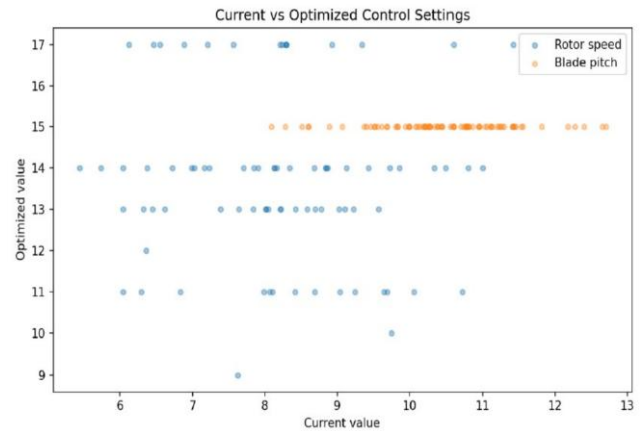


Fig. 4. Current and optimized values of rotor speed and blade pitch.

Fig. 4 examines the control settings that produce this result, plotting the optimized value of each parameter against the value currently applied. The rotor speed points are widely dispersed, indicating that the search adjusted the rotational speed dynamically; the optimized values nevertheless concentrate at the upper end of the admissible range, between approximately 13 and 17, largely irrespective of the starting value. The blade pitch points display a different and more striking pattern: the optimized values converge on a value of approximately 15 across the whole range of current settings, which span roughly 8 to 13.

The convergence of the pitch angle suggests that the baseline control regime was operating away from the point that the model identifies as optimal for the prevailing conditions, and that a fixed angle of approximately 15 units represents the most favourable configuration for energy capture in this wind regime. Taken together, the two panels describe a transition from a broadly distributed, reactive control pattern to a narrow and targeted one.

The aggregate effect of the optimization is summarised in Table II. The average instantaneous gain of 73.70 kW corresponds, through Eq. (3), to an estimated annual energy benefit of 645.60 MWh.

TABLE II. SUMMARY OF THE ENERGY GAIN OBTAINED FROM OPERATIONAL OPTIMIZATION

Indicator	Value	Unit
Mean baseline power (test segment)	35	kW
Mean optimized power (test segment)	105	kW
Average power gain	73.70	kW
Estimated annual energy gain	645.60	MWh/year
Optimized blade pitch (modal)	15	deg
Optimized rotor speed (range)	13–17	rpm

C. Comparative Statistical Performance

Fig. 5 compares the forecasting model and the optimization model across the three evaluation metrics on a normalised scale on which higher values denote better performance. The forecasting model attains the maximum normalised value of 1.0 for both MAE and RMSE, whereas the optimization model operates close to 0.0 on the same axes.

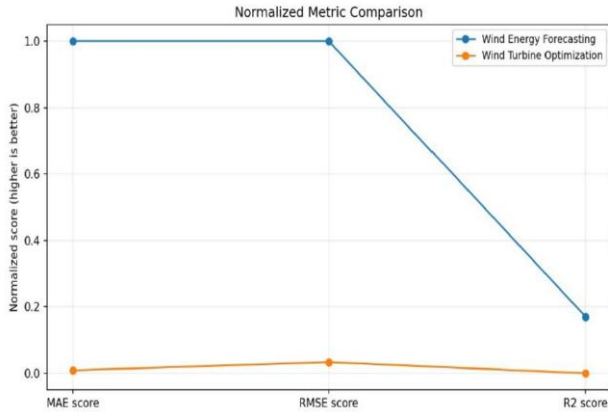


Fig. 5. Normalized comparison of statistical performance between the forecasting model and the optimization model.

This contrast should be interpreted with care. It does not indicate that the optimization model is defective, but rather that the two models are solving different problems: reproducing an observed series is an intrinsically easier task than searching a control space in which variability is substantially higher. The comparatively low R^2 values obtained by both models, relative to their error-based scores, indicate that although the absolute magnitude of the errors is controlled, a large share of the variance of the target remains unexplained. This is consistent with the stochastic character of wind.

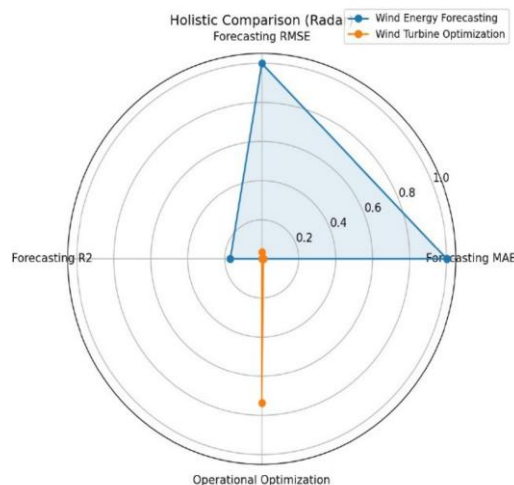


Fig. 6. Radar summary of forecasting accuracy and operational improvement for the two models.

Fig. 6 presents the same comparison in radar form, adding an axis for operational improvement. The forecasting model occupies a large area on the accuracy axes but contributes little on the operational axis, whereas the optimization model displays

the complementary profile, extending strongly toward operational improvement while remaining modest on the statistical axes.

The two profiles are best read as describing a division of labour rather than a ranking. The forecasting model supplies the calibrated representation of the environment in which the turbine operates; the optimization layer converts that representation into an operating decision. Neither component delivers the reported gain on its own.

D. Comparison of the LSTM and Random Forest Models

Fig. 7 reports a quantitative comparison of the Long Short-Term Memory network used for forecasting and the Random Forest model used for optimization, in the original units of the error metrics. The LSTM attains error values close to zero on both MAE and RMSE, confirming the suitability of gated recurrent architectures for wind energy time series. The Random Forest model records higher errors, of the order of 38 for MAE and 47 for RMSE. The numerical values are summarised in Table III.

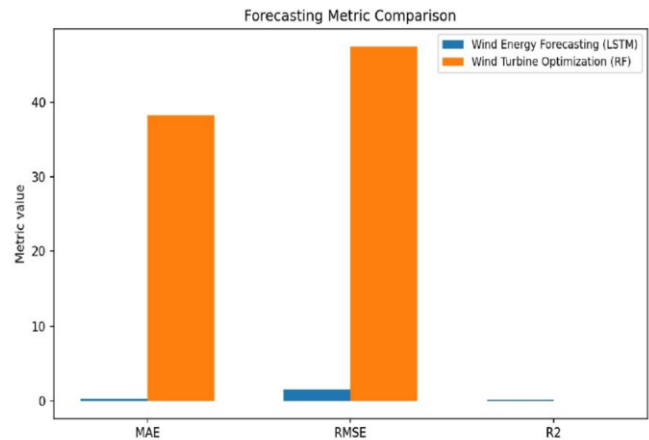


Fig. 7. Comparative error performance of the LSTM and Random Forest models.

TABLE III. COMPARATIVE FORECASTING PERFORMANCE OF THE TWO MODELS

Model	MAE	RMSE	R ²
LSTM (forecasting)	0	0	low
Random Forest (optimization)	38	47	low

The difference in error magnitude is expected given the roles assigned to the two models. The recurrent network is trained exclusively to reproduce a temporal signal, whereas the ensemble model is required to generalise across a control space that is explored counterfactually, at operating points that were never observed. The lower R^2 values recorded by both models reflect the difficulty of explaining the full variance of wind-driven output rather than a deficiency specific to either architecture.

V. DISCUSSION

A. Forecasting Accuracy Versus Operational Value

The results reported above are consistent with the established finding that nonlinear models are better suited to

wind power prediction than linear alternatives, because atmospheric processes are themselves nonlinear and require flexible functional forms to be represented adequately. Linear specifications remain valuable for their simplicity and interpretability, but they do not capture the full complexity of the system.

At the same time, the study shows that even advanced models do not achieve uniformly strong performance: the negative coefficient of determination obtained in the case study reflects a high level of uncertainty in the wind data and confirms that forecasting remains a difficult problem in this setting. The significance of this observation lies in what follows from it. Although the forecasting model is imperfect in a statistical sense, it still supports an operational decision that produces a measurable gain. The value of a model in this context is therefore not exhausted by its statistical accuracy; its practical applicability constitutes a distinct and equally relevant criterion.

B. Integration of the Two Approaches

A forecasting-oriented study is concerned with building and comparing models, evaluates them with statistical metrics, and aims at improving predictive accuracy. Such work is methodologically strong, well suited to benchmarking and provides a solid theoretical foundation, but it does not guarantee any direct improvement in energy production and remains at an analytical level. An optimization-oriented study combines forecasting with a search over the control space, focuses on operational decision-making and generates measurable energy gains, at the cost of depending on the quality of the underlying forecasting model and of requiring greater implementation complexity. Table IV summarises this comparison.

TABLE IV. FORECASTING-ORIENTED AND OPTIMIZATION-ORIENTED APPROACHES

Aspect	Forecasting-oriented	Optimization-oriented
Objective	Predictive accuracy	Operational decision
Evaluation	Statistical metrics	Energy gain in kW, MWh
Strength	Benchmarking, theory	Practical applicability
Limitation	No direct gain	Depends on forecast quality
Output	Predicted series	Control set-points

The central claim of this study is that these two approaches are complementary rather than competing. Forecasting supplies an estimate of future conditions; optimization uses that estimate to select an operating point. Their combination constitutes an integrated decision-support system in which the model predicts production, the optimization layer adjusts the operational parameters, and the resulting configuration delivers higher performance than either component would achieve in isolation.

C. Implications for the Kosovo Power System

In the context of Kosovo, where the renewable generation portfolio is still developing, the framework has practical relevance. Data-driven operational management of the kind described here can contribute to increasing the efficiency of installed turbines, reducing energy losses, improving the stability of the network and supporting decarbonisation policy. These considerations apply in particular to existing installations

such as the Bajgora and Kitka wind farms, provided that sufficient operational and meteorological data are made available for model calibration.

D. Limitations

Four limitations should be borne in mind when interpreting the results. First, complete SCADA records were not available, and the power output, rotor speed and blade pitch series were therefore modelled rather than measured; the reported gains consequently describe the behaviour of the model under a simulated operating regime and require confirmation against instrumented data. Second, the optimization is performed within the trained model, so the estimated gains are conditional on the fidelity of that model and cannot be interpreted as measured production increases. Third, the coefficient of determination obtained in the case study is low and in places negative, which places a clear bound on the confidence that can be attached to individual point forecasts. Fourth, the analysis covers a single site and a single turbine configuration, so the generalisability of the reported set-points to other wind regimes remains to be established.

VI. CONCLUSION

This study has integrated machine learning-based forecasting with the operational optimization of wind turbines, providing a coherent framework for the data-driven management of renewable energy systems. The results indicate that advanced nonlinear models, and in particular Long Short-Term Memory networks, capture the complex temporal dependencies and high variability of wind power generation more effectively than conventional statistical approaches. The superior performance of the recurrent architecture demonstrates that forecasting can be substantially improved when temporal dependence and nonlinearity are modelled explicitly.

The addition of an operational optimization layer over the rotor speed and blade pitch parameters further demonstrates that the practical value of a forecasting model should not be assessed solely by its statistical accuracy. Even where forecasting uncertainty persists, the predicted operating conditions can be used to identify beneficial turbine configurations. In the present case, the optimization yielded an average increase in power output of 73.70 kW, corresponding to an estimated annual energy gain of approximately 645.60 MWh.

Forecasting and data-driven optimization should therefore not be regarded as competing approaches. Their integration provides a decision-support framework capable both of predicting future generation and of identifying the operating conditions associated with improved turbine performance. The findings are particularly relevant in the context of Kosovo, where the efficient use of renewable resources supports energy security, grid integration and the reduction of losses and emissions.

Future work should extend the framework to larger and instrumented datasets, incorporate additional turbine operating parameters, adopt uncertainty-aware forecasting methods, and examine real-time optimization under changing environmental and grid conditions. Such developments would strengthen both the applicability and the scalability of integrated machine

learning and optimization approaches for sustainable wind energy management.

REFERENCES

- [1] T. Hong, P. Pinson, S. Fan, and Y. Ji, "Probabilistic energy forecasting: Global review and future directions," *Renewable and Sustainable Energy Reviews*, vol. 91, pp. 1012–1029, 2018.
- [2] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, NJ, USA: Wiley, 2015.
- [3] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 2nd ed. Melbourne, Australia: OTexts, 2018.
- [4] S. J. Taylor and B. Letham, "Forecasting at scale," *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.
- [5] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [6] T. Wei, X. Li, and Y. Li, "Random forest-based short-term wind power forecasting using SCADA data," *Scientific Reports*, vol. 9, art. no. 56202, 2019.
- [7] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," in *Advances in Neural Information Processing Systems*, vol. 9, 1997, pp. 155–161.
- [8] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, 2004.
- [9] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks*, vol. 17, no. 1, pp. 113–126, 2004.
- [10] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [11] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, 2017.
- [12] H. Liu, J. Zhang, and Y. Li, "Wind power forecasting using LSTM neural network," *Energies*, vol. 12, no. 10, art. no. 1888, 2019.
- [13] Y. Zhang, J. Wang, and H. Wang, "A hybrid forecasting approach based on LSTM neural network and ARIMA model for wind power," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 3, pp. 1085–1095, 2018.
- [14] X. Shi, Z. Chen, H. Wang, D. Y. Yeung, W. K. Wong, and W. C. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Advances in Neural Information Processing Systems*, vol. 28, 2015, pp. 802–810.
- [15] A. Kusiak and Z. Song, "Design of wind farm layout for maximum wind energy capture," *Renewable Energy*, vol. 35, no. 3, pp. 685–694, 2010.
- [16] L. J. Tashman, "Out-of-sample tests of forecasting accuracy: An analysis and review," *International Journal of Forecasting*, vol. 16, no. 4, pp. 437–450, 2000.