Evaluation of Weather Parameters for Renewable Energy Forecasting with Echo State Networks

Samuel G. Dotson[a,*], Kathryn D. Huff[a]

[a] *Dept. of Nuclear, Plasma, and Radiological Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801*

Abstract

The abstract goes here. As a general guide, you should provide a concise (150-250 words) summary of your article - introduction, methodology, results, and conclusion. Avoid using abbreviations and acronyms unless the abbreviation/acronym is used repeatedly in the abstract. There should be no references in the abstract.

Keywords: FIXME, key words, go here, like:, simulation, spent nuclear fuel

3.07242in

1. Introduction

1.1. Motivation

In response to the rising threat of climate change many countries have prioritized reducing carbon emissions. The goal set by the 2015 Paris Agreement is to prevent the global temperature from rising more than 1.5 °C above pre-industrial levels [1]. Virtually all current plans to reduce carbon emissions depend on increasing the share of energy production by renewable and clean energy sources, especially solar and wind [2, 3, 4, 5]. While solar and wind are low-carbon sources, these forms of electricity generation increase variability, which can lead to blackouts and power system failures [6]. Further, even modest penetrations of renewable energy negatively affect the economics of other types of clean energy, such as nuclear power [2, 7, 8]. This may force nuclear plants to shut down prematurely, at the precise moment all clean sources of energy are most needed. There has been some work done to quantify the economic benefit of improving forecasts of renewable energy [9, 10, 11]. Improving renewable energy forecasts can mitigate some of the negative side effects of variability. The economic benefits of better forecasts include: reduced costs compared to building storage devices [9]; curtailment reduction and more efficient use of non-renewable sources [10]; and a slight, but important, amount of load-following from nuclear and biomass generators, which are unable to follow rapid changes in demand [11]. Most proposed forecasting improvements involve new algorithms or machine learning techniques. However, one of the simplest approaches to improving forecasts is to improve the training data for such algorithms. There is a veritable zoo of weather parameters that can supplement target training data, but we don't know *a priori* which of these parameters will be helpful or detrimental to model performance. In this paper, we evaluate several common parameters for use in renewable energy forecasting with Echo State Networks (ESNs).

1.2. Why Echo State Networks

ESNs have several appealing features. They are simple, consisting only of a large, sparse, reservoir and a single output layer [12]; flexible and generalizable, while other network architectures require significant fine tuning [13]; and fast, due to their simple structure and few trainable weights relative to other neural networks. The ESN network architecture eliminates the need for complicated data pre-processing, such as feature extraction, that is required for other machine learning and statistical algorithms [14, 15]. ESNs can also outperform other prediction techniques [16, 17, 18, 19, 20].

Classical ESNs have previously been used to forecast demand, wind energy, and solar energy [21, 17, 20]. ESNs are typically used to make extreme short term predictions, on the order of seconds

*Corresponding Author
Email address: `sgd2@illinois.edu` (Samuel G. Dotson)

or minutes [22, 23, 19], one-hour ahead [18], and up to a single day ahead [21]. Forecasts must be multiple-hours to a couple of days ahead to aid unit commitment and grid-scale energy economy [9, 10, 11]. In this work we use a classic ESN architecture to forecast total demand, wind production, and solar production, 4-hours and 48-hours ahead.

There has been a lot of work to improve the forecasting capability of the basic ESN. Approaches include adding multiple reservoirs [20, 24, 25, 26]; including non-linear units [27, 19]; combining with other network architecture [22, 28]; and using a particle swarm approach [29, 23]. Some works mention that including weather parameters may be useful for renewable energy forecasting [30, 19], but none have demonstrated the effect each parameter has on model performance. The primary goal of this work is to fill that gap.

1.3. Contributions

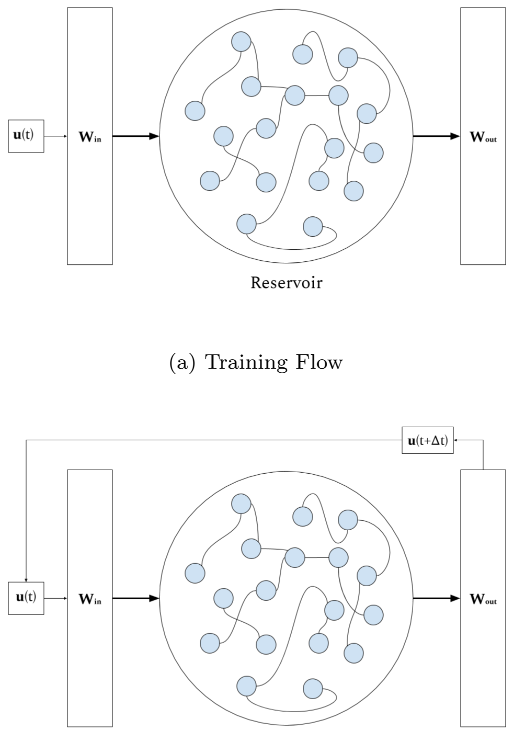
In this work, we use ESNs for three main prediction tasks: total electricity demand, wind energy production, and solar energy production. We split these tasks into further sub tasks; predicting 4-hours ahead and 48-hours ahead. These predictions facilitate scheduling and grid planning because current market rules put renewable energy on the grid first, forcing conventional power generators to work around this variability [9]. Using ESNs to make predictions two-days ahead is unique to this paper since the longest predictions by ESNs in the literature only reach one-day ahead [21]. Finally, we repeat these tasks with several commonly used weather parameters and evaluate their effect on model performance. The need to consider exogenous meteorological inputs has been noted previously. Suprisingly, sun elevation is seldom used as a correlated quantity for energy demand and wind power.

The structure of the paper is as follows. In Section 2, we discuss how data were selected and processed, and we review ESNs. Section 3 shows a benchmarking exercise for our ESN implementation and presents the results. We discuss the results and future implications in Section 4.

2. Methodology

2.1. Echo State Networks

An ESN, sometimes called a "reservoir computer,"[31, 32, 33] is a type of recurrent neural network that replaces the many hidden



(a) Training Flow

(b) Predicting Flow

Figure 1: (a) Shows the behavior of an ESN during the training phase. (b) Shows ESN behavior during the predicting phase. The output $u(t + \Delta t)$ is used as the next input value.

layers of a conventional feed-forward neural network with a reservoir that is:

1. sparse,
2. connected by uniformly random weights, centered at zero,
3. and large (i.e. has many neurons).

The reservoir is therefore a randomly instantiated adjacency matrix, $\boldsymbol{W}$, of size $N \times N$. The input vector, $U(t)$, of K units is mapped onto the reservoir by an input matrix, W^{in} of size $N \times K$. The activation states of the reservoir are calculated by

$$x(t) = \tanh\left(W^{in} \cdot U(t) + \mathbf{W}x(t-1)\right) \qquad (1)$$

Where $x(t)$ is the collection of reservoir activations [18, 32, 12]. The output is read by an output weight matrix, W^{out}.

$$U(t + \Delta t) = \left(W^{out}\right)^T \cdot x(t) \qquad (2)$$

In the training phase, the output, $U(t + \Delta t)$, is discarded and the next training input is passed to the network. During the prediction phase, the output is kept and used as the next input. Figure 1 illustrates this behavior. The speed of ESNs is owed to this structure— only W^{out} has tunable weights. Everything else is fixed. In this work, we adapted the open source Python package `pyESN` [34] to construct and train the network.

2.2. Hyper-Parameter Optimization

ESNs are fast because the hidden layer in a conventional feed-forward neural network is replaced by a large reservoir that does not require training. The trade off is that ESNs are sensitive to various hyper-parameters that need to be optimized [12]. These hyper-parameters are summarized in Table 1. The spectral radius (ρ) should satisfy the "echo state property" which means that previous reservoir activations have a decaying influence on future states. This is usually guaranteed for $\rho < 1$, but is not a requirement [12].

The hyper-parameters are optimized by performing a grid search over the test values specified in Table 1. The following steps were taken for each prediction task:

1. Select a hyper-parameter or pair of parameters.
2. Generate ESN prediction with the specified parameters.
3. Calculate and record the root mean squared error (RMSE).
4. Continue until last entry in the parameter set is reached.
5. Set the network parameters to hyper-parameter value that minimizes the RMSE.

This algorithm generates an error surface where the coordinates of the absolute minimum correspond to the indices of values in the hyper-parameter test sets that minimized the RMSE.

2.3. Prediction Tasks

We first performed a benchmarking task by making a prediction for the Lorenz 1963 model [35]. Then we optimized predictions for univariate time-series representing total demand, solar energy, and wind energy 4-hours ahead and 48-hours ahead. Finally, those same six tasks were repeated with an additional predictor. The tasks are summarized in Table 2.

2.4. Data Selection and Processing

All data predicting demand, wind energy, and solar energy on the University of Illinois at Urbana-Champaign (UIUC) campus are from the UIUC Solar Farm 1.0 dashboard [36] and proprietary data shared with us courtesy of the UIUC Facilities and Services Department. All data had hourly resolution. Weather data was retrieved from the National Oceanic and Atmospheric Administration (NOAA)[37] for two locations: Champaign, IL, where UIUC is located, and Lincoln, IL, where Railsplitter Windfarm is located. UIUC has a power purchase agreement with Railsplitter Windfarm [38]. In the case of UIUC solar data, significant portions were missing due to instrument failure. In order to fill in this missing data, we calculated the theoretical solar energy production based on irradiance data from OpenEI [39]. The solar output is given by [40]

$$P = G_T \eta_{ref} \tau_{pv} A \left[1 - \gamma\left(T - 25\right)\right] \quad [W] \qquad (3)$$

where

$$\begin{aligned} G_T &= P_{DNI} * \cos\left(\beta + \delta - lat\right) \\ &\quad + P_{DHI} * \left(\frac{180 - \beta}{180}\right) \left[\frac{W}{m^2}\right] \end{aligned} \qquad (4)$$

Table 1: Description of Model Hyper-Parameters

Hyper-parameter	Purpose	Tested Values
`noise`	Neuron regularization	[0.0001, 0.0003, 0.0007, 0.001, 0.003, 0.005, 0.007, 0.01]
ρ	Spectral radius	[0.5, 0.7, 0.9, 1, 1.1, 1.2, 1.3, 1.5]
N	Size of reservoir, $\mathbf{W}$	[600, 800, 1000, 1500, 2000, 2500, 3000, 4000]
`sparsity`	The density of connections in $\mathbf{W}$	[0.005, 0.01, 0.03, 0.05, 0.1, 0.12, 0.15, 0.2]
Training Length	Size of the training set before prediction	$L \in$ [5000, 25000], step size = 300

Table 2: Summary of Prediction Tasks

Target	Future	Additional Predictor
		None
Total Demand		Solar Elevation
	4 hours ahead	Humidity
Solar Energy		Pressure
	48 hours ahead	Wet Bulb Temp.
Wind Energy		Dry Bulb Temp.
		Wind Speed

where

$$\delta = 23.44 \sin\left(\left(\frac{\pi}{180}\right)\left(\frac{360}{365}\right)(N+284)\right) \text{[degrees]} \tag{5}$$

η, τ, γ are solar panel properties
P_{DNI} is the direct normal irradiance
P_{DHI} is the diffuse horizontal irradiance
β is the tilt angle of the solar panels

The solar elevation angle, α, was also calculated [41, 42] using coordinates for the UIUC Solar Farm 1.0.

$$\alpha = \sin^{-1}\left[\sin(\delta)\sin(\phi) + \cos(\delta)\cos(\phi)\cos(\omega)\right] \tag{6}$$

where

δ is the declination angle
ϕ is the latitude of interest
ω is the hour angle

Finally, we normalized all of the data using the infinity norm

$$\|\mathbf{x}\|_\infty \equiv \max |x_i| . \tag{7}$$

The infinity norm is equivalent to normalizing by the system capacity. This is useful because it simplifies the comparison of our results between tasks whose training data have vastly different magnitudes. This normalization also makes it possible to compare results with other works and is consistent with the recommendation from Kobylinski et al. (2020) [43].

2.5. Performance Metrics

We measure the accuracy of the model using two error metrics: mean absolute error (MAE) and root mean squared error (RMSE). These are defined as

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i| \tag{8}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2} \tag{9}$$

The MAE measures the expected error throughout the forecast horizon. The RMSE indicates the presence of large but infrequent errors. Since the data were normalized by system capacity [9], the error metrics are easily interpretable. In order to compare how each individual weather input either improved or worsened the forecast we calculated a "percent improvement" over the univariate case (i.e. a demand prediction based only on historical demand data). This percent improvement is calculated by

$$\% \text{ Improvement} = \frac{\hat{e} - e}{e} \times 100, \ [\text{-}] \quad (10)$$

where e is the univariate error and $\hat{e}$ is the duovariate error. The sign indicates the direction of change in error.

3. Results

Below we show the best prediction for each task.

3.1. Benchmark: Lorenz 1963

We first verified that our choice of implementation for ESNs produces similar results to those found in the literature [31]. The hyper-parameters that minimized the RMSE of the model can be found in Table 3. Our optimized values are somewhat different from the literature, but our ESN implementation succesfully replicated the climate of the Lorenz Attractor similar to Pathak et. al 2017.

Table 3: Hyper-parameters for the Lorenz 1963 Model

Parameter	This paper	Literature [31]
N	2000	300
ρ	0.9	1.2
`sparsity`	0.1	0.1
`noise`	0.001	0
Training Length	3200	Not Specified

4. Discussion

The forecast accuracy of our ESN for the Lorenz model does not persist for quite as long as in other works [31]. However, our model successfully replicates the environment that produces the Lorenz Attractor. Further, optimal parameters may be unique for each randomly instantiated reservoir. It is impossible to replicate the exact conditions of other works without information about a seed for the random state. We have included this information for future work to compare with our results.

For each target variable – demand, wind, and solar – we found that air pressure was the only meteorological factor that improved the forecast error in every case. Solar elevation angle also decreased the error in most cases with one exception, 48-hour ahead solar production. Table 5 and Table 4 show that adding humidity, air temperature, or wind speed as a model input weakened the forecast for electricity demand. Yet, adding air temperature as a model input improves the forecast for solar and wind energy, as shown in Table 7, Table 6, Table 9, and Table 8, respectively. This behavior is caused by relative complexity between model inputs. Electricity demand, for example, is quite "predictable," and therefore has low complexity. Air temperature and other weather related variables are less predictable. Thus, adding air temperature as a model input increases the total complexity of the system and weakens performance. Conversely, solar and wind energy are both nonlinear functions of many weather variables and consequently have greater complexity than air temperature, humidity, or air pressure. This means that adding any of those variabes as a model input will decrease the total complexity of the system and improve the forecast. Including wind speed only improved the wind energy forecasts, likely because it has greater complexity than solar energy and less than wind energy. Quantifying the predictability and complexity of these systems is in progress. A good measure for this type of complexity is the *weighted permutation entropy* [44, 45, 46].

Relative complexity also explains why adding solar elevation angle as a model input generally improved the forecast. Solar angle is deterministic and periodic which makes it perfectly predictable. Adding solar angle to any of the models decreases the total complexity of the system. One notable exception to this is shown in Table 6, where solar angle weakened the forecast for solar energy. This is possibly due to ESNs inability to properly handle the zero values at night. Even the best solar energy forecast from this model, shown in Figure 5, shows that the model struggles have zero energy production at night. Air pressure likely improved forecasts, not because it is a correlated quantity, but because it has slightly lower complexity than each of the target variables and because it has very small fluctuations around a mean value.

These results point to an important disadvantage of using ESNs to forecast renewable energy. This network architecture is simple and fast, but remains a black box. We assume that there exists some underlying dynamics that can be "learned" but cannot observe the learning process nor extract important features from ESNs.

The forecast lengths were decided based on the requirements for improved economics and planning

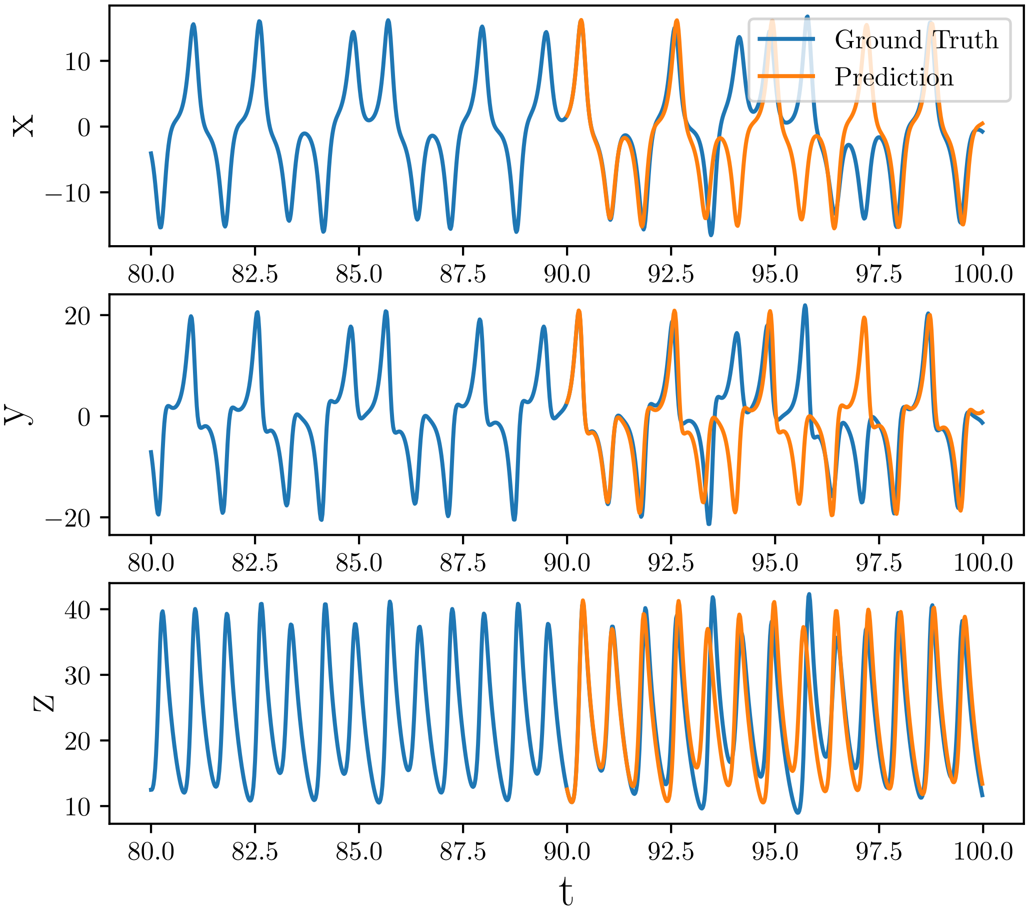


Figure 2: Using an ESN to replicate the climate of the Lorenz Attractor.

mentioned in the literature [9, 10, 11]. The ESN model performed reasonably well at predicting four hours ahead but is not an improvement over the state-of-the-art [9, 47]. The model did not perform well at the 48-hour ahead forecasts. This could be due to the lack of higher resolution data. ESNs are known for their ability to predict highly non- linear systems [48, 49] yet using hourly data could add spurious complexity that confounds the model [45]

4.1. Future Work

One appealing avenue of continued work is to leverage ESNs to generate synthetic data that respects real dynamics. Sythetic data is often useful for other machine learning or optimization algorithms. Typically, these data are produced by sampling from an Auto-Regressive Moving Average (ARMA) model [50, 51], which tacitly assumes the training data can be made stationary. ESNs have been shown to replicate the environment of a dynamical system, although it remains to be seen how far in the future this behavior persists [31, 32]. Future work will also explore the effect of data resolution on model performance, as well as evaluate some of the improvements to the ESN algorithm.

5. Conclusion

Improving renewable energy forecasting is important for grid-planning and unit commitment. Especially as the share of variable renewable resources increases, challenging the baseload power from nuclear plants. We first demonstrated that our implementation of the ESN algorithm is consistent with the literature. Then we applied this model to prediction tasks for total demand, solar energy, and wind

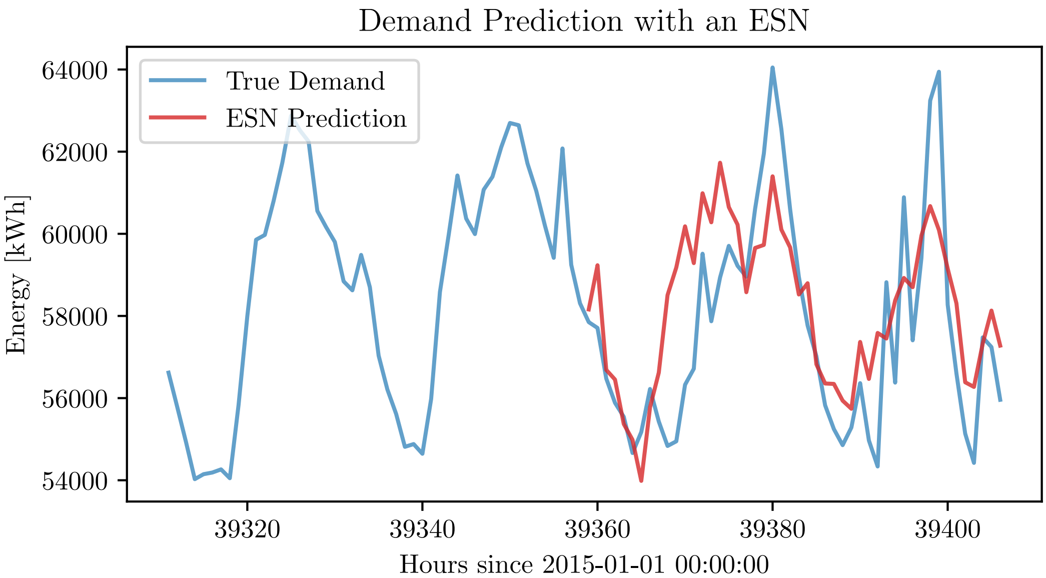


Figure 3: The optimized 48-hour ahead demand prediction. The inputs for this forecast were hourly demand and relative humidity. *Hyperparameters*: Reservoir Size:1500, Sparsity: 0.2, Spectral Radius: 1.5, Noise: 0.0007, Training Length: 5000, Prediction Window: 48, Random state: 85

Table 4: Tabulated error for 48-hour ahead total electricity demand forecasts with various coupled quantities. Improvement indicates the percentage improvement over the base case of forecasting electricity demand alone.

Scenario	MAE	RMSE	Improvement MAE (%)	Improvement RMSE (%)
Total Demand	0.0189	0.0241	[-]	[-]
Demand + Sun Elevation	0.0191	0.0240	+1.0582	-0.4149
Demand + Humidity	0.0180	0.0223	-4.7619	-7.4689
Demand + Pressure	0.0176	0.0245	-6.8783	+1.6600
Demand + Wet Bulb Temp.	0.0241	0.0314	+27.5132	+30.2904
Demand + Dry Bulb Temp.	0.0218	0.0273	+15.3439	+13.2780
Demand + Wind Speed	0.0197	0.0245	+4.2328	+1.6600

energy, and evaluated the influence of several meteorological factors on model performance. Our results show that additional inputs must be chosen carefully to avoid increasing the system complexity. The conventional ESN used here did not demonstrate an improvement over the state-of-the-art. Nor was it accurate enough to improve grid-scale energy economy. Future work will explore other applications of ESNs and evaluate improvements to the model algorithm.

6. Acknowledgments

This work was made possible with the support from the people at UIUC Facilities & Services. In particular, Morgan White, Mike Marquissee, and Mike Larson. It was also aided by other members of the Advanced Reactors and Fuel Cycles (ARFC) group, in particular Nathan Ryan, Gavin Davis, and Nataly Panczyk. This work is supported by the Nuclear Regulatory Commission Fellowship Program. Prof. Huff is supported by the Nuclear Regulatory Commission Faculty Development Program (award NRC-HQ-84-14-G-0054 Program B), the Blue Waters sustained-petascale computing project supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois, the DOE ARPA-E MEITNER Program (award DE-AR0000983), and the DOE H2@Scale Program (Award Number: DE-

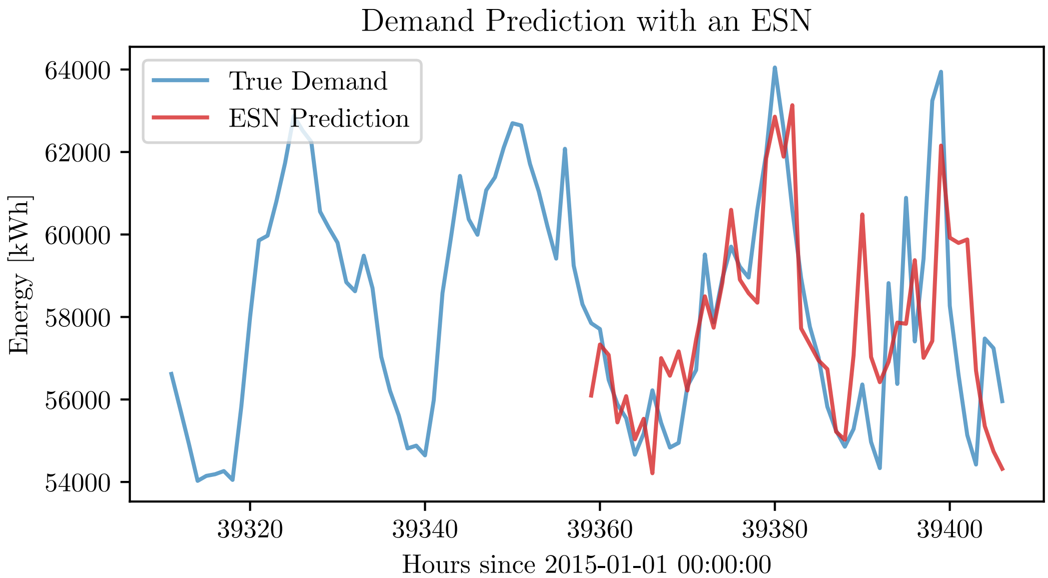


Figure 4: The optimized 4 hour ahead demand prediction. The inputs for this forecast were hourly demand and solar elevation angle. *Hyperparameters*: Reservoir Size:2500, Sparsity: 0.01, Spectral Radius: 1.5, Noise: 0.003, Training Length: 5000, Prediction Window: 4, Random state: 85

Table 5: Tabulated error for 4-hour ahead electricity demand forecasts with various coupled quantities. Improvement indicates the percentage improvement over the base case of forecasting solar energy alone.

Scenario	MAE	RMSE	Improvement MAE (%)	Improvement RMSE (%)
Total Demand	0.0193	0.0263	[-]	[-]
Demand + Sun Elevation	0.0183	0.0239	-5.1831	-9.1255
Demand + Humidity	0.0219	0.0290	+13.4715	+10.2662
Demand + Pressure	0.0186	0.0273	-3.6269	+3.8023
Demand + Wet Bulb Temp.	0.0196	0.0253	+1.5544	-3.8023
Demand + Dry Bulb Temp.	0.0208	0.0270	+7.7720	+2.6616
Demand + Wind Speed	0.0201	0.0268	+4.1451	+1.9011

EE0008832)

References

[1] The paris agreement | UNFCCC. URL `https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement`

[2] C. Cany, C. Mansilla, G. Mathonnière, P. da Costa, Nuclear contribution to the penetration of variable renewable energy sources in a french decarbonised power mix 150 544–555. `doi:10.1016/j.energy.2018.02.122`. URL `http://www.sciencedirect.com/science/article/pii/S0360544218303566`

[3] J. Chilvers, T. J. Foxon, S. Galloway, G. P. Hammond, D. Infield, M. Leach, P. J. Pearson, N. Strachan, G. Strbac, M. Thomson, Realising transition pathways for a more electric, low-carbon energy system in the united kingdom: Challenges, insights and opportunities 231 (6) 440–477, publisher: IMECHE. `doi:10.1177/0957650917695448`. URL `https://doi.org/10.1177/0957650917695448`

[4] 99th General Assembly, Illinois general assembly - full text of SB2814. URL `http://www.ilga.gov/legislation/fulltext.asp?DocName=&SessionId=88&GA=99&DocTypeId=SB&DocNum=2814&GAID=13&LegID=96125&SpecSess=&Session=`

[5] iSEE, Illinois climate action plan (iCAP). URL `https://sustainability.illinois.edu/campus-sustainability/icap/`

[6] H. Haes Alhelou, M. E. Hamedani-Golshan, T. C. Njenda, P. Siano, A survey on power system blackout and cascading events: Research motivations and challenges 12 (4) 682, number: 4 Publisher: Multidisciplinary Digital Publishing Institute. `doi:10.3390/en12040682`. URL `https://www.mdpi.com/1996-1073/12/4/682`

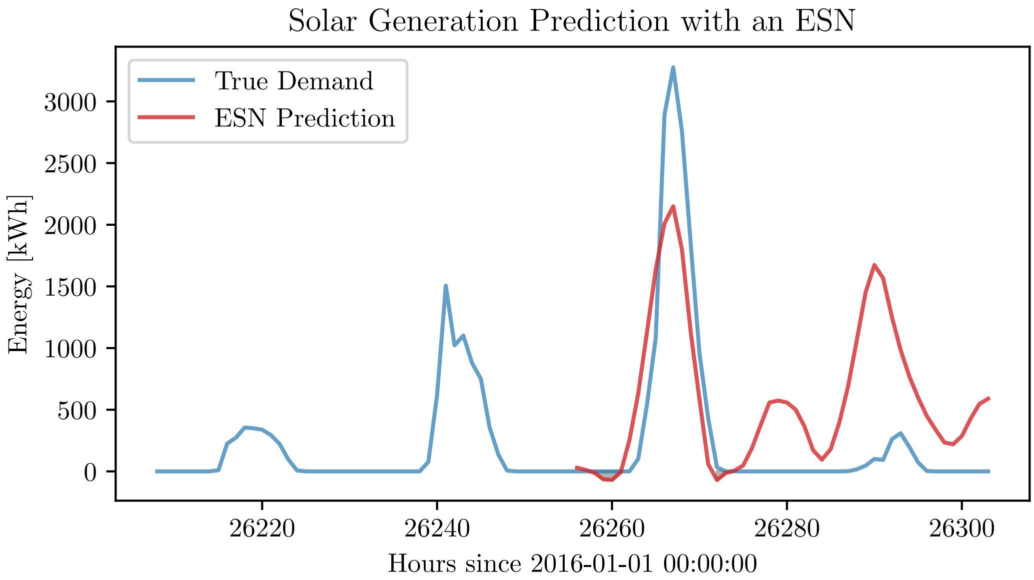


Figure 5: The optimized 48-hour ahead solar energy prediction. The inputs for this forecast were solar energy and relative humidity. Hyperparameters: Reservoir Size:800, Sparsity: 0.2, Spectral Radius: 1.5, Noise: 0.0001, Training Length: 5000, Prediction Window: 48, Random state: 85

Table 6: Tabulated error for 48-hour ahead solar energy forecasts with various coupled quantities. Improvement indicates the percentage improvement over the base case of forecasting solar energy alone.

Scenario	MAE	RMSE	Improvement MAE (%)	Improvement RMSE (%)
Solar Energy	0.1433	0.2062	[-]	[-]
Solar + Sun Elevation	0.0957	0.1375	-33.2170	-33.3172
Solar + Humidity	0.1001	0.1297	-30.1465	-37.1000
Solar + Pressure	0.1910	0.2158	+33.2868	+4.6557
Solar + Wet Bulb Temp.	0.1519	0.1884	+6.0014	-8.6324
Solar + Dry Bulb Temp.	0.1080	0.1512	-24.6336	-26.6731
Solar + Wind Speed	0.2136	0.2500	+49.0579	+21.2415

[7] J. H. Keppler, C. Marcantonini, O. N. E. Agency, O. for Economic Co-operation {and} Development, Carbon pricing, power markets and the competitiveness of nuclear power, Nuclear development, Nuclear Energy Agency, Organisation for Economic Co-operation and Development.

[8] Illinois Commerce Commision (ICC), I. P. A. (IPA), I. E. P. A. (IEPA), I. D. of Commerce and Economic Opportunity (IDCEO), Potential nuclear power plant closings in illinois.
URL `http://www.ilga.gov/reports/special/Report_Potential%20Nuclear%20Power%20Plant%20Closings%20in%20IL.pdf`

[9] Q. Wang, C. B. Martinez-Anido, H. Wu, A. R. Florita, B.-M. Hodge, Quantifying the economic and grid reliability impacts of improved wind power forecasting 7 (4) 1525–1537, conference Name: IEEE Transactions on Sustainable Energy. `doi:10.1109/TSTE.2016.2560628`.

[10] E. V. Mc Garrigle, P. G. Leahy, Quantifying the value of improved wind energy forecasts in a pool-based electricity market 80 517–524. `doi:10.1016/j.renene.2015.02.023`.
URL `http://www.sciencedirect.com/science/article/pii/S0960148115001135`

[11] C. Brancucci Martinez-Anido, B. Botor, A. R. Florita, C. Draxl, S. Lu, H. F. Hamann, B.-M. Hodge, The value of day-ahead solar power forecasting improvement 129 192–203. `doi:10.1016/j.solener.2016.01.049`.
URL `http://www.sciencedirect.com/science/article/pii/S0038092X16000736`

[12] M. Lukoševičius, A practical guide to applying echo state networks, in: G. Montavon, G. B. Orr, K.-R. Müller (Eds.), Neural Networks: Tricks of the Trade: Second Edition, Lecture Notes in Computer Science, Springer, pp. 659–686. `doi:10.1007/978-3-642-35289-8_36`.
URL `https://doi.org/10.1007/978-3-642-35289-8_36`

[13] H. Liu, C. Chen, X. Lv, X. Wu, M. Liu, Deterministic

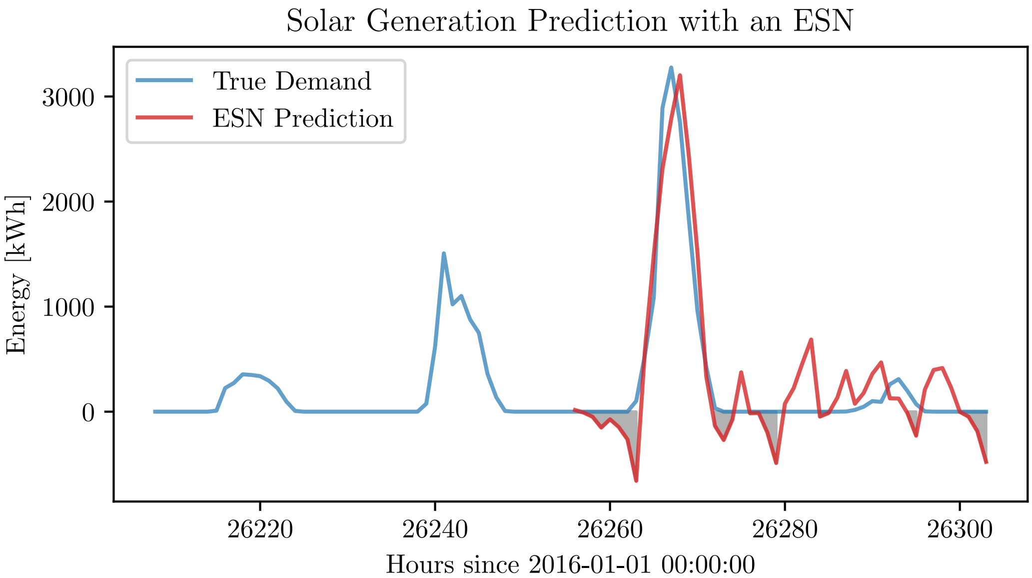


Figure 6: The optimized 4 hour ahead solar energy prediction. The inputs for this forecast were solar energy and hourly wet bulb temperature. *Hyperparameters*: Reservoir Size:800, Sparsity: 0.01, Spectral Radius: 0.9, Noise: 0.0001, Training Length: 5000, Prediction Window: 4, Random state: 85

Table 7: Tabulated error for 4-hour ahead solar energy forecasts with various coupled quantities. Improvement indicates the percentage improvement over the base case of forecasting solar energy alone.

Scenario	MAE	RMSE	Improvement MAE (%)	Improvement RMSE (%)
Solar Energy	0.0614	0.0958	[-]	[-]
Solar + Sun Elevation	0.0554	0.0832	-9.7720	-13.1524
Solar + Humidity	0.0663	0.0971	+7.9804	+1.3570
Solar + Pressure	0.0925	0.1263	+50.6515	+31.8372
Solar + Wet Bulb Temp.	0.0526	0.0673	-14.3322	-29.7954
Solar + Dry Bulb Temp.	0.0682	0.0993	+11.0749	+3.6534
Solar + Wind Speed	0.0723	0.1137	+17.7524	+18.6848

wind energy forecasting: A review of intelligent predictors and auxiliary methods 195 328–345, publisher: Pergamon. doi:10.1016/j.enconman.2019.05.020. URL http://www.sciencedirect.com/science/article/pii/S0196890419305655

[14] D. Lazos, A. B. Sproul, M. Kay, Optimisation of energy management in commercial buildings with weather forecasting inputs: A review 39 587–603. doi:10.1016/j.rser.2014.07.053. URL https://www.sciencedirect.com/science/article/pii/S136403211400505X

[15] Y. Chen, H. Tan, U. Berardi, Day-ahead prediction of hourly electric demand in non-stationary operated commercial buildings: A clustering-based hybrid approach 148 228–237. doi:10.1016/j.enbuild.2017.05.003. URL https://www.sciencedirect.com/science/article/pii/S0378778816313792

[16] I. Jayawardene, G. K. Venayagamoorthy, Comparison of echo state network and extreme learning machine for PV power prediction, in: 2014 IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG), pp. 1–8, ISSN: 2326-7690. doi:10.1109/CIASG.2014.7011546.

[17] I. Jayawardene, G. Venayagamoorthy, Comparison of adaptive neuro-fuzzy inference systems and echo state networks for PV power prediction 53 92–102. doi:10.1016/j.procs.2015.07.283.

[18] G. Shi, D. Liu, Q. Wei, Energy consumption prediction of office buildings based on echo state networks 216 478–488. doi:10.1016/j.neucom.2016.08.004. URL http://www.sciencedirect.com/science/article/pii/S0925231216308219

[19] M. A. Chitsazan, M. S. Fadali, A. Trzynadlowski, Wind speed and wind direction forecasting using echo state network with nonlinear functions 131 879–889, publisher: Pergamon. doi:10.1016/j.renene.2018.07.060. URL http://www.sciencedirect.com/science/article/pii/S0960148118308577

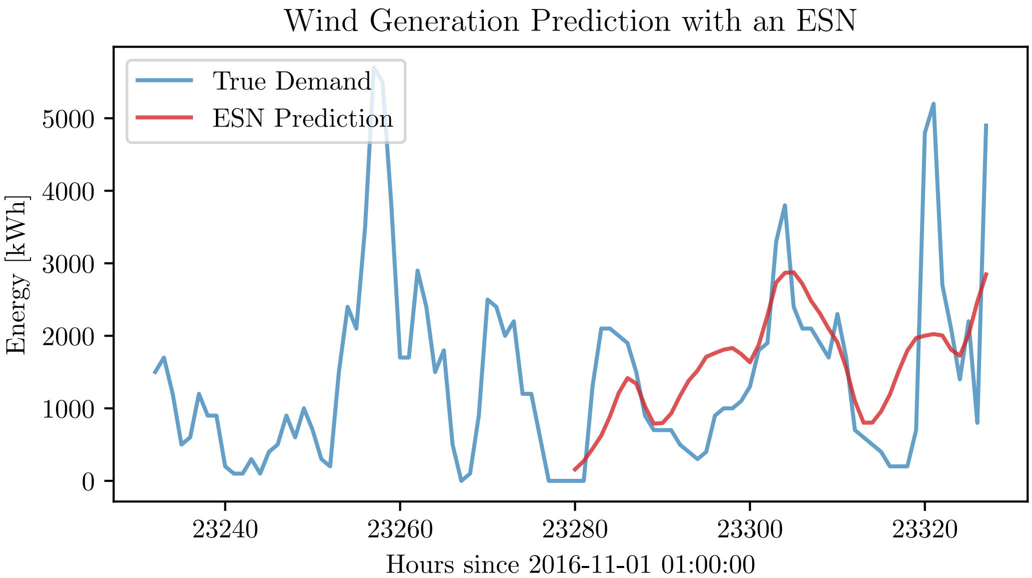


Figure 7: The optimized 48-hour ahead wind energy prediction. The inputs for this forecast were wind energy and solar elevation angle. *Hyperparameters*: Reservoir Size:1000, Sparsity: 0.1, Spectral Radius: 0.9, Noise: 0.0001, Training Length: 19100, Prediction Window: 48, Random state: 85

Table 8: Tabulated error for 48-hour ahead wind forecasts with various coupled quantities. Improvement indicates the percentage improvement over the base case of forecasting wind energy alone.

Scenario	MAE	RMSE	Improvement MAE (%)	Improvement RMSE (%)
Wind Energy	0.1035	0.1308	[-]	[-]
Wind + Sun Elevation	0.0857	0.1141	-17.1981	-12.7676
Wind + Humidity	0.0952	0.1193	-8.0193	-8.7620
Wind + Pressure	0.1076	0.1381	+3.9614	+5.5810
Wind + Wet Bulb Temp.	0.0886	0.1184	-14.3961	-9.4801
Wind + Dry Bulb Temp.	0.0815	0.1213	-21.2560	-7.2630
Wind + Wind Speed	0.0763	0.1182	-26.2802	-9.6330

[20] H. Hu, L. Wang, S.-X. Lv, Forecasting energy consumption and wind power generation using deep echo state network 154 598–613. doi:10.1016/j.renene.2020.03.042.
URL http://www.sciencedirect.com/science/article/pii/S0960148120303645

[21] A. Deihimi, H. Showkati, Application of echo state networks in short-term electric load forecasting 39 (1) 327–340. doi:10.1016/j.energy.2012.01.007.
URL https://linkinghub.elsevier.com/retrieve/pii/S0360544212000126

[22] Y. Chen, Z. He, Z. Shang, C. Li, L. Li, M. Xu, A novel combined model based on echo state network for multi-step ahead wind speed forecasting: A case study of NREL 179 13–29. doi:10.1016/j.enconman.2018.10.068.
URL https://linkinghub.elsevier.com/retrieve/pii/S0196890418311968

[23] H. Wang, Z. Lei, Y. Liu, J. Peng, J. Liu, Echo state network based ensemble approach for wind power forecasting 201 112188. doi:10.1016/j.enconman.2019.112188.
URL http://www.sciencedirect.com/science/article/pii/S019689041931194X

[24] C. Gallicchio, A. Micheli, Deep echo state network (DeepESN): A brief surveyarXiv:1712.04323.
URL http://arxiv.org/abs/1712.04323

[25] X. Yao, Z. Wang, H. Zhang, A novel photovoltaic power forecasting model based on echo state network 325 182–189. doi:10.1016/j.neucom.2018.10.022.
URL http://www.sciencedirect.com/science/article/pii/S0925231218312104

[26] Q. Li, Z. Wu, R. Ling, L. Feng, K. Liu, Multi-reservoir echo state computing for solar irradiance prediction: A fast yet efficient deep learning approach 95 106481. doi:10.1016/j.asoc.2020.106481.
URL https://linkinghub.elsevier.com/retrieve/pii/S1568494620304208

[27] G. Holzmann, H. Hauser, Echo state networks with filter

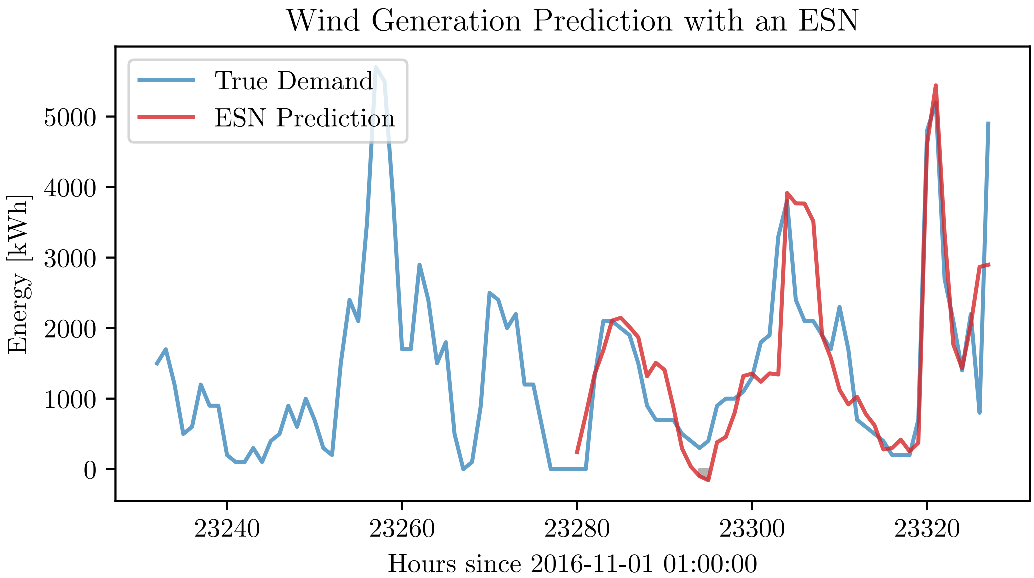


Figure 8: The optimized 4 hour ahead wind energy prediction. The inputs for this forecast were wind energy and hourly windspeed. *Hyperparameters*: Reservoir Size:1000, Sparsity: 0.15, Spectral Radius: 0.9, Noise: 0.001, Training Length: 14300, Prediction Window: 4, Random state: 85

Table 9: Tabulated error for 4-hour ahead wind forecasts with various coupled quantities. Improvement indicates the percentage improvement over the base case of forecasting wind energy alone.

Scenario	MAE	RMSE	Improvement MAE (%)	Improvement RMSE (%)
Wind Energy	0.0903	0.1243	[-]	[-]
Wind + Sun Elevation	0.0705	0.1171	-21.9269	-5.7924
Wind + Humidity	0.0813	0.1201	-9.9668	-3.3789
Wind + Pressure	0.0866	0.1244	-4.0974	+0.0804
Wind + Wet Bulb Temp.	0.0731	0.1070	-19.0476	-13.9179
Wind + Dry Bulb Temp.	0.0747	0.1123	-17.2757	-9.9654
Wind + Wind Speed	0.0571	0.0837	-36.7663	-32.6629

and a delay&sum readout.

[28] E. López, C. Valle, H. Allende, E. Gil, H. Madsen, Wind power forecasting based on echo state networks and long short-term memory 11 (3) 526, number: 3 Publisher: Multidisciplinary Digital Publishing Institute. `doi:10.3390/en11030526`. URL `https://www.mdpi.com/1996-1073/11/3/526`

[29] N. Chouikhi, B. Ammar, N. Rokbani, A. M. Alimi, PSO-based analysis of echo state network parameters for time series forecasting 55 211–225. `doi:10.1016/j.asoc.2017.01.049`. URL `https://linkinghub.elsevier.com/retrieve/pii/S1568494617300649`

[30] Q. Li, Z. Wu, R. Ling, M. Tan, Echo state network-based spatio-temporal model for solar irradiance estimation 158 3808–3813, publisher: Elsevier. `doi:10.1016/j.egypro.2019.01.868`. URL `http://www.sciencedirect.com/science/article/pii/S1876610219309105`

[31] J. Pathak, Z. Lu, B. R. Hunt, M. Girvan, E. Ott, Using machine learning to replicate chaotic attractors and calculate lyapunov exponents from data 27 (12) 121102. `arXiv:1710.07313`, `doi:10.1063/1.5010300`. URL `http://arxiv.org/abs/1710.07313`

[32] J. Pathak, B. Hunt, M. Girvan, Z. Lu, E. Ott, Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach 120 (2) 024102, publisher: American Physical Society. `doi:10.1103/PhysRevLett.120.024102`. URL `https://link.aps.org/doi/10.1103/PhysRevLett.120.024102`

[33] P. R. Vlachas, J. Pathak, B. R. Hunt, T. P. Sapsis, M. Girvan, E. Ott, P. Koumoutsakos, Backpropagation algorithms and reservoir computing in recurrent neural networks for the forecasting of complex spatiotemporal dynamics 126 191–217. `doi:10.1016/j.neunet.2020.02.016`. URL `http://www.sciencedirect.com/science/`

article/pii/S0893608020300708

[34] C. Korndörfer, pyESN.
URL https://github.com/cknd/pyESN

[35] E. N. Lorenz, Deterministic nonperiodic flow 20 (2) 130–141, publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences. doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.
URL https://journals.ametsoc.org/view/journals/atsc/20/2/1520-0469_1963_020_0130_dnf_2_0_co_2.xml

[36] AlsoEnergy, University of illinois solar farm dashboard, http://go.illinois.edu/solar.
URL http://go.illinois.edu/solar

[37] N. C. for Environmental Information, Find a station | data tools | climate data online (CDO) | national climatic data center (NCDC).
URL https://www.ncdc.noaa.gov/cdo-web/datatools/findstation

[38] S. Breitweiser, Wind power: University of illinois at urbana-champaign.
URL https://www.fs.illinois.edu/docs/default-source/news-docs/newsrelease_windppa___factsheet.pdf?sfvrsn=43aaffea_0

[39] National solar radiation data base - NSRDB viewer - OpenEI datasets.
URL https://openei.org/datasets/dataset/national-solar-radiation-data-base/resource/b2074dd9-36a4-4382-a12f-e795b578404c

[40] H. E. Garcia, J. Chen, J. S. Kim, M. G. McKellar, W. R. Deason, R. B. Vilim, S. M. Bragg-Sitton, R. D. Boardman, Nuclear hybrid energy systems regional studies: West texas & northeastern arizona. doi:10.2172/1236837.
URL https://www.osti.gov/biblio/1236837-nuclear-hybrid-energy-systems-regional-studies-west-texas-northeastern-arizona

[41] N. US Department of Commerce, ESRL global monitoring laboratory - global radiation and aerosols.
URL https://www.esrl.noaa.gov/gmd/grad/solcalc/calcdetails.html

[42] J. Meeus, Astronomical Algorithms, 2nd Edition, Willmann-Bell, Inc.

[43] P. Kobylinski, M. Wierzbowski, K. Piotrowski, High-resolution net load forecasting for micro-neighbourhoods with high penetration of renewable energy sources 117 105635. doi:10.1016/j.ijepes.2019.105635.
URL http://www.sciencedirect.com/science/article/pii/S0142061518335257

[44] B. Fadlallah, B. Chen, A. Keil, J. Príncipe, Weighted-permutation entropy: A complexity measure for time series incorporating amplitude information 87 (2) 022911, publisher: American Physical Society. doi:10.1103/PhysRevE.87.022911.
URL https://link.aps.org/doi/10.1103/PhysRevE.87.022911

[45] J. Garland, R. James, E. Bradley, Model-free quantification of time-series predictability 90 (5) 052910, publisher: American Physical Society. doi:10.1103/PhysRevE.90.052910.
URL https://link.aps.org/doi/10.1103/PhysRevE.90.052910

[46] F. Pennekamp, A. C. Iles, J. Garland, G. Brennan, U. Brose, U. Gaedke, U. Jacob, P. Kratina, B. Matthews, S. Munch, M. Novak, G. M. Palamara, B. C. Rall, B. Rosenbaum, A. Tabi, C. Ward, R. Williams, H. Ye, O. L. Petchey, The intrinsic predictability of ecological time series and its potential to guide forecasting 89 (2) e01359, _eprint: https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1002/ecm.1359. doi:https://doi.org/10.1002/ecm.1359.
URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/ecm.1359

[47] J. G. Powers, J. B. Klemp, W. C. Skamarock, C. A. Davis, J. Dudhia, D. O. Gill, J. L. Coen, D. J. Gochis, R. Ahmadov, S. E. Peckham, G. A. Grell, J. Michalakes, S. Trahan, S. G. Benjamin, C. R. Alexander, G. J. Dimego, W. Wang, C. S. Schwartz, G. S. Romine, Z. Liu, C. Snyder, F. Chen, M. J. Barlage, W. Yu, M. G. Duda, The weather research and forecasting model: Overview, system efforts, and future directions 98 (8) 1717–1737, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society. doi:10.1175/BAMS-D-15-00308.1.
URL https://journals.ametsoc.org/view/journals/bams/98/8/bams-d-15-00308.1.xml

[48] H. Jaeger, Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication 304 (5667) 78–80. doi:10.1126/science.1091277.
URL https://www.sciencemag.org/lookup/doi/10.1126/science.1091277

[49] M. Lukoševičius, H. Jaeger, Reservoir computing approaches to recurrent neural network training 3 (3) 127–149. doi:10.1016/j.cosrev.2009.03.005.
URL http://www.sciencedirect.com/science/article/pii/S1574013709000173

[50] T. E. Baker, A. S. Epiney, C. Rabiti, E. Shittu, Optimal sizing of flexible nuclear hybrid energy system components considering wind volatility 212 498–508.

[51] H. E. Garcia, J. Chen, J. S. Kim, R. B. Vilim, W. R. Binder, S. M. Bragg Sitton, R. D. Boardman, M. G. McKellar, C. J. J. Paredis, Dynamic performance analysis of two regional nuclear hybrid energy systems 107 234–258. doi:10.1016/j.energy.2016.03.128.
URL http://www.sciencedirect.com/science/article/pii/S0360544216303759