DSCC2015-9784

UNDERSTANDING WIND TURBINE INTERACTIONS USING SPATIOTEMPORAL PATTERN NETWORK

Zhanhong Jiang and Soumik Sarkar[*]
Department of Mechanical Engineering
Iowa State University, Ames, IA 50011
Emails: {zhjiang@iastate.edu, soumiks@iastate.edu}

ABSTRACT

This paper presents a data-driven modeling framework to understand spatiotemporal interactions among wind turbines in a large scale wind energy farm. A recently developed probabilistic graphical modeling scheme, namely the spatiotemporal pattern network (STPN) is used to capture individual turbine characteristics as well as pair-wise causal dependencies. The causal dependency is quantified by a mutual information based metric and it has been shown that it efficiently and correctly captures both temporal and spatial characteristics of wind turbines. The causal interaction models are also used for predicting wind power production by one wind turbine using observations from another turbine. The proposed tools are validated using the Western Wind Integration data set from the National Renewable Energy Laboratory (NREL).

INTRODUCTION

Wind power is a significant alternate source of energy that is clean and pollution-free and accurate wind power prediction is a critical need for the wind energy industry today in order to compete with the more traditional energy sources in the energy market where both lower and surplus energy productions often get penalized. Furthermore, reliable models are required to maximize global energy production. However, due to the stochastic nature and intermittence of wind source, improving modeling accuracy, reliability assessment and characterization of the wind energy is a critical technical challenge. Various methods have been proposed to predict overall farm-wide wind energy produc-

tion that are primarily either physics-based or data-driven methods. Physics-based methods rely on local meteorological service or numerical weather prediction (NWP) [1–4] model data associated with specific geographical information [5] and wind turbine fluid mechanics [6, 7]. However, they often lack scalability and robustness while modeling complex spatiotemporal interactions among a large number of wind turbines. Data-driven methods are also widely used especially for short-term wind power prediction such as, artificial neural networks (ANNs) [8–14], support vector machines [15], Kalman filtering [16] and Bayesian methods [17]. More recently, hybrid methods have been put forward to enhance the accuracy of short-term wind energy prediction, such as ANN and imperialistic competitive algorithm (ICA) [18]; Kalman filtering and NWP [19]; Gaussian process (GP) and NWP [2]; genetic algorithm (GA) and ANN [20]; hybrid neural network and computational fluid dynamics (CFD) [21]; wavelet transform, GA and SVM [22]; ANNs and Bayesian methods [23]; adaptive Bayesian learning and GP [24]. Moreover, He et al. [25, 26] proposed a spatiotemporal analysis framework in which finite state Markov models were developed for wind farm energy generation forecast where a short-term distributional forecasts and a point forecast were derived by using Markov chains and ramp trend information. Similar models were explored by Xie el al. [27] for a short-term spatiotemporal wind power forecast with the Trigonometric Direction Diurnal (TDD) and geostrophic wind information (TDDGW) to improve the forecasting performance. They further integrated the improved performance and economic dispatch framework to improve economic benefits.

Based on the above discussion, it is evident that the current state-of-the-art techniques mostly focus on predicting short-term

---

[*]Address all correspondence to this author.

farm-wide energy production. Therefore, they don't necessarily capture various complex spatiotemporal interactions at local levels, such as turbine-turbine and turbine-wind pattern interactions. While physics-based methods explore such interactions using first principles, it becomes significantly complicated as the number of turbines increases. This is because although physics-based relationships can be constructed for a single turbine or a small number of them, scaling such nonlinear relationships for a large number turbines becomes intractable due to high degree of uncertainty. Furthermore, such techniques may succumb to the *curse of dimensionality*. However, using data-driven methods can effectively avoid directly figuring out the complex relations involving many physical variables. Modeling spatiotemporal interactions is crucial in order to monitor performance and diagnose issues related to individual turbines as well as to develop supervisory control schemes for individual wind turbines at a local level to maximize global energy production. Accurate forecasting of overall farm production is of great importance as it can improve ability to manage dispatchable energy sources and enable greater wind energy penetration. Although it is out of the scope of this paper currently, integrating the improved performance of wind turbines over wind farms by using novel data-driven methods and economic dispatch framework is essential to improve economic benefits and will be the future focus. Note, high-fidelity modeling of isolated single turbine will not be sufficient in this context as the behavior of a turbine changes significantly based on the joint behavior of other turbines and wind patterns [26].

This paper proposes a novel way of capturing spatiotemporal interaction patterns among multiple wind turbines in the form of a recently proposed probabilistic graphical model, called the spatiotemporal pattern network (STPN) [28]. STPN is built on the concept of Symbolic Dynamic Filtering (SDF) that has been shown to be an efficient time-series feature extraction tool for various applications, e.g., parameter estimation in electronic circuits [29], fatigue damage detection in polycrystalline alloys [30, 31], activity recognition [32], fault detection and isolation in nuclear power plants [33], coal gasification systems [34] and aircraft gas turbine engines [35]. While most of the traditional uses of SDF tool focus on modeling individual sub-systems using D-Markov machines [36], the key advancement made in the STPN development was development of a variant, called the xD-Markov machine [28, 35] to capture the causal dependencies between multiple sub-systems. Such models are shown to be sufficient to capture a fairly general class of causal dependencies while correlation-based analysis fails to perform the task [37]. Furthermore, information theoretic measures can be used to quantify the spatiotemporal causality among various sub-systems captured by the the xD-Markov machines. Hence, the STPN technique is highly suited to capture spatiotemporal interaction patterns among multiple wind turbines and the paper validates the proposed method using wind farm data from the

National Renewable Energy Laboratory (NREL) database [38].

The paper is organized in five sections including the present one. In Section II, a brief background on SDF along with D-Markov and xD-Markov machines is provided for completeness. Detail description of STPN construction and information theoretic causality measure in the context of the wind turbine problem is presented in Section III. Section IV presents the validation results along with a comprehensive discussion. The paper is summarized and concluded in Section V with recommendations of future work.

BRIEF BACKGROUND ON SDF AND D-MARKOV MACHINES

This section presents a brief background on SDF and D-Markov machines for completeness. Further details can be found in [39]. SDF is based on the concept of symbolic dynamics where discretization and symbolization are two key steps which yield symbol sequences from observed continuous data. This enables study of dynamical systems in the symbolic space using a language-theoretic approach with tools such as shift-maps and sliding block codes in both deterministic and probabilistic setting. However, although there have been various methods proposed in literature for discretization or partitioning, there is not a single standard way to approach it. In the present study, a supervised multi-variate discretization scheme called the maximally bijective discretization (MBD) [39] is used that was proposed for modeling complex dynamical systems. MBD aims at maximally preserving the input-output relationship present in the continuous domain for a dynamical system.

The discretization of time-series data is followed by the symbolization process and construction of D-Markov machines. An important assumption in SDF is that the symbol sequence generated from a system observation can be approximated by a Markov chain of order D (a positive integer), which is called the D-Markov machine. A D-Markov machine can be used to model the individual time-series data as it captures the temporal characteristics embedded in the symbol sequence. More formally,

Definition 1. *[28] (DFSA) A deterministic finite state automaton (DFSA) is a 3-tuple $G = (\Sigma, Q, \delta)$ where:*

1. *Σ is a non-empty finite set, called the symbol alphabet, with cardinality $|\Sigma| < \infty$;*
2. *Q is a non-empty finite set, called the set of states, with cardinality $|Q| < \infty$;*
3. *$\delta : Q \times \Sigma \to Q$ is the state transition map;*

and $\Sigma^\star$ is the collection of all finite-length strings with symbols from Σ including the (zero-length) empty string ε.

Definition 2. *[28] (PFSA) A probabilistic finite state automaton (PFSA) is constructed upon a DFSA $G = (\Sigma, Q, \delta)$ as a pair $K = (G, \tilde{\Pi})$, i.e., the PFSA K is a 4-tuple $K = (\Sigma, Q, \delta, \tilde{\Pi})$, where:*

1. Σ, Q, and δ are the same as in Definition 1;
2. $\tilde{\Pi} : Q \times \Sigma \rightarrow [0,1]$ is the symbol generation function (also called probability morph function) that satisfies the condition $\sum_{\sigma \in \Sigma} \tilde{\Pi}(q, \sigma) = 1 \ \forall q \in Q$, and $\tilde{\pi}_{ij}$ is the probability of occurrence of a symbol $\sigma_j \in \Sigma$ at the state $q_i \in Q$.

Definition 3. *[28] (D-Markov) A D-Markov machine is a PFSA in which each state is represented by a finite history of D symbols as defined by:*

1. *D is the depth of the Markov machine;*
2. *Q is the finite set of states with cardinality $|Q| \leq |\Sigma|^D$, i.e., the states are represented by equivalence classes of symbol strings of maximum length D where each symbol belongs to the alphabet Σ;*
3. *$\delta : Q \times \Sigma \rightarrow Q$ is the state transition function that satisfies the following condition if $|Q| = |\Sigma|^D$, then there exist $\alpha, \beta \in \Sigma$ and $x \in \Sigma^\star$ such that $\delta(\alpha x, \beta) = x\beta$ and $\alpha x, x\beta \in Q$.*

Remark 4. *It follows from Definition 3 that a D-Markov machine is a statistically stationary stochastic process $S = \cdots s_{-1} s_0 s_1 \cdots$, where the probability of occurrence of a new symbol s_n depends only on the last D symbols, i.e., $P[s_n | s_{n-1} \cdots s_{n-D} \cdots s_0] = P[s_n | s_{n-1} \cdots s_{n-D}]$.*

The details of *D*-Markov machine construction can be found in [28, 36], which may include state splitting that generates symbol blocks of different lengths and state merging that results in the reduction of the number of states.

While $\tilde{\Pi}$ denotes the symbol generation function or the probability morph function, Π denotes the state transition matrix, elements of which captures the transition probabilities from one symbolic state to another. For example, let the symbolic state of wind turbine A at the kth instant be denoted by q_k^A. Then, the ij^{th} matrix element π_{ij}^A of the (stationary) state transition matrix Π^A is the probability of q_{k+1}^A being i provided the past observation q_k^A was j, i.e.,

$$\pi_{ij}^A \triangleq P\left(q_{k+1}^A = i \mid q_k^A = j\right) \text{ for an arbitrary instant } k$$

D-Markov machines can be used for modeling individual sub-systems. Recently, concept of x*D*-Markov machine was proposed to capture the causal dependencies between multiple sub-systems. Such models are shown to be sufficient to capture a fairly general class of causal dependencies while correlation-based analysis fails to perform the task [37]. Formal definition of x*D*-Markov machine is as follows:

Definition 5. *[28] (xD-Markov) Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be the PFSAs corresponding to symbol streams $\{s_1\}$ and $\{s_2\}$ respectively. Then a xD-Markov machine is defined as a 5-tuple $\mathcal{M}_{1 \rightarrow 2} \triangleq (\mathcal{Q}_1, \Sigma_1, \Sigma_2, \delta_1, \tilde{\Pi}_{12})$ such that:*

1. *$\Sigma_1 = \{\sigma_0, ..., \sigma_{|\Sigma_1|-1}\}$ is the alphabet set of symbol sequence $\{s_1\}$*
2. *$Q_1 = \{q_1, q_2, ..., q_{|\mathcal{Q}_1|}\}$ is the state set corresponding to symbol sequence $\{s_1\}$*
3. *$\Sigma_2 = \{\sigma_0, ..., \sigma_{|\Sigma_2|-1}\}$ is the alphabet set of symbol sequence $\{s_2\}$*
4. *$\delta_1 : Q_1 \times \Sigma_1 \rightarrow Q_1$ is the state transition mapping that maps the transition in symbol sequence $\{s_1\}$ from one state to another upon arrival of a symbol in $\{s_1\}$*
5. *$\tilde{\Pi}_{12}$ is the symbol generation matrix of size $|Q_1| \times |\Sigma_2|$; the ij^{th} element of $\tilde{\Pi}_{12}$ denotes the probability of finding the symbol σ_j in the symbol string $\{s_2\}$ while making a transition from the state q_i in the symbol sequence $\{s_1\}$*

From the above definition, it is evident that while a regular *D*-Markov machine encodes the probability of occurrence of a new symbol given the last *D* symbols for one symbol sequence, a x*D*-Markov machine encodes the probability of occurrence of a new symbol in one symbol sequence given the last *D* symbols for another different symbol sequence (possibly from a different sub-system). Thus it can capture the causal effects of one symbol sequence (sub-system) on another symbol sequence (sub-system).

STPN CONSTRUCTION FOR WIND ENERGY FARMS

This section presents the spatiotemporal pattern network (STPN) construction process for the specific problem of understanding wind turbine interactions based on the concepts of SDF outlined above. We begin with data space abstraction that involves data partitioning/discretization and symbolization.

Data Space Abstraction

A wind turbine in a wind energy farm can be seen as a sub-system in a multi-agent complex system. Among various measured variables, two key observations for each wind turbine are wind speed and associated wind power. In the context of symbolic modeling, a wind turbine A is characterized by two sensor observations namely, wind speed and wind power. Note, other variables such as wind direction and humidity that may effect power can be included in the sub-system modeling as well. For the data discretization step, there are many methods that can be applied, such as the symbolic false nearest neighbor partitioning [40], maximum entropy partitioning and uniform partitioning [41]. In this paper, maximally bijective discretization (MBD) [39] is used to maintain the functional relationship between wind speed and wind power present in the continuous domain. In this case, wind speed is discretized as an independent variable and wind power discretization depends on it to maximize bijectivity between discrete states of the two variables. The second part of the abstraction step is symbolization of continuous data using the discretization. The symbolization process converts the temporal evolution of A in the multi-dimensional (2-D in this
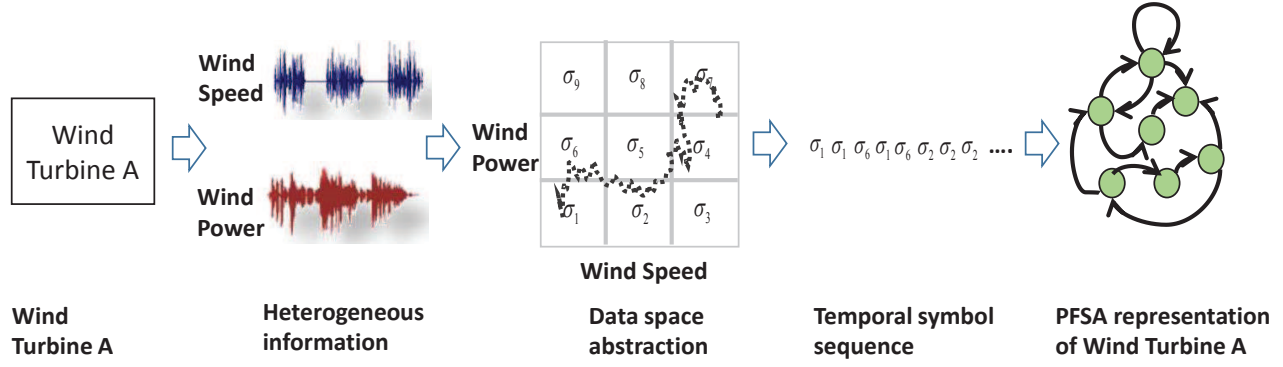
FIGURE 1. *Illustration of Symbolic Dynamic Filtering (SDF) process applied to a wind turbine: Steps to generate a D-Markov machine (PFSA model) using heterogeneous data of wind power and wind speed from a wind turbine*

case) space into a symbol sequence. An illustration is presented in fig. 1: heterogeneous data for wind speed and wind power is obtained from wind turbine *A* and using a discretization policy time-series data are converted into a temporal symbol sequence. Note, although the same discretization may be used for different wind turbines, typically they would correspond to different symbol sequence characteristics. Also, this symbolization process implicitly performs data-level information fusion with wind speed and wind power data.

Symbolic Modeling of Wind Turbines and Interactions

The symbol sequences generated at the abstraction step are modeled in the form of PFSA following the SDF approach as shown in fig. 1. Note, A *D*-Markov machine is a PFSA in which each state is represented by a finite history of *D* symbols. Now we consider the two different wind turbines and quantify their spatiotemporal relations. According to the definition of *D*-Markov machine, the regular state transition matrices Π^A and Π^B represent wind turbines *A* and *B* respectively. Similarly, cross state transition matrices Π^{AB} and Π^{BA} are defined for x*D*-Markov machines representing the causal dependencies of *B* on *A* and of *A* on *B* respectively. Note, that such dependencies may not be symmetric, i.e., Π^{AB} and Π^{BA} are not necessarily the same. Features (e.g., state transition matrix or stationary distribution) from *D*-Markov machines are known as atomic patterns (AP) and those from x*D*-Markov machines are known as relational patterns (RP) [28]. The elements of the cross-state transition matrices Π^{AB} and Π^{BA} are shown as follows:

$$\pi_{k\ell}^{AB} \triangleq P\left(q_{n+1}^B = \ell \mid q_n^A = k\right) \ \forall n$$

$$\pi_{ij}^{BA} \triangleq P\left(q_{n+1}^A = j \mid q_n^B = i\right) \ \forall n$$
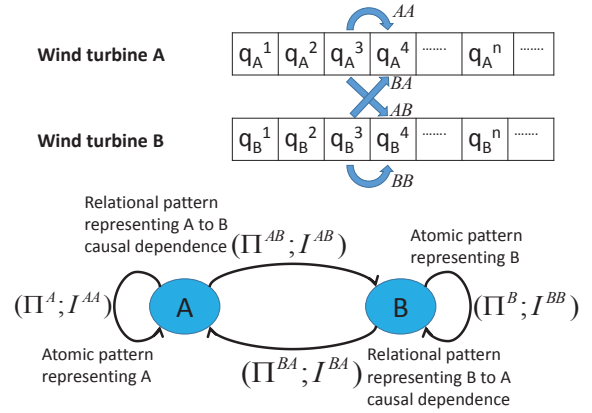

FIGURE 2. *Extraction of atomic and relational patterns (using D-Markov and xD-Markov machines respectively) to characterize individual wind turbine behavior and interaction behavior among different wind turbines*

where $j, k \in Q^A$ and $i, \ell \in Q^B$. The cross-state transition matrix is established from symbol sequences generated from two wind turbines and every element in the matrix identifies the probability of transition from a state in one wind turbine to another state in the second wind turbine.

Furthermore, an information-theoretic measure can be used quantify the information content of the atomic and relational features. As shown in fig. 2, I denotes the mutual information based importance metric for a feature. For example, mutual information for the atomic pattern of wind turbine *A* is expressed as:

$$I^{AA} = I(q_{n+1}^A; q_n^A) = H(q_{n+1}^A) - H(q_{n+1}^A \mid q_n^A)$$

4

where

$$H(q_{n+1}^A) = -\sum_{i=1}^{Q_A} P(q_{n+1}^A = i)\log_2 P(q_{n+1}^A = i)$$

$$H(q_{n+1}^A|q_n^A) = -\sum_{i=1}^{Q_A} P(q_n^A = i)H(q_{n+1}^A|q_n^A = i)$$

$$H(q_{n+1}^A|q_n^A = i) = -\sum_{i=1}^{Q_A} P(q_{n+1}^A = l|q_n^A = i)\cdot$$
$$\log_2 P(q_{n+1}^A = l|q_n^A = i)$$

Note, the quantity I^{AA} signifies the temporal self-prediction capability (self-loop) of the wind turbine A.

Similarly, mutual information for the relational pattern between wind turbines A and B is represented as:

$$I^{AB} = I(q_{n+1}^B; q_n^A) = H(q_{n+1}^B) - H(q_{n+1}^B|q_n^A)$$

where

$$H(q_{n+1}^B|q_n^A) = -\sum_{i=1}^{Q_A} P(q_n^A = i)H(q_{n+1}^B|q_n^A = i)$$

$$H(q_{n+1}^B|q_n^A = i) = -\sum_{i=1}^{Q_B} P(q_{n+1}^B = l|q_n^A = i)\cdot$$
$$\log_2 P(q_{n+1}^B = l|q_n^A = i)$$

Note, the quantity I^{AB} signifies wind turbine A's capability of predicting wind turbine B's outputs and vice versa for I^{BA}. Such mutual information based metric can be used as the weights assigned for the patterns, that can be used for rejecting patterns with low information content (for network pruning) as well as decision fusion. Interested readers can find more details in [28].

As described above, STPN provides an effective tool for learning the spatiotemporal interactions among individual wind turbines. In order to validate the model quality, this paper presents prediction results using the learnt Markov models in both symbolic and continuous domain. In the symbolic domain, most likely symbol sequence for a turbine A given the symbol sequence B can be obtained via running the xD-Markov model
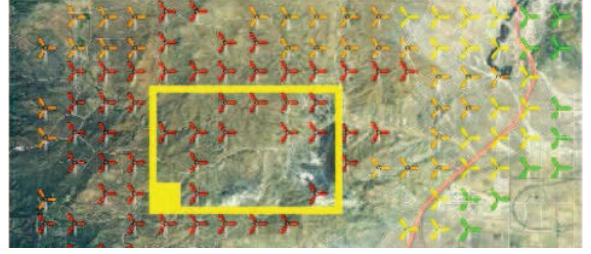

FIGURE 3. *Geographical information for the studied wind turbines: the selected wind turbines are located in the state of california, between 35.28-35.33n and 118.09-118.17w. they are numbered as 1-12 and their power capacities are more than 40%*
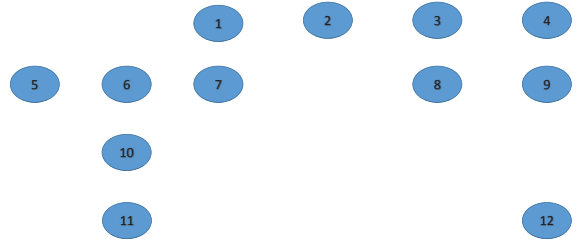

FIGURE 4. *Wind turbines as the sub-system nodes for the abstract spatiotemporal network representation*

(from B to A) a large number of times. The symbolic prediction can be further converted to a continuous domain prediction in an expected sense as follows:

$$E(Power_k) = \sum_{j=1}^{m} Pr_k(j)E(Power|j)$$

where, $E(Power_k)$ is the expected value of power at the kth instant, $Pr_k(j)$ is the probability of jth symbol at the kth instant evaluated using the large number of simulation runs of the learnt Markov models and $E(Power|j)$ is the expected value of power in the discrete bin denoted by symbol j (there are m discrete symbols).

RESULTS AND DISCUSSION

This section presents validation results and discussions based on the Western Wind Integration data set available from the National Renewable Energy Laboratory (NREL) [38]. There are five types of wind turbines based on power capacity: larger than 40%, 35% to 40%, 30% to 35%, 25% to 30% and less than 25%. In this study, 12 wind turbines (located in California) with power capacity larger than 40% are used; IDs: 4494, 4495, 4496,
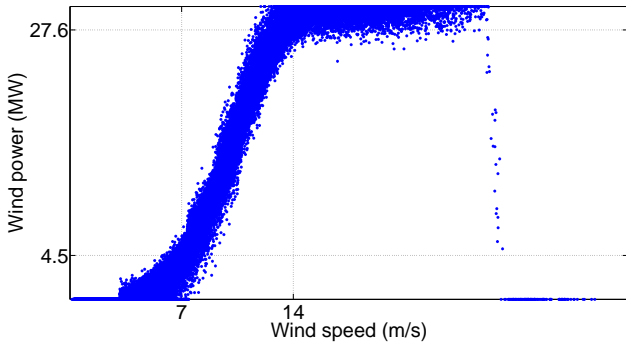
FIGURE 5. *Scatter plot of wind power and wind speed and the partitioning using maximally bijective discretization (MBD)*

FIGURE 7. *Mutual information between multiple wind turbines decreases with the increment of time lag: seven pairs of wind turbines are studied here*
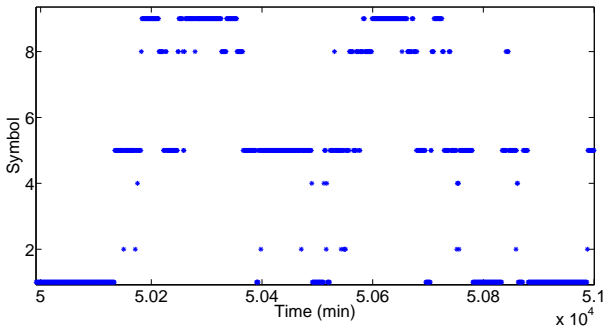

FIGURE 6. *Symbolization of a typical wind turbine time series data*

4497, 4423, 4424, 4425, 4426, 4427, 4361, 4313 and 4314 (numbered 1 - 12) and the power capacities range approximately from 41% to 45%. The wind turbines are located in an area where the annual average wind speed is about 9 m/s, with the elevation of 1019 m - 1207 m. All the conditions for the data set correspond to those of the year 2006.

Figure 3 provides the geographical information of twelve wind turbines distributed in different locations. They are abstracted as nodes of the spatiotemporal pattern network (STPN) as shown in fig. 4 for visualization simplicity. Figure 5 shows the scatter plot with wind power output and wind speed of one wind turbine (other wind turbines also follow nearly the same pattern). The figure further shows the MBD where the partitioning of the wind power axis can be found given the partitioning of the wind speed axis such that their correspondence in the continuous domain is maximally preserved in the symbolic domain. In this context, the discretization of the power curve is tailored to create the symbols in order to find out the spatial and temporal characteristics of wind turbines and interaction patterns between them in the symbolic domain for wind power prediction task. There are nine symbols generated based on the data partitioning. Figure 6 shows an example symbol sequence for a typical wind
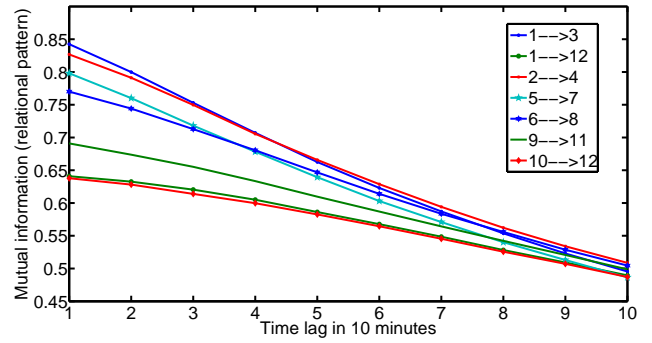
turbine. As it is evident from the discretization plot, most of the symbols are 1, 5, 8 and 9.

The mutual information based causality metric (for the relational pattern) is calculated using the state transition matrices of the xD-Markov machines. As the depth of Markov machine is fixed as 1 in this paper, the current symbol for one wind turbine is only dependent on the past one symbol of another wind turbine. With this setup, we first investigate the effect of time lag on the causal impact between two turbines. As shown in fig. 7, mutual information between any pair of wind turbines decreases along with increment of time lag from 1 to 10. This shows that the choice of depth to be 1 maximizes the causality at time lag 1 for every case.

Next we investigate the effect of spatial distance on the causal dependency between turbines. We choose wind turbines 5, 6, 7, 1 and 10 for this purpose. The results in fig. 8 show that with increase in spatial distance between wind turbines (along any direction - latitude or longitude), causality quantified by mutual information of the relational pattern decreases. Figure 9 shows the general decreasing trend of mutual information based causality metric and the Euclidean distance between pairs of wind turbines. Based on the above two observations, it is evident that the mutual information based causality metric is able to capture both temporal and spatial characteristics.

The efficacy of the xD-Markov machines in capturing causal dependencies is further evaluated using prediction of wind power production by one wind turbine using observed symbol sequence of another turbine. Using the prediction process as described in the previous section, fig. 10 and fig. 11 show the comparisons between predicted and actual symbol sequences of wind turbine 5 using wind turbines 6 and 7 respectively. The results are obtained by using half year data of 2006 to train and then using the rest of the year data for testing. While the plots show strong prediction capability of xD-Markov machines, most of the errors come from the transient symbols as expected. Furthermore, vi-
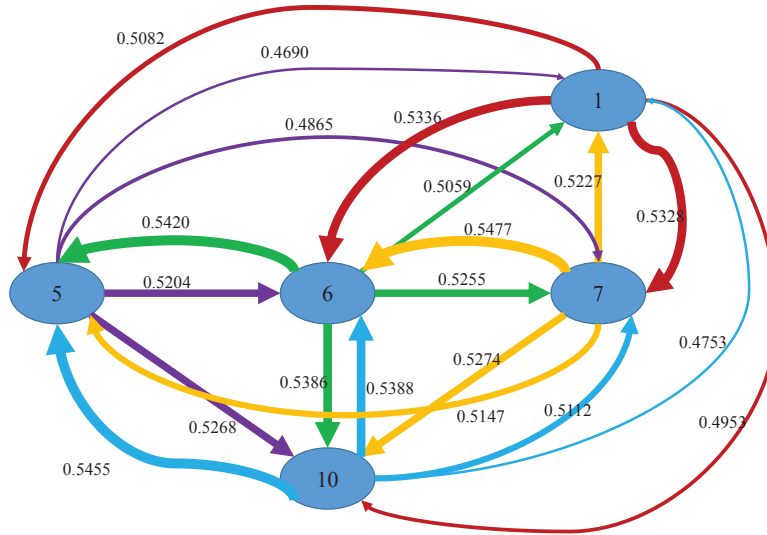
FIGURE 8. *Mutual information of relational pattern between pairs of wind turbines at different locations: overall observation is that Causality decreases with increase in spatial distance*
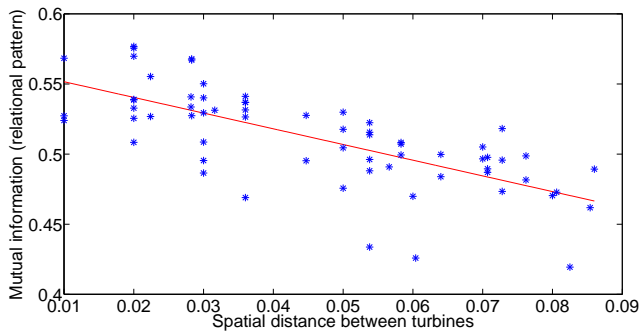

FIGURE 9. *Mutual information of relational pattern decreases with increase in Euclidean distance between turbine pairs*
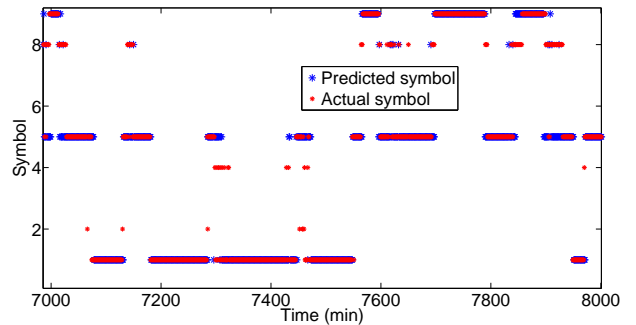

FIGURE 10. *Prediction of wind turbine 5 power using wind turbine 6 observation in the symbolic domain using the half year data of 2006 for training and the rest of the year data for testing*

sual inspection can verify that turbine 7 performs slightly worse compared to predicting power generation by turbine 6 as suggested by the mutual information based causality metric. Continuing the same prediction process with turbines 8 and 9, fig. 12 shows the monotonic increase in mean square errors (MSE) as spatial distance between turbine pairs increases (or mutual information decreases). Finally, an example of continuous domain prediction (as discussed in the previous section) of wind turbine 5 power generation (using observed symbol sequence for wind turbine 6) is shown in fig. 13. The plot shows that even with significant coarse graining of the data space, the continuous domain prediction catches the major trends in data quite well. This can be attributed to the efficacy of the MBD scheme as it aims to preserve the functional relationship present in the continuous domain while discretizing. While finer discretization improves

prediction quality in the continuous domain, it increases model and computational complexity as well as the need for data.

SUMMARY, CONCLUSIONS AND FUTURE WORK

This paper has proposed the use of a novel spatiotemporal pattern network (STPN) framework to capture the interaction characteristics between multiple wind turbines. While the discretization and symbolization steps of SDF performs data level fusion of wind power and wind speed for a single wind turbine system, a *D*-Markov machine captures its stationary temporal dynamics. Causal dependency between two turbines is modeled using a variant called the x*D*-Markov machine. Moreover, the causal dependency is quantified by a mutual information based
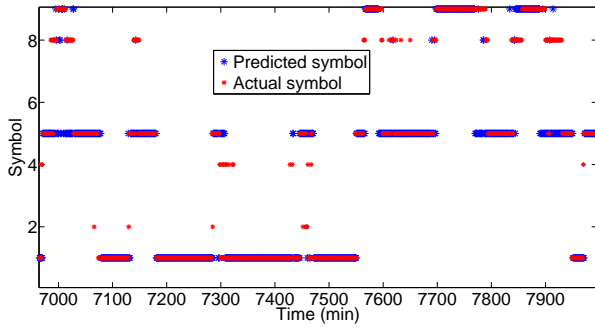
FIGURE 11. *Prediction of wind turbine 5 power using wind turbine 7 observation in the symbolic domain with same training and testing data sets as in fig. 10*
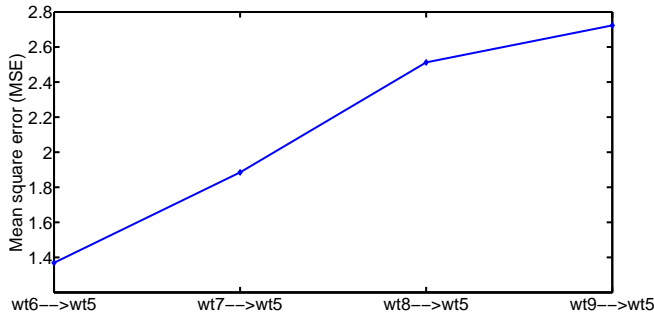

FIGURE 13. *Expected continuous domain prediction of wind power for wind turbine 5 using observed symbol sequence from wind turbine 6 with same training and testing data sets as in fig. 10*

FIGURE 12. *Prediction MSEs increase as spatial distance between turbine pairs increases or mutual information decreases*

metric. The proposed scheme is validated using the Western Wind Integration data set from NREL. The primary observation made is that the Markov machines and the mutual information based causality metric are able to capture both temporal and spatial characteristics as causality decreases with increase in both temporal lags and spatial distances. The STPN scheme is further evaluated using prediction of wind power production by one wind turbine using observed symbol sequence from another turbine. Some of the future research directions currently being pursued are:

1. Impact analysis of other physical variables, e.g., wind direction on model quality;
2. Systematic (short and long term) farm-wide wind power prediction using STPN;
3. Simulating "what if" scenarios using STPN for distributed optimization of farm-wide wind power production.
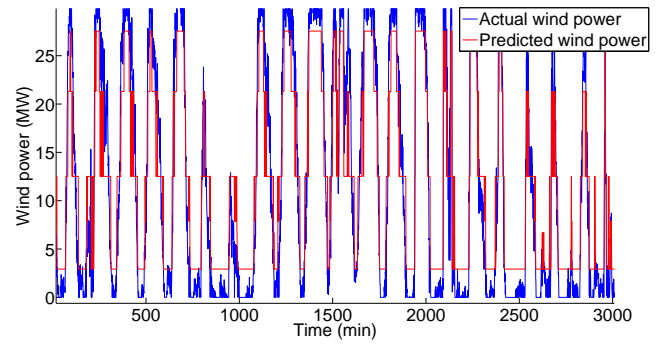
REFERENCES

[1] Nor, K. M., Shaaban, M., and Rahman, H. A., 2014. "Feasibility assessment of wind energy resources in malaysia based on nwp models". *Renewable Energy,* **62**, pp. 147–154.
[2] Chen, N., Qian, Z., Nabney, I. T., and Meng, X., March,2014. "Wind power forecasts using gaussian process and numerical weather prediction". *Power Systems, IEEE Transactions on,* **29**(2).
[3] De Giorgi, M. G., Ficarella, A., and Tarantino, M., 2011. "Assessment of the benefits of numerical weather predictions in wind power forecasting based on statistical methods". *Energy,* **36**, pp. 3968–3978.
[4] Vladislavleva, E., Friedrich, T., Neumann, F., and Wagner, M., 2013. "Predicting the energy output of wind farms based on weather data: Important variables and their correlation". *Renewable Energy,* **50**, pp. 236–243.
[5] Heimiller, D., and Haymes, S., January, 2011. "Geographic information systems in support of wind energy activities at nrel". *in Proceedings of 39th AIAA Aerospace Sciences Meeting, Reno, Nevada*.
[6] Stevens, R., Gayme, D. F., and Meneveau, C., 2015. "Coupled wake boundary layer model of wind-farms". *J. Renew. Sust. Energy,* **7, 023115**.
[7] Tian, L. L., Zhu, W. J., Shen, W. Z., Srensen, J. N., and Zhao, N., 2014. "Investigation of modified ad/rans models for wind turbine wake predictions in large wind farm". *Journal of Physics: Conference Series,* **524**(1), p. 012151.
[8] Li, S., Wunsch, D. C., O'Hair, E., and Giesselmann, M. G., 2001. "Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation". *Journal of Solar Energy Engineering,* **123**, pp. 327–332.
[9] Ren, C., An, N., Wang, J., Li, L., Hu, B., and Shang, D., 2014. "Optimal parameters selection for bp neural network

on particle swarm optimizaiton: A case study of wind speed forecasting". *Knowledge-Based Systems, 56*, pp. 226–239.

[10] Hong, C.-M., Cheng, F.-S., and Chen, C.-H., 2014. "Optimal control for variable-speed wind generation systems using general regression neural network". *Electrical Power and Energy Systems, 60*, pp. 14–23.

[11] Sideratos, G., and Hatziargyriou, N. D., 2012. "Probabilistic wind power forecasting using radial basis function neural networks". *Power Systems, IEEE Transactions on, 27*(4), pp. 1788–1796.

[12] Mabel, M. C., and Fernandez, E., 2008. "Analysis of wind power generation and prediction using ann: A case study". *Renewable Energy, 33*, pp. 986–992.

[13] Kelouwani, S., and Agbossou, K., September, 2004. "Nonlinear model identification of wind turbine with a neural network". *Energy Conversion, IEEE Transactions on, 19*(3).

[14] Kramer, O., Gieseke, F., and Satzger, B., 2013. "Wind energy prediction and monitoring with neural computation". *Neurocomputing, 109*, pp. 84–93.

[15] Salcedo-Sanz, S., Ortiz-Garcia, E. G., Perez-Bellido, A. M., Portilla-Figueras, A., and Prieto, L., 2011. "Short-term wind speed prediction based on evolutionary support vector regression algorithms". *Expert Systems with Applications, 38*, pp. 4052–4057.

[16] Poncela, M., Poncela, P., and Peran, J. R., 2013. "Automatic tuning of kalman filters by maximum likelihood methods for wind energy forecasting". *Applied Energy, 108*, pp. 349–362.

[17] Li, G., and Shi, J., 2012. "Applications of bayesian methods in wind energy conversion systems". *Renewable Energy, 43*, pp. 1–8.

[18] Imaie, E., Sheikholeslami, A., and Ahangar, R. A. "Improving short-term wind power prediction with neural network and ica algorithm and input feature selection". *Jounal of Advances in Computer Research, 5*(3).

[19] Cassola, F., and Burlando, M., 2012. "Wind speed and wind energy forecast through kalman filtering of numerical weather prediction model output". *Applied Energy, 99*, pp. 154–166.

[20] Koihe, M., Lin, T. C., and Maunuksela, J., March, 2011. "Ga-ann for short-term wind energy prediction". *in Proceedings of Power and Energy Engineering Conference, 2011 Asia-Pacific*, pp. 1–6.

[21] Castellani, F., Burlando, M., Taghizadeh, S., Astolfi, D., and Piccioni, E., 2014. "Wind energy forecast in complex sites with a hybrid neural network and cfd based method". *Energy Procedia, 45*, pp. 188–197.

[22] Shi, J., Liu, Y., Yang, Y., and Lee, W., May, 2011. "Short-term wind power prediction based on wavelet transform-support vector machine and statistic characteristics analysis". *in Proceedings of Industrial and Commercial Power Systems Technical Conference, 2011 IEEE*, pp. 1–7.

[23] Blonbou, R., 2011. "Very short-term wind power forecasting with neural networks and adaptive bayesian learning". *Renewable Energy, 36*, pp. 1118–1124.

[24] Blonbou, R., Monjoly, S., and Jean-Francois, D., 2011. "An adaptive short-term prediction for wind energy storage management". *Energy Conversion and Management, 52*, pp. 2412–2416.

[25] He, M., Yang, L., Zhang, J., and Vittal, V., 2013. "Spatiotemporal analysis for smart grids with wind generation integration". *in Proceedings of 2013 International Conference on Computing, Networking and Communications*, pp. 1107–1111.

[26] He, M., Yang, L., Zhang, J., and Vittal, V., July 2014. "A spatio-temporal analysis approach for short-term forecast of wind farm generation". *Power Systems, IEEE Transactions on, 29*, pp. 1611–1622.

[27] Xie, L., Gu, Y., Zhu, X., and Genton, M. G., January, 2014. "Short-term spatio-temporal wind power forecast in robust look-ahead power system dispatch". *Smart Grid, IEEE Transactions on, 5*(1), pp. 511–520.

[28] Sarkar, S., Sarkar, S., Virani, N., Ray, A., and Yasar, M., 2014. "Sensor fusion for fault detection & classification in distributed physical processes". *frontiers in Robotics and AI - Sensor Fusion and Machine Perception*.

[29] Rao, C., Ray, A., Sarkar, S., and Yasar, M., 2009. "Review and comparative evaluation of symbolic dynamic filtering for detection of anomaly patterns". *Signal, Image and Videl Processing, 3*, pp. 101–114.

[30] Gupta, S., Ray, A., and Keller, E., 2007. "Symbolic time series analysis of ultrasonic data for early detection of fatigue damage". *Mechanical Systems and Signal Processing, 21*, pp. 866–884.

[31] Sharan Singh, D., Gupta, S., and Ray, A., 2010. "Symbolic dynamic analysis of surface deformation during fatigue crack initiation". *Measurement Science and Technology, 21*, pp. 1–7.

[32] Jin, X., Gupta, S., Ray, A., and Damarla, T., July, 2011. "Symbolic dynamic filtering of seimic sensors for target detection and classification". *in Proceedings of American Control Conference, San Francisco, CA, USA*, pp. 5151–5156.

[33] Jin, X., Guo, Y., Sarkar, S., Ray, A., and Edwards, R. M., February, 2011. "Anomaly detection in nuclear power plants via symbolic dynamic filtering". *Nuclear Science, IEEE Transactions on, 58*(1).

[34] Chakraborty, S., Sarkar, S., Gupta, S., and Ray, A., October, 2008. "Damage monitoring of refractory wall in a generic entrained-bed slagging gasification system". *Proceedings of I Mech E Part A: Journal of Power and Energy, 222*(8), pp. 791–807.

[35] Sarkar, S., Sarkar, S., Mukherjee, K., Ray, A., and Srivas-

tav, A., 2012. "Multi-sensor information fusion for fault detection in aircraft gas turbine engines". *Proc IMechE Part G: J Aerospace Engineering,* **227**, pp. 1988–2001.

[36] Gupta, S., and Ray, A., 2007. "Symbolic dynamic filtering for data-driven pattern recognition". *In: Pattern Recognition: Theory and Application*, pp. 1–71.

[37] Chattopadhyay, I., 25 Jun, 2014. "Causality network". *arXiv:1406.6651v1[cs.LG]*.

[38] http://wind.nrel.gov/webnrel/.

[39] Sarkar, S., Srivastav, A., and Shashanka, M., June, 2013. "Maximally bijective discretization for data-driven modeling of complex systems". *in Proceedings of Americal Control Conference, Wahsington, DC, USA*, pp. 2674–2679.

[40] Buhl, M., and Kennel, M., 2005. "Statistically relaxing no generating partitions for observed time-series data". *Physical Review E,* **71**(4).

[41] Sarkar, S., Mukherjee, K., and Ray, A., June, 2009. "Generalization of hilber transform of symbolic analysis of noisy signals". *Signal Processing,* **89**, pp. 1245–1251.