Short-Term Wind Forecasting Using Spatio-Temporal Covariance Models

Tianxia Jia

Supervisor: Dr. Deniz Sezer
Department of Mathematics and Statistics
University of Calgary

Workshop on Forecasting and Mathematical Modeling for Renewable Energy

July 26, 2023

Table of Contents

Table of Contents

UNIVERSITY OF CALGARY

Background

Why is wind speed/power forecasting important?

- It helps to integrate wind power effectively into the energy system.
- It enables system operators to anticipate fluctuations in wind power generation and maintain grid stability and reliability.
- It enables market participants to make informed decisions and manage risks associated with the variability of wind power generation.
- In general, it can ultimately reduce the levelized cost of energy and increase the overall system value.

Background

What are the tools and techniques are used for wind forecasting?

- Time series models
  - Browell and Gilbert (2017) used regime-switching (RS) AR to predict daily wind power.
  - Browell et al. (2018) adopted RS AR and VAR models with exogenous variables to predict wind speed at multiple locations.
  - Jia, Sezer, & Wood (2022) applied RS mixture AR and VAR models to predict wind speed at 23 weather stations in Alberta.
- Covariance models
  - Gneiting (2002); Gneiting et al. (2006) proposed a parametric family of covariance models. The prevailing wind was considered as fixed in a Lagrangian framework.
  - Ezzat et al. (2018, 2019) proposed a calibrated RS covariance method where the prevailing wind was considered a random variable.

Background

Covariance models:

- Tobler's law: everything is related to everything else, but near things are more related than distant things.
- They are used when wind is modeled by spatio-temporal (ST) Gaussian processes which are stationary and completely determined by mean and covariance functions.
- The spatio-temporal covariance functions must be positive semi-definite (p.s.d.) and there are limited choices: e.g. Matérn class, Gneiting class, etc.
- We study the Gneiting class and 1) introduce a new covariance function for accounting asymmetry, 2) allow the model to be regime-switching, 3) relax the constraint of the interaction parameter.

UNIVERSITY OF CALGARY

Table of Contents

ST Gaussian processes

- A ST process $\{Y(\mathbf{x}, t) : \mathbf{x} \in \mathbf{S}, t \in [0, T]\}$ is Gaussian if for any $n \geq 1$ and $(\mathbf{x}_1, t_1), \ldots, (\mathbf{x}_n, t_n)$, $Y(\mathbf{x}_1, t_1), \ldots, Y(\mathbf{x}_n, t_n)$ has multivariate Gaussian distribution with

$$\mathrm{E}[Y(\mathbf{x}_i, t_i)] = \mu(\mathbf{x}_i, t_i) \tag{1}$$

$$\mathrm{Cov}(Y(\mathbf{x}_i, t_i), Y(\mathbf{x}_j, t_j)) = C(\mathbf{x}_i, \mathbf{x}_j; t_i, t_j) \tag{2}$$

for some functions

$$\mu : \mathbf{S} \times [0, T] \to \mathbb{R},$$
$$C : (\mathbf{S} \times [0, T]) \times (\mathbf{S} \times [0, T]) \to \mathbb{R}.$$

ST Gaussian processes

- Often we need second-order stationarity and isotropy to make inference on the function C:

$$\text{Cov}(Y(\mathbf{x}_1, t_1), Y(\mathbf{x}_2, t_2)) = C(\|\mathbf{x}_1 - \mathbf{x}_2\|, |t_1 - t_2|) \tag{3}$$
$$= C(\|\mathbf{h}\|, |u|). \tag{4}$$

- It is difficult to come up with a stationary ST covariance function from scratch. Hence we use known parametric covariance functions.

Table of Contents

UNIVERSITY OF CALGARY

Gneiting Class

- **Gneiting class** (Gneiting, 2002) is built on top of purely spatial and purely temporal correlation functions.
- Purely spatial model:

$$C_S(\mathbf{h}) = (1-\nu)\exp(-c\|\mathbf{h}\|^{2\gamma}) + \nu\delta_{\mathbf{h}=0}, \tag{5}$$

  where $\delta_{\mathbf{h}=0} = 1$ if $\mathbf{h} = 0$ and 0 o.w., $\nu \in [0,1], c > 0, \gamma \in (0, 1/2]$.
- Purely temporal model:

$$C_T(u) = (1 + a|u|^{2\alpha})^{-1}, \tag{6}$$

  where $a > 0$ and $\alpha \in (0, 1]$.

Gneiting Class

- Separable model:

$$C_{\mathsf{SEP}}(\mathbf{h}, u) = C_{\mathsf{S}}(\mathbf{h}) C_{\mathsf{T}}(u) \tag{7}$$

$$= \left[(1-\nu)\exp(-c\|\mathbf{h}\|^{2\gamma}) + \nu\delta_{\mathbf{h}=0}\right](1 + a|u|^{2\alpha})^{-1}. \tag{8}$$

- Fully symmetric model:

$$C_{\mathsf{FS}}(\mathbf{h}, u) = \frac{1-\nu}{1 + a|u|^{2\alpha}}\left(\exp\left(\frac{c\|\mathbf{h}\|^{2\gamma}}{(1 + a|u|^{2\alpha})^{\beta\gamma}}\right) + \frac{\nu}{1-\nu}\delta_{\mathbf{h}=0}\right), \tag{9}$$

where $\beta \in [0, 1]$.

Asymmetric Covariance Functions - Lagrangian Framework

- Lagrangian reference frame is used to achieve asymmetry in correlation due to transport effects of prevailing wind flows.
- For a spatial random field on $\mathbb{R}^d$ with stationary covariance function C_S, suppose the entire field moves time-forward with velocity $\mathbf{V} \in \mathbb{R}^d$, then the resulting ST random process has the covariance

$$C(\mathbf{h}, u) = \mathsf{E}_{\mathbf{V}}[C_S(\mathbf{h} - \mathbf{V}u)]. \tag{10}$$

- The above function is generally asymmetric.

Asymmetric Covariance Functions - Lagrangian Framework

- General stationary model (Gneiting et al., 2006):

$$C_{\mathrm{STAT}}(\mathbf{h}, u) = (1-\lambda) C_{\mathrm{FS}}(\mathbf{h}, u) + \lambda C_{\mathrm{LGR}}(\mathbf{h}, u), \tag{11}$$

$$C_{\mathrm{LGR}}(\mathbf{h}, u) = C_{\mathrm{S}}(\mathbf{h} - \mathbf{V} u). \tag{12}$$

- Lagrangian correlation functions:
  - Triangular:

$$\left(1 - \frac{1}{2v} |\mathrm{h}_1 - vu| \right)_+, \tag{13}$$

  - New:

$$\left(1 - \frac{1}{2\|\boldsymbol{v}\|} \left| \frac{\mathbf{h}^\top \boldsymbol{v}}{\|\boldsymbol{v}\|} - \|\boldsymbol{v}\| u \right| \right)_+, \tag{14}$$

  - Askey (Askey, 1973):

$$\left(1 - \frac{1}{2\|\boldsymbol{v}\|} \|\mathbf{h} - \boldsymbol{v} u\| \right)_+^{3/2}. \tag{15}$$

Covariance Models - Regime-Switching

- The regime-switching models show better performance in wind forecasting than traditional time-series and covariance models.
- Regime-switching general stationary model:

$$C_{\mathrm{STAT},k}(\mathbf{h}, u) = (1 - \lambda_k) C_{\mathrm{FS}}(\mathbf{h}, u) + \lambda_k C_{\mathrm{LGR},k}(\mathbf{h}, u_k), \tag{16}$$

$$C_{\mathrm{LGR},k}(\mathbf{h}, u) = \left(1 - \frac{1}{2\|\boldsymbol{v}_k\|} \left| \frac{\mathbf{h}^\top \boldsymbol{v}_k}{\|\boldsymbol{v}_k\|} - \|\boldsymbol{v}_k\| u_k \right| \right)_+, \tag{17}$$

where k is the regime index.

Table of Contents

Markov Chain Gaussian Field

- A more practical framework is a p-th order Markov Chain Gaussian Field (MCGF) where time is discrete and we only model covariances up to time lag p. The system state is updated one time-step at a time using a transition probability conditional on the history up to p time lags.
- Markov chain Gaussian field: $\{\boldsymbol{Y}(\mathbf{S}, t) : t \in [0, \infty)\}$.
  - For a Gaussian process defined on field $\mathbf{S}$ with time lag p,

$$P(\boldsymbol{Y}(\mathbf{S}, t) | \boldsymbol{Y}(\mathbf{S}, t-1), \ldots, \boldsymbol{Y}(\mathbf{S}, 0)) = P(\boldsymbol{Y}(\mathbf{S}, t) | \boldsymbol{Y}(\mathbf{S}, t-1), \ldots, \boldsymbol{Y}(\mathbf{S}, t-p)). \quad (18)$$

  - It is different from Gaussian Markov random fields where the Markov property is applied spatially on $\mathbf{S}$.

Markov Chain Gaussian Field

Reasons of considering MCGF:

- For short-term wind speed modeling, future relies on short time lags.
- For RS models, defining the transition probabilities is more elegant in a Markovian framework.
- The covariance functions only need to be valid for temporal domain $[-p, p]$, instead of $\mathbb{R}$.

Covariance Models

- Fully symmetric model:

$$C_{\mathsf{FS}}(\mathbf{h}, u) = \frac{1-\nu}{1+a|u|^{2\alpha}} \left(\exp\left(\frac{c\|\mathbf{h}\|^{2\gamma}}{(1+a|u|^{2\alpha})^{\beta\gamma}} \right) + \frac{\nu}{1-\nu} \delta_{\mathbf{h}=0} \right), \tag{19}$$

where $\beta \in [0, 1]$.

- During parameter estimation, we found that the constraint $\beta \leq 1$ is binding and results will be improved if we allow $\beta > 1$.
- While letting $\beta > 1$ may not give a valid correlation function, we can prove that the constraints can be relaxed under certain conditions.

Covariance Models

- Goal: allow $\beta > 1$ in C_{FS}; that is, prove that C_{FS} gives a valid covariance matrix for **S** when $\beta > 1$.
- Steps:
  1. Let $\mathbf{G} = \{\mathbf{g}_1, \ldots, \mathbf{g}_m\}$ be a finite discrete rectangular grid that covers **S**.
  2. Let C be a function on $\mathbf{G} \times \mathbf{G} \times [-p, p]$, and we build the block-Toeplitz matrix $\boldsymbol{\Sigma}$ consisting of blocks

$$\Sigma_{i,j,u} = C(\mathbf{g}_i, \mathbf{g}_j, u).$$

  We assume:

  1. C is non-negative.
  2. $\boldsymbol{\Sigma}$ is p.s.d. This can be verified numerically.
  3. $C(\mathbf{g}_i, \mathbf{g}_i, |u|) = C_{\mathsf{T}}(u)$ is a p.s.d. function on $\{-p, -p+1, \ldots, p\}$ such that $C_{\mathsf{T}}(u) = C_{\mathsf{T}}(-u)$ and $C_{\mathsf{T}}(0) = 1$.

UNIVERSITY OF CALGARY

Covariance Models

Theorem

Under Assumptions 1-3, the following function is a valid covariance function for MCGF $\boldsymbol{W}(\cdot, t)$.

$$C_M(\mathbf{x}_1, \mathbf{x}_2, t_1, t_2) = \begin{cases} C_T(|t_1 - t_2|), & \text{if } \mathbf{x}_1 = \mathbf{x}_2 \\ (1-\nu)C(G(\mathbf{x}_1), G(\mathbf{x}_2), |t_1 - t_2|), & \text{otherwise,} \end{cases}$$

where $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{S}$ *and* $t_1, t_2 \in \{t-p, \ldots, t\}$. *That is, for any* $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subseteq \mathbf{S}$, *and any time lag* $u \in \{0, 1, \ldots, p\}$, C_M *gives a valid covariance matrix for the joint distribution of* $\boldsymbol{w} = (\boldsymbol{W}(\mathbf{X}, t), \ldots, \boldsymbol{W}(\mathbf{X}, t-u))$.

UNIVERSITY OF CALGARY

The proof is based on the following construction of a ST process. Let $\mathbf{Y}(\cdot, t)$ be a Gaussian process with covariance matrix $\mathbf{\Sigma}$ on $G(\mathbf{X})$. We let

$$\widetilde{\boldsymbol{W}}(\mathbf{x}, t) = \boldsymbol{Y}(G(\mathbf{x}), t) + \boldsymbol{Z}(\mathbf{x}, t), \tag{20}$$

where $\boldsymbol{Z}(\mathbf{x}, t)$ is a ST Gaussian process with no spatial correlation and temporal correlation given by $\nu C_{\mathrm{T}}(u)$. Then $\{\widetilde{\boldsymbol{W}}(\mathbf{x}, t)\}$ is a valid ST process and its covariance matrix is given by C_M. Therefore, C_M must be positive semi-definite.

Remarks:

- Function C can be any functions that satisfy Assumptions 1-3.
- Example:
  - Let

$$C(\mathbf{h}, u) = \frac{1}{1 + a|u|^{2\alpha}} \exp\left(\frac{c\|\mathbf{h}\|^{2\gamma}}{(1 + a|u|^{2\alpha})^{\beta\gamma}} \right), \tag{21}$$

then

$$C_M = \begin{cases} (1 + a|u|^{2\alpha})^{-1}, & \text{if } \mathbf{h} = \mathbf{0} \\ \dfrac{1-\nu}{1 + a|u|^{2\alpha}} \exp\left(\dfrac{c\|\widetilde{\mathbf{h}}\|^{2\gamma}}{(1 + a|u|^{2\alpha})^{\beta\gamma}} \right), & \text{otherwise,} \end{cases} \tag{22}$$

is the fully symmetric model.

UNIVERSITY OF CALGARY

Covariance Models - Summary

- Propose a new correlation function to account both prevailing wind speed and direction.

$$C_{\mathrm{LGR}}(\mathbf{h}, u) = \left(1 - \frac{1}{2\|\boldsymbol{v}\|}\left|\frac{\mathbf{h}^\top \boldsymbol{v}}{\|\boldsymbol{v}\|} - \|\boldsymbol{v}\| u\right|\right)_+ . \qquad (23)$$

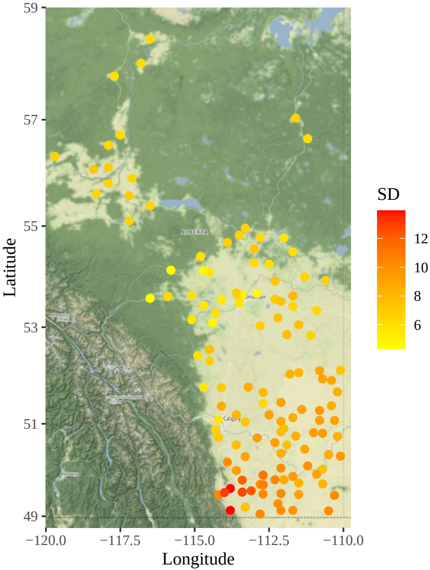
- Expend the covariance models to be regime-switching.

$$C_{\mathrm{STAT},k}(\mathbf{h}, u) = (1 - \lambda_k) C_{\mathrm{FS}}(\mathbf{h}, u) + \lambda_k C_{\mathrm{LGR},k}(\mathbf{h}, u_k). \qquad (24)$$

- Allow β to be higher than 1.

Table of Contents

- We fit above models to predict hourly wind speed at 131 Alberta weather stations for up to 6-h ahead.
- Time span: 6-year
  - Training: 2016-2019
  - Test: 2020-2021
- To fit regime-switching models, we identify atmospheric regimes with ERA5 database on four atmospheric variables (Browell et al., 2018) using k-means and hidden Markov models.



Wind Speed Forecasting

- Candidate models:
  - Empirical model
  - Fully symmetric model (FS) and that with relaxed constraints on β (FS2)
  - Regime-switching general stationary model (STAT) and that with relaxed constraints on β (STAT2)
  - For regime-switching models, we also obtain soft forecasts: the weighted averages of forecasts under each regime.

Modeling Results - Summary

- All parametric models are better than the empirical model, on average.
- The STAT2 model with soft forecast under HMM performs the best, followed by the FS2 model.
- Compared with the FS model, all general stationary models that take regime-dependent prevailing wind into consideration perform better.
- Considering atmospheric information and prevailing wind is not beneficial for short forecast horizons (1-2 h), and the Empirical model works the best.

Modeling Results

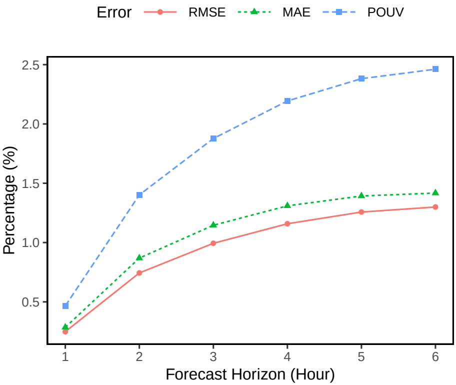


Figure: Percentage Improvement of STAT-soft Under HMM Over FS

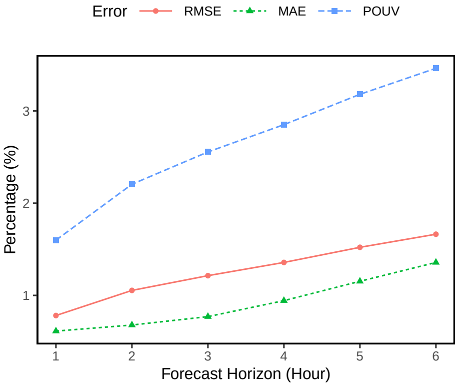


Figure: Percentage Improvement of FS2 Over FS

Modeling Results - Station

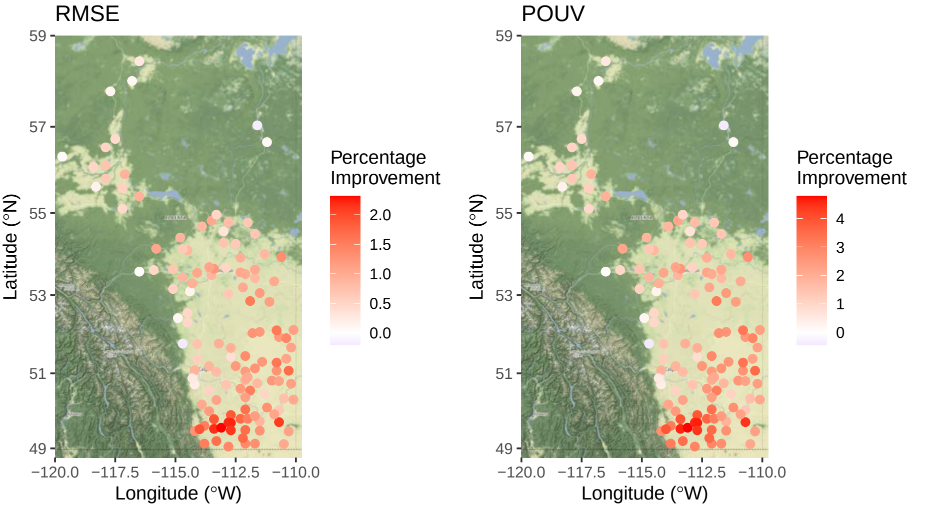


Figure: Percentage Improvement of STAT-soft Over FS by Weather Stations.

Modeling Results - Station

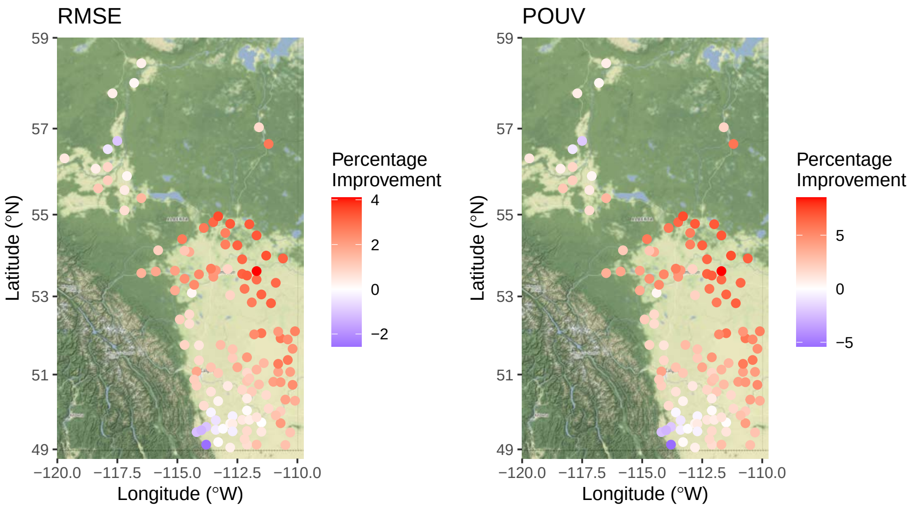


Figure: Percentage Improvement of FS2 Over The FS by Weather Stations.

Future Work

- Allow prevailing wind speed and direction to vary by time and location.
- Allow different models for different areas.
- Propose new models for areas with more complex wind patterns.

Future Work

- We are writing an R package called `mcgf` to
  - simulate MCGF with aforementioned (regime-switching) covariance structures,
  - estimate parameters of the covariance functions by weighted least squares or maximum likelihood estimation.
  - predict future values by simple kriging.
- We have finished simulation and parameter estimation parts of the package, and now we are finishing the rest.
- The package is available on: `https://github.com/Tianxia-Jia/mcgf`.

UNIVERSITY OF CALGARY

References I

Askey, R. (1973). Radial characteristic functions. *Wisconsin University Madison Mathematics Research Center*.

Browell, J., Drew, D. R., and Philippopoulos, K. (2018). Improved very-short-term spatio-temporal wind forecasting using atmospheric regimes. *Wind Energy*, 21(11):968–979.

Browell, J. and Gilbert, C. (2017). Cluster-based regime-switching ar for the eem 2017 wind power forecasting competition. pages 1–6.

Ezzat, A., Ahmed, Jun, Mikyoung, and Ding, Y. (2018). Spatio-temporal asymmetry of local wind fields and its impact on short-term wind forecasting. *IEEE Transactions on Sustainable Energy*, 9(3):1437–1447.

Ezzat, A. A., Jun, M., and Ding, Y. (2019). Spatio-temporal short-term wind forecast: A calibrated regime-switching method. *The Annals of Applied Statistics*, 13(3):1484–1510.

References II

Gneiting, T. (2002). Nonseparable, stationary covariance functions for space–time data. *Journal of the American Statistical Association*, 97(458):590–600.

Gneiting, T., Genton, M. G., and Guttorp, P. (2006). Geostatistical space-time models, stationarity, separability and full symmetry. Technical report, Department of Statistics, University of Washington.

Jia, T., Sezer, D., and Wood, D. (2022). Short-term wind speed forecasting with regime-switching and mixture models at multiple weather stations over a large geographical area. *Journal of Renewable and Sustainable Energy*, 14(4):043305.

UNIVERSITY OF CALGARY

Thank You!

UNIVERSITY OF CALGARY