

# Spatio-temporal modeling of wind power generation of Alberta

Alberta

*(joint work with Yilan Luo, Mingkuan Wu, David Wood, Hamidreza Zareipour)*

Deniz Sezer

University of Calgary

May 24, 2019



**UNIVERSITY OF  
CALGARY**

## Definition:

- An infinite family of random variables:  $\{Y(x, t), x \in S, t \in [0, \infty)\}$ .  $S$  is a subset of  $\mathbb{R}^2$  or  $\mathbb{R}^3$ .
- $t$  represents time.
- $S$  represents a geographical or atmospherical region.
- $Y(x, t)$  represents a physical quantity measured at time  $t$  and at location  $x$ . For example: Wind speed, or wind power.

# Main benefits of spatio-temporal processes

- We can model probabilistic dependence due to spatial and temporal distance.
- We can make inference for unobserved locations from the ones for which a history of observations is available.
- We can incorporate physical laws that affect the evolution of a spatio-temporal process through space and time.

# Spatio-temporal Processes Tool Box:

- Gaussian spatio-temporal Processes,
- Markov Chain Markov Random Fields,
- Marked Point Processes,
- Hierarchical Bayesian Models,
- Stochastic Partial Differential Equations.

# Wind as a spatiotemporal process

- Traditional approach: Multivariate time series are used to model the evolution of wind speed at the observable locations. Inference is limited to future time series of the observable locations.
- Most studies in the current literature on wind and wind power uses the traditional approach!
- Gneiting et.al (2005): First study to use a Gaussian spatio-temporal process to model wind speed.

## Overview:

- 1479 MW current installed capacity
- 901 wind turbines
- Approximately 8% of electricity demand
- The capacity will be tripled by 2030, intermittency of wind will strongly affect the electricity markets and the power system.

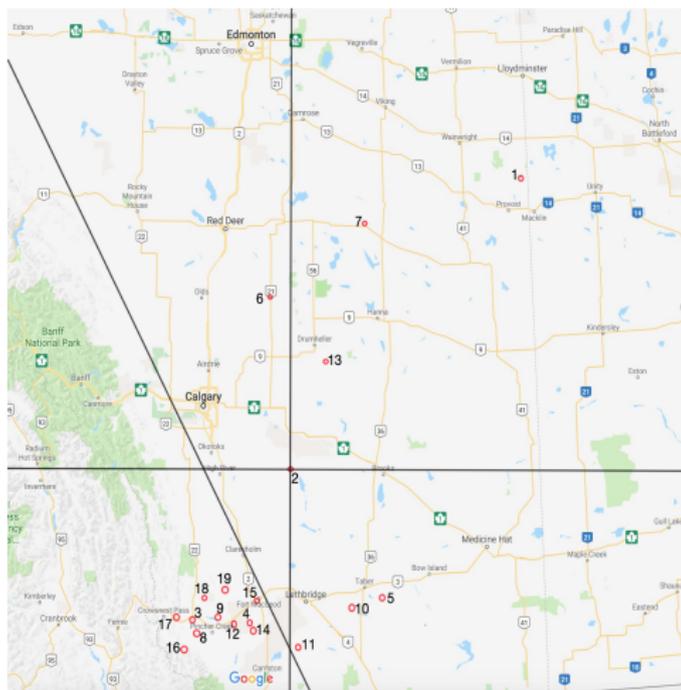
# Our goal

- To model the wind energy generation in Alberta as a spatio-temporal process.
- To make inference on the future aggregate wind power generation incorporating the known locations of future sites.

# Wind Energy Generation Data

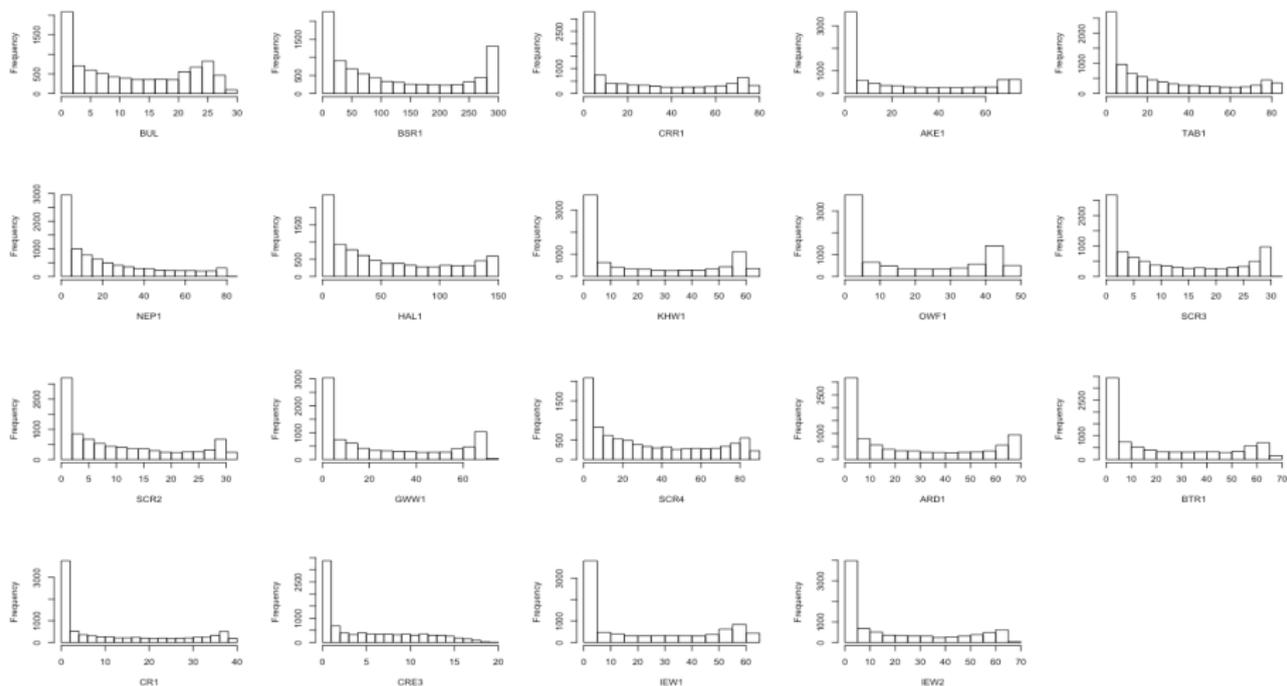
- Data set is obtained from the Alberta Electric System Operator (AESO), consisting of wind power production values in megawatts (MW) averaged over one hour periods from January 2016 to December 2017 for 19 wind farms in Alberta. The number of farms was constant over those two years.
- The data in 2016 are used as the training set and the data in 2017 are used as the testing set.

# Alberta wind farms



**Figure:** Locations of the 19 wind farms in Alberta. The region below the slanted line is subject to strong winds associated with Rocky Mountains. The rest of the region is flat

# Data Preprocessing



**Figure:** Histograms of hourly wind power production in 19 wind farms in the training test.

- Two visible modes: One is at zero power (i.e. when the wind speed is below the cut-in speed), and the other is at maximum capacity (i.e. when all turbines are operating at their maximum power).
- Wind power has a mixed distribution with non-zero probabilities at those conditions.
- Mixed distributions can be obtained by truncation which can be estimated effectively in a Bayesian framework. Related applications: the modeling of precipitation.
- A simpler alternative: We transform our data to daily averages which softens the effect of truncation. The resulting data are skewed to the right; we fix that by applying a square root transform.

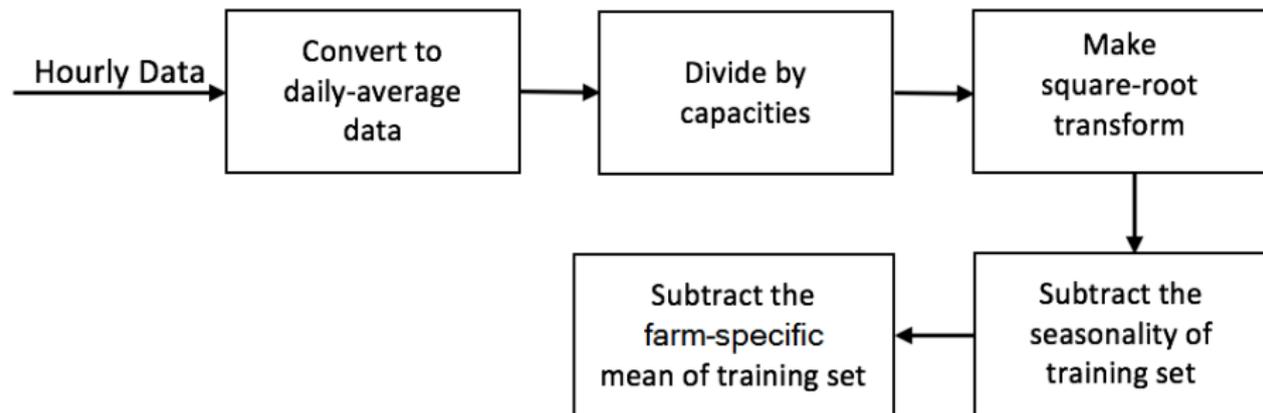
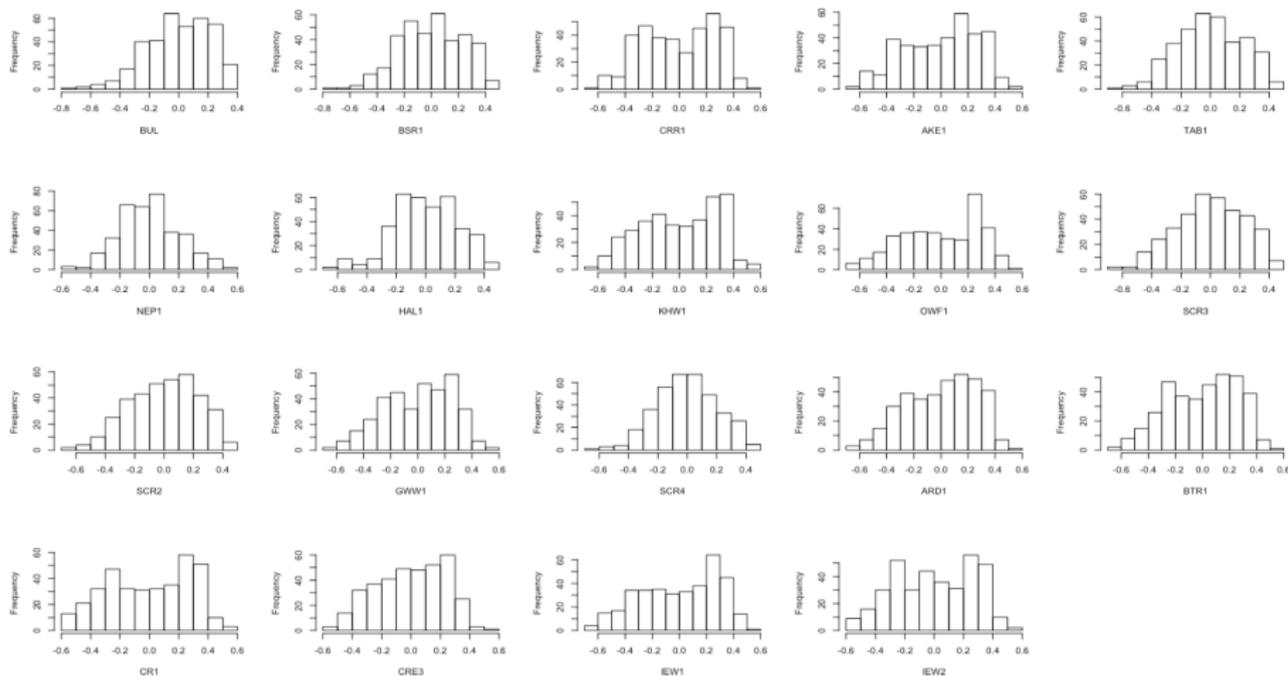


Figure: The flow chart of data preprocessing

# Data after preprocessing



**Figure:** Histograms of daily wind power production in 19 wind farms after preprocessing in the training set

## Assumptions:

- The preprocessed data corresponds to discrete measurements of a Gaussian spatio-temporal process  $\{Y(x, t), x \in S, t \in [0, \infty)\}$ .  $S$  is a subset of  $\mathbb{R}^2$  corresponding to the geographical region containing the 19 wind farms and possible locations for future farms.
- $E[Y(x, t)] = 0$ ,  $\text{Var}[Y(x, t)] = \sigma^2(x)$  for some function  $\sigma : S \mapsto (0, \infty)$ .
- We also assume that there exists a function  $C : S \times T \mapsto \mathbb{R}$ , called the correlation function, such that
$$\text{Cov}[Y(x_1, t_1), Y(x_2, t_2)] = \sigma(x_1)\sigma(x_2)C(x_2 - x_1, t_2 - t_1).$$

Our goal is to uncover  $\sigma$  and  $C$  from the data.

**Table:** Alberta wind farms, their capacity in Megawatts (second row), and specific variances in daily production (third row)

BUL	BSR1	CRR1	AKE1	TAB1	NEP1	HAL1	KHW1	OWF1	SCR3
1	2	3	4	5	6	7	8	9	10
29	300	77	73	81	82	150	63	46	30
0.046	0.055	0.072	0.074	0.048	0.041	0.048	0.079	0.084	0.053
SCR2	GWW1	SCR4	ARD1	BTR1	CR1	CRE3	IEW1	IEW2	
11	12	13	14	15	16	17	18	19	
30	71	88	68	66	39	20	66	66	
0.053	0.064	0.042	0.067	0.066	0.08	0.055	0.082	0.072	

# Estimation of $\sigma(\cdot)$

We assume that  $\sigma$  is of the form:

$$\sigma(x) = \begin{cases} 0.05 & \text{if } x \text{ is above the slanted line,} \\ 0.07 & \text{if } x \text{ is below the slanted line.} \end{cases}$$

# Estimation of $C(.,.)$

- $C$  must be positive definite. *Positive definiteness* is a necessary and sufficient condition for a correlation function  $C$ , since

$$\text{Var}\left(\sum_{i=1}^m \frac{a_i}{\sigma(x_i)} Y(x_i, t_i)\right) = \sum_{i=1}^m \sum_{j=1}^m a_i a_j C(x_i - x_j, t_i - t_j) > 0$$

- Following the approach of Gneiting et.al.(2005), we use three embedded parametric families which are known to be positive definite:

separable  $\subset$  symmetric  $\subset$  stationary

# Estimation of $C$

## Candidate Models

- Separable correlation function

$$C_{Sep}(\mathbf{h}, u) = C_S(\|\mathbf{h}\|) \times C_T(|u|) \quad (1)$$

- Non-separable but fully symmetric (isotropic) correlation function

$$C_{FS}(\mathbf{h}, u) = C(\|\mathbf{h}\|, |u|) \quad (2)$$

- Lagrangian correlation function

$$C_{LGR}(\mathbf{h}, u) = C_S(\mathbf{h} - \mathbf{V}u) \quad (3)$$

- General stationary correlation function by taking convex combination of the above types.

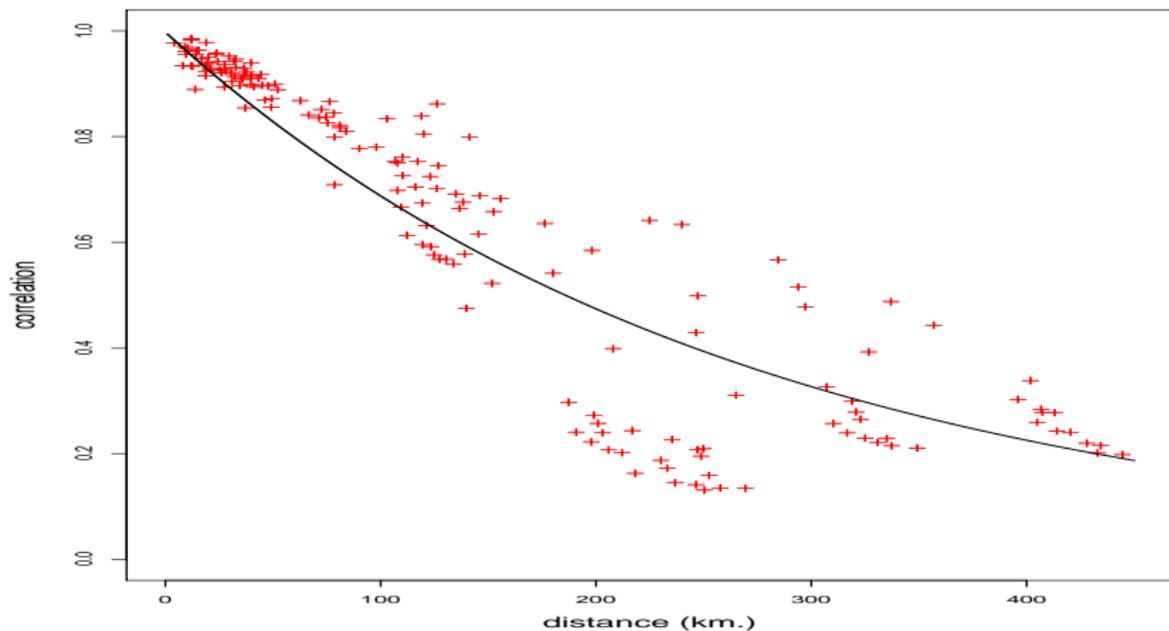
$$C_{STAT}(\mathbf{h}; u) = (1 - \lambda)C_{FS}(\mathbf{h}; u) + \lambda C_{LGR}(\mathbf{h}; u) \quad (4)$$

- As in Gneiting et.al, we estimate the parameters in the models by the method of weighted least squares (WLS):

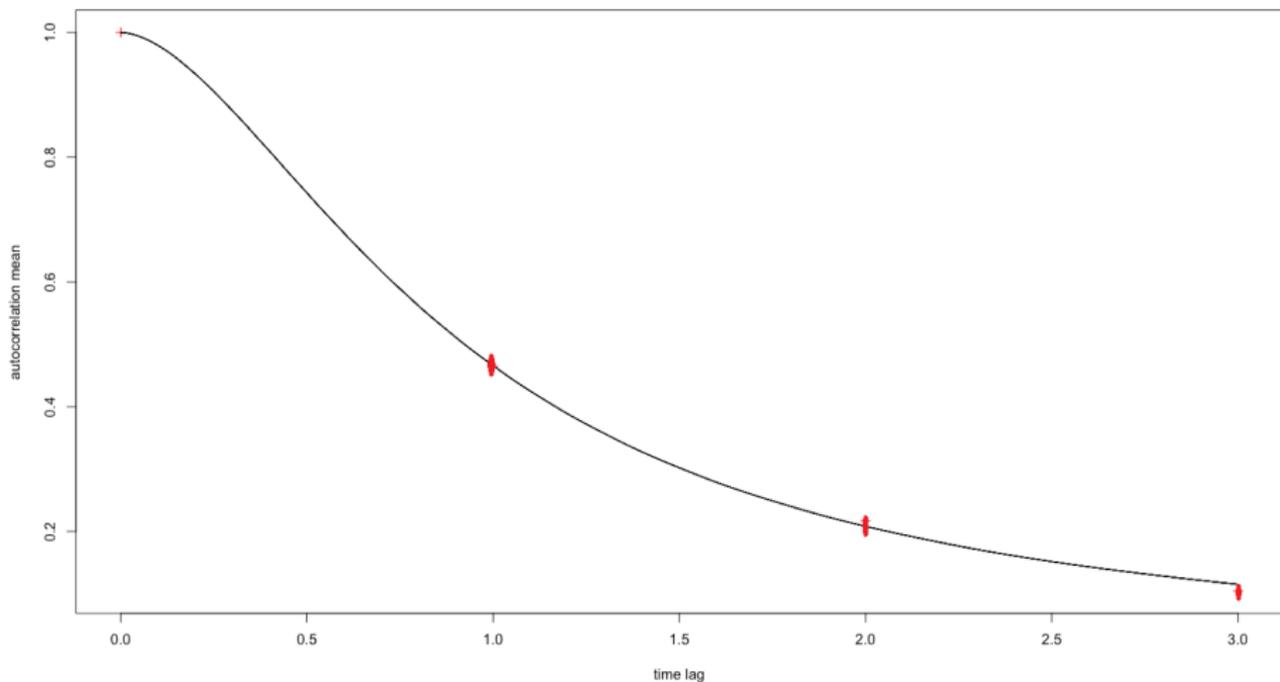
$$\hat{\theta}_{WLS} = \underset{\theta}{\operatorname{argmin}} \sum_{\mathbf{h}} \sum_u \left( \frac{\hat{C}or(\mathbf{h}; u) - Cor(\mathbf{h}; u|\theta)}{1 - Cor(\mathbf{h}; u|\theta)} \right)^2 \quad (5)$$

where  $\hat{C}or(\mathbf{h}; u)$  is the empirical correlation and  $Cor(\mathbf{h}; u|\theta)$  is the fitted function with parameter vector  $\theta$ .

- No significant improvement is achieved by relaxing the separability assumption while enforcing symmetry.
- We achieve significant improvement if we relax the symmetry.

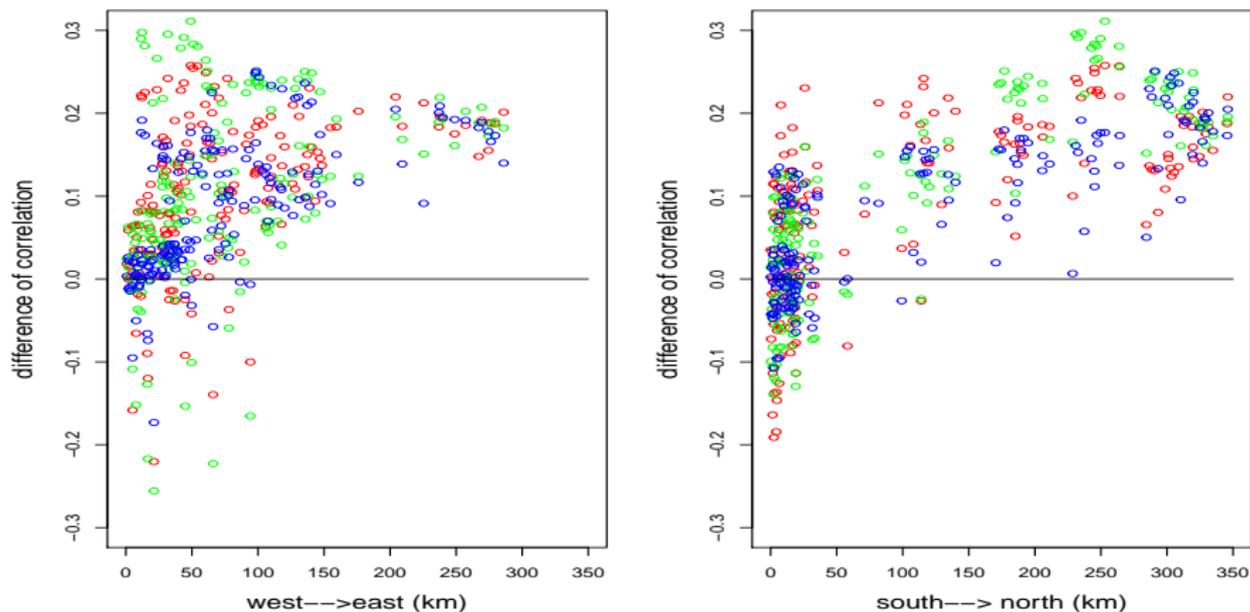


**Figure:** Empirical correlations against distance with fitted pure spatial correlation.



**Figure:** Empirical auto-correlations against time lags with fitted pure temporal correlation.

# Asymmetry



**Figure:** The left (resp. right) plot indicates the difference between the empirical west-to-east (south-to-north) and east-to-west (north-to-south) cross-correlations

- Asymmetry is violated in both the west-east direction and the south-north direction. More precisely, for most pairs  $x_1, x_2$  such that  $x_1$  is to the west of  $x_2$

$$\widehat{\text{Cor}}(Y(x_1, t), Y(x_2, t + u)) - \widehat{\text{Cor}}(Y(x_1, t), Y(x_2, t - u)) > 0. \quad (6)$$

- This is also true also for most pairs  $x_1, x_2$  such that  $x_1$  is to the south of  $x_2$ .
- To account for the lack of symmetry, we add the prevailing wind influence in the model by using a Lagrangian correlation function

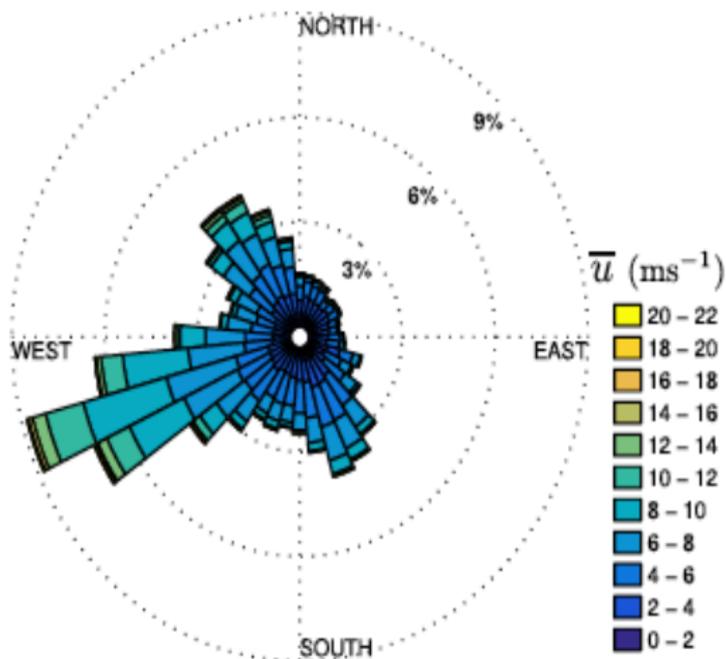
$$C_{LGR}(\mathbf{h}; u) = \left(1 - \frac{1}{2\|\mathbf{v}\|} \|\mathbf{h} - \mathbf{v}u\|\right)_+ \quad (7)$$

where we take  $\mathbf{v}$  as a two dimensional vector.

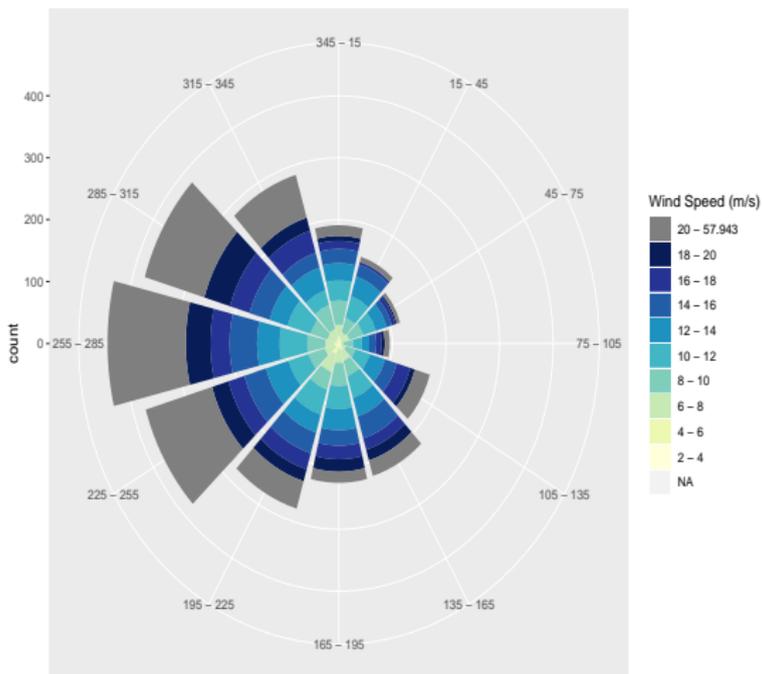
- The model identifies the prevailing wind direction as southwesterly, which is consistent with the findings of Sherry & Rival (2015) on the wind patterns near the Rocky mountains.
- The southerly component is not prevalent in the entire region but the westerly component is. Thus, we also consider a Lagrangian correlation function where the wind direction is assumed to be westerly:

$$C_{LGR,w}(\mathbf{h}; u) = \left(1 - \frac{1}{2|v_w|} \|h_1 - v_w u, h_2\|\right)_+ \quad (8)$$

where we take  $v_w$ , where the subscript “w” indicates a westerly value, as a scalar.



**Figure:** Wind rose plot taken from Sherry & Rival (2015) . The measurements were made at a height of 50 m using a wind mast on the northern outskirts of Calgary.



**Figure:** Wind rose plot from 8 randomly selected weather stations in Alberta.  
 Data: Alberta Agriculture and Forestry

We identify three models for comparison:

- Model 1

$$C_{SEP}(\mathbf{h}; u) = C_S(\mathbf{h})C_T(u) \quad (9)$$

with

$$C_S(\mathbf{h}) = 0.998 \exp(-0.0037||h||) + 0.0025\delta_{\mathbf{h}=\mathbf{0}},$$
$$C_T(u) = \left(1 + 1.1472|u|^{2(0.8635)}\right)^{-1}.$$

- Model 2

$$C_{STAT}(\mathbf{h}; u) = 0.744C_{SEP} + 0.256C_{LGR}(\mathbf{h}; u) \quad (10)$$

with

$$C_{LGR}(\mathbf{h}; u) = \left(1 - \frac{1}{2||v||}||h - vu||\right)_+$$

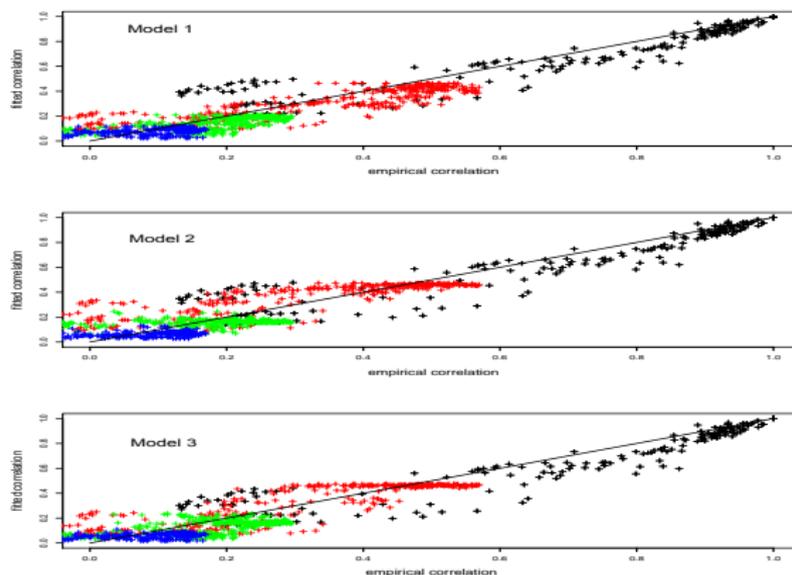
and  $v = (143.52, 74.57)$  km/day and  $C_{SEP}$  as in Model 1.

- Model 3

$$C_{STAT,w}(\mathbf{h}; u) = 0.764C_{SEP} + 0.236C_{LGR,w}(\mathbf{h}; u) \quad (11)$$

with

# Comparison of the models: Goodness of fit



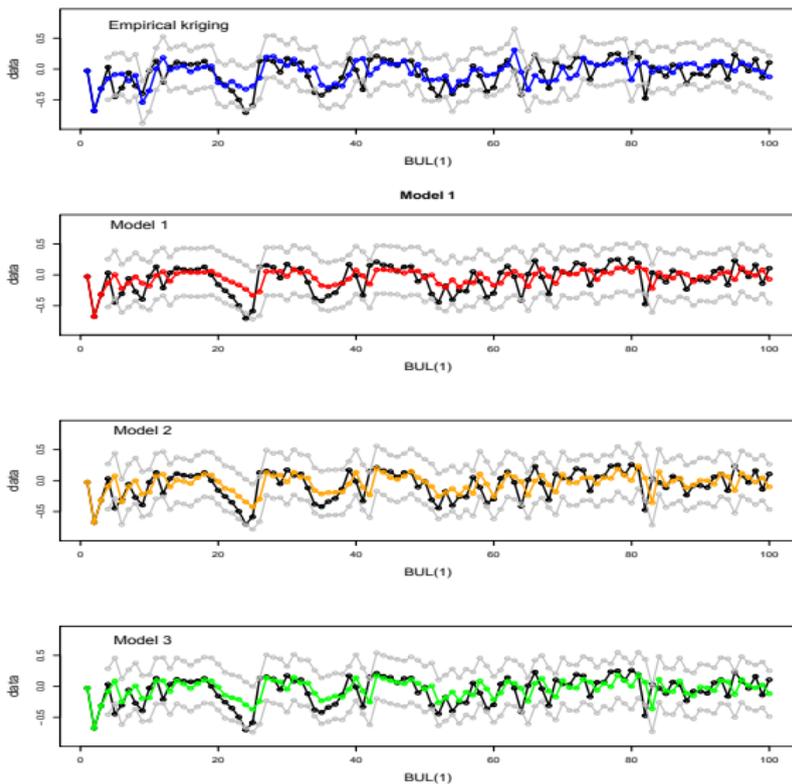
**Figure:** black, red, green and blue colors correspond to cross correlations of any two locations at time lags 0, 1, 2 and 3 days, respectively. The cluster of red and green points near the line  $y = x$  is larger and distributed more symmetrically for Model 2.

# Comparison of the models: Kriging predictions-Scenario 1

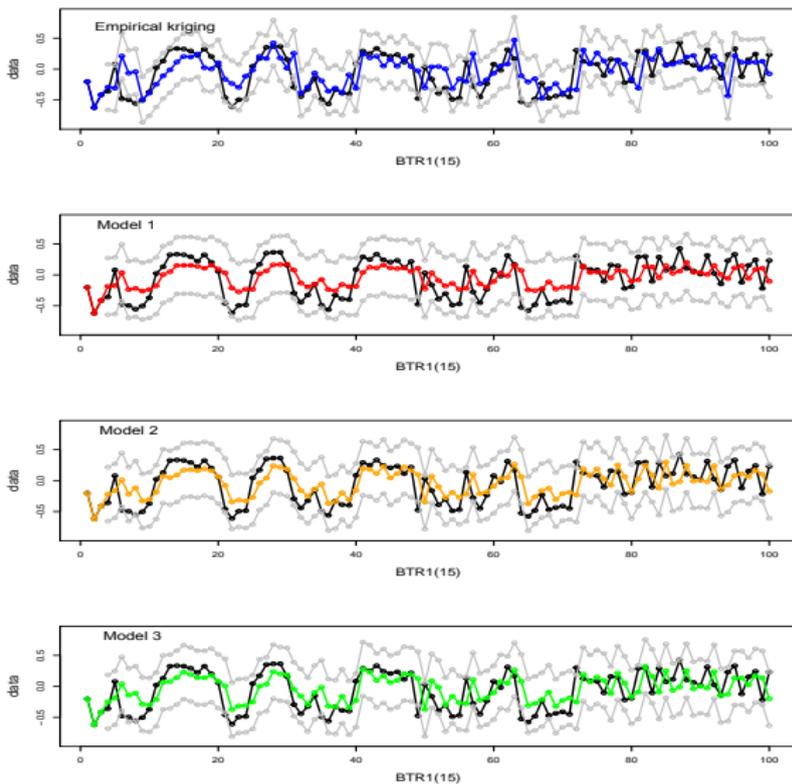
- In scenario 1, we predict the power output of a given wind farm on a given day based on the observations from all the wind farms (including the same farm) in the past three days.
- Various measures of error are calculated :the root mean square error (RMSE) and the mean absolute error (MAE) as a function of the site. Since both of these measures scale with the variability of observations, we also use coefficient of determination  $R^2$ .
- We also construct 95% prediction intervals (PI) calculate the realized percentage of observations falling outside the 95% PI (POPI). Ideally this percentage should be close to 5%.
- Results indicate that Model 2 is the best model overall.

Table: Mean prediction errors for all wind farms

	$\overline{RMSE}$	$\overline{MAE}$	$\overline{R^2}$	$\overline{POPI}$
Empirical	0.2379	0.1920	0.1484	0.1234
Model 1	0.2242	0.1886	0.2440	0.0429
Model 2	0.2238	0.1869	0.2446	0.0584
Model 3	0.2255	0.1873	0.2326	0.0608



**Figure:** Actual daily power production (black) versus predicted values when BUL is treated as an existing wind farm. The x-axis is in days. Prediction intervals are in color gray.



**Figure:** Actual daily power production (black) versus predicted values when CR1 is treated as an existing wind farm. The x-axis is in days. Prediction intervals are in color gray.

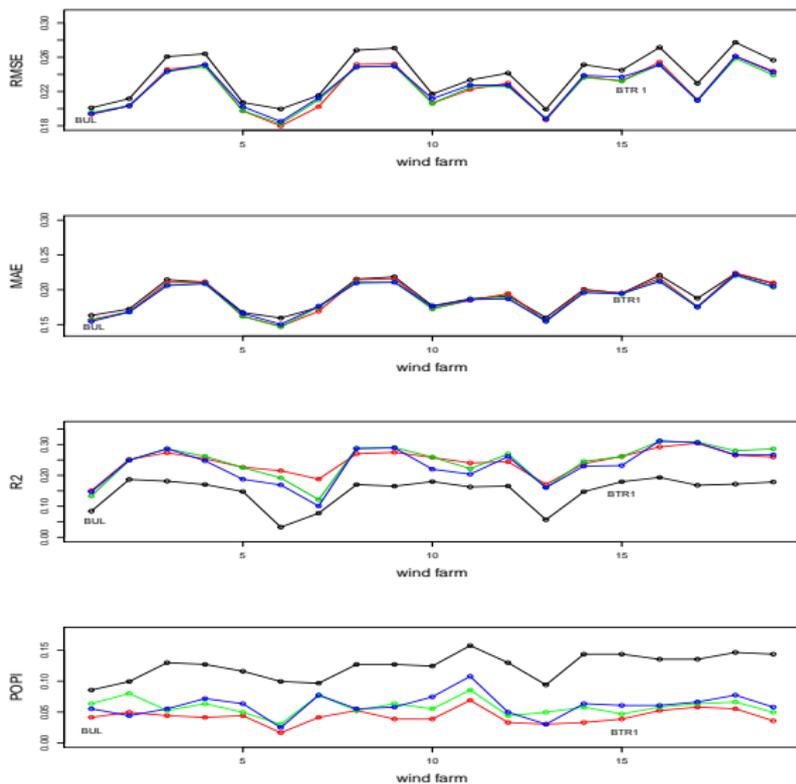
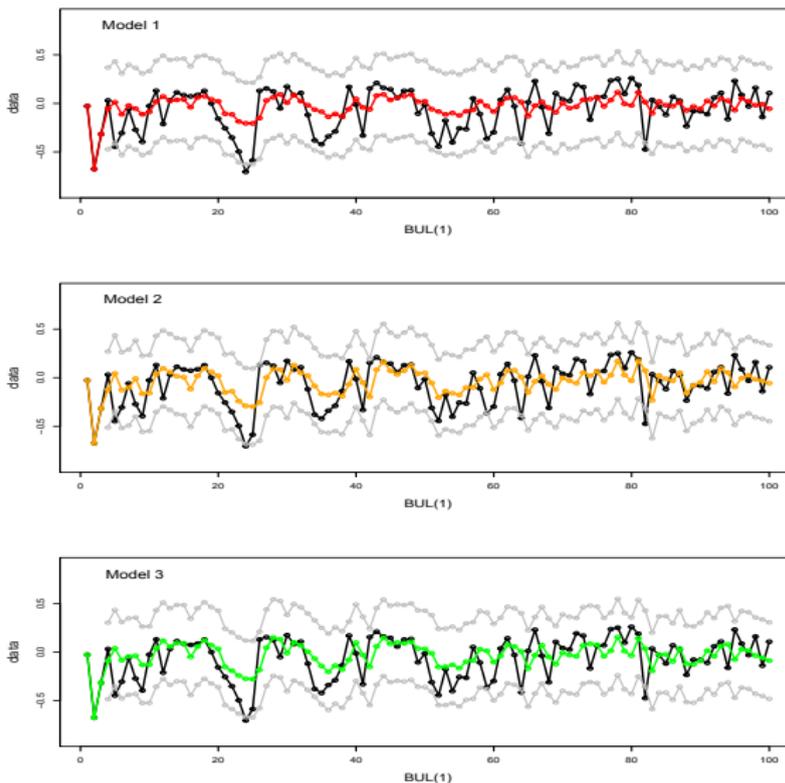


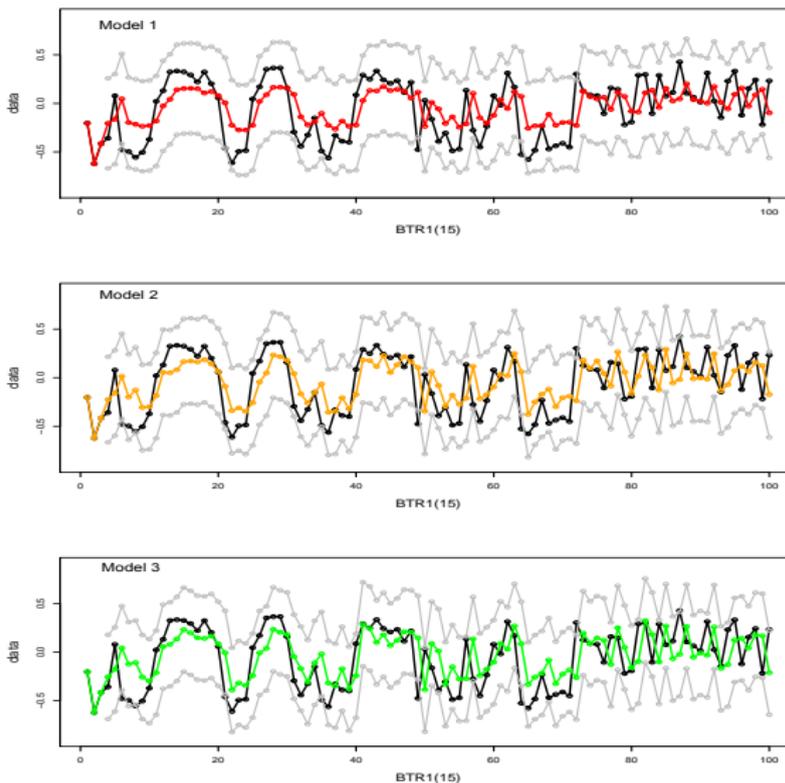
Figure: RMSE, MAE,  $R^2$  and POPI for predicting power output from all wind farms. Empirical: red. Model 1: red. Model 2: green. Model 3: blue.

## Comparison of the models: Kriging predictions-Scenario 2

- We select a wind farm to be treated as a new site and perform the estimation procedure by removing all its historical data,
- We predict the observation at the selected site on a given day using the past three days of data for all other sites.
- Results indicate that Model 2 is also the best model for predictions for a new site.



**Figure:** Actual daily power production (black) versus predicted values when BUL is treated as a new wind farm. The x-axis is in days. Prediction intervals are in color gray.



**Figure:** Actual daily power production (black) versus predicted values when BTR1 is treated as a new wind farm. The x-axis is in days. Prediction intervals are in color gray.

**Table:** Mean prediction errors for a new wind farm

	$\overline{RMSE}$	$\overline{MAE}$	$\overline{R^2}$	$\overline{POPI}$
Model 1	0.2245	0.1892	0.2432	0.0366
Model 2	0.2246	0.1878	0.2397	0.0538
Model 3	0.2260	0.1880	0.2306	0.0621

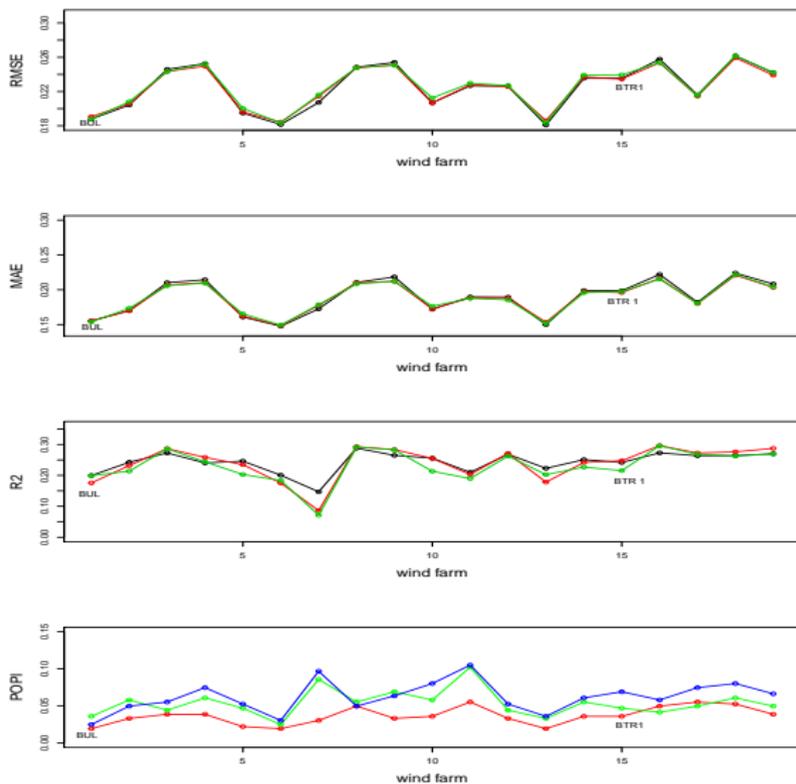


Figure: RMSE, MAE,  $R^2$  and POPI for predicting a new wind farm as a function of the farm site. Model 1: black. Model 2: red. Model 3: green.

# Aggregate Wind Power Generation

- Let  $W(x, t)$  represent the wind power output at location  $x$  and time  $t$ .
- Then,

$$Y(x, t) = \sqrt{W(x, t) - (S(t) + m(x))}, \quad (12)$$

or equivalently,

$$W(x, t) = (Y(x, t) + S(t) + m(x))^2, \quad (13)$$

where  $Y(x, t)$  is a second order stationary Gaussian process with mean 0, variance  $\sigma^2(x)$  and correlation function

$C(s_2 - s_1, t_2 - t_1) = C(h, u)$ .  $S(t)$  and  $m(x)$  represent the seasonality at time  $t$  and the station-specific mean at location  $x$ .

- Let  $X^c = \{x_1^c, x_2^c, \dots, x_k^c\}$  be the vector of locations of current wind farms with corresponding maximum capacities  $C^c = \{c_1^c, c_2^c, \dots, c_k^c\}$ .
- We also consider  $p$  new wind farms with locations  $X^f = \{x_1^f, x_2^f, \dots, x_p^f\}$  and capacities  $C^f = \{c_1^f, c_2^f, \dots, c_p^f\}$ .
- Let  $t^*$  be a time point in the future, then  $S(t^*)$  is a constant in this case. Thus, the aggregate wind power generation at time  $t^*$  is

$$G_{agg}(t^*) = \sum_{i=1}^k c_i^c W(x_i^c, t^*) + \sum_{j=1}^p c_j^f W(x_j^f, t^*) \quad (14)$$

When  $x, t$  are fixed,  $S(t)$  and  $m(x)$  are constant.

- We derive explicit formulas for the mean and the variance of  $G_{agg}(t^*)$ .

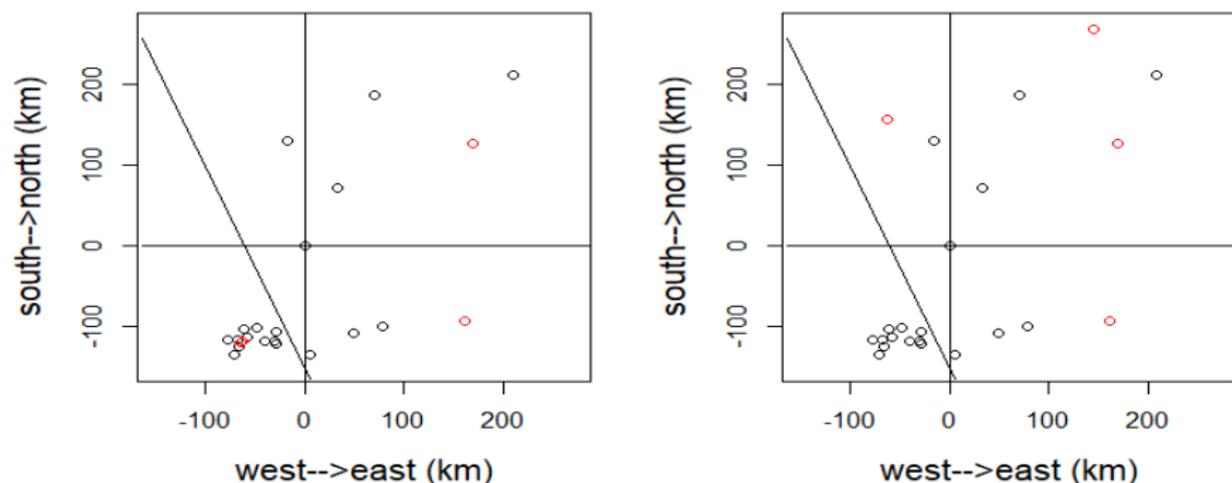
Table: Future wind farms

Sites	Capacity (MW)	Coordinates
Sharp Hills, Oyen	248.4	(51.74, -110.66)
Riverview, Pincher Creek	115	(49.53, -113.92)
CRR2 Pincher Creek	30.6	(49.55, -113.89)
Whitla Wind, Medicine Hat	201.6	(49.76, -110.77)

**Table:** Current and future aggregate daily production: **With the new wind farms the total power generation will increase by 40.8%, whereas the variability as measured by the standard deviation, of the aggregate generation will increase by 31.5%.**

	Total capacity (MW)	Mean (MW)	Std deviation (MW)
Current	1445	737.23	381.34
Future	2040.6	1037.93	502.34

# The effect of geographic dispersion



**Figure:** Hypothetical relocation of the planned sites to increase geographic dispersion results an increase in the mean wind power generation by 40.4%, an increase in the variability of the wind power generation by 25.47%.

# Summary

- We modeled the wind power as a Gaussian spatio-temporal processes, using publicly available data
- Physical effects such as prevalent wind directions can be captured by asymmetric correlation functions.
- We estimated the mean and the variance of the future aggregate power generation of Alberta.
- Our model can demonstrate the effect of geographic dispersion on the variability of the aggregate power generation.

- Assumption of Gaussian Process
- Increasing the time resolution
- Modeling of Prevalent Wind Directions

- [1] Marc G. Genton, Tilmann Gneiting and Peter Guttorp. (2005) Geostatistical space-time models, stationarity, separability and full symmetry, Department of Statistics University of Washington, Technical Report no. 475.
- [2] Y. Luo, M. Wu, D. Sezer, D. Wood, H. Zareipour, (2019), Estimation of the daily variability of aggregate wind power generation in Alberta, Canada. To appear in *Energies*.

# THANK YOU!