UNIVERSITY OF CALGARY

Spatial-Temporal Gaussian Processes for Wind Power Generation

Yilan Luo

（work with Dr. Deniz Sezer, University of Calgary）

Sep. 19, 2018

Department of Mathematics and Statistics

University of Calgary

Overview

01

INTRODUCTION

- Wind energy in Alberta
- Motivation and goal

02

MODELING

- Spatial-temporal processes
- Covariance functions

03

APPLICATION

- Kriging prediction
- Aggregate wind power generation

04

R PACKAGE

- Design of R package

Overview

01 INTRODUCTION

- Wind energy in Alberta
- Motivation and goal

02 MODELING

- Spatial-temporal processes
- Covariance functions

03 APPLICATION

- Kriging prediction
- Aggregate wind power generation

04 R PACKAGE

- Design of R package

01

INTRODUCTION

Wind Energy in Alberta

(From *CanWEA*)

- Third largest wind market in Canada
- 1483 MW installed capacity (Aug. 2018)
- 901 wind turbines
- Approximately 8% of electricity demand

link: https://canwea.ca/wind-energy/alberta/

01

INTRODUCTION

Wind Farms in Alberta

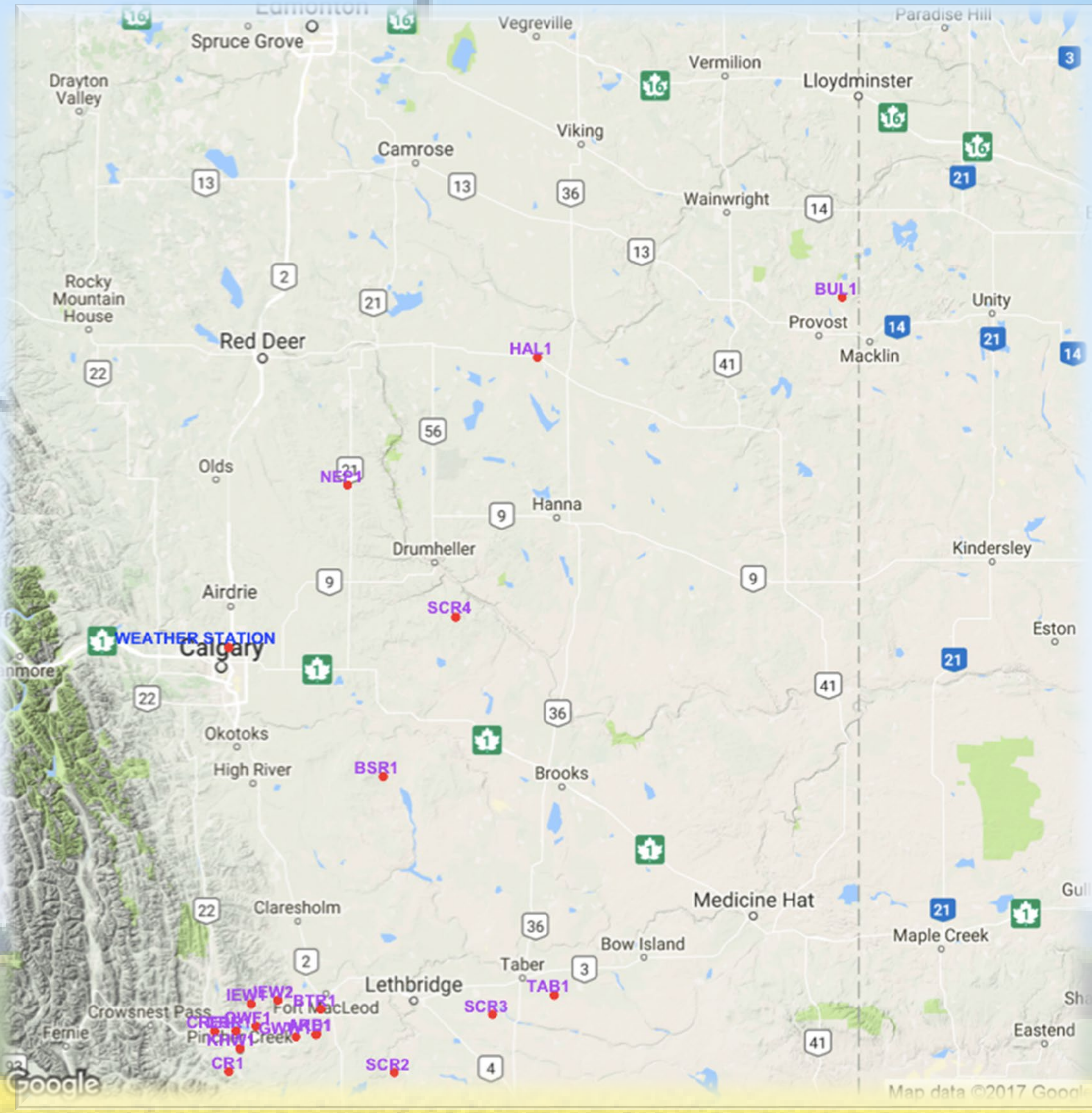
WIND		
ASSET	MC	TNG
Ardenville Wind (ARD1)*	68	64
BUL1 Bull Creek (BUL1)*	13	11
BUL2 Bull Creek (BUL2)*	16	14
Blackspring Ridge (BSR1)*	300	41
Blue Trail Wind (BTR1)*	66	6
Castle River #1 (CR1)*	39	21
Castle Rock Wind Farm (CRR1)*	77	29
Cowley Ridge (CRE3)*	20	4
Enmax Taber (TAB1)*	81	11
Ghost Pine (NEP1)*	82	12
Halkirk Wind Power Facility (HAL1)*	150	19
Kettles Hill (KHW1)*	63	59
McBride Lake Windfarm (AKE1)*	73	50
Oldman 2 Wind Farm 1 (OWF1)*	46	41
Soderglen Wind (GWW1)*	71	69
Summerview 1 (IEW1)*	66	55
Summerview 2 (IEW2)*	66	44
Suncor Chin Chute (SCR3)*	30	14
Suncor Magrath (SCR2)*	30	29
Wintering Hills (SCR4)*	88	38

(From *AESO*)

- ❖ 20 wind farms
- ❖ MC: Maximum Capacity (NW)
- ❖ TNG: Total Net Generation (NW)

link: http://ets.aeso.ca/ets_web/ip/Market/Reports/CSDReportServlet

INTRODUCTION



Wind Farms in Alberta

Fact

- Strong westerly winds buffet the Rocky Mountains creating an excellent wind-energy resource in the plains of Alberta.

Result

- Alberta exhibits substantial wind-energy deployment near the complex terrain of the Rocky Mountains.

INTRODUCTION

Motivation

Wind Farm Operator

- Wind energy is the lowest-cost source of new electricity in Alberta.
- Alberta is the only province in Canada with a fully deregulated electricity market and operates on a spot price.

Alberta Government

- For economic and environmental benefits, government expects at least 4,500 MW of new wind energy capacity.

Challenge: The intermittency of wind

INTRODUCTION

Goal

Wind Farm Operator

- Probabilistic modeling of wind energy generation in wind farms in Alberta.

Alberta Government

- Inference on the future aggregate wind energy generation in wind power system.

02

MODELING

- Spatial-temporal processes
- Covariance functions

Spatio-Temporal Process

Definition

- A collection of random variables

$$Y(s,t), \qquad s \in D \subset \mathbb{R}^d, \mathrm{t} \in [0,T) \subset \mathbb{R}$$

- D: A geographical region such as Alberta.
- s: Geographical coordinates of sites in D. $s=(x,y)$ if $d=2$.
- $Y(s,t)$: Wind power generated at location s and time t.

Spatio-Temporal Process

Hourly Energy Generation in Alberta (from AESO)

Select Report: --- Metered Volumes (All)
Select Format: html
Begin Date:(MM/DD/YYYY) 09 01 2018
End Date:(MM/DD/YYYY) 09 01 2018 OK

Current Historical

Metered Volumes (All)

Report Date: September 10, 2018.

September 01, 2018.

Pool Participant ID	Asset Type	Asset ID	Hour 1	Hour 2	Hour 3	Hour 4	Hour 5	Hour 6	Hour 7	Hour 8	Hour 9	Hour 10
			29.86	30.43	30.28	28.03	27.30	26.12	27.18	26.36	27.18	30.31
9496	RETAILER	941A	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941C	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941E	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941L	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941P	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941R	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941U	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941X	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9558	IPP	G035	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9558	IPP	G036	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0026	0.0000
9558	RETAILER	951A	3.0417	2.9107	2.8636	2.8354	2.8315	2.8451	2.9521	3.1326	3.4374	3.3876
9558	RETAILER	951E	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9558	RETAILER	951L	0.0592	0.0545	0.0518	0.0500	0.0496	0.0500	0.0543	0.0579	0.0661	0.0733
9558	RETAILER	951R	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9558	RETAILER	951U	11.1334	10.8675	10.7171	10.5166	10.5538	10.7230	11.1560	11.7478	12.0838	12.8100

Pool Participant ID	Asset Type	Asset ID	Hour 11	Hour 12	Hour 13	Hour 14	Hour 15	Hour 16	Hour 17	Hour 18	Hour 19	Hour 20
			31.60	30.77	32.33	33.25	32.20	31.75	32.33	33.25	33.25	33.25
9496	RETAILER	941A	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941C	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941E	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941L	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941P	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941R	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941U	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9496	RETAILER	941X	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9558	IPP	G035	0.0432	0.4026	0.4835	0.5072	0.4881	0.3402	0.2600	0.1257	0.0022	0.0000
9558	IPP	G036	0.0338	0.3354	0.3339	0.2362	0.3660	0.3624	0.3994	0.2469	0.0597	0.0122
9558	RETAILER	951A	3.6601	3.7839	3.5619	3.4721	3.5040	3.5860	3.6458	3.5667	3.5851	3.4824
9558	RETAILER	951E	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9558	RETAILER	951L	0.0777	0.0798	0.0804	0.0796	0.0809	0.0827	0.0858	0.0879	0.0845	0.0803
9558	RETAILER	951R	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9558	RETAILER	951U	12.9287	12.9821	12.8970	12.7351	12.5166	12.4268	12.6568	12.9530	12.9076	12.5838

Download link: http://ets.aeso.ca/ets_web/docroot/Market/Reports/HistoricalReportsStart.html

Gaussian Process

Definition

- A collection of random variables, any finite number of which have a multivariate normal distribution.

Property

- A Gaussian process $\{f(x): x \in \mathcal{X}\}$ is completely specified by its mean function $m(x) = E[f(x)]$,
- and covariance function $C(x, y) = E[f(x) - m(x)]E[f(y) - m(y)]$.

- Thus, $f(x) \sim \mathcal{GP}(m(x), C(x, y))$.

Gaussian Process

Example

- For any finite subset of elements $x_1, \ldots, x_k \in \mathcal{X}$, the associated finite set of random variables $f(x_1), \ldots, f(x_k)$ have distribution,

- $f(x_\mathrm{i}) \sim N\left(m(x_\mathrm{i}), \sigma^2(x_\mathrm{i})\right)$, $\quad \mathrm{i} = 1, \ldots, k$

- $\begin{bmatrix} f(x_1) \\ \vdots \\ f(x_k) \end{bmatrix} \sim N\left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_k) \end{bmatrix}, \begin{bmatrix} C(x_1, x_1) & \cdots & C(x_1, x_k) \\ \vdots & \ddots & \vdots \\ C(x_k, x_1) & \cdots & C(x_k, x_k) \end{bmatrix}\right)$

Gaussian Process

Advantages

- If a Gaussian process is assumed to have mean zero, defining the covariance function completely defines the process' behaviour.

- In a spatial-temporal process, covariance function is written as:

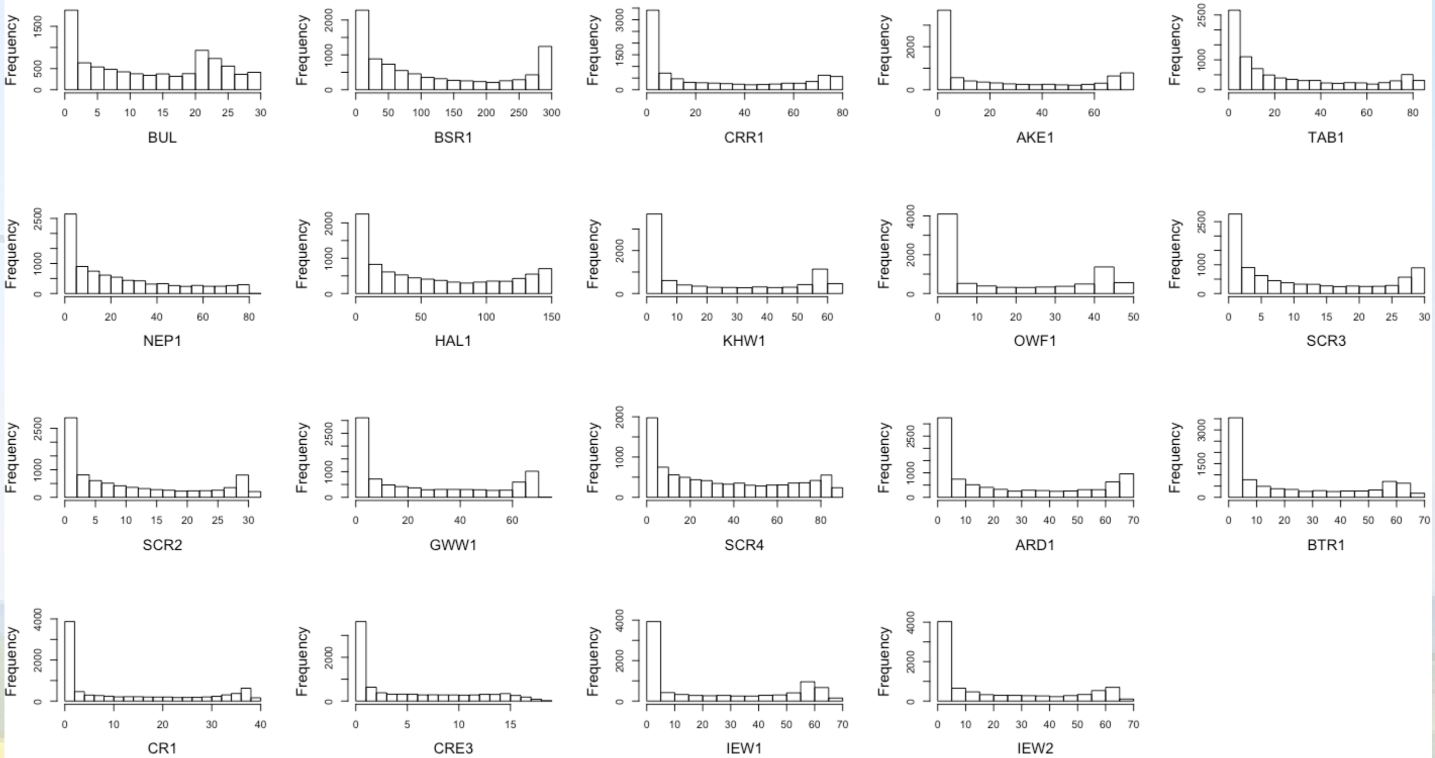
$$C(s_1, s_2, t_1, t_2) = Cov\{Y(s_1, t_1), Y(s_2, t_2)\}$$

02

MODELING

Spatio-Temporal Process

Frequency Histograms of Hourly Wind Energy Generation



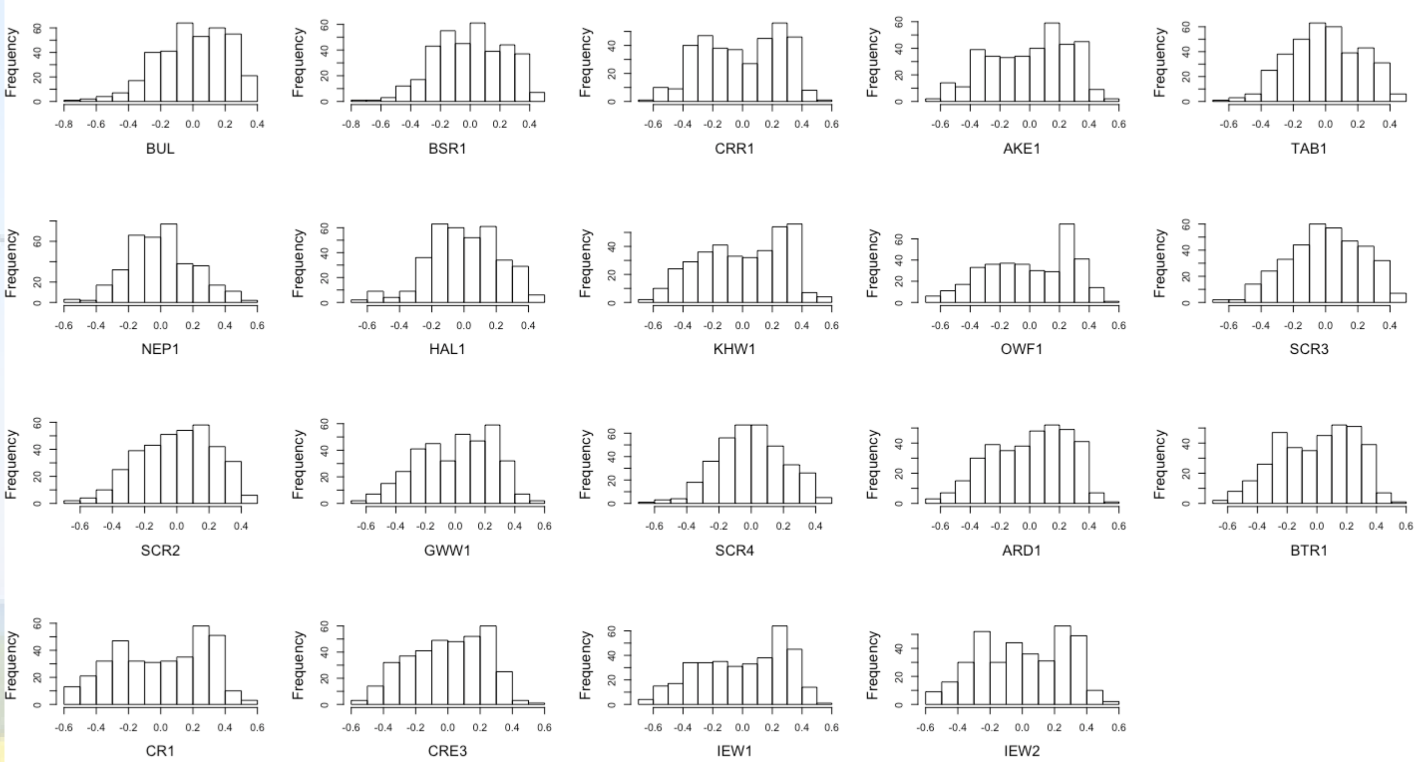
Gaussian Process

Data Preprocessing

1. Hourly Data
2. Convert to daily-average data
3. Divide by capacities
4. Make square-root transform
5. Subtract the seasonality
6. Subtract the station-specific mean
7. Normally Distributed Data

Gaussian Process

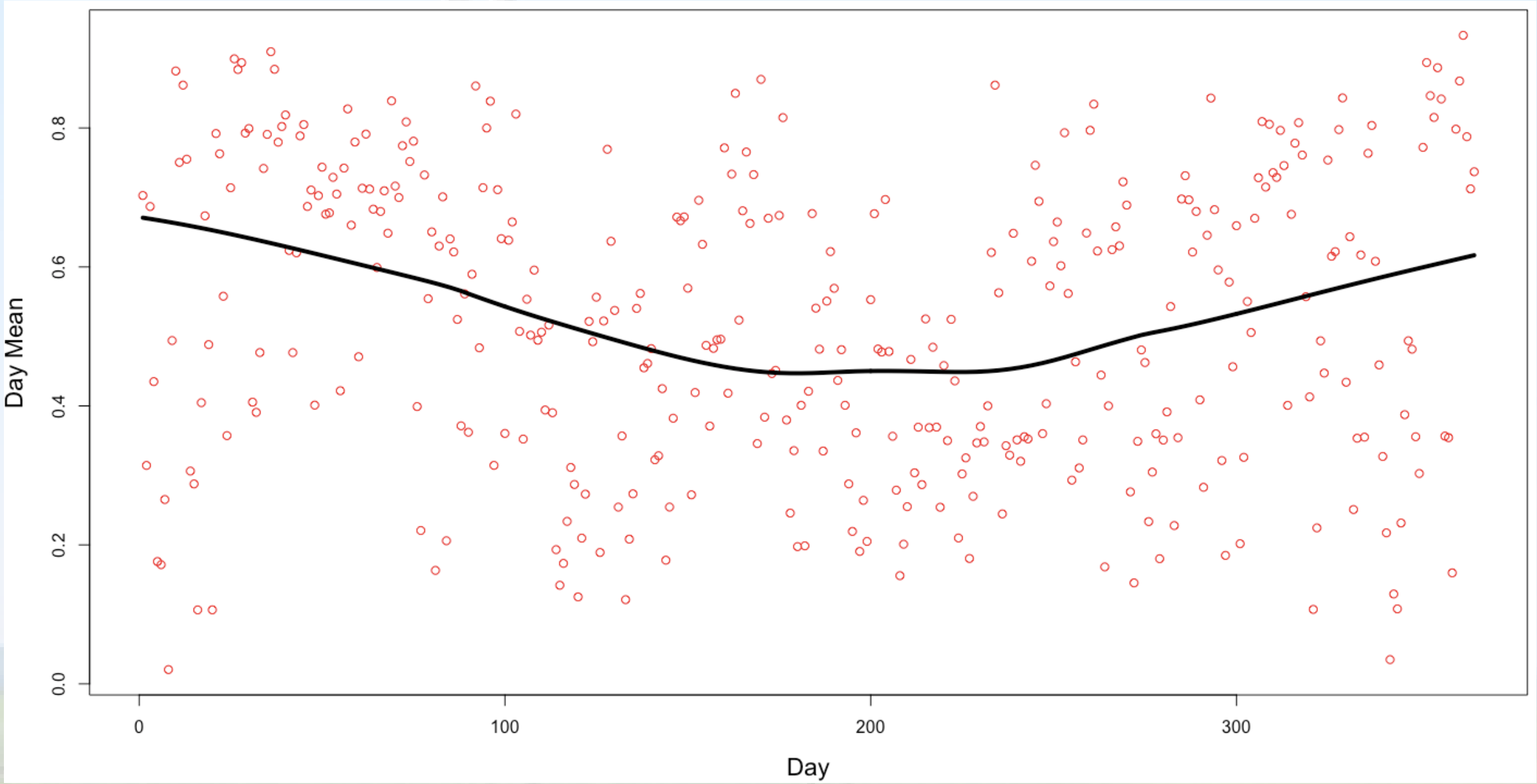
Preprocessed Daily-Average Wind Energy Generation (from AESO)



Gaussian Process

Seasonality

R Code:
stats::lowess(y ~ x, f = 0.5)



Covariance Function

Example of spatial-temporal covariance matrix

- Data in 4 sites during 6 days (24 elements):

$Y(s_1,t_1),\dots,Y(s_1,t_6),Y(s_2,t_1),\dots,Y(s_2,t_6),Y(s_3,t_1),\dots,Y(s_3,t_6),Y(s_4,t_1),\dots,Y(s_4,t_6)$

- Covariance matrix (24 X 24):

$$\begin{bmatrix} C(Y(s_1,t_1),Y(s_1,t_1)) & \dots & C(Y(s_1,t_1),Y(s_1,t_6)) & \cdots & C(Y(s_1,t_1),Y(s_4,t_1)) & \dots & C(Y(s_1,t_1),Y(s_4,t_6)) \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ C(Y(s_1,t_6),Y(s_1,t_1)) & \dots & C(Y(s_1,t_6),Y(s_1,t_6)) & \cdots & C(Y(s_1,t_6),Y(s_4,t_1)) & \dots & C(Y(s_1,t_6),Y(s_4,t_6)) \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ C(Y(s_4,t_1),Y(s_1,t_1)) & \dots & C(Y(s_4,t_1),Y(s_1,t_6)) & \cdots & C(Y(s_4,t_1),Y(s_4,t_1)) & \dots & C(Y(s_4,t_1),Y(s_4,t_6)) \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ C(Y(s_4,t_6),Y(s_1,t_1)) & \dots & C(Y(s_4,t_6),Y(s_1,t_6)) & \cdots & C(Y(s_4,t_6),Y(s_4,t_1)) & \dots & C(Y(s_4,t_6),Y(s_4,t_6)) \end{bmatrix}$$

Covariance Function

Constraint

- Positive Definiteness:

$$Var\left(\sum_{i=1}^{m} a_i Y(s_i, t_i)\right) = \sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j C(s_1, s_2, t_1, t_2) > 0$$

Selection

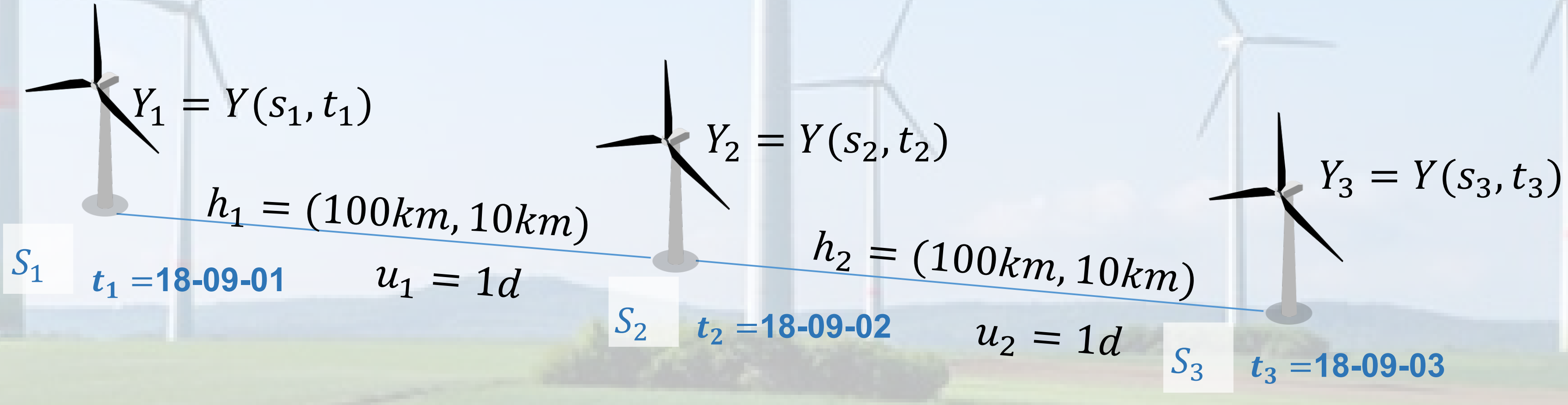
- To ensure that a covariance function is valid, we should consider a parametric family whose members are known to be positive definite functions.

Covariance Function

Stationarity

$$C(s_1, s_2, t_1, t_2) = C(s_2 - s_1, t_2 - t_1) = C(h, u)$$

$$C(s_1, s_2, t_1, t_2) = C(h_1, u_1) = C(h_2, u_2) = C(s_2, s_3, t_2, t_3)$$



Covariance Function

Stationarity

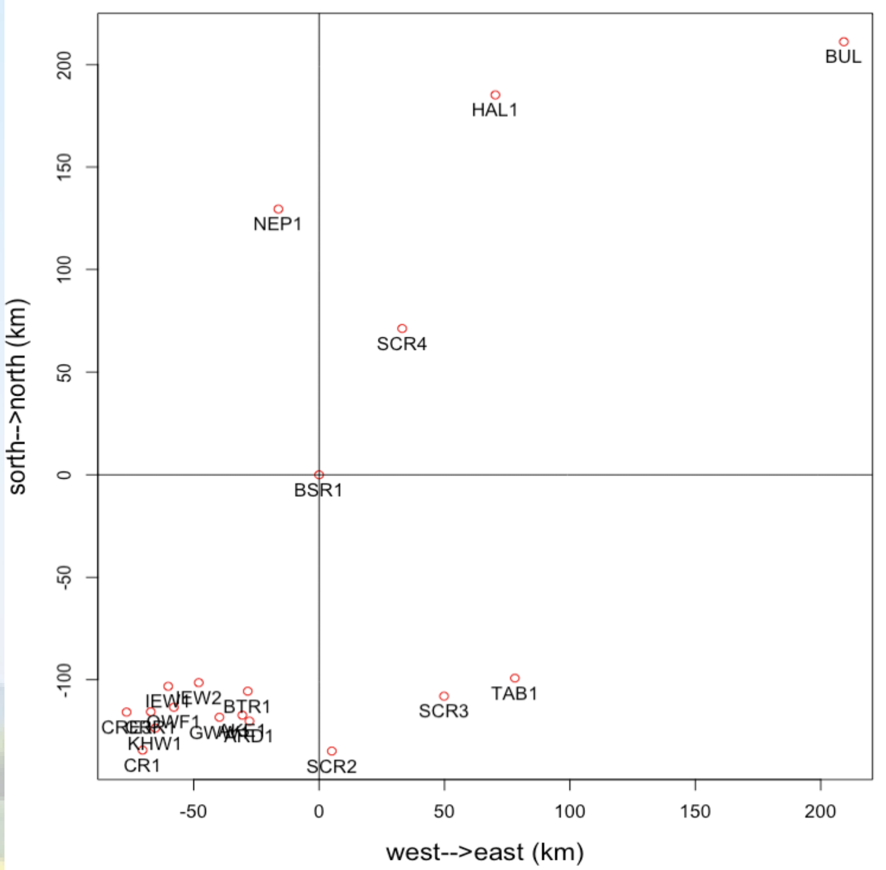
- Spatial-temporal stationarity is hard to justify.
- Temporal stationary is reasonable to assume by subtracting the seasonality of time series.
- However, spatial stationarity is not always realistic because of the variation of geographical conditions between different sites, but it is always ignored in most statistical modeling.

02 MODELING

Covariance Function

Stationarity

R Code:
sp::spDists(longlat, longlat=TRUE)



Covariance Function

Separable covariance function

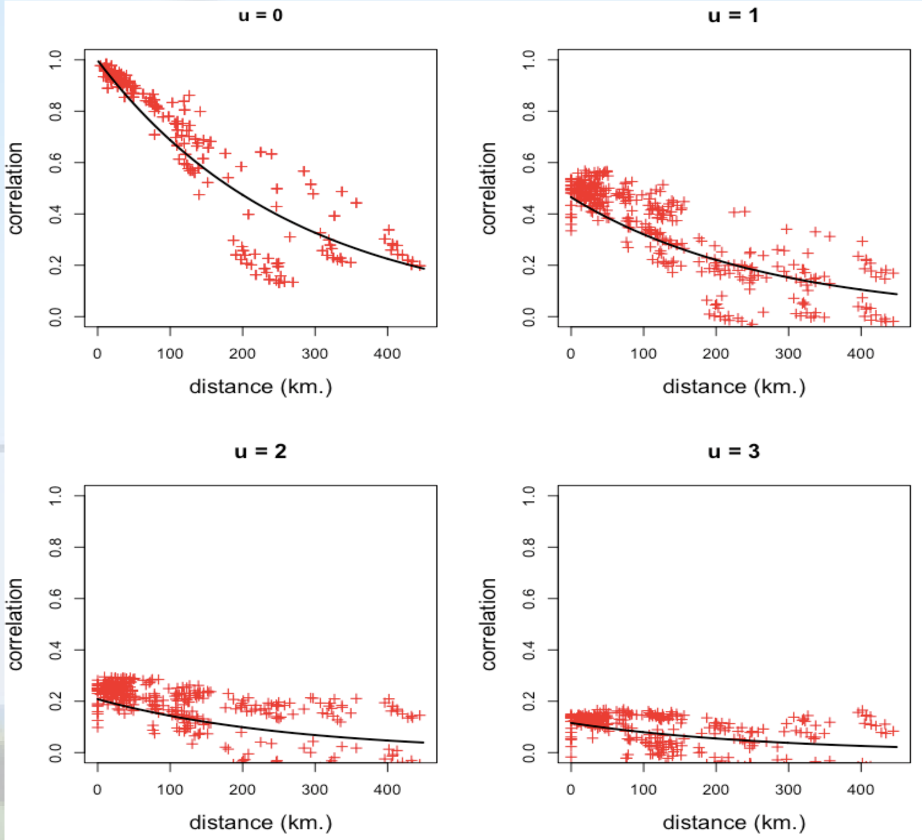
$$C_{sep}(h,u) = C_s(\|h\|) \times C_T(|u|)$$

- $C_s(\|h\|) = (1-\vartheta)\exp(-c\|h\|) + \vartheta\boldsymbol{\delta_{h=0}}$
- $C_T(|u|) = (1 + a|u|^{2\alpha})^{-1}$

Covariance Function

Empirical Cross-covariance

R Code:
stats::ccf(x, y, lag.max = 3, type = c("correlation", "covariance"), …)



Covariance Function

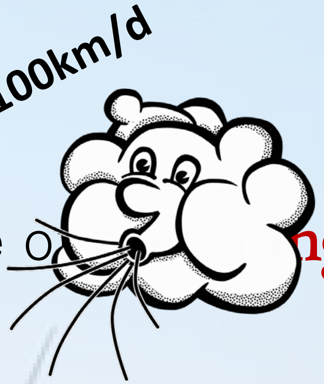
- Non-separable but fully symmetric covariance function

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

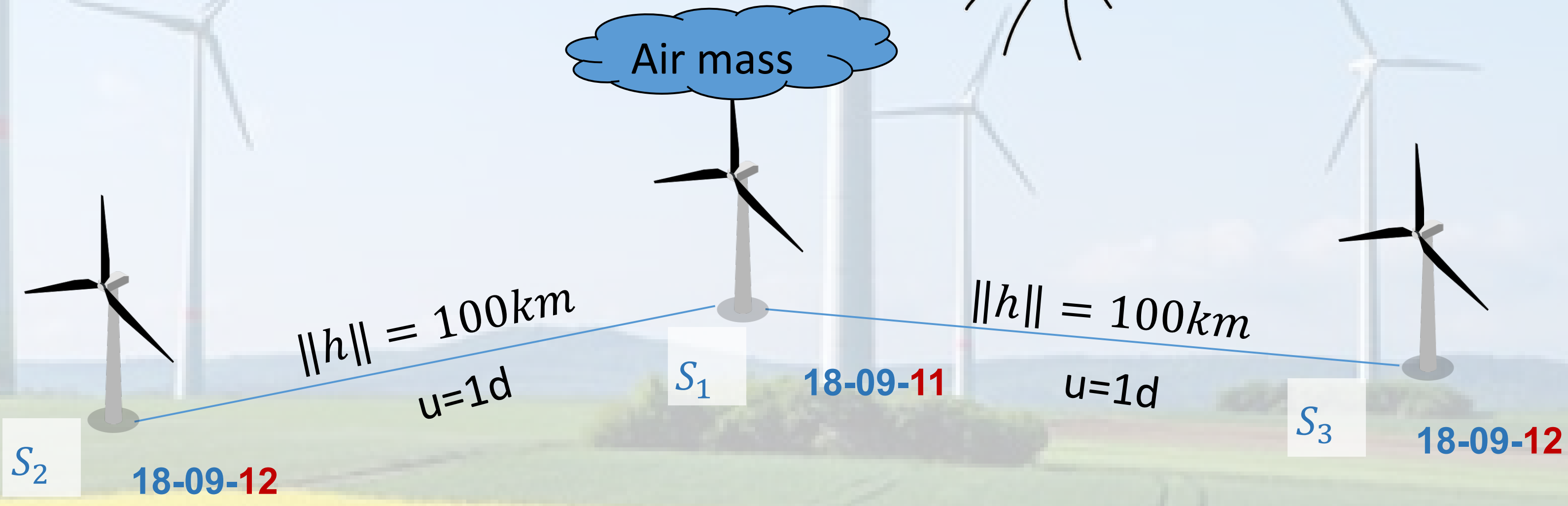
$$C_{FS}(h,u) = \frac{1-\vartheta}{(1+a|u|^{2\alpha})}\left(\exp\left(\frac{-c\|h\|}{(1+a|u|^{2\alpha})^{\frac{\beta}{2}}}\right) + \frac{\vartheta}{1-\vartheta}\boldsymbol{\delta}_{\boldsymbol{h}=\mathbf{0}}\right)$$

Covariance Function

Lack of fully symmetry



- $C(h,u) = C(\|h\|,|u|)$ is violated because o ng wind



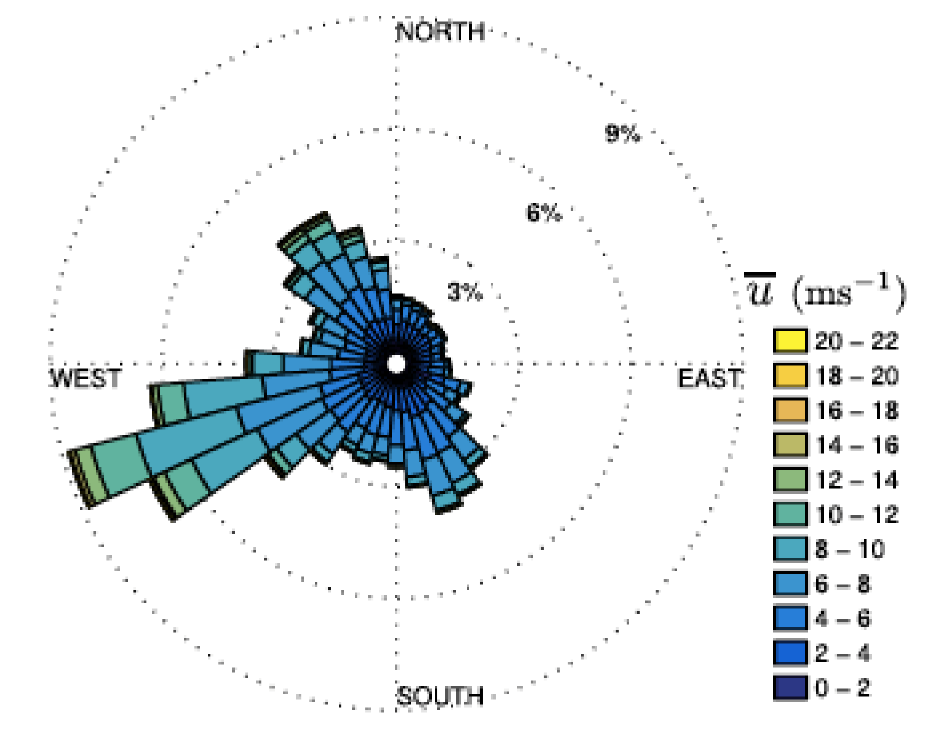
Covariance Function

Lack of fully symmetry

- Prevailing wind:

West-southwesterly wind

R Code:
openair::windRose(mydata, ws = "ws", wd = "wd",…)

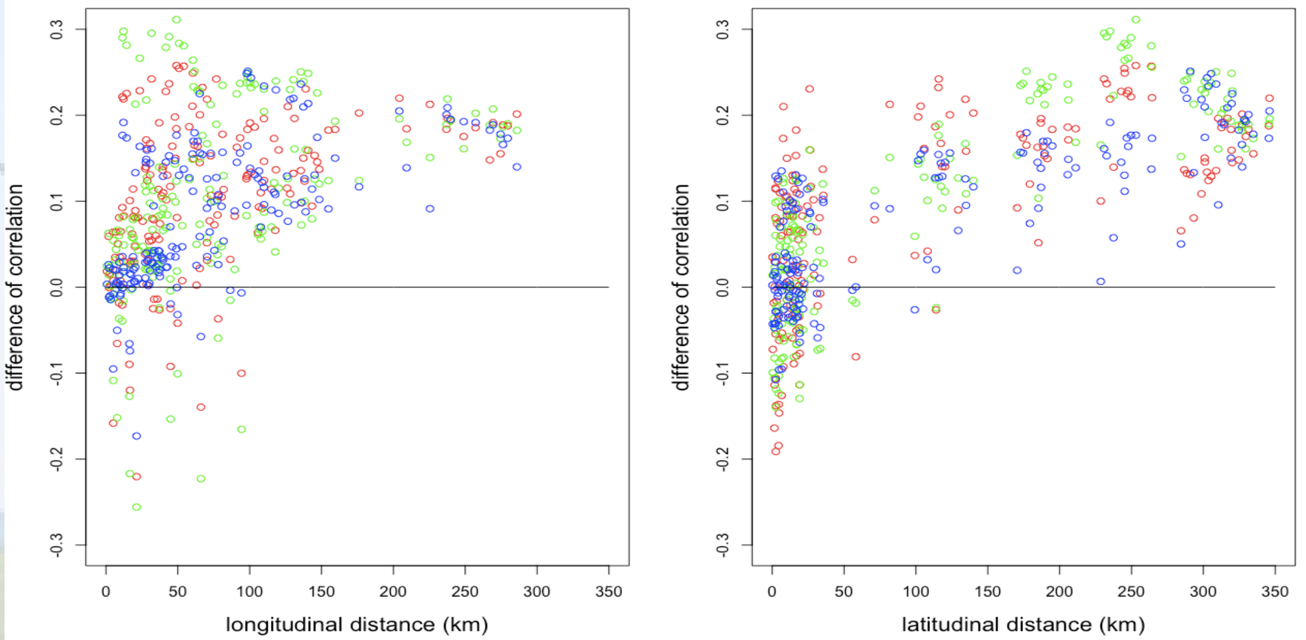


Covariance Function

Lack of fully symmetry

$C(h,u) - C(h,-u)$, u = 1,2,3 days

Red: $u = 1$
Green: $u = 2$
Blue: $u = 3$

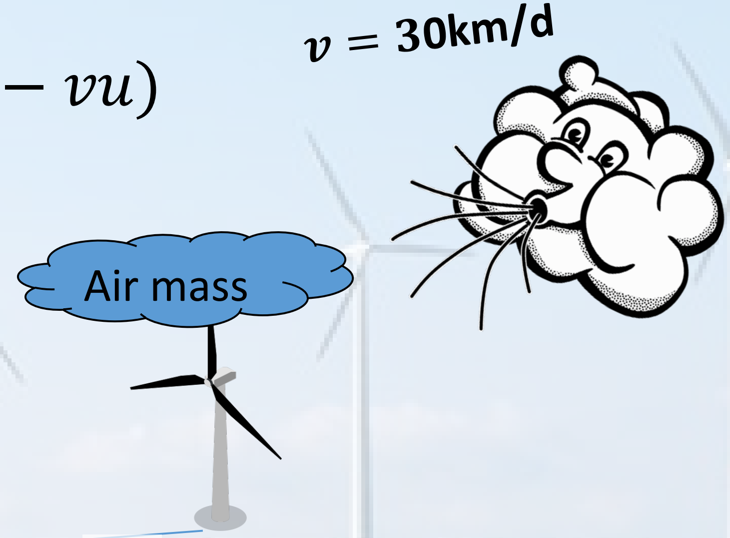


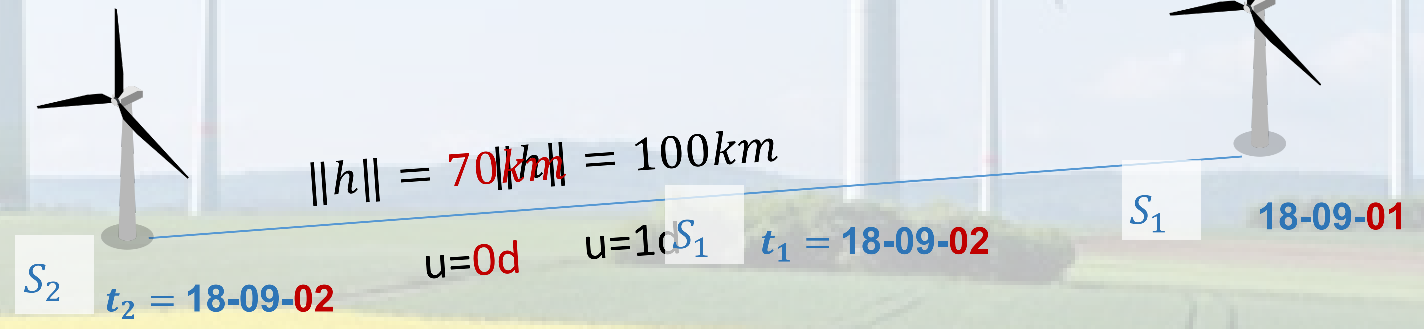
Covariance Function

Lagrangian covariance function

$$C_{LRG}(h,u) = C_s(h - vu)$$

$$C(100,1) = C(100 - 30 \times 1,0)$$
$$= C_s(70)$$



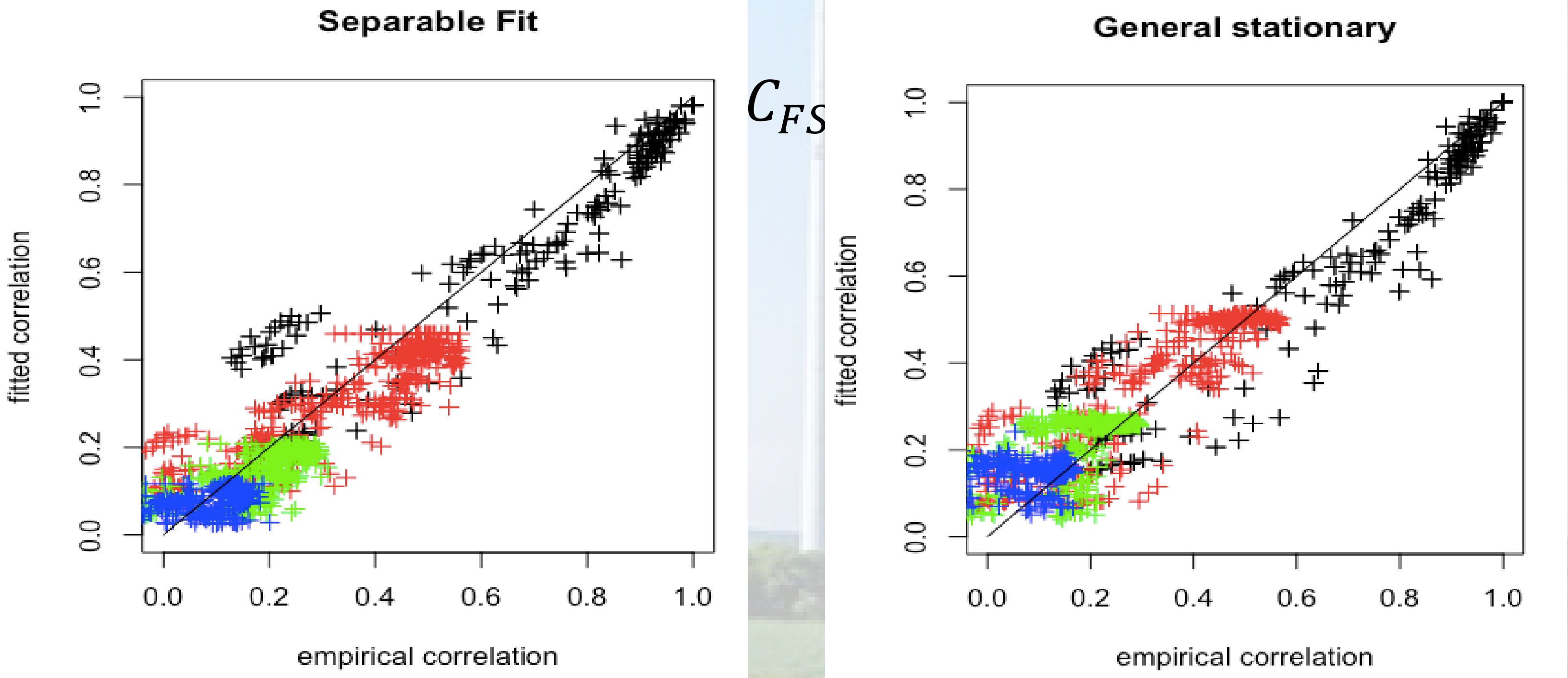


Covariance Function

General stationary covariance function

Black: $u = 0$
Red: $u = 1$
Green: $u = 2$
Blue: $u = 3$

C_{FS}



03

APPLICATION

- Kriging prediction
- Aggregate wind power generation

Prediction for Individual Wind Farm

Kriging Prediction

- An interpolation algorithm, widely used to predict spatial data at an unobserved point in geostatistics.

$$\hat{Y}_{N+1} - m_{N+1} = \sum_{I=1}^{N} w_i (Y_i - m_i)$$

- Simple Kriging $\hat{Y}_{N+1} = \sum_{I=1}^{N} w_i Y_i$

Prediction for Individual Wind Farm

Kriging Prediction

- Weight Selection:
  - Always generated from the covariance function
  - Less weight to a site having a low correlation with the predictand.
  - Determine the weights by minimizing
    $$Var(\hat{X}_{N+1} - X_{N+1})$$
  - Under the unbiasness constraint
    $$E\left(\hat{X}_{N+1} - X_{N+1}\right) = 0$$

Prediction for Individual Wind Farm

Kriging Prediction

- Simple Kriging : $\hat{Y}_{N+1} = C'_{N+1}\Sigma^{-1}Y$
- Σ : Covariance matrix of observations
- C_{N+1}: Covariance matrix between observations and predictand

Observations

(Farm 1, day 1)	(Farm 1, day 2)	(Farm 1, day 3)		*Predictand*
(Farm 2, day 1)	(Farm 2, day 2)	(Farm 2, day 3)	$\Rightarrow$	(Farm k, day 4)
⋮	⋮	⋮		
(Farm 19, day 1)	(Farm 19, day 2)	(Farm 19, day 3)		

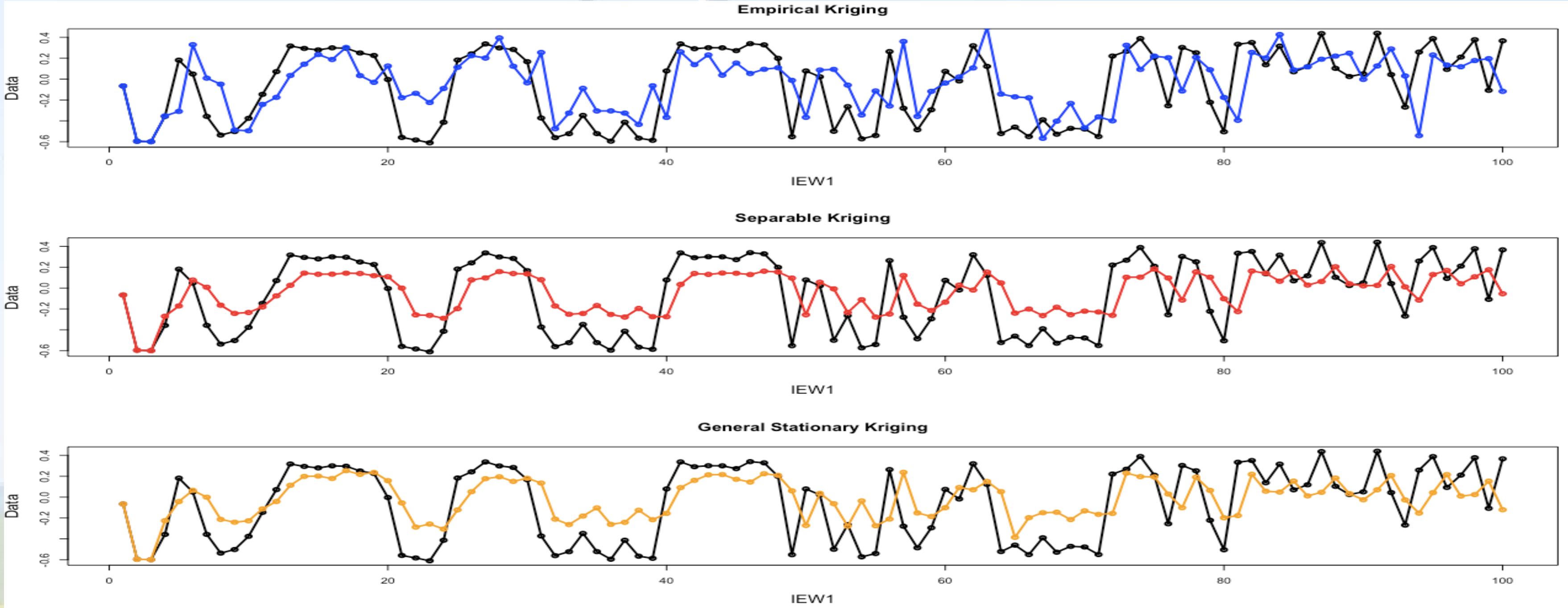
03

APPLICATION

Prediction for Individual Wind Farm

Kriging Prediction （existing wind farm）

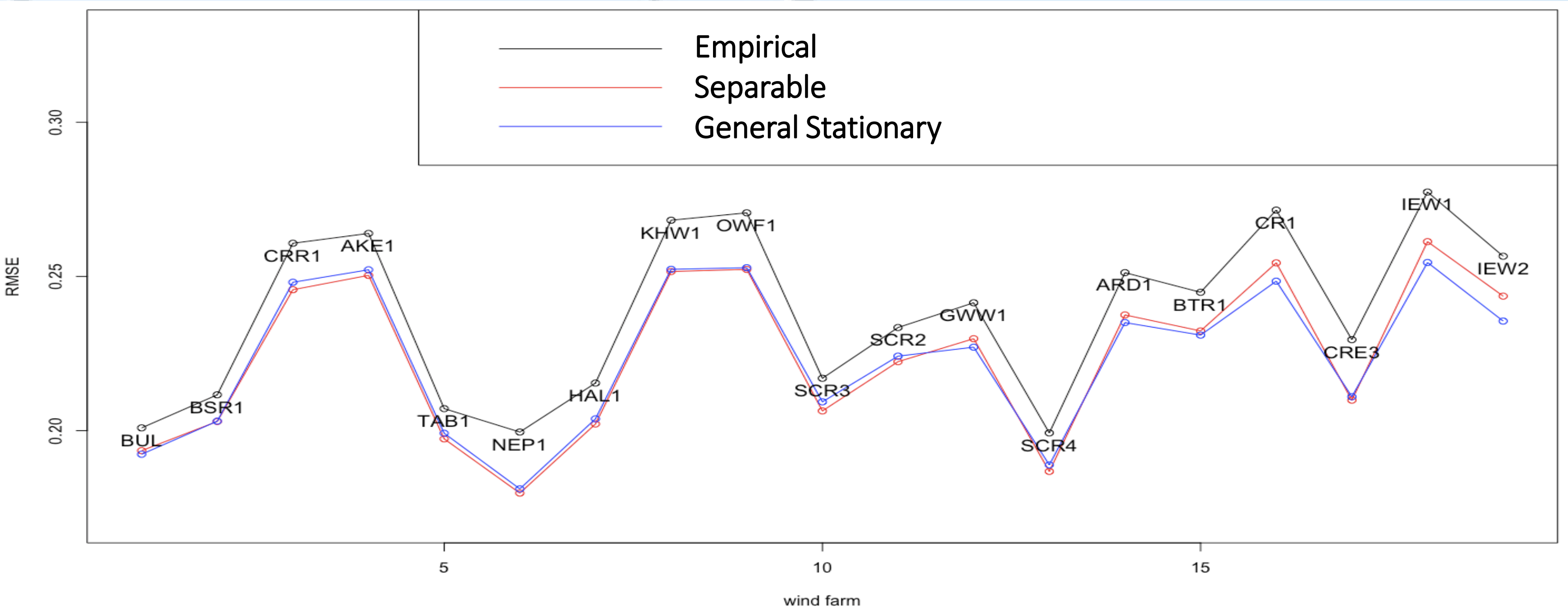
- Empirical
- Separable
- Stationary



03 APPLICATION

Prediction for Individual Wind Farm

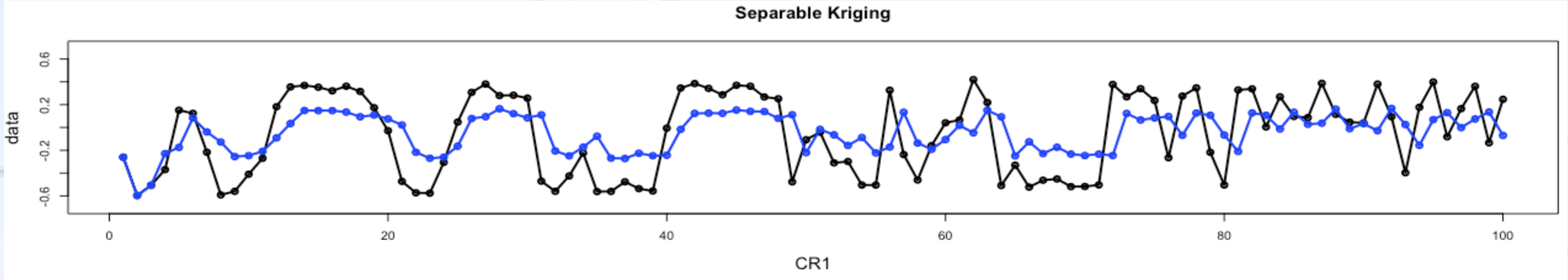
Kriging Prediction （existing wind farm）



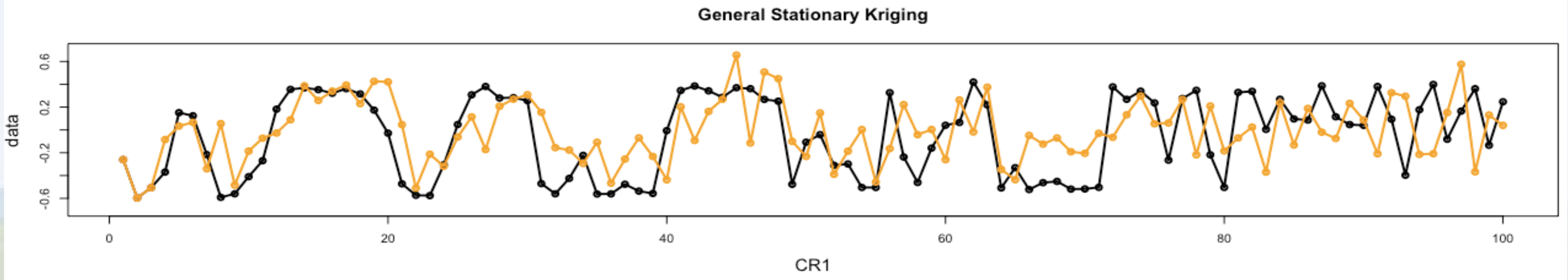
Prediction for Individual Wind Farm

Kriging Prediction（new wind farm）

- Separable



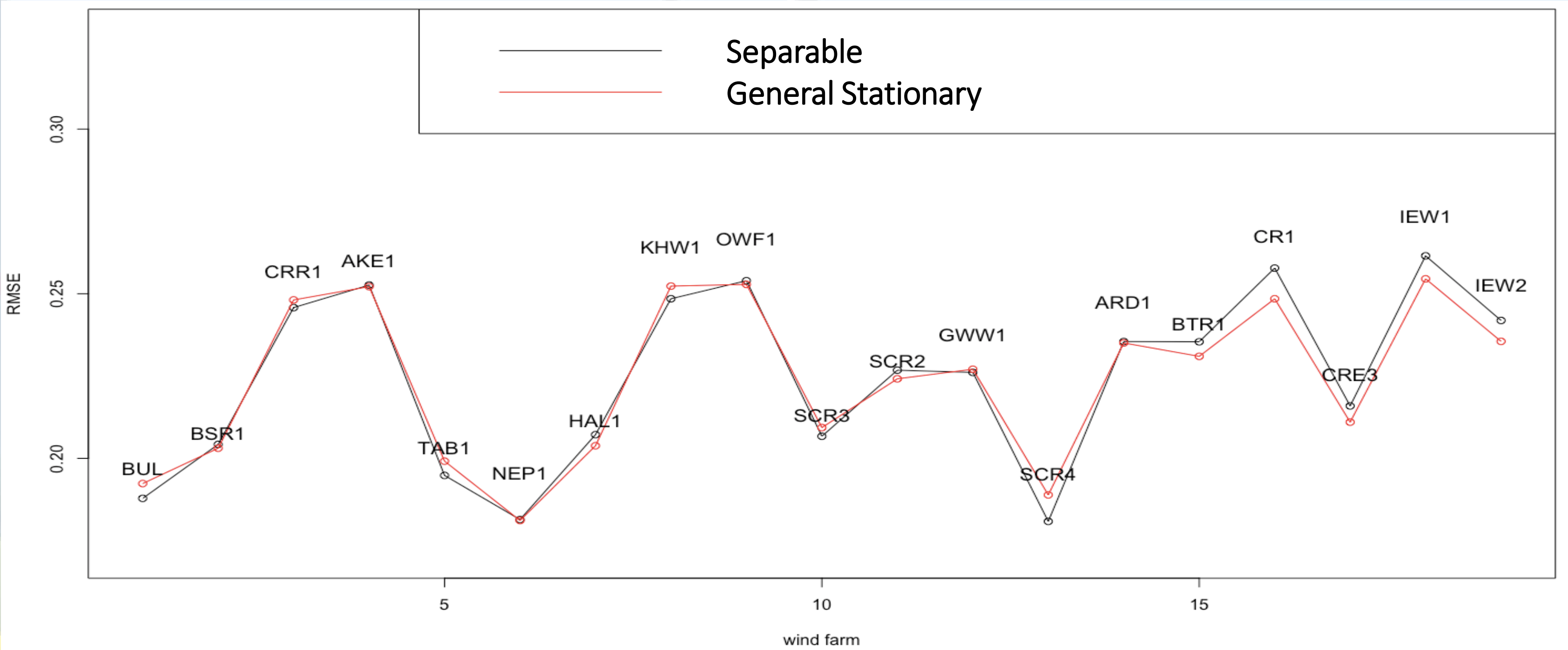
- Stationary



03 APPLICATION

Prediction for Individual Wind Farm

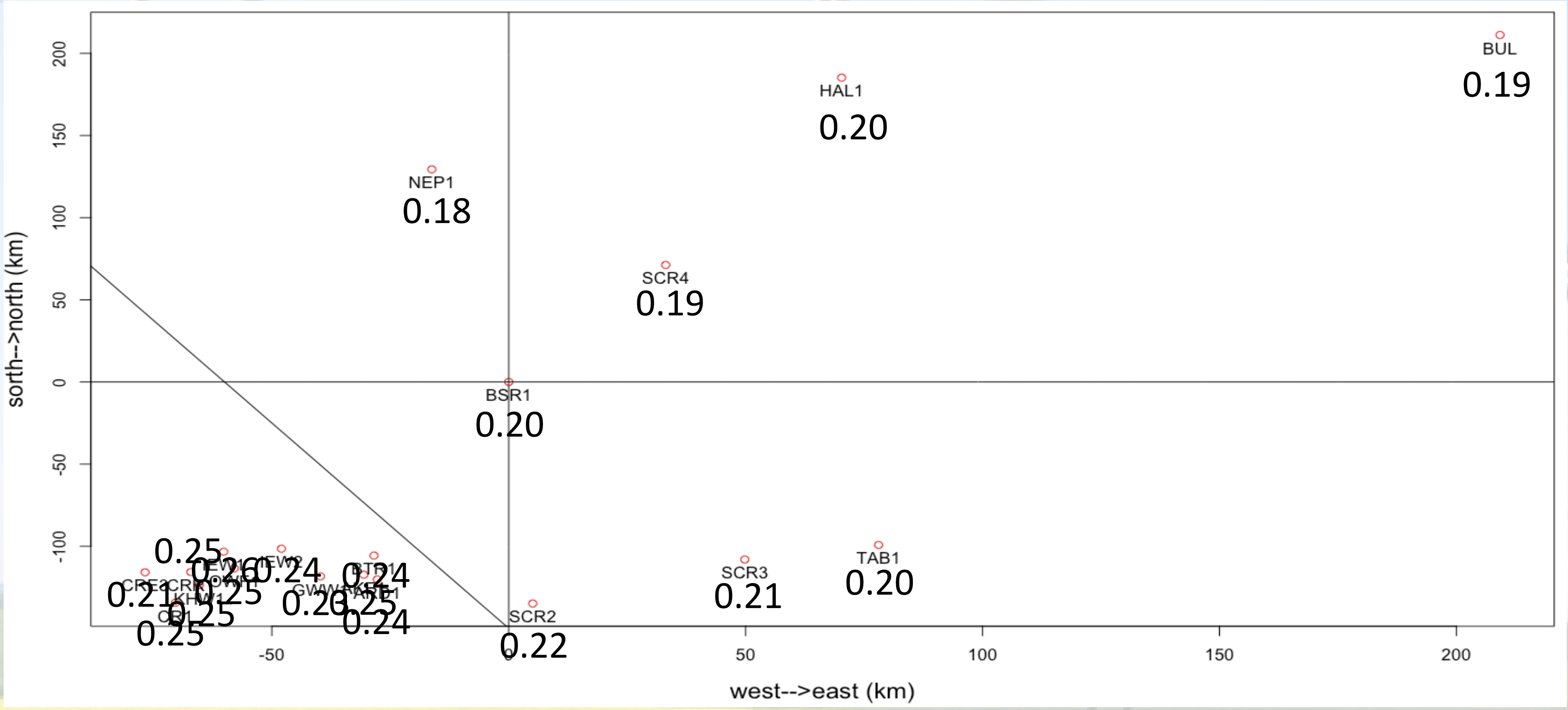
Kriging Prediction（new wind farm）



03

APPLICATION

Prediction for Individual Wind Farm



Aggregate Wind Power Generation

Definition

- $W(x,t)$: Actual daily-average wind power generation divided by capacities.
- $Y(x,t)$: Spatial-temporal stationary Gaussian process
- $S(t)$: Seasonality at time t
- $m(x)$: Station-specific mean at site x

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

Aggregate Wind Power Generation

Definition

- Aggregate wind power generation at time t^*

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

- Locations of current wind farms: $X^c = \{x_1^c, \dots, x_k^c\}$
- Locations of future wind farms: $X^f = \{x_1^f, \dots, x_p^f\}$
- Capacities of current wind farms: $C^c = \{c_1^c, \dots, c_k^c\}$
- Capacities of future wind farms: $C^f = \{c_1^f, \dots, c_p^f\}$

Aggregate Wind Power Generation

"Best" Scenario of new wind farms

Locations	Means	Variances
one farm above the split line	1076.8	683931.5
two farms above the split line	1073.8	654072.1
three farms above the split line	1071.3	627529.3
four farms above the split line	1068.8	601594
five farms above the split line	1066.2	575885.5

04

R PACKAGE

- Design of R package

Design of R Package

R Package: YLWind

Data Input:

1. Wind energy generation

STdata = matrix(data, dimnames=list(c(time),c (names of wind farms))

2. Locations of wind farms

Loc = data.frame(longitude, latitude)

3. Wind data (optional)

Wind = data.frame(direction, speed)

Design of R Package

R Package: YLWind

Preprocessing Functions:

1. Wind energy generation → Gaussian process

Fn: PreST(object = STdata, train = [1: n,], test = [-1:n,], histogram = T)

Return: two matrices (trainset & testset)

2. Locations of wind farms → Distance

Fn: Predists(object = Loc, longlat = T)

Return: distance matrices (dists & londists & latdists)

3. Wind data (optional) → Average wind direction and wind speed

Fn: Avgwind(object = Wind)

Return: two numbers (wv & sv)

Design of R Package

R Package: YLWind

Class: ywind

$Data

List(STdata, Loc, Wind (optional))

$model

list(“separable”, “Fully Symmetric”, “General Stationary”))

Summary.ywind: parameter optimization.

Predict.ywind: Kriging prediction values, RMSE and MAE errors.

Plot.ywind: Goodness of fit plots, Kriging prediction plots.

Agg.ywind: mean and variance of the aggregate wind power generation.

Thank You !

Contact Info:

Yilan Luo
Email: yilan.luo1@ucalgary.ca
Phone: (587) 969-3444