Changepoint modelling in spatio-temporal processes with application to Ireland wind data

StatScale ECR Meeting, 14-16 Dec 2022, Brighton

Gaurav Agarwal
Coauthors: Idris Eckley and Paul Fearnhead

Mathematics & Statistics
Lancaster University

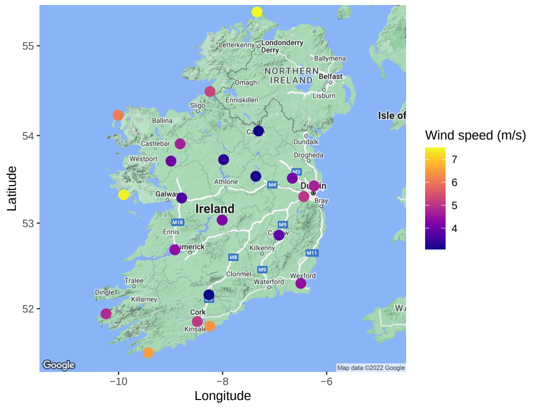
Motivation: Ireland Wind Data

- global goal of **sustainable environment**, wind power clean energy source
- vital to evaluate country's wind resource

StatScale Lancaster University

Motivation: Ireland Wind Data

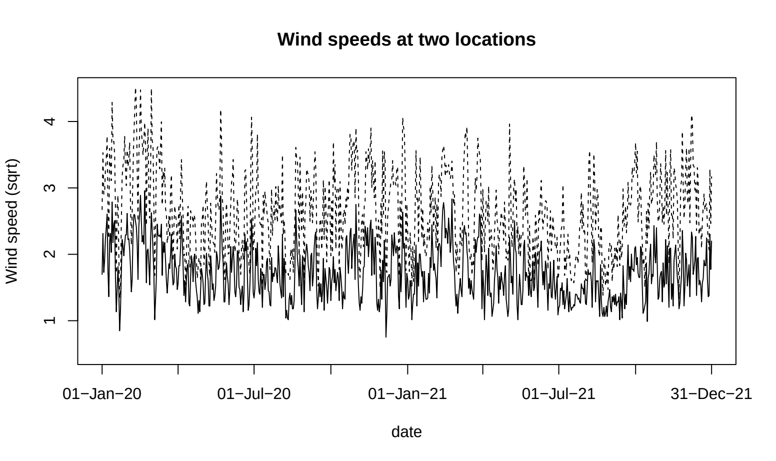
Spatio-temporal data: data indexed over space and time



Spatial distribution of wind speed, averaged over a two year period (1-Jan-2020 to 31-Dec-2021); may depict a change in pattern of wind speed over time.

Motivation: Ireland Wind data

Time series plots



evidence of spatial and temporal dependencies

Motivation

- Changepoints extensively studied for time series, but **limited literature** for change in spatio-temporal processes.
- Increasingly, spatio-temporal data is becoming **routinely available**.
- Assumption of stationarity for practical convenience.
- Spatio-temporal data exhibits **nonstationarity**, requiring methods for detecting changes in the process.

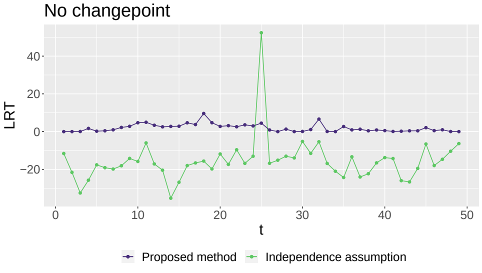
StatScale | Lancaster University

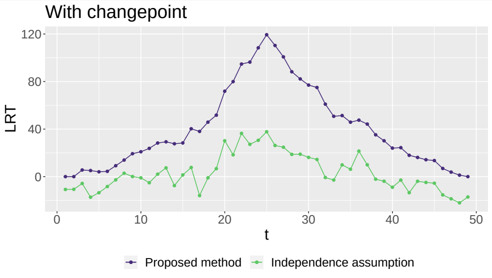
Existing Methods

- Bayesian methods such as Majumdar et al. (2005) and Altieri et al. (2015) usually have very **specific model** structures for spatio-temporal changes.
- Gromenko et al. (2017) considered at most one change point, assuming **separability** of space and time.
- Recently, Zhao et al. (2019) defined a nonseparable pairwise likelihood-based approach.
- However, for simplicity, assume **data across segments are independent**.

Motivation- Issues of independent segment assumptions

statistical inconsistencies





StatScale Lancaster University

Methodology: Background

- Let $Y(\boldsymbol{s}, t)$ be a stochastic process, where $\boldsymbol{s} \in \mathbb{R}^d$: location, and $t \in \mathbb{R}$: time point.

- For each time $t = t_1, \ldots, t_T$, there are m observations.

- $\boldsymbol{Y} = \left\{Y(\boldsymbol{s}_1, t_1), \ldots, Y(\boldsymbol{s}_m, t_1), \ldots, Y(\boldsymbol{s}_1, t_T), \ldots, Y(\boldsymbol{s}_m, t_T)\right\}^\top$, with $n = mT$ observations.

- The vector $\boldsymbol{Y} \sim \mathcal{MVN}(\boldsymbol{\mu}_{n\times 1}, \boldsymbol{\Sigma}_{n\times n})$, where $\boldsymbol{\Sigma} = [\text{Cov}\{Y(\boldsymbol{s}_i, t_i), Y(\boldsymbol{s}_j, t_j)\}]_{i,j=1}^{n}$ is the **covariance** matrix and $\boldsymbol{\mu} = [E\{Y(\boldsymbol{s}_1, t_1)\}, \ldots, E\{Y(\boldsymbol{s}_m, t_T)\}]$.

Model

Consider a spatio-temporal process

$$Y(\boldsymbol{s}, t) = \mu(\boldsymbol{s}, t) + \epsilon(\boldsymbol{s}, t),$$

$t = 1, \ldots, T$ and $\boldsymbol{s} \in \mathcal{S}$, where $\mu(\boldsymbol{s}, t)$ is deterministic mean and $\epsilon(\boldsymbol{s}, t)$ is the mean-zero error process.

- Use a parametric nonseparable **nonstationary** space-time **covariance function**, which allows spatial process to evolve over time (Qadir, G.A. and Sun, Y., 2022).
- Considers **time-varying** spatial parameters.
- Cov$\{Y(\boldsymbol{s}, t_i), Y(\boldsymbol{s} + \boldsymbol{h}, t_j)\} = C(\boldsymbol{h}, t_i, t_j)$

Change in spatial process

Denote all the spatial covariance parameters by $\boldsymbol{\theta}$, nuisance parameters by γ, the null hypothesis of no change is

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \ldots \boldsymbol{\mu}_T \quad \text{and} \quad \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2 = \ldots \boldsymbol{\theta}_T = \boldsymbol{\theta}_0,$$

against change in mean or spatial dependence

$$H_{A_1} : \boldsymbol{\mu}_1 = \ldots\ldots \boldsymbol{\mu}_\tau = \boldsymbol{\mu}^{(1)} \neq \boldsymbol{\mu}_{\tau+1} = \ldots \boldsymbol{\mu}_T = \boldsymbol{\mu}^{(2)}, \quad \text{or}$$

$$\boldsymbol{\theta}_1 = \ldots\ldots \boldsymbol{\theta}_\tau = \boldsymbol{\theta}^{(1)} \neq \boldsymbol{\theta}_{\tau+1} = \ldots \boldsymbol{\theta}_T = \boldsymbol{\theta}^{(2)}.$$

Likelihood ratio test statistics:

$$\text{LR}_\tau = -2\ln\left[\frac{\max_{\boldsymbol{\mu},\boldsymbol{\theta_0},\gamma_0} f_{\boldsymbol{Y}}(\boldsymbol{y};\boldsymbol{\mu},\boldsymbol{\theta}_0,\gamma_0)}{\max_{\boldsymbol{\mu}^{(1)},\boldsymbol{\mu}^{(2)},\boldsymbol{\theta}^{(1)},\boldsymbol{\theta}^{(2)},\gamma_1} f_{\boldsymbol{Y}}\{\boldsymbol{y};\boldsymbol{\mu}^{(1)},\boldsymbol{\mu}^{(2)},\boldsymbol{\theta}^{(1)},\boldsymbol{\theta}^{(2)},\gamma_1\}}\right].$$

StatScale Lancaster University

Markov Likelihood and Estimation

- Log-likelihood of $\boldsymbol{Y}$: $\ell(\boldsymbol{\mu}, \boldsymbol{\theta} \mid \boldsymbol{Y}) = -\{\log\det\Sigma(\boldsymbol{\theta}) + (\boldsymbol{Y} - \boldsymbol{\mu})^\top \Sigma(\boldsymbol{\theta})^{-1}(\boldsymbol{Y} - \boldsymbol{\mu}) + n\log 2\pi\}/2$
- **computationally challenging**: it involves the inverse of a high-dimensional covariance matrix, complexity $O(m^3T^3)$
- storing very large matrices can exhaust the memory of the machine
- To **overcome** the computational issues, we implement a **Markov likelihood** approach for optimization

Markov Likelihood and Estimation

- The joint distribution of $\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_T$:

$$\{\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_T\} = \{\boldsymbol{Y}_T \mid \boldsymbol{Y}_{T-1}, \dots, \boldsymbol{Y}_1\} \dots \{\boldsymbol{Y}_3 \mid \boldsymbol{Y}_2, \boldsymbol{Y}_1\}\{\boldsymbol{Y}_2 \mid \boldsymbol{Y}_1\}\{\boldsymbol{Y}_1\}$$

- most real-world spatio-temporal processes can be characterized conditional on the process in the **recent past**.
- First-order Markov assumption, joint distribution

$$\{\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_T\} = \prod_{t=2}^{T} \{\boldsymbol{Y}_t \mid \boldsymbol{Y}_{t-1}\} \times \{\boldsymbol{Y}_1\}.$$

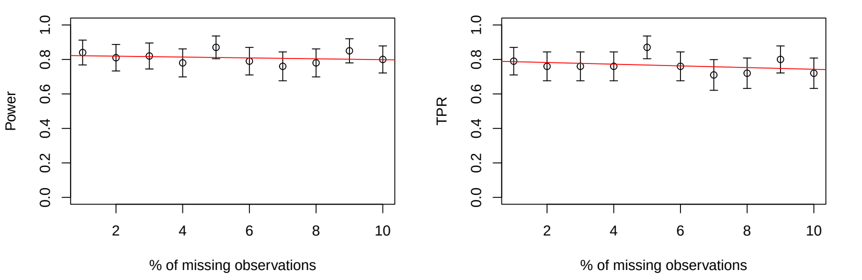
- relatively simple conditional distributions, **complexity reduced** to $O(m^3 T)$ from $O(m^3 T^3)$.

Optimistic search and multiple changepoint

- Searching through all candidate changepoints $\tau = 1, \ldots, T-1$, computationally demanding.
- Each point requires fitting a spatio-temporal model.
- Use optimistic search strategy (Kovács et al., 2020), adaptively determine the next search point, $O(\log T)$ evaluations instead of $O(T)$.
- Combine with binary segmentation for multiple changepoints.

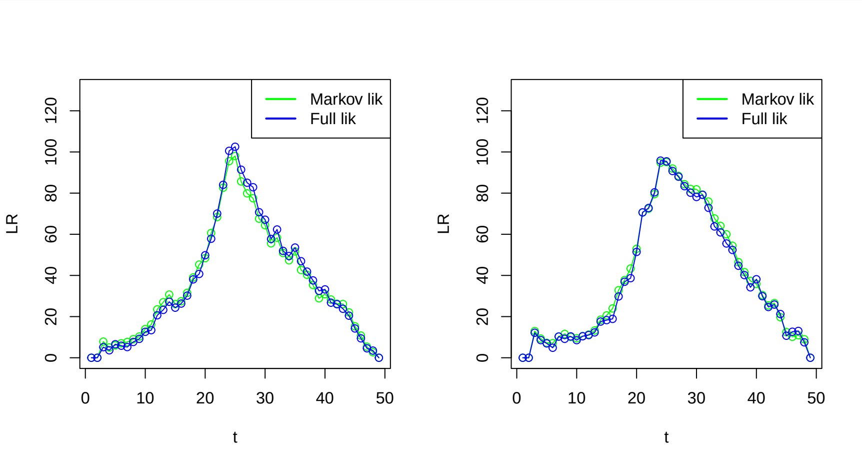
Simulation study: missing observations

method works with missing observations, quite common in spatio-temporal data



Simulation study: Markov approximation

comparison of Markov likelihood with Full likelihood



Simulation results

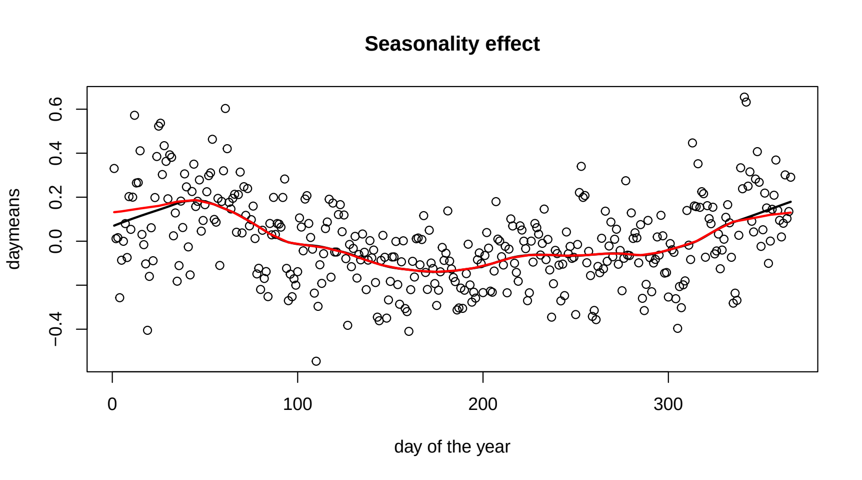
change in spatial dependence

size of change	Method	Power	TPR
0.025	Proposed (Markov Lik)	0.7	0.4
	Proposed (Full Lik)	0.74	0.43
	Ind Segment (Pair Lik)	0	0
0.05	Proposed (Markov Lik)	1	0.91
	Proposed (Full Lik)	1	0.95
	Ind Segment (Pair Lik)	0.4	0.1
0.2	Proposed (Markov Lik)	1	1
	Proposed (Full Lik)	1	1
	Ind Segment (Pair Lik)	0.7	0.2
0.3	Proposed (Markov Lik)	1	1
	Proposed (Full Lik)	1	1
	Ind Segment (Pair Lik)	0.95	0.5

Application: Ireland Wind data

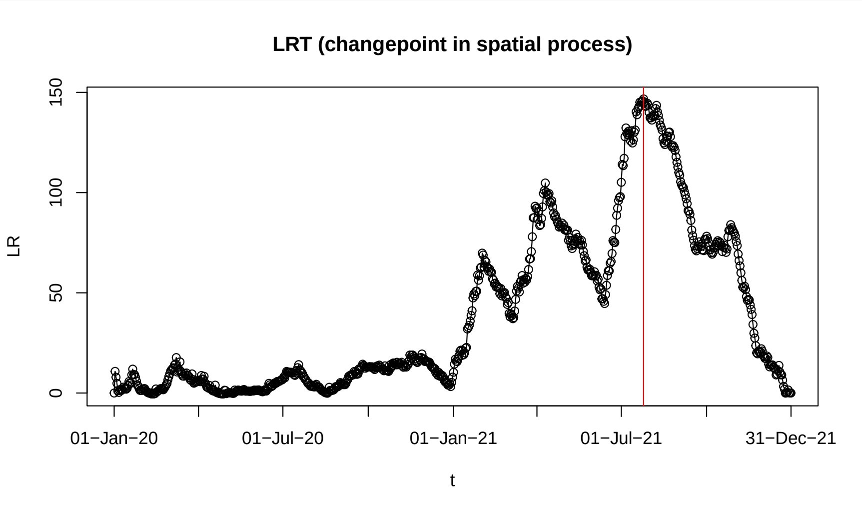
- The Irish Meterological Service (**Met Éireann**) provides wind speed data at **22 synoptic weather** stations across the country.
- This data can be used to monitor, analyse and predict Ireland's wind power resource.
- We consider daily **wind speed** data over a **2 year** period for the 22 synoptic weather stations.
- The data is available at https://www.met.ie/climate/available-data.

Application: Ireland Wind data



seasonality estimated using 5 years historical data from 2015-2019, smoothed at start and end of year (in red)

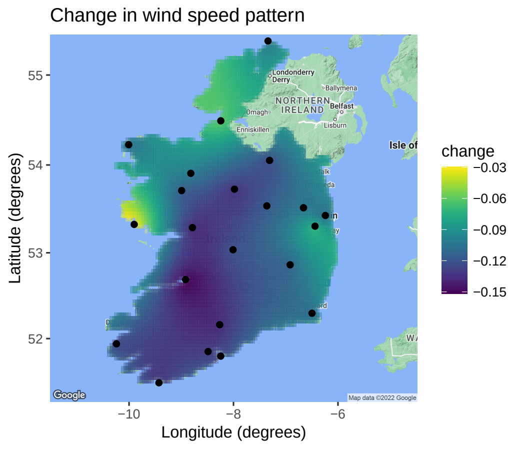
Application: Ireland Wind data



changepoint: 24 july 2021, threshold 20-30

reason: climate statement about heatwaves

Application: Ireland Wind data



changes higher at some locations than others.

Application: Ireland Wind data

- Use spatio-temporal kriging for **prediction**.
- Left 20% data out to check the prediction accuracy with changepoint model (CP) and compare with a no-change model (No CP).
- Point prediction: Root Mean Squared Error (RMSE)
- Probabilistic prediction: Continuous Rank Probability Scores (CRPS), Logarithmic Score (logS)

Metric	No CP	CP
RMSE	0.54	0.52
CRPS	0.32	0.29
logS	0.87	0.83

Summary

- A likelihood-based methodology for estimation of **changepoints** and model parameters of **spatio-temporal processes**.
- Fit a nonstationary changepoint model **without any independent segment** assumption
- computationally efficient **Markov approximation**
- changepoint detection and missing data prediction in daily **wind speeds** across different synoptic weather stations in Ireland over two years

Thank you for your attention (g.agarwal@lancaster.ac.uk).