Time-series analyis
Space-time geostatistics

D G Rossiter

College of Agriculture and Life Sciences, Cornell University

April 23, 2018

Outline

Outline

Topic: Concepts of space and time

1. space
2. time
3. space and time

Space

- Refers to processes that occur in **space** (1D, 2D, 3D)
- Often there is **spatial dependency**
  - "nearby" locations have more similar attributes than "far away"
- 2D or 3D dependency may be **isotropic** (same in all directions) or **anisotropic** (stronger in one direction)
- Anisotropic dependency is **symmetric**: the same "forward" or "backward" in each 1D
  - e.g., main axis of long-range 2D dependence 30° (NNE) is the same as $30^\circ + 180^\circ = 210^\circ$ (SSW)
- This is captured in the definition of **spatial autocorrelation**
  - differences are squared; there is no "head" or "tail" of a point pair

Time

- Refers to processes that occur in **time**
- Usually there is **temporal dependency**
  - attribute values do not change (completely) randomly over time
  - they **evolve**, and closely-spaced events have similar values
- Time is 1D and **asymmetric** because time **only flows in one direction**
  - It is a deep philosophical question as to why, but it is a consistent observation
  - So, "forward" and "backward" cross-dependencies may be different
  - However, "forward" and "backward" auto-dependencies are identical

Space-time

- Refers to processes which occur in **both** space and time
- There may be either or **both** spatial and temporal dependence
- These may be **independent or not**

Concepts of time

time stamp	a moment in time, to some resolution
snapshot	the state of nature at one moment in time
interval	time between two snapshots
linear	a series of moments
cyclical	time considered as a repeating cycle
discrete	separate moments in time
continuous	moments in time follow each other continuously

Time-related variables

instantaneous measured directly at one point in time

- e.g., air T, ground-water depth
- e.g., location of an observation

interval directly measured over some known interval

- e.g. stream flow; mean air T using an integrating recorder

cumulative measured at one time from an accumulation over time

- e.g., daily rainfall from an accumulating rain gauge

Temporal aggregation

- Upscaling by a representative time
  - e.g., minimum daily T from a single measurement at 0500
- Upscaling by aggregation
  - e.g., monthly GDD from daily GDD
- How to aggregate?
  - sum (cumulative)
  - mean, median . . . (central tendency)
  - min, max . . . (extremes or quantiles)
  - . . .

Outline

Topic: space-time processes

All processes occur over some time period (long or short); many processes take place in geographic space.

Space-time processes may be viewed as:

1. **spatial** only: time is not important
   - consider at one "instant" (conceptually 0-dimensional time)
   - reduce to one temporal point by a summary statistic (average, maximum . . .)
2. **temporal** only: space is not important
3. **spatio-temporal**: space and time must both be considered

Space-time processes and their related variables

processes dynamic, cause effects at locations in space

- e.g., atmospheric processes cause weather

variables what is observed at defined coördinates in space and/or time

- e.g., measured temperature, barometric pressure, precipitation . . .
- measurements referred to a specific time may be "instantaneous", time-averaged, or cumulative

A taxonomy of space-time processes (1)

- Purely **spatial**
  - observations all refer to **one time** or time period
    - represented by a single **time stamp**, which can represent any length of time
  - this can be an **aggregate** time, e.g., average or cumulative over some time period
  - any difference in time of observation is considered unimportant ("nothing" has changed in the intervening time)
    - or, temporal differences are negligibly small compared to spatial differences

...

A taxonomy of space-time processes (2)

. . .

- Purely **temporal**
  - observations all refer to **"one" location**
    - represented as a point or "homogeneous" area
  - this can be an **aggregate** location, e.g., average of a set of weather stations
  - spatial differences (e.g., small offsets) are negligible
    - or, spatial differences are negligibly small compared to temporal differences
  - or we only care about the spatial area as a unit

. . .

A taxonomy of space-time processes (3)

. . .

- **Spatio-temporal**
  - observations have **both** a location and a time stamp
  - observations at different locations may have been made at different times
  - or there may be the same time-series of observations at each location
  - space and/or time may be aggregates

Example of space-time processes: Soil organic carbon (SOC) 1 – purely spatial

- soil samples in an **agricultural field**, all collected at the **“same”** time
  - time differences in sampling are insignificant compared to process time (e.g., SOC decomposition)
  - aim is to **map** SOC distribution in the field and relate to other soil or land properties
- soil samples in a **forest**, collected over a **number of years**
  - aim is to assess SOC stocks in the area
  - must **assume** no drivers of SOC change in this time, no seasonal effects – or at least these are negligibly small compared to spatial variation
  - must justify this assumption from literature or small time-series studies

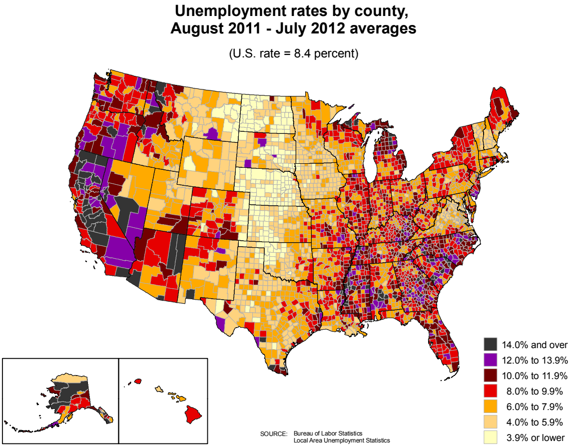
Example: SOC 2 – purely temporal

- **Repeat** soil samples at **one** location before and after manure and fertilizer applications, crop growth, crop harvest, residue incorporation, winter weather . . .
  - Aim is to reveal SOC dynamics as influenced by weather and management
  - Since soil sampling is inherently **destructive** it is impossible to sample repeatedly at the same exact location
    - If using close-by locations micro-scale spatial variation and small support will cause problems.
    - Solution: use composite sampling from the one location (i.e., support of several tens of m^2)

Example: SOC 3 – Spatio-temporal

- **Repeat** samples at **multiple** sites in an agricultural field
  - Aim is to discover if there are different temporal dynamics at different locations
  - Can also **map** more efficiently (more information), and produce maps for different times
  - May relate to spatio-temporal **covariables** (e.g., weather)

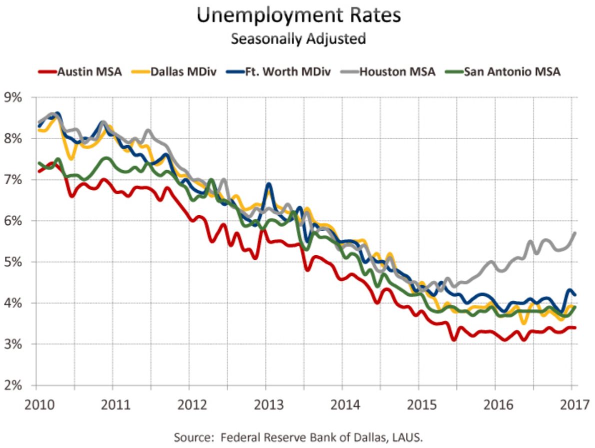
USA unemployment rate – spatial



source:

http://moneybasicsradio.com/2012/09/us-unemployment-rate-by-county/

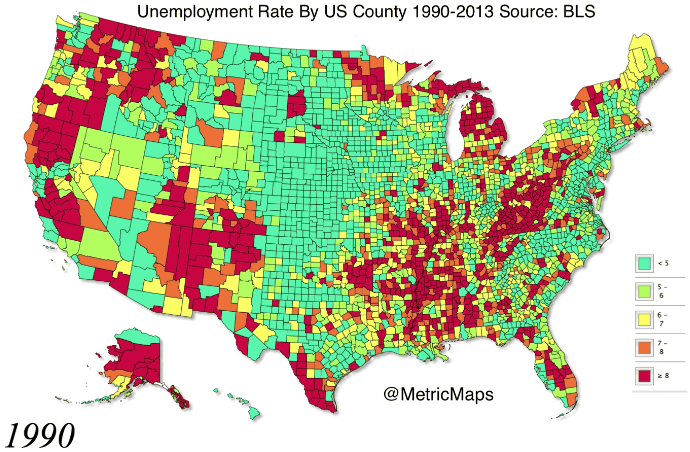
USA unemployment rate – temporal



source:

https://www.austinchamber.com/blog/03-14-2017-job-growth-unemployment

USA unemployment rate – spatio-temporal



source: https://www.huffingtonpost.com/2014/08/31/america-unemployment-map_n_5744656.html

What could be analyzed?

1. **Time-series** of unemployment; could examine each county's time series separately
2. **Spatial dependence** at each time slice
3. **Spatio-temporal interaction**: The spatial pattern is *not the same* at each time slice

A taxonomy of spatio-temporal analysis

Following the taxonomy of **processes**, how do we **analyze** them?

1. **spatial** at each time slice, **independently**
   - point geostatistics, point-pattern analysis, spatial autoregressive models ...
2. **temporal** at each location, **independently**
   - time-series analysis [9]; see next slide
3. spatial at each time slice, but using a **temporally-pooled** estimate of spatial dependence
   - spatial statistics as (1) but each time of observation is considered a **replication** of the spatial dependence
4. spatio-temporal **as a combined model**
   - both spatial and temporal statistics

Outline

Topic: Time-series analysis

Only considering time:

1. a series of observations over time at **one location**
   - **single** time series
     - one variable measured at different times
   - **multiple** (also called "multivariate") time series
     - several variables, measured together at the same observation times
2. same, at **multiple locations**
   - single or multiple time series at each location
   - but with **no explicit spatial dimension**
     - no coördinates, no distance
     - so it is not geostatistical
   - a type of multiple time series: attribute is the same but at different (non-georeferenced) locations

Single time-series

A series of readings of the **same** variable at the **same** location, at **different** times

- Often at regular intervals (15-min, hourly, daily, . . .)
- But can be at irregular times, although analysis becomes much harder

Components of a single time-series

Up to six components *may* be present:

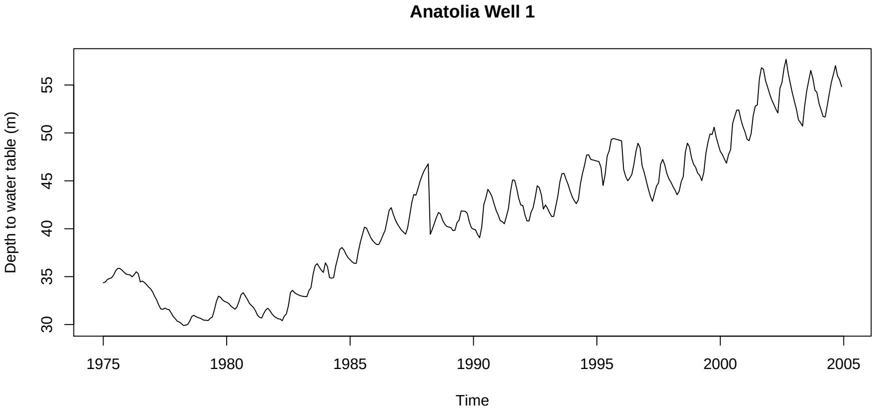
- **trend**: the variable changes **systematically** with time
  - e.g., increasing daily mean T
- **periodic**: the variable **fluctuates** around a central value, with a fixed period
  - e.g., hourly T over a day; daily T over a year
- **cyclic but non-periodic**: the **fluctuates** around a central value, with a variable period
  - e.g., predator-prey abundances
- **anomalies**: unusual values of the variable that do not fit a pattern ("spikes")
  - e.g., daily T during an unusual weather event
- **correlated noise**: "small" variations after accounting for the above, with **temporal dependence**
- **white noise**: **uncorrelated** "small" variations

Main analysis problem

How to identify, separate and quantify these?

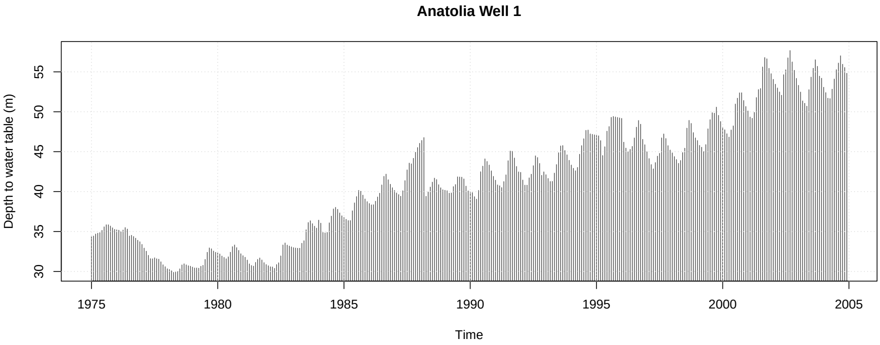
1. decomposition
2. autocorrelation analysis
3. modelling

Example: Groundwater level: time series



Monthly 1975-2005; irrigation well in Anatolia (Turkey)

Time series as bar graph

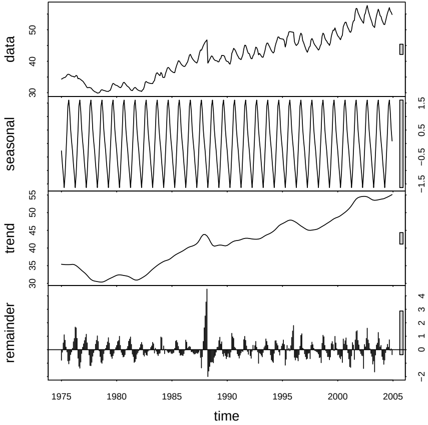


Monthly 1975-2005; irrigation well in Anatolia (Turkey)

Time-series analysis – (1) decomposition

- **Decomposition**: identify the components of the series
  1. identify and remove **periodic** component
     - single amplitude or moving amplitude, user-specified “span”
  2. fit a **trend** to the non-periodic component by local polynomial regression (“Locally-weighted scatterplot smoother” = “lowess”)
     - user-specified “span” (smoothing window) *to* degree of smoothing
  3. the residual “noise” = **anomalies** from trend and cycles
- These components should correspond to **processes** that produced the time series

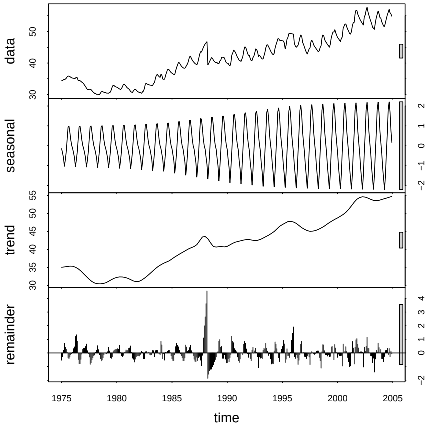
Groundwater level: decomposition (1)



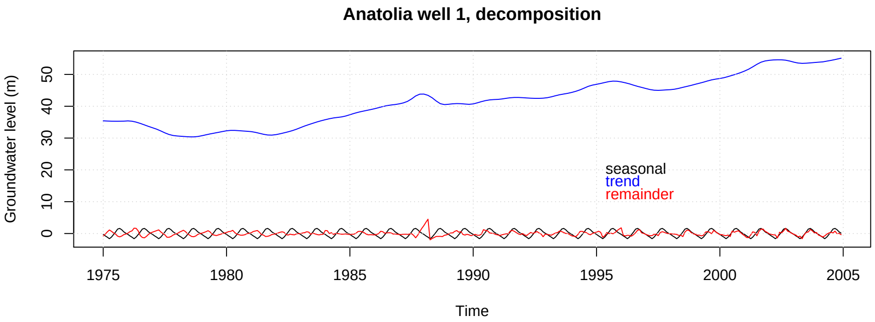
Trend modelled with a moving window

Yearly cycle assumed the same – does this seem correct?

Groundwater level: decomposition (2)



Magnitude of yearly cycle increases with time – reduces remainder



Trend and remainder after subtracting trend (note absolute values)
Note large anomaly in 1988

Interpretation of time-series decomposition

- Aim: understand the underlying process that produced the time-series
  - Example: daily T cycle driven by insolation (clue: the cycle is 24 hours!)
- Aim: determine the magnitude of changes (trend) or persistence of phenomena (anomalies)
  - and from that, infer causes

Interpretation: groundwater example (1)

- Groundwater was closer to the surface from 1975–1980 and then began a steady **trend** to become deeper
  - Increased groundwater extraction for irrigation? Decreased winter rainfall for recharge?
- There is a **yearly cycle**
  - explained by extraction for irrigation (summer) and recharge from rainfall and irrigation excess (winter)
- This **cycle seems to be getting stronger**
  - suggesting increasing irrigation demand
- A strong **anomaly** of draw-down and then recharge was observed for 1988 – why?
- **Residual noise** may have temporal autocorrelation – see next slides

Time-series analysis – (2) auto-correlation analysis

- **Auto-correlation analysis**: time dependence of observations
  - of the original series
  - after de-trending and/or removing any cycles (i.e., after decomposition)
- Question: how strongly are observations linked in time?
- Question: how long does this autocorrelation last ("range")?

Serial autocorrelation

If the random process that generated the time series is **2nd order stationary** (mean and variance are constant over time), at lag k the autocorrelation is:

$$\rho_k = \frac{E\left[(z_t - \mu)(z_{t+k} - \mu)\right]}{\sigma_z^2} \quad (1)$$

ρ_k can be estimated as r_k:

$$r_k = c_k / c_0 \quad (2)$$

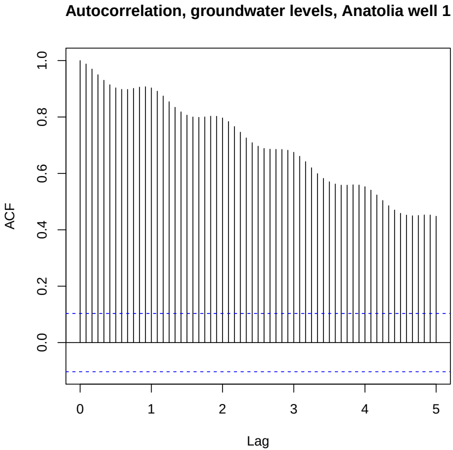
$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (z_t - \bar{z})(z_{t+k} - \bar{z}), k = 0, 1, 2 \ldots K, \quad (3)$$

$$\mathrm{var}[r_k] \approx \frac{1}{N} \left(1 + 2 \sum_{v=1}^{q} r_q^2 \right), k > q \quad (4)$$

Use of autocorrelation graphs

- **Original** series: assume first- and second-order stationarity
- **Remainder** series (after subtracting a trend): only assume constant variance and co-variance for all lags
- Use the approximate formula for variance (previous slide) to establish confidence limits: is the observed correlation at a given lag significantly non-zero?

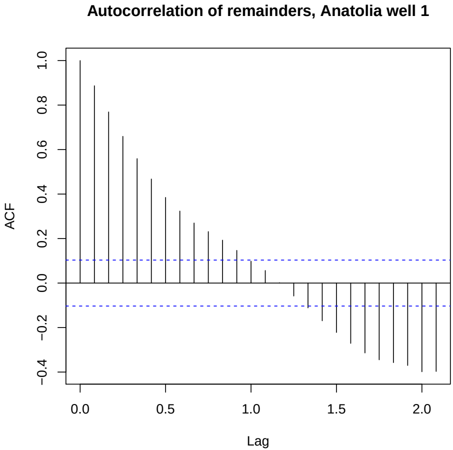
Groundwater level: autocorrelation – original series



Lag 1 is one year

Autocorrelation increases exactly on the year, but less each year

Groundwater level: autocorrelation – remainders



No autocorrelation at one year; negative at two years!

Artefact of moving average?

Interpretation: groundwater example (2)

- Original time series shows strong positive autocorrelation
  - decreases over a year, but then increases a bit on the yearly cycle
  - steadily decreases over the years, but still $\rho \approx 0.5$ after five years
- **Residual noise** is positively autocorrelated for one year
  - this reflects **continuity** month-to-month

Partial autocorrelation

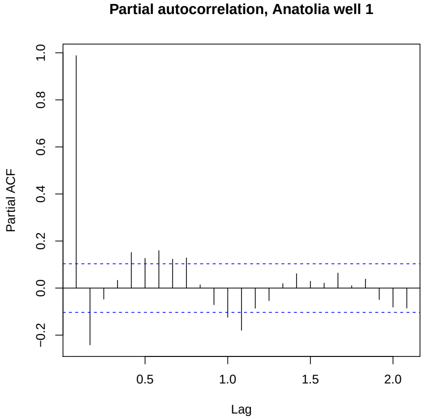
- Autocorrelation *after* accounting for previous lags; symbol $\phi_{k,j}$: coefficient j of an autoregressive process of order k
- Example: if all autocorrelation can be explained at lag 1, then there is no partial autocorrelation at lags 2, 3, ..., so that the apparent autocorrelation at these lags can be explained by repeated lag-1 correlations

$$\mathbf{P_k}\phi_\mathbf{k} = \rho_\mathbf{k} \quad (5)$$

$$\mathbf{P_k} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ . & . & . & \cdots & . \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \quad (6)$$

$$\rho_j = \phi_{k,1}\rho_{j-1} + \cdots + \phi_{k,k}\rho_{j-k},\ j = 1, 2, \ldots k \quad (7)$$

Groundwater level: partial autocorrelation



Lag 1 month strongly positive
Lag 2 slightly negative after accounting for lag 1
Lag 4–8 (half-season), 12 (?), 13 (full season) significant

Time-series modelling: AR, MA, ARMA, ARIMA

- Aim: a mathematical description of the series
  - Generally used for **forecasting**
    - assuming that the statistical characteristics of the time-dependent process are the same in the future as in the past
  - Also can be used to **understand** the process from the form of the best-fitted model
  - Also can be used for **gap filling** of incomplete series
- Model types: AR (auto-regressive), MA (moving average), ARMA, ARIMA
  - I = "integrated", for the degree of differencing applied to the series before ARMA)
- See Shumway and Stoffer [9], Wilks [11, Ch. 9], Box [1]

Autoregressive models (AR)

- **Current** value of a process is expressed as a **finite** and **linear** sum of **previous** values of the process
- this is the **temporally autocorrelated** noise

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \cdots + \phi_p \tilde{z}_{t-p} + a_t \tag{8}$$

- ϕ_i are the autocorrelation parameters (strength of dependence at each lag)
- a_t is the white noise, sometimes called "shock" at time t
- simplest is AR(1): $\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + a_t$: dependence only on previous value

Time-series modelling: spectral analysis

- Fourier (1807): any *second-order stationary* time-series can be **decomposed** into **sums of sines and cosines** with increasing **frequencies**, each of varying **amplitude** or "power"
  - trend removed and auto-covariance not dependent on position in the series
- **Frequency** ω: the number of divisions of one **cycle** per unit time
  - e.g., a one-year cycle (as in the groundwater levels), $\omega = 12$ is a monthly frequency
  - By the Nyquist-Shannon sampling theorem, a function is completely determined by sampling at a rate of $1/2\omega$
  - a time-series with n samples per cycle $\rightarrow$ can estimate spectral densities for $n/2$ frequencies.
- **Period** T: inverse of frequency: number of divisions required for one full cycle; $T = (1/\omega)$

Why spectral analysis?

- Reveals the relative strength of the frequencies of periodic time series
  - Examples: sunspot intensity; El Niño/La Niña cycles vs. annual cycles
- Usually applied after **de-trending** – that is a separate feature of the time-series.

Spectral decomposition

Covariance sequence γ_t decomposed into **spectral densities** f:

$$\gamma_t = \frac{1}{2\pi} \int_{-1/2}^{+1/2} e^{2\pi i \omega_f t} f(2\pi\omega_f) d\omega_f$$

where ω_f is the frequency (cycles per unit of time).

Recall: $e^{ix} = \cos x + i \sin x$

This can be inverted to give the density at each frequency:

$$f(\omega) = \gamma_0 \left[1 + 2 \sum_{i}^{\infty} \rho_t cos(\omega t) \right]$$

where γ_0 is the overall covariance.

Spectral analysis as a linear regression problem

Time series of length n, gives $(n-1)$ predictors $a_0 \ldots a_{n/2}, b_1 \ldots b_{n/2-1}$ of the response variable:

$$x_t = a_0 + \sum_{k=1}^{n/2-1} \left[a_k \cos(2\pi kt/n) + b_k \sin(2\pi kt/n)\right] + a_{n/2} \cos(\pi t)$$

Note how at each value of k in the sum the frequency is increased.

a_0 is the intercept, it centres the series around its mean $\overline{x}$.

The highest frequency is $\omega = 0.5$ cycles per sampling interval, the lowest is $\omega = 2\pi$, i.e., one cycle.

(Note: the computation is usually with the Fast Fourier Transform).

The periodogram

This is used to estimate the spectral density (R `spectrum` function)

The periodogram at a given frequency ω: the squared correlation between the time series X and the sine/cosine waves at that frequency:

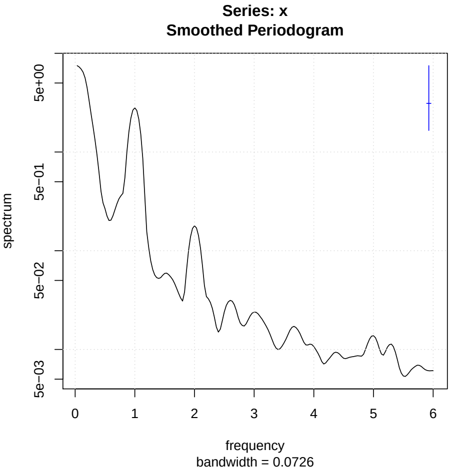
$$I(\omega) = \frac{1}{n} \left| \sum_t e^{-i\omega t} X_t \right|^2 = a_k^2 + b_k^2$$

- The spectrum is scaled by $1/\omega$, i.e. the inverse of the frequency.
- So the spectral density is computed over $-\omega/2 \ldots + \omega/2$
- Since the function is symmetric, it is displayed for $0 \ldots + \omega/2$.
- Groundwater example: $\omega = 12$; the decomposition is from $0 \ldots \omega/2 = 6$ periods per cycle.
- One period per cycle is annual, 2 per cycle is semi-annual, 6 periods is bi-monthly, etc.

Accounting for noise

- The time-series is usually noisy; we want to reveal the various periodicities w/o the noise
- So, smooth the series with **Daniell windows**, trial-and-error spans m
  - each spectral density estimate is computed as the mean of the $m/2$ preceding and following periodogram values

Example periodogram – log scale



Phase

- The series can also be expressed as a sum of cosines, each term with a **phase**: i.e., a lag from $\phi_k \in [-\pi/2 \ldots 3\pi/2]$
- The phase angle $\phi_k = \tan^{-1}(b_k/a_k)$
- Then $f(t) = \frac{1}{2}a_0 + \sum_{k=1}^{n} \sqrt{a_k^2 + b_k^2} \cos\left(\frac{2\pi n t}{\omega} - \phi_k\right)$
- The phase angle gives the offset from the centre of the period
- Reference: Jakubauskas, M. E., Legates, D. R., & Kastens, J. H. (2001). *Harmonic analysis of time-series AVHRR NDVI data*. **photogrammetric Engineering and Remote Sensing**, 67(4), 461–470.

Time-series modelling: PCA

- Principal Components Analysis (PCA): convert a number of inputs into the same number of outputs
  - Often applied to time-series of images
  - **orthogonal** – i.e., uncorrelated
  - in descending order of **variance explained** – importance of each component
  - contribution of each **original** image to each **synthetic** image
- Reveals correlation among dates, separates overall value (first component) from contrasts between dates
- If there are **periods** this should come out in one or more components
- Can also find **anomalies** – not explained by any overall value or periods

What does PCA do?

1. The **vector space** made up of the original variables is **projected** onto another space;
2. The new space has the **same dimensionality** as the original[1], i.e., there are as many variables in the new space as in the old;
3. In this space the new **synthetic variables**, also called **principal components** are **orthogonal** to each other, i.e. **completely uncorrelated**;
4. The synthetic variables are arranged in **decreasing order of variance explained**; and the total variance is unchanged;
5. The contribution of each original variable to each synthetic variables is given;
6. Each observation can be **re-projected** into the new (PC) space.

[1] unless the original was rank-deficient

Mathematics: Eigen decomposition

The key insight is that the **Eigen decomposition**[2] of the covariance or correlation matrix **C** orders the synthetic variables into descending amounts of variance, and ensures they are **orthogonal** (Hotelling 1933).

- Decompose a square, symmetric positive-definite matrix, e.g., the correlation matrix **C** formed from a data matrix such that $\mathbf{AC} = \lambda \mathbf{C}$
- **Eigenvalues**: a diagonal matrix λ; off-diagonals 0, i.e., no covariances, so orthogonal
- **Eigenvectors**: the transformation matrix **A**, a **coördinate transformation** such that the matrix multiplied by the diagonal eigenvalues matrix is the same as multiplication by a matrix made up of the eigenvectors
- Eigenvectors span an orthogonal **vector space** onto which we can **project** the original data.

[2](German *eigen* $\approx$ English "own, belonging to oneself")

Computation

- $|\mathbf{C} - \lambda \mathbf{I}| = 0$: a determinant to find the **eigenvalues** of the correlation matrix
  - these are sometimes called the **characteristic values**
  - their relative magnitude is the proportion of the original covariance explained
- Then the axes of the new space, the **eigenvectors** γ_j (one per dimension) are the solutions to $(\mathbf{C} - \lambda_j I)\gamma_j = \mathbf{0}$
- Obtain **synthetic variables** by projection: $\mathbf{Y} = \mathbf{PC}$ where $\mathbf{P}$ is the row-wise matrix of eigenvectors (rotations).

Example: Time series of diurnal temperature differences

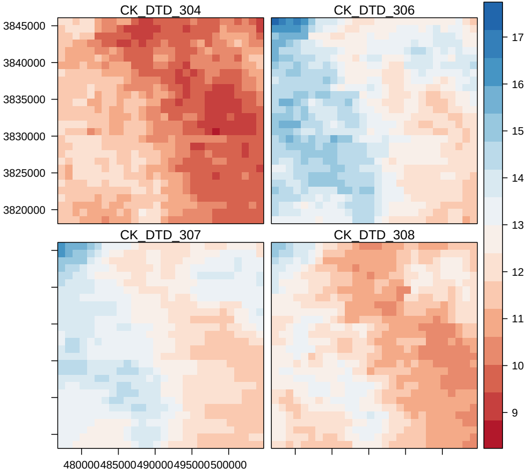
- 中国江苏沛县 Pei county in JiangSu province, PRC
- MODIS[3] daily land-surface temperature product[4]
- Images are **diurnal temperature differences** (day – night) between two MODIS products, units are Δ°C
- 30 Oct – 03 Nov 2000 (Julian days 304 ff.); Soil is drying after a heavy rain
- Objective: **try to relate DTD to soil texture and organic matter**
- Reference: Zhao, M.-S., Rossiter, D. G., Li, D.-C., Zhao, Y.-G., Liu, F., & Zhang, G.-L. (2014). *Mapping soil organic matter in low-relief areas based on land surface diurnal temperature difference and a vegetation index*. **Ecological Indicators**, 39, 120–133.[5]

[3] http://modis.gsfc.nasa.gov

[4] available at http://ladsweb.nascom.nasa.gov/data/search.html

[5] http://doi.org/10.1016/j.ecolind.2013.12.015

Original images



Δ°C
(day - night)

Low DTD on first day after rain; increases overall then decreases (closer day/night air T?)

But: similar overall pattern of high vs. low DTD (**redundancy**)

PCA results – Importance of components

```
> summary(pca)
Importance of components:
                          PC1     PC2    PC3     PC4
Standard deviation     1.8232 0.40406 0.3230 0.20810
Proportion of Variance 0.9145 0.04492 0.0287 0.01191
Cumulative Proportion  0.9145 0.95938 0.9881 1.00000
```

The four DTD images are highly-correlated, 92% of the information is in common, i.e., over the four days the same areas tend to have narrow and wide DTD ranges

PCA results – loadings

Contributions of each **original** image to each **synthetic** image

```
> pc$rotation
            PC1     PC2     PC3     PC4
DTD304 -0.4447  0.5893  0.5870  0.3322
DTD306 -0.6084  0.3379 -0.5200 -0.4952
DTD307 -0.4717 -0.3857 -0.3677  0.7025
DTD308 -0.4578 -0.6243  0.4998 -0.3884
```

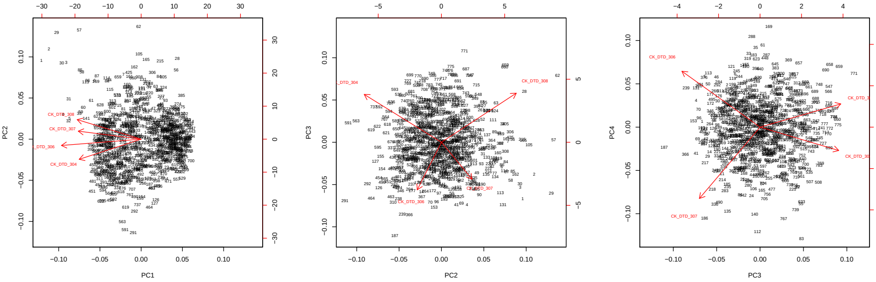
PC1 "intensity" of the phenomenon over all days

PC2 contrast between first two and second two days *after* accounting for the mutual correlation of PC1

PC3 contrast between middle two and end two days *after* . . .

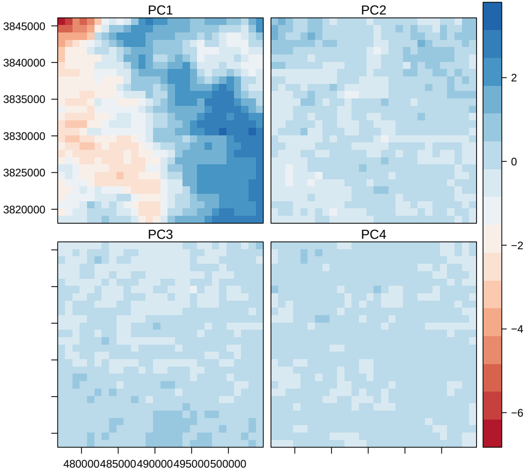
Biplots

These show positions of the observations as synthetic variables (bottom, left axes) and the correlations/variances of the original standardized images (top, right axes):



First PC is **overall intensity** across all 4 dates; other PCs show **contrasts** between dates

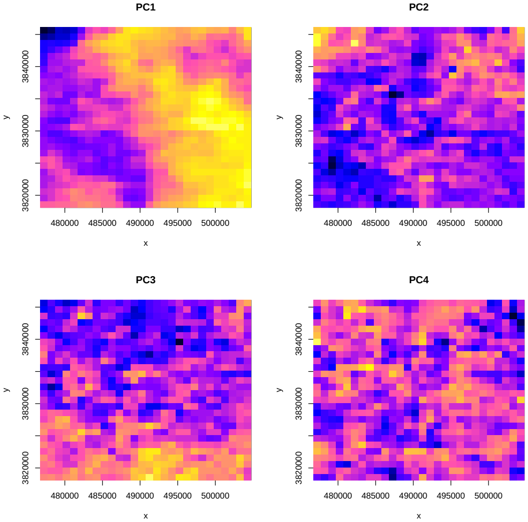
PCs (synthetic bands) – same stretch



Note decreasing information content with PC

But there still seems to be some pattern in PC2, 3, 4

PCs (synthetic bands) – individual stretch



This allows to visualize contrasts within a PC

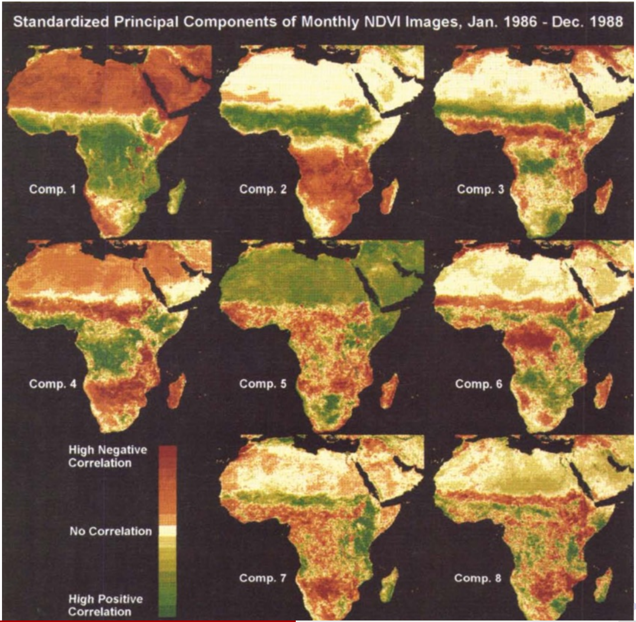
PCA of long time series of imagery

Groundbreaking paper, 200+ citations:

> *Eastman, J., & Fulk, M. (1993). Long Sequence Time-Series Evaluation Using Standardized Principal Components (reprinted from Photogrammetric Engineering and Remote-Sensing , Vol 59, Pg 991–996, 1993).* **Photogrammetric Engineering and Remote Sensing**, *59(8), 1307–1312.*

- Sensor was AVHRR
- Images were monthly maximum NDVI, 1986-1988 (three full years)

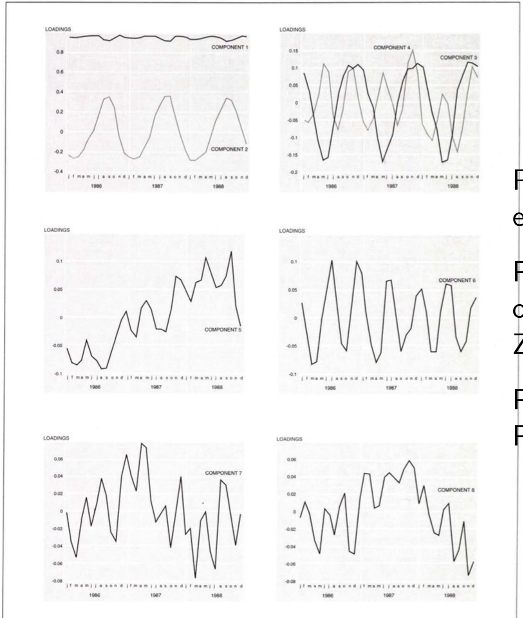
Synthetic images



Green = "+" score,
Red = "-"

Sign is arbitrary, here chosen to show high NDVI in green.

Loadings



PC1 All bands contribute $\approx$ equally:(overall intensity)

PC2 shows annual movement of Intertropical Convergence Zone

PC3/4 show deviations from PC2 seasonality

Interpretation

Synthetic bands images summarizing all original images, according to the PCs

Loadings relation between original images (dates, x-axis) and contribution to the PC (y-axis).

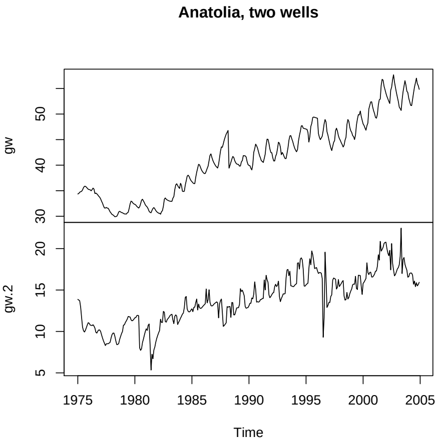
- Note decreased absolute loadings at higher PCs, this is because they represent less of the total variability of the 36 images

- PC1: $\approx$ equal contribution of all dates (overall vegetation vigour averaged over the 3 years)
- PC2: + correlation in N. hemisphere summer, - in winter (seasonality) – Intertropical Convergence Zone
- PC3/4: deviations from PC2 seasonality – note time lag of greening/senescence
- PC5: detecting sensor drift

Multiple time-series

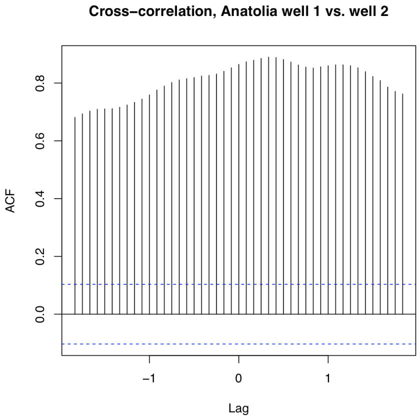
- Time-series of **several variables** measured at the **same times**
  - e.g., T, relative humidity, precipitation . . . at a single weather station
- These may be **cross-correlated** at various **lags**
- In spectral analysis, this is reflected as their **coherency** (correlation at a given period) and **phase difference** (coherent but offset)

Groundwater level: two wells



Clearly they are related – but at what lags? and how closely?

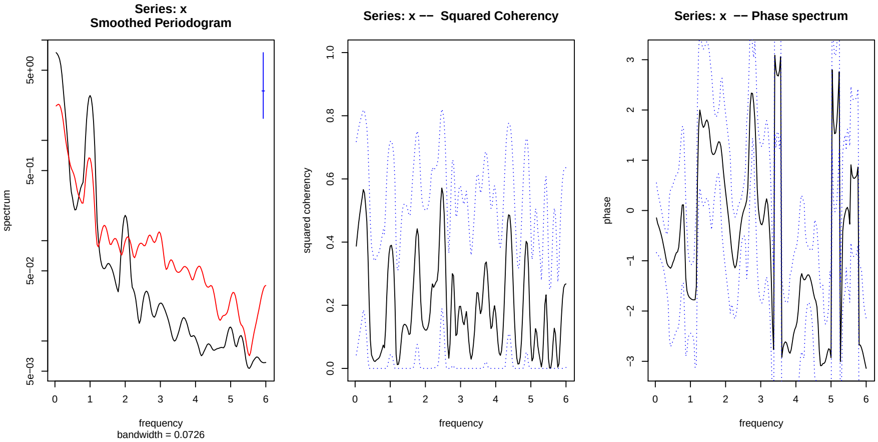
Groundwater level: cross-correlation



May be **asymmetric** – one station may be “ahead of” the other

Here highest cross-correlation is at +4 months, well 1 leads well 2

Groundwater level: cross-spectral analysis



Well 1 "leads" well 2 at some periods

Clear phase differences and loss of coherence.

Time-series analysis: computing

- CRAN Task View
  `cran.r-project.org/web/views/TimeSeries.html`
- base R has many functions, class `ts` for time-series objects
  - decomposition `stl`
  - autocorrelation `acf`, `pacf`
  - modelling `ar`, `arima`
  - spectral decomposition (Fourier analysis) `spectrum`
- specialized packages for advanced analysis
- R tutorial "Time series analysis in R" `http://www.css.cornell.edu/faculty/dgr2/pubs/list.html#pubs_m_R`

Outline

Topic: Spatial analysis

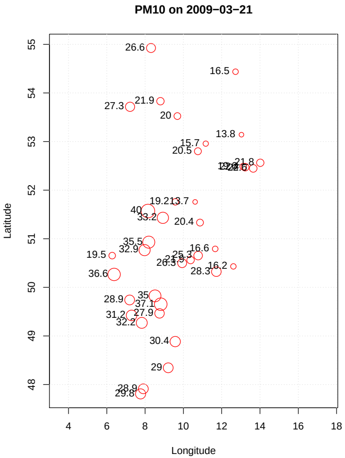
This has been covered in other lectures; included here just to emphasize that at one time slice there is usually some spatial structure.

1. trend surfaces
2. relation to other spatially-distributed attributes
3. local spatial dependence, also after accounting for (1) and (2)

Similarities between time series analysis and geostatistics

- TSA uses the correlation function, geostats the the variogram to examine auto-dependence
  - for **second-order stationary** series there is a direct conversion
- TSA uses autoregressive moving-average (ARMA) models to describe the time-evolution of a process
  - however there are also 2D spatial ARMA processes using the von Neumann neighbourhood (a diamond-shaped neighbourhood that defines a set of grid cells)

Example of spatial structure



Particulate matter (PM10, μg m^{-3}) at stations in Germany, one date; shows regional trend and residual local spatial dependence

PM10 air quality standards

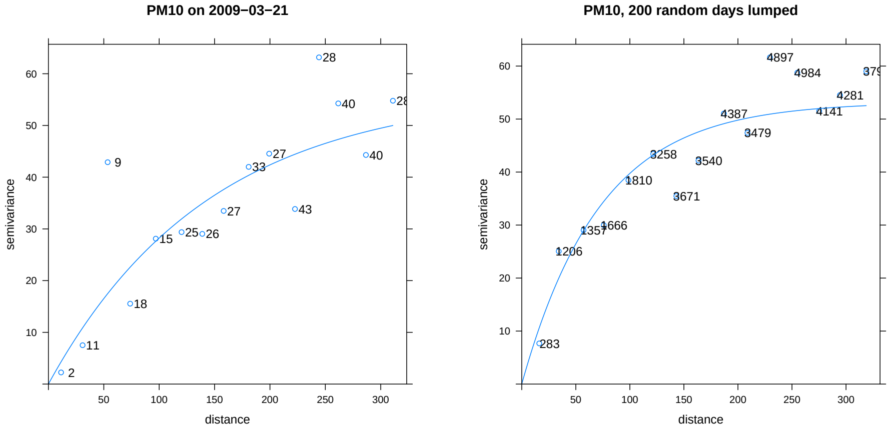
From the EU[6]:

- 50 μg m^{-3} over any single 24-hour period; can be exceeded 35 times per year
- 40 μg m^{-3} averaged over a year

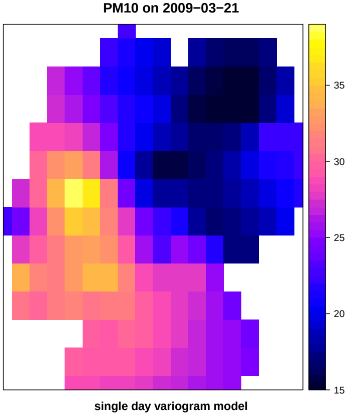
Previous slide shows all (rural) stations below the one-day threshold on 2009-03-31.

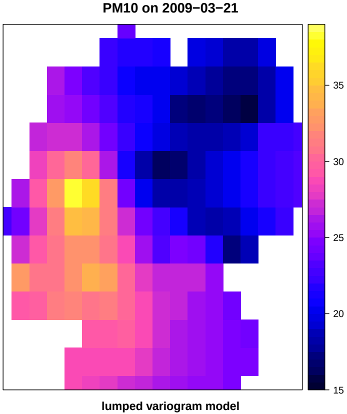
[6] http://ec.europa.eu/environment/air/quality/standards.htm

Spatial variograms



Ordinary Kriging predictions





one date variogram model

lumped, 200 days variogram model

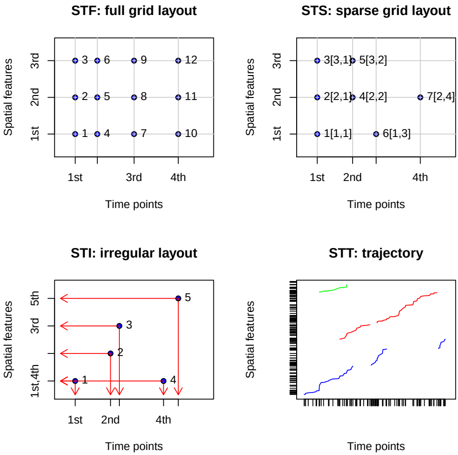
Ordinary Kriging onto the centres of 0.25°x 0.25°grid cells.

Outline

Topic: Spatio-temporal kriging

- Both space and time drive the observed process
- These may be independent or interacting
- Observations have **both** a location and time stamp
  - various ways to organize, see next slide
  - special data structures are required
- We can compute **spatio-temporal empirical variograms**;
  - These can be **modelled** with authorized variogram models
  - The fitted models can be used to interpolate in space and time by **kriging**
- Examples using R package `gstat`, many other applicable packages
  - CRAN Task View `https://cran.r-project.org/web/views/SpatioTemporal.html`

Organization of space-time data



Source: Pebesma [8]

Explanation

1. **grid**: all combinations have been observed
2. **sparse grid**: regular lattice but some missing
3. **irregular**: each observation has its own time stamp and location
4. **trajectory**: individual can move in both space and time
    - e.g., animals with tracking collars

Space-time dependence

The value in **s**pace and *t*ime of some target variable Z can be expressed as:

$$Z(\mathbf{s}, t) = Z^*(\mathbf{s}, t) + \varepsilon'(\mathbf{s}, t) + \varepsilon(\mathbf{s}, t) \tag{9}$$

s: spatial vector (commonly 2D); t: time

- Z^* is the **deterministic**, also called "structural" or "trend" component in both space and time
    - This can be modelled physically (e.g., process model) or empirically (e.g., regression on coördinates or covariates)
- ε' is a **spatio-temporally correlated** space-time process of the **residuals**
- ε is pure noise; it assumed to be purely random ("white noise")

Local space-time dependence

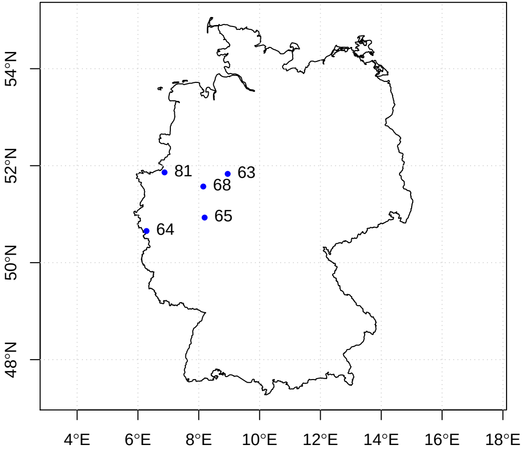
Ignoring any trend in space and time, Z^* becomes a constant mean in both space and time:

$$Z(\mathbf{s}, t) = \mu + \varepsilon'(\mathbf{s}, t) + \varepsilon(\mathbf{s}, t) \tag{10}$$

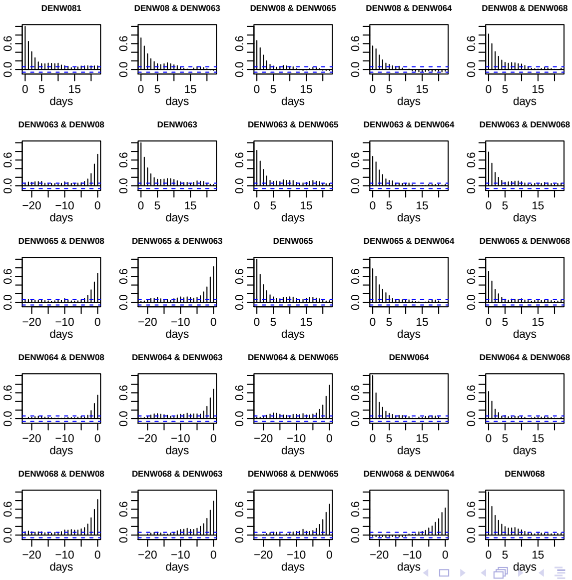
There is no way to objectively decide if there is a structural component, or if present, how to model it.

In practice local approaches can be used even if there is a structure; we lose information but the predictions may be satisfactory.

Location of PM10 monitoring stations in NRW (D)



Direct & cross-correlations, PM10, NRW



Explanation

1. Each station has temporal dependence: one day's PM10 is "similar to" the next, to some lag $\rightarrow$ a somewhat persistent atmospheric phenomenon
2. Stations can be **cross-correlated** to others at time lag 0 (i.e., **spatial correlation**) – we can see if this depends on distance or a trend
3. Stations can also be cross-correlated at **different time lags**, and this may be **asymmetric**
    - e.g., station A may lag behind station B, for example process could be wind blowing PM10 from B to A

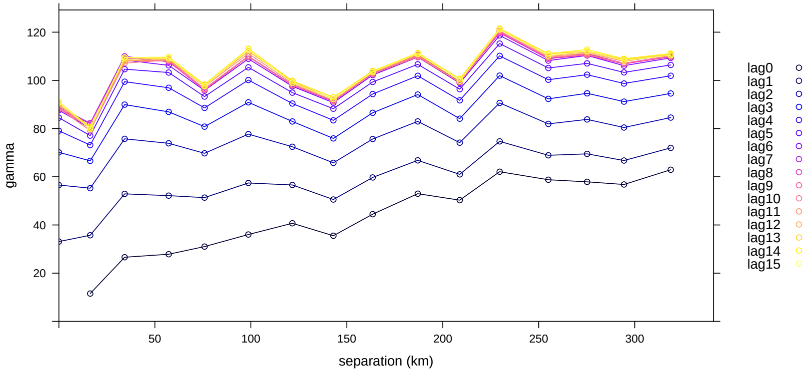
Spatio-temporal variograms

- Compute semi-variance at **combined** lags
  - **h** spatial lag, usually 2D, can include a direction (anisotropy)
  - u temporal lag, 1D
- At all combinations of space and time:

$$\gamma(\mathbf{h}, u) = \frac{1}{2} \sum (z_{\mathbf{s},t} - z_{\mathbf{s}+\mathbf{h},t+u})^2 \qquad (11)$$

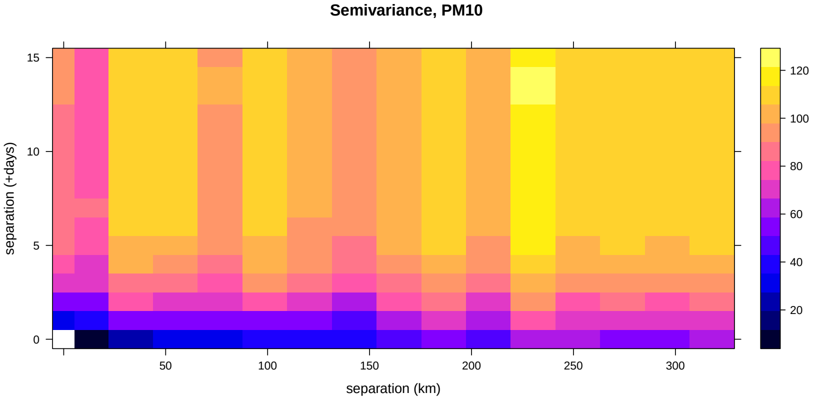
- Summarized in **lags** (bins) to have enough point-pairs
- Isotropic in 2D space → 2D variogram (isotropic spatial + temporal separations)
- Anisotropic in 2D space → 3D variogram! (anisotropic spatial + temporal separation)

Spatio-temporal variograms – view 1



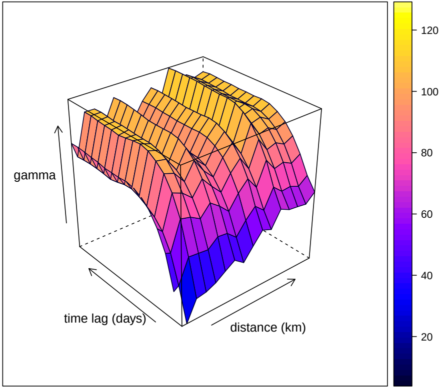
PM10 concentrations, one **spatial** variogram per **temporal** lag

Spatio-temporal variograms – view 2



PM10 concentrations, distance/time lag matrix

Spatio-temporal variograms – view 3



PM10 concentrations, distance/time lag matrix, wireframe

The two **marginal** (space and time) variograms have **different sills**

Explanation

- **Zonal** anisotropy: different sills in the two dimensions
- Variograms are erratic because of few stations
- Time and space lag 0 **marginal** variograms show typical empirical variogram (exponential model)?
- Marginal total sill for space is about half that for time
  - i.e., more time than space variability in PM10 over Germany
- As temporal lags increase:
  - spatial dependence becomes less (shorter range; higher nugget/sill ratio)
  - total variability (total sill) increases
- Interesting anomaly at longer time lags, 25 km separation – wind?

Spatio-temporal combined model – viewpoints

1. time as an extra dimension **equivalent** to space
   - this is called the "metric" model
   - time asymmetry can not be modelled
   - time must be **re-scaled** with the **space-time anisotropy ratio** to match space
   - spatial and temporal dependence structure must be the same
   - **rarely realistic**
2. **separate** covariance structures for space and time
   - this is called the "separable covariance" model
   - no interaction between space and time
3. **interacting** covariance structures for space and time
   - this is called the "sum-metric" model
   - a **metric** model plus a **purely spatial** and also a **purely temporal** component added to it

Modelling space-time variograms

Models forms, with their assumptions:

metric
: time is equivalent to an additional "spatial" dimension, re-scaled to match the spatial units
  - single model form, nugget and partial sill
  - unrealistic if sills in space vs. time not the same

separable
: time and space are modelled separately, each in their own units; the variogram is a proportional product
  - structure of temporal variation must be the same at all locations
  - structure of spatial variation must not vary over time
  - not realistic in most contexts

sum-metric
: also has a space-time term, allowing for **interaction**

Metric covariance structure

$$C(h, u) = C\left(\sqrt{h^2 + (\alpha u)^2}\right) \tag{12}$$

- One covariance structure C
- h is the distance lag, generally 2D
- u is the time lag, here another "dimension"
- α is a scaling factor ("metric") to match spatial and temporal units: the **space-time anisotropy ratio**

This represents **geometric anisotropy**, i.e., same structure, nugget and partial sill but the range varies in different dimensions (here, the time vs. space dimensions)

e.g., $\alpha = 20\ km\ d^{-1}$ means that a lag of 20 km in space is equivalent to a lag of 1 day in time.

Separable covariance structure

$$C(h, u) = C_s(h) \cdot C_t(u) \tag{13}$$

- Two covariance structures, one each for space and time
- No interaction; the covariance at a given spatial and temporal lag is just the product of the two marginal covariances
- There is only one joint sill; the two sills of the two components are given as **proportions** of this.

Sum-metric covariance structure

$$C(h,u) = C_s(h) + C_t(u) + C_{st}\left(\sqrt{h^2 + (\alpha u)^2}\right) \tag{14}$$

- Three covariance structures
- Space, time, and joint each have their own sill, range, and nugget
- C_{st} is the metric structure of the residuals, after accounting for space and time separately
- The last term has **geometric anisotropy**

Sum-metric parameters

spatial marginal variogram 3: partial sill, nugget, range

temporal marginal variogram 3: partial sill, nugget, range

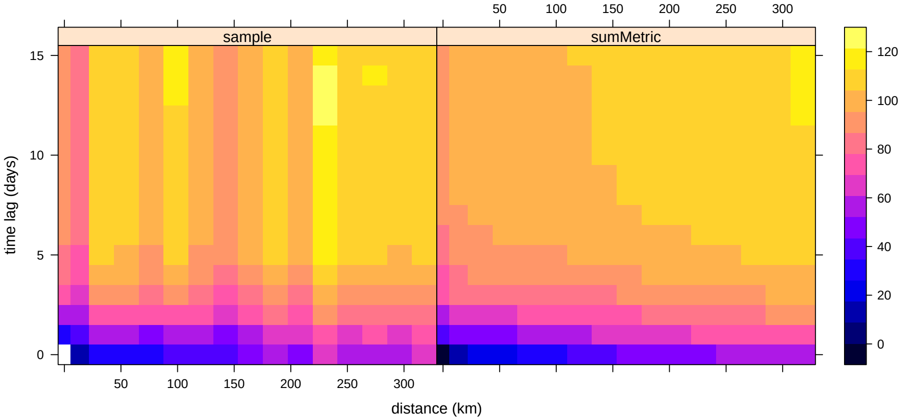
space-time variogram 4: partial sill, nugget, range, **anisotropy**

- The space and time sills have to match, but only for the **residuals** after accounting for the marginal variograms (space and time)

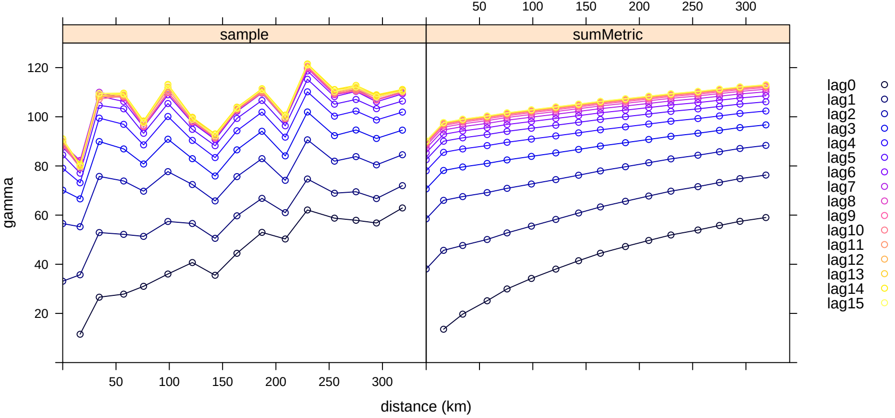
To estimate the three different nuggets it is required to have **repeated** measurements at the same **locations**, as well as **repeated** measurements at the same **times**.

Parameters are fit with `fit.StVariogram`, which uses the generic `optim` optimization function.

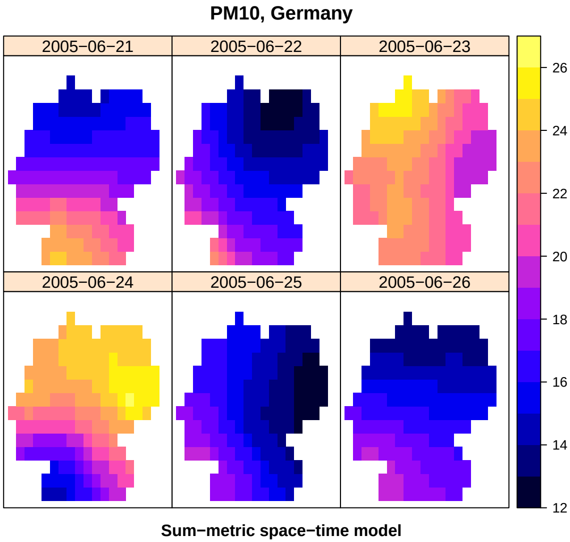
Fitted sum-metric variogram – view 1



Fitted sum-metric variogram – view 2



Results of space-time kriging



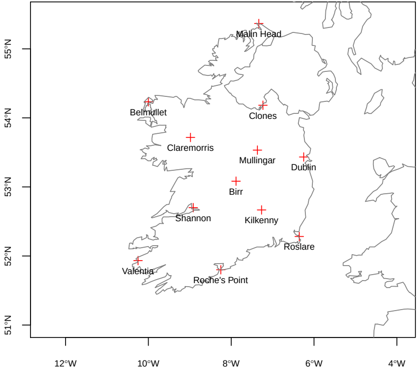
Outline

Topic: Modelling spatio-temporal structure with EOF

- EOF: **Empirical Orthogonal Functions**
- mathematically the same as PCA (Principal Component Analysis)
- finds the spatial or temporal patterns of variability of **one variable**
  - i.e., spatial patterns over time, or temporal patterns over space
- measures the importance of each contribution
- clusters variables (either time instances or spatial locations) in PC space

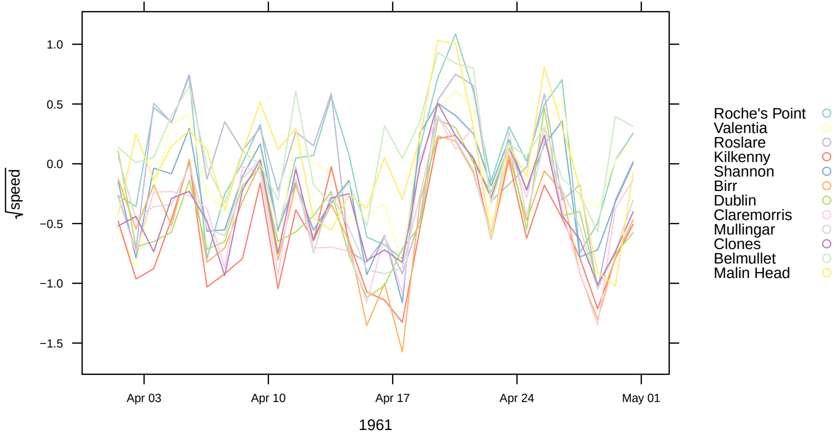
Example: Republic of Ireland wind speed

Source: Pebesma [8]



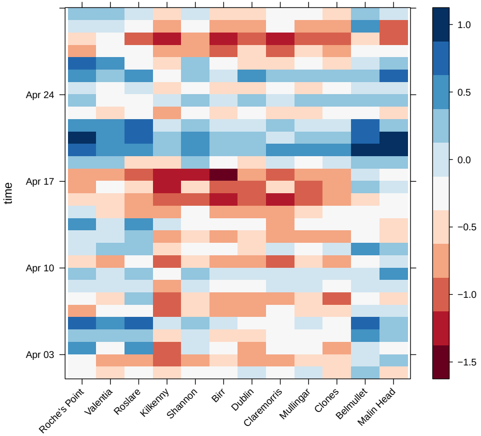
Station locations

Ireland: time-series at different stations



does not show spatial distances

Ireland: another space-time plot

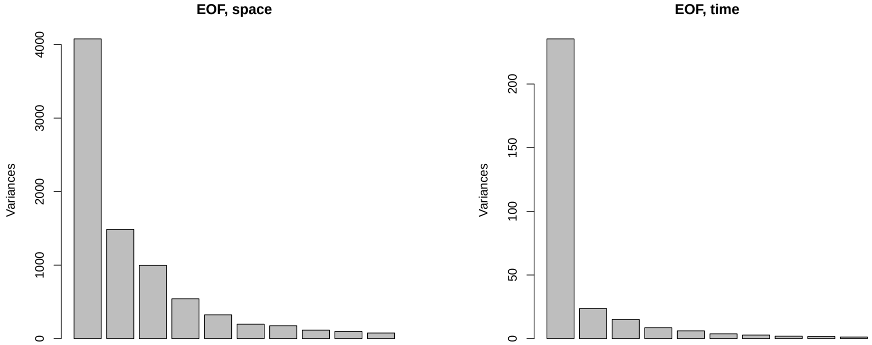


does not show spatial distances

Explanation

1. Obvious correlation between stations at each time period
   - Low-speed days, e.g., 17-April
   - High-speed days, e.g., 20-April
2. But sometimes the correlation is **lagged**: the highest speed at a station comes a day or two after that at another
   - e.g., 21, 20, 19 April
   - Is this from a synoptic event, e.g., Atlantic storm?
3. Some temporal dependence, although erratic

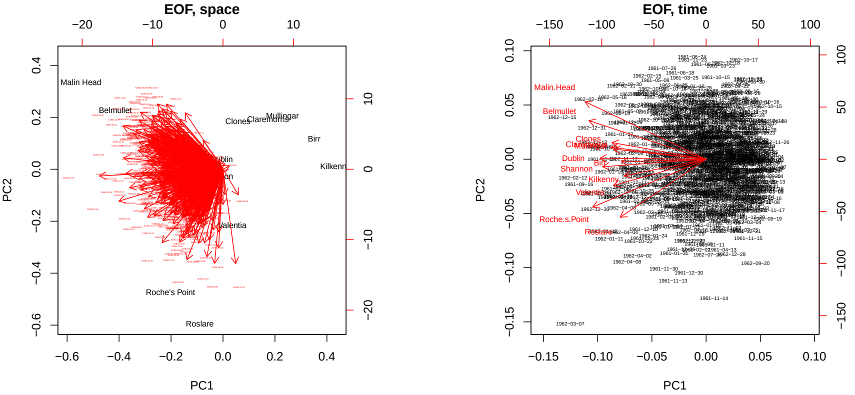
EOF: scree plots



These show the relative importance of each PC

The variance is higher in the spatial EOF because many more time replications (two years) vs. space replications (12)

EOF: biplots



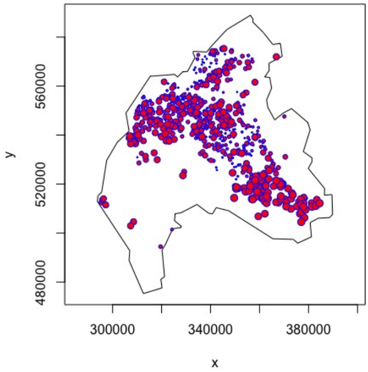
These show the loadings of each variable on the first two PCs (red) and the position of each observation in PC space (black).

Outline

Spatio-temporal point patterns

- Aim: analyze point-patterns which may change over time
  - e.g., locations of live trees in a forest plot (some die, some new ones grow)
  - e.g., locations of crime or disease incidences, each with a **time stamp**
- Q1: Does the structure of the point-pattern change over time?
  - intensity, kernel density, G, F, K, L functions ...
- Q2: Does the point-pattern at one time affect the pattern at a later time?
  - crossed K function

Example: time of occurrence

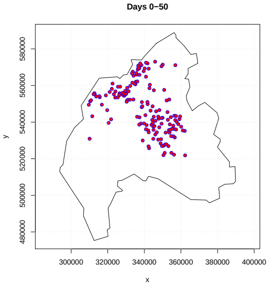


Foot-and-mouth disease, northern Cumbia (England), 2001; from R package `stpp`, dataset `fmd`

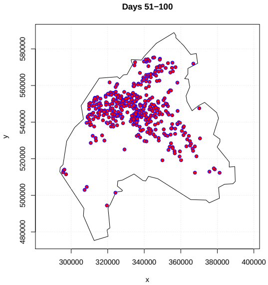
more recent → larger symbol; total 648 cases (points)

Analysis

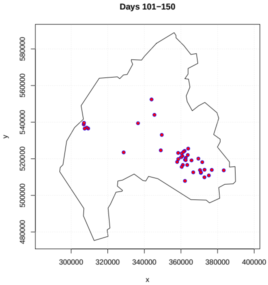
1. Divide into **time slices**, each with enough cases
   - here, 50-day slices (156, 404, 40, 48 cases)
   - could also divide by "epidemic stage" from expert opinion
2. Display point-pattern in the slice, compute **intensity**
3. Compute point-pattern functions G (**closest neighbour**), F (**empty space**), look at pattern over time
4. Compute **crossed K-function** for adjacent time slices, to see if they are independent, attracting, or repulsing
5. From all of this, try to infer **mechanisms**



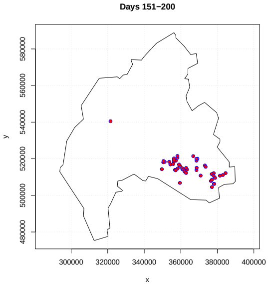
36 km^2 case^{-1}



14 km^2 case^{-1}

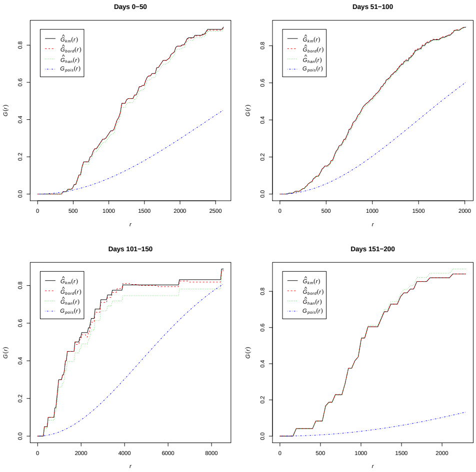


139 km^2 case^{-1}



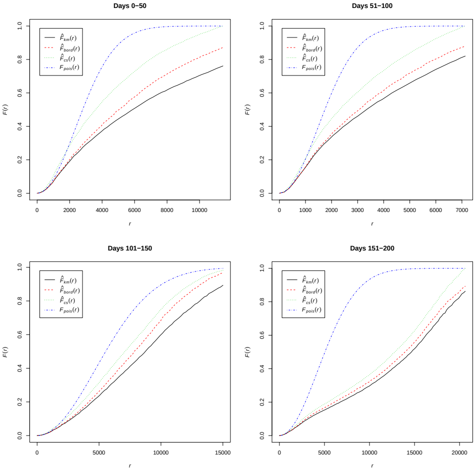
116 km^2 case^{-1}

G function at each time slice



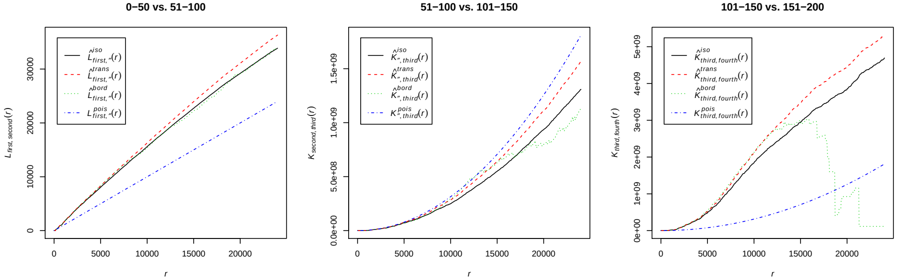
Strong clustering at each time slice, but amount changes with time

F function at each time slice



Much more empty space than expected, amount changes with time

Crossed K function between time slices



0-50 to 51-100: clustered

51-100 to 101-150 ≈ independent

101-150 to 151-200 clustered

Interpretation

- Point-pattern moves over time
- Intensity changes over time
- Nearest-neighbour G function changes over time
- empty-space function F changes over time
  - i.e., distance from an arbitrary position to an occurrence of foot-and-mouth disease
- patterns are *not* independent between the 0-50 and 51-100, and the 101-150 and 151-200 slices
  - in both cases strongly dispersed
  - no interaction between the 51-100 and 101-150 slices

Further analysis of spatio-temporal point patterns

" Log-Gaussian Cox Processes"

- window $W \subset \mathbb{R}^2$, time slice $T \subset \mathbb{R}_{\geq 0}$
- Cases occur at spatio-temporal positions $(x, t) \in W \times T$ according to an **inhomogeneous spatio-temporal Cox process**, i.e., a Poisson process with intensity $R(x, t)$
  - i.e., number of cases $X_{S,[t_1,t_2]}$ is Poisson-distributed conditional on R.
- $R(s, t) = \lambda(s)\mu(t) \exp\{\mathcal{Y}(s, t)\}$
  - i.e., fixed spatial, fixed temporal, and interaction term
- see Taylor et al. [10] (`lgcp` R package)

Outline

Concluding remarks

- All processes occur in both space and time.
- Unless we restrict to one moment in time or one location in space, or if the variability in one of them is negligible compared to the other, both space and time need to be considered:
  - when modelling the process, and then ...
  - ...interpreting the model to look for causes
- Space and time may be **separable** elements of the analysis, but very commonly the variations in space and time are **not independent**.

Outline

Further reading

- Time series: Shumway and Stoffer [9], Wilks [11, Ch. 9], Box [1]
  - these also as e-books via CU library
- Description of R package `spacetime`: Pebesma [8]
- Spatio-temporal point patterns: Diggle [2], Taylor et al. [10]
- Theory: Kyriakidis and Journel [7], Gneiting et al. [3]
- Applications: Heuvelink and Griffith [5]; Jost et al. [6]; Hengl et al. [4]

Web pages

- CRAN Task View: Handling and Analyzing Spatio-Temporal Data: http://cran.r-project.org/web/views/SpatioTemporal.html
- Benedikt Gräler (was in Münster, now at Ruhr University Bochum): http://ben.graeler.org

References I

[1] George E. P. Box. *Time series analysis: forecasting and control*. John Wiley & Sons, Inc., fifth edition / edition, 2016. ISBN 978-1-118-67492-5. URL http://ebookcentral.proquest.com/lib/cornell/detail.action?docID=2064681.

[2] P. J. Diggle. *Statistical analysis of spatial and spatio-temporal point patterns*. CRC Press, Boca Raton, 3rd edition, 2013. ISBN 978-1-4665-6023-9.

[3] Tilmann Gneiting, Marc G. Genton, and Peter Guttorp. *Geostatistical Space-Time Models, Stationarity, Separability and Full Symmetry*. Number TR475 in Department of Statistics, University of Washington. 2005. URL http://www.stat.washington.edu/research/reports/2005/tr475.pdf.

References II

[4] Tomislav Hengl, Gerard B. M. Heuvelink, Melita Percec Tadic, and Edzer J. Pebesma. Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images. *Theoretical and Applied Climatology*, 107(1-2):265–277, Jan 2012. doi: 10.1007/s00704-011-0464-2.

[5] Gerard B. M. Heuvelink and Daniel A. Griffith. Space-time geostatistics for geography: A case study of radiation monitoring across parts of Germany. *Geographical Analysis*, 42(2):161–179, 2010. doi: 10.1111/j.1538-4632.2010.00788.x.

[6] G. Jost, G. B. M. Heuvelink, and A. Papritz. Analysing the space-time distribution of soil water storage of a forest ecosystem using spatio-temporal kriging. *Geoderma*, 128(3-4):258–273, 2005.

[7] Phaedon C. Kyriakidis and André G. Journel. Geostatistical space–time models: A review. *Mathematical Geology*, 31(6):651–684, 1999. doi: 10.1023/A:1007528426688.

References III

[8] Edzer Pebesma. spacetime: Spatio-temporal data in R. *Journal of Statistical Software*, 51(7), 2012. URL `http://www.jstatsoft.org/v51/i07`.

[9] Robert H. Shumway and David S. Stoffer. *Time series analysis and its applications with R examples*. Springer texts in statistics. Springer, New York, 3rd ed edition, 2011. ISBN 9781441978653.

[10] Benjamin M. Taylor, Tilman M. Davies, Barry S. Rowlingson, and Peter J. Diggle. lgcp: an R package for inference with spatial and spatio-temporal log-Gaussian Cox processes. *Journal of Statistical Software*, 52(4), 2013. ISSN 1548-7660. doi: 10.18637/jss.v052.i04. URL `http://www.jstatsoft.org/v52/i04/`.

[11] Daniel S. Wilks. *Statistical methods in the atmospheric sciences*. International geophysics series. Academic Press, 3rd edition, 2011. ISBN 9780123850225.