Module INM433 – Visual Analytics

Lecture 03

Visual analytics of spatio-temporal data

given by

prof. Gennady Andrienko and

prof. Natalia Andrienko

Content and objectives

- We shall consider the types of data that have spatial and temporal components: spatial events, spatial time series, origin-destination moves, and trajectories of moving objects.
- Analysis of spatio-temporal data often requires transformation from one data type to another. We shall consider transformations from spatial events and movement data to spatial time series.
- We shall introduce two possible perspectives for looking at spatial time series and show two complementary ways of applying partition-based clustering (PBC) to them. You will see the principal differences between the kinds of the results and learn how to interpret the results depending on the way of using PBC.
- Another transformation to be considered is extraction of spatial events from spatial time series.

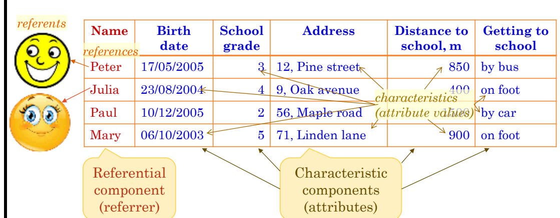
Types of spatio-temporal data

Data with spatial and temporal components

Semantic roles of data components

A reminder

- **Reference**: What is described?
- **Characteristic**: What is known about it?

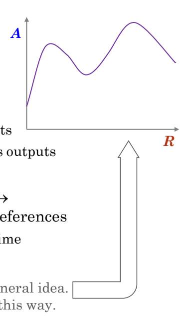


Name	Birth date	School grade	Address	Distance to school, m	Getting to school
Peter	17/05/2005	3	12, Pine street	850	by bus
Julia	23/08/2004	4	9, Oak avenue	400	on foot
Paul	10/12/2005	2	56, Maple road		by car
Mary	06/10/2003	5	71, Linden lane	900	on foot

Data components viewed as variables

A reminder

- Referrers ~ independent variables
- Attributes ~ dependent variables
- Data represent (are generated by) a *function* Referrers → Attributes
  - References (values of referrers): function's inputs
  - Characteristics (values of attributes): function's outputs
- Function's ***behaviour***: how the outputs vary over the set of inputs → how the characteristics vary over the set of references
  - E.g., how the crime rates vary over space and time

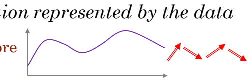


Note:
This is a schematic representation of the general idea. Not all kinds of data can be represented in this way.

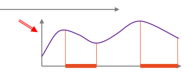
General aim of data analysis:

study the behaviour of the function represented by the data

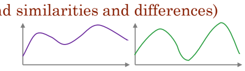
- Describe the behaviour of one or more attributes



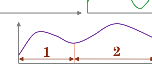
- Locate particular behaviour: find subsets of references where attributes have it



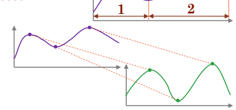
- Compare two or more behaviours (find similarities and differences)
  - Different attributes over the same set of references



  - Same attributes over different subsets of references



- Relate behaviours of two or more attributes



Data typology: a reminder

- Data typology is based on distinguishing references and attributes.
- Major types of references: **time**, **space**, discrete **objects**
- Data may have two references or even more.
  - Frequent combinations: time + space, time + objects
- Major data types are defined according to the types of references.
- Discrete objects are classified based on their relation to space and time.

Major data types

according to the types of references

- *Object-referenced data*:
  - attribute values refer to discrete objects
- *Time-referenced data*, a.k.a. *time series*:
  - attribute values refer to different times (moments or intervals)
- *Space-referenced data*, a.k.a. *spatial data*:
  - attribute values refer to different spatial locations (points, lines, areas, volumes in 3D space)
- *Object-referenced time series*:
  - attribute values refer to discrete objects and to different times
- *Spatial time series*:
  - attribute values refer to different spatial locations and to different times

Classes of discrete objects

according to their relation to space and time

- Generic objects: no relation to space or time (or the existing relation is not taken into account)
- Temporal objects (events): exist at some moment or interval in time
  - Time (existence time) is an *attribute* of the objects
- Spatial objects: have location in space
  - Space (spatial location) is an *attribute* of the objects
  - Moving objects are spatial objects that may have different spatial locations at different times
- Spatio-temporal objects (spatial events): exist in time and have location in space
  - Time (existence time) and space (spatial location) are *attributes* of the objects

Spatio-temporal data

- Have both spatial and temporal components
- The spatial and temporal components may play different semantic roles: references or attributes

Types of spatio-temporal data

According to the semantic roles of space and time

- *Spatial time series*:
  - attribute values refer to different spatial locations and different times; space and time are *references*
- *Spatial event data*: data characterising spatial events (a specific kind of discrete objects)
  - attribute values, including spatial location and existence time, refer to objects; space and time are *attributes*
- *Movement data*: data characterising moving objects (a specific kind of discrete spatial objects)
  - Space (spatial location) is an *attribute* of moving objects
  - The spatial location may change over time ⇒ the data include two or more spatial locations referring to different times ⇒ time is a *reference*
  - Hence, movement data are a kind of object-referenced time series: attribute values, including spatial location, refer to objects and times

Spatial time series (an example)

Referrer 1: time

Referrer 2: place

Attributes

year	id	State	Population	Index offenses	Violent crime	Murder	Forcible rape	Robbery	Aggravated assault	Property crime	Burglary	Larceny-theft	Motor vehicle theft
1960	1	Alabama	3266740	39920	6097	406	281	898	4512	33823	11626	19344	2853
1960	2	Alaska	226167	3730	236	23	47	64	102	3494	751	2195	548
1960	4	Arizona	1302161	39243	2704	78	209	706	1711	36539	8926	23207	4406
1960	5	Arkansas	1786272	18472	1924	152	159	443	1170	16548	5399	10250	899
1960	6	California	15717204	546069	37558	616	2859	15287	18796	508511	143102	311956	53453
1960	8	Colorado	1753947	38103	2408	73	229	1362	744	35695	9996	21949	3750
1960	9	Connecticut	2535234	29321	928	41	103	236	548	28393	8452	16653	3288
1960	10	Delaware	446292	9642	375	33	41	157	144	9267	2661	5867	739
1960	11	District of Co	763956	20725	4230	81	111	1072	2966	16495	4587	9905	2003
1960	12	Florida	4951560	133919	11061	527	403	4005	6126	122858	39966	73603	9289
…													
1972	54	West Virgini	1781000	25584	2299	109	146	562	1482	23285	7356	13976	1953
1972	55	Wisconsin	4520000	133382	4358	126	376	1661	2195	129024	28862	89642	10520
1972	56	Wyoming	345000	10461	511	14	48	117	332	9950	2057	7190	703
1973	1	Alabama	3539000	91389	12390	468	751	2809	8362	78999	31754	39206	8039
1973	2	Alaska	330000	16313	1269	33	147	221	868	15044	3852	9456	1736
1973	4	Arizona	2058000	137966	9877	167	637	3031	6042	128089	40301	76560	11228
1973	5	Arkansas	2037000	56149	5905	180	398	1456	3871	50244	18088	29204	2952
1973	6	California	20601000	1298872	116563	1862	8357	49531	56813	1182309	407824	643488	130997
1973	8	Colorado	2437000	133933	10088	193	944	3970	4981	123845	38963	70931	13951
1973	9	Connecticut	3076000	112717	6421	102	342	2589	3388	106296	31661	58742	15893
…													
2000	44	Rhode Island	1048319	36444	3121	45	412	922	1742	33323	6620	22038	4665
2000	45	South Caroli	4012012	209482	32293	233	1511	5883	24666	177189	38888	123094	15207
2000	46	South Dakot	754844	17511	1259	7	305	131	816	16252	2896	12558	798
2000	47	Tennessee	5689283	278218	40233	410	2186	9465	28172	237985	56344	154111	27530
2000	48	Texas	20851820	1033311	113653	1238	7856	30257	74302	919658	188975	637522	93161
2000	49	Utah	2233169	99958	5711	43	863	1242	3563	94247	14348	73438	6461
2000	50	Vermont	608827	18185	691	9	140	117	425	17494	3501	13184	809
2000	51	Virginia	7078515	214348	19943	401	1616	6295	11631	194405	30434	146158	17813
2000	53	Washington	5894121	300932	21788	196	2737	5812	13043	279144	53476	190650	35018
2000	54	West Virgini	1808344	47067	5723	46	331	749	4597	41344	9890	28139	3315
2000	55	Wisconsin	5363675	172124	12700	169	1165	4537	6829	159424	25183	119605	14636
2000	56	Wyoming	493782	16285	1316	12	160	70	1074	14969	2078	12318	573

References are not always in columns

Referrer 1: place

Referrer 2: time →

identifiers	financial year=2001(__ 1 Total offences (Offences rates)	financial year=2002(__ 1 Total offences (Offences rates)	financial year=2003(__ 1 Total offences (Offences rates)	financial year=2004(__ 1 Total offences (Offences rates)	financial year=2005(__ 1 Total offences (Offences rates)	financial year=2006(__ 1 Total offences (Offences rates)	financial year=2007(__ 1 Total offences (Offences rates)	financial year=2008(__ 1 Total offences (Offences rates)	financial year=2009(__ 1 Total offences (Offences rates)	financial year=2010(__ 1 Total offences (Offences rates)	financial year=2011(__ 1 Total offences (Offences rates)
E05000026 Abbey	278	303	279	300	248	255	209	215	221	206	183
E05000027 Alibon	79	84	86	81	90	97	89	95	97	96	91
E05000028 Becontree	86	89	105	108	112	110	101	102	109	98	99
E05000029 Chadwell Heath	118	133	157	157	153	138	127	129	112	119	97
E05000030 Eastbrook	75	86	77	76	91	92	87	85	84	67	74
E05000031 Eastbury	100	96	109	119	105	134	121	105	97	106	81
E05000032 Gascoigne	277	239	220	190	170	162	133	123	107	129	122
E05000033 Goresbrook	78	75	99	103	99	86	89	89	93	94	84
E05000034 Heath	104	107	104	117	115	128	119	107	99	97	103
E05000035 Longbridge	77	67	82	89	71	89	77	73	75	81	79
E05000036 Mayesbrook	90	74	72	92	95	102	96	100	100	94	80
E05000037 Parsloes	77	79	70	92	85	97	92	89	84	75	76
E05000038 River	113	103	114	102	121	115	109	99	96	86	93
E05000039 Thames	263	226	254	227	215	209	180	178	176	163	130
E05000040 Valence	81	75	83	87	86	97	89	76	80	74	82
E05000041 Village	111	124	143	120	115	134	130	129	130	89	100
E05000042 Whalebone	109	93	96	97	114	119	99	102	107	93	86
E05000043 Brunswick Park	70	68	84	78	66	61	55	54	56	52	56
E05000044 Burnt Oak	88	108	102	114	95	95	79	70	76	67	64
E05000045 Childs Hill	119	125	129	142	136	127	118	106	106	105	100
E05000046 Colindale	93	114	114	107	95	93	76	76	70	77	71
E05000047 Coppetts	99	104	115	113	100	89	76	75	81	82	77
E05000048 East Barnet	61	62	80	88	82	61	66	65	61	57	59
E05000049 East Finchley	85	87	95	92	81	84	59	67	61	57	65
E05000050 Edgware	107	108	107	121	114	106	89	84	86	97	91
E05000051 Finchley Church	63	59	58	64	62	56	56	51	58	59	51
E05000052 Garden Suburb	87	83	88	105	70	72	74	65	66	71	79
E05000053 Golders Green	94	98	101	112	96	96	79	70	68	70	76
E05000054 Hale	58	63	65	70	72	70	54	54	65	53	54
E05000055 Hendon	88	83	101	94	108	86	84	93	82	73	72
E05000056 High Barnet	78	92	105	109	101	75	80	80	72	68	81
E05000057 Mill Hill	79	90	103	93	85	90	84	74	71	73	77
E05000058 Oakleigh	57	63	72	79	72	61	58	62	51	49	59

Spatial event data (an example)

Referrer: objects

identifier	date time	latitude	longitude	Height	Type	Amplitude
7635	20110818 14:00:05	50.614	3.5492	0	1	-5.5
7636	20110818 14:00:05	50.6091	3.5469	0	1	-48.4
7637	20110818 14:00:05	50.6083	3.5433	0	1	-20.4
7638	20110818 14:00:05	50.6086	3.5407	0	1	-5.7
7639	20110818 14:00:06	50.6111	3.5504	0	1	-12.8
7640	20110818 14:00:06	50.6105	3.5482	0	1	-22.8
7641	20110818 14:00:06	50.6115	3.5504	0	1	-6.1
7642	20110818 14:00:06	50.6114	3.5504	0	1	-7.5
7643	20110818 14:00:08	51.5461	3.4626	0	1	17.7
7644	20110818 14:00:08	50.5266	3.3881	6.7	2	-4.4
7645	20110818 14:00:10	50.7414	3.7894	6.9	2	-10.5
7646	20110818 14:00:10	50.7388	3.7909	7.6	2	-17.5
7647	20110818 14:00:10	50.763	3.7868	10.5	2	-5.6
7650	20110818 14:00:15	50.1491	3.3555	0	1	-23
7651	20110818 14:00:18	50.7346	3.789	7.1	2	-6.3
7652	20110818 14:00:18	50.7349	3.7881	0	1	-8.7
7653	20110818 14:00:19	50.84	4.1534	5.3	2	6.8

time (attribute)

spatial location (attribute)

Movement data (an example)

Referrer 1: objects

Referrer 2: time

trID	X	Y	time	S
1	-0.13	-9.41	20/11/2015; 20:47:25.000	0
1	0.04	-9.12	20/11/2015; 20:47:25.200	0.55
1	0.33	-8.7	20/11/2015; 20:47:25.400	2.07
1	0.74	-8.17	20/11/2015; 20:47:25.600	10.31
1	1.27	-7.55	20/11/2015; 20:47:25.800	14.82
1	1.92	-6.88	20/11/2015; 20:47:26.000	16.49
1	2.68	-6.2	20/11/2015; 20:47:26.200	17.66
1	3.52	-5.58	20/11/2015; 20:47:26.400	18.1
1	4.41	-5.02	20/11/2015; 20:47:26.600	18.69
1	5.34	-4.57	20/11/2015; 20:47:26.800	19.07
…				
1	34.3	13.91	20/11/2015; 21:35:34.400	4.99
1	34.14	13.65	20/11/2015; 21:35:34.600	5.64
1	33.99	13.37	20/11/2015; 21:35:34.800	5.15
2	-5.03	-15.03	20/11/2015; 20:47:25.000	0
2	-4.93	-15.07	20/11/2015; 20:47:25.200	0.8
2	-4.76	-15.11	20/11/2015; 20:47:25.400	1.4
2	-4.53	-15.15	20/11/2015; 20:47:25.600	2.78
2	-4.21	-15.18	20/11/2015; 20:47:25.800	3.47
2	-3.82	-15.19	20/11/2015; 20:47:26.000	5
2	-3.37	-15.15	20/11/2015; 20:47:26.200	7.85
2	-2.92	-15.07	20/11/2015; 20:47:26.400	8.72
…				

spatial location (attribute)

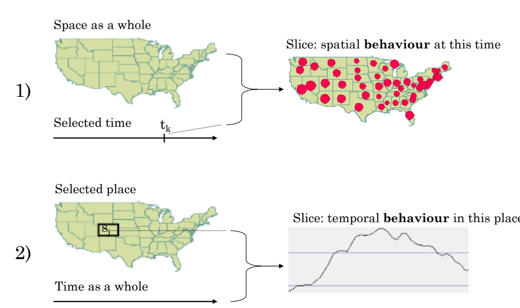
Spatial time series

Attribute values referring to spatial locations and times

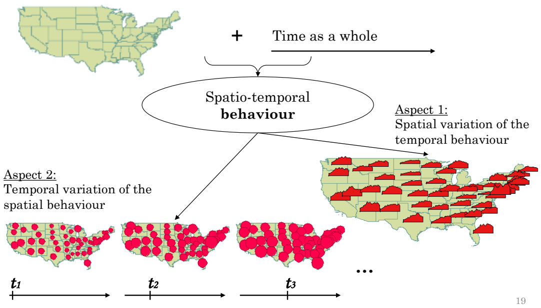
Complex behaviour

- In case of 2 (or more) referrers, we need to analyse the behaviour of the attributes over a complex reference set consisting of all available combinations of values of the referrers.
  - I.e., the behaviour of A over the Cartesian product $R_1 \times R_2 \times \ldots$
- In case of spatial time series: analyse the distribution of the attribute values over space and time
- Such a complex behaviour cannot be represented by a single image and observed as a whole.
- To study and describe a complex behaviour, we need to decompose it into slices and aspects

Decomposing a complex behaviour: slices

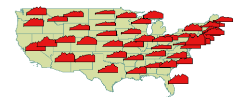


Decomposing a complex behaviour: aspects

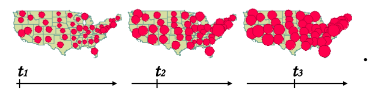


Two complementary views of spatial time series

1. As local time series: a set of time series of attribute values in different locations

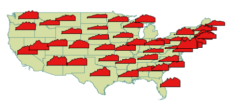


2. As time-varying spatial situations: sequence of spatial distributions of attribute values at different times



These views require different visualisation and analysis techniques.

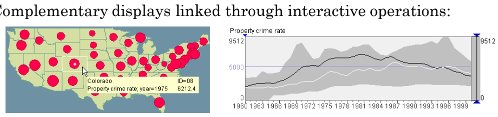
Visualisation of local time series



Limitations of a diagram map:

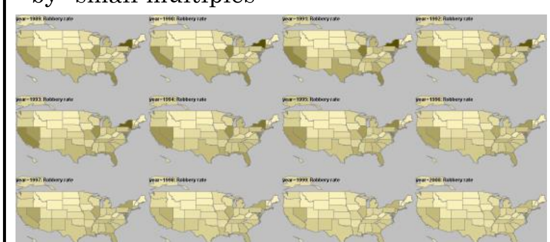
- Not applicable to long time series.
- Not applicable when spatial locations are numerous (not enough space for diagrams).
- Even when spatial locations are relatively few but irregularly distributed, diagrams in dense areas may overlap.

Complementary displays linked through interactive operations:



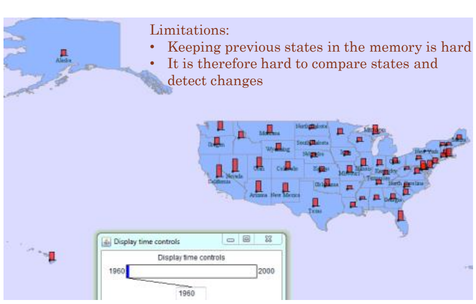
- Limitation: no overall view of all time series and their spatial distribution.

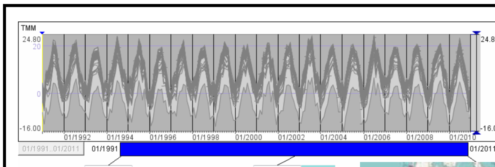
Visualisation of spatial situations by "small multiples"



Limitations:
- Applicable to a small number of different time references (time steps)
- Hardly applicable to more than one attributes

Visualisation of spatial situations by map animation





Example: time series of weather attributes (monthly averages of temperature, precipitation, wind, …) for 71 weather stations over Germany. Length of the time series: 241 time steps (months from January 1991 till January 2011).

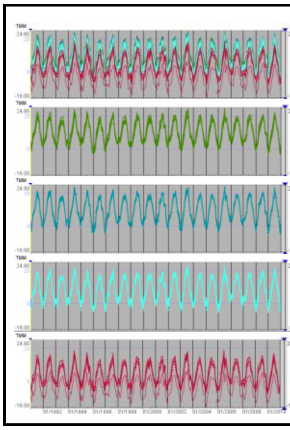
Analysis task: study how the values of the weather attributes are distributed over Germany and over time.

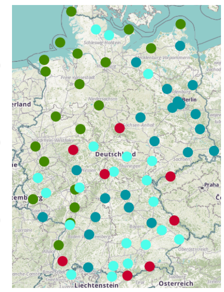
The data are too large for purely visual exploration. This is a case for using partition-based clustering.

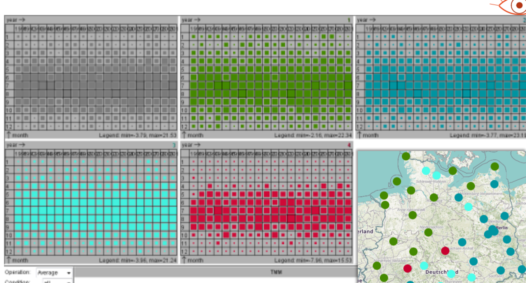
Use of partition-based clustering for analysis of spatial time series

Clustering of local time series and spatial situations

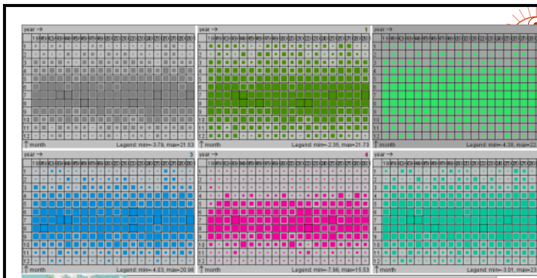
Subtask 1: study the spatial distribution of the local time series



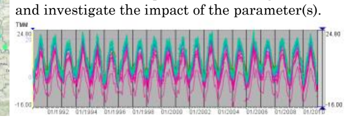


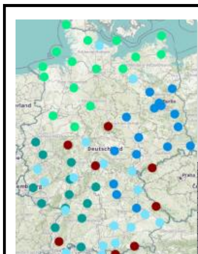


2D time histogram is convenient for exploration of time series involving a repeating time cycle. Here we have the cycle of 12 months of a year.
A histogram is an aggregated representation of multiple time series. The sizes of the symbols in the cells may be proportional to sums, averages, etc.



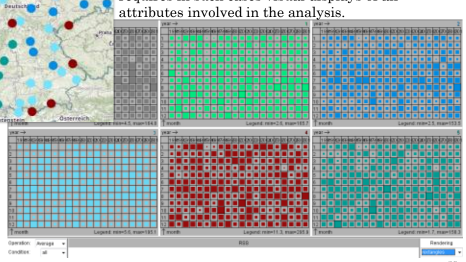
As usual, we need to perform clustering several times and investigate the impact of the parameter(s).





Clustering can be applied simultaneously to time series of several attribute. In this example: monthly averages of the minimal, mean, and maximal daily temperatures and monthly total precipitation amounts.

Exploration and interpretation of clustering results requires in such cases visual displays of all attributes involved in the analysis.

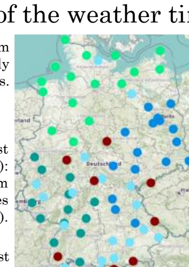


What have we learned about the spatial distribution of the weather time series?

Northwest: warm winters, moderately warm summers.

West, southwest (dark cyan): warm winters, warm summers (sometimes too warm).

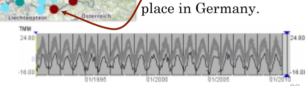
West, southeast (light cyan): slightly cooler winters and summers, higher precipitation.



East: colder winters, warmer summers.

Higher altitudes: generally colder climate, high precipitation.

Zugspitze: the coldest place in Germany.



Subtask 2: study the temporal variation of the spatial situations

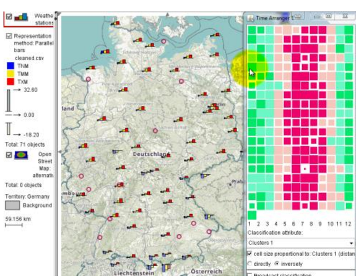
- A spatial situation is the distribution of attribute values over a set of spatial locations. We would like to simplify the task by putting the spatial situations into groups by similarity. PBC is a helpful means for this.
- Input to PBC consists of sequences of attribute values. The nature and meaning of the sequences are irrelevant to the clustering algorithm.
- In using PBC for subtask 1, we composed each sequence from attribute values associated with a single location but referring to different times. PBC output: groups of similar locations.
- For subtask 2, we can compose sequences from attribute values referring to the same times but to different locations. PBC output: groups of similar times.

Technically speaking

identifiers	TimeID=08/1998: TMM	TimeID=09/1998: TMM	TimeID=10/1998: TMM	TimeID=11/1998: TMM	TimeID=12/1998: TMM	TimeID=01/1999: TMM	TimeID=02/1999: TMM	TimeID=03/1999: TMM	TimeID=04/1999: TMM
Koeln-Bonn	17.1	14.7	9.7	3.2	3.6	4.8	2.6	7.5	10.1
Konstanz	19.4	13.9	10.3	2.3	1.3	1.7	0	6.3	9.6
Angermuende	16.2	13.9	8.5	0.7	0.3	2.5	0.7	4.9	9.1
Lahr	18.9	14.7	11	3.1	2.9	4.2	1.6	7	10.9
Leipzig/Halle	17.9	14.1	9.2	1.9	1.1	3.3	0.9	6	9.7
Lindenberg	16.8	13.9	8.6	0.9	0.3	2.4	0.9	5.3	9.9
Magdeburg	17.1	14	8.6	1.6	1.2	3.3	1.1	5.7	9.9
Meiningen	16.3	12.3	7.6	0.4	-0.4	1.1	-0.9	4.4	8.1
Muenchen/Flughafen	18.4	13.2	9.7	1.3	-0.1		-1	5.4	9
Muenster/Osnabrueck	16.7	14.4	9.2	3.1	3.3	4.8	2.6	6.9	10.1
Norderney	16.2	14	9.8	3.7	3.3	4.5	3.1	5.9	8.9
Augsburg	17.5	12.8	9.4	1.1	0	0.9	-1.3	4.9	8.1
Nuerburg-Barweiler	16	12.6	7.4	0.8	1.4	2.5	-0.2	5.1	7.7
Nuernberg	18.1	13.5	9.4	1.7	0.6	1.7	-0.2	5.7	9.1
Oberstdorf	15.8	11	7.8	-1.7	-2.5	-2	-3.4	2.1	5.3
Lippspringe	16.2	14.3	8.7	2.8	2.3	4.4	1	6.4	9.6
Potsdam	17.1	13.9	8.3	1.2	0.5	2.8	1.1	5.4	10
Regensburg	18.7	13.5	9.2	1.3	-0.6	0.1	-0.9	5.6	9.2
Rostock-Warnemuende	16.5	14.3	9.2	2.5	1.4	3.3	1.8	5.2	8.5
Saarbruecken-Ensheim	17.7	13.5	9	1.6	1.6	3.3	1.2	6.2	9.2
Schleswig	15.1	13.5	8.7	1.8	1.1	2.9	1	4.8	8
Stuttgart-Echterdingen	18.2	13.8	9.8	1.8	1.1	2.8	0.2	6	9.4
Trier-Petrisberg	18.1	14	9.2	2.3	2.5	3.7	1.7	6.7	9.7
Wasserkuppe	13.4	9.6	4.5	-2.2	-2.1	-0.4	-3.7	1.6	5
Wendelstein	10.6		2.7	-5.1	-3.4	-2.3	-8.1	-1.5	0.1
Wuerzburg	18.8	13.9	9.6	2.2	1.4	2.6	0.6	6.2	9.9
Bamberg	17.8	13.4	9.7	2.1	0.6	2	0.3	5.7	9.1
Berlin-Tegel	17.6	14.8	9.1	1.9	1.2	3.6	1.6	6	10.6
Karlsruhe	20.3	15.4	11	3.5	3.2	4.5	2.1	7.7	11.4

Sort by: No selection | Ascending | TableLens | condensed | Attribute...

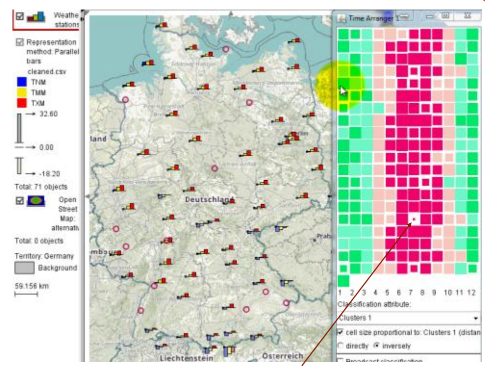
Data with 2 references can be organised in a table where rows correspond to one referential component and columns to the other. PBC can be applied either to rows (as we did before) or to columns. In our example, application of PBC to the columns, which correspond to times, will group the times by similarity of the values in the different locations.



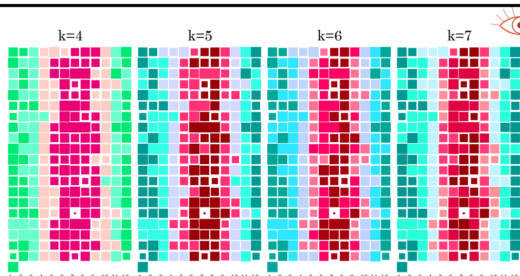
PBC has been applied to the values of three temperature attributes. The output is grouping of the months, which can be explored using a temporal display, such as a calendar view.

The temporal display shows how months with similar spatial situations are distributed over the whole time period.

To see what spatial situations correspond to each time cluster, we use a dynamic link between the calendar display and an animated map. By clicking on the calendar display, we select representatives of the clusters. The map is updated to show the spatial situation for the selected time.

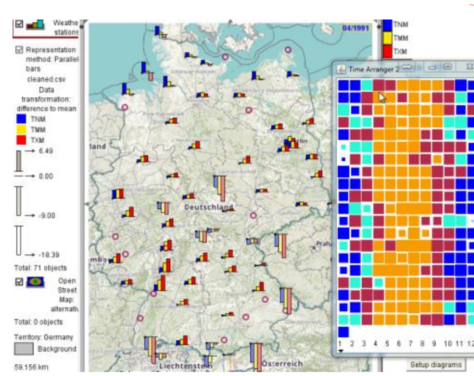


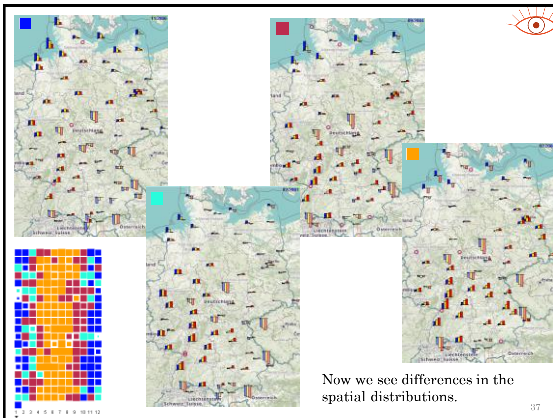
Outlier: extremely warm weather



The pattern of the temporal variation expectably adheres to the seasonal cycle, with some variations from year to year.
The spatial situations differ mostly quantitatively (higher or lower values) rather than in terms of the spatial arrangement of the values.

To see whether the spatial patterns change, we transform the absolute values of the temperatures into differences from the country's mean.





Now we see differences in the spatial distributions.

Analysis of spatial time series using PBC

A summary

- Spatial time series have 2 referrers: space (set of locations) and time (set of time steps – moments or intervals). The function represented by the data has complex behaviour (dependence on two variables).
- To study the complex behaviour, we decompose it into two aspects:
  - Spatial distribution of local temporal variations
  - Temporal variation of spatial distributions
- PBC is helpful in studying both aspects:
  - Apply PBC to the set of the local time series, then investigate the spatial distribution of the cluster membership.
  - Apply PBC to the set of the spatial distributions (values associated with different locations at the same time step), then investigate the temporal distribution of the cluster membership.
- General rules: try different parameters, visually investigate results, transform data whenever reasonable

Questions?

Spatial time series, aspects of a complex behaviour, two-way application of partition-based clustering

Spatial events

Objects having spatial locations and existence times

Spatial event data

- Spatial event data structure:
  - 1 referrer: set of objects
  - 2 mandatory attributes: *spatial location + time of existence*
  - any other attributes, further called *thematic attributes*
- Spatial location and existence time are *attributes* of the objects
  - ⇒Analysis tasks address the behaviour (distribution) of the spatial locations and existence times over the set of objects.
- However, space and time can also be considered as independently existing *containers* of the objects.
  - ⇒Analysis tasks may be equivalently re-formulated in terms of the distribution of the objects over the space and time, i.e., the *spatio-temporal distribution* of objects

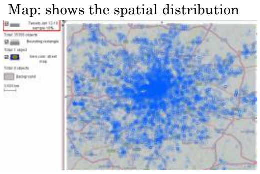
Space and time as object containers

- Considering space and time as containers of objects is quite intuitive.
  - We can easily imagine space and time without objects. It is difficult to do the same for other attributes (e.g., size).
- In visualisation, it is typical to represent space and time by display dimensions and objects by marks located within the display space.
  - I.e., the display conveys the idea of the objects being contained in space and/or time.
  - This representation is usual for people and therefore easily understandable.

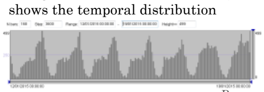
➢ We take this *absolute view* of time and space as object containers.

Visual displays of spatial events

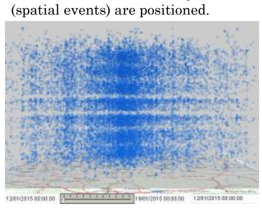
Map: shows the spatial distribution



Temporal display (e.g., frequency histogram of the occurrence times): shows the temporal distribution



Space-time cube*: represents 2D space + time as a single 3D continuum, in which the objects (spatial events) are positioned.

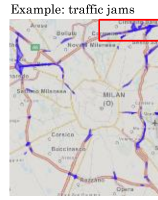


Hägerstrand. T. (1970). "What about people in regional science?" *Papers of the Regional Science Association*; 24:7-21.

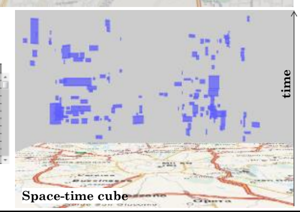
43

Spatial events may be extended in time and/or space

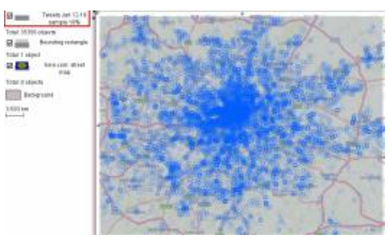
Example: traffic jams



	extent, m	Begin time	End time	Duration, minutes
1	1909.82	04/04/2007 05:04:56	04/04/2007 08:08:19	183.4
2	291.75	04/04/2007 05:22:46	04/04/2007 05:39:20	16.6
3	734.22	04/04/2007 05:06:35	04/04/2007 05:41:35	35.0
4	576.55	04/04/2007 05:08:45	04/04/2007 05:46:38	37.9
5	491.55	04/04/2007 05:33:43	04/04/2007 06:03:21	29.6
6	143.53	04/04/2007 05:35:34	04/04/2007 05:45:41	10.1
7	459.94	04/04/2007 05:38:42	04/04/2007 06:07:08	28.4
8	473.88	04/04/2007 15:11:05	04/04/2007 15:56:18	45.2
9	358.59	04/04/2007 15:41:36	04/04/2007 16:16:29	34.9
10	1256.72	04/04/2007 10:32:13	04/04/2007 12:45:12	133.0
11	270.41	04/04/2007 12:14:09	04/04/2007 12:23:01	8.9
12	371.49	04/04/2007 06:00:20	04/04/2007 06:06:14	5.9



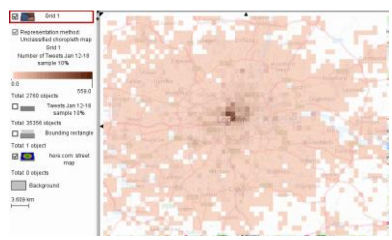
Spatial distribution of events



Point events (i.e., having no spatial extents) are typically represented on a map by dot symbols (small circles). Very often the dots overlap ⇒ semi-transparent drawing is recommended. The degree of transparency is adjusted to the dot density. Such a way of rendering helps us to observe the variation of the event density over the territory and detect spatial clusters of events.

45

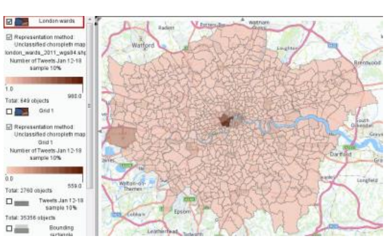
Spatial aggregation by a regular grid



Semi-transparent drawing does not tell us how many events occurred at different locations. Investigation of the spatial distribution can be supported by spatial aggregation of the events, e.g., by cells of a regular grid (here: 1×1 km). For each cell, the number of events is counted; additionally, thematic attributes can be summarized by computing the mean, mode, median, etc.

46

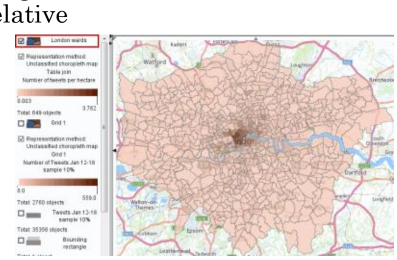
Spatial aggregation by arbitrary areas



Events can also be aggregated by arbitrary areas, such as administrative districts (e.g., to relate event numbers and/or attributes to other characteristics of the areas). However, arbitrary areas may significantly differ in sizes, which distorts the perception of the event density. ⇒ Choropleth maps, like this one, should **not** be used for representing counts by areas.

47

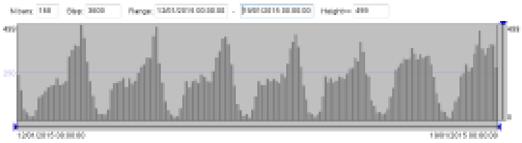
Aggregation: absolute counts transformed to relative



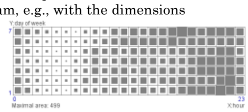
A correct image of the variation of the event density can be obtained by transforming the absolute event counts to relative w.r.t. the areas of the districts. When the events are related to lives and activities of people (e.g., tweets), it can also make sense to compute the relative counts w.r.t. the population of the districts.

48

Temporal distribution of events

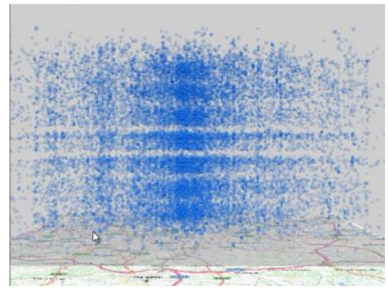


A frequency histogram of the event occurrence times shows the overall temporal distribution of the events and exhibits temporal patterns such as temporal trends and periodicity. A periodic temporal distribution can be additionally explored using a 2D histogram, e.g., with the dimensions corresponding to the days of the week and hours of the day.



49

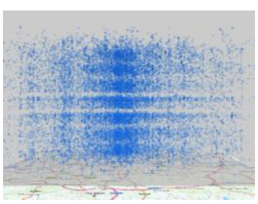
Spatio-temporal distribution of events

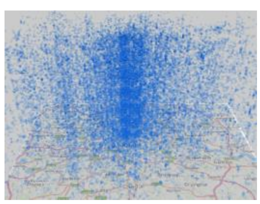


A static perspective view of a 3D representation is insufficient for observing and exploring the content. Interactive operations for changing the perspective (rotation, shifting, and tilting) are necessary.

50

Spatio-temporal distribution of events



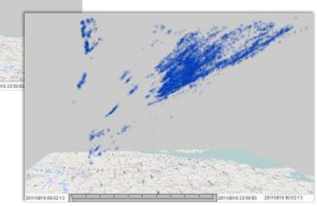


A space-time cube (STC) can exhibit several patterns of spatio-temporal distribution (types of spatio-temporal behaviour) appearing as horizontal "layers" and/or "gaps" (= periods of high and low event density), vertical "columns" (= high event density in some area for a long time), and "lumps" (= spatio-temporal clusters, i.e., groups of events that occurred closely in space and time).

51

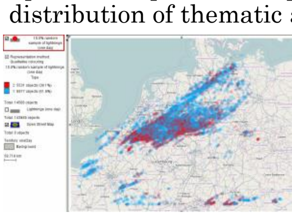
Spatio-temporal distribution of events

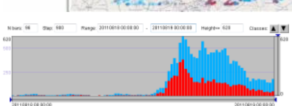
Diagonally elongated clusters may mean that the events were caused by a moving phenomenon, such as a storm or parade.
This example: lightning events.

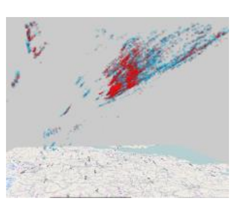


52

Spatial, temporal, and spatio-temporal distribution of thematic attribute values

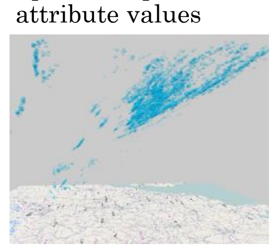


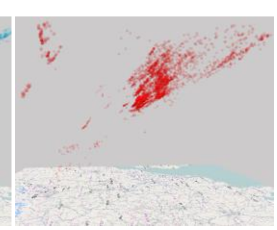




53

Spatio-temporal distributions of different attribute values

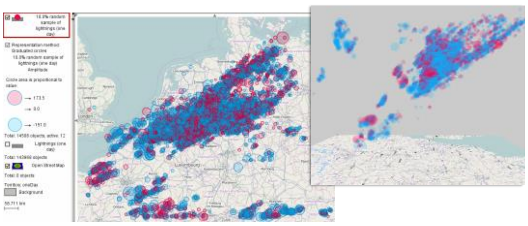




Different attribute values are selected using interactive attribute-based filtering.

54

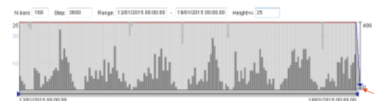
Visualisation of thematic attributes of events



Although retinal visual variables can be used for representing the values of thematic attributes of spatial events on a map and in a space-time cube, the visual clutter and overlapping of marks make the displays illegible and practically useless for the analysis.

55

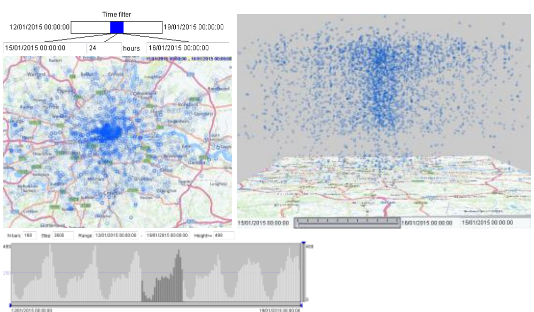
Spatial filtering of events



Note the focuser position

56

Temporal filtering of events



57

Insufficiency of visualisation and interaction techniques

- Map and STC do not work well when the events are numerous and dense.
  - The spatial and spatio-temporal distribution of the events themselves and their thematic attributes cannot be effectively analysed due to display clutter and greatly overlapping marks.
- Besides, STC may not work well when the time span of the data is long.
  - Spatio-temporal distribution patterns are not well visible.
- Interactive filtering can only partly reduce the display clutter while eliminating the overall view.
- Data transformations and computational techniques are strongly needed.
  - E.g., computational detection of clusters – to be considered later

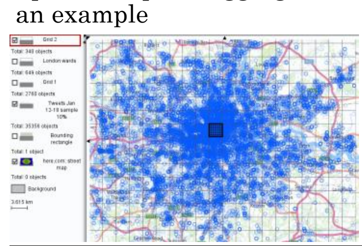
58

Spatio-temporal aggregation of events

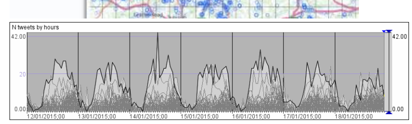
- Spatial events and their thematic attributes can be aggregated spatially by areas (as considered before) and, simultaneously, by time intervals.
- Resulting data type: **spatial time series** of
  - event counts, densities, counts per capita, …
  - statistical summaries of thematic attributes: mean, median, mode, minimum, maximum, quantiles, …
- The time series can be analysed using partition-based clustering.

59

Spatio-temporal aggregation of events: an example



Geolocated tweets have been aggregated by cells of a regular grid and hourly time intervals. For each grid cell, there is a time series of hourly tweet counts.

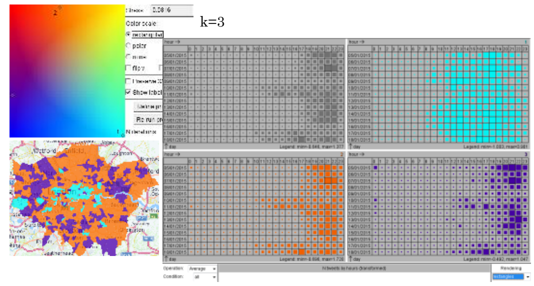


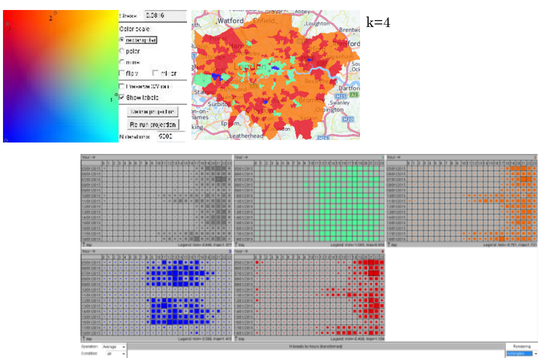
60

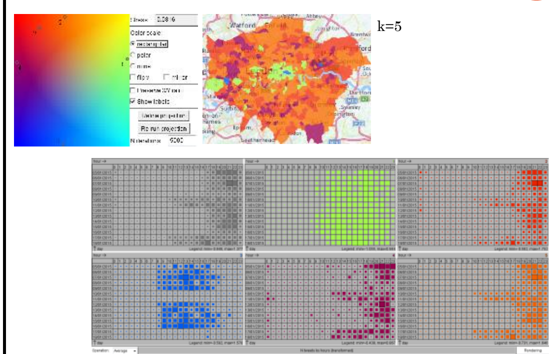
Techniques for analysis of spatio-temporal event data: a summary

- Visual displays: map, space-time cube, time-based frequency histograms, various display types for thematic attributes.
- Interactive techniques: selection, classification, filtering.
- Data transformations: spatial aggregation, spatio-temporal aggregation.

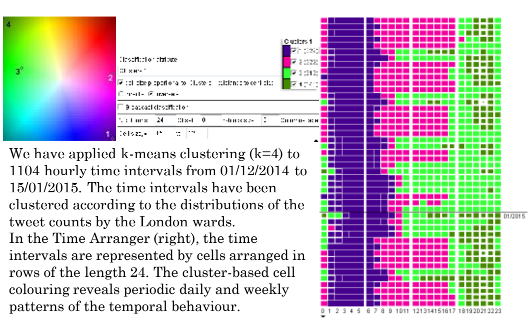
Interactive visual analysis of aggregated event data by PBC





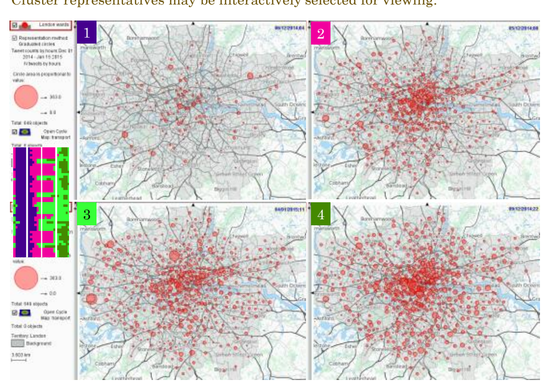


PBC of spatial situations

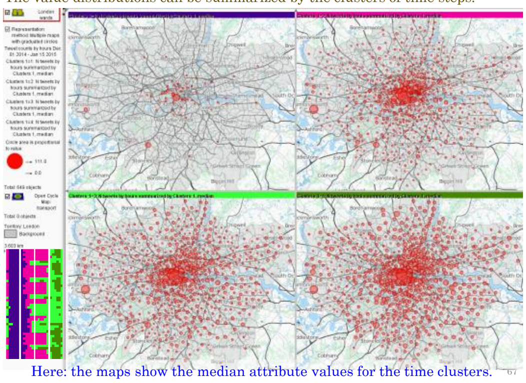


We have applied k-means clustering (k=4) to 1104 hourly time intervals from 01/12/2014 to 15/01/2015. The time intervals have been clustered according to the distributions of the tweet counts by the London wards.
In the Time Arranger (right), the time intervals are represented by cells arranged in rows of the length 24. The cluster-based cell colouring reveals periodic daily and weekly patterns of the temporal behaviour.

Cluster representatives may be interactively selected for viewing.



The value distributions can be summarized by the clusters of time steps.



Here: the maps show the median attribute values for the time clusters.

Questions?

Spatial events; spatio-temporal aggregation of spatial events

Movement data (trajectories)

Time series of spatial locations of moving objects

Trajectories: data and visual representation

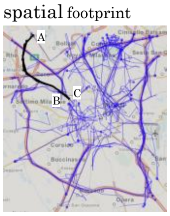
Sequences of position records

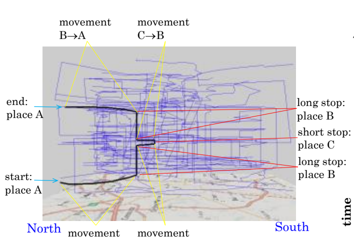
	Car id	pointN	longitude	latitude	time
1	104876	1	9.119127	45.558304	04/04/2007 06:45:15
2	104876	2	9.142448	45.559753	04/04/2007 06:48:49
3	104876	3	9.156955	45.554982	04/04/2007 06:54:54
4	104876	4	9.156504	45.55017	04/04/2007 07:00:12
5	104876	5	9.156504	45.55017	04/04/2007 07:11:08
6	104876	6	9.158044	45.547703	04/04/2007 07:13:26
7	104876	7	9.156909	45.547688	04/04/2007 07:19:23
8	104876	8	9.162037	45.554867	04/04/2007 07:40:02
9	104876	9	9.167628	45.55907	04/04/2007 08:02:32
10	104876	10	9.172845	45.555725	04/04/2007 08:05:38
11	104876	11	9.172696	45.555492	04/04/2007 10:03:31
12	104876	12	9.188886	45.54488	04/04/2007 10:08:38
13	104876	13	9.163299	45.557983	04/04/2007 10:12:05
14	104876	14	9.162168	45.554855	04/04/2007 10:13:51
15	104876	15	9.162158	45.55487	04/04/2007 11:38:23
16	104876	16	9.162622	45.557976	04/04/2007 12:08:17
17	104876	17	9.16232	45.55496	04/04/2007 12:09:19
18	104876	18	9.162361	45.554943	04/04/2007 15:30:22
19	104876	19	9.122161	45.55825	04/04/2007 15:38:51
20	110800	1	9.266509	45.386322	04/04/2007 05:21:45
21	110800	2	9.261211	45.40307	04/04/2007 05:22:57
22	110800	3	9.247442	45.418125	04/04/2007 05:24:13
23	110800	4	9.254333	45.43362	04/04/2007 05:29:45
24	110800	5	9.257282	45.451482	04/04/2007 05:32:44
25	110800	6	9.252168	45.468788	04/04/2007 05:34:21
26	110800	7	9.251433	45.48671	04/04/2007 05:35:48
27	110800	8	9.258238	45.504066	04/04/2007 05:37:05
28	110800	9	9.280647	45.522255	04/04/2007 05:38:26
29	110800	10	9.278728	45.53516	04/04/2007 05:39:48
30	110800	11	9.274316	45.533176	04/04/2007 11:57:53
31	110800	12	9.261258	45.519483	04/04/2007 11:59:21
32	110800	13	9.256271	45.502083	04/04/2007 12:00:51
33	110800	14	9.250924	45.484287	04/04/2007 12:02:11
34	110800	15	9.251606	45.465813	04/04/2007 12:03:32
35	110800	16	9.258425	45.44919	04/04/2007 12:04:53
36	110800	17	9.251738	45.43149	04/04/2007 12:06:14
37	110800	18	9.251131	45.415356	04/04/2007 12:07:39
38	110800	19	9.260662	45.399742	04/04/2007 12:08:54
39	110800	20	9.270314	45.382904	04/04/2007 12:10:10
40	116291	1	9.234817	45.508648	04/04/2007 18:01:18
41	116291	2	9.257177	45.51305	04/04/2007 18:08:57
42	116291	3	9.255232	45.498352	04/04/2007 18:09:05

Space-time cube

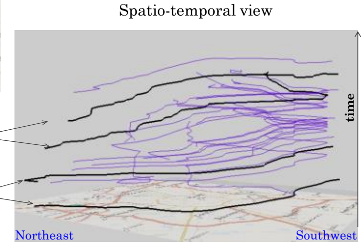
Spatio-temporal view

spatial footprint



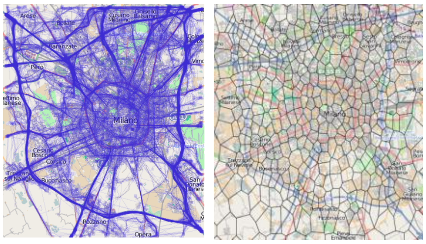


Spatio-temporal view



The interpretation of the line slope is the same as for OD moves.

Spatio-temporal aggregation of trajectories



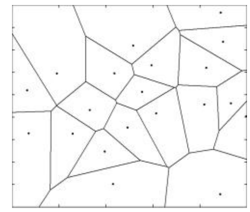
Analogously to spatial events, analysis of trajectories is concerned with the spatio-temporal variation of the movements of objects. Like in case of spatial events, spatio-temporal aggregation may be useful, especially for large data. ST aggregation requires division of the territory into compartments and division of the time into intervals.

73

Voronoi tessellation (a.k.a. Voronoi diagram)

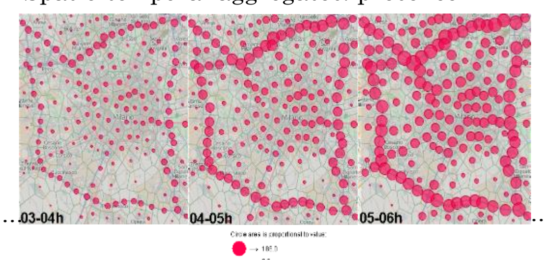
Used for building irregular grids

- The partitioning of a plane with N points into convex polygons *(cells),* such that
  - each polygon contains exactly one generating point
  - every point in a given polygon is closer to its generating point than to any other.
- The generating points are also called *seeds*.
- A Voronoi diagram is also known as a Dirichlet tessellation.
- The cells are called Dirichlet regions, Thiessen polytopes, or Voronoi polygons.



74

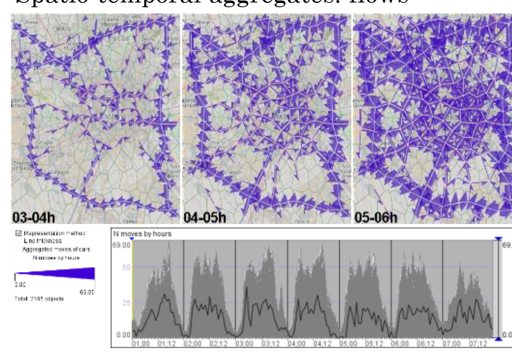
Spatio-temporal aggregates: presence



This example: hourly counts of cars in the compartments (cells of an irregular grid produced by means of Voronoi tessellation).

75

Spatio-temporal aggregates: flows



Flows

Link *(short for 'spatial link')*::= a spatial object representing directed relation, such as movement, between 2 locations.

A link is specified by a pair (origin place, destination place).

Links may have attributes such as number of moving objects, number of transitions, movement speed, ...

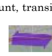
Links with attributes describing collective movements are called **flows**.

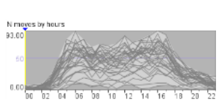
Flows may be represented on a map by half-arrow symbols with widths proportional to numeric attribute values.

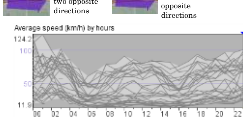
Spatial time series of flows::= attribute values of the links in different time moments or intervals:

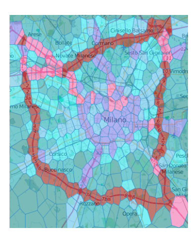
((origin, destination), time) → object count, transition count, speed, ...

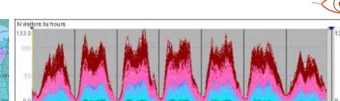

equal flows in two opposite directions


unequal flows in two opposite directions



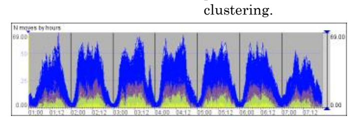


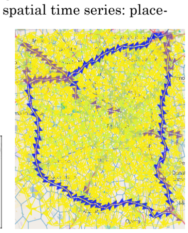




Spatio-temporal aggregation of movement data produces two types of spatial time series: place-based and link-based.

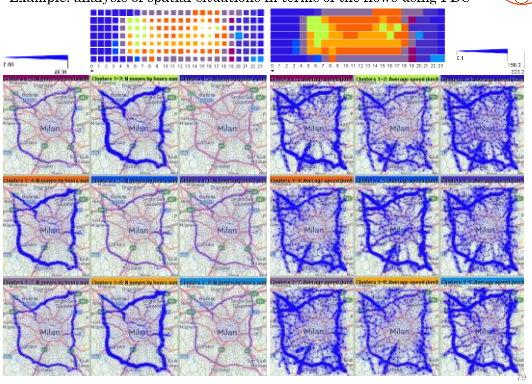
Both types of spatial time series can be analysed using partition-based clustering.





78

Example: analysis of spatial situations in terms of the flows using PBC



Spatio-temporal aggregation

A summary

- Spatio-temporal aggregation can be applied to spatial events and movement data; result: spatial time series.
- For ST aggregation:
  - Divide the territory into suitable compartments (such as regular grid, administrative division, irregular grid built by data-driven tessellation*).
  - Divide the time range of the data into intervals (typically equal length).
  - For each compartment and time interval:
    - Count the objects (events or moving objects) that were there at that time
    - Compute statistical summaries of thematic attributes
  - In case of movement data**, for each pair of compartments and time interval:
    - Count the objects that moved from the first to the second compartment
    - Compute statistical summaries of the moves: speed, duration, etc.

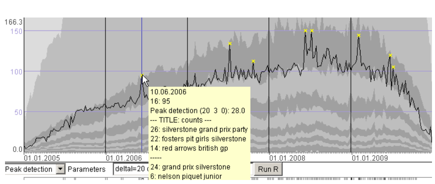
* to be introduced in following lectures; ** considered in more detail in following lectures

Questions?

Spatio-temporal aggregation of movement data; place-based and link-based time series

Extraction of spatial events from spatial time series

Events in time series



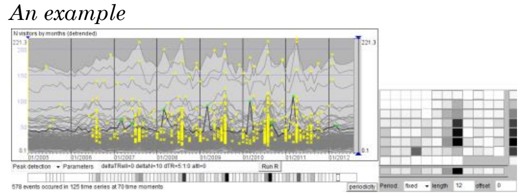
Time series of attribute values may contain abrupt changes of the values. Such changes can be treated as events, i.e., objects appearing at some time moment and having limited time of existence.

When time series refer to spatial objects or locations, their events are also located in space, i.e., they are spatial events.

Spatial events of interest, such as peaks, can be extracted from time series (e.g., using computational processing) and explored using visualisations and tools suitable for spatial events.

Extraction of events from time series

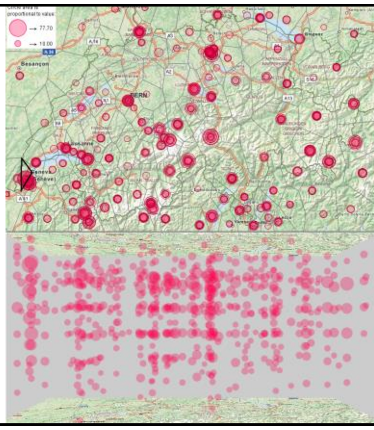
An example

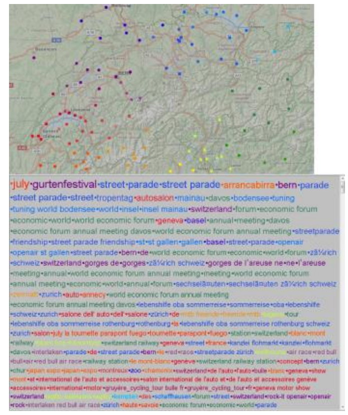


Spatial time series: counts of authors of photos published in Flickr by monthly time intervals and spatial compartments in Switzerland.

Events: peaks with minimal amplitude 10; extracted using an algorithm implemented in package R.

The crosses on the time graph mark the detected events. The calendar display (rows: years, columns: months) shows the temporal distribution of the number of the extracted events.



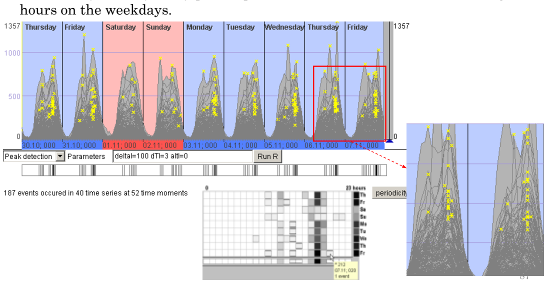


The text cloud display shows summarised titles of the photos from the places and times where the peak events occurred. The texts have been summarised by extracting frequent terms and combinations.

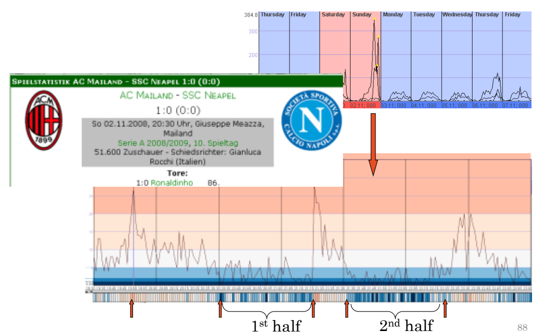
The symbols representing the peak events on the map are coloured based on their spatial locations. The colours are propagated to the text cloud display, which enables the viewer to locate the texts to areas on the map.

Another example

Peaks in time series of mobile phone call counts in mobile network cells over Milan, Italy. In very many places, peaks occur after the end of the working hours on the weekdays.



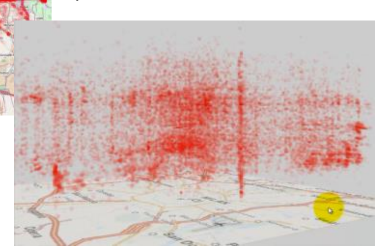
An area with peaks in unusual times



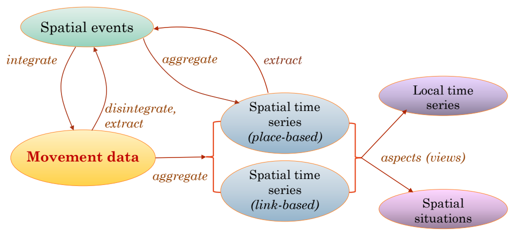
Spatial events and movement data

- Any item of movement data describes presence of some moving object in some spatial location at some time. It can be treated as a spatial event (event of presence), i.e., an object having spatial location and time of existence.
- A trajectory, therefore, can be considered as a sequence of events.
- Some events in trajectories may be of interest for analysis:
  - Trip starts and ends, stops
  - Too fast or too slow movement, too high acceleration, …
  - Coming close to a particular location or object
  - …
- Events of interest can be extracted from trajectories and analysed using visualisations and tools suitable for spatial events.

Example: low speed events extracted from trajectories of cars



Transformations of spatio-temporal data



91

Purposes of data transformations

- **Aggregation**
  - Supports abstraction, gaining an overall view of characteristics and behaviour
  - Reduces large data
  - Simplifies complex data
- **Extraction** *of events, etc.*
  - Selects a portion of data relevant to a task, enables focusing
  - Allows dealing with complex data portion-wise
- **Integration, disintegration, projection** (taking one of possible aspects)
  - Adapts data to analysis tasks

Questions?

Transformations of spatio-temporal data

93

Summary of the lecture

- Different types of spatio-temporal data: spatial time series, spatial events, movement data
- Spatial time series: two referrers, complex behaviour, need to decompose into aspects
  - Set of local time series distributed over space
  - Spatial situations (attribute value distributions) changing over time
- Analysis of spatial time series by two-way application of partition-based clustering
- Transformation of spatial events and movement data to spatial time series
- Extraction of spatial events from spatial time series and movement data

94