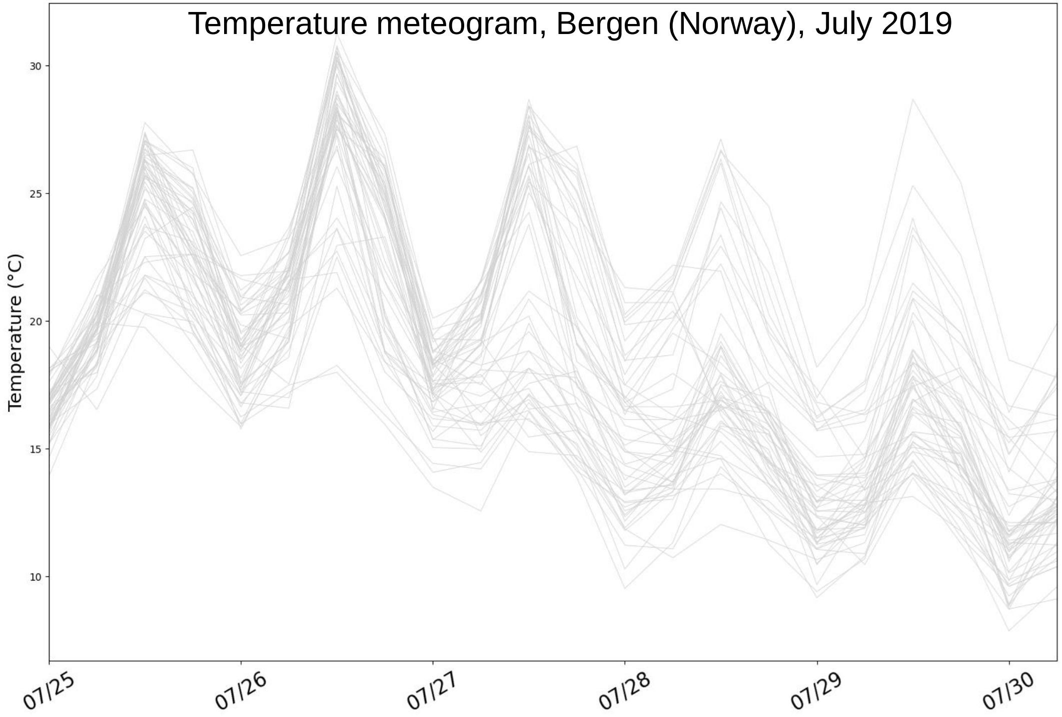
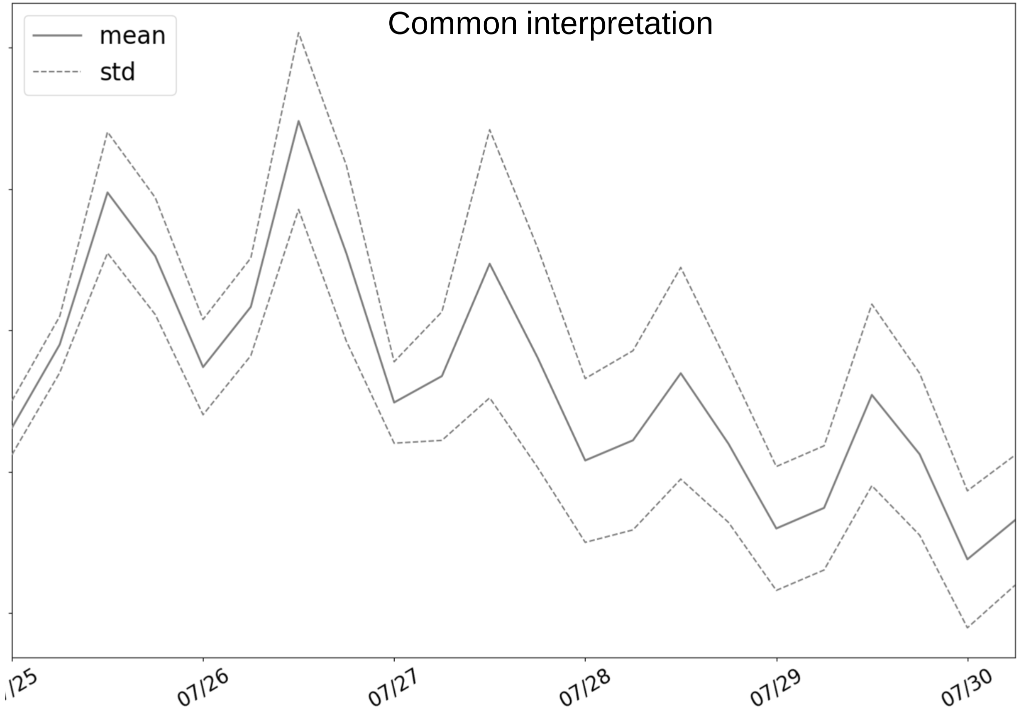
Revealing Multimodality in Ensemble Weather Prediction

Thomas Spengler

Presentation prepared by Natacha Galmiche

N. Galmiche[1], N.Blaser[1], T.Spengler[2], C.Spensberger[2], H.Hauser[1] and M.Brun[3]
[1] Department of Informatics [2] Geophysical Institute and [3] Department of Mathematics
of the University of Bergen, Norway





Ensemble data:

N (= 50) univariate time series (each called ***member***) evolved out of:

- Perturbed initial condition
- Different models

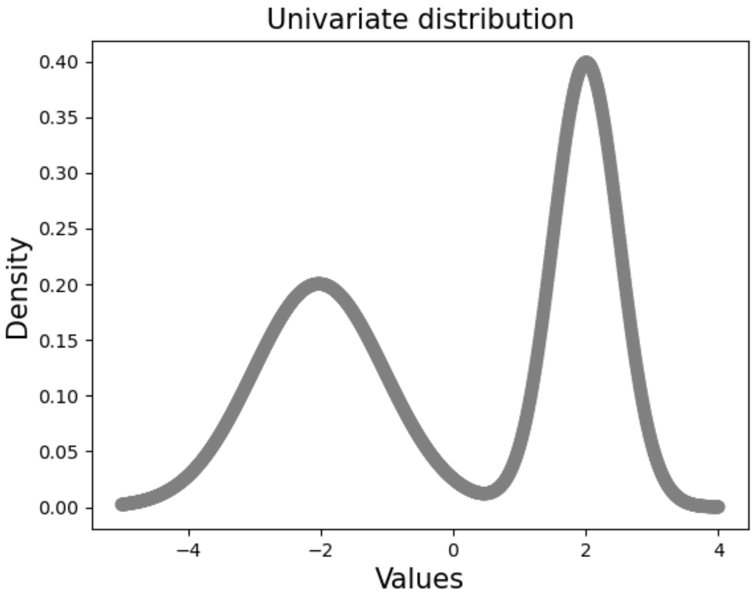
Common interpretation method:

- **Mean** as expected value
- **Standard deviation** as uncertainty

What is a multimodal distribution?

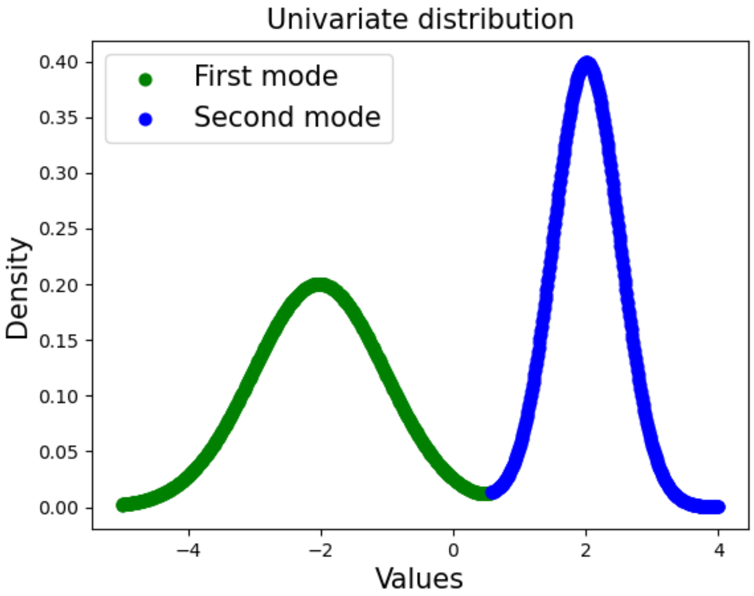
Multimodality:

Number of modes (i.e likely outcomes) ***k>1***



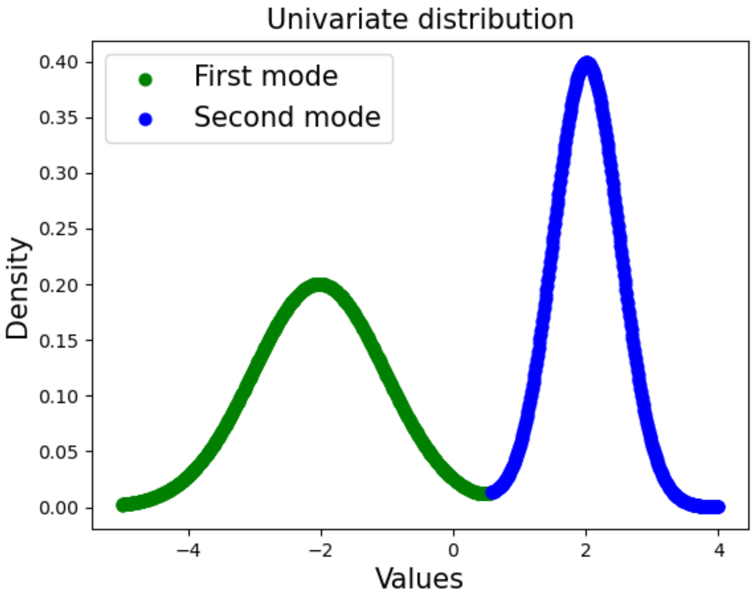
What is a multimodal distribution?

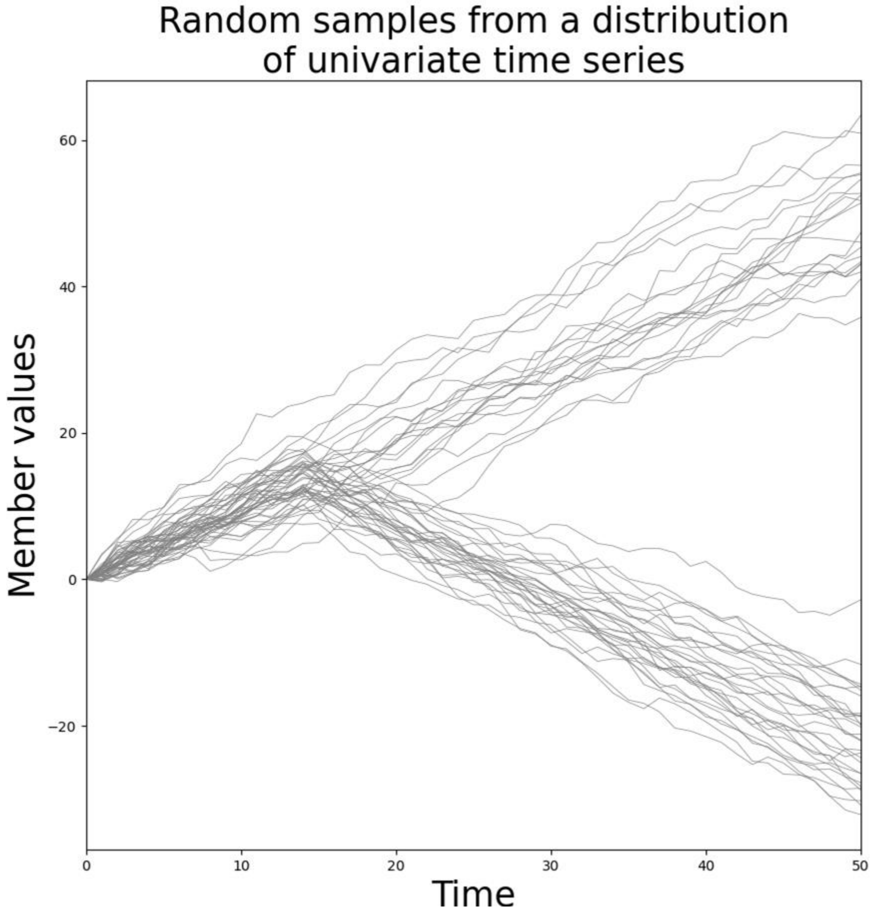
Multimodality:
Number of modes (i.e likely outcomes) ***k>1***



What is a multimodal distribution?

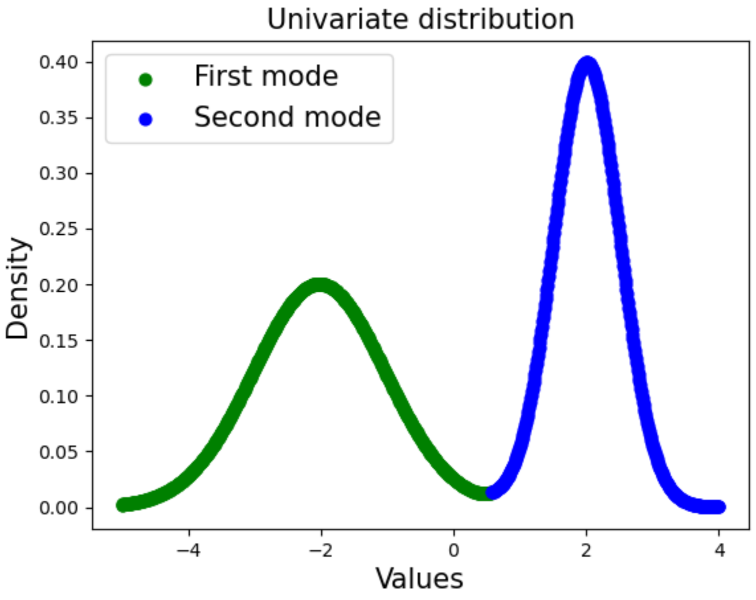
Multimodality:
Number of modes (i.e likely outcomes) ***k>1***

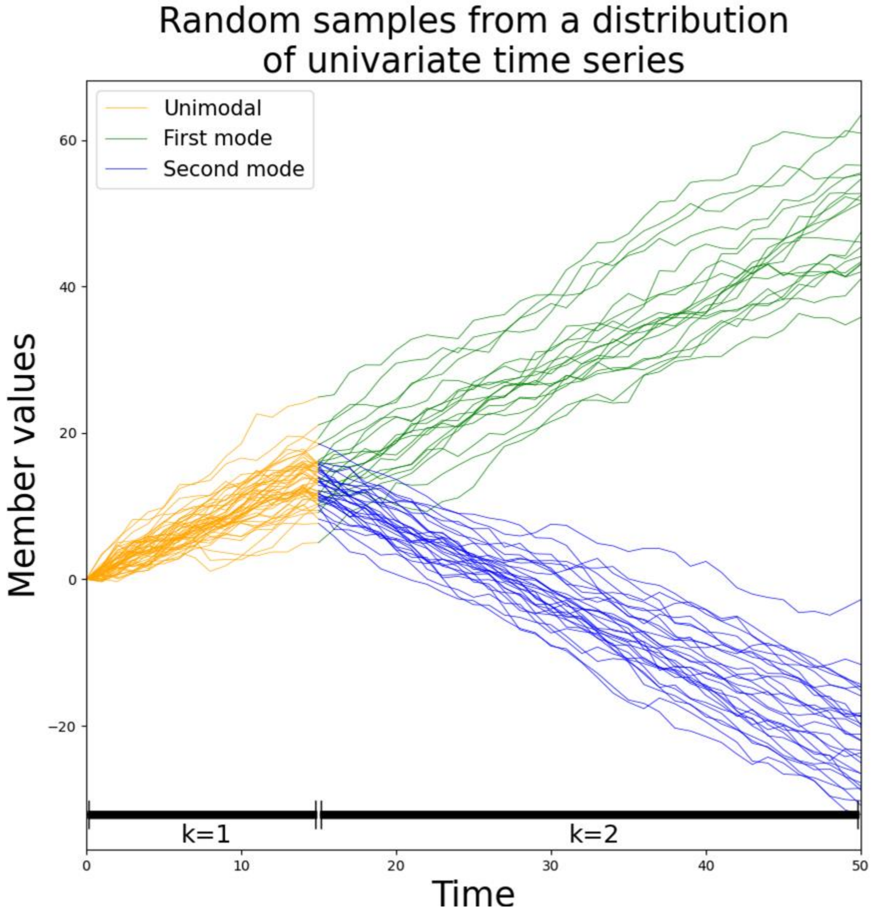




What is a multimodal distribution?

Multimodality:
Number of modes (i.e likely outcomes) ***k>1***





What is a multimodal distribution?

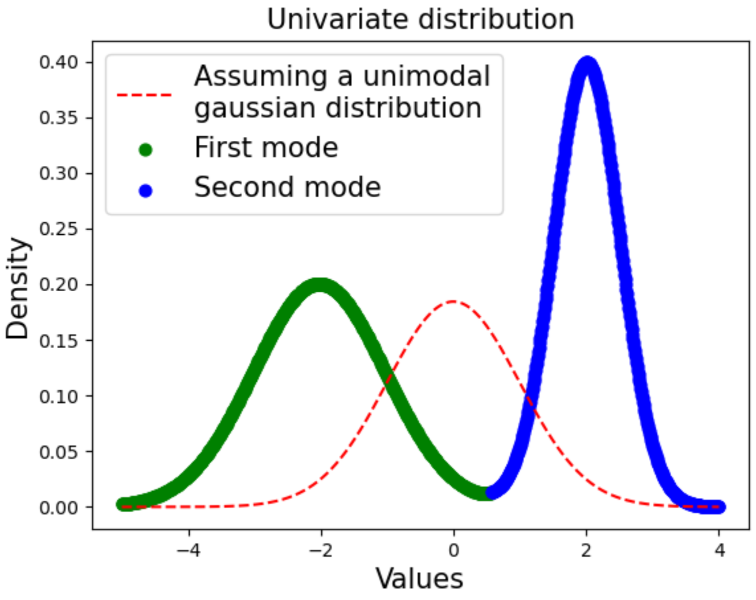
Multimodality:

Number of modes (i.e likely outcomes) ***k>1***

Problem?

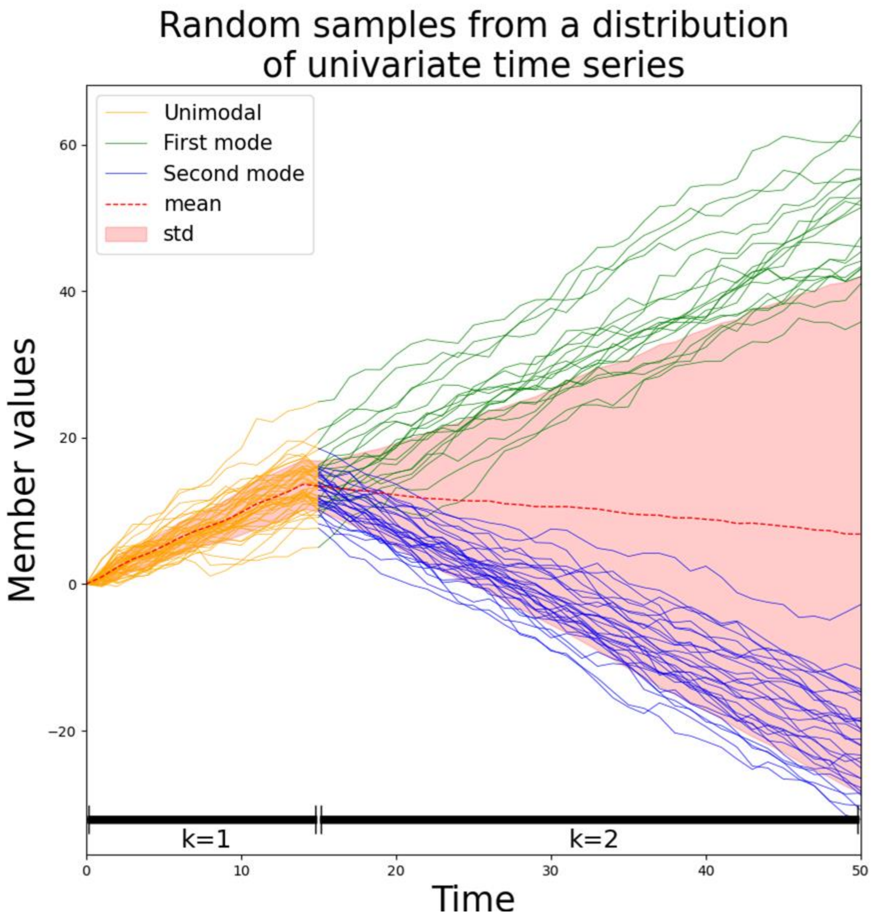
“mean + std” method:

Assumes a unimodal Gaussian distribution



Result:

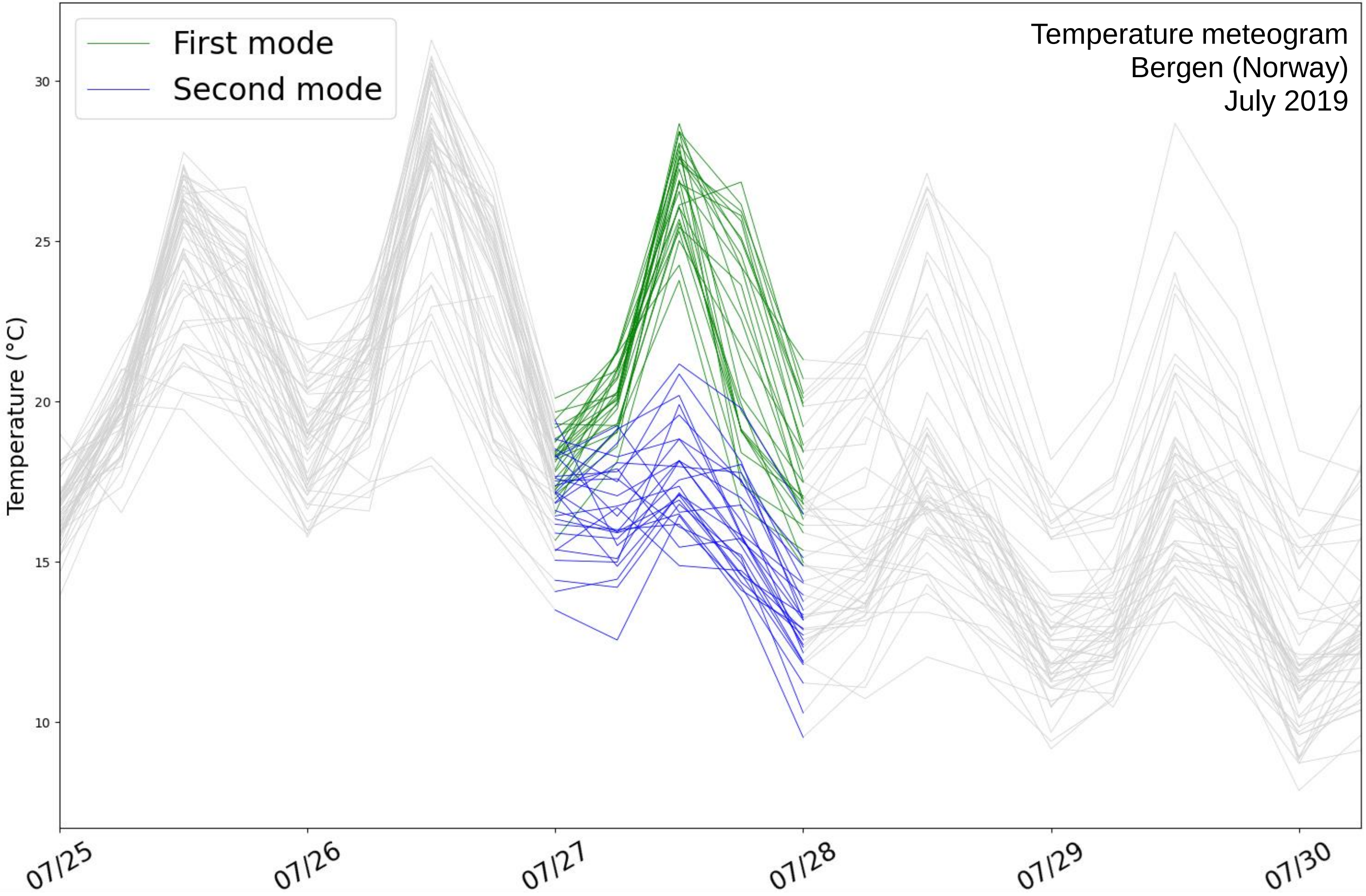
- Likely values are ignored
- Unlikely values are emphasized



Atmospheric system:

Chaos + Stability

=> prone to multimodality



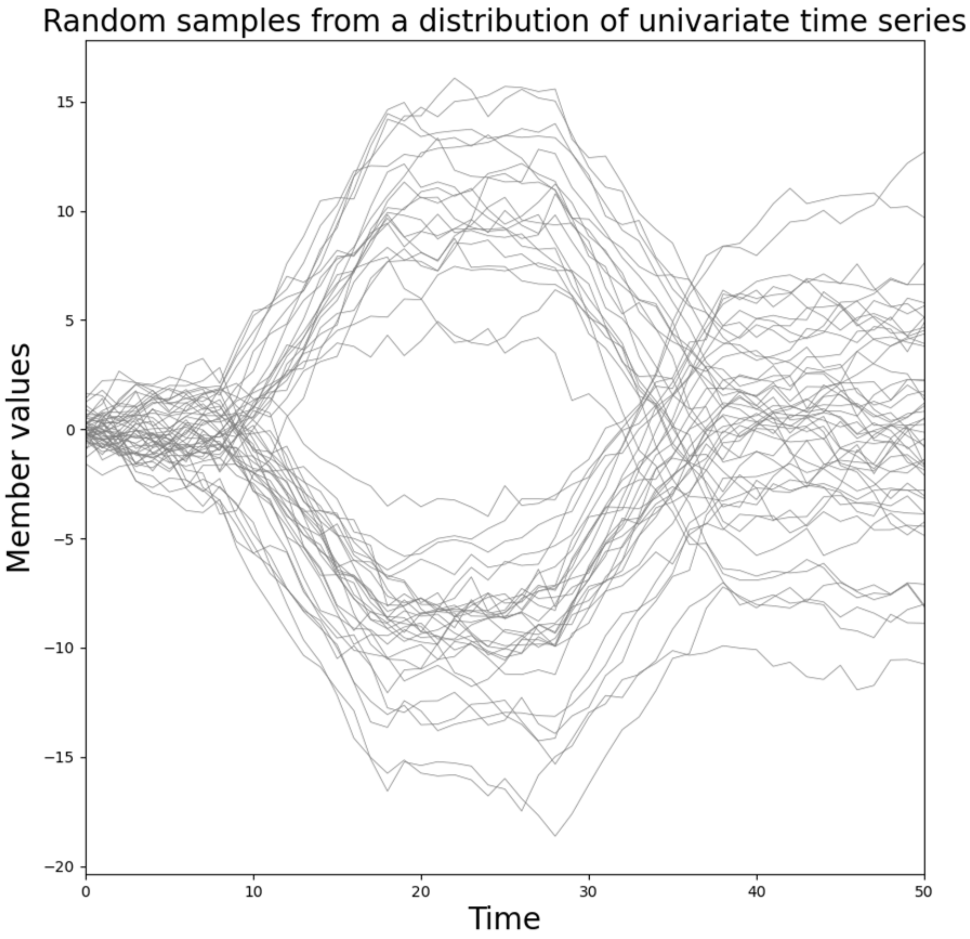
Tasks:

T1. Estimate the number of modes **k** at each time step **t**

T2. Summarize each mode at each **t**

T3. Determine when modes appear and/or disappear

T4. Determine all possible connections between modes



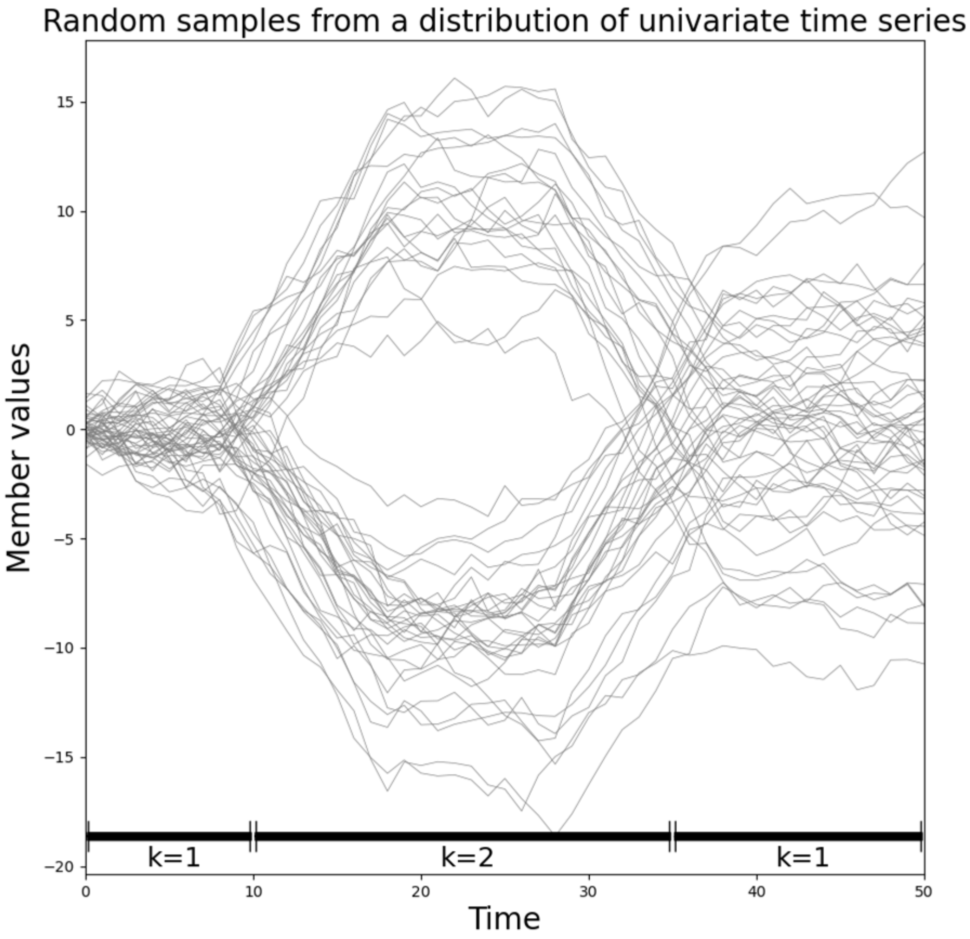
Tasks:

T1. Estimate the number of modes **k** at each time step **t**

T2. Summarize each mode at each **t**

T3. Determine when modes appear and/or disappear

T4. Determine all possible connections between modes



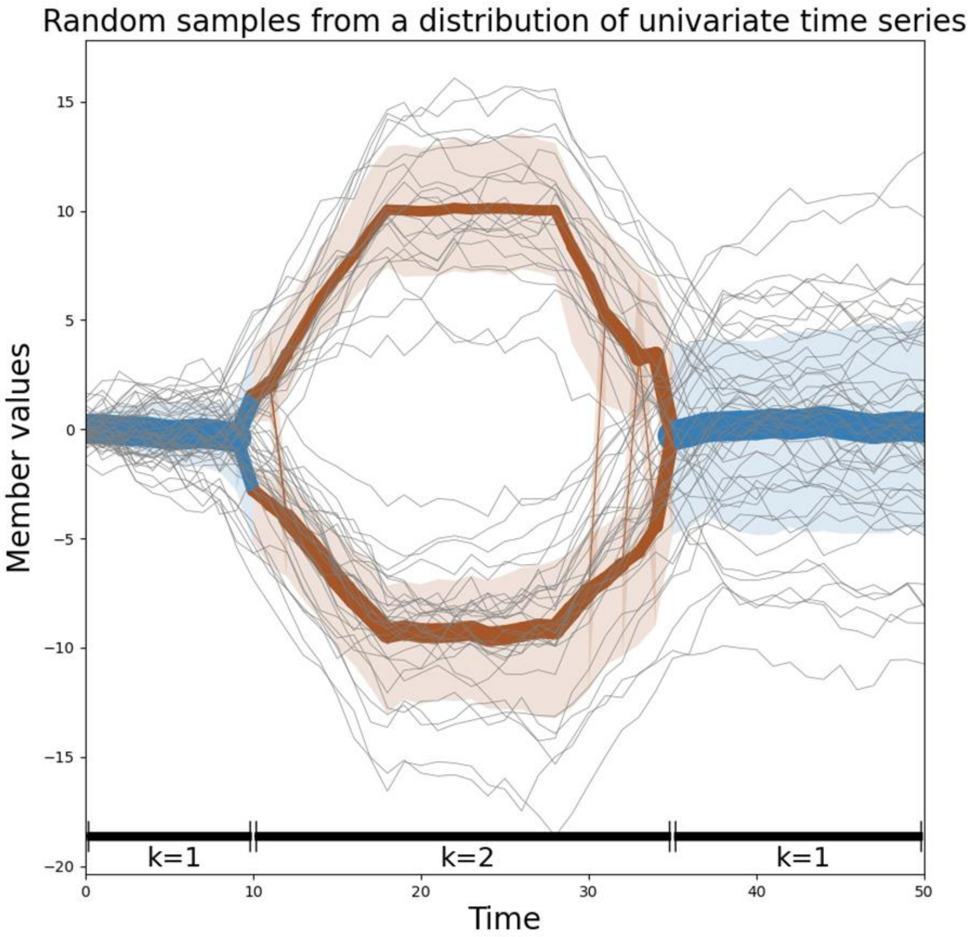
Tasks:

T1. Estimate the number of modes **k** at each time step **t**

T2. Summarize each mode at each **t**

T3. Determine when modes appear and/or disappear

T4. Determine all possible connections between modes



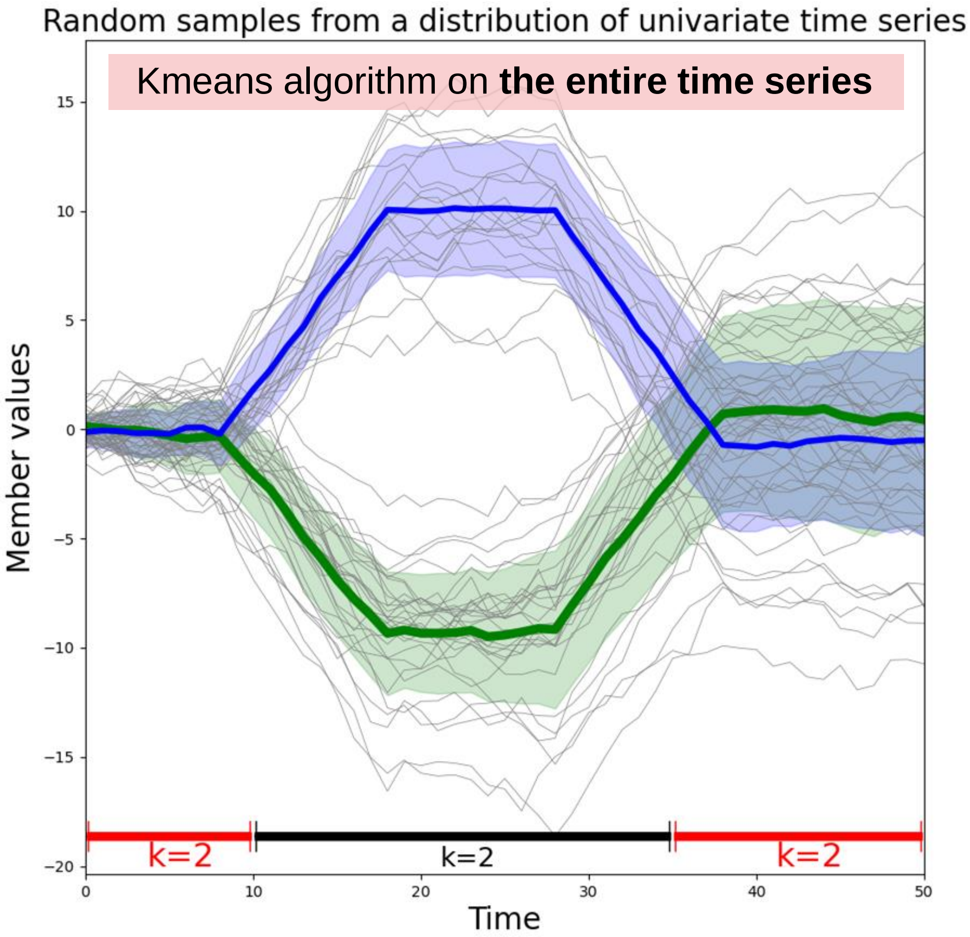
Tasks:

T1. Estimate the number of modes **k** at each time step **t**

T2. Summarize each mode at each **t**

T3. Determine when modes appear and/or disappear

T4. Determine all possible connections between modes



Tasks:

T1. Estimate the number of modes **k** at each time step **t**

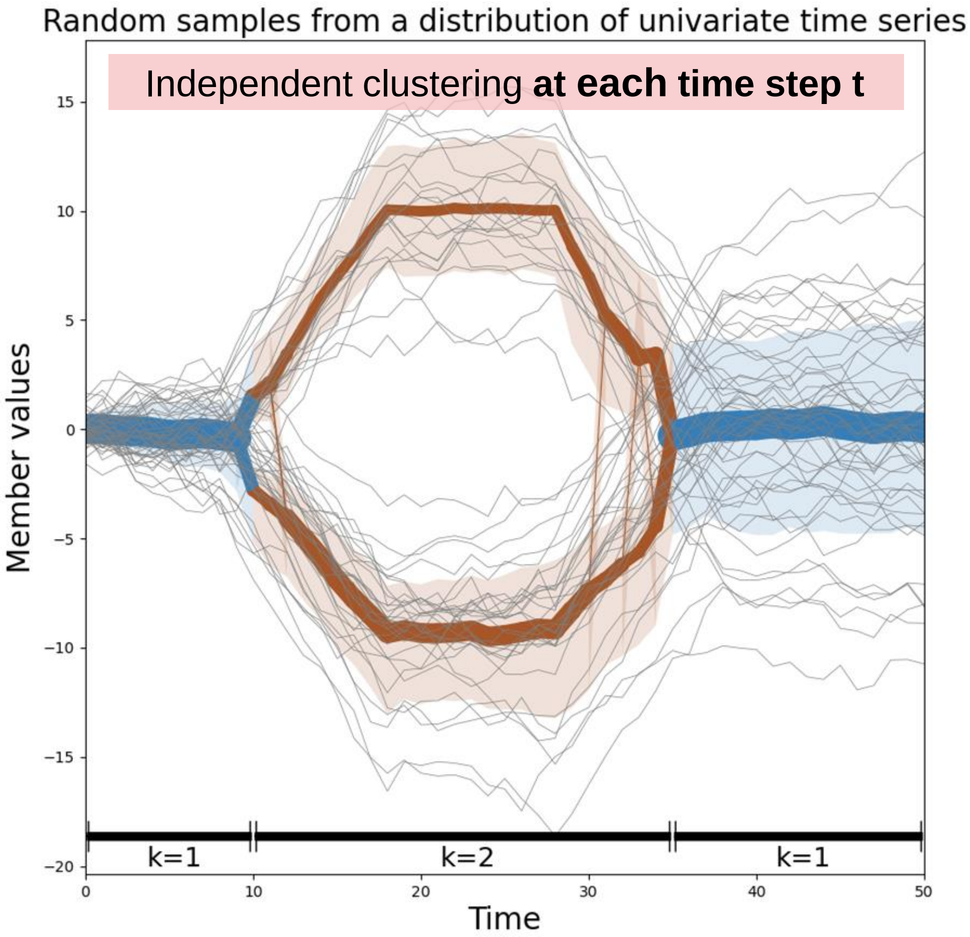
T2. Summarize each mode at each **t**

T3. Determine when modes appear and/or disappear

T4. Determine all possible connections between modes

Consequence:

We independently cluster the members at each time step



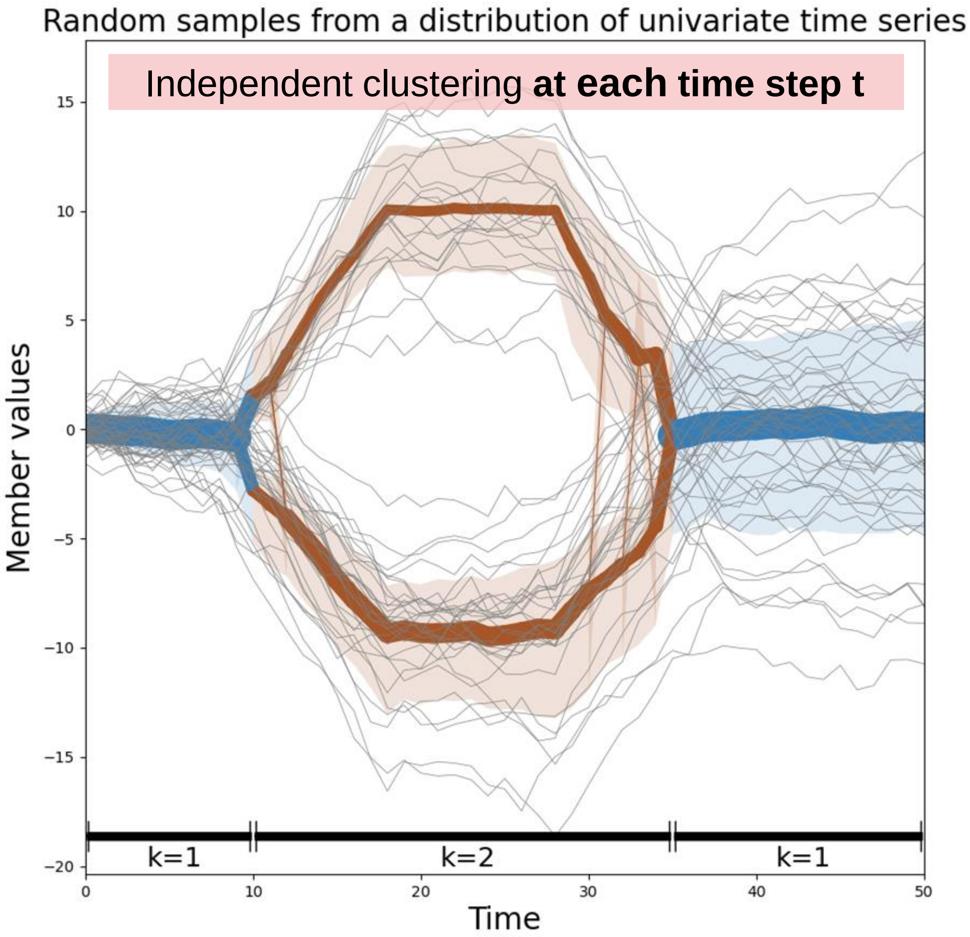
Design rationales:

1. The distribution can have any shape
→Number of modes **k** and type of distribution unknown

2. All ensemble members matter
→Outliers should not be ignored

3. All scales matter
→Avoid thresholds

4. Not a black box
→The process of determining **k**,
→its uncertainty
→and the consequences of choosing a particular **k** should remain transparent



Graph construction

For **each time step t:**

1) Cluster the data **assuming successively**

2)k=1, 2, 3, .., N

3) Associate **each assumption k** with a **score**

4) Define a **life span** for **each assumption k**

5) For each cluster, create a vertex

6) Create edges between vertices

Graph construction

For **each time step t:**

1) Cluster the data **assuming successively**

2)k=1, 2, 3, .., N

3) Associate **each assumption k** with a **score**

4) Define a **life span** for **each assumption k**

5) For each cluster, create a vertex

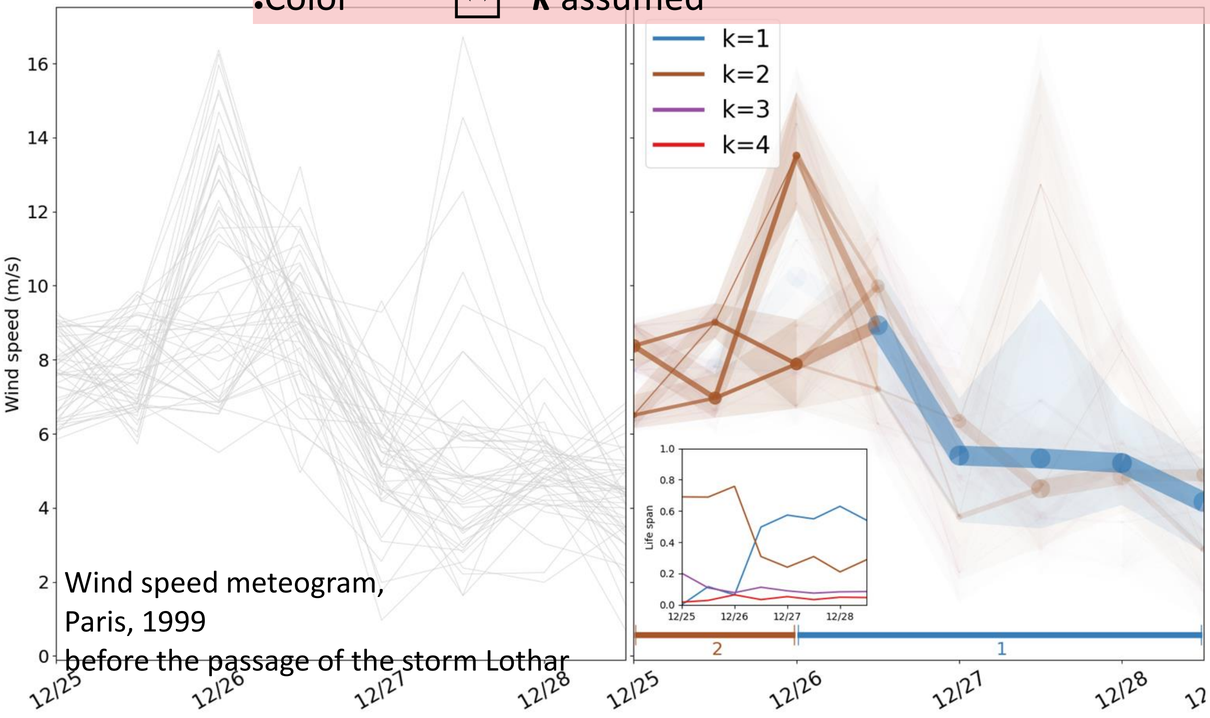
6) Create edges between vertices

A proxy for their relevance

Entire graph view:

Concurrently displays the clustering outcome **for all *k***

- Opacity ↔ Life span
- Color ↔ ***k*** assumed



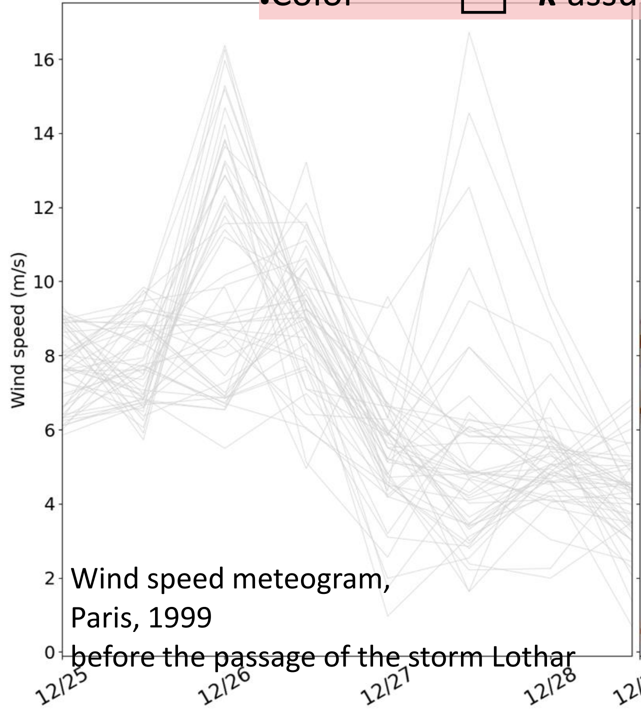
Wind speed meteogram,
Paris, 1999
before the passage of the storm Lothar

Effortless exploration & exposition of a threatening scenario

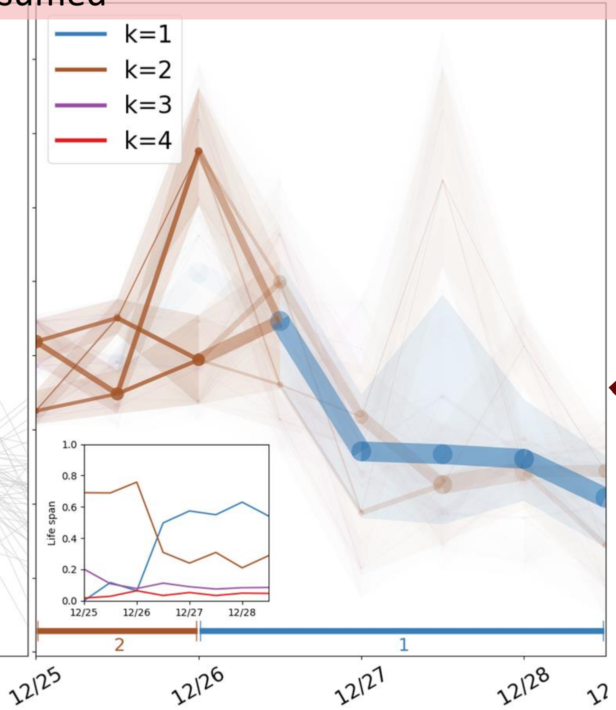
Entire graph view:

Concurrently displays the clustering outcome **for all *k***

- Opacity ↔ Life span
- Color ↔ *k* assumed



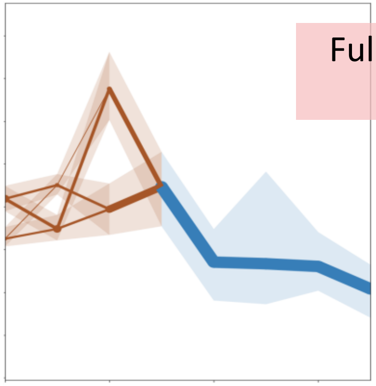
Wind speed meteogram,
Paris, 1999
before the passage of the storm Lothar



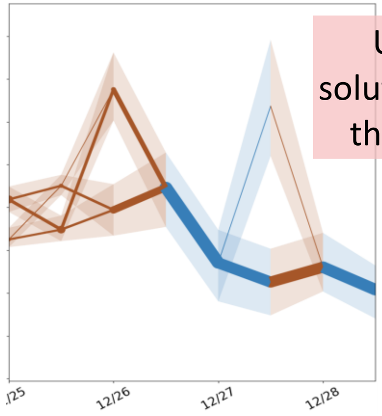
Most relevant components view

Most relevant interpretation

(**automatic or user-defined**)

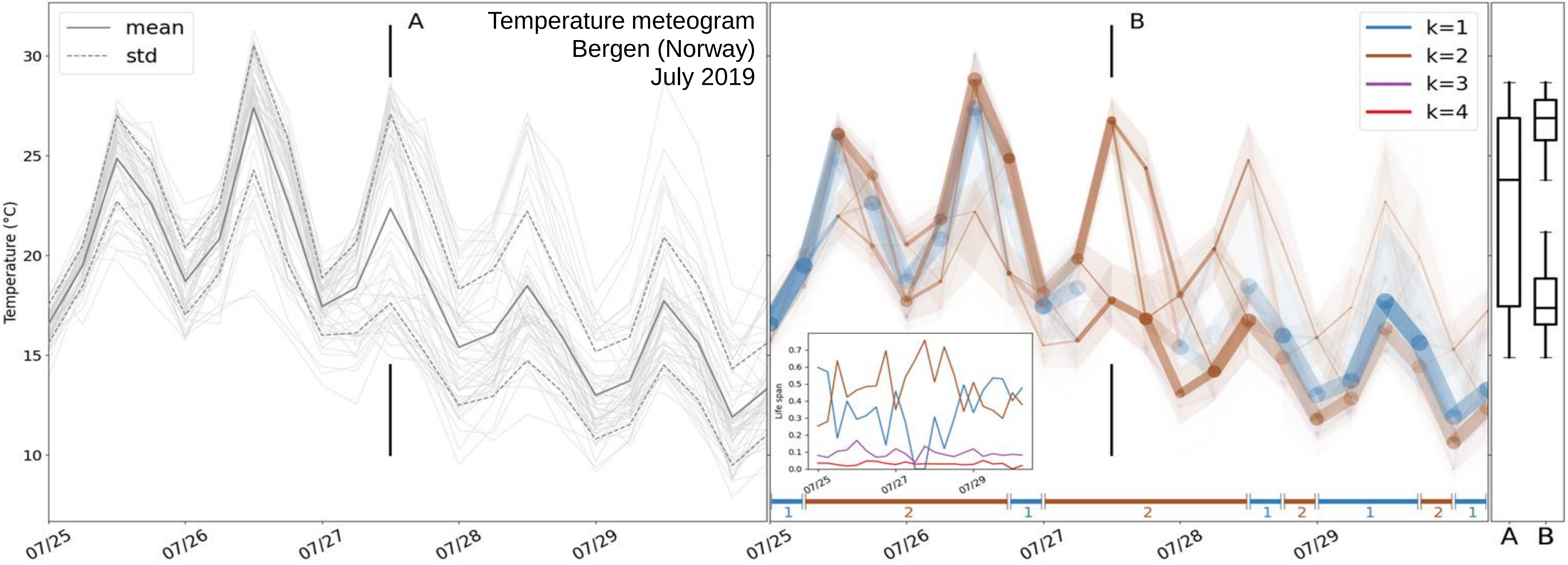


Fully automated solution



User-adjusted solution emphasizing the potential risk

Effortless exploration & exposition of a threatening scenario



On July 27 at noon

Mean and standard deviation:

"22.3 ± 4.7°C"

Our method:

"51% probability of 18.0 ± 1.6°C and 49% prob of 26.2 ± 1.0°C"

On July 28 at noon

Mean and standard deviation:

Upper bound = 22.8°C... but 20% of the members are > 22.8!

Ensemble weather prediction:

- Complex and chaotic system
- Prone to multimodality
- Large and small scales co-exist
- Of great socio-economic importance

Our method:

- Reveals multimodality & its uncertainty
- Aids the understanding of ensemble weather prediction
- Provide fully automated solutions
- Provide quantitative and qualitative information