EuroVis 2017
Machine Learning Methods in Visualization for Big Data 2017
12 June 2017, Barcelona, Spain

Dynamic Data:

Modelling Time Series

Ian Nabney[1]

1 Aston University

Outline

- Nature of time series data
- Capturing temporal dependencies: GTM through time
- Case study: condition monitoring of helicopter airframes

Nature of Time Series Data

- Measurements are taken repeatedly from a single system.
- In most data, there are dependencies between one time step and the next.
- Many different types of behaviour: stationary/non-stationary, seasonality, trends, regimes.

Capturing Dependencies

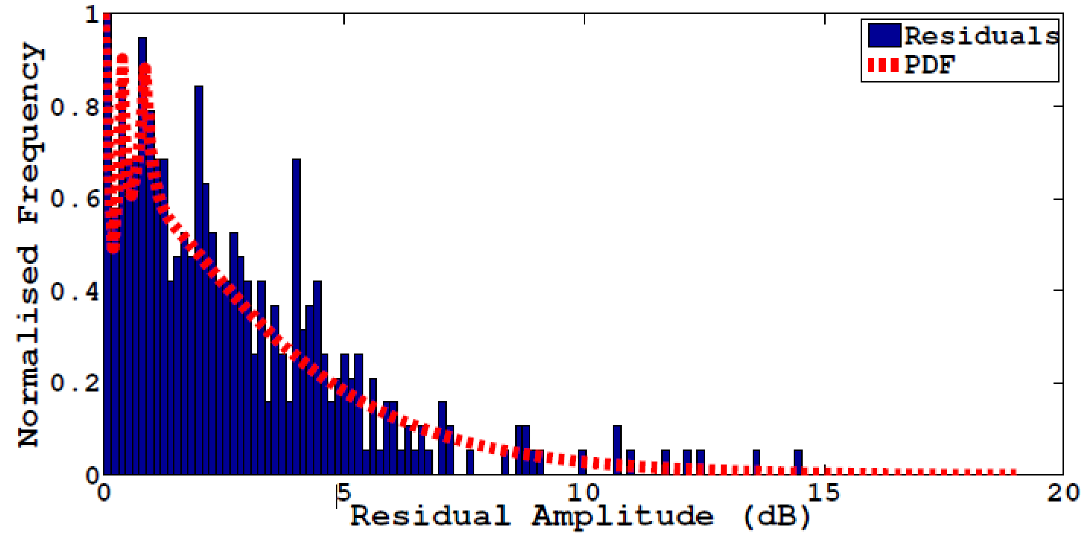
- Simplest way to capture dependencies is through signal processing – features that model dynamics.
  - Autoregressive model parameters
  - Fourier Transforms
  - Wavelets
- Then apply visualisation in the feature space.
- Alternative is to build dynamics into the visualisation model.

SONAR Data

- Sensor array consisted of 32 hydrophones configured in an approximately linear array with a target ship transmitting a signal whilst traversing the length of the array.
- Shallow water scenario with low levels of rain and thermal noise expected.
- Frequency response of the hydrophones was 124 – 249Hz.
- The sensor array across all beams can be considered either as a set of one-dimensional time series or groups of vector time series processes.
- Overlapping groups of 5 beams via a shifting window, i.e. for the 32 beams analysed there are 28 groups.

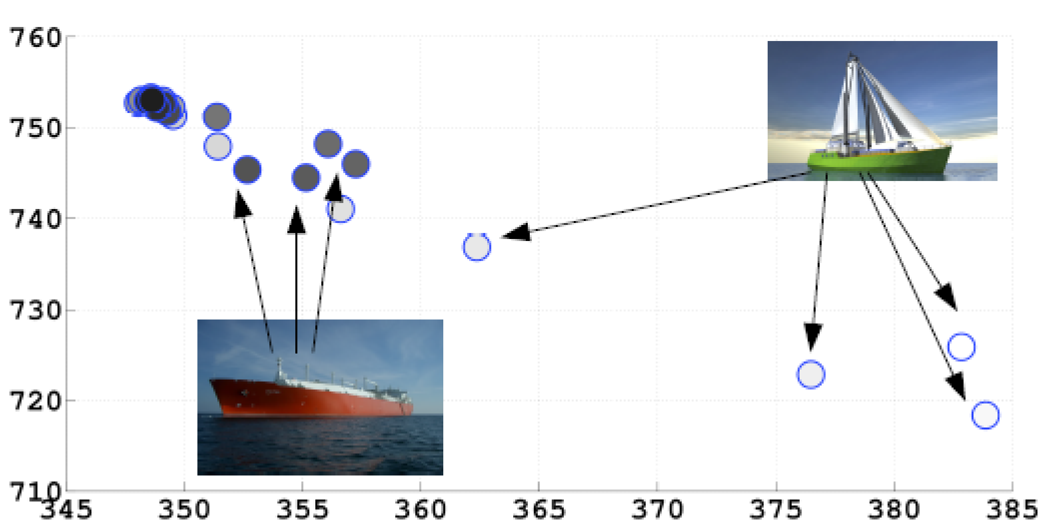
Visual Analytics of SONAR

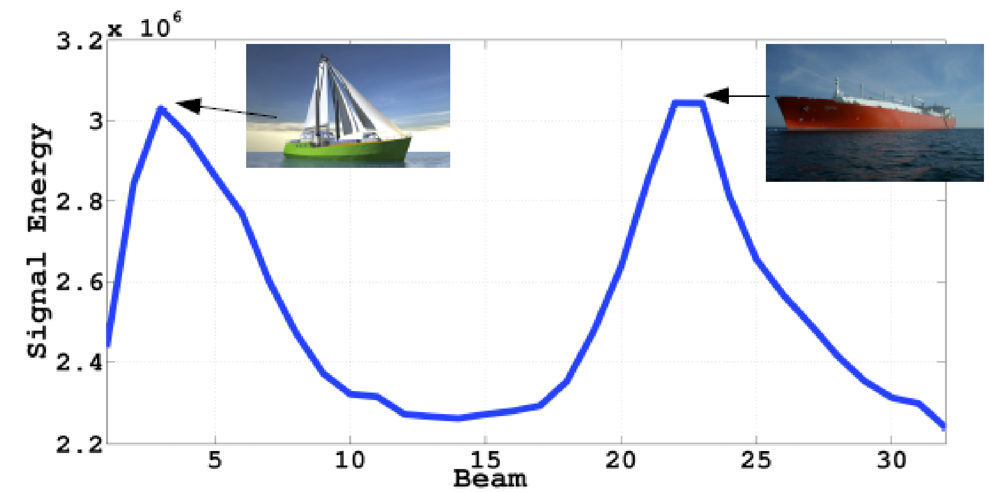
- $$s_b = u_b + \epsilon_b$$
- Signal is target tonals plus noise
  - Fit a non-linear autoregressive model (RBF) with mixture of noise processes and use KL-divergence for dissimilarity:
    - Residual signal characteristics: Laplace distribution,
    - Thermal noise characteristics: Rayleigh distribution,
    - Surface wave scatter: K distribution (two-gamma compound),
    - Rain characteristics: Gamma distribution,
    - Miscellaneous characteristics: Normal distribution.



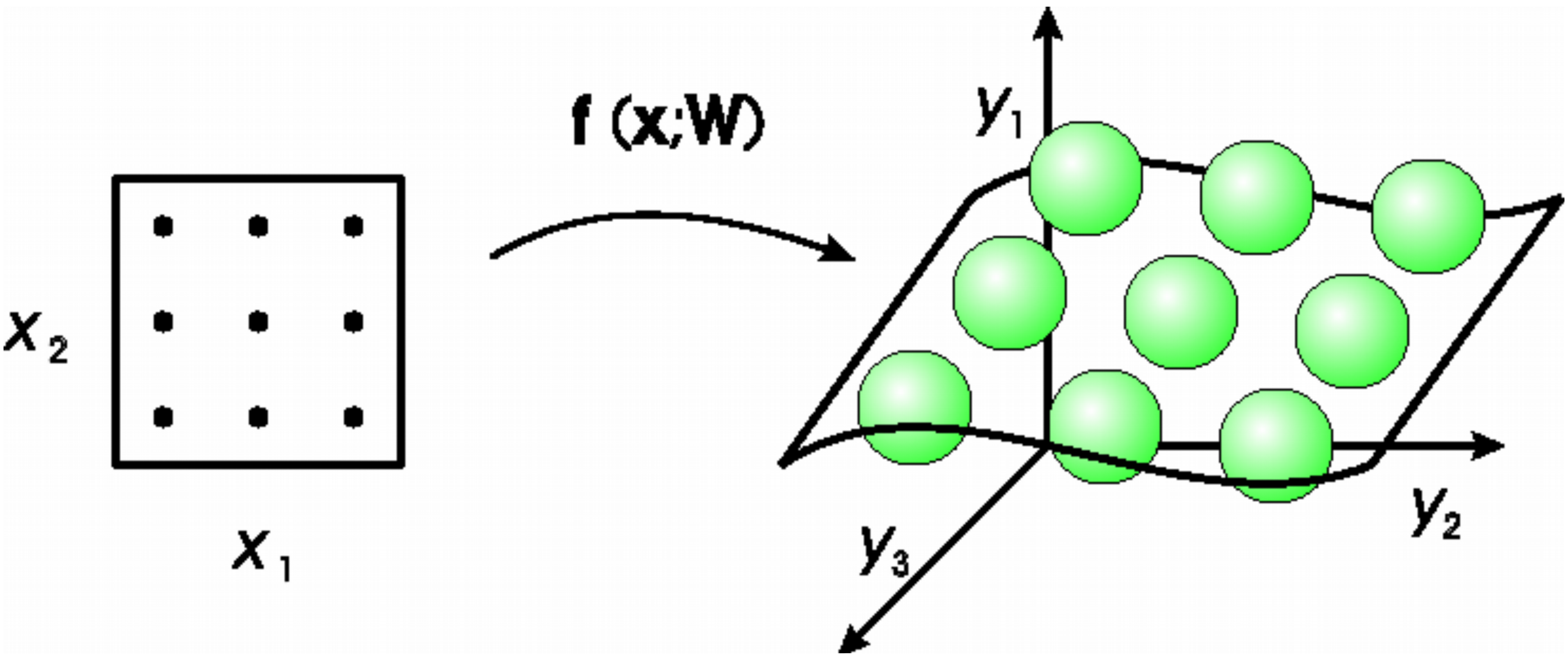
Beam Grouping

- Visualisation shows the relative similarity of the groups of beams where each point in the plot corresponds to one of the beam groups.
- Cluster of beams containing only noise and a separate group of 'outlier' beams which contain target signal elements.
- Shading of beams is white to black ordered from group 1 (beams 1-5) to 28 (beams 28-32) respectively. Targets are present in beams 1-6 and 21-25, which appear as anomalies.



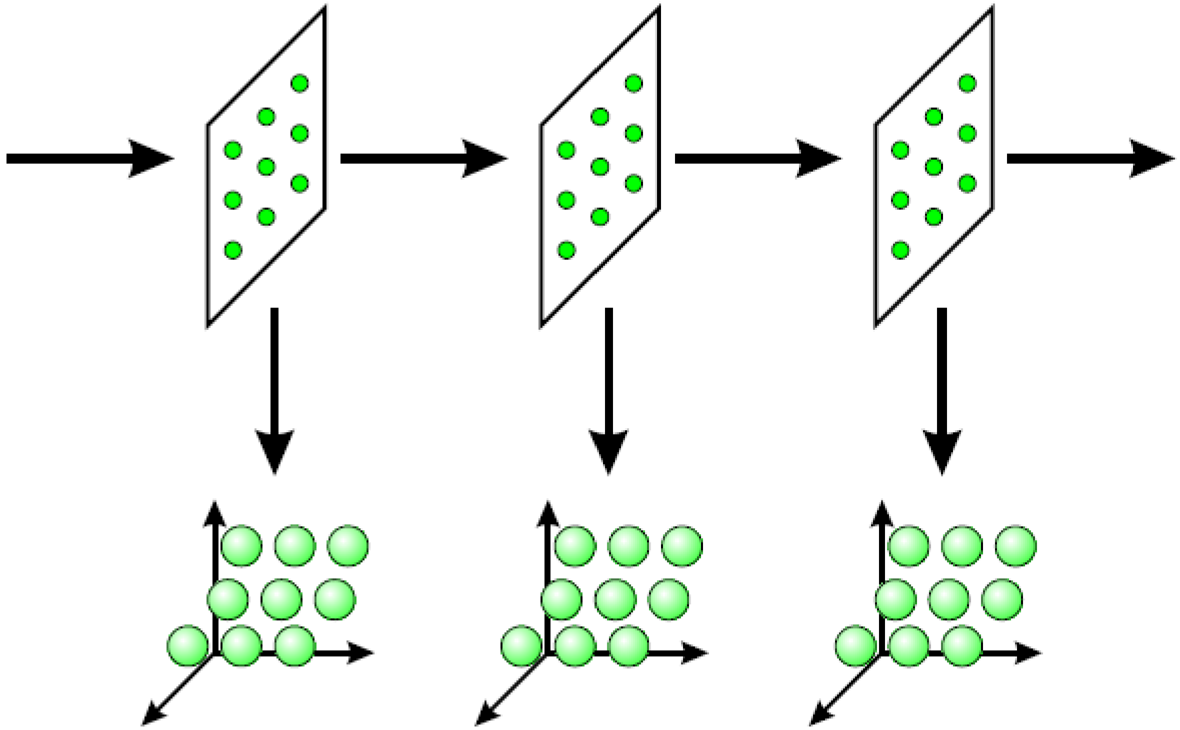


GTM through Time

- Original GTM assumes that all data points are independently generated.
- 

GTM through Time

- The GTMTT consists of a hidden Markov model in which the hidden states are given by the latent points of the GTM model, and the emission probabilities are governed by the GTM mixture distribution.
- A topologically-constrained HMM, as GTM is a topologically-constrained GMM.
- The parameters of the GTM model, as well as the transition probabilities between states, are tied to common values across all time steps.

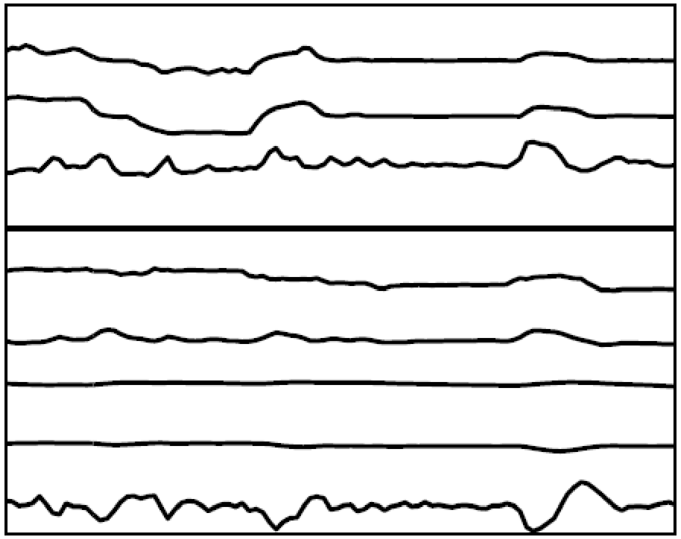


GTMTT practicalities

- If we allow a fully connected matrix of independent transition probabilities connecting every state at time n to every state at time n + 1, then the number of independent parameters would be prohibitively large.
- If we have, for example, 100 hidden states in the GTM model then we would have 10^4 independent transition probability parameters to be determined.
- In many applications we expect different regions of the latent space to correspond to different regimes. We also expect smooth changes in latent space within a regime and relatively rare jumps to other regimes. An approximate way to capture this prior knowledge is to allow groups of transitions to be governed by a common parameter. More recently, Bayesian methods with a GP prior have been used.
- Both GTM and HMMs are trained using an EM algorithm, so (with some work) there is an EM algorithm for GTMTT.

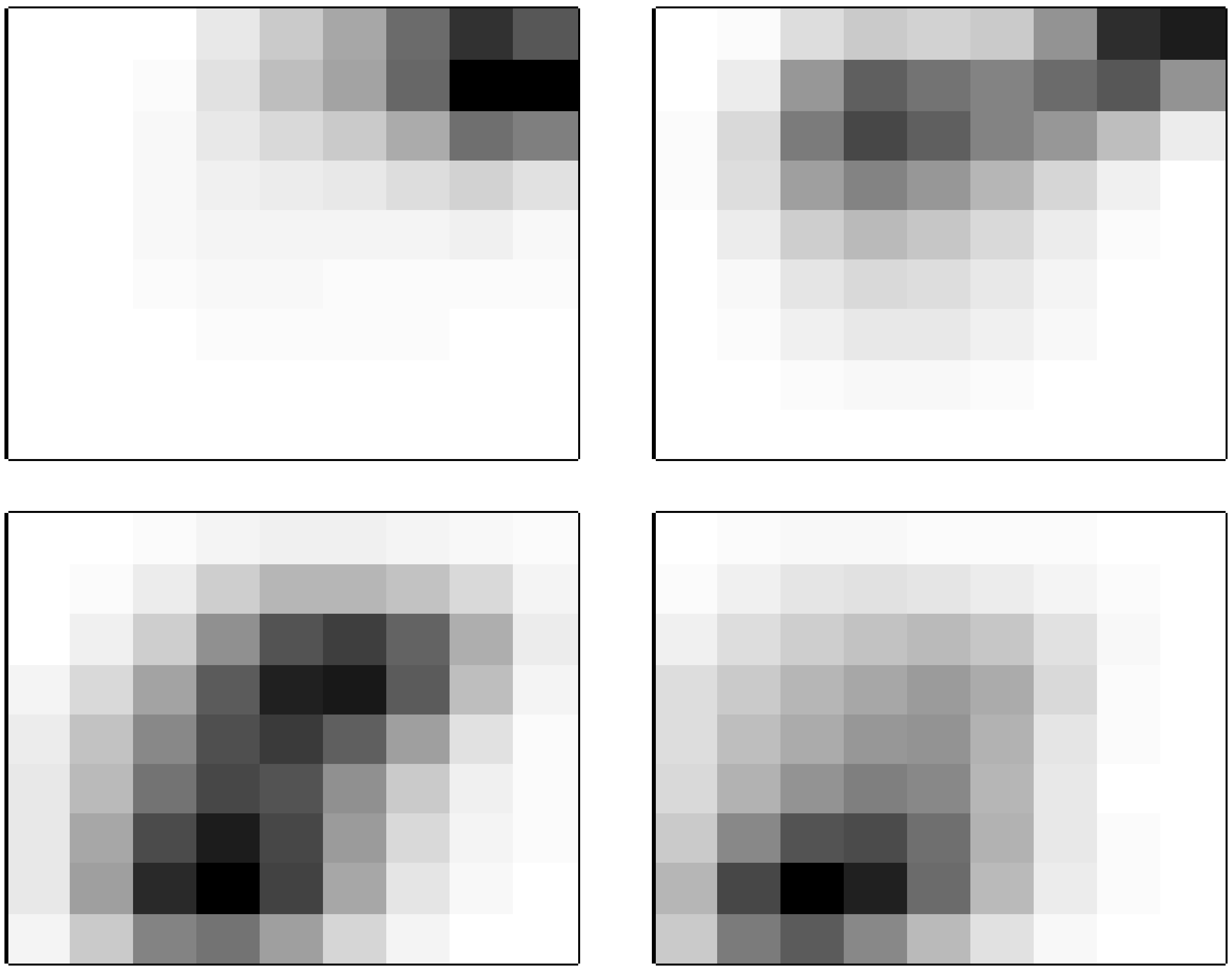
(Simple) Helicopter Example

- 9 variables (sampled every two seconds) measuring quantities such as acceleration, rate of change of heading, speed, altitude and engine torque.
- GTM with a 15x15 grid in latent space. For each latent state *i*, the transition probabilities to states at the next time step are collected into 10 separate groups, in which 9 of the groups correspond to those states *j* which are within 1 unit from state *i*, while the 10th group consists of all other states *j*.
- In the trained model, different regions of the latent space will correspond to different flight regimes.



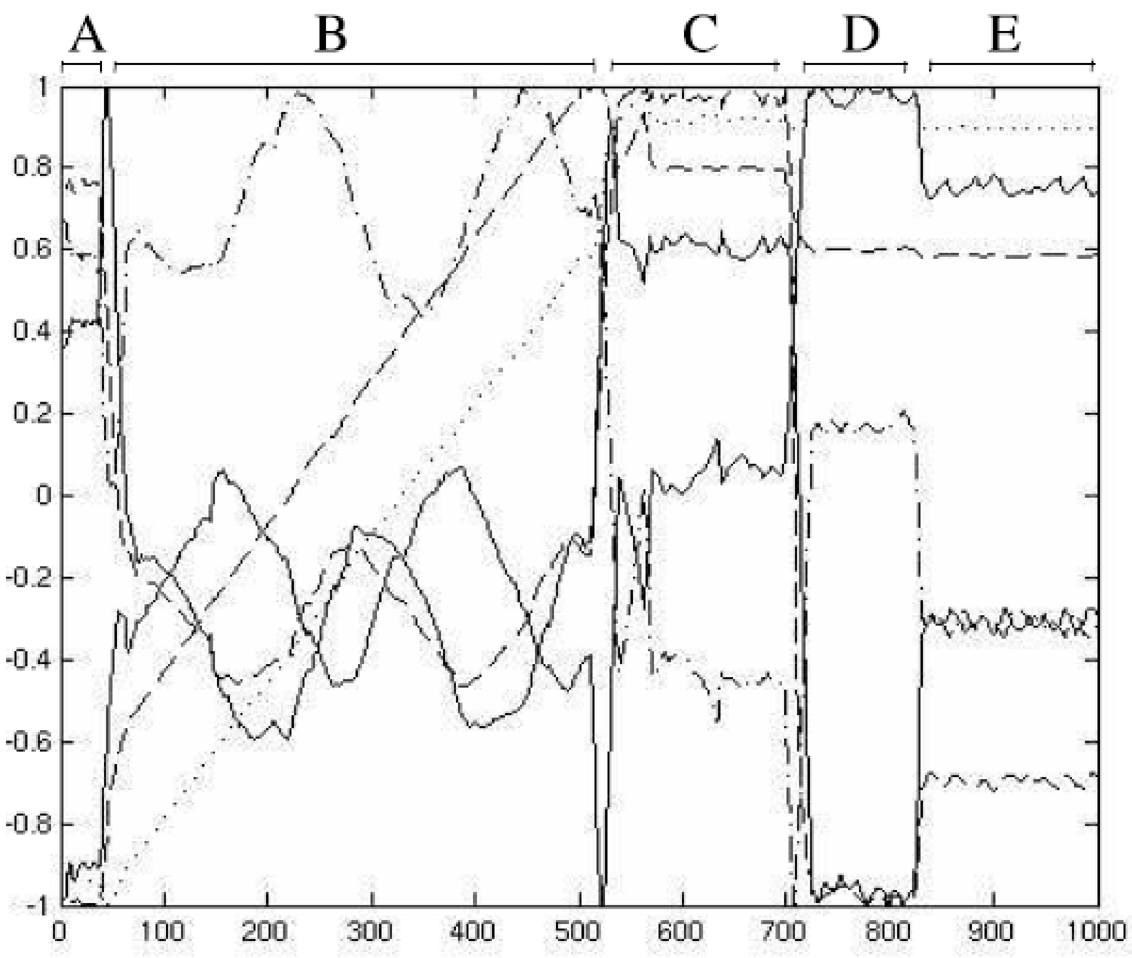
GTMTT: Helicopter

- Plots of posterior distribution in latent space at 4 time points.
-



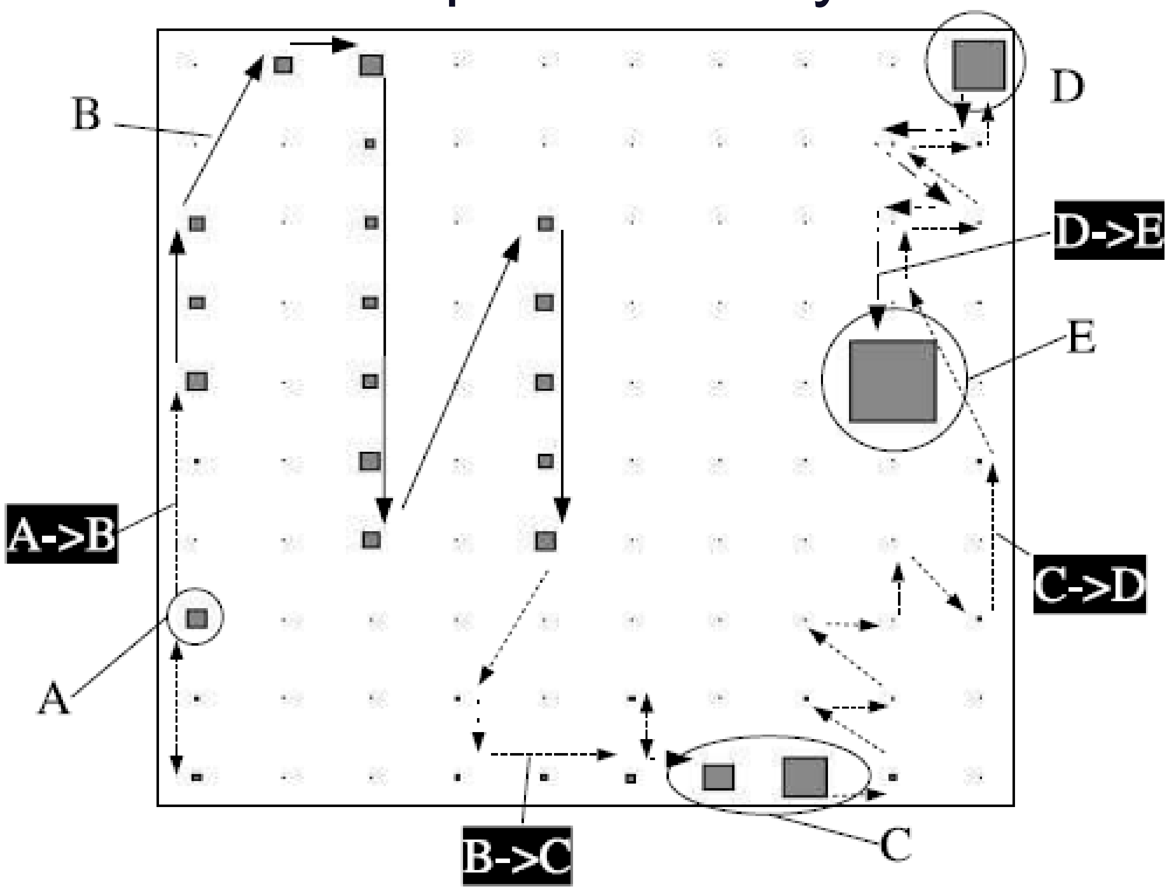
Shuttle Example (Vellido)

- 6-variate time series consists of 1000 data points obtained from various inertial sensors from Space Shuttle mission STS-571.
- Contains sub-sequences of little variability followed by sudden transition periods.



GTMTT Latent Space

- 10x10 latent space.
- Latent states representing low-variability periods are circled, and sudden transition intervals are represented by discontinuous oriented lines.
- The state transitions of period B are represented by a continuous oriented line.

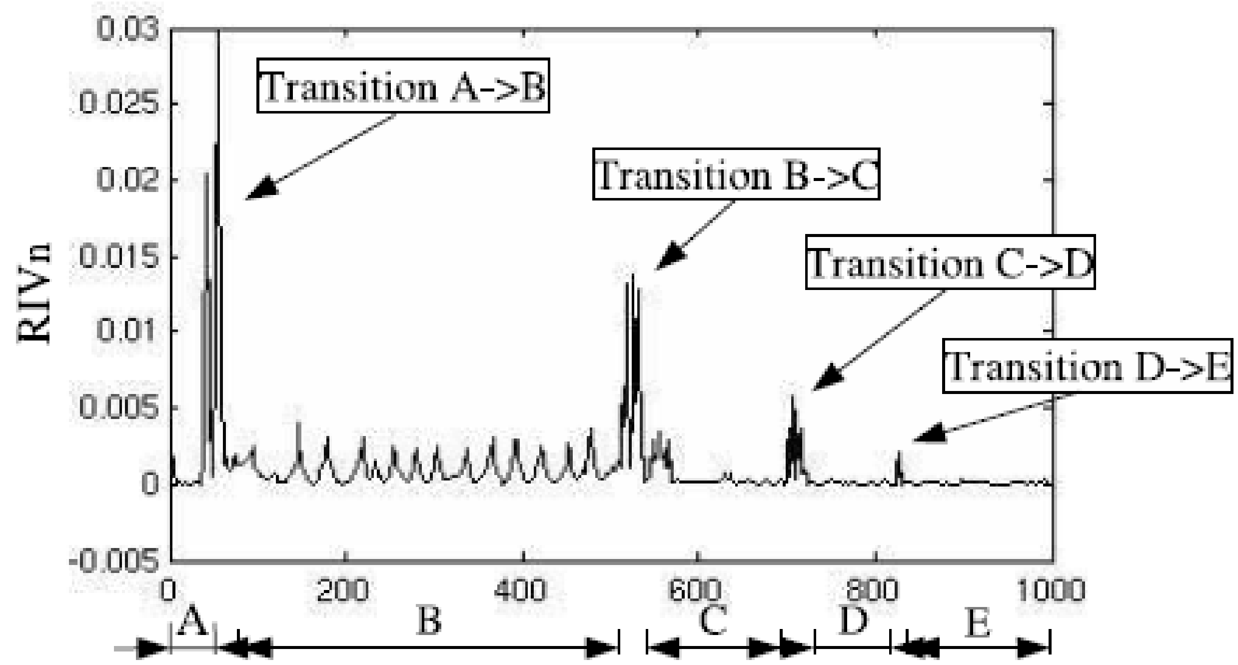


Measuring transitions

-
-

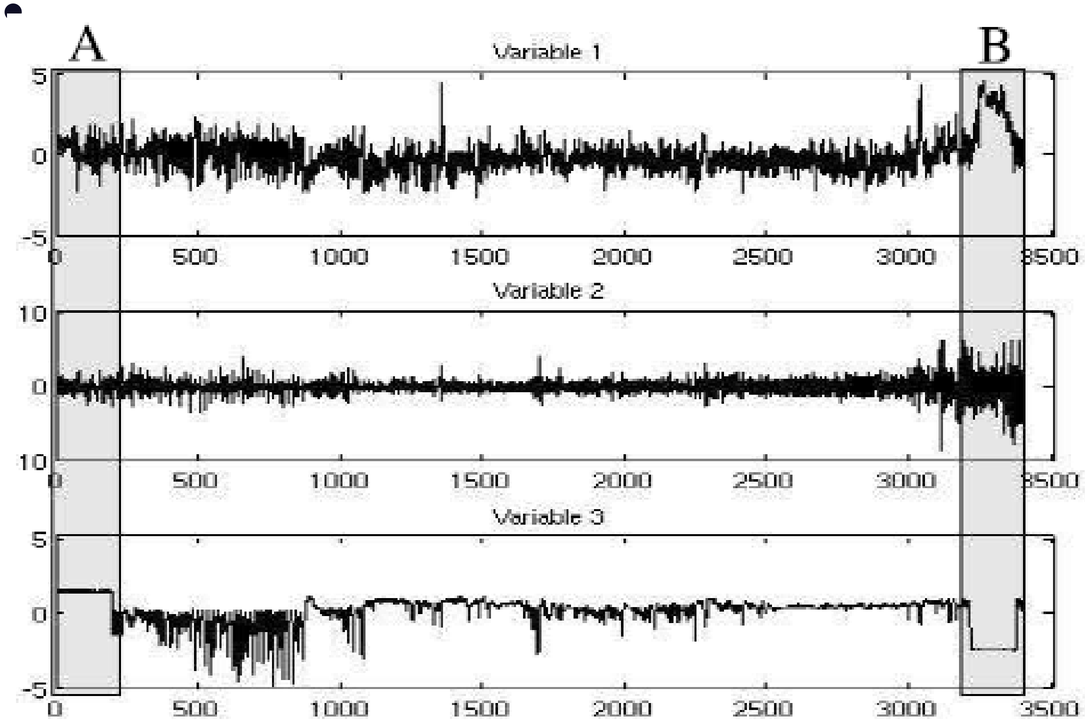
$$RIV_n = \frac{-\{\log P(X_n) - \log P(X_{n-1})\}}{\sum_n \{\log P(X_n) - \log P(X_{n-1})\}}$$

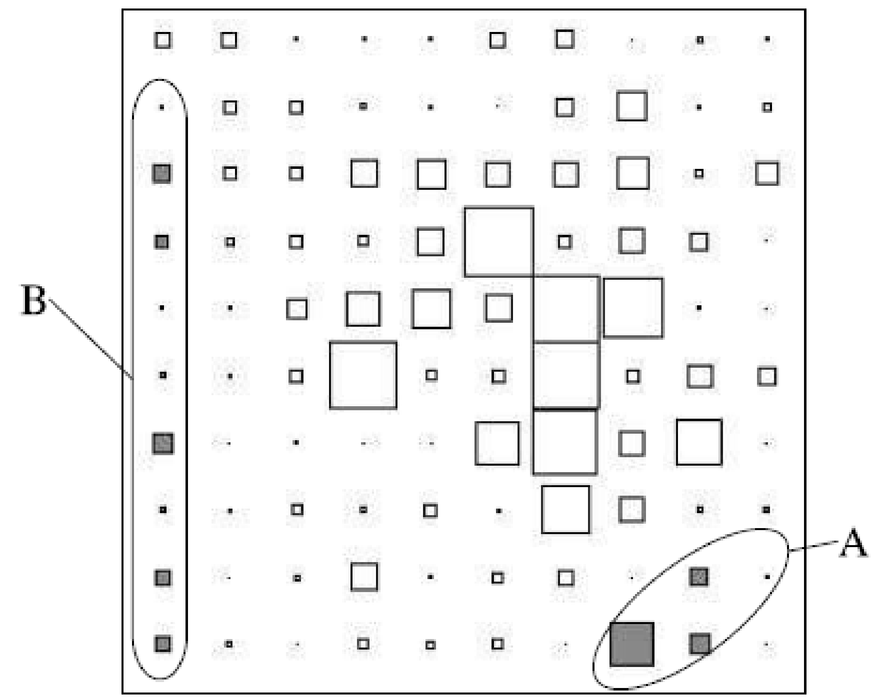
- Suddenness of transitions is proportional to RIV_n (based on sub-sequences X_n. Lower bound of zero.
- Interval B is represented by a sequence of small fluctuations in RIV_n



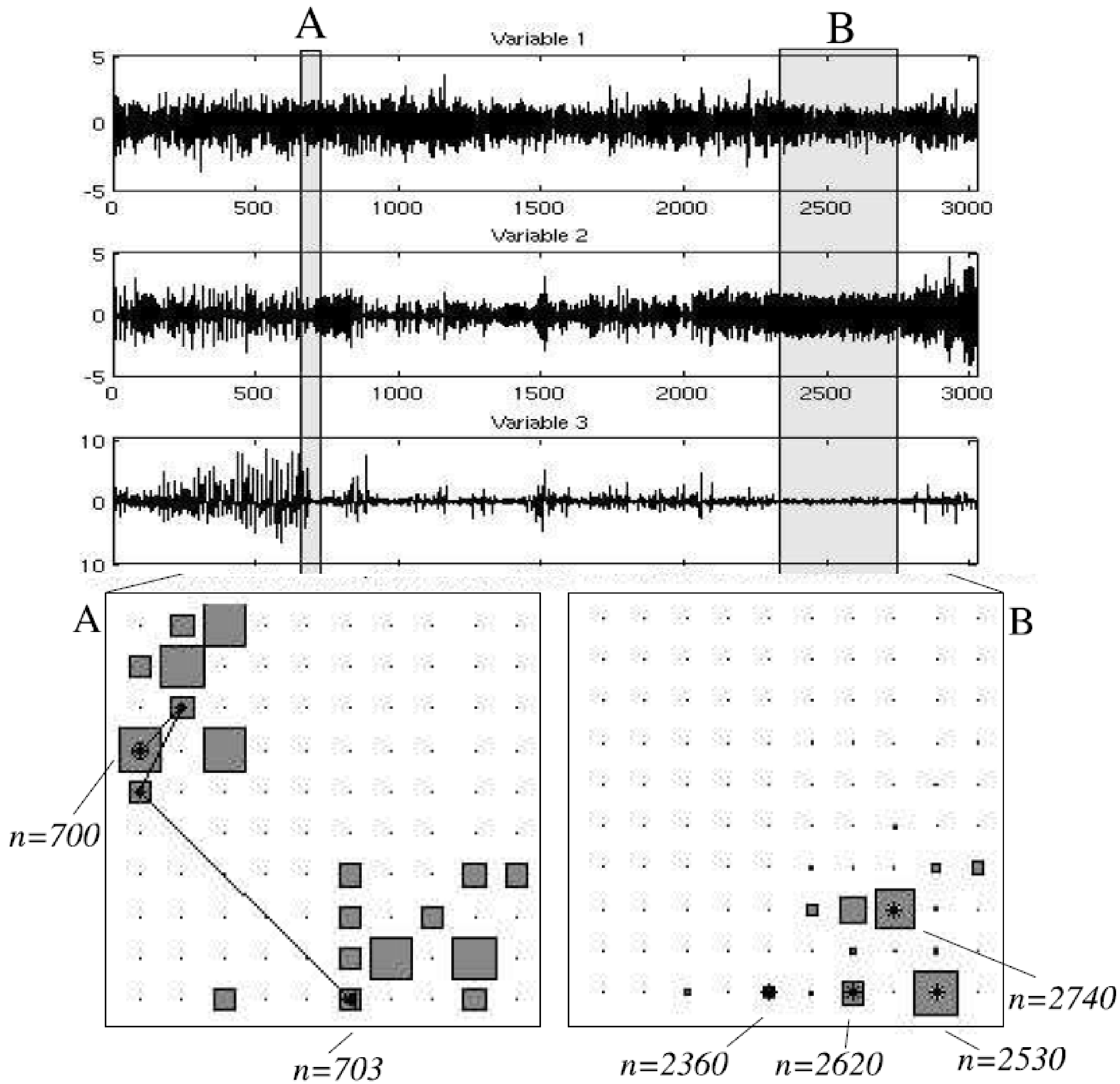
Physiological Data

- 3-variate time series consisting of 3400 samples of physiological data, used in the Santa Fe Competition in 1991. They consist of three physiological variables measured in a subject while sleeping, and contain clearly atypical sub-sequences due to a measurement error (failure in a sensor).
-





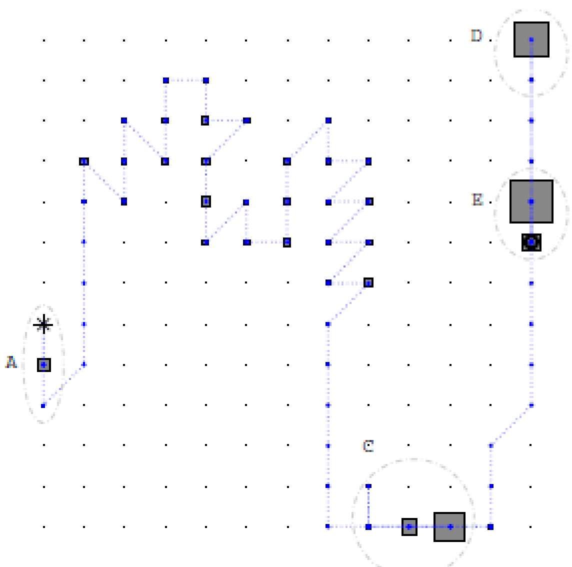
Regime Changes

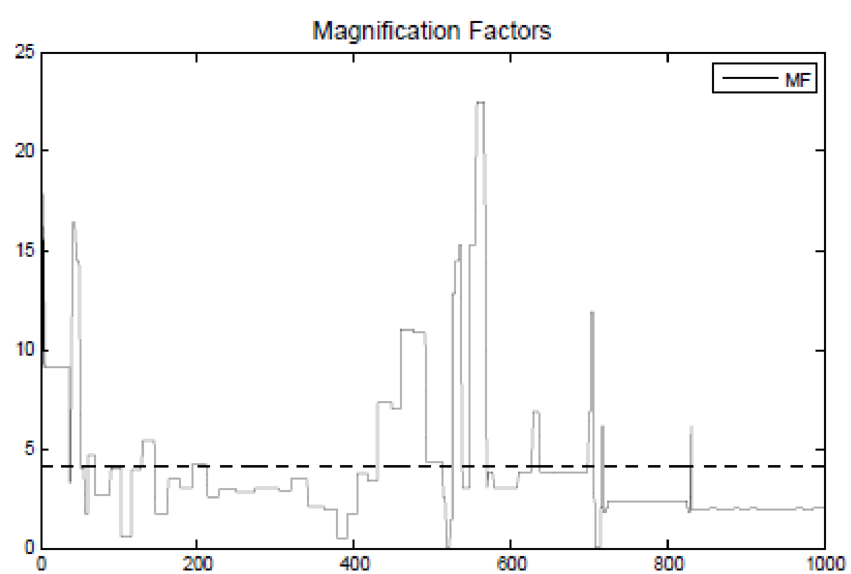


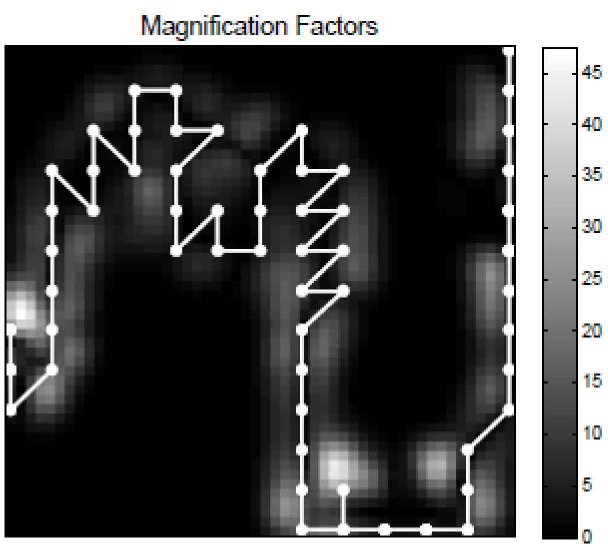
- Atypical sequences removed
- Trends taken out with signal processing.
-

GTMTT Latent Space

- Variational GTMTT.
- Square size proportional to time spent in state.
- Peaks of distortion correspond to transitions.



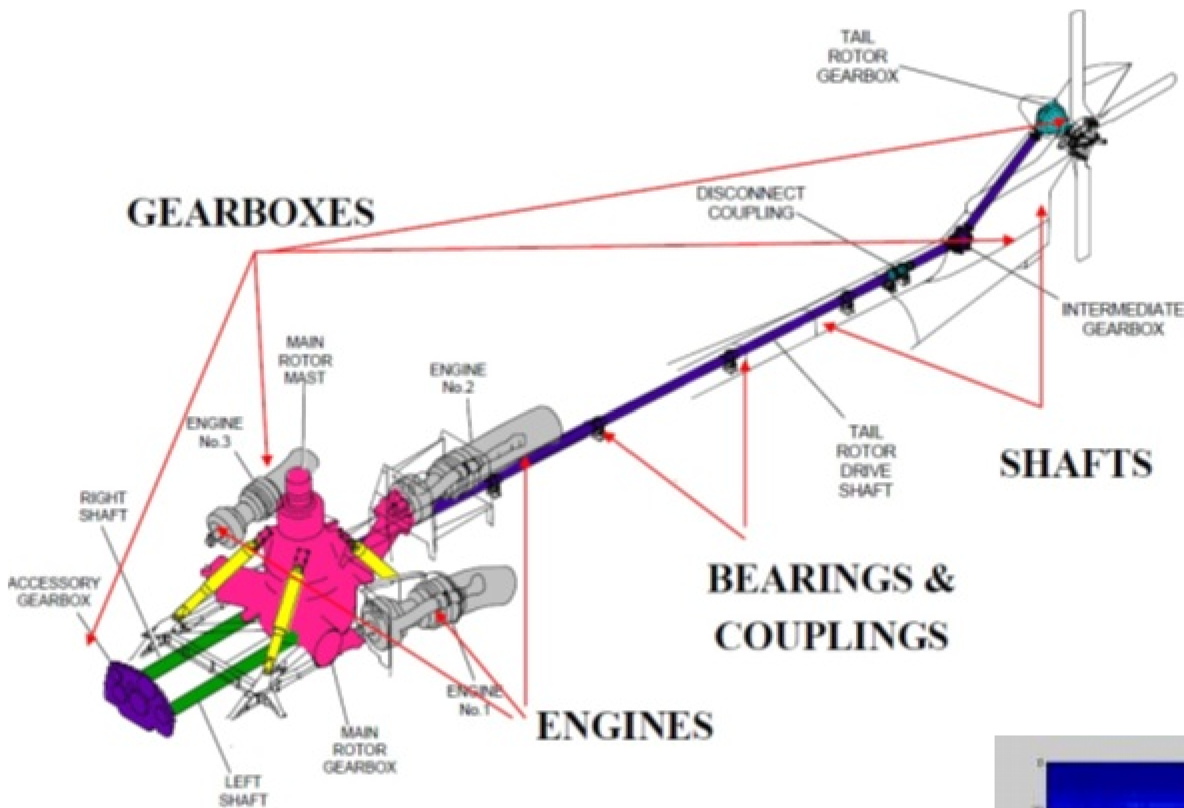




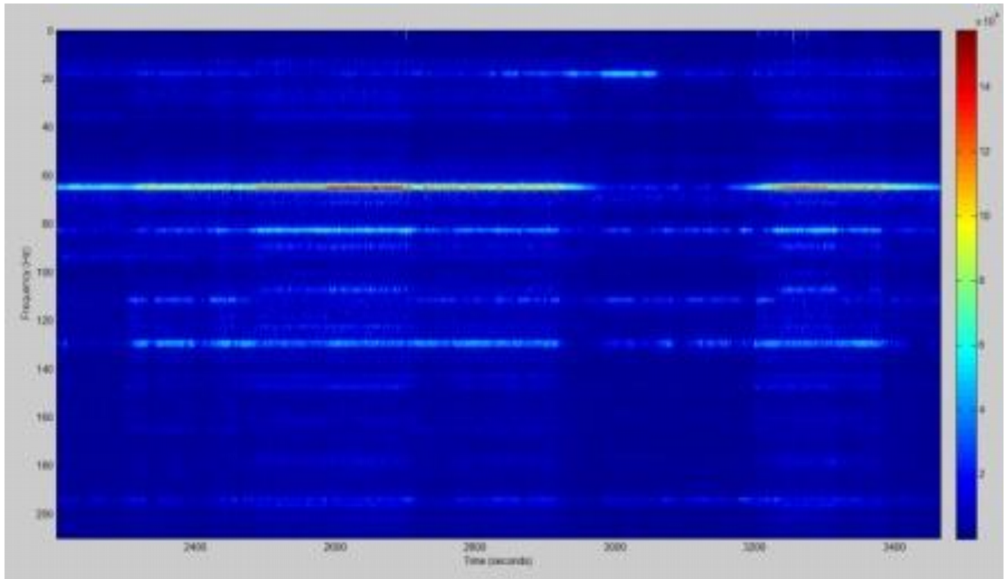
Case Study: Agusta Westland

- AW has pioneered CVM, the continuous recording of airframe vibration (0-200Hz), to improve the investigation of unusual occurrences and monitor airframe integrity.
- Develop a probabilistic framework for inferring flight mode and key parameters from multiple streams of vibration data.
- Improve indicators of airframe condition: the wavelet transform and kernel entropy to assess the dynamics (i.e. non-stationary characteristics) of the vibration signal.
- Integrated diagnosis based on probabilistic models of normality and using a belief network to model prior knowledge about the domain and interactions between key variables.
-

Understanding the Data

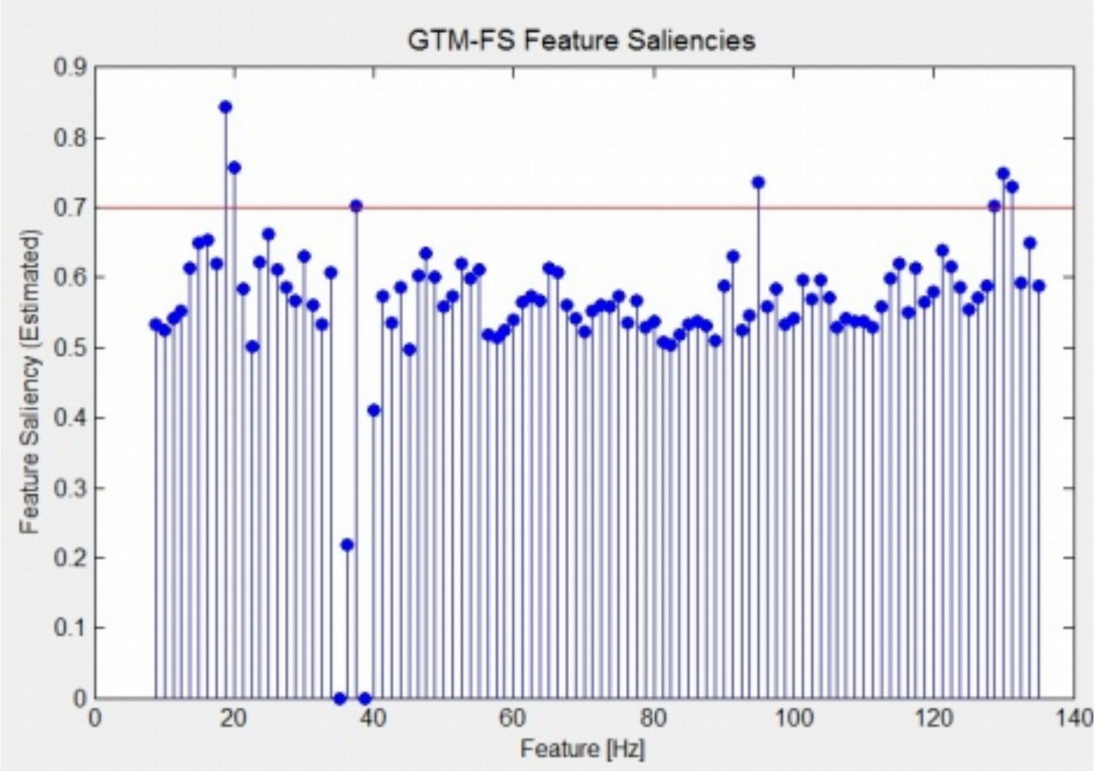


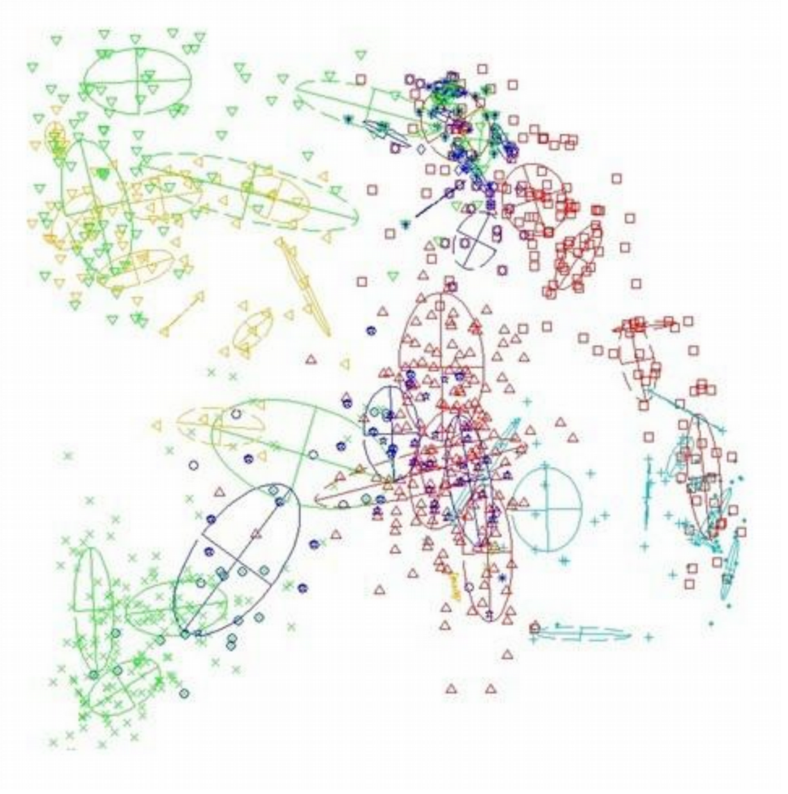
- 8 sensors measuring vibration
- 108 frequency bands (STFFT) for each sensor
- Too much data to build a model from.



Feature Selection

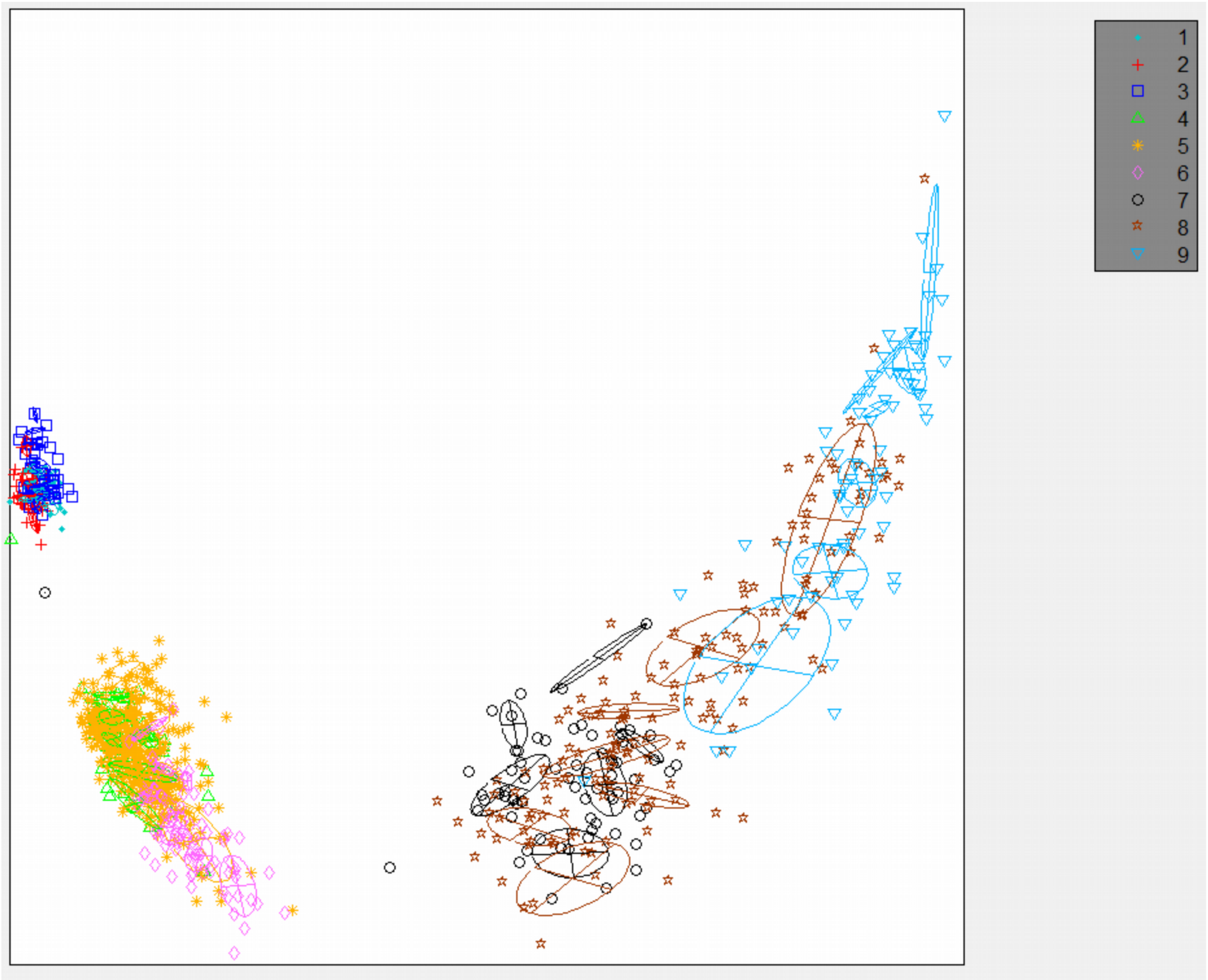
- Features are selected using GTM with Feature Saliencies.
- Sensors are selected by comparing inter-class separation in different plots.





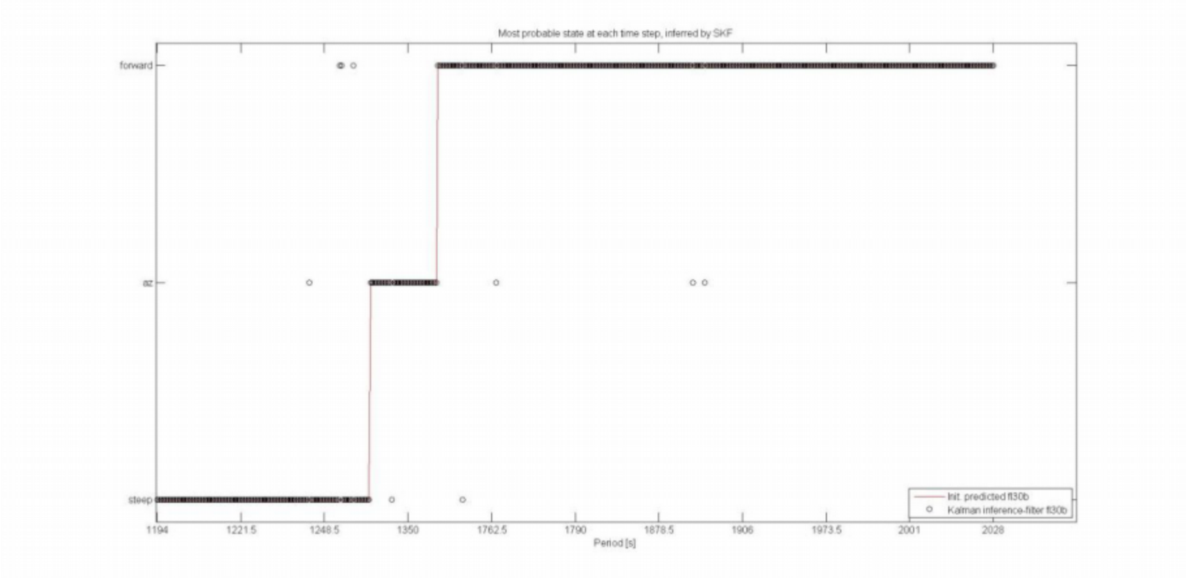
Flying through the Visualisation

- Flight went through a number of different states.
- The sequence can be recaptured from the visualisation

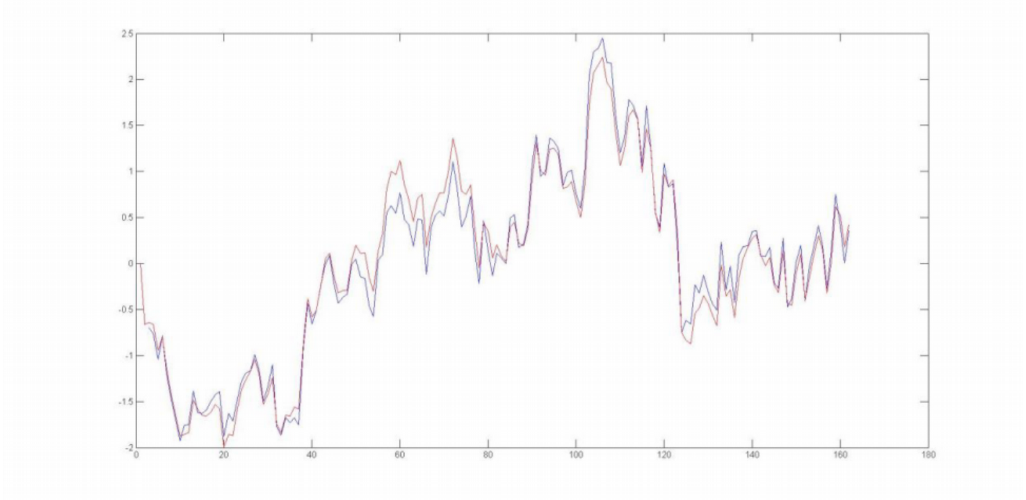


Novelty Detection

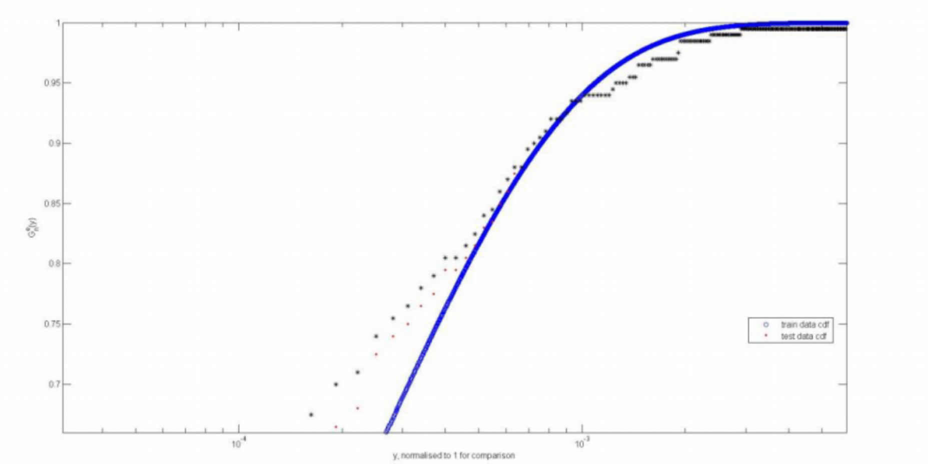
•



Inference of flight state



Non-linear signal prediction



Outlier detection using Extreme Value Theory

Case Study: Agusta Westland

- AW has pioneered CVM, the continuous recording of airframe vibration (0-200Hz), to improve the investigation of unusual occurrences and monitor airframe integrity.
- Develop a probabilistic framework for inferring flight mode and key parameters from multiple streams of vibration data.
- Improve indicators of airframe condition: the wavelet transform and kernel entropy to assess the dynamics (i.e. non-stationary characteristics) of the vibration signal.
- Integrated diagnosis based on probabilistic models of normality and using a belief network to model prior knowledge about the domain and interactions between key variables.
-

Conclusions

- Exploration of Animation, Small Multiples, and Drawing Stability
- Guidelines on their use
- In future work, other definitions for the mental map?