A Bayesian Approach to Regime Assignment

Swinda Falkena, Ted Shepherd, Antje Weisheimer, Jana de Wiljes

EPSRC Centre for Doctoral Training

Mathematics of Planet Earth

Imperial College London

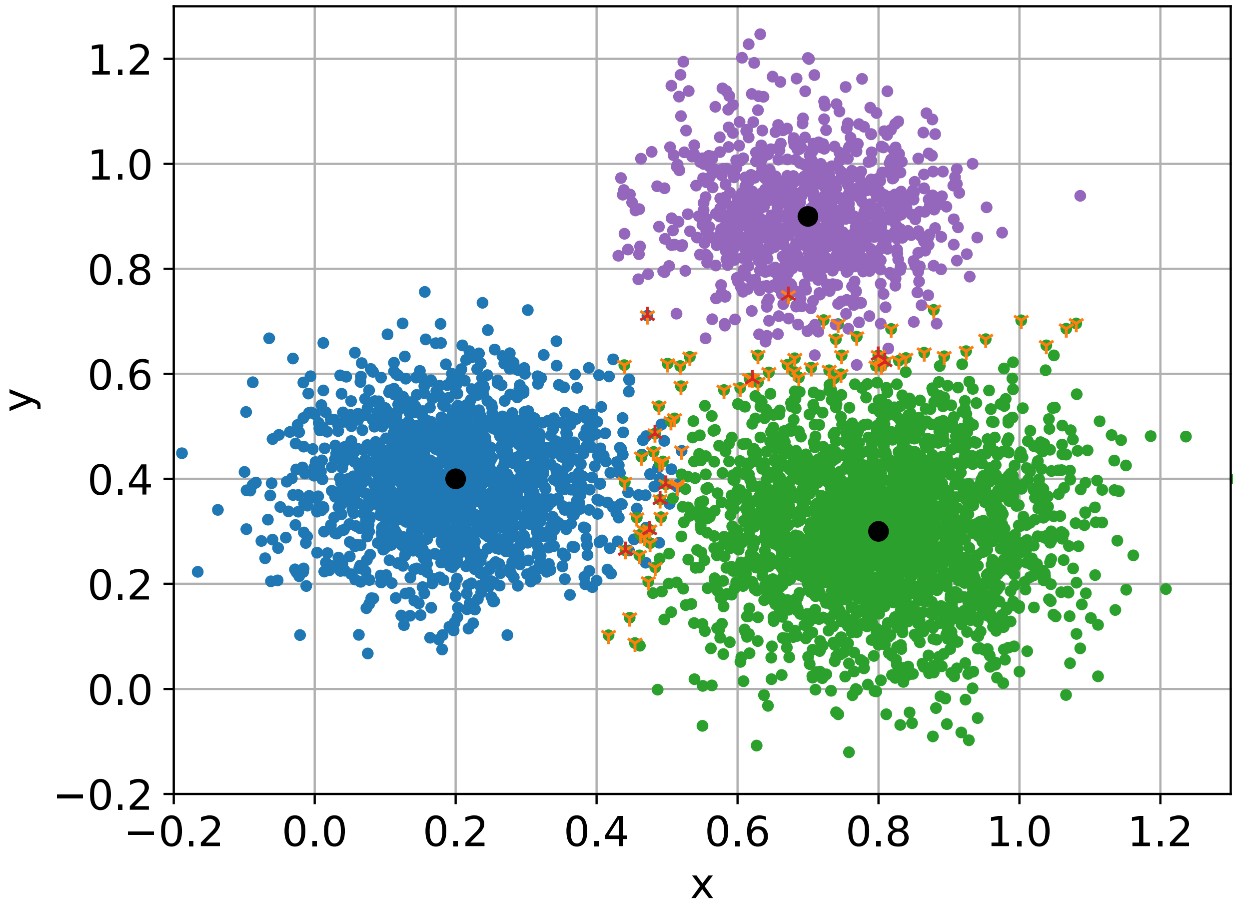
University of Reading

Regimes, or clusters, are subsets of data which are similar within a regime, but different between regimes.

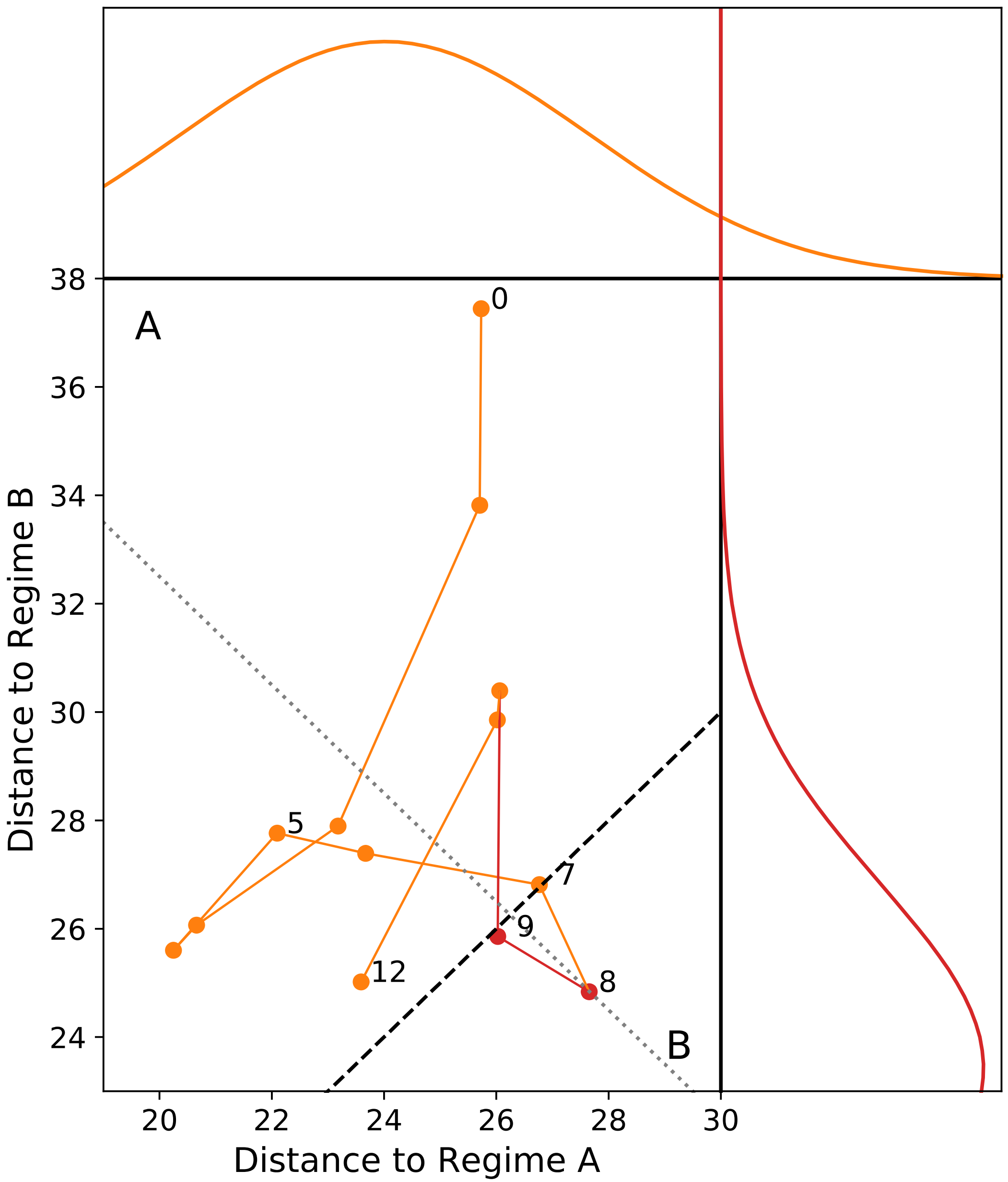
K-means clustering:
Split the data in k clusters or regimes

Drawbacks:

- Every data point has to be assigned to a regime
- Sensitive to small perturbations



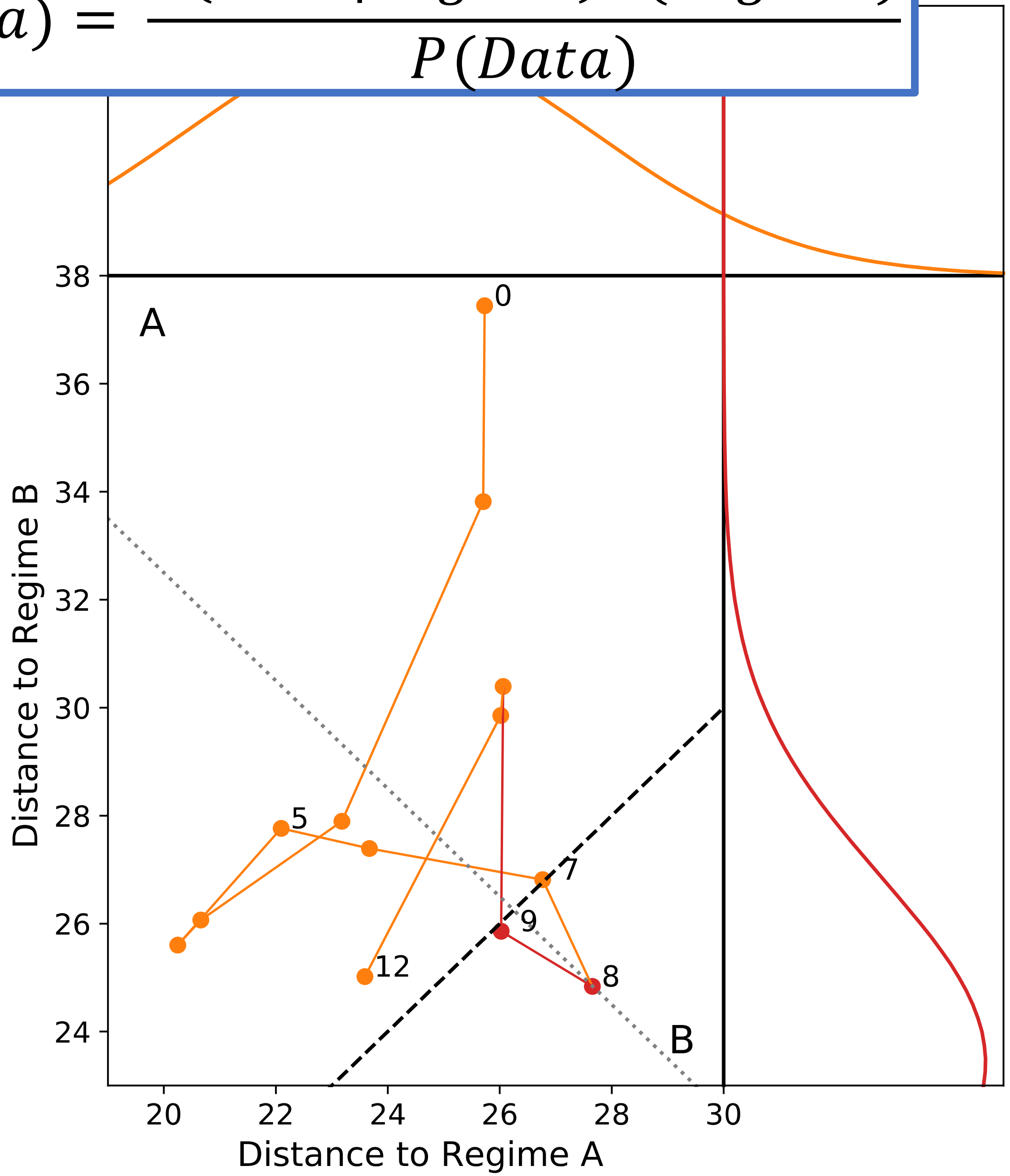
Noise can affect the distance to the regimes and this way alter the regime assignment.

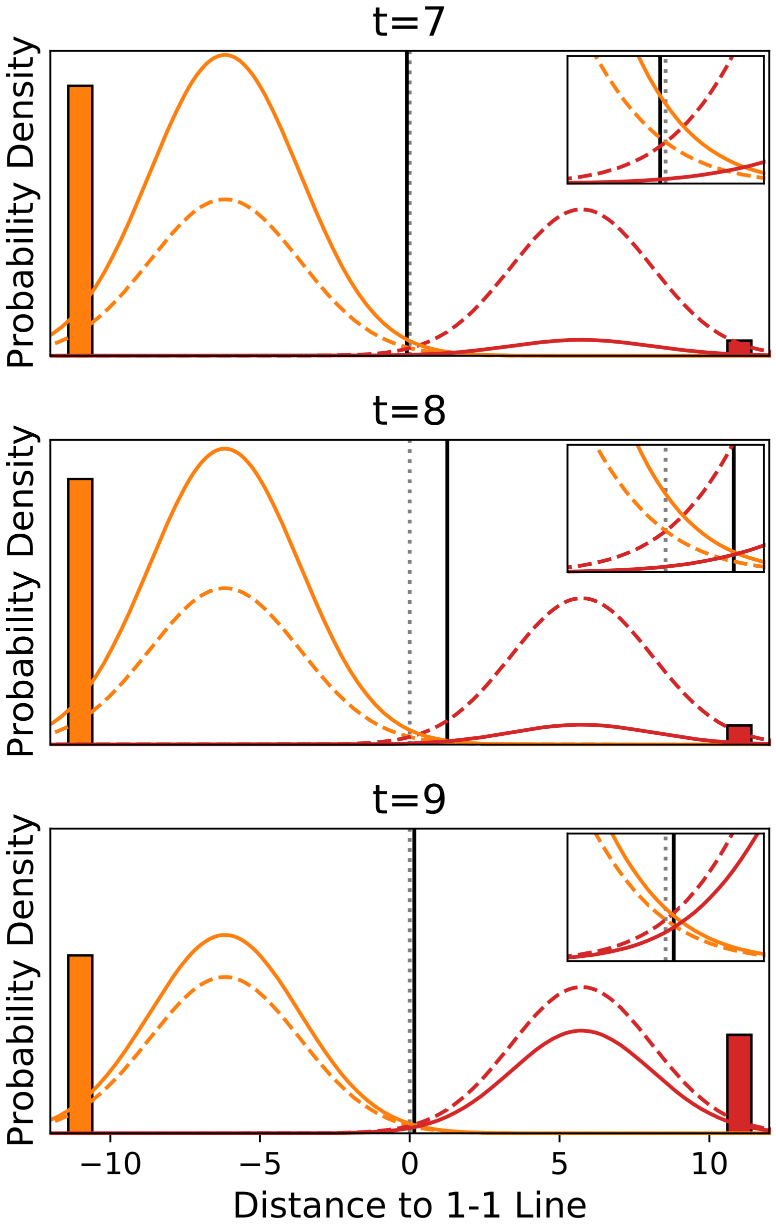


Probabilistic regime assignment

$$P(Regime|Data) = \frac{P(Data|Regime)P(Regime)}{P(Data)}$$

We use Bayes Theorem to obtain a probabilistic regime assignment.





Apply this approach to atmospheric circulation regimes.

Recurrent and persistent patterns

Low-frequency variability

Wintertime Euro-Atlantic sector

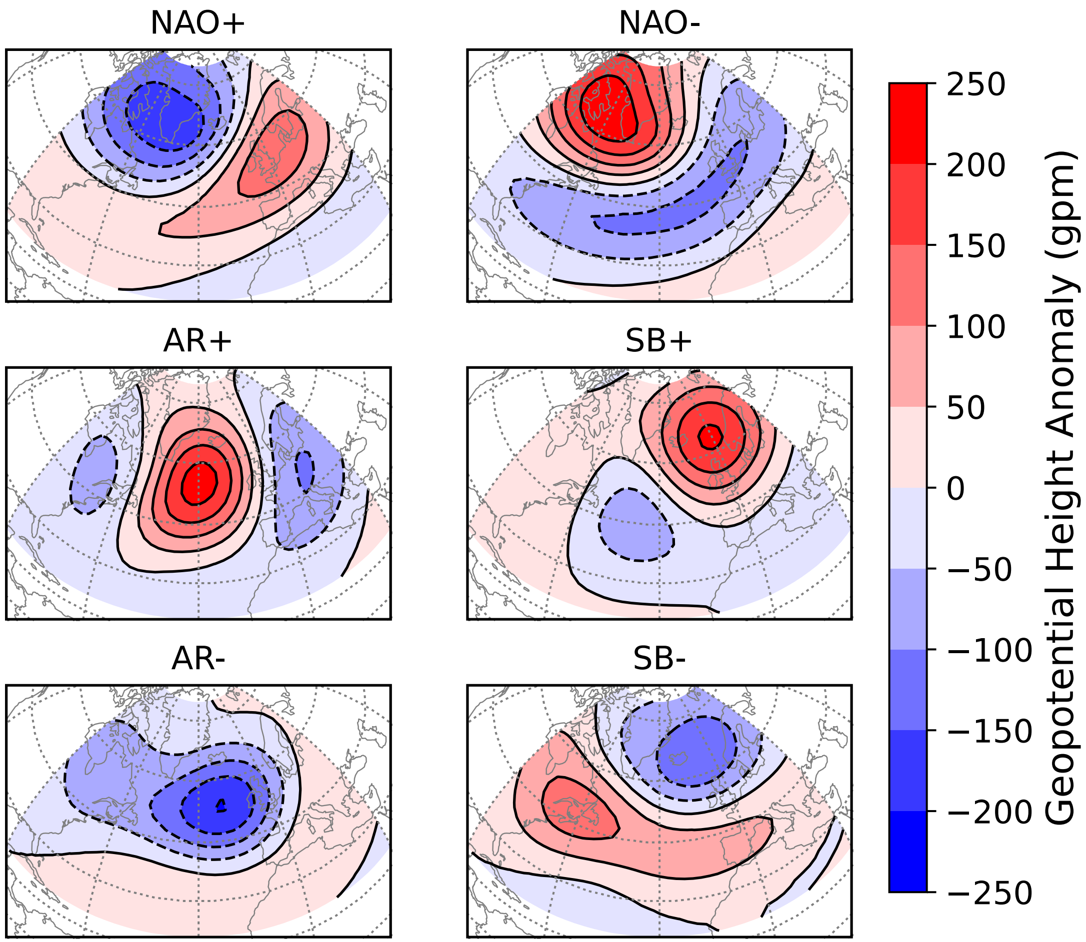
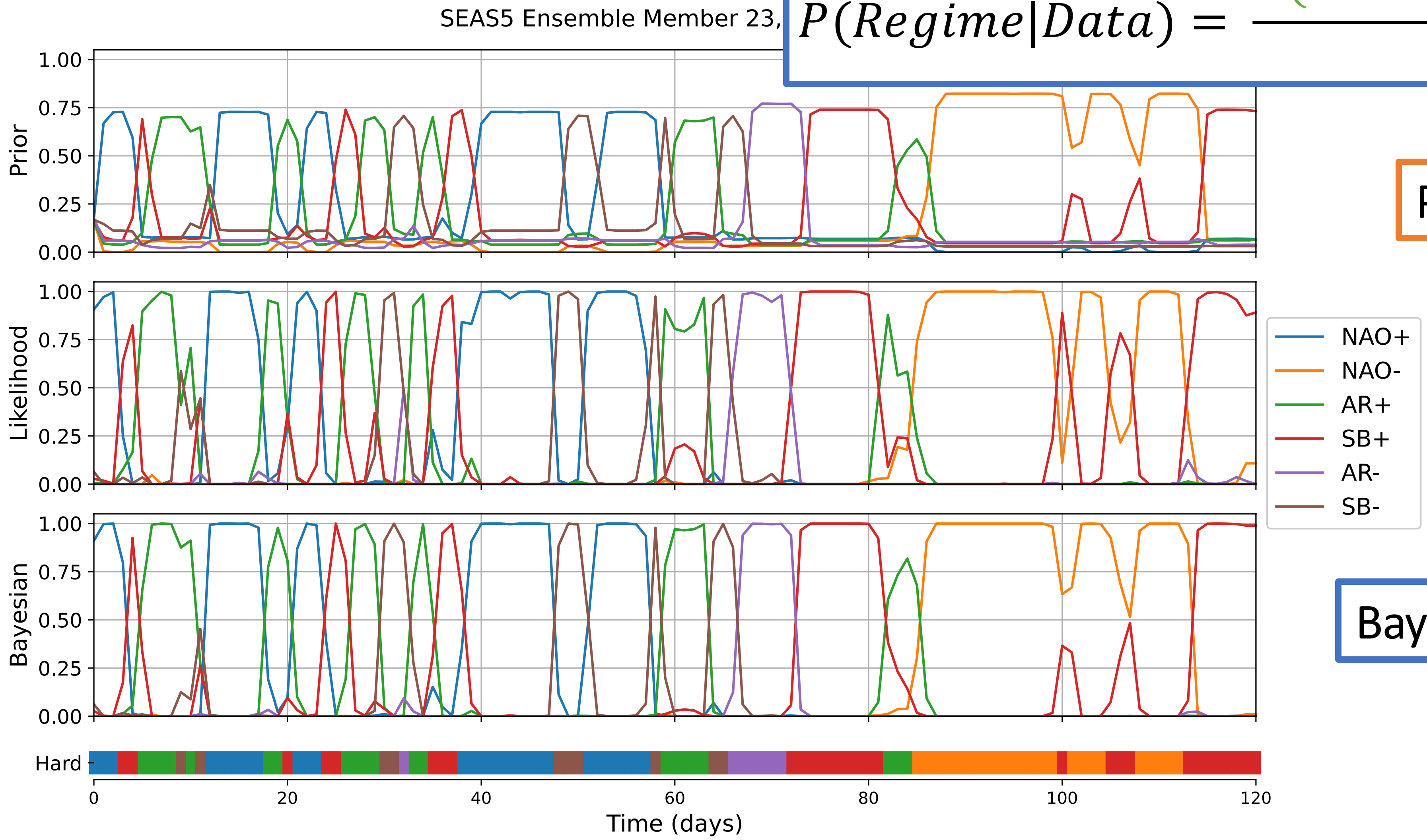


Fig. Six circulation regimes for the Euro-Atlantic sector (North Atlantic Oscillation (NAO) + and -, Atlantic Ridge (AR) + and -, Scandinavian Blocking + and -

How does the sequential Bayesian regime assignment work?

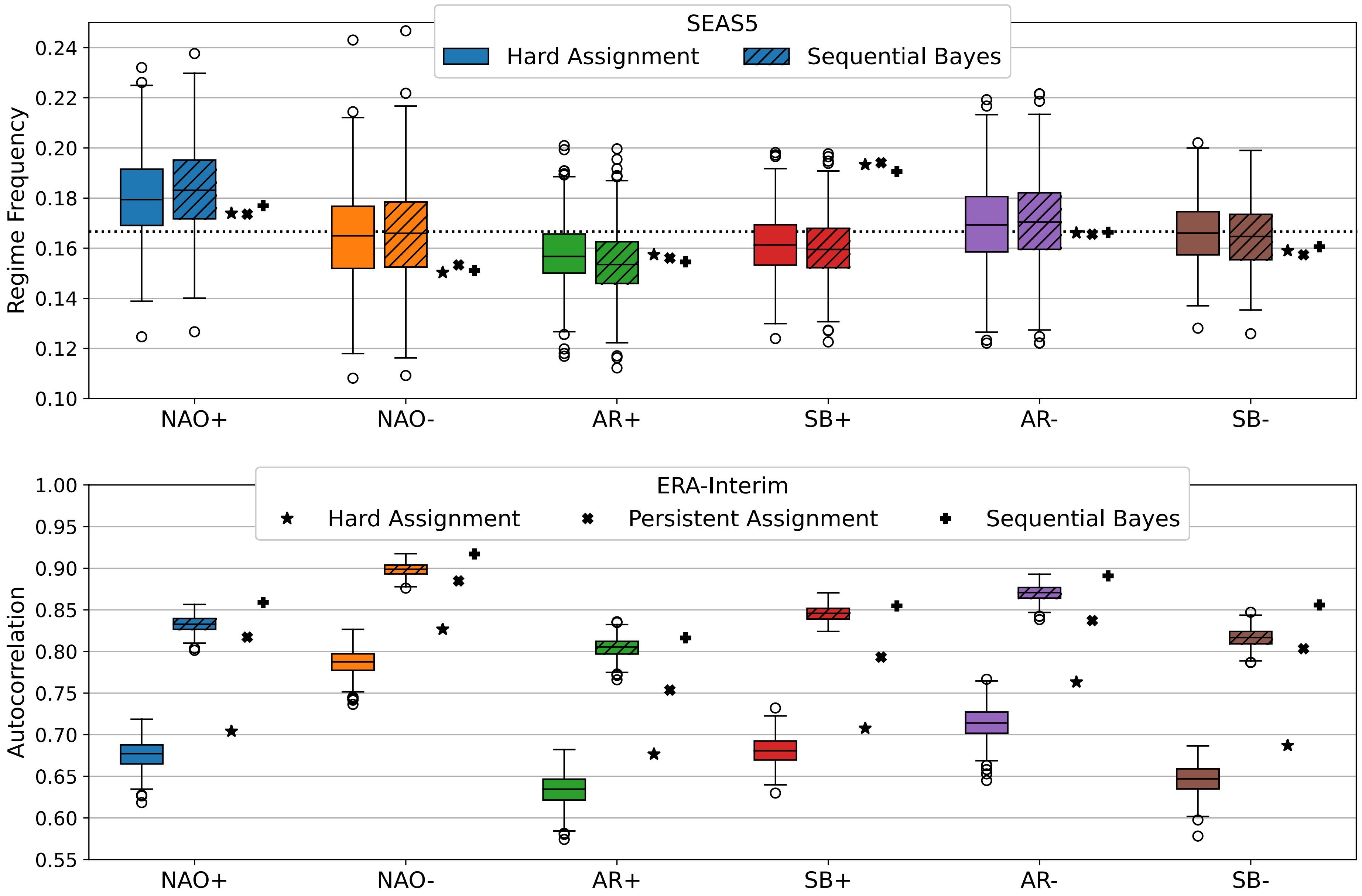
$$P(Regime|Data) = \frac{P(Data|Regime)P(Regime)}{P(Data)}$$



Prior: $P_{t+1} = TP_t$

Observed Likelihood

Bayesian probabilities



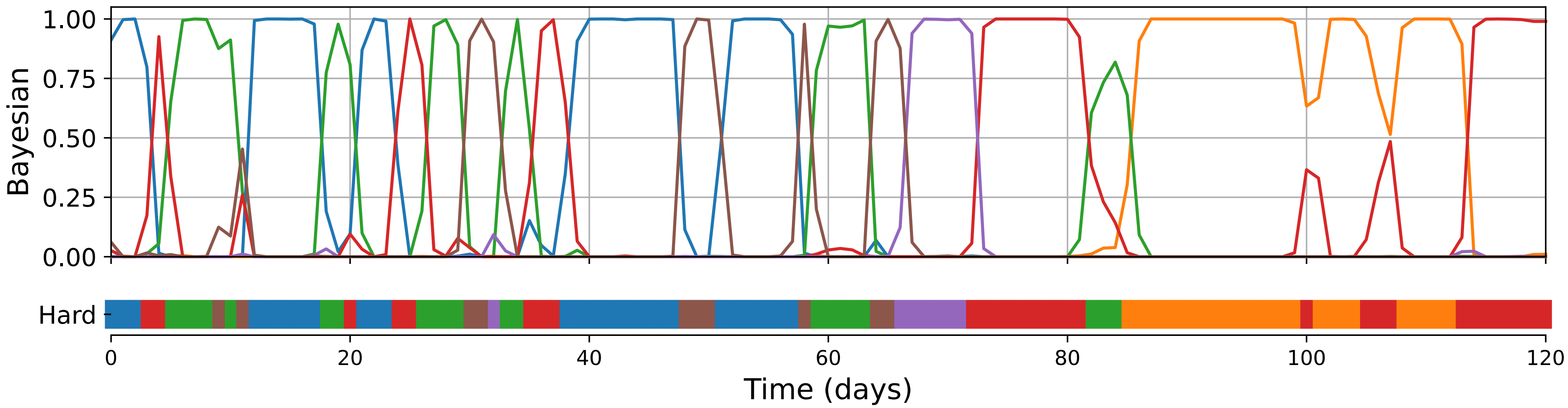
The sequential Bayesian approach yields more persistent regimes without affecting the regime frequencies.

The sequential Bayesian approach yields a more informative regime assignment compared to standard categorical methods.

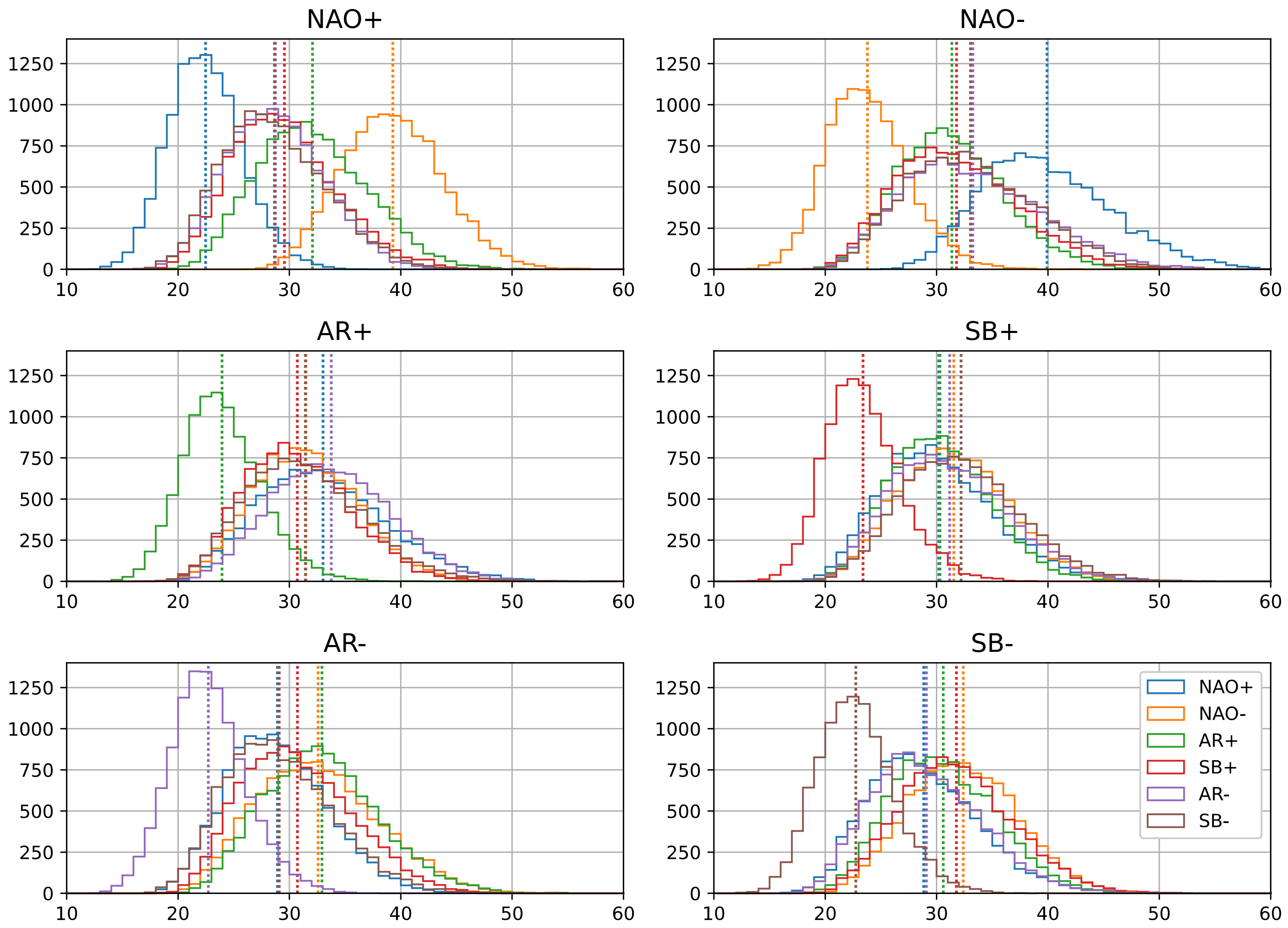
QUESTIONS?

Bayes Theorem:

$$P(Regime|Data) = \frac{P(Data|Regime)P(Regime)}{P(Data)}$$



A Bayesian approach treats the regime assignment in a probabilistic way



$\boxed{\text{Posterior}}$

Bayes Theorem:

$$\boxed{P(r|D)} = \frac{\boxed{P(D|r)}\boxed{P(r)}}{\boxed{P(D)}}$$

Prior: $P_t(r) = TP_{t-1}(r|D)$

Data: $P(D) = \sum_{r=1}^{k} P(D|r)P(r)$

Observations: $P(D|r) = \frac{-\frac{1}{2}(D-\mu_r)^T \Sigma_r^{-1}(D-\mu_r)}{\sqrt{(2\pi)^k |\Sigma_r|}}$