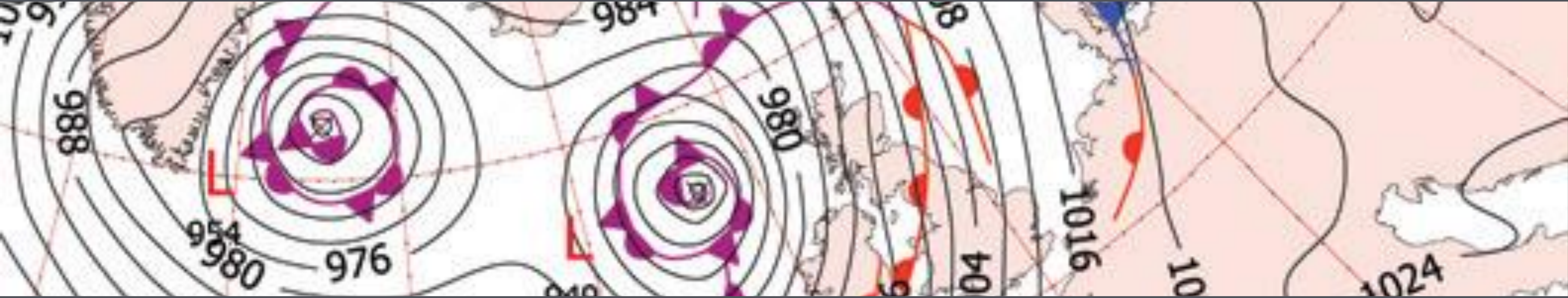
Department of Meteorology

University of Reading

EXTRACTING LIKELY SCENARIOS FROM HIGH RESOLUTION FORECASTS IN REAL-TIME



By Kris Boykin

Supervised by John Methven and Tom Frame, University of Reading, Nigel Roberts and Stephen Moseley, Met Office

Figure: Mean SLP and fronts for 12:00 12/10/2018 (top). Adapted from "Weather Log", by B. Prichard, 2018, *Weather*, *73*(12), p. ii-iii.

Scenario DOCTORAL TRAINING PARTNERSHIP | NERC SCIENCE OF THE ENVIRONMENT | Met Office

INTRODUCTION

- High-resolution ensemble forecasts for local and high impact weather produce large amounts of data
- The quantity and complexity of data makes it difficult for forecasters to digest quickly in real time
- Forecasters must be able to quickly group forecasts into scenarios

RESEARCH QUESTION

How do we group high resolution forecasts into likely scenarios for high impact events?

- Relevant to users
- Cluster members should indicate similar outcomes in terms of high impact weather
- Allow forecasters to assign probabilities to each scenario

BACKGROUND

- Previous clustering has been done on:
  - Large scale smooth fields such as Z500, i.e. North Atlantic patterns or variability and central Europe Großwetterlagen
  - Large scale pattern clustering to compare to local precipitation, using horizontal wind, geopotential height, and specific humidity at 500, 700, and 850 hPa (Marsigli et al., *Nat. Hazards Earth Syst. Sci.*, 2004)

CASE STUDY

- The events – heavy rain over the UK during the second week of October 2018 and an extratropical cyclone developing over the North Atlantic
- The model – Met Office Unified Model
- The ensemble – MOGREPS-G
- The area – 40° N to 70° N, 45 W° to 45° E
- The time – 8 days starting 00:00 10/10/2018
- The variable – θ_w at 850 hPa

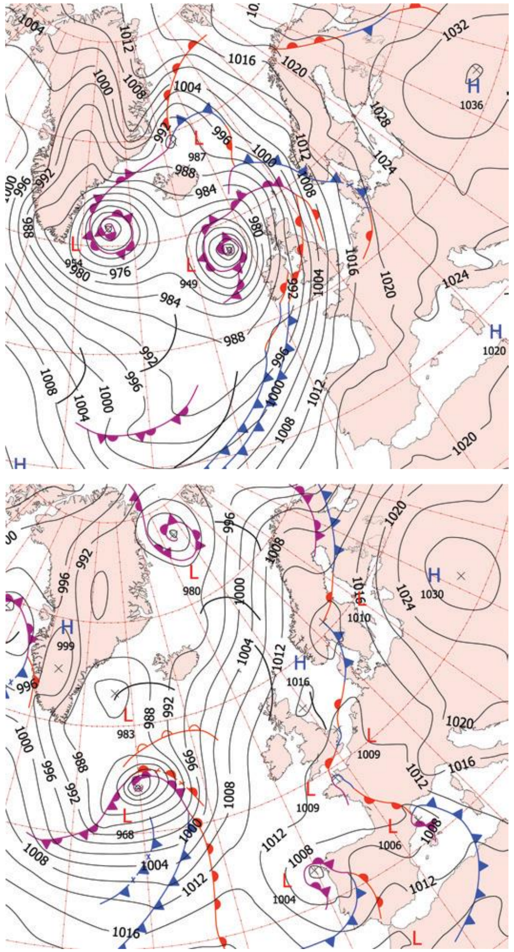
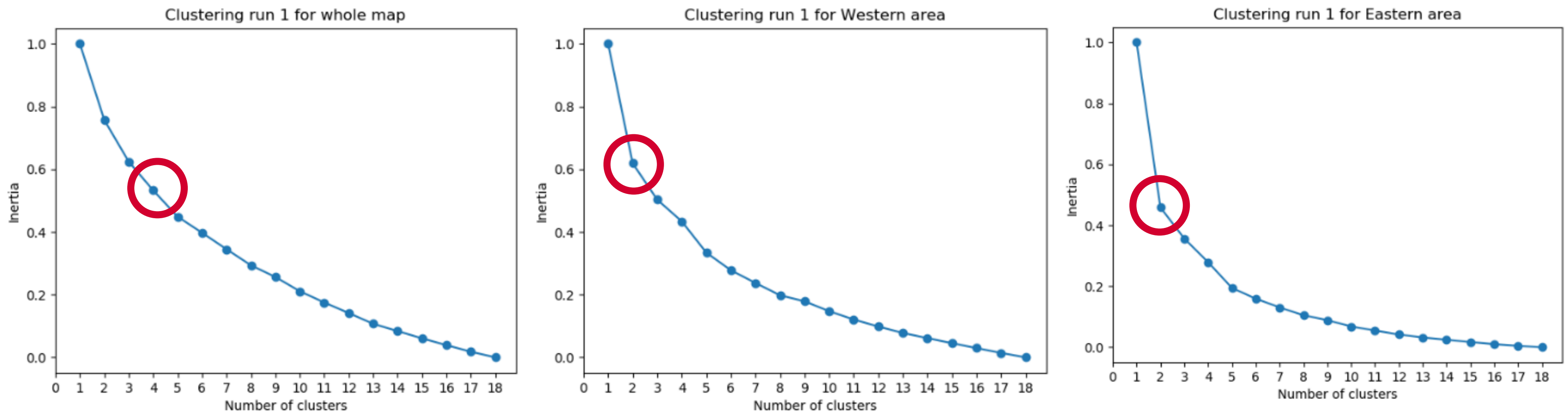


Figure: Mean SLP and fronts for 12:00 12/10/2018 (top) and 12:00 15/10/2018 (bottom). Adapted from "Weather Log", by B. Prichard, 2018, *Weather*, *73*(12), p. ii-iii.

RESEARCH METHOD

- We began with K-means (Pedregosa et al., *JMLR,* 2011.) clustering of ensemble forecasts using the RMS difference of θ_w fields to define the "distance" between forecasts.
- Global ensemble members are clustered using θ_w over different domains: the North Atlantic, a Western region and an Eastern region. These regions are held constant through time.
  - Western region: 40 N° to 55 N°, 45 W° to 30 W°
  - Eastern region: 46 N° to 61 N°, 7 W° to 7 E°

CHOOSING THE NUMBER OF CLUSTERS



Inertia per cluster: $I_n = \sum_i d_{i,n}^2$

Inertia: $I = \sum_n \sum_i d_{i,n}^2$

$d_{i,n} = RMS(\theta_i, \theta_n)$

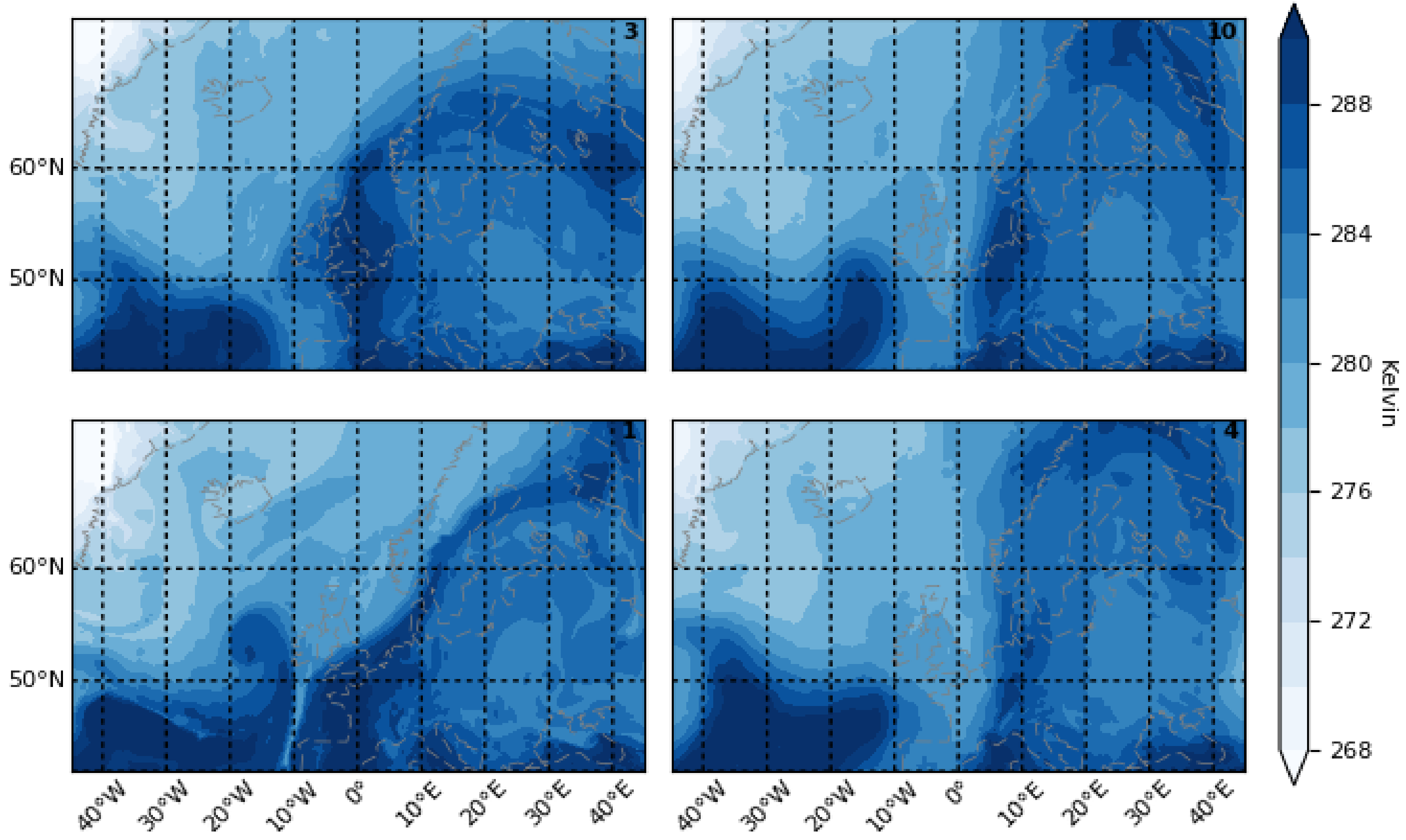
i =member label

n =cluster label

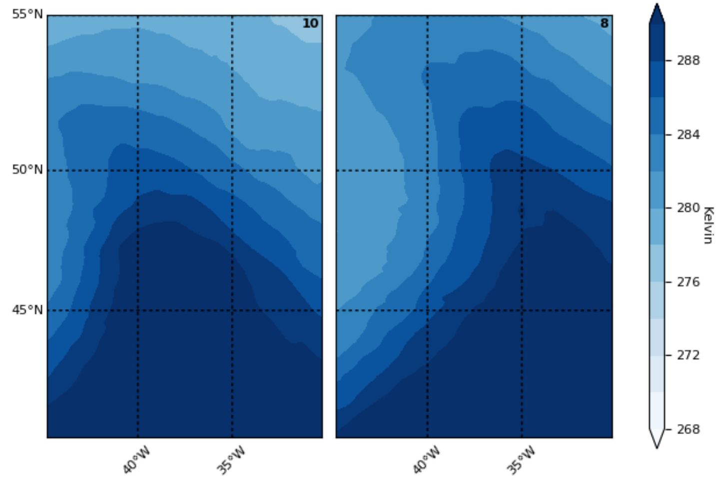
d =distance between member i and the centroid of cluster n in terms of RMS

CLUSTER CENTROIDS

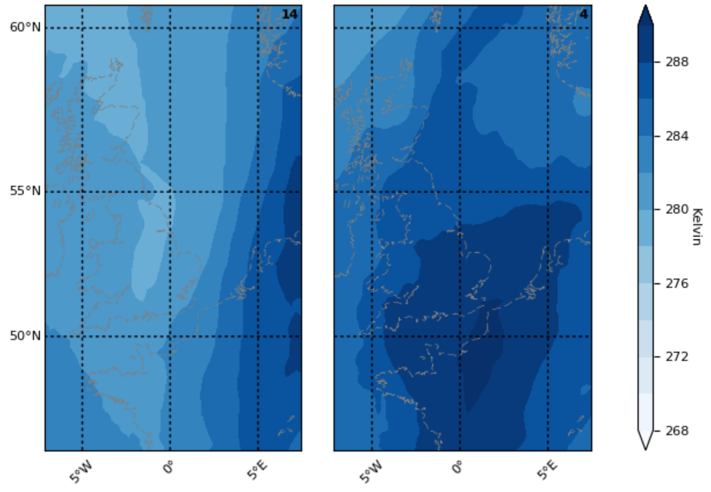
Whole θ_w centroids at t + 120 hours



West θ_w centroids at t + 120 hours

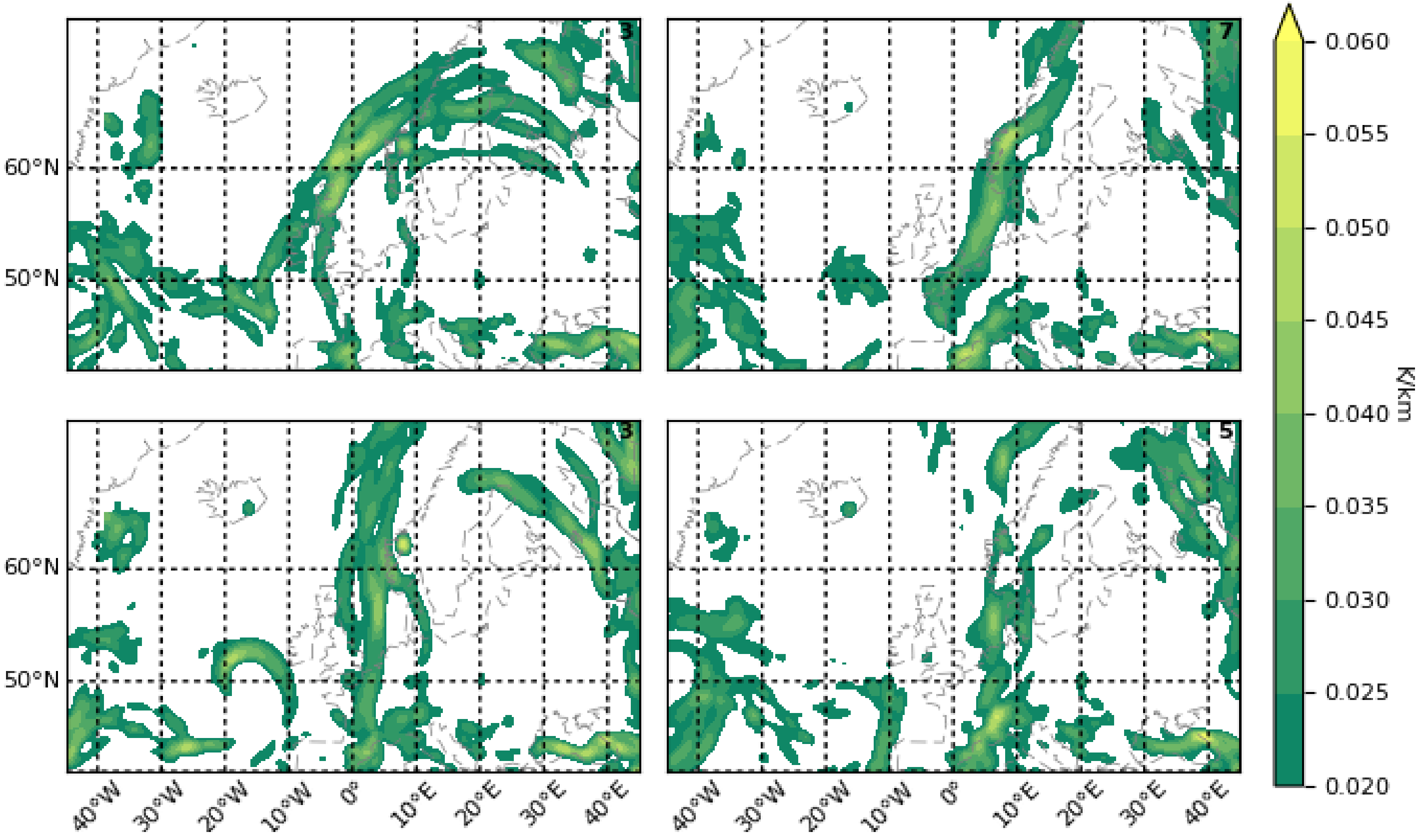


East θ_w centroids at t + 120 hours

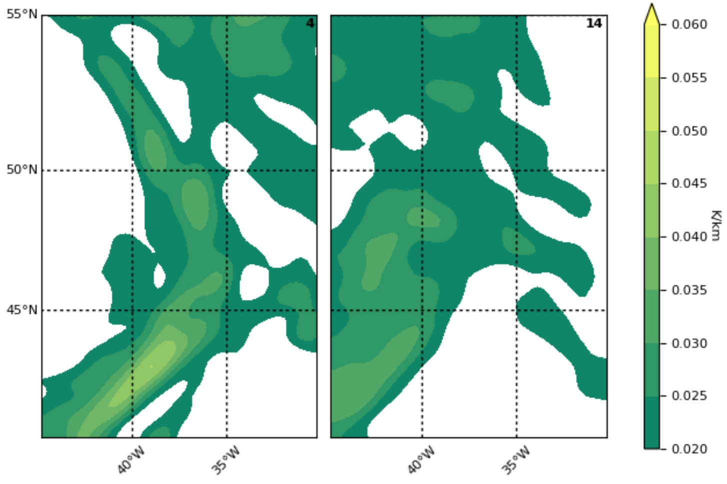


CLUSTER CENTROIDS

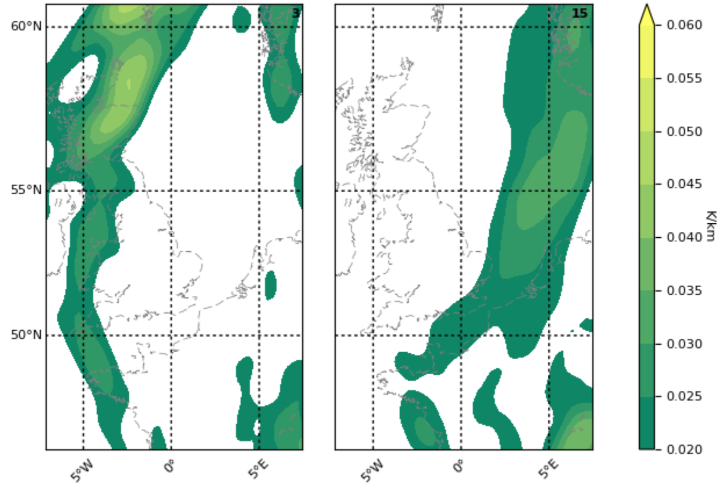
Whole $|\nabla\theta_w|$ centroids at t + 120 hours

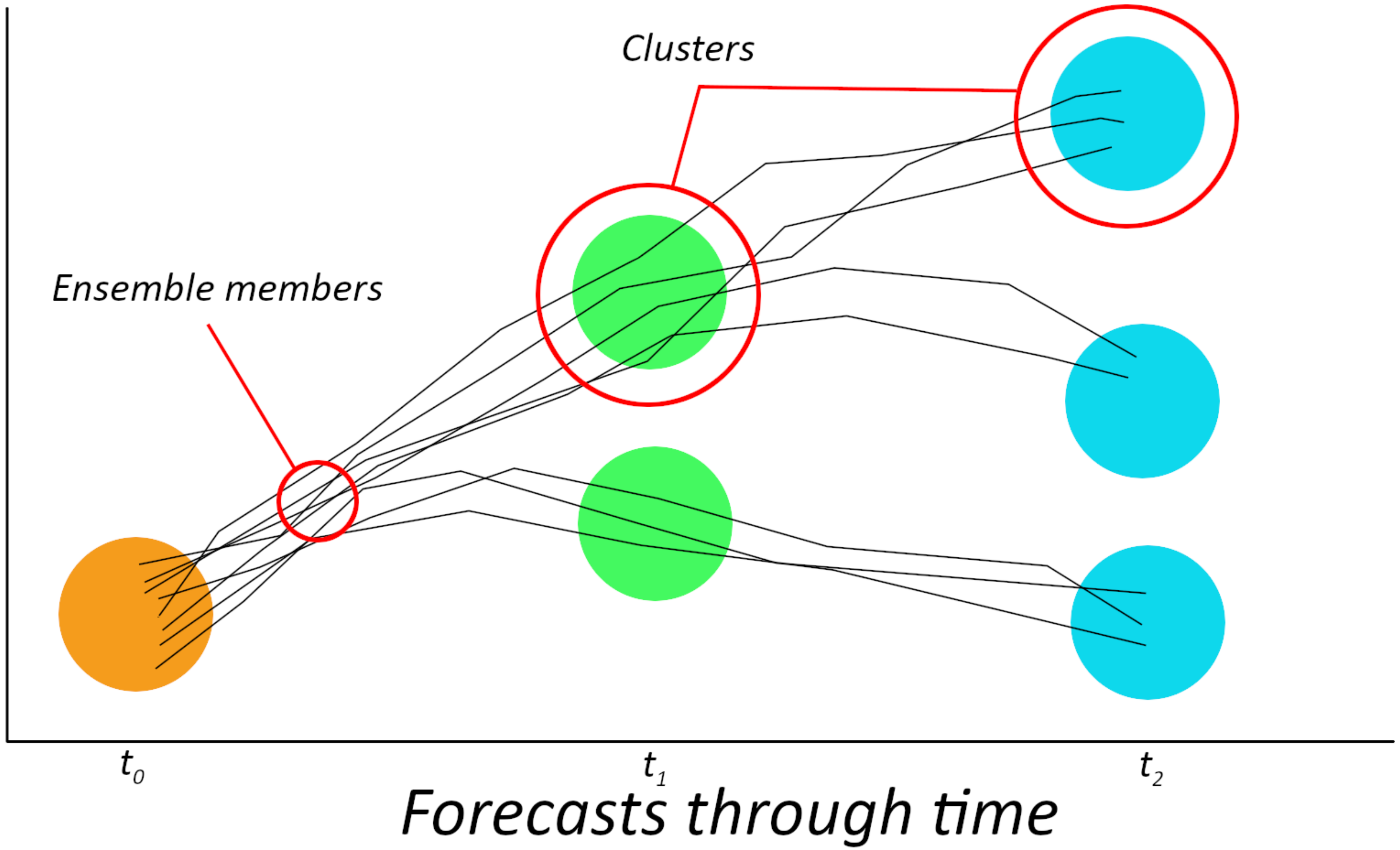


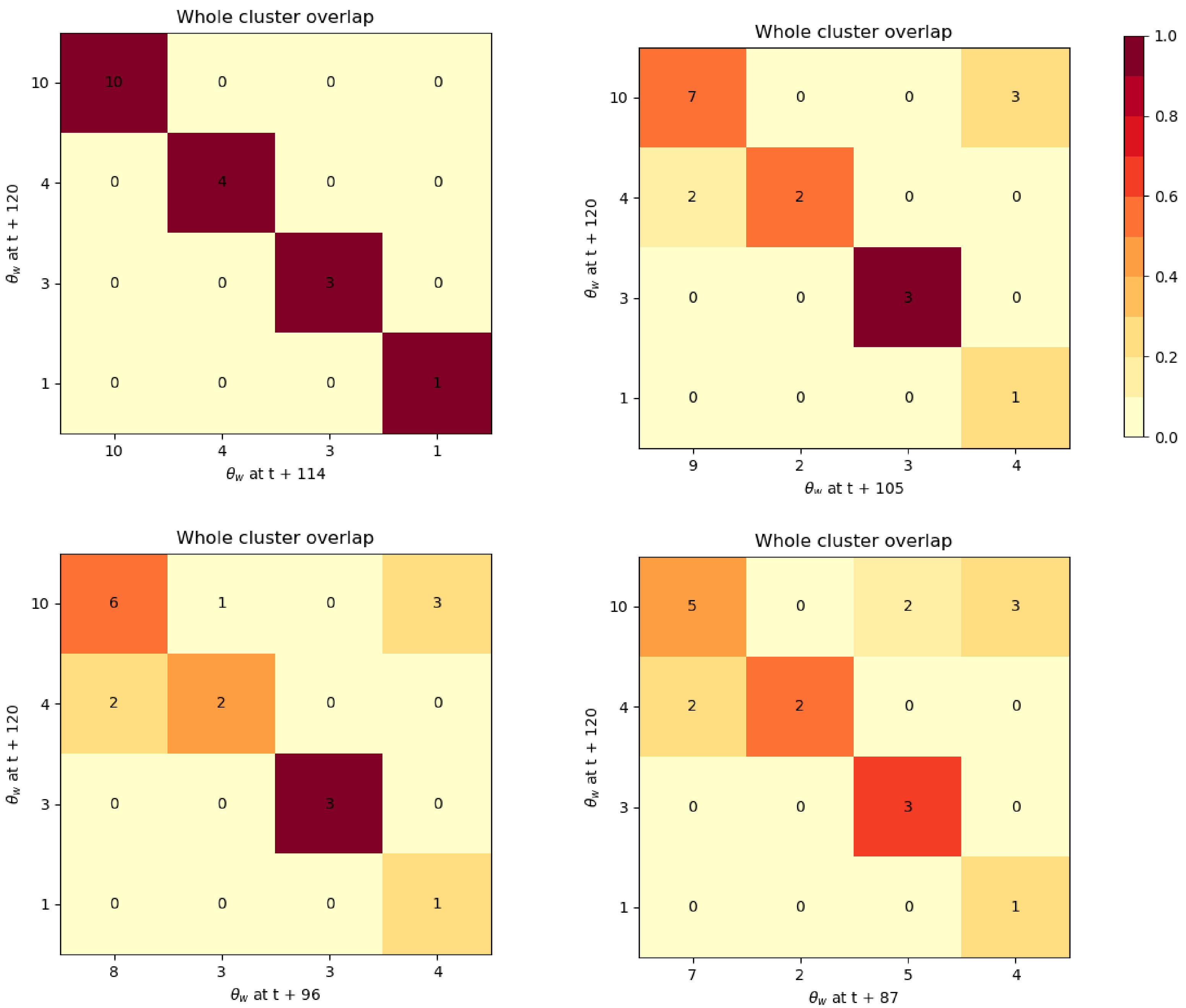
West $|\nabla\theta_w|$ centroids at t + 120 hours



East $|\nabla\theta_w|$ centroids at t + 120 hours




Clusters
Ensemble members
t_0
t_1
t_2
Forecasts through time



The Jaccard Index

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

FUTURE RESEARCH

- Do clusters identified using smooth, more predictable, fields produce scenarios relevant to heavy precipitation? What fields are most useful and why?
- Compare “distance” between forecasts of “weather variables” (i.e. precipitation rate)
    - Considerations – how to measure distance (FSS - Roberts, *Meteorol. Appl.,* 2008).
- Investigate success of clustering methods in terms of high impact weather outcomes.
    - Verification with observations (radar)
- Develop clusters that are relevant to operational forecasters and their communications with users.

THANK YOU FOR YOUR ATTENTION.

QUESTIONS?