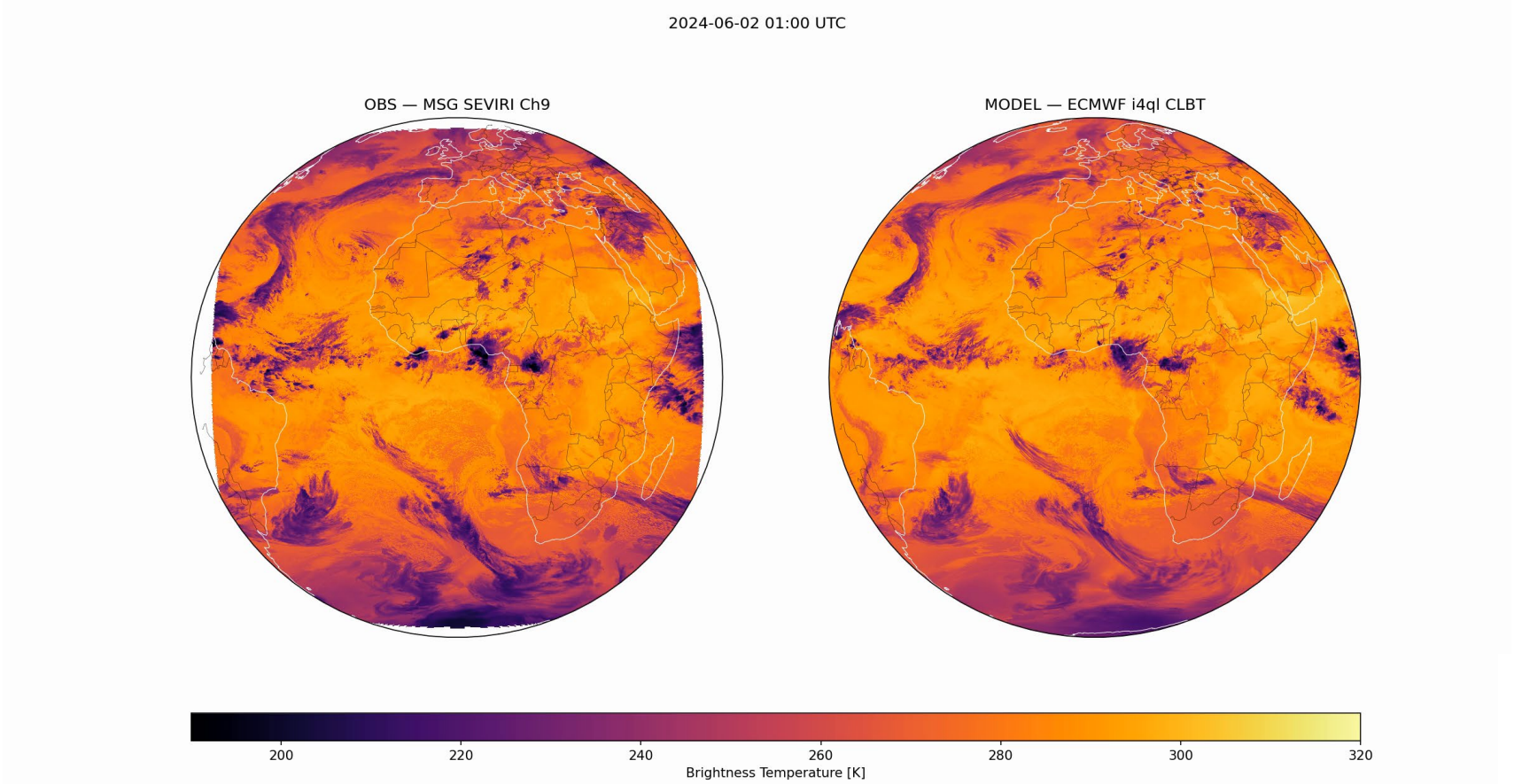
A New Concept for Comparing Satellite Observations and km-Scale Atmospheric Simulations using Self-Supervised Machine Learning

D. Chatterjee (KIT), P. Knippertz (KIT), N. Raabe (UoC), S. Crewell (UoC),
P. Dueben (ECMWF), B. Vannière (ECMWF)



Karlsruher Institut für Technologie

Dataset

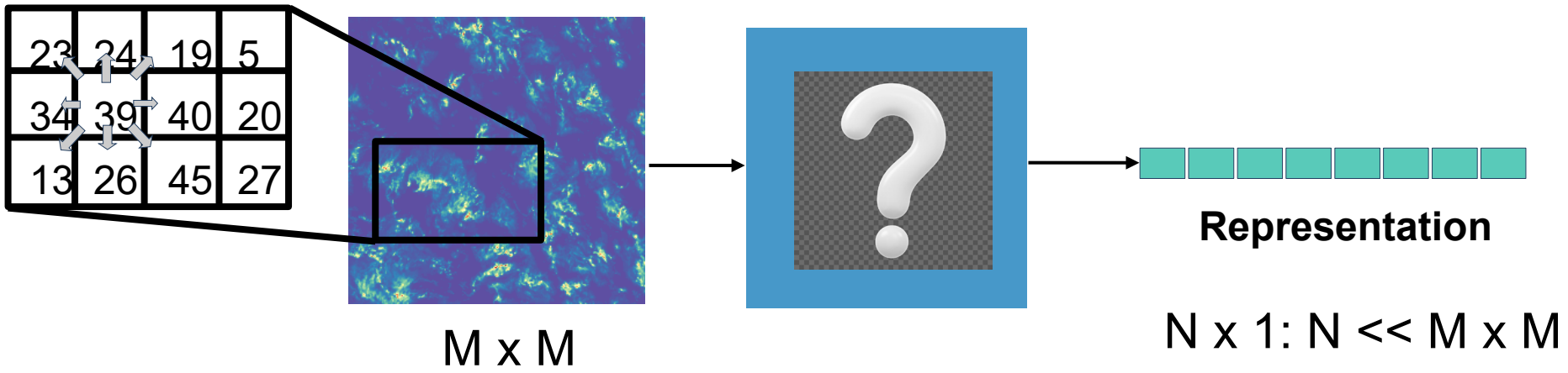
MSG 10.8 μm channel
Apr – Sep, 2020 – 2024
3.2 km × 3.2 km, 15 min
128x128, 350k samples

ICON–EU RTTOV 10.8 μm
April -Sep 2024
6.5 km, hourly
88x88, 18k samples

IFS-i4ql RTTOV 10.8 μm
April – Sep 2024
4.4 km, hourly
100x100, 18k samples

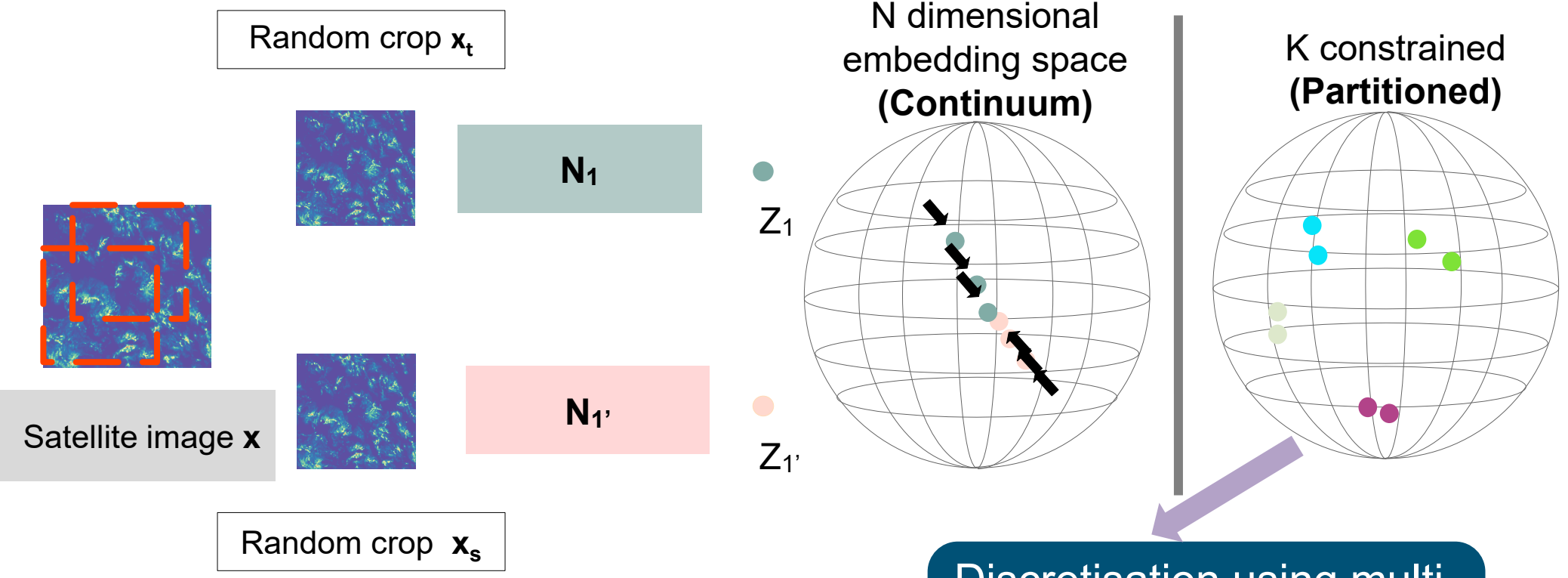
Icon data: Dr. Vera Schemann, uni köln

Embedding vector representation



Extract contextual information from high dimensional space to build low-dimensional representations

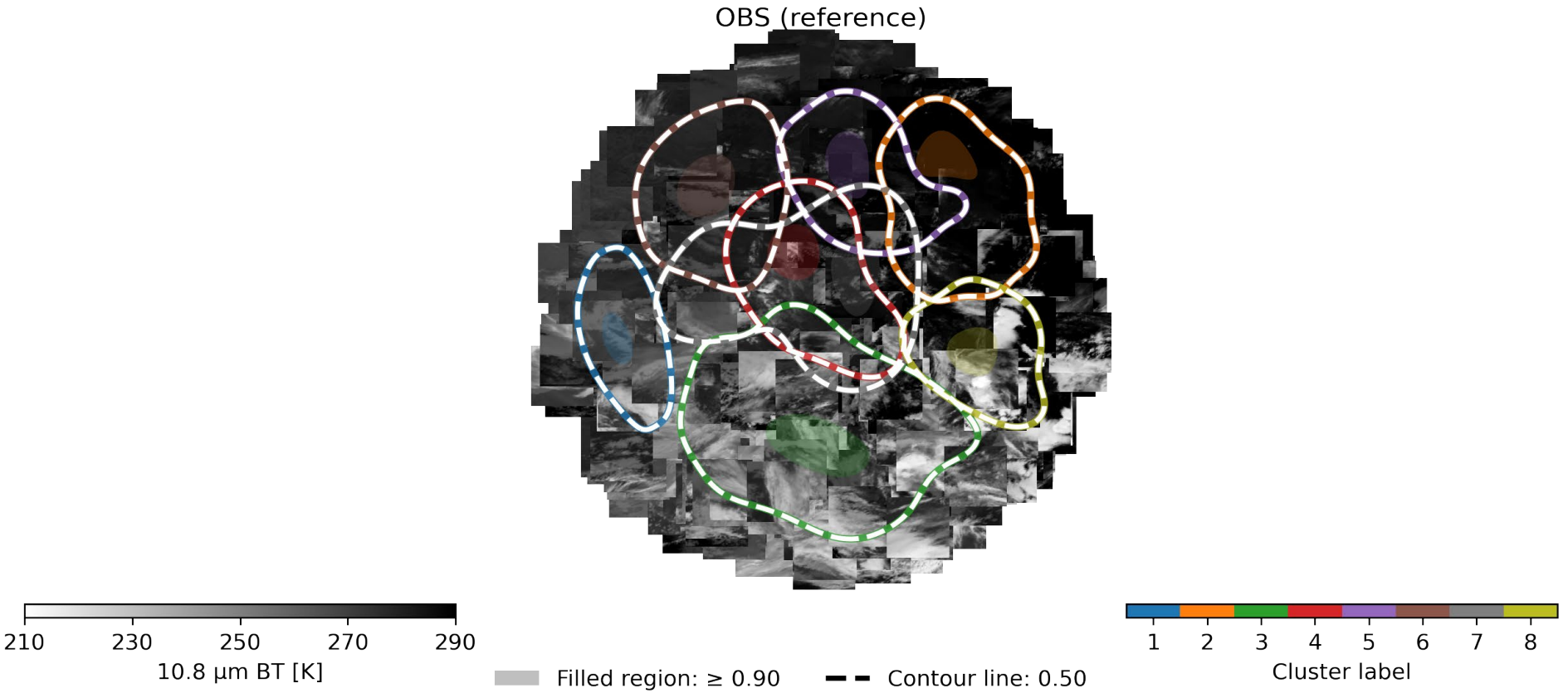
Representation learning approach



N = Vision transformer / Convolutional nets

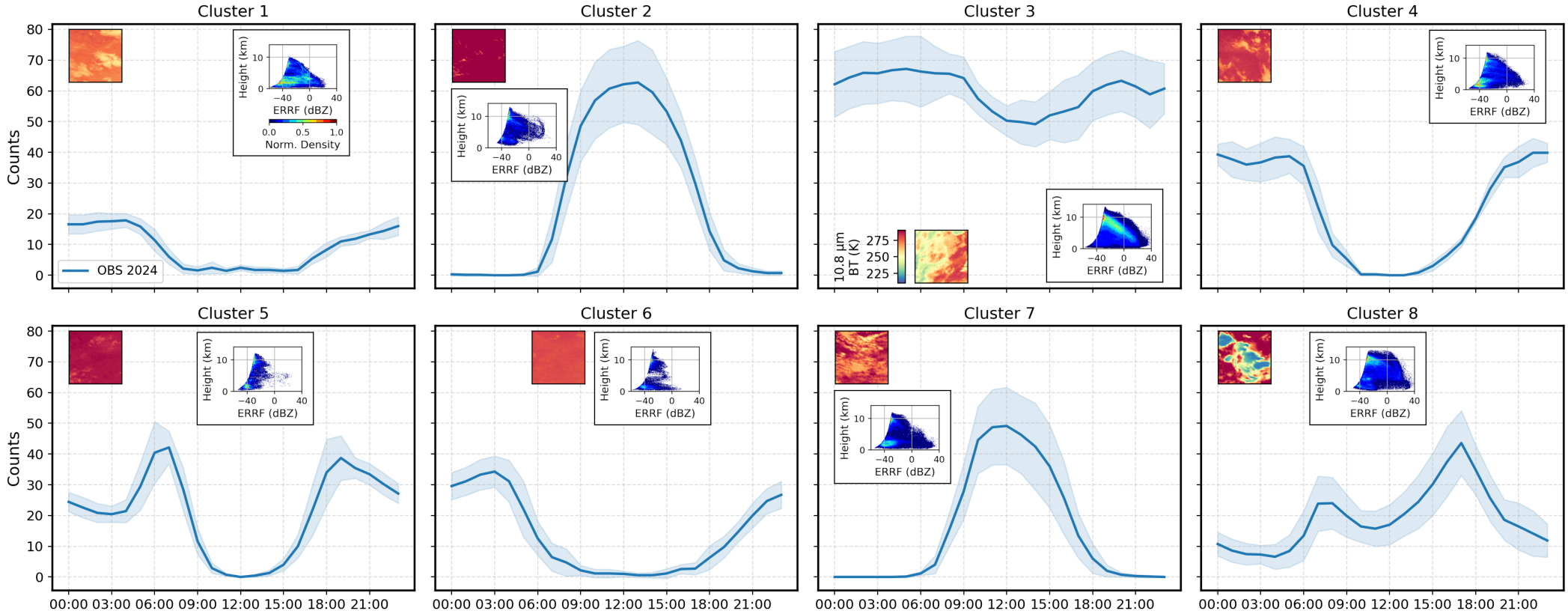
Facebook artificial intelligence research (FAIR) (Caron, 2020, 2022)
FAIR open source Vision Self Supervised Learning library (Priya Goyal, 2021)

Observation latent space

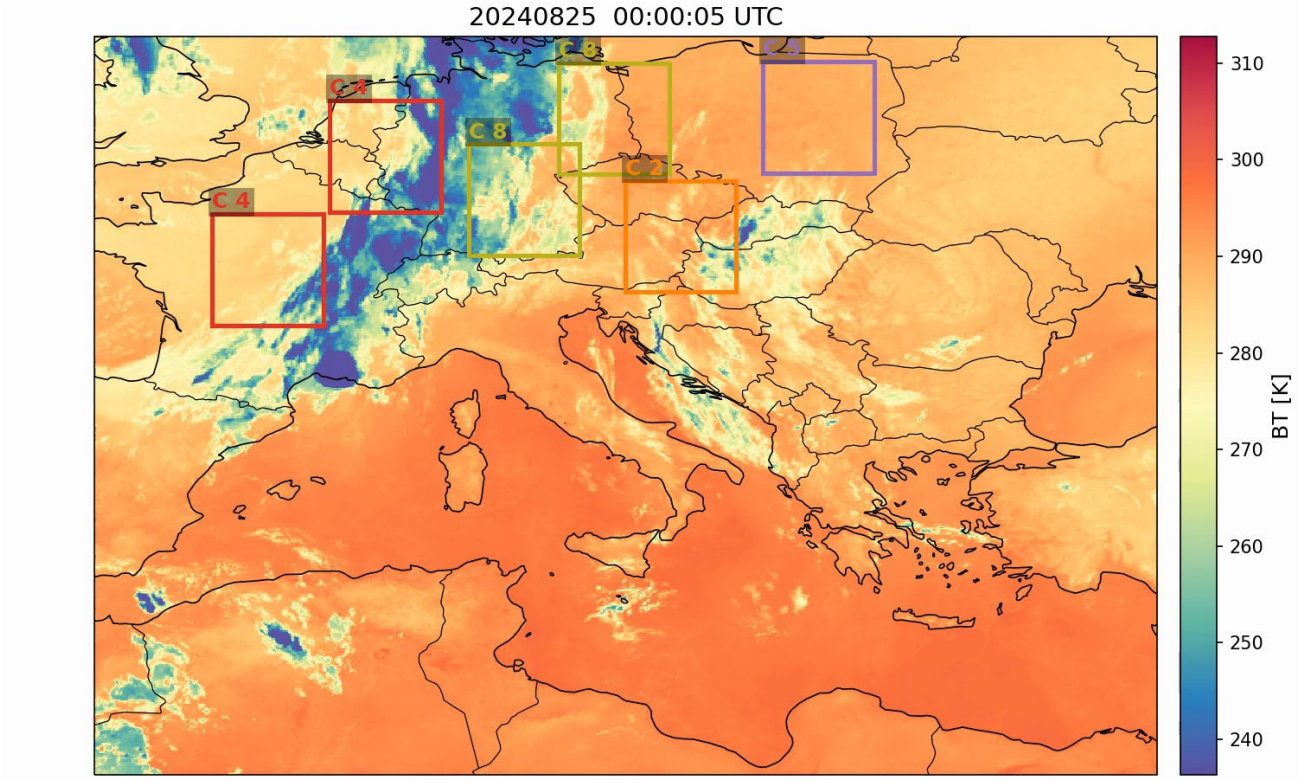


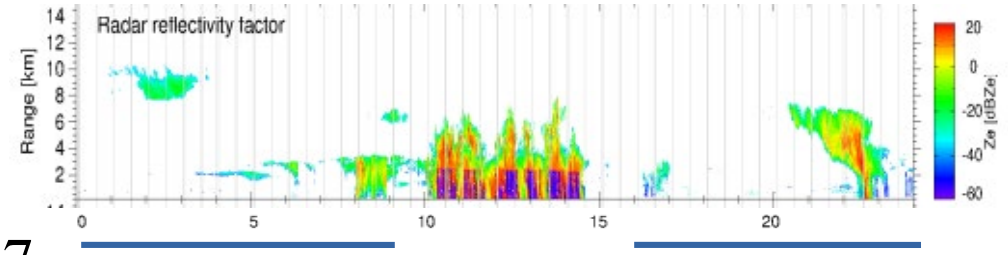
Chatterjee et al., 2023, 2024, 2025

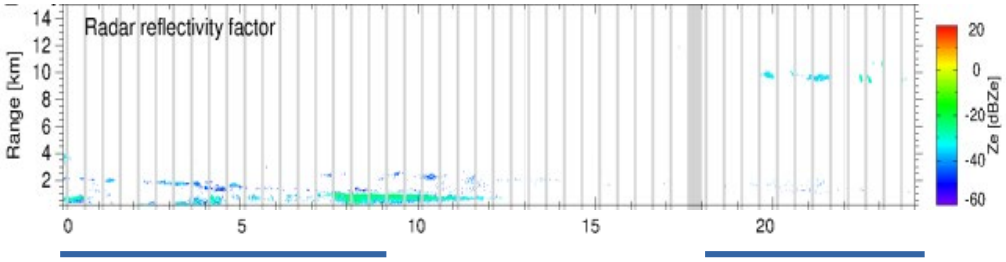
Physical characteristics



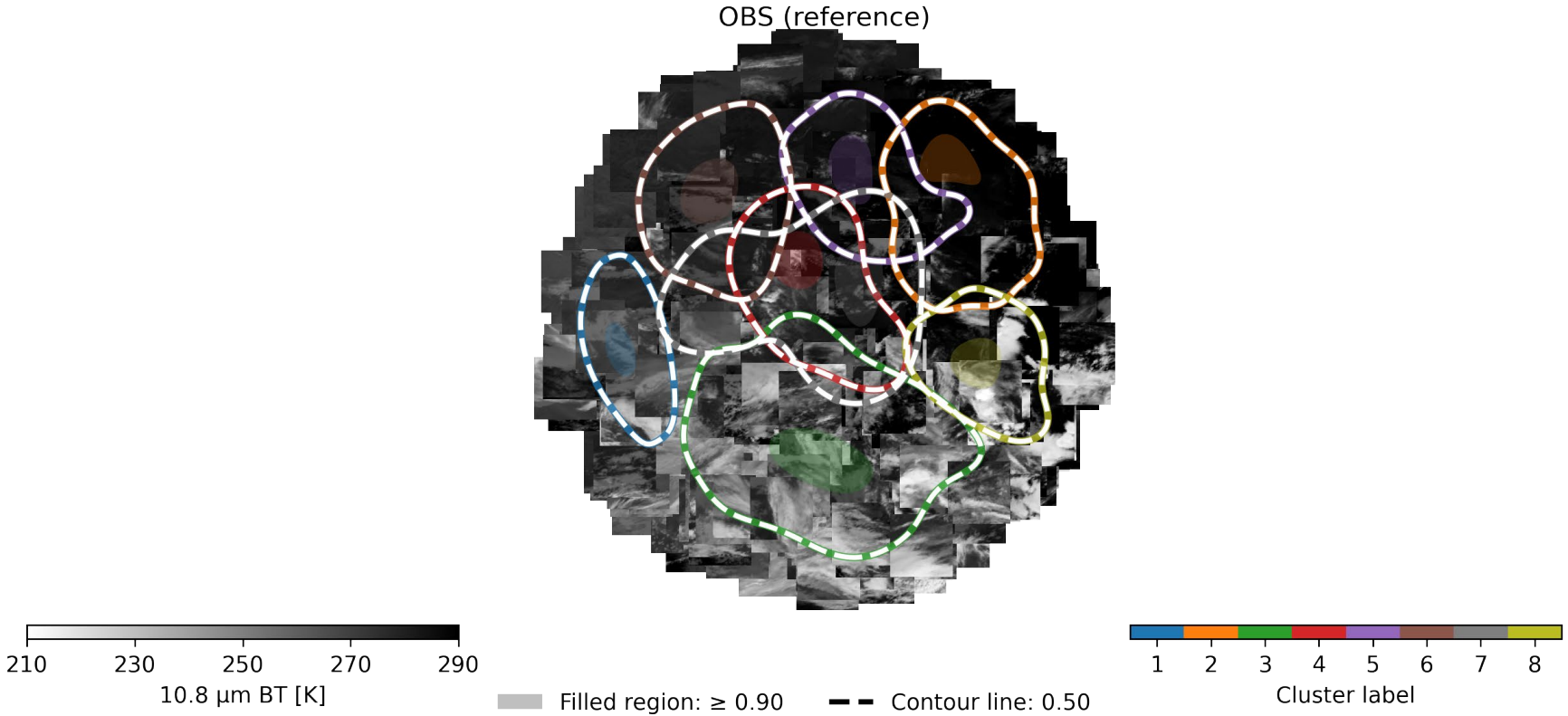
C4 example day





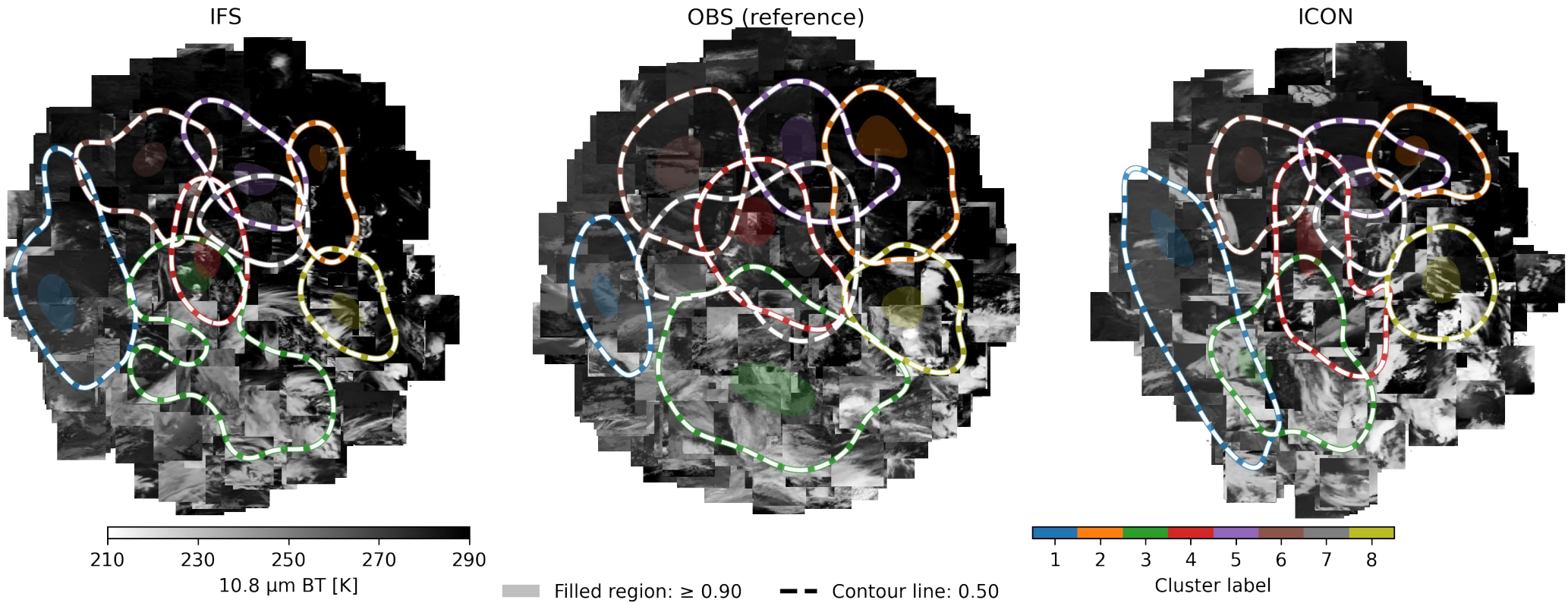


Observation latent space



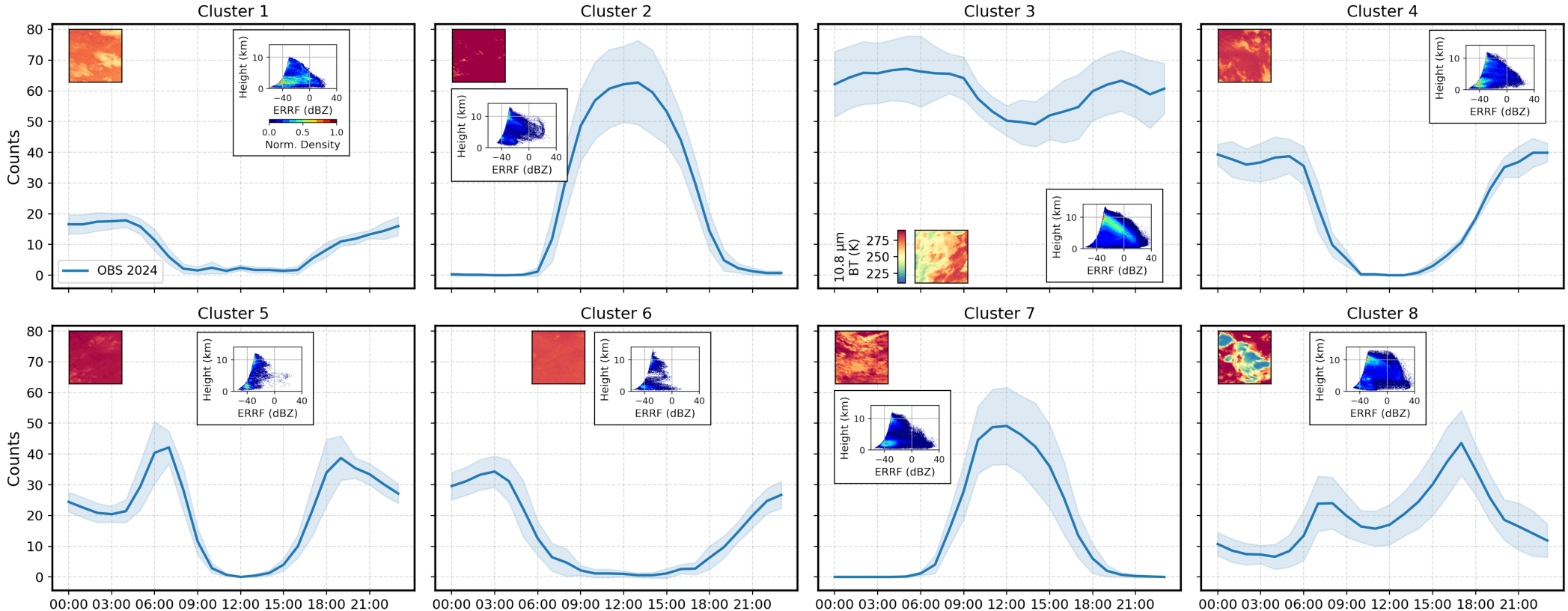
Chatterjee et al., 2023, 2024, 2025

Observation latent space

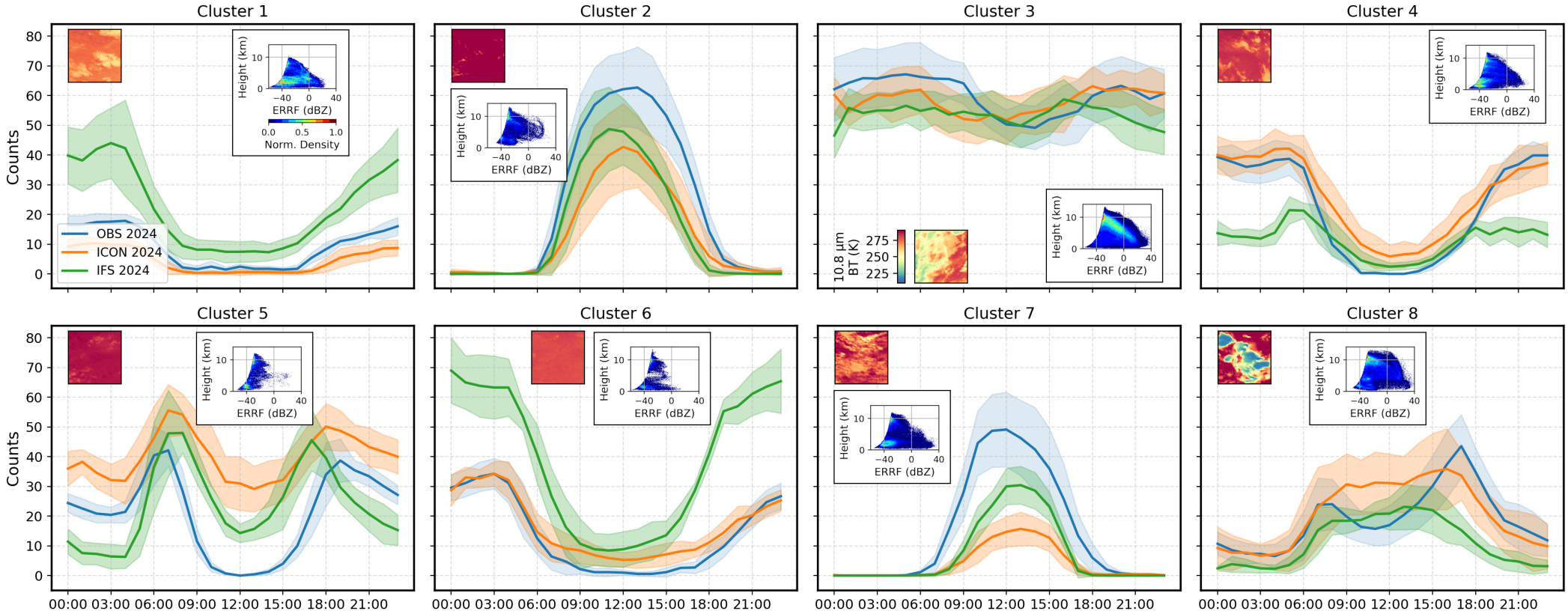


Chatterjee et al., 2023, 2024, 2025

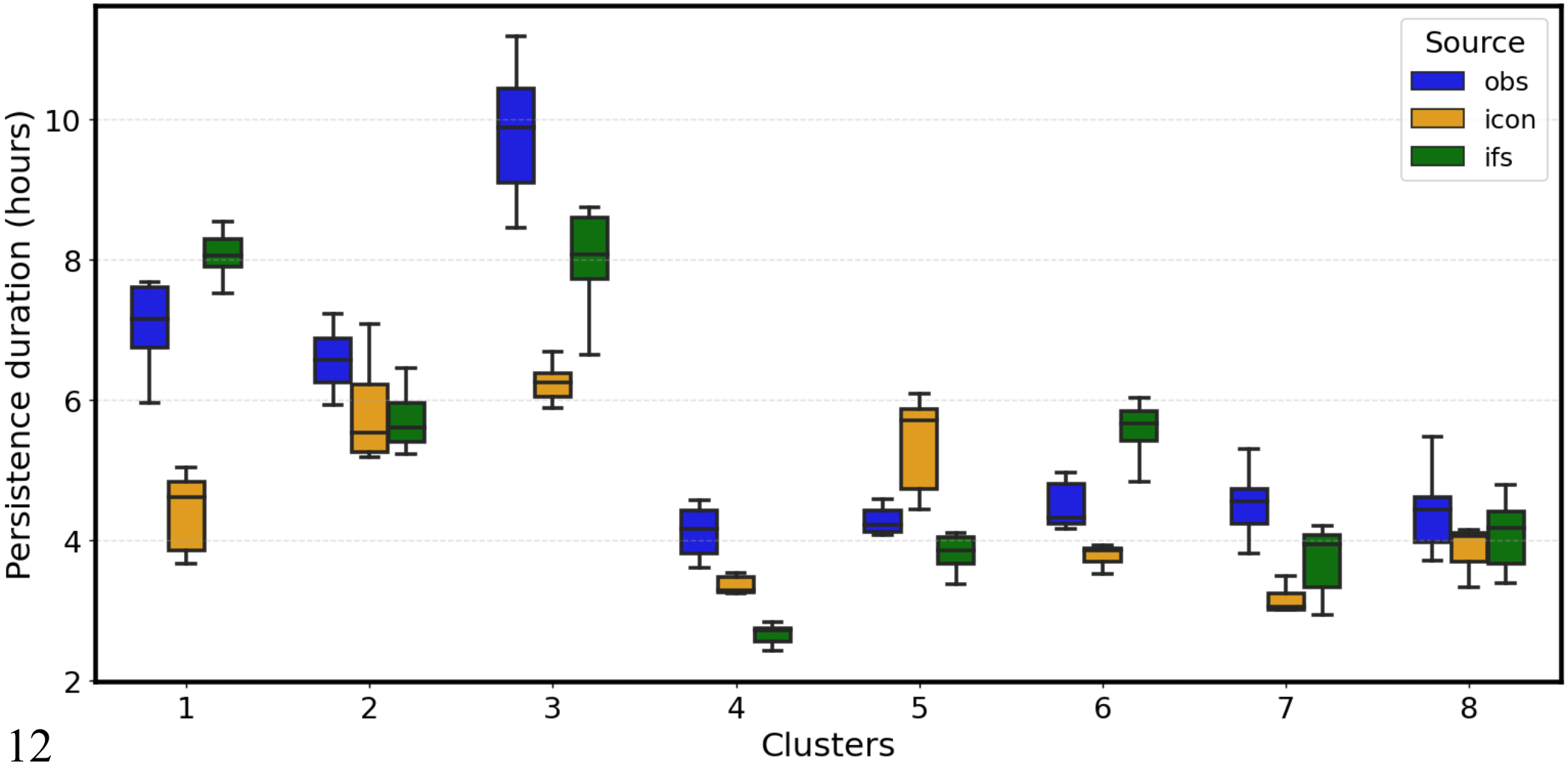
Physical characteristics



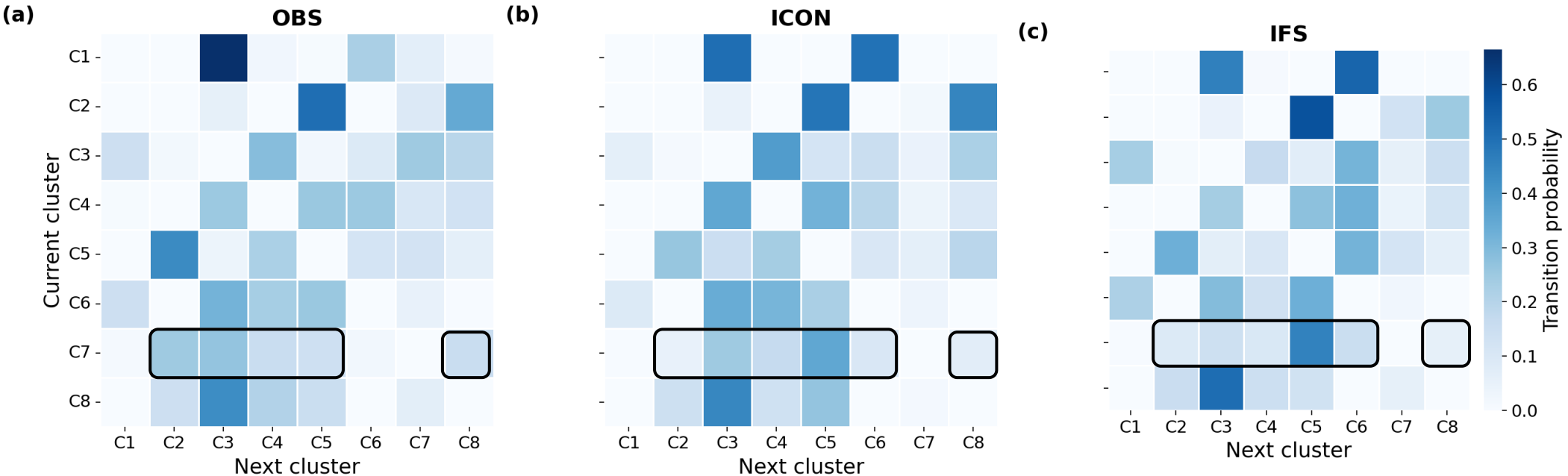
Physical characteristics



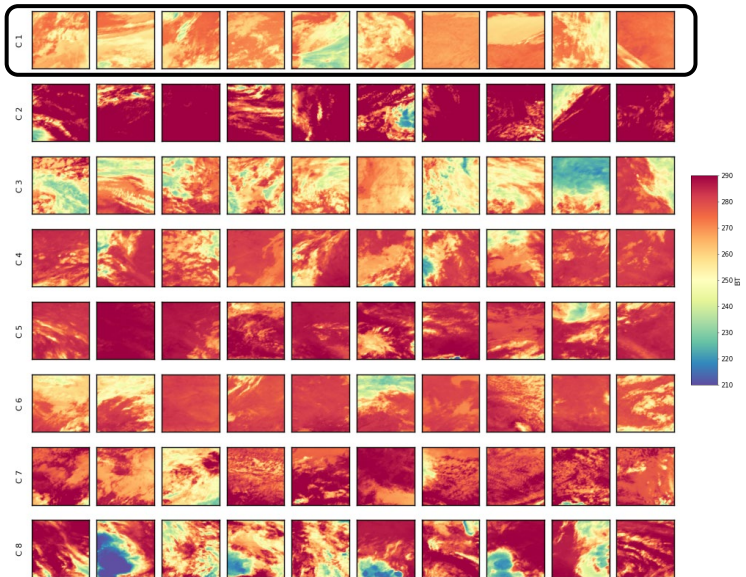
Evolution Tendency: Persistence



Evolution Tendency: Transition Probability

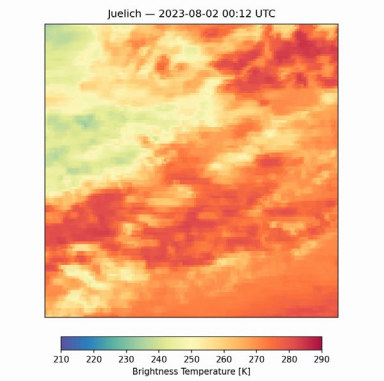


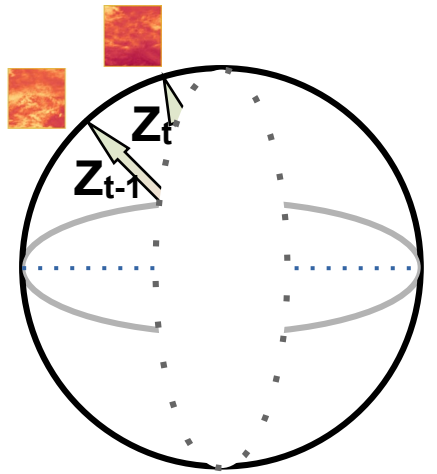
Coupled response & models decision making



OUTLOOK

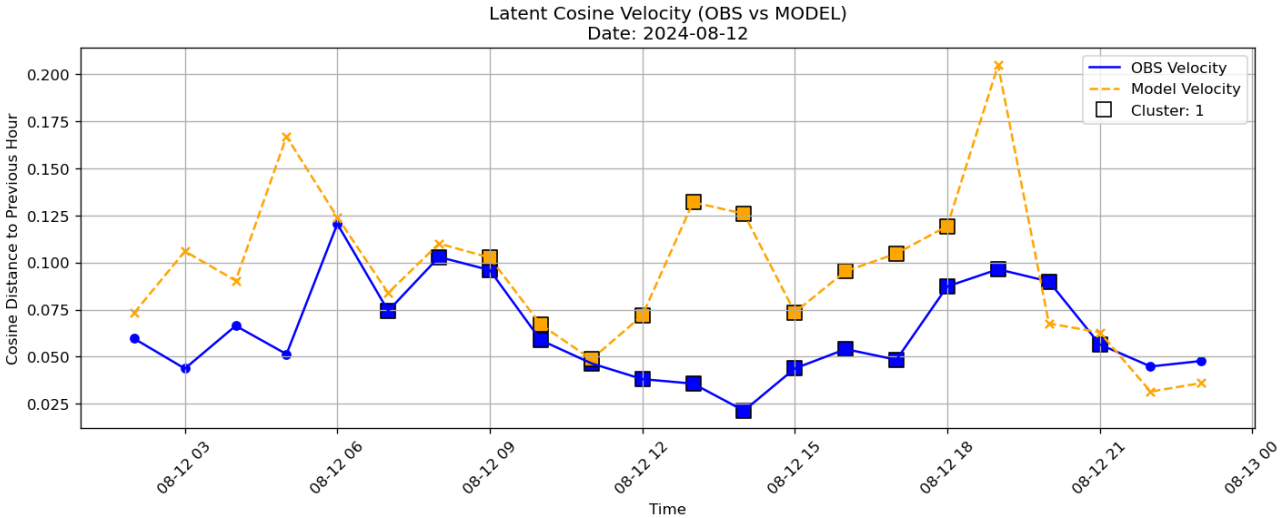
Latent velocity: Concept



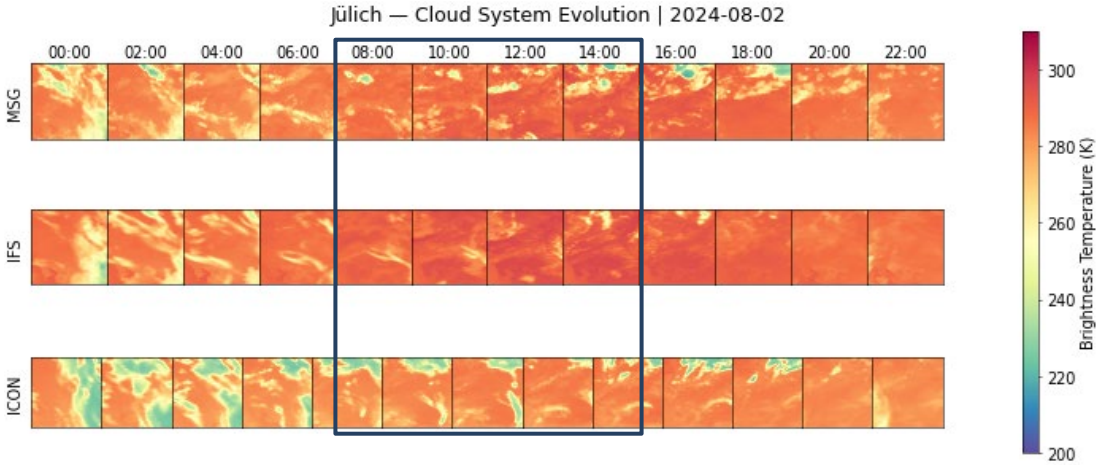


$$\text{Cos } \theta = \{Z_t, Z_{t-1}\}$$

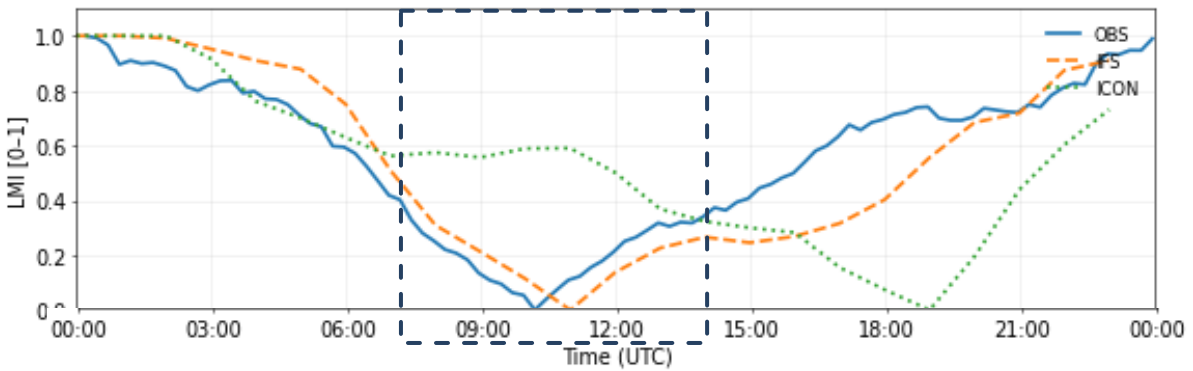
$$V_t = 1 - \cos \theta$$



Latent maturity index

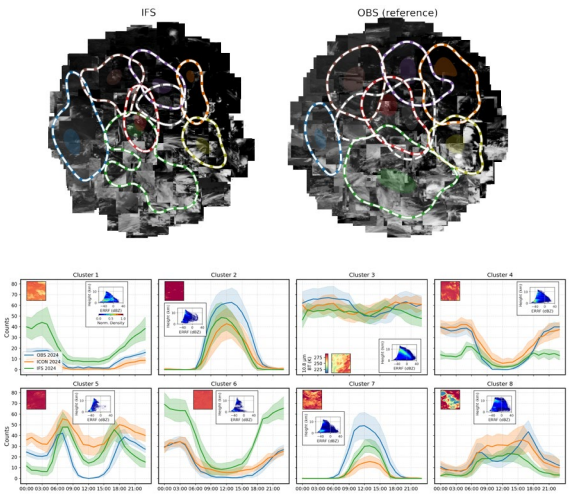


$$LMI(t) = \frac{d_{\cos}\left(Z_t, Z_{immature}\right)}{d_{\cos}\left(Z_{mature}, Z_{immature}\right)} = \frac{d_{\cos}\left(Z_t, Z_{warmest\,TB}\right)}{d_{\cos}\left(Z_{coldest\,TB}, Z_{warmest\,TB}\right)}$$

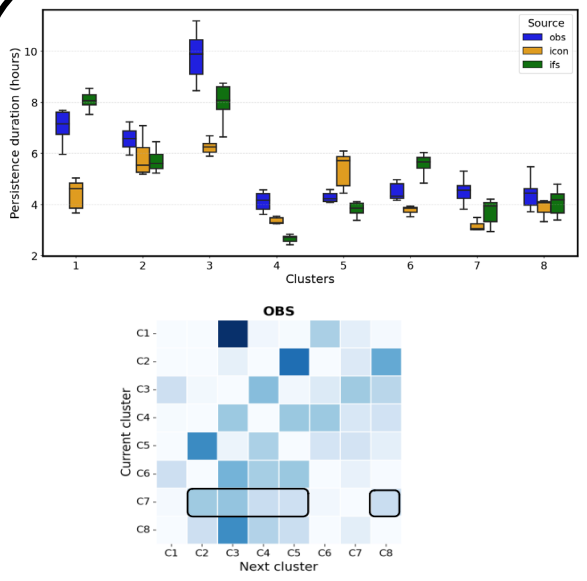


Conclusion

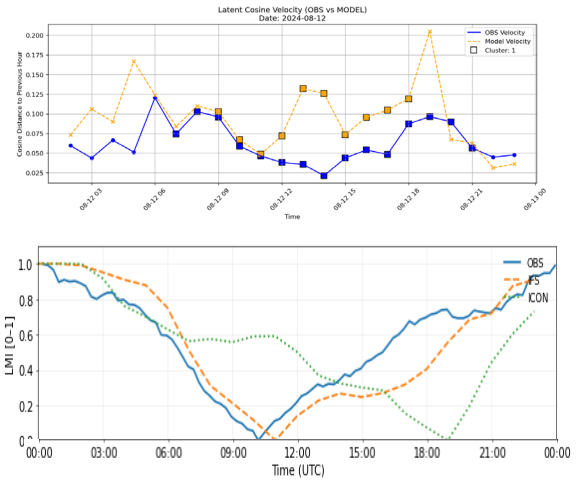
For the first time, this work demonstrates that SSL latent spaces derived from MSG observations reveal systematic limitations when mapped onto km-scale climate model simulations.



1) Physical Characteristics



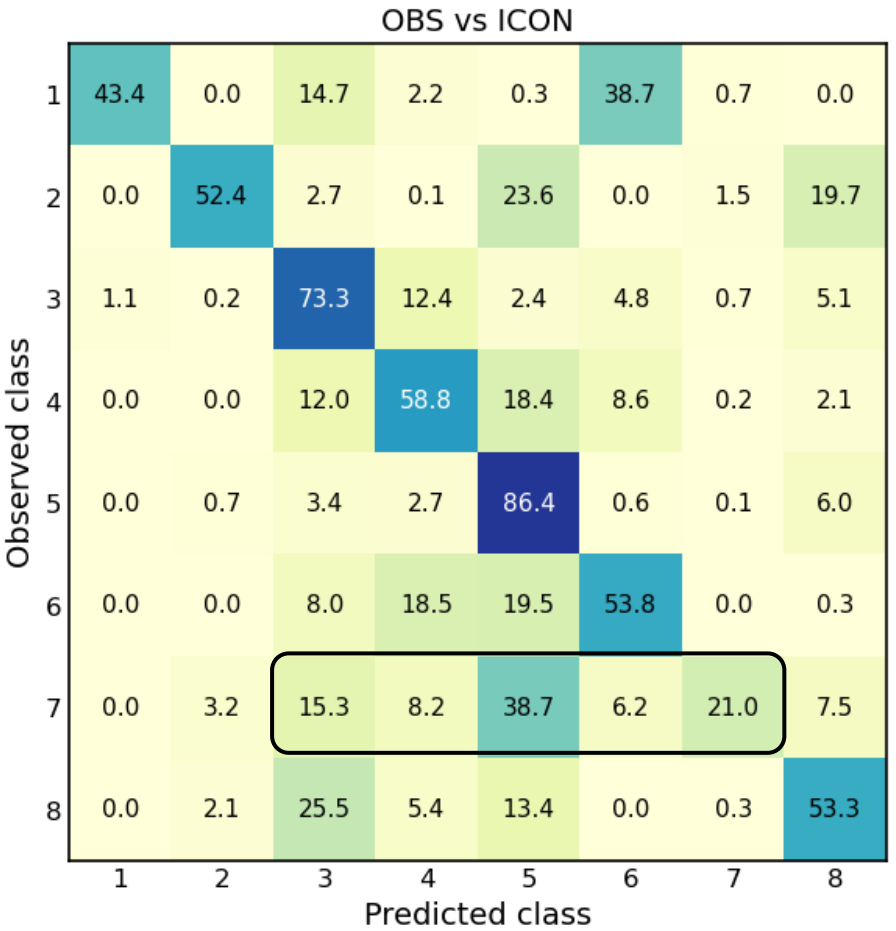
2) Evolution tendencies

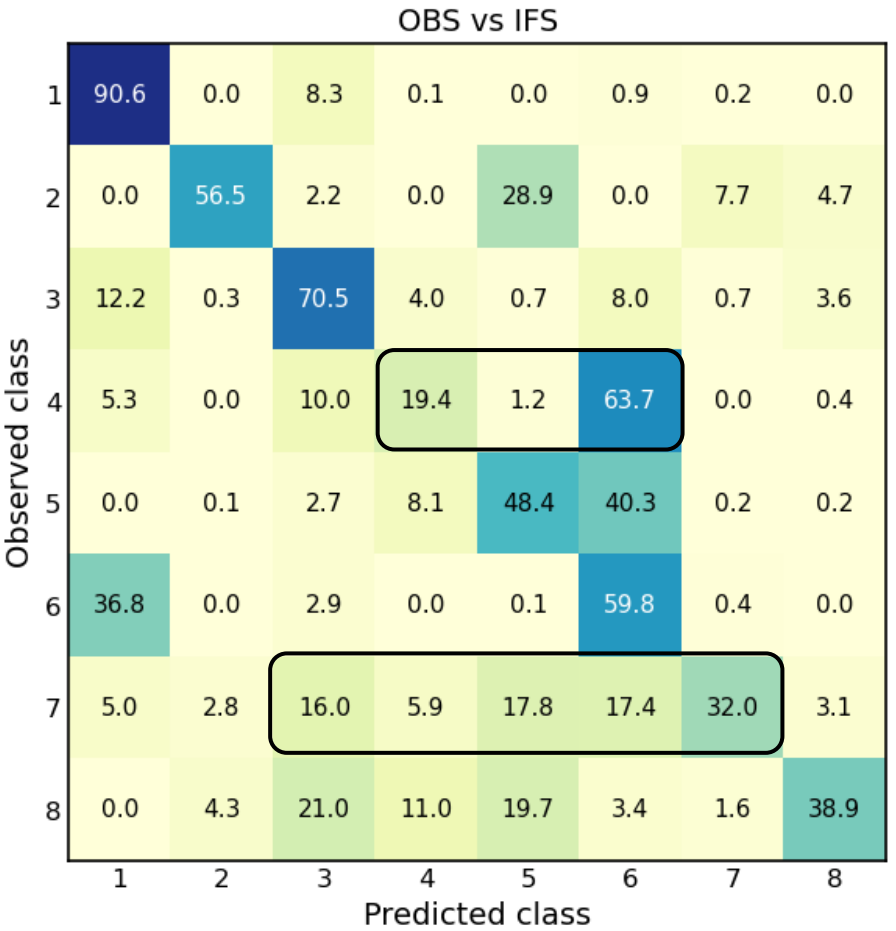


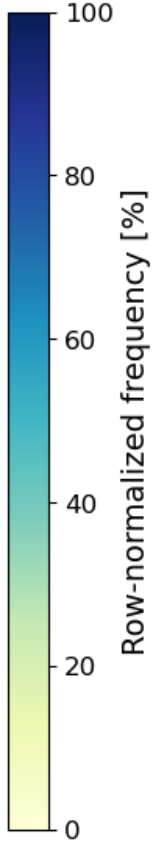
3) Latent velocity & maturity index

EXTRA SLIDES

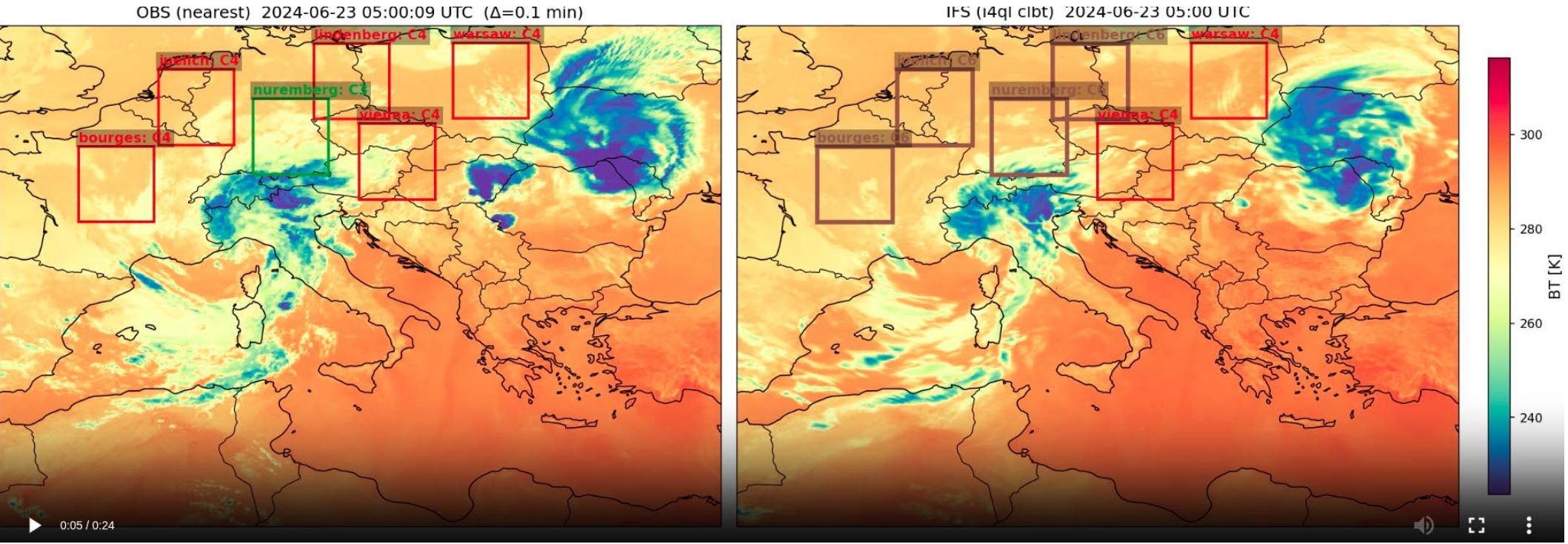
Confusion Matrix







Example class from Confusion Matrix



Final effective systems