Development of Ocean Atmosphere Foundation Models

flowershift

justin.freeman@flowershift.com.au

flowershift

Foundation models are machine learning models trained on large datasets that can be adapted to many different tasks.

flowershift

Data → Large and diverse training data

Model → General purpose representations

Task → Adaptable to many applications

flowershift

Training data → ERA5[1], APS3 ACCESS-G[2], BRAN2020[3]

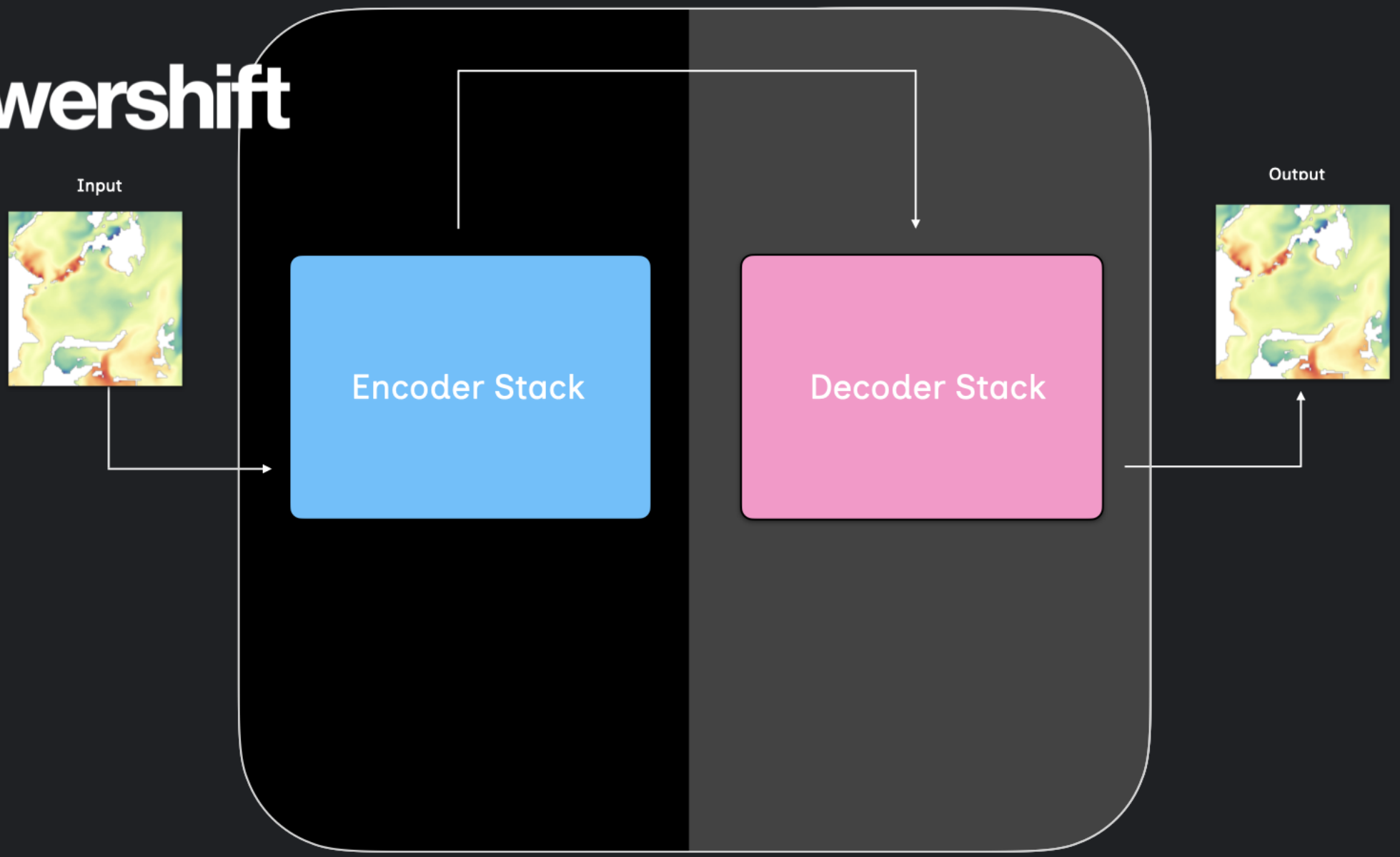
Coverage → Global, multi-variable, multi-level

Types → 2D and 3D fields

1 https://cds.climate.copernicus.eu/stac-browser/collections/reanalysis-era5-complete?.language=en
2 https://geonetwork.nci.org.au/geonetwork/srv/eng/catalog.search#/metadata/f3307_5503_1483_3079
3 https://geonetwork.nci.org.au/geonetwork/srv/eng/catalog.search#/metadata/f9372_7752_2015_3718

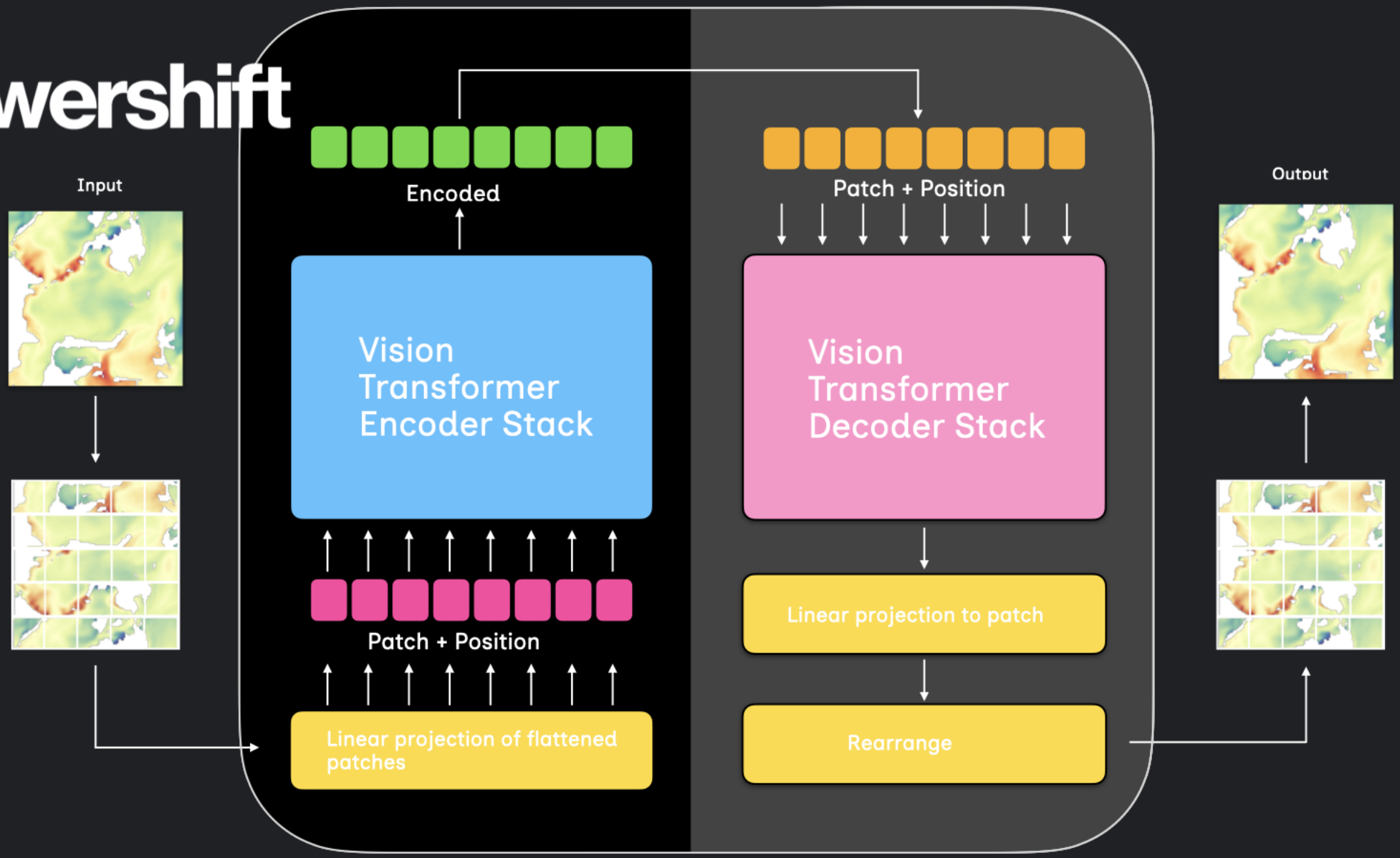
flowershift

Pre-training



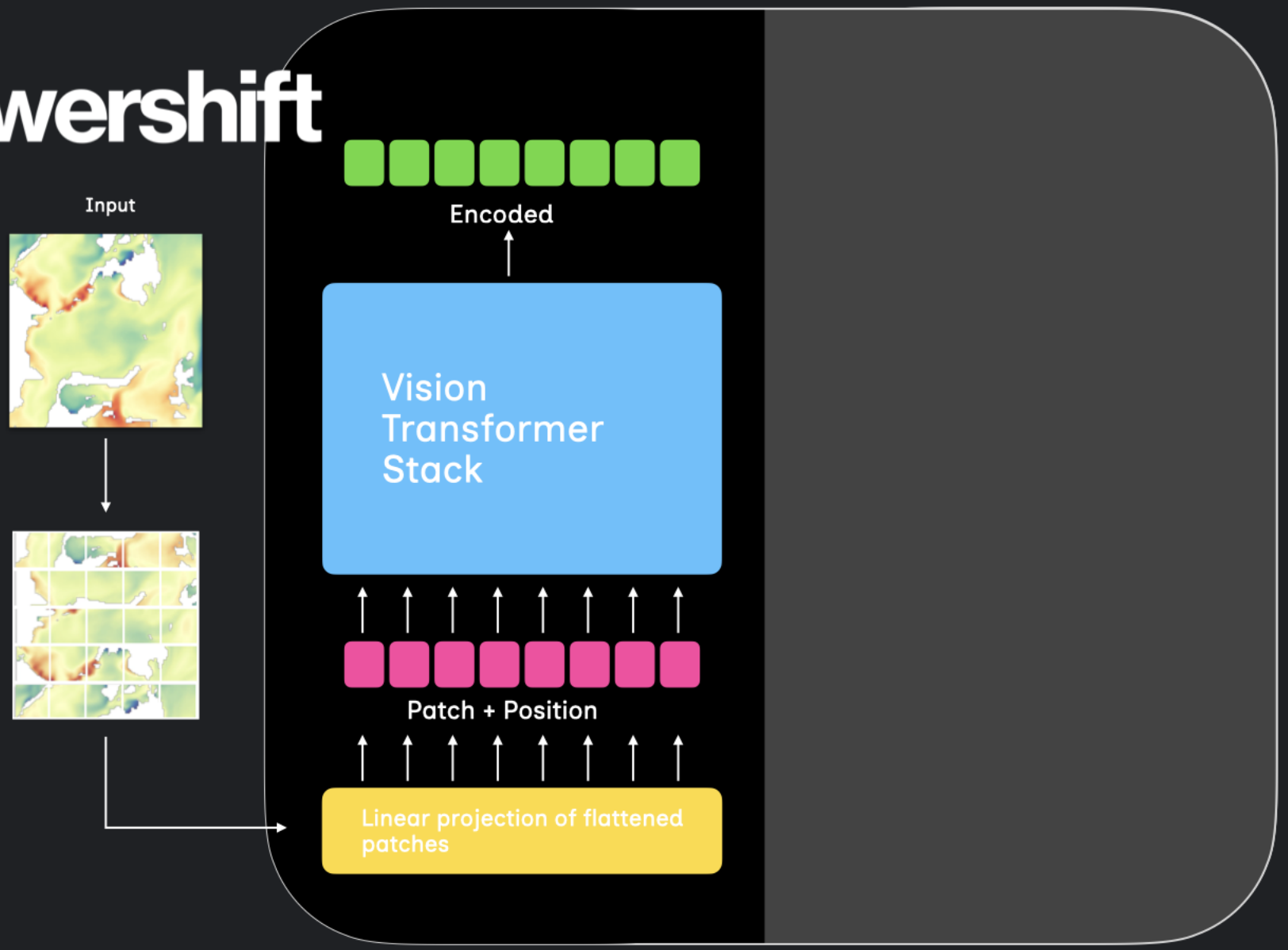
flowershift

Pre-training



flowershift

Trained




Foundation Model Applications
flowershift

flowershift

Patch Similarity

→ Reference patch shown in red

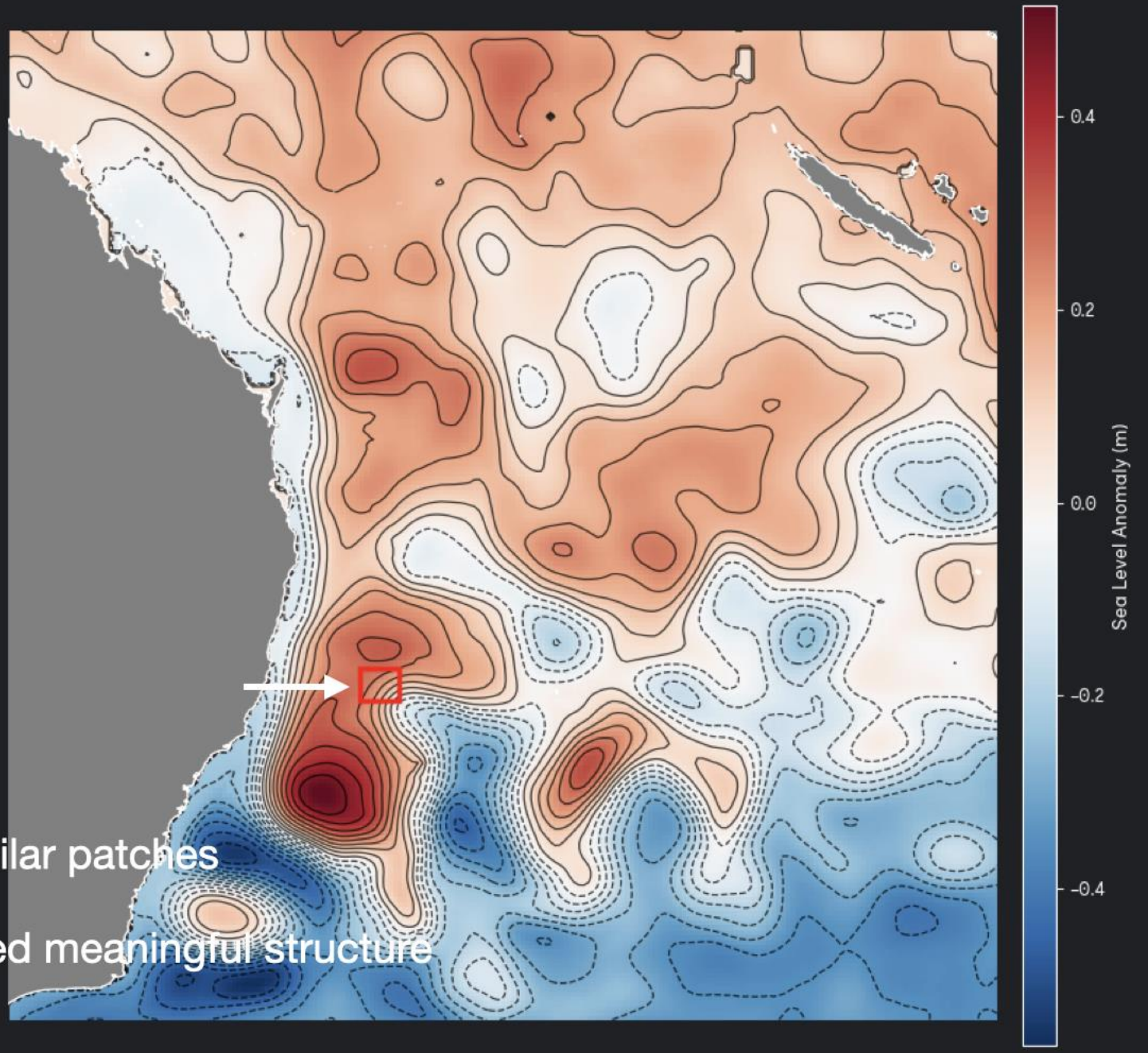
→ Sliding window over the data

→ Embedding computed for every patch

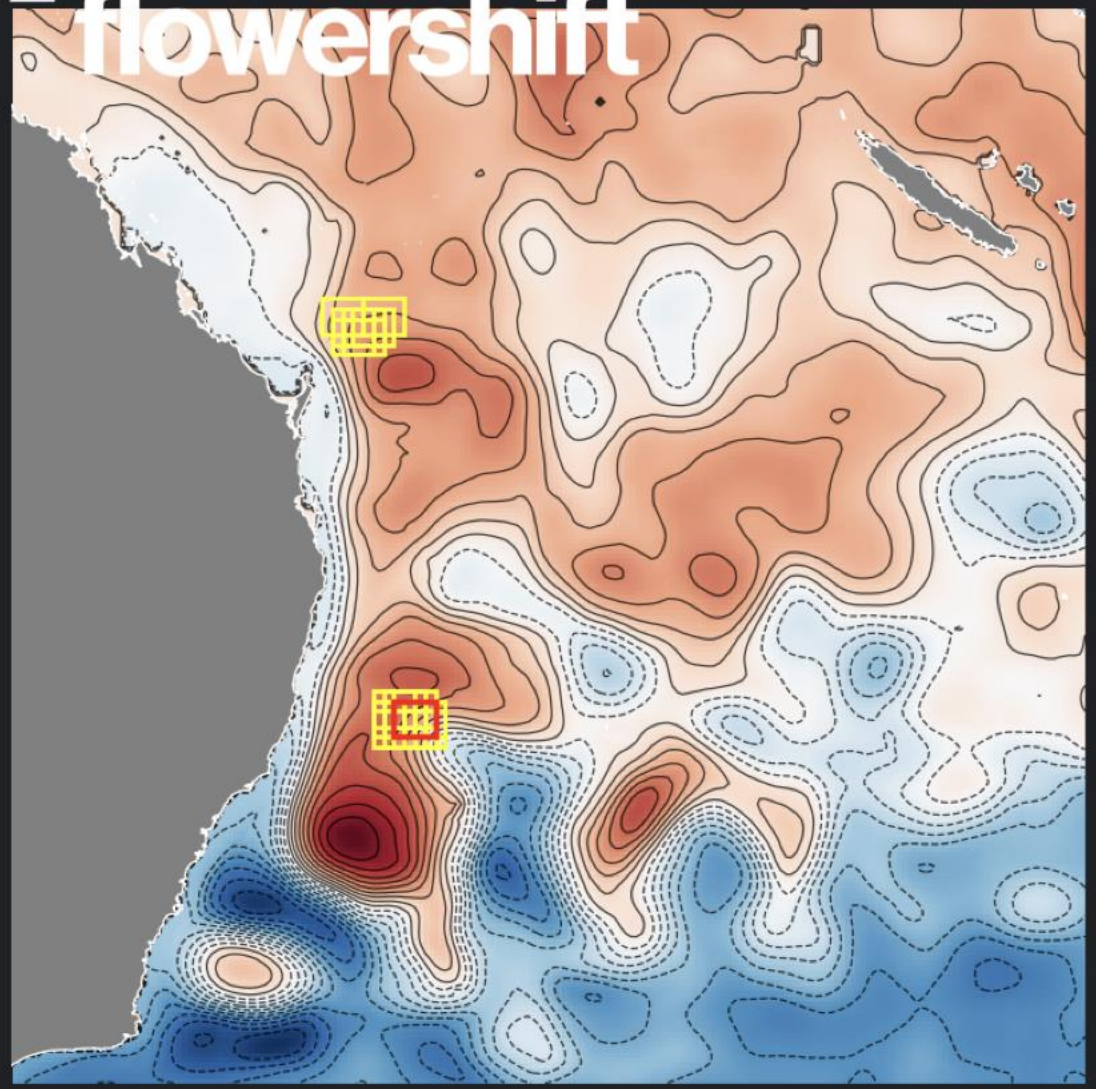
→ Cosine similarity used to rank matches

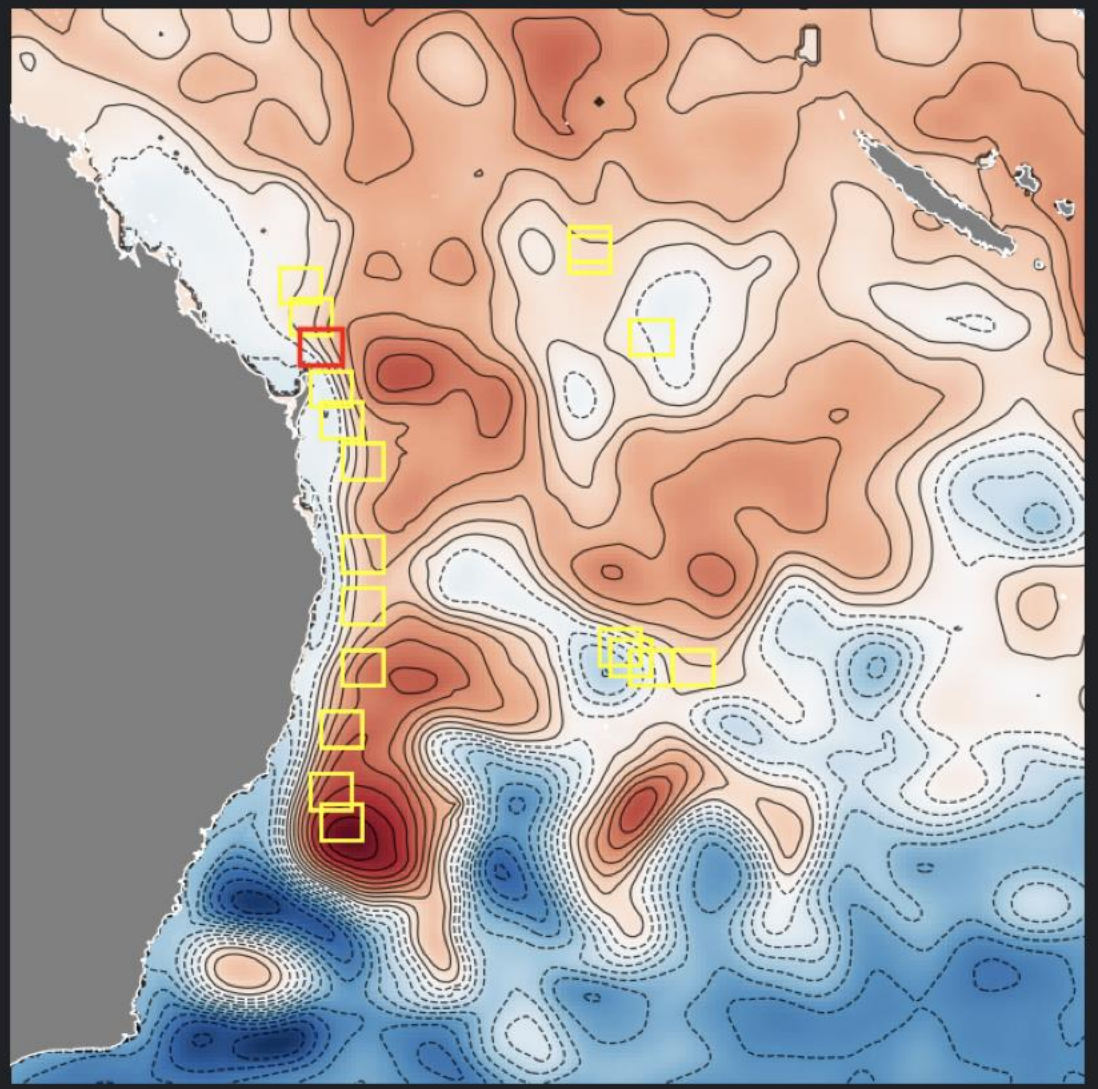
→ Yellow boxes show the top 20 most similar patches

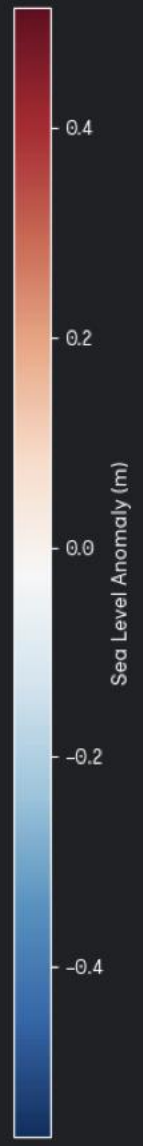
→ Demonstrates that the model has learned meaningful structure



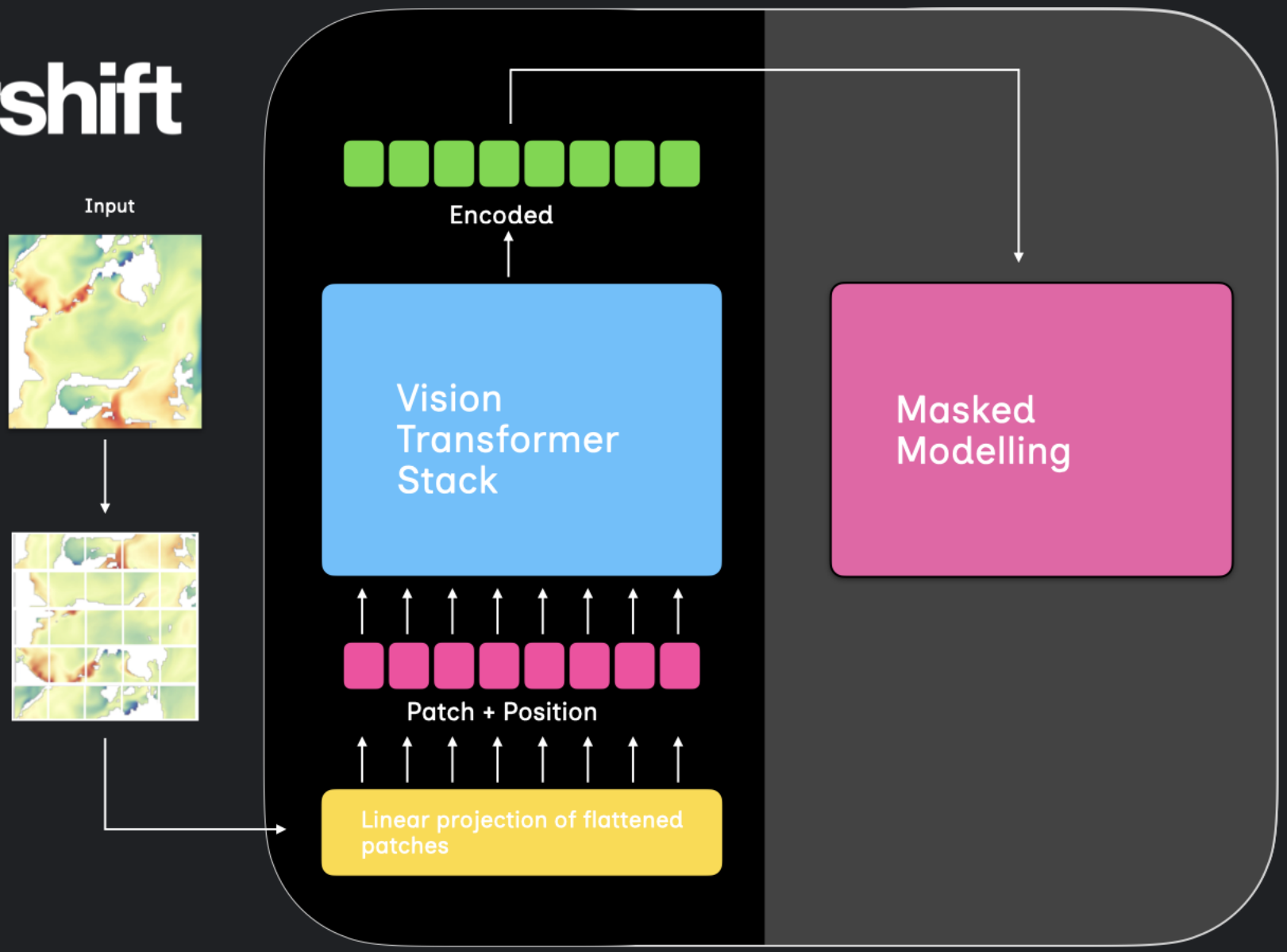
flowershift







flowershift



flowershift

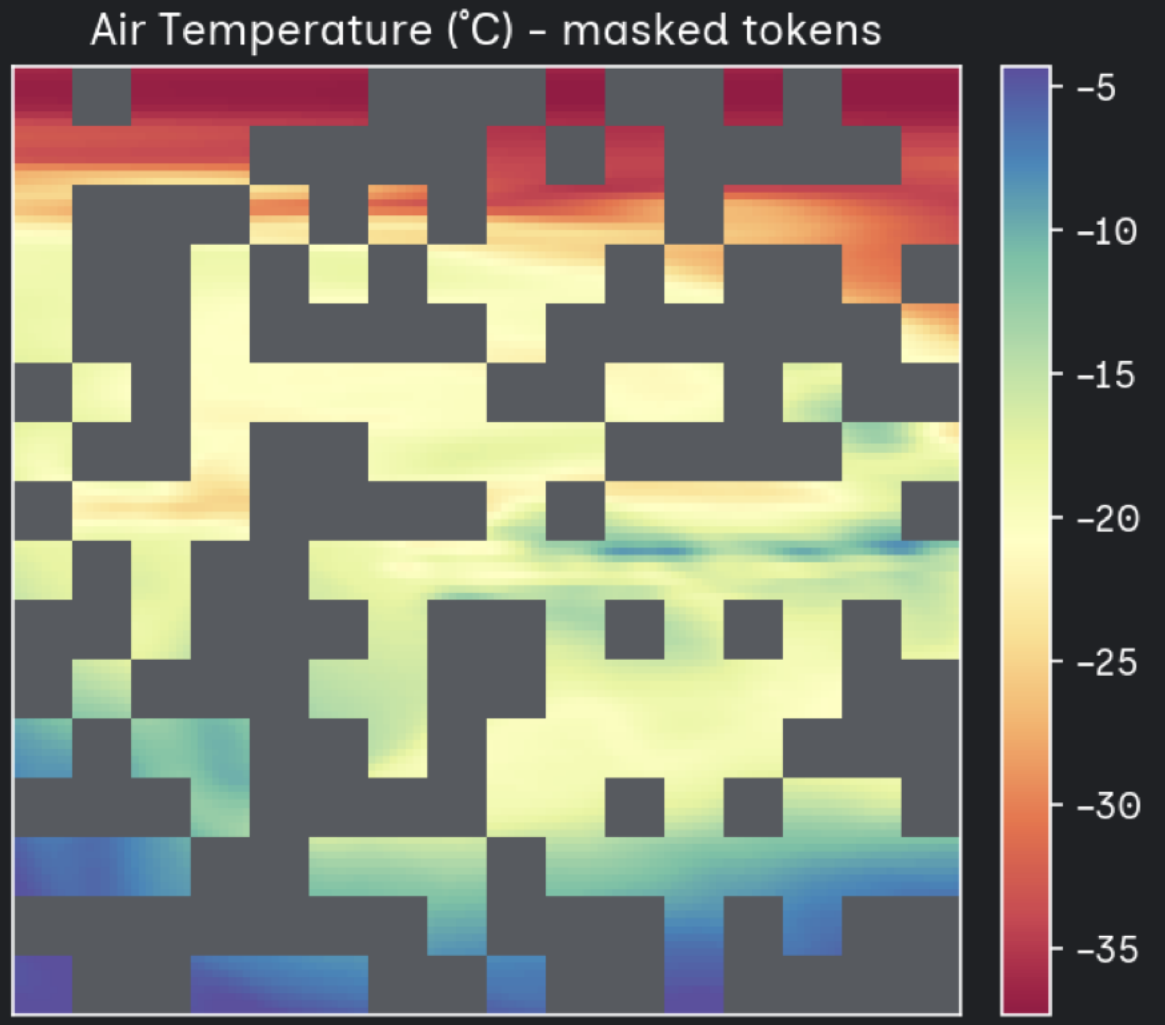
Masked Modelling

→ Mask out part of the field

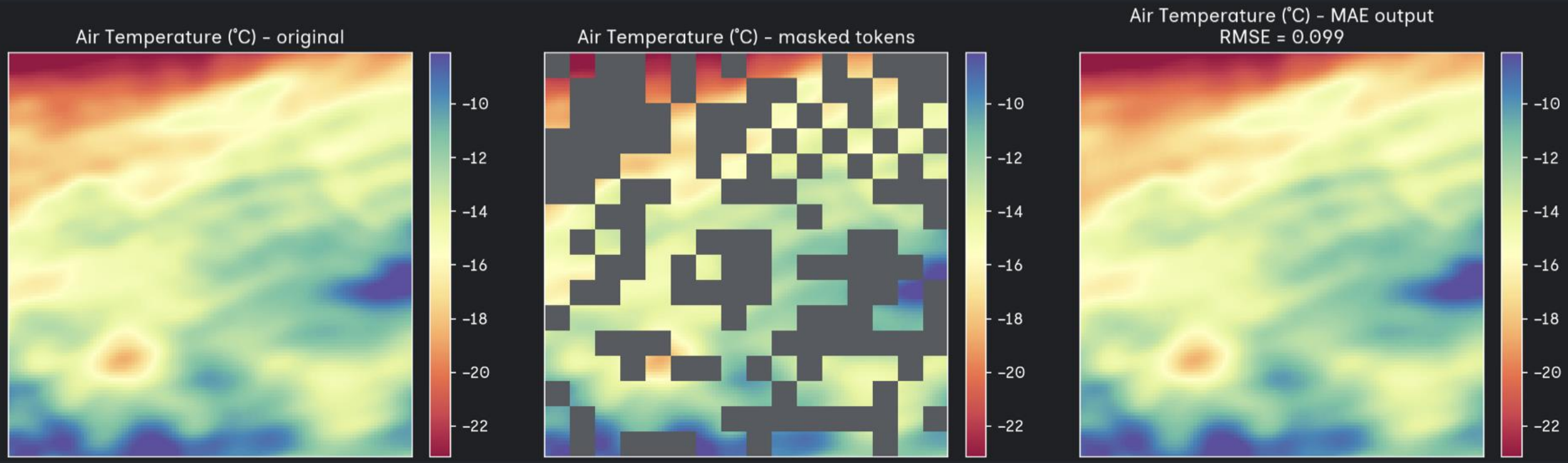
→ Train a small task-specific head on top of the foundation model

→ The head learns to reconstruct the missing region using the foundation model's learned features

→ Demonstrates that the model has learned meaningful ocean-atmosphere structure

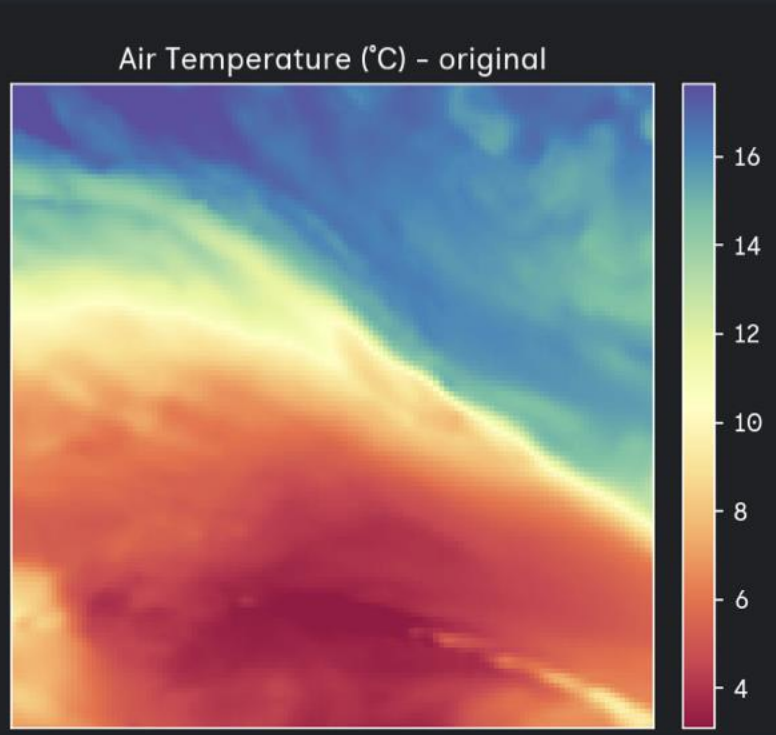


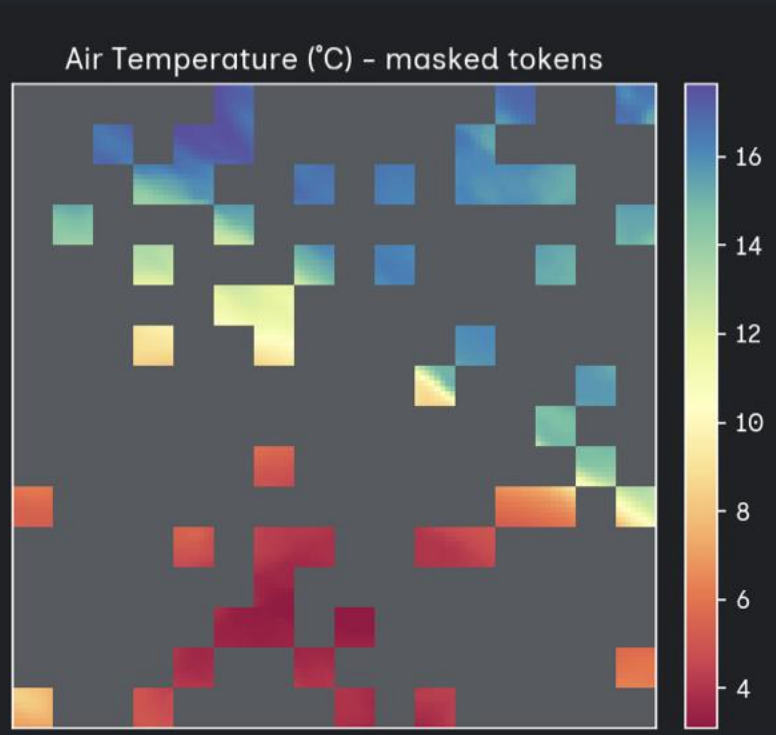
flowershift

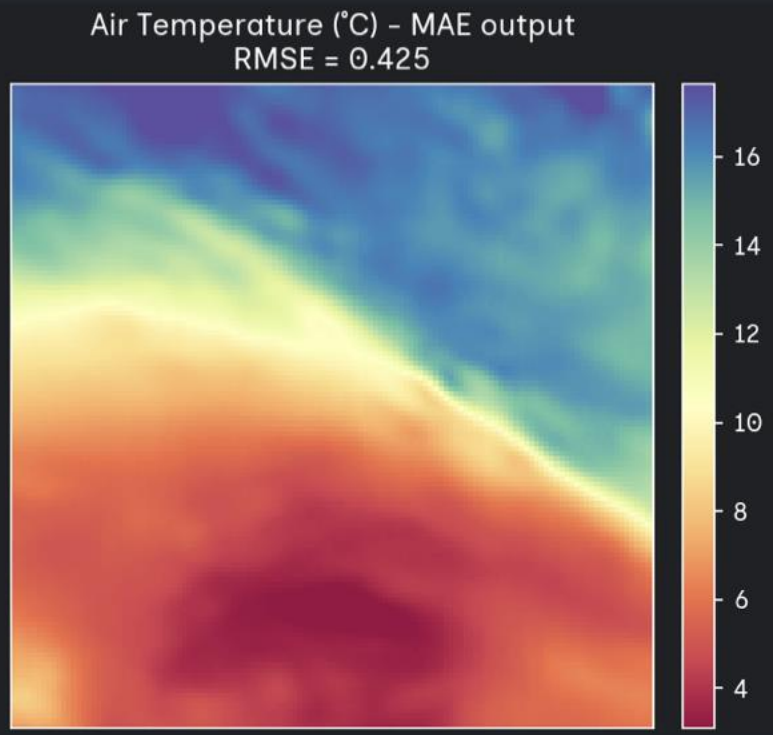


50% Masked

flowershift

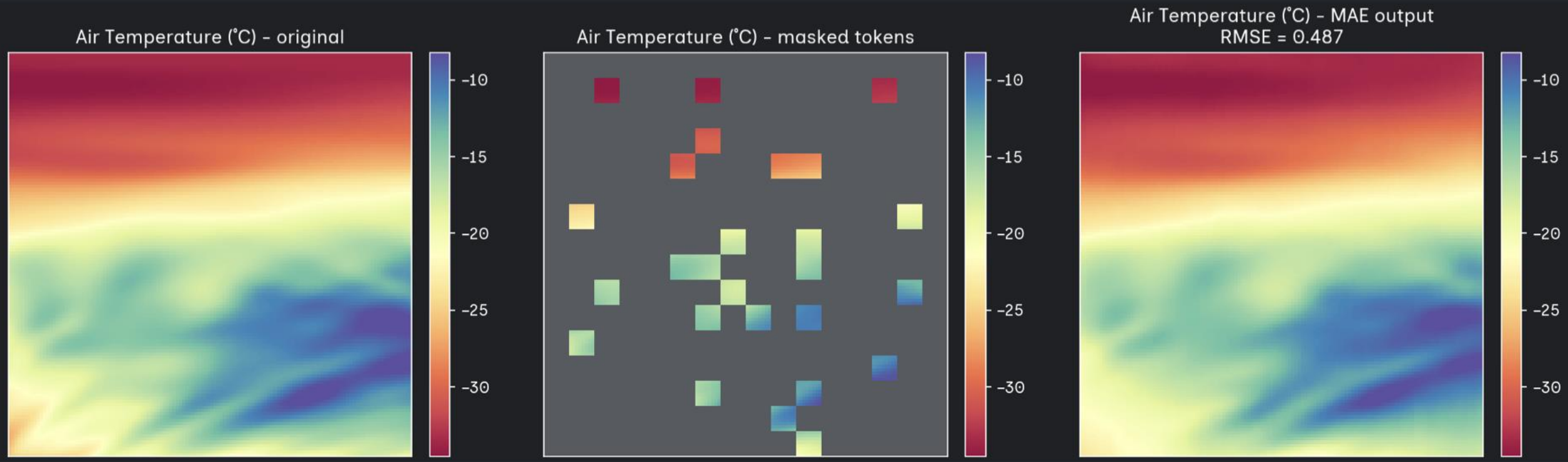






80% Masked

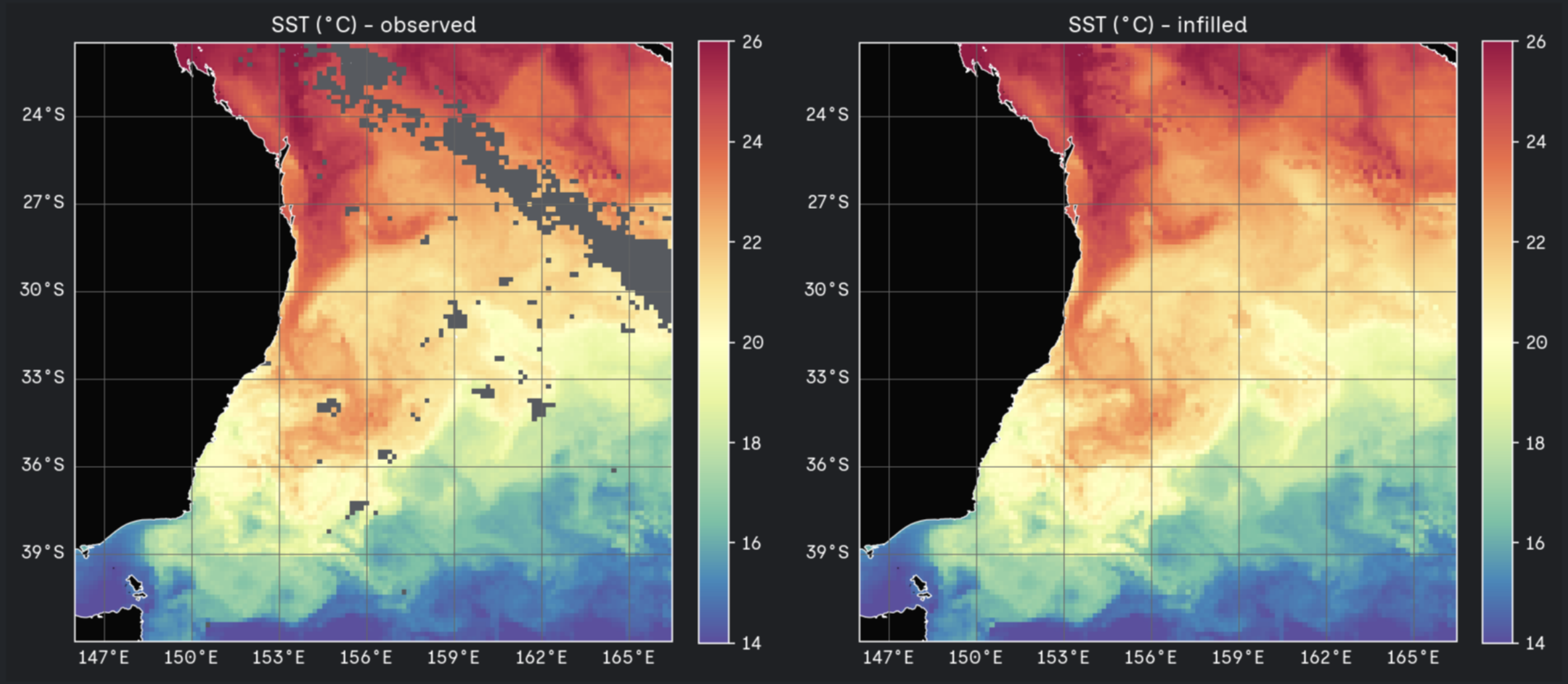
flowershift



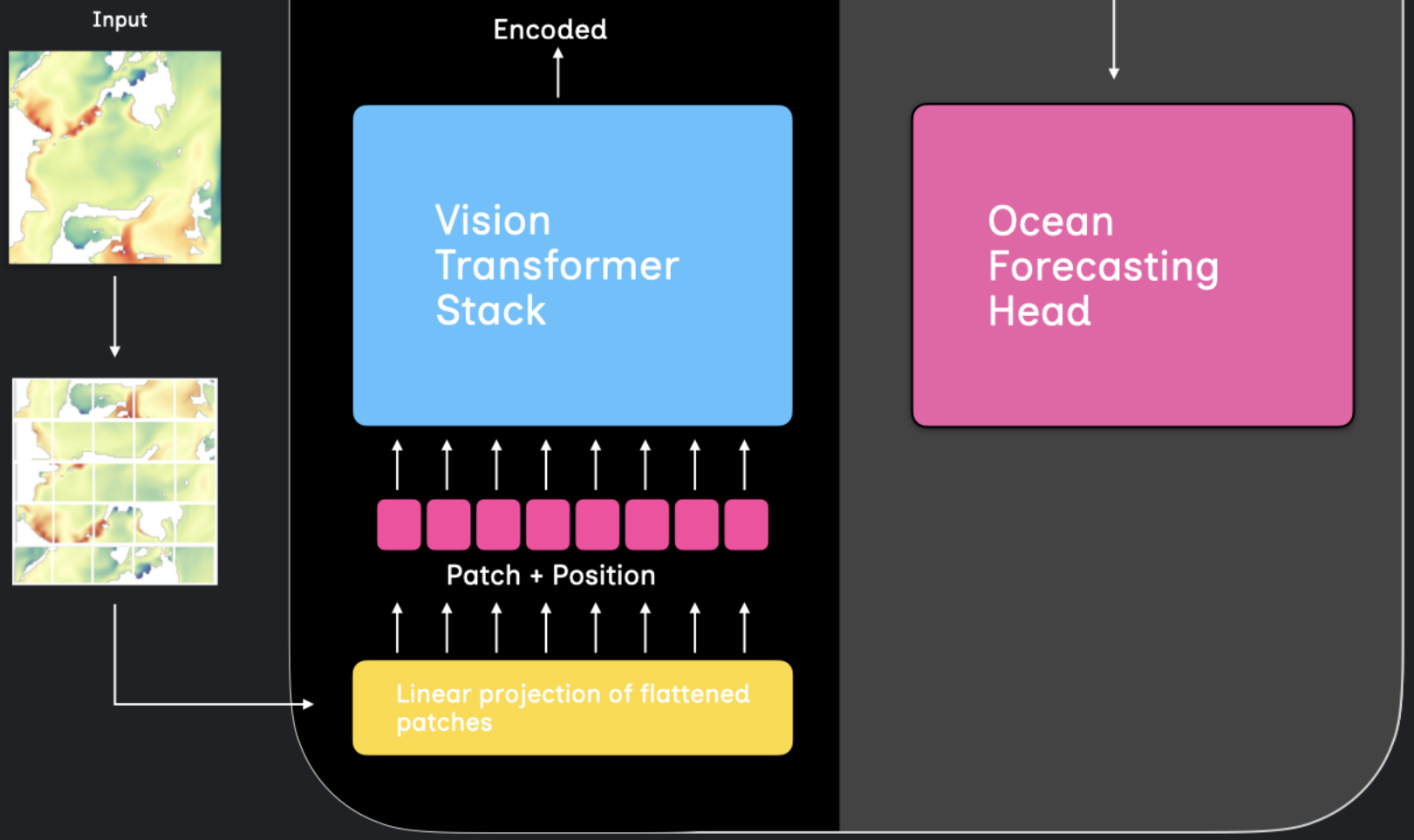
90% Masked

flowershift

Skin sea surface temperature retrievals from the VIIRS sensor on the N20 satellite, composited over multiple swaths and gridded over a 0.02 degree rectangular grid over the Australian domain.



flowershift



flowershift

Ocean Forecasting

Inputs → 25 ocean variables

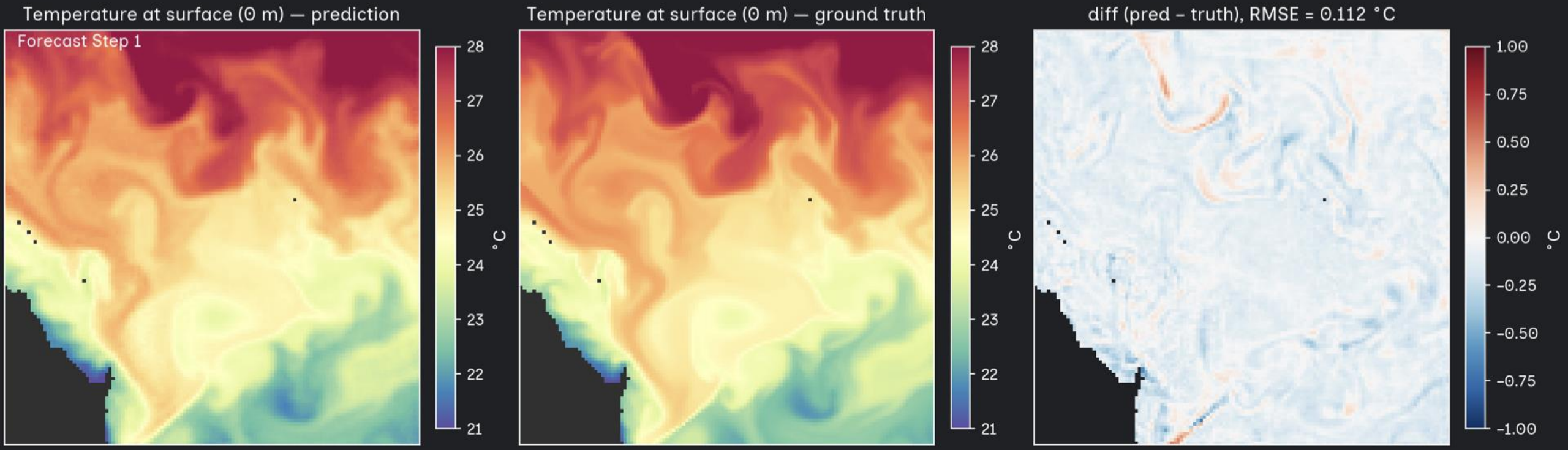
T, S, U, V at 0-1000 m depth + SSH

Output → Next day fields

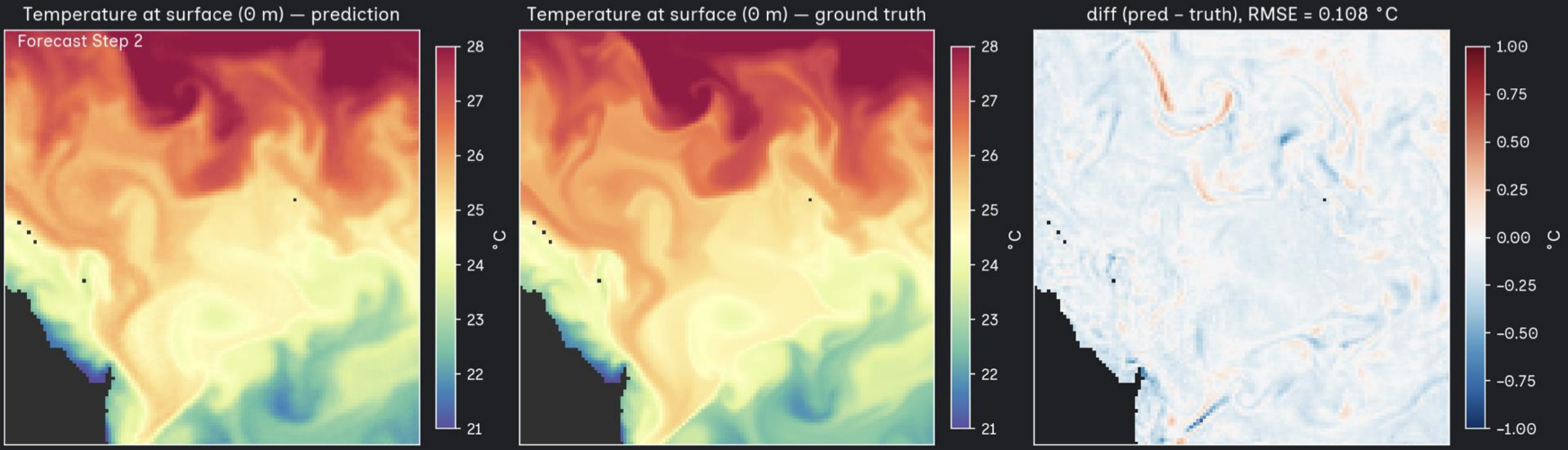
Forecast head trained on top of the foundation model

Autoregressive rollout for multi-day forecasts

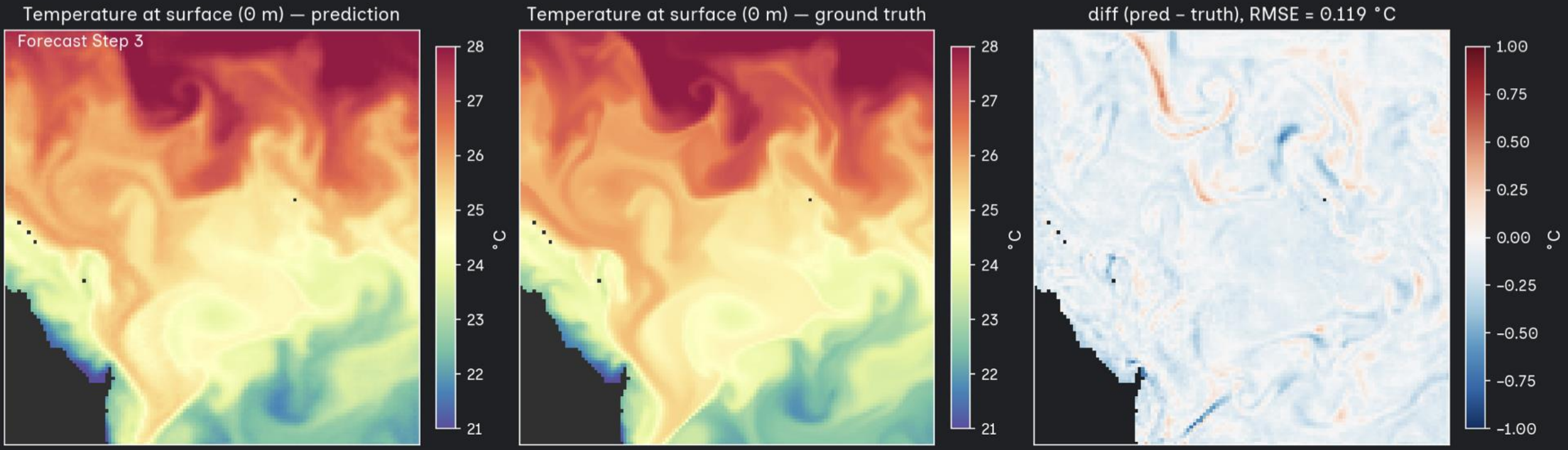
flowershift



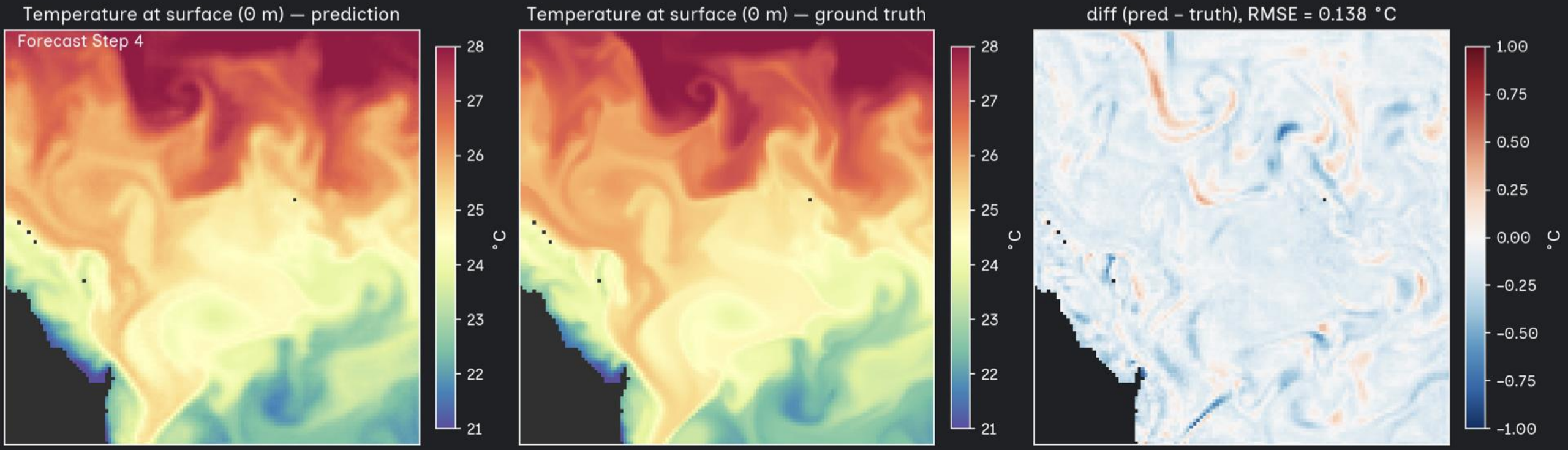
flowershift



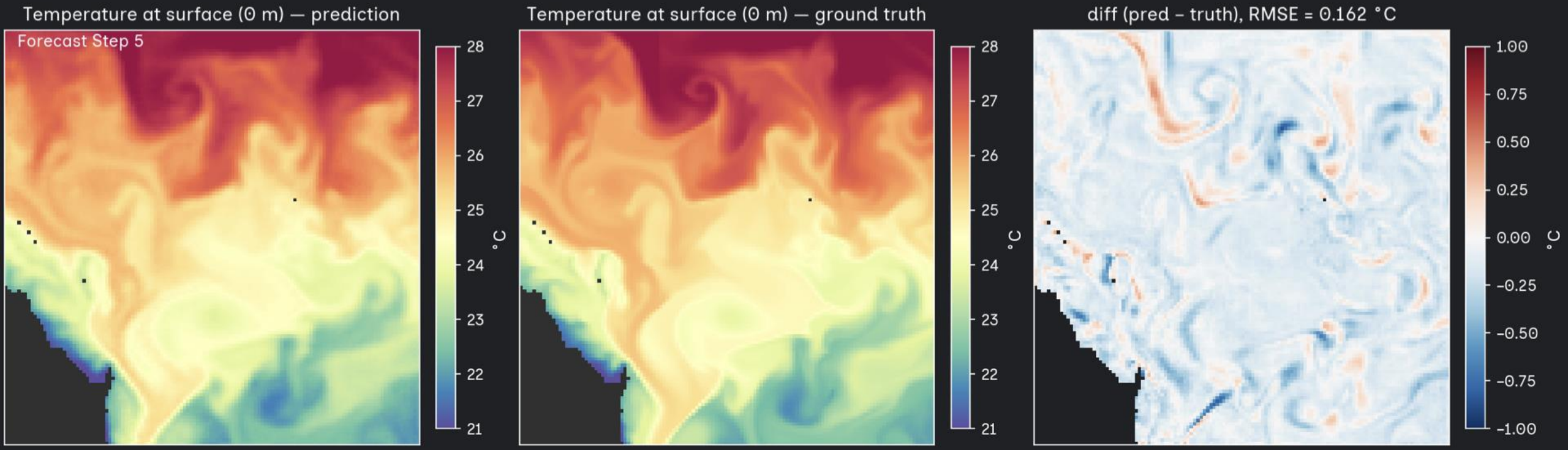
flowershift



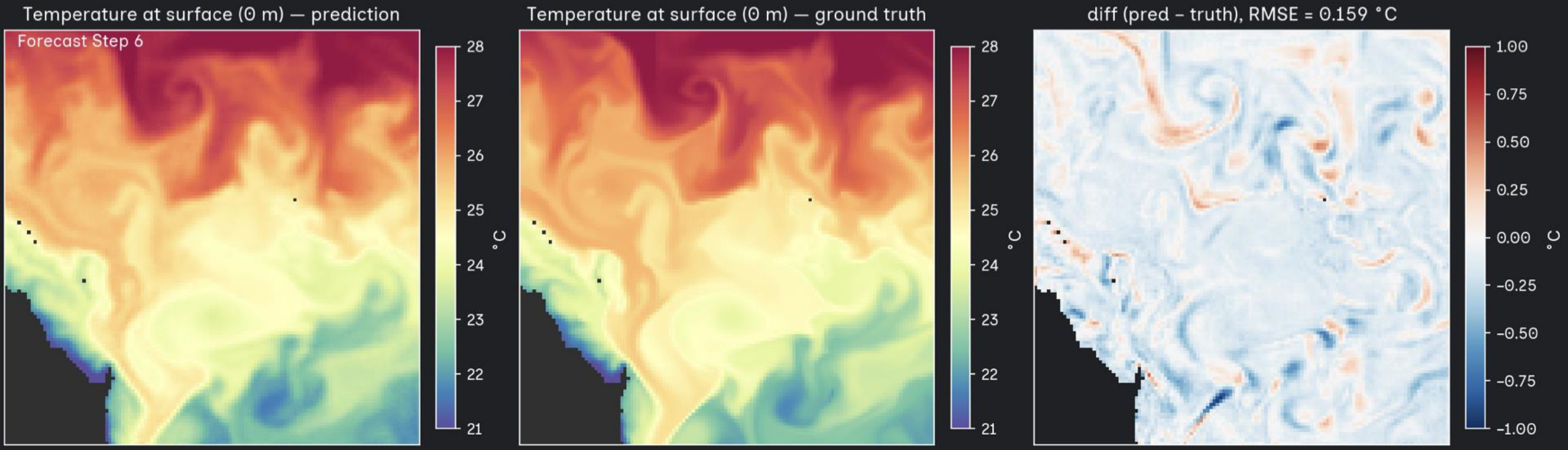
flowershift



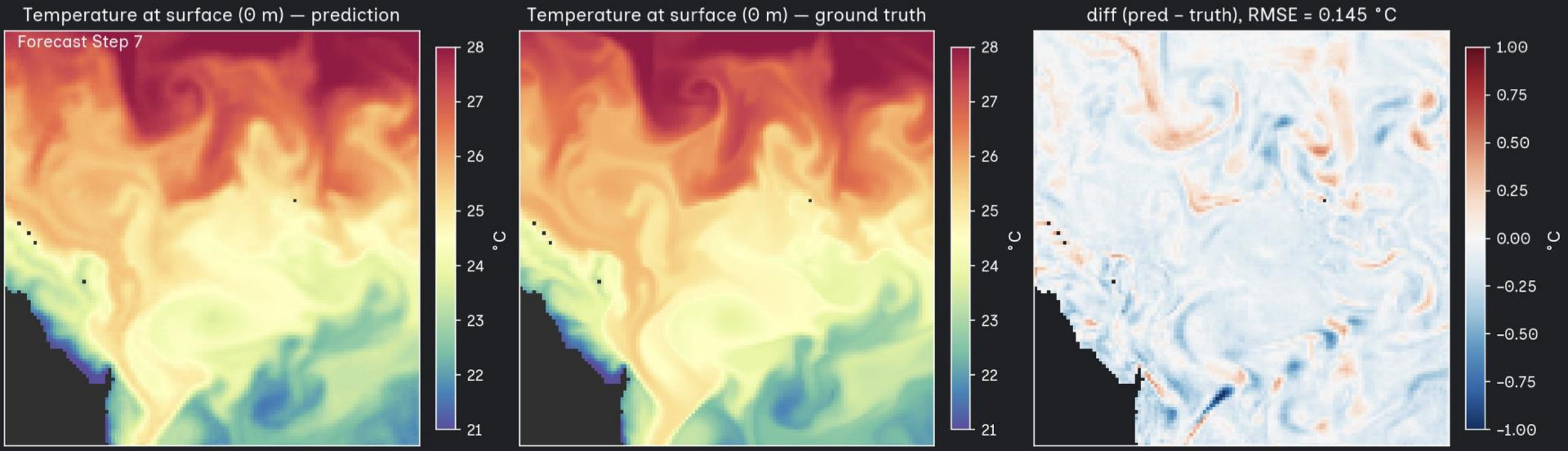
flowershift



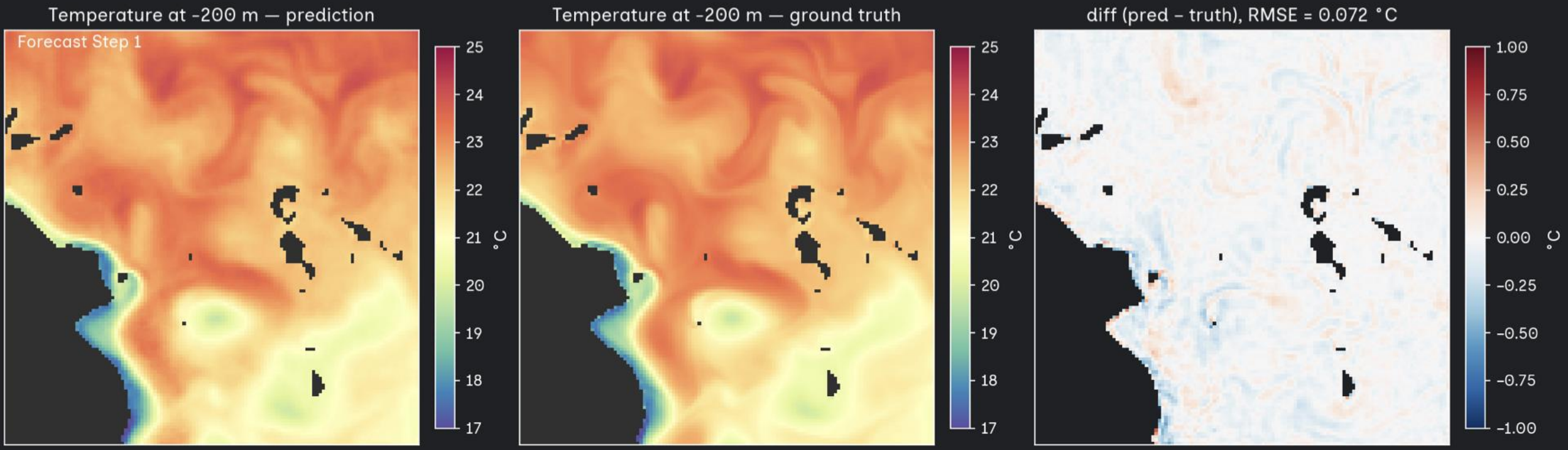
flowershift



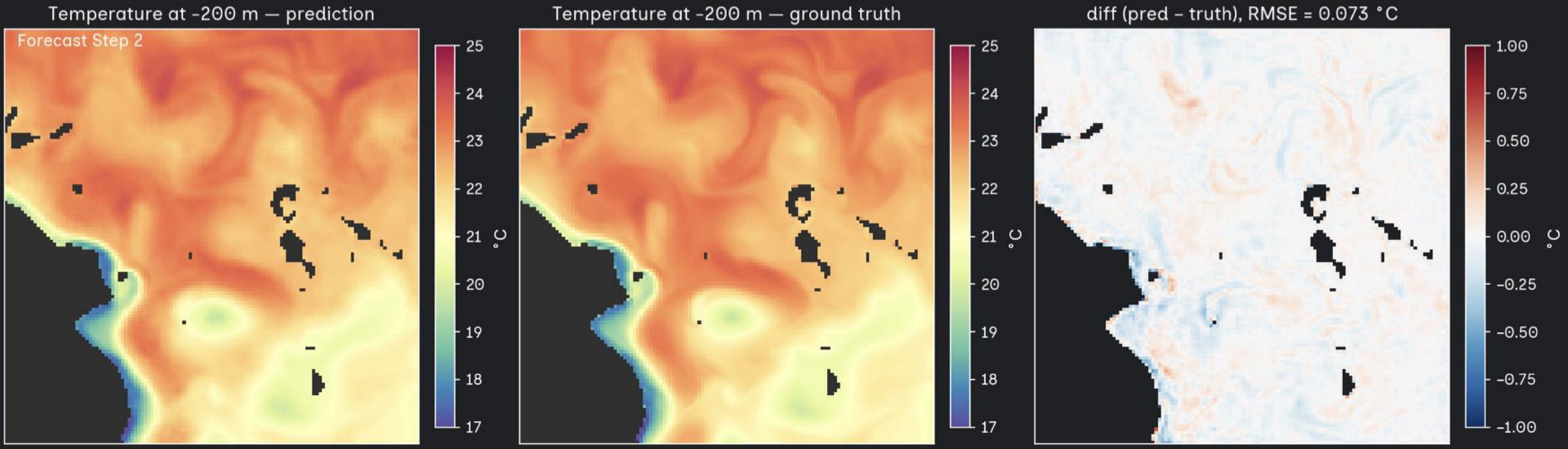
flowershift



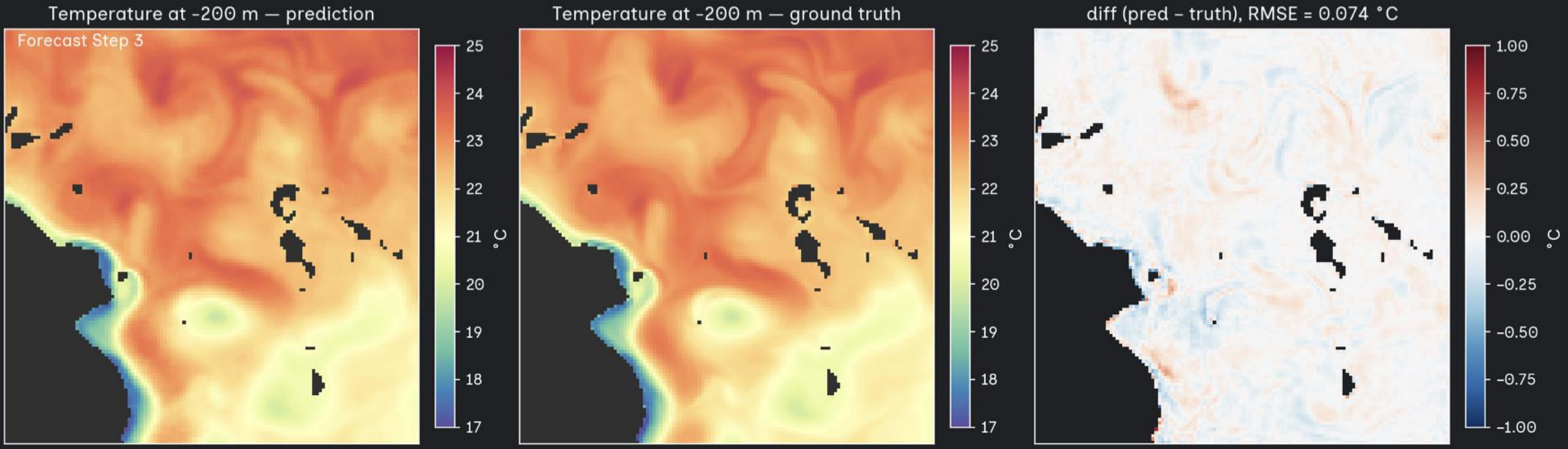
flowershift



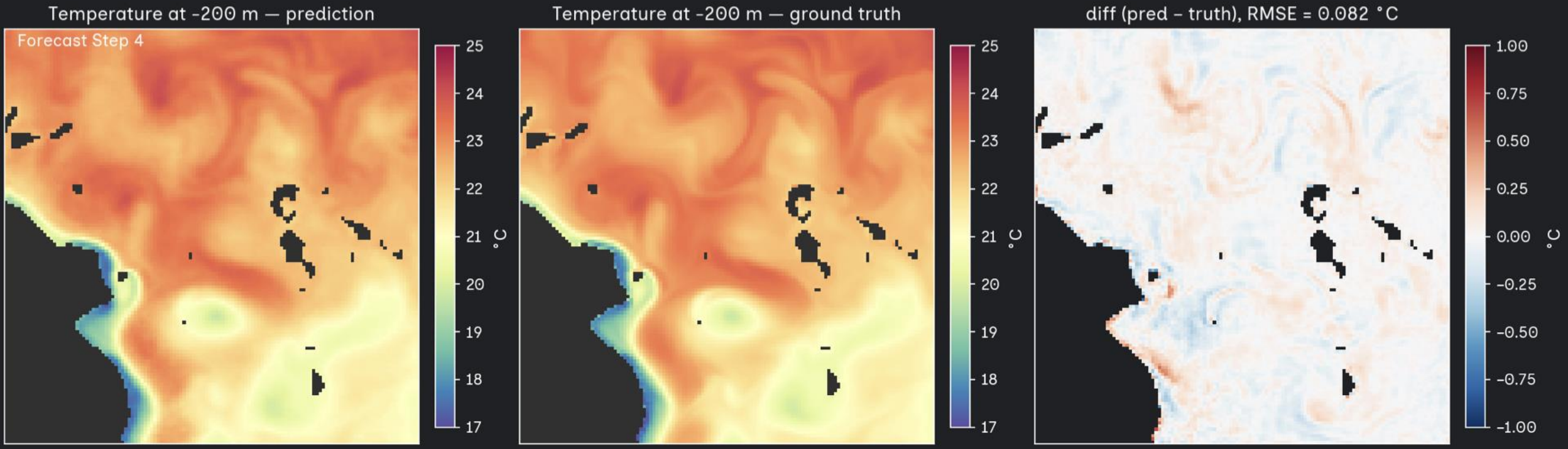
flowershift



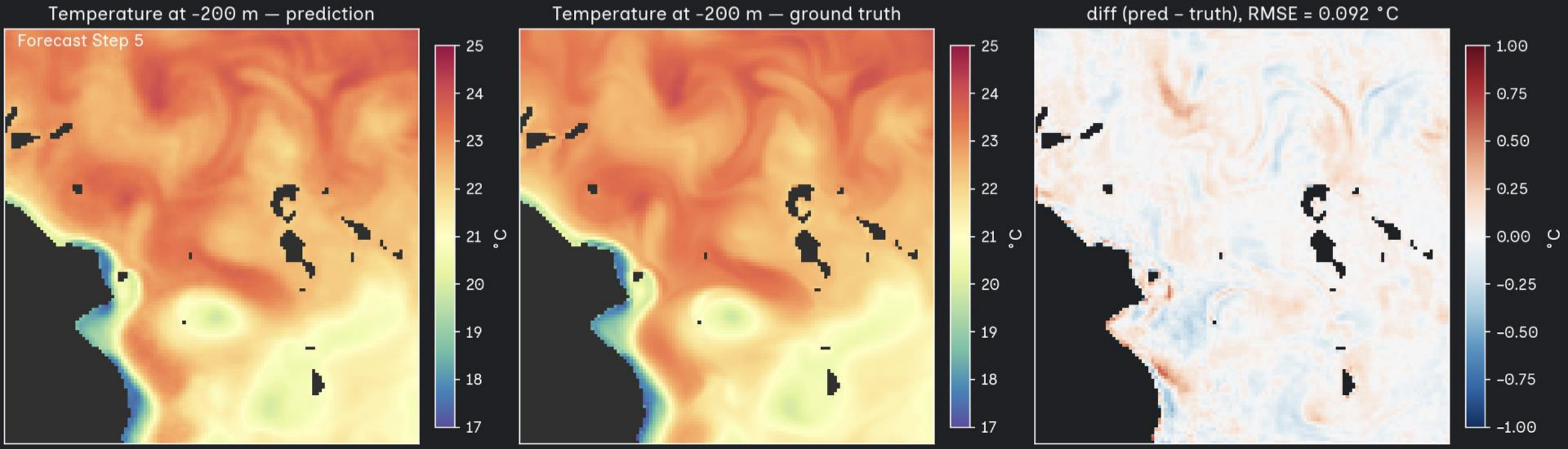
flowershift



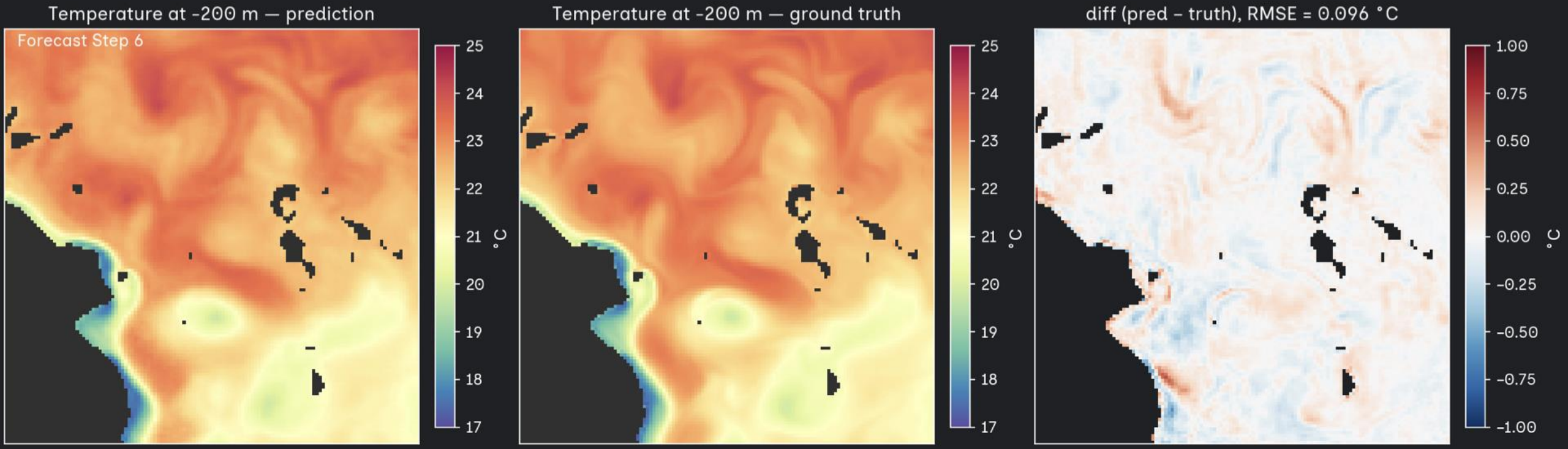
flowershift



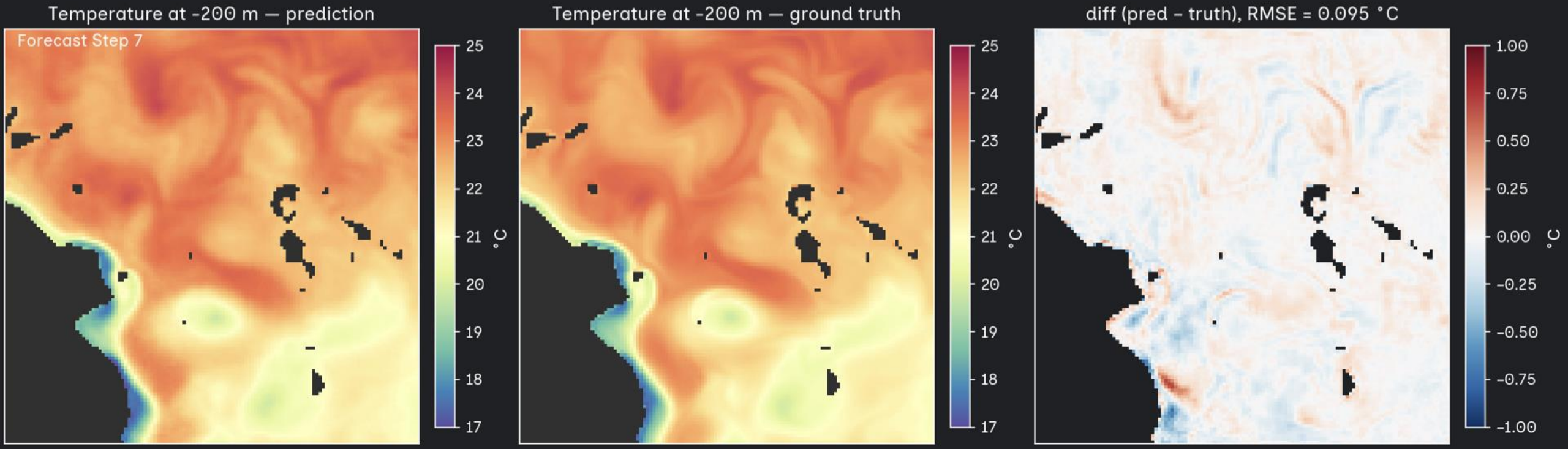
flowershift



flowershift



flowershift



flowershift

Foundation models learn general purpose representations from large, diverse datasets.

The same pretrained model can be adapted to support many tasks.

The flowershift foundation models were trained only on model data.

→ The opportunity is combining our observations and models into a shared capability.

flowershift

connect@flowershift.com.au