Atmospheric Blocking Pattern Recognition in Global Climate Model Simulation Data

Grzegorz Muszynski[1,2], Prabhat[2], Jan Balewski[2],
Karthik Kashinath[2], Michael Wehner[2], Vitaliy Kurlin[1]

1) Department of Computer Science, University of Liverpool, Liverpool, United Kingdom
2) Lawrence Berkeley National Laboratory, Berkeley, United States

National Energy Research
Scientific Computing Center

BERKELEY LAB

Paper ID: 368

Milan, Italy 10-15 January 2021

Introduction and Scientific Motivation

- We address a problem of **Atmospheric Blocking (AB)** pattern recognition in global climate datasets, such as climate reanalyses and climate model simulations.
- ABs are quasi-stationary pressure systems that cause abrupt changes in the westerly air circulation. ABs often lead to extremes weather events, such as heat waves and cold snaps in Europe, North America, and Australia.
- The main objectives are the classification and localisation of AB patterns in mid-latitude regions.

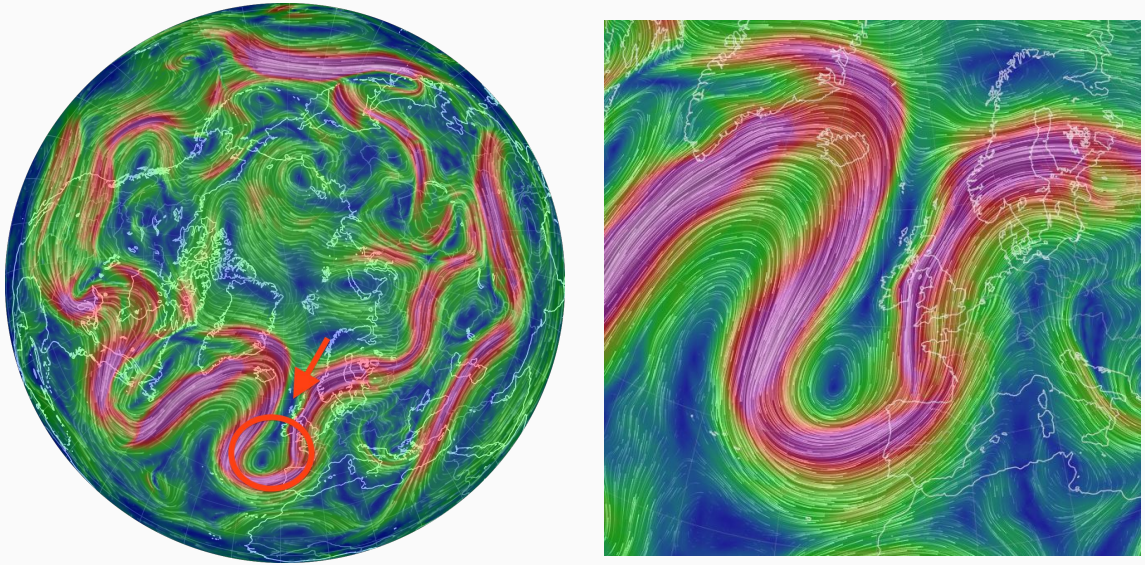

Fig. 1. Left: an example of Atmospheric Blocking (AB) pattern over the Atlantic European coast. **Right**: a close-up of the AB pattern. It is shown the wind speed field (m/s). Source: the US NWS data visualised by http://earth.nullschool.net.

Data: Climate Reanalysis Dataset

- **ERA-Interim reanalysis** of the European Center for Medium-Range Weather Forecasting, and binary masks (i.e., AB or non-AB labels) are provided by the Institute for Atmospheric and Climate Science at the ETH Zurich, Switzerland.
- We use **temperature**, **meridional** and **zonal wind speed**, **geopotential height**, and **potential vorticity** variables at **eight pressure levels** (150 - 500 millibars), **6-hourly** at approx. **80 km spatial resolution** in the period 1979-2018.
- Each geographic region of interest is roughly centered over a local maximum of blocking frequency and has a size of 60 px x 120 px x 40 channels.

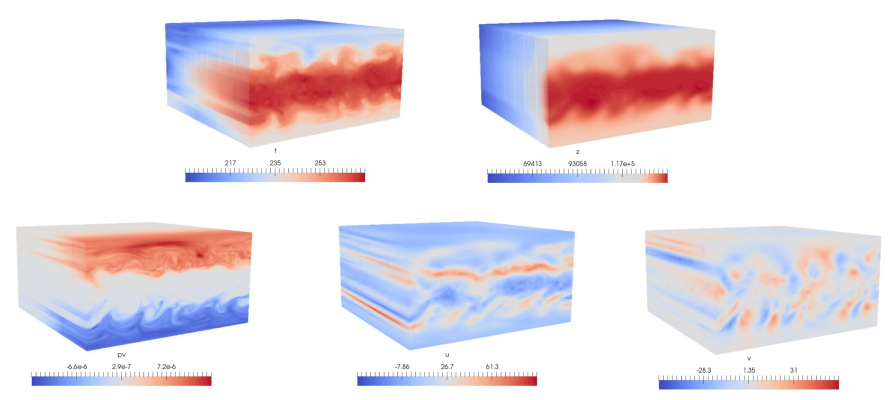

Fig. 2. Shown are examples of global images of five physical variables in ERA-Interim reanalysis product. **Top row**: t - temperature; z - geopotential height. **Bottom row**: pv- potential vorticity; u - zonal velocity; v- meridional velocity.

Fig. 3. An example of the globe with six regions used in this study: the North Pacific Region (**NP**); the North Atlantic Region (**NA**); the North Continental Region (**NC**); the South Pacific Region (**SP**); the South Atlantic Region (**SA**), and the South Indian Ocean Region (**SI**).

Methodology: a hierarchical pattern recognition method based on Convolutional Neural Network (CNN) architectures

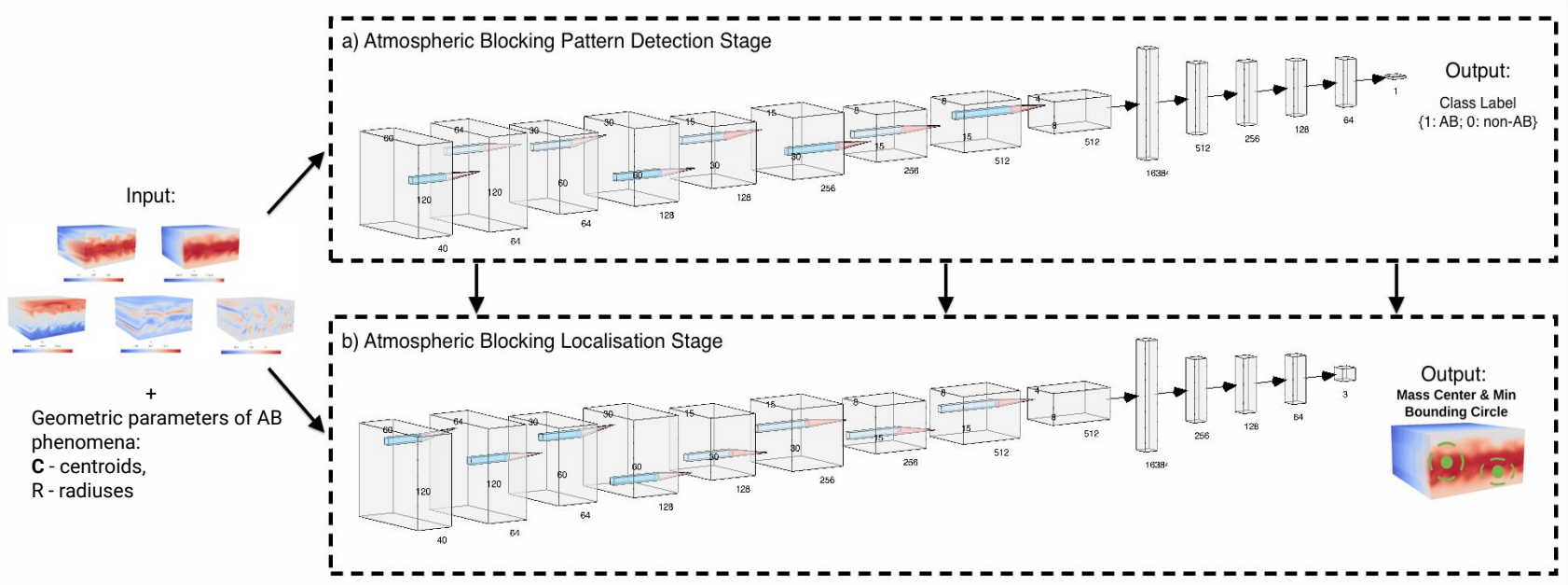

Fig. 4. The big picture of a hierarchical AB pattern recognition method. The method consists of two stages: (a) illustrates architecture of CNN-based classifier that distinguishes an atmospheric blocking events (AB label) and otherwise non-blocking events (non-AB label); and (b) illustrates architecture of CNN-based regressor that predicts three values describing the AB location, i.e. a latitudinal position, a longitudinal position, and a radius of a minimum enclosing circular box.

Methodology: pattern recognition method based on convolutional neural network (CNN) architectures

- Both stages employ customised CNN architectures inspired by VGG architectures developed in Visual Geometry Group
- We investigate five different architectures of the generic CNN (A) designed for each stage of the method, as outlined in Table Ⅰ and Table Ⅱ.

Stage 1: classification

CNN Architectures				
A	B	C	D	E
Input: (60 × 120 × 40)				
conv-64	conv-64 conv-64	conv-64 conv-64	conv-64 conv-64	conv-64 conv-64
Maxpooling				
conv-128	conv-128	conv-128 conv-128	conv-128 conv-128	conv-128 conv-128
Maxpooling				
conv-256	conv-256	conv-256	conv-256 conv-256	conv-256 conv-256
Maxpooling				
conv-512	conv-512	conv-512	conv-512	conv-512 conv-512
Maxpooling				
FC-512				
FC-256				
FC-128				
FC-64				
Output: FC-1				

Stage 2: localisation

CNN Architectures				
A	B	C	D	E
Input: (60 × 120 × 40)				
conv-64	conv-64 conv-64	conv-64 conv-64	conv-64 conv-64	conv-64 conv-64
Maxpooling				
conv-128	conv-128	conv-128 conv-128	conv-128 conv-128	conv-128 conv-128
Maxpooling				
conv-256	conv-256	conv-256	conv-256 conv-256	conv-256 conv-256
Maxpooling				
conv-512	conv-512	conv-512	conv-512	conv-512 conv-512
Maxpooling				
FC-256				
FC-128				
FC-64				
Output: FC-3				

Table Ⅰ and Table Ⅱ: CNN architectures are shown in columns. The depth of the architectures increases from the left (A) to the right (E). The parameters of both types of layers are denoted as follows: conv-(number of filters); FC-(number of channels). Maxpooling is a pooling operation that calculates the maximum, or largest, value in each patch of each feature map representation.

Classification Results: Accuracy and F1 score

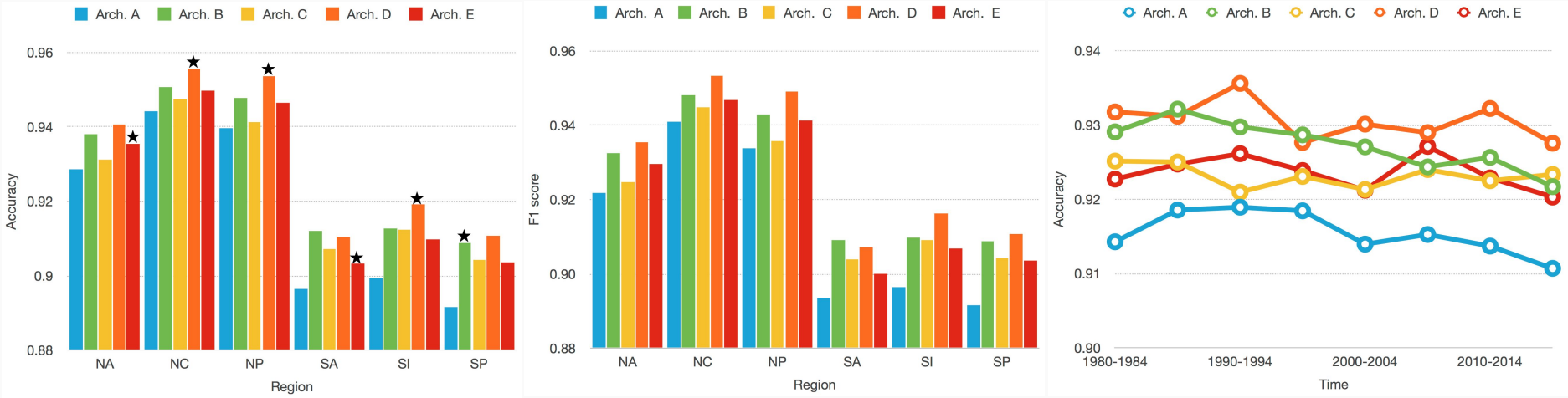

Fig. 7. Performance of CNN models architecture A, architecture B, architecture C, architecture D, and architecture E; for regions NA, NC, NP, SA, SI and SP. The left bar chart (a) illustrates classification accuracy for each architecture per region and the middle chart (b) displays F1 score for each architecture per region. The right bar chart (c) illustrates the mean classification accuracy of each architecture per region over five year periods.The ★ symbol stands for a p-value $\ll 0.05$.

Localisation Results: Concordance Correlation Coefficient & Mean Percentage Error

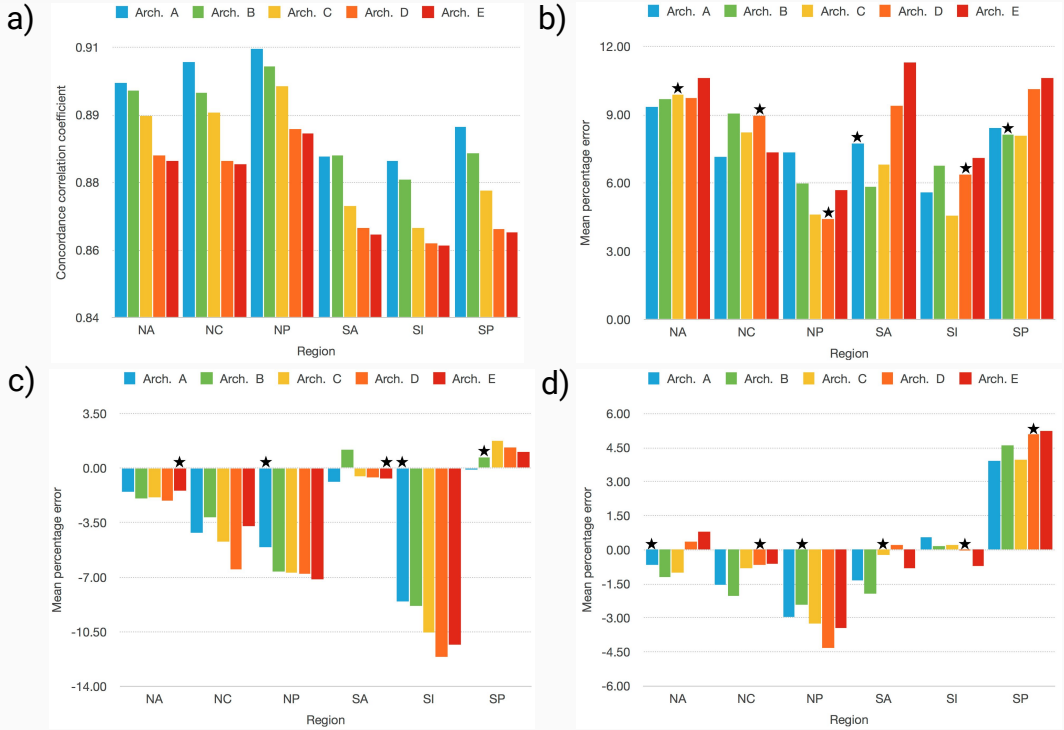

Fig. 8. Performance of CNN architectures: architecture A, architecture B, architecture C, architecture D, and architecture E; for regions NA, NC, NP, SA, SI and SP. Top left bar chart illustrates the Lin's concordance correlation coefficient for each architecture per region. The top right bar chart (a) illustrates the Lin's concordance correlation coefficient for each architecture per region. The top left chart (b) displays illustrates mean percentage error for each architecture in estimating the latitudinal position of the mass centre of atmospheric blocks (ABs) per region. The bottom left chart (c) shows shows mean percentage error for each architecture in estimating the longitudinal position of the mass centre of ABs per region. The bottom right chart (d) illustrates mean percentage error for each architecture in estimating the radius of the mass centre of ABs per region. The ★ symbol stands for a p-value $\ll 0.05$.

Conclusion & Future Work

- We propose **a two-stage hierarchical pattern recognition method** for detection and localisation of AB patterns in various regions over the globe.
- For both classification stage and localisation stage, we **train five different architectures of a CNN-based classifier and regressor.**
- The results show the general pattern of the AB classification **performance (ACC) significantly increases for the deep CNN architectures.**
- In contrast, we see **the estimation error of AB location (MPE) significantly decreases in the localisation task for the shallow CNN architectures.**
- CNN architectures achieve **higher ACC for blocking event detection** and **lower estimation error of AB localisation** in regions of the Northern Hemisphere than in regions of the Southern Hemisphere.
- In the future work, we will explore 3D CNN architectures to study the AB pattern recognition problem. We also plan to use customized loss function (e.g., generalized intersection over union function) that can be beneficial in this problem.

Acknowledgements

- **Colorado State University**: Marie McGraw
- **NERSC, LBNL**: Steve Farwell
- **ETH Zurich**: Stephan Pfahl & Steinfeld Daniel
- **University of California, Davis**: Marielle Pinheiro & Professor Paul Ullrich
- **Intel & Big Data Center at NERSC**

Thank you for your attention!