Can we predict dry air intrusions using an artificial neural network?

Stav Nahum, Shira Raveh-Rubin, Jonathan Shlomi, Vered Silverman

Weizmann Institute of Science, Rehovot, Israel

Introduction

Dry air intrusions (DIs) are one of the main air streams associated with extra-tropical cyclones which govern the weather in the mid-latitudes, and can potentially cause extreme surface weather. This work demonstrates how deep learning can be used to predict the occurrence of DIs using a few commonly used atmospheric parameters (GPH) of current immediate time interval, overcome the limitations of the current lengthy and costly identification method and at the same time expand our limited knowledge about DIs triggering mechanisms.

Motivation and Methods

Dry-air intrusions are often associated with high impact weather events[×]. Understanding the triggering mechanism of DIs is important to predict the likelihood of their occurrence in weather forecasts.

The current identification of DIs (air trajectories descending >400hPa in 48h[××], Fig. 1) is based on a systematic costly Lagrangian method that requires knowledge of the wind field for 48 hours ahead, with high spatial and temporal resolution, which is not available for all atmospheric datasets. Currently, the DIs trajectories can be found only after reaching lower levels, limiting the predictability of these events.

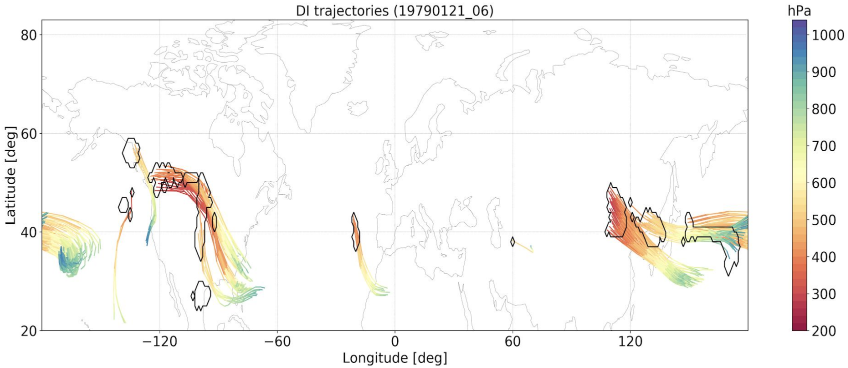


Figure 1: DI trajectories (color) and initiations (black) indicate regions of strong air descent.

× *Catto, J.L. and Raveh-Rubin, S., 2019. Climatology and dynamics of the link between dry intrusions and cold fronts during winter. Part I: global climatology. Climate dynamics, 53(3-4), pp.1873-1892.*

×× Raveh-Rubin, S., 2017. Dry intrusions: Lagrangian climatology and dynamical impact on the planetary boundary layer. Journal of Climate, 30(17), pp.6661-6682.

Objectives

Can deep neural networks (DNNs) be used to:

- Identify and quantify an atmospheric phenomenon using only few meteorological parameters, which are available in most atmospheric reanalyses/ models ?
- Improve our physical understanding of Dry Intrusions, and their environment of initiations in particular, by learning which parameters are essential to identify their origins
- Predict the initiation regions of strong air decent without knowledge of the future ?

Approach - Convolutional Neural Network (CNN)

CNN: *Convolutional Neural Network,*

is a **neural network** that performs its image learning process through a series of linear convolution operations followed by nonlinear functions. The iterative learning process forms a complex function that relates the **input** to the **output** of the network. Future use of this function will enable us to predict outputs when we don't have their matching inputs

Approach - Image Segmentation: an unconventional use

Several convolutional neural network (CNN) were developed for computer vision purposes - **Semantic Segmentation (**labelling each **pixel** of an image with a corresponding **class, Figure 2a**)

Conceptually, training of the network is done by 'showing it' many different images of dogs coupled with an overlaid mask of where the dog is located in the image. If we performed the training process right, the network would learn adequately to identify dogs in images it hasn't seen before - image segmentation task.

Our work implements an image segmentation task in an unconventional way, labeling pixels in 2D maps of geopotential heights surfaces to predict where DIs are likely to be found (Fig. 2b).

Unlike in Fig 2a, where you can clearly identify the dog in the input image, in Fig 2b, you **can not** identify the DI origin with your own eyes.

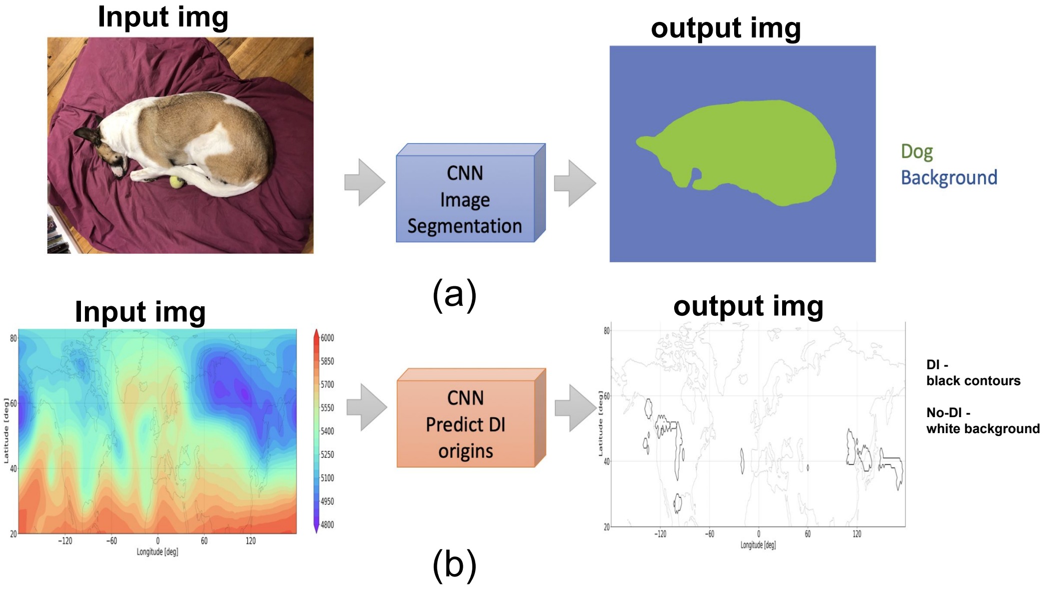
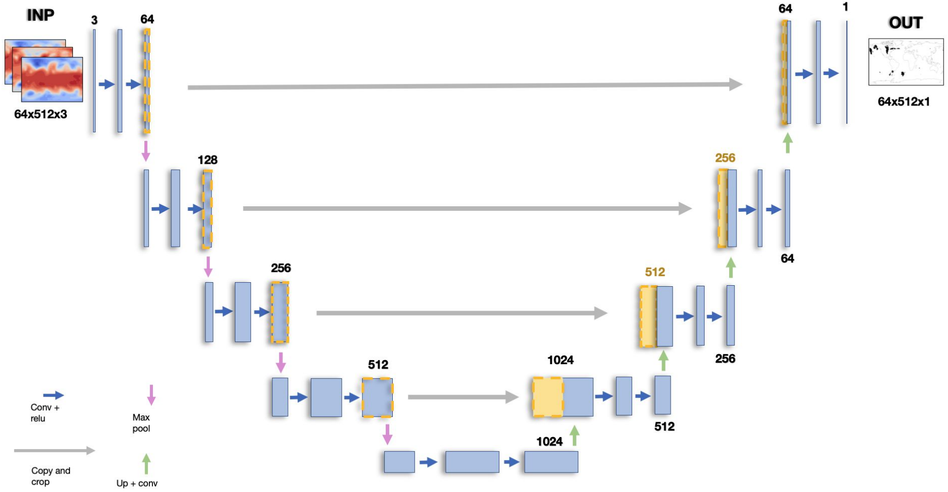


Figure 2: Illustration using CNN's for (a) image segmentation (pixel labeling), in this case, labeling pixels belonging to dogs in images it hasn't seen before, and (b) identifying atmospheric phenomena - our implementation of an image segmentation model to relate atmospheric parameters for identification DI origins.

Approach - Image Segmentation: a word about U-NET

Figure 3:
U-NET Scheme



The scheme of our CNN is based on the U-NET[×] architecture. It was designed to detect and differentiate between objects in a simple RGB image (or similar) and execute a training process similarly to a conventional CNN. Its main **advantage** is in its **distinct structure** which is composed of 2 main branches and an additional unique connecting component. A `**descending reducing'** branch (encoder) which detects the meaningful features but loses their location, an `**ascending increasing'** branch (decoder) that is responsible for increasing the features' resolution and creating the segmentation map. The connecting component, **copy and concatenate**, is used to mitigate between the two branches by carrying information about the features' location, assisting in creating the high resolution segmentation map.

[×] *Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention 2015 Oct 5 (pp. 234-241). Springer, Cham.*

Approach - Image Segmentation: an unconventional use

- An Unconventional Use

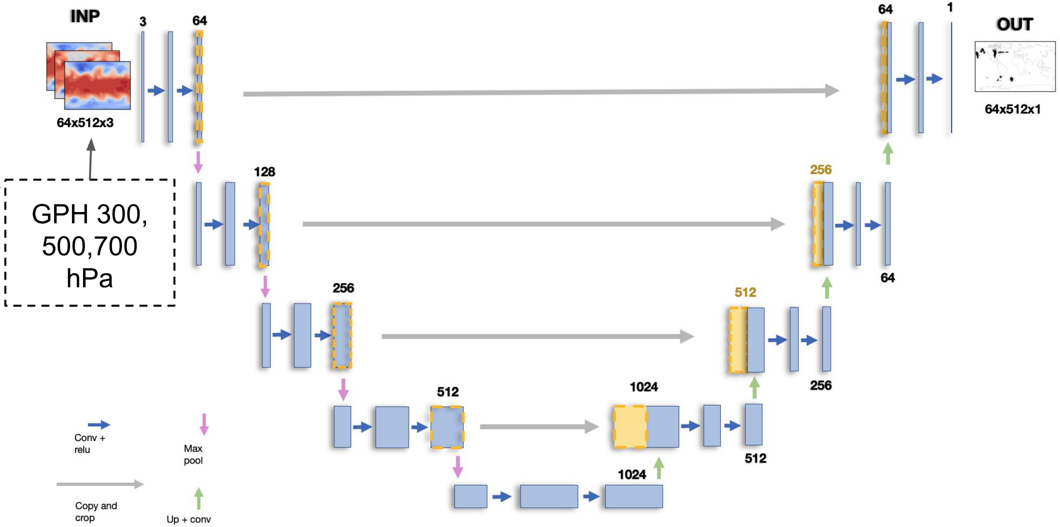
When training the network, we feed the neural network with **2D maps of Geopotential Surfaces (300,500,700-hPa,** in a certain timestep) coupled with **a mask representing where DI origins located.**

- The task for the network

to learn the meteorological patterns and hopefully, if a pattern is found, it would be able to identify in which pixel a DI is likely to occur in meteorological patterns it hasn't seen before (Figure 2b).

This is a **non trivial use** of a model designed for image segmentation to **identify meteorological features** which are not necessarily apparent to the naked eye.

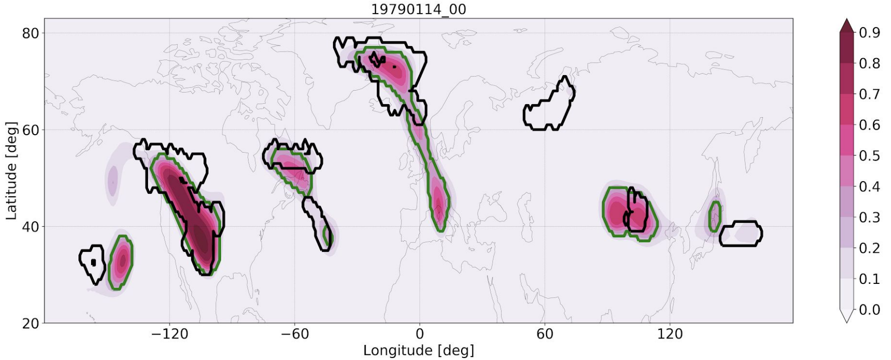
Figure 3: U-NET Scheme



Results

Initial analysis of the results: how well the model did in detecting specific events?

Purple shading	Model output, prediction (probability map[x])
Black contour	Ground Truth - where DIs where actually located
Green contour	Threshold□ contour for probability map



[x]probability map -
the model prediction output is in fact, a probability map. Values between 0-to-1 for each grid-point.
Higher values mean higher likelihood of DI origins existing in those grid-points.

□Threshold -
An acceptable practice is to reasonably choose a threshold for the probability map in order to dichotomously differentiate between the two classes (DI, no DI)

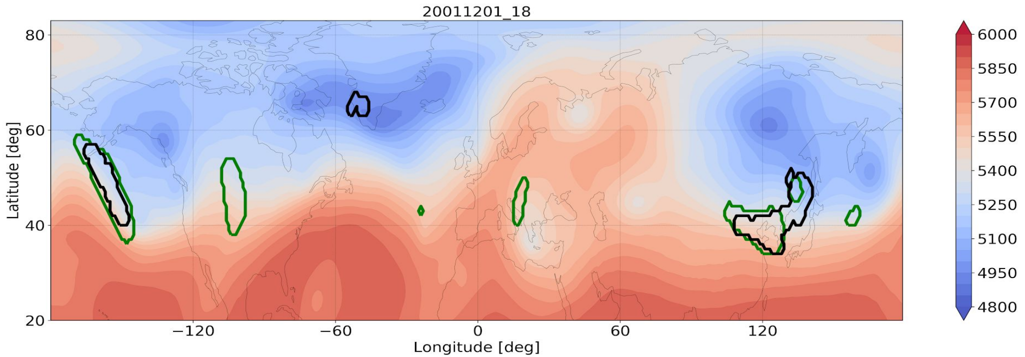
We can observe that overall, for this specific time step, the model produced a very similar probability map[x] (purple shading) that overlaps fairly well with the ground-truth (black contour)

We can notice that over Asia the model failed to capture the event, and that over Japan it recognized an event but couldn't position it correctly.

Results

Initial analysis of the results:
how well the model did in detecting specific events?

Red-Blue shading	500-hPa Geopotential surface
Black contour	Ground Truth - where DIs where actually located
Green contour	Threshold□ contour for probability map



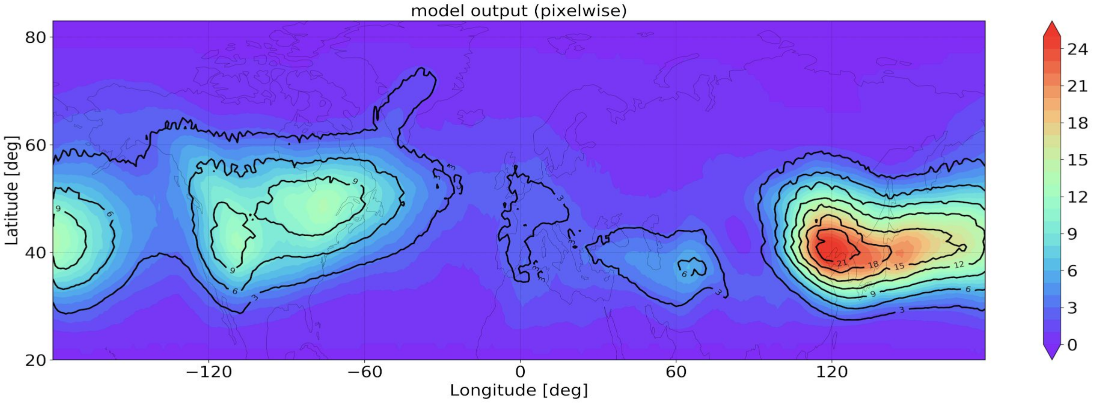
For this different timestep, the model (green contour) also caught quite accurately the ground truth (black contour). There are some variations in orientations and exact location but the overall performance is good.

From previous work, we know that DI origins are usually found on the western side of a trough (NH) having similar orientation as the trough. This leads us to another interesting element that we can draw from the results, the **model's False Positive** (FP) patch over Central North America is situated in a typical location DI origins. Since, the behavior, in which the model signals a strong FP DI origin patch, is widely observed in our results we can carefully assume that model had learned some interesting patterns of the dynamical environment typical with DI initiations.

Results - Climatology

Can the model re-create the DI origins climatology? (psstt... yes it can)

Rainbow shading	Model's predictions
Black contour	Ground Truth - Real DI climatology



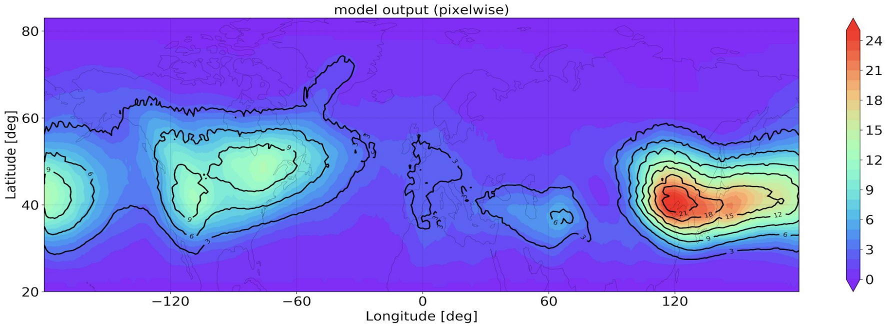
The model was set to evaluate the entire period from 1979-2017 (DJF) 6h-time intervals and we produced a climatology based on its predictions. While the rainbow coloured shading show the climatology based entirely on the model's predictions, the black contours shows the real climatology based on our conventional tools. At a first glance, it looks promising.

The model recreates the climatology quite accurately, both in orientation and magnitude of the events found along the NH mid-latitudes. It captures the irregular shape and the occurrences of its main hot-spots. On average, we do not observe any strong biases.

Discussion and Summary (1)

In this work we tested the feasibility of using a neural network in the atmospheric field.

We used common atmospheric fields (produced by most GCM and forecast models) as our inputs, and showed the network is also able to formulate a complex function to identify features we cannot objectively identify with the naked eye.
Doing so, we demonstrate that this method is not fixed to a specific model but can be implemented for other tasks, and inspire others to implement and create similar approaches in our field.



- In a close case-to-case observation we can find some differences between the model's prediction and the ground-truth. This can stem from similar GPH patterns that either (1) confuse the model, (2) induce DIs of different scale (different descent criteria) thus resembling in appearance to a DI but not quite with a magnitude of a DI.

- An examination of the climatology produced by the model is very encouraging, if we can safely say that the model learned to detect, on average, atmospheric patterns that are conducive to DIs
- We suggest that a possible implementation of this model is onto CMIP future projection data, to create a future DI climatology that can teach us much more about future circulation changes, associated future DI occurrences and their related impact

Discussion and Summary (2)

- Prediction of DI origins is now possible using ***instantaneous*** and commonly available meteorological data, and is computationally ***fast*** using the presented U-Net model.
- We have demonstrated a non trivial use of a model designed for image segmentation (computer vision) to identify meteorological features which are not necessarily apparent to the naked eye. The methodology can be useful to study other weather phenomena.
- A closer examination of false positive (fp) and false negative (fn) events could yield some interesting insights and to broaden our knowledge in our research;
    - **False Positive Events**; DIs were predefined with a strict constraint (400hPa/48hrs). This threshold was selected after thorough diagnostic of all other possible criteria. We are aware that other different ‘magnitudes of descents’ (quicker/slower than 400hPa/48hrs) are occurring all the time. Our thoughts are that perhaps the model catches some of these irregularities and is indicating us about the existence of a similar phenomena but just in a different magnitude that might be taking place prior a DI event or in other locations other than the current familiar pattern.
    - **False Negative Events;** in times when the model fails to detect a DI origin event it arises a few questions -
        - maybe the GPH fields aren’t sufficient
        - perhaps there is more information ‘hidden’ in other parameters
        - maybe a different model architecture would be more suitable, i.e, a Dynamical Filter Network this network structure considers the geographical differences and learns to evaluate differently each domain.