

Nonlin. Processes Geophys. Discuss., author comment AC1  
<https://doi.org/10.5194/npg-2021-6-AC1>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC1

Camille Besombes et al.

---

Author comment on "Producing realistic climate data with generative adversarial networks" by Camille Besombes et al., Nonlin. Processes Geophys. Discuss.,  
<https://doi.org/10.5194/npg-2021-6-AC1>, 2021

---

First of all, we would like to thank the referee for their review of our paper and for giving us the opportunity to improve it.

The structure of the response will be the detailed answers to the questions raised by the referee and then the list of the changes made compared to the preprint version :

### Answers to the referee's questions :

RC1-C1.

How tolerant are the wrap padding layers to discontinuity?

We have added in supplementary material Fig. 6 that is a plot of the value at z100 and z500 levels along the longitude for a given sample where the dashed line shows the x index corresponding to the sides of the image. It can be observed, in this sample, that no discontinuity can be seen. This test was reproduced for other samples where no discontinuities were detected as well.

We also note that the original simulator, PLASIM conserves continuity across the boundary intrinsically. It was also added in the text of the article that "*PLASIM simulator is a spectral model run on a gaussian grid that consequently enforces the periodic boundary condition.*" (line 174, page 7)

To which extent, the padding process may affect the training ?

The discontinuities can be seen as unwanted spatial patterns. In the early training phase, the generator will produce noisy images. During training iterations, the generator will remove unwanted spatial patterns. The wrap padding layer can be seen as a way to underline those unwanted spatial patterns (here discontinuities) between points in the images that are far from one another. Thus, authors argue that the padding is necessary – but not that the effect is insignificant. The computational cost is too important to repeat the analysis for a range of padding values.

RC1-C2 :

By looking at the inputs/outputs in figures 1 to 5, I assume that some convolutional blocks have pooling/ upsampling layers and others do not. Could you explain the reason for that and present these blocks differently in the figures?

The upsampling and pooling layers are here to respectively increase and decrease the size of the image until it reached the desired output size. Once the image has the right size it can be useful to add convolutional block in order to increase the capacity of the network but these additional convolutional layers will not need to modify the image shape.

In our case, the pooling layer is a strided convolution that can be seen as a learnable pooling layer (see Fig 1 to 5 in supplementary material). The same could be done for upsampling layer, where it can be replaced by a strided transposed convolution. In the sake of simplicity, it was decided to keep upsampling layer for our case.

Added sentence in the article : *“At each strided convolutional block,  $s=2$  in Fig. 3 the image size is divided by a factor 2. It is equivalent to a learnable pooling layer that can increase the result (Springenberg 2014).” (line 190, page 7)*

RC1-C4 :

The stopping criteria of the training is a maximum number of iterations. Could this condition be combined with the convergence of the loss function, for example loss function derivative very small or training set and validation set having almost equal loss ?

In theory, because the wasserstein distance, and its derivative, has a mathematical meaning, it could be used as a stopping criteria. But in application it is not yet proven that the WGAN will converge the Nash Equilibrium in a finite number of iterations. Moreover, the instability of the GAN training is a difficulty for using the derivative of the loss function. Thus, to define a threshold under which the loss function should be under to stop the training, it would be necessary to first define metrics related to the data domain (as we have done in our study).

The concept of training and validation set for the generator is not pertinent because it is a non-supervised training. However, it could be useful to perform such dataset split to evaluate the discriminator performance on images it has never seen.

Remark about this question was added in the manuscript :

*“Although it is possible to use the loss of the critic as convergence criteria because the Wasserstein loss is used and has a mathematical meaning such as the distance between synthetic and real data distributions and should converge to 0. But WGAN-GP is not yet proven to be locally convergent under proper conditions (Nagarajan 2017), the*

*consequence is that it can cycle around equilibrium points and never reach a local equilibrium. Condition on loss derivative is also difficult because of the instability of the GAN training procedure. Consequently, a quality check using metrics adapted to the domain on which the GAN is applied is still necessary.” (line 221, page 13)*

### **Minor changes :**

A paragraph was added in the introduction that presents the different data-driven methods applied to numerical weather prediction :

*“Data driven approaches and numerical weather prediction are two domains that share important similarities, Watson-Parris (2021) explains that both domains use the same methods to answer different questions. This study and Boukabara et al. (2019) also show that numerical weather prediction contains lots of interesting challenges that could be tackled by machine learning methods. It clarifies the growing literature about data driven techniques applied to weather prediction. Scher (2018) used variational autoencoders to generate the dynamics of a simple general circulation model conditioned on a weather state. Weyn et al. (2019) trained a convolutional neural network (CNN) on gridded reanalysis data in order to generate 500-hPa geopotential height fields at forecast lead times up to 3 days. Lagerquist et al. (2019) developed a CNN to identify cold and warm front and post-processing method to convert probability grids to objects. Weyn et al. (2020) built a CNN able to forecast some basic atmospheric variables using a cubed-sphere remapping in order to alleviate the task of*

*the CNN and impose simply boundary conditions.” (line 33, page 2)*

The formatting of the equations 5, 10, 13 was reviewed.

Figure 1 modified in order to specify the layers where strided convolution were used. (see Fig. 1 in supplementary material)

Figure 2 and 3 : Captions were modified to differentiate convolutional and identity blocks. (see Fig. 2 and 3 in supplementary material)

Figure 3 was modified to introduce the parameter  $s$  representing the stride value of the convolution that change between different convolutional blocks in the critic architecture. (see Fig. 3 in supplementary material)

Figure 4 : The mention of the upsampling layers, when used, were added on the convolutional blocks. (see Fig. 4 in supplementary material)

Figure 5 : Caption modified to precise that upsampling layer is not in all convolutional blocks. (see Fig. 5 in supplementary material)

The merging of figure 10 and 11 was not done due to a loss of readability of the figures.

Figure 14 was removed because of the redundancy of the information it shows.

Figure 22, 23 and 25 : Descriptions of the panels (a), (b)... were added.

Figure 22 : *"Temperature  $(K)$  and zonal wind  $(m/s)$  latitude zonals from a boreal winter situation : the thermal wind balance. Left panels correspond to a situation taken from dataset (a) zonal temperature, (c) zonal wind, right panels correspond to a situation taken from generator (b) zonal temperature and (d) zonal wind."*

Figure 23 : *"Thermal wind balance from the boreal winter situation shown in Fig. 22 : (a) sample from the dataset, (b) sample generated by the generator. The temperature  $(K)$  is from pressure level 800 hPa and the wind  $(m/s)$  from 200 hPa."*

Figure 24 : *"Temperature  $(K)$  and zonal wind  $(m/s)$  latitude zonals averaged on 30 years subsample. Left panels correspond to a situation taken from dataset (a) zonal temperature, (c) zonal wind, right panels correspond to a situation taken from generator (b) zonal temperature and (d) zonal wind."*

Figure 30 : The equation was referenced instead of described in the caption.

Please also note the supplement to this comment:

<https://npg.copernicus.org/preprints/npg-2021-6/npg-2021-6-AC1-supplement.pdf>