Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-128-RC2, 2019

Geoscientific Model Development Discussions

Interactive comment on "DINCAE 1.0: a convolutional neural network with error estimates to reconstruct sea surface temperature satellite observations" *by* Alexander Barth et al.

Anonymous Referee #1

Received and published: 13 December 2019

The author has clarified a lot. Thanks.

1."We actually have good results using the present network structure on the Adriatic Sea and we have been contacted by a researcher using the same network architecture on the South China Sea and West Philippine Sea providing a convincing reconstruction."

This is indeed interesting to see

2. "However, the cross-validation error for this experiment is 0.433 C and the results are not improved. Increasing the number of input features can aggravate the potential for overfitting as the number of parameters in the neural network is increased."

The channels of input data increased from 8 to 10. But the filter size, the number of output feature maps, and layer size, number of layers stay the same. Thus, the parameters of the network should stay the same. Right?

3. "The effective range of the error standard deviation is thus from..."

This is helpful. Previously, the author introduced two variables with no explaination of what and why. Previously, from the formula only, it seems like the value of these two variables will affect strongly the computation. e.g. delta = 100 vs. delta = 0.01

4."The RMSE of the average DINCAE reconstruction is smaller than the RMSE at any given epoch because computing the RMSE is a non-linear operation. The DINCAE reconstruction at a given epoch included..."

I am not very sure I fully understand it. But I will leave it to other reviewers!

5. "the additional data marked as missing within a minibatch is a spatial subset which enhances these fluctuations but allows us to define the cost function more closely to our objective..."

I would guess the fundamental reason why the RMSE and Loss fluctuate so much is that the random mark missing data in every mini batch. Because in every epoch, the spatial correlation of missing and available data is disrupted due to random marking, hence what the network has learned in previous epoch is disrupted as well, which eventually is reflected in RMSE and Loss. The fluctuations may not have so much to do with mini-batch optimization. Perhaps one way to check is to use same random mark missing data for every 20 epochs, and average at every 20 epochs. Just my opinion.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-128, 2019.