Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-128-RC1, 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: 3 December 2019

DINCAE 1.0: a convolutional neural network with error estimates to reconstruct sea surface temperature satellite observations

General Comments: It is an interesting paper overall. The author uses Auto Encoder to reconstruct the missing data commonly found in optical satellite remote sensing caused by instrument failure or cloud cover. The author uses an interesting way to handle missing data in training image. As compared to the widely used DINEOF method, the author showed that DINCAE can, on some degree, produce better results measured in RMSE metrics, as well as spatial distributions of SST. From the technique side, Auto Encoder is a commonly used machine learning method in semantic segmentation and

object detection. The author uses this method to solve, particularly SST, reconstruction obtained from satellite remote sensing. It is an interesting and meaningful problem to tackle. But as for developing a new methodology, I have some concerns.

(1)The applicability of this method to other satellite measurements. Variables such as SST, have low frequency variability both in space and time (If I am right). This nature suggest that it is relatively easier for CNN to estimate the spatial correlation (e.g. for an image in which there are multiple people, it is harder to do segmentation than for an image with only lawn and sky). This also gives ground that average pooling turns out to achieve better results than max pooling, as stated in paper. For variables, especially those on land, such as plant reflectance, usually have high frequency variability both in space and time due to heterogeneous growth stage, backgroud, and so on. These feature creates additional challenges, which I think, cannot be handled with the method configuration stated in paper. It will be interesting to see how it does (This may not directly related to the topic of this paper). Additionally, the method is tested at one site, which hardly persuasive to show its applicability over the globe. Will a model trained at one site be able to use at another site, or it is needed to develop a new model to a new site, which usually needs a lot of work to prepare data, training model, parameter tuning and so on? If so, from model deployment side, what the advantage of using it?

(2)Temporal feature of reconstructed variables EOF method is essentially PCA analysis. DINEOF method does take into consideration of both temporal and spatial correlation of variables, to my understanding. Though DINCAE, as described in the paper, also uses the spatial and temporal correlation of variables, it only uses correlation presented in 3 days (the day, the day before and the day after). In other words, spatial information is what it uses mainly for reconstructing. Do you have persuasive arguments that 3 days correlation in time are enough to capture temporal dependency? However, longer time dependency, e.g. seasonality, may also be important on estimating missing values. In this case, net work configuration both capture spatial and temporal structure of variables (e.g. LSTM + CNN) could be more general and powerful.

Technique Comments: Page 1 line 2: ‘A method to reconstruct missing data in satellite data using a neural network is presented’ The first sentence is not as precise as it should be. As the first impression that this paper is going to introduce a neural network based method to reconstruct/interpolate gappy satellite images caused by cloud coverage, instrument failures (e.g. LandSat 7) and so on. However the following paper mostly discussed an AutoEncoder method to reconstruct SST and tested only on SST.

Page 2 line 31: ‘effectively reducing....’ What is the meaning of putting this sentence here

Page 4. Figure 1 caption ‘The arrow represent...’ There is no arrow on figure

Page 4 line 6: ‘so that for a given date also the satellite’ Delete ‘also’

Page 5 line 12, ‘...in the following’ Delete ‘in the following’

Page 5 line 19 ‘assimilation of data’ Change to ‘data assimilation’

Page 7 line 20 ‘skip connection’ Does the resolution of SST data have effect on how you use skip connections? How large scale is called large scale for resolution of 4KM by 4KM, how about SST with resolution 1KM by 1KM. From another point of view, this operation again consolidate to use the spatial information for reconstruction, while temporal information somehow is ignored.

Page 8 line 5 The two parameters here seemly have profound effect on reconstruction result, how does these two parameter chosen?

Page 9 line 15 ‘the output of the neural network is a Gaussian probability distribution’ The author assume the output is a Gaussian distribution, ‘is a Gaussian distribution’ means the author know it is Gaussian.

Page 10 line 18-21 ‘As mentioned before, ...neural network’ Not quite understand the training procedure here. ‘a random subset of data is marked as missing’? Since the missing data is marked randomly for each epoch, it is possible that at epoch = k, some

part of data is marked as missing, while at epoch = k+1, the same part of data of the same image is marked as available. If this is the case, it essentially means the model was told what it should predict randomly? This is somewhat contradictory with Page 9 line 10.

Page 10 line 21-22 ‘we average ...intermediate result’ Why do not average multiple runs?

Page 11 Figure 2 caption ‘red dash line ...’ How come the average DINCAE reconstruction is smaller than RMSE at any given epoch? Also, the error curve indicates that the model has no sign of convergence. I bet if you continue training the model for another 1000 epochs, the cross validation error curves will not converge. This also indicates that there might be something wrong in the training procedure. Can you plot your loss function here as well?

Page 14 line 16 ‘also tried ...’ The max pooling operation tries to extract distinct signals from neighbors, while average pooling operation tries to extract common signals from neighbors. For SST, which has low frequency variation in space, it makes sense average pooling should do better than max pooling.

Page 17 line 14 ‘...reconstruction is it thus...’ Change ‘is it’ to ‘it is’

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