Gmail

Elizabeth Barnes <eabarnes.earth@gmail.com>

track data

Mark DeMaria - NOAA Affiliate <mark.demaria@noaa.gov> Tue, Oct 26, 2021 at 10:46 AM
To: eabarnes@colostate.edu

Libby:

Thanks again for continuing the interesting discussion today. Below are the details of how to calculate the intensity and track errors for individual cases using the nnfit_vlist_20-Oct-2021.dat file I sent before and the NN predictions from that data. I also summarized what would be good to show with the plots. Let me know if you have questions.

Mark

Intensity error:

1. Let DVNN be the deviation from the intensity consensus for an individual case predicted by the Neural Network, given the inputs (the deviations of the models from the consensus DSDV, LGDV, etc, and the storm environment variables, DV12, SLAT, SSTN, etc). The actual intensity forecast is DVNN + VMXC, but we don't need to calculate that for the intensity error analysis.
2. The deviation of the best track (observed) intensity from the consensus is the column labeled OBDV in the input file.
3. The intensity error for a case is then IE = DVNN-OBDV.
4. If you calculate IE for each case, then you can calculate the mean absolute error, bias, RMSE etc for the total sample of cases.
5. The intensity error of the consensus forecast (IEC) for an individual case is just OBDV, so IEC=OBDV. You also calculate the MAE of the individual IEC values.

Track error:

1. Let DXNN and DYNN be the x and y deviations from the consensus predicted by the NNs for lon and lat for a single case, analogous to DVNN for intensity error. The actual track forecast would be DXNN and DYNN converted to lon and lat and then added to the consensus LON and LAT (LONC and LATC), but we don't need to calculate those for the error analysis in terms of distances.
2. The x and y deviations of the best track (observed) position from the consensus forecast position are the columns labeled OBDX and OBDY in the nnfit_vlist input file.
3. The x and y intensity errors for each case are then XE = DXNN-OBDX and YE = DYNN-OBDY
4. The distance error for a case is then DE = sqrt(XE*XE + YE*YE)
5. You can then calculate the MAE and RMSE from DE for each case for the total sample of cases.
6. Although we probably won't show it, you could also calculate the bias of XE and YE, which would show if the NN predicted tracks were systematically east or west and north or south of the observed track.
7. For comparison, if we want the track error of the consensus forecast, the DX and DY consensus errors (XEC, YEC) are just the OBDX and OBDY values (XEC=OBDX, YEC=OBDY).
8. The distance error of the consensus track forecast (DCE) is then DCE=sqrt(XEC*XEC + YEC*YEC)
9. You can then calculate MAE from the DCE values for each case.

Uncertainty estimation:

For the uncertainty information for intensity, we need to have a graph that shows there is a relationship between the actual error (IE) and the uncertainty estimated as part of the NN. This could be shown by binning the predicted IE and showing statistics of the actual IE for each bin. A secondary use would be the uncertainty in the deviation from the consensus. If that is low, then we are better off sticking with the consensus.

The track uncertainty is similar. We want to show that the uncertainty estimated by the network for DE (using the 2-step procedure you described) is related to actual distance error DE. A plot binned by estimated uncertainty of DE verus the actual DE would be good for that too. There is also a secondary use for track, where if the NN is not confident that the predicted deviation from the consensus, we would stick with the consensus.
[Quoted text hidden]