The 45 Days CWBGFS Ensemble Prediction System Based on Singular Vectors

John Chien-Han Tseng, Jen-Her Chen, Ting-Huai Chang, Chong-Wei Allen Huang, Young-Sin Lai, Chi-Cha Wang
Central Weather Bureau, Taiwan

1. The ensemble prediction system

The Taiwan Central Weather Bureau global forecast system (CWBGFS) has been providing 45 days 20 ensemble member prediction once per day on resolution T319L60 model since 1 July 2017. The ensemble perturbations are given by singular vectors (Buizza *et al*., 1993) which are calculated from northern/southern hemisphere (0^0-360^0 , 20^0N/S-80^0N/S) and three tropical areas (around 45^0E-165^0E, 20^0S-20^0N) dry total energy norm separately. Then, the 20 ensemble perturbations are composed by these five domains singular vector linear combination. The CWBGFS ensemble prediction system (CWBGEPS) setup is shown in Table 1. Moreover, the Stochastically Perturbed Parameterization Tendency (SPPT) process (Palmer *et al*., 2009) is added in CWBGEPS since 19 July 2018. The SPPT uses three different horizontal scales 500 km, 1000 km, and 2000 km random patterns which effectively increase the original ensemble spread (Fig. 1).

Basically, the smaller differences between the root mean square error (RMSE) and the ensemble spread, the better ensemble prediction system is. But, in Fig. 1, there are still about 15% differences between RMSE (solid line) and ensemble member spread (dash line). If we compare recently CWBGEPS results with NCEP GEFS (Fig. 2), we can find the RMSE of CWBGEPS (black solid line) is larger than NCEP GEFS (red solid line) and the spread of CWBGEPS (black dash line) is far smaller than NCEP GEFS (red dash line). That

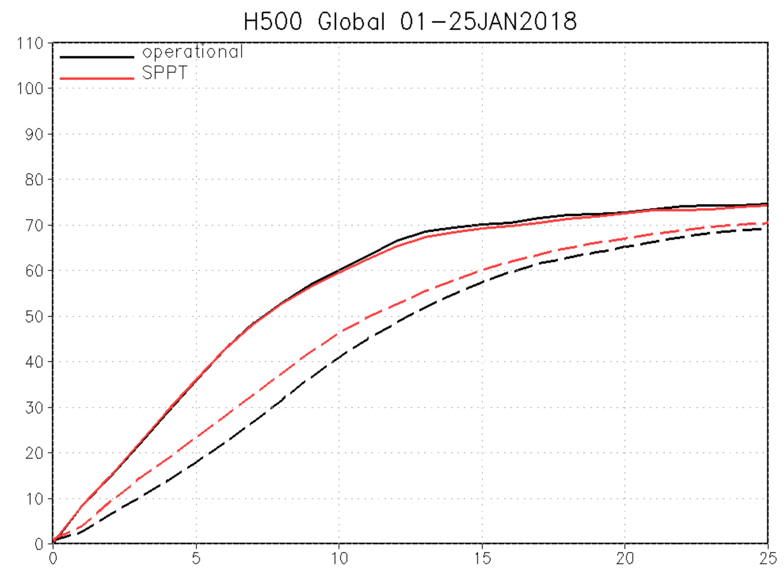

Figure 1. CWBGEPS with the SPPT (red) and w/o (black). The RMSE is shown by solid line and the ensemble spread is dash line.

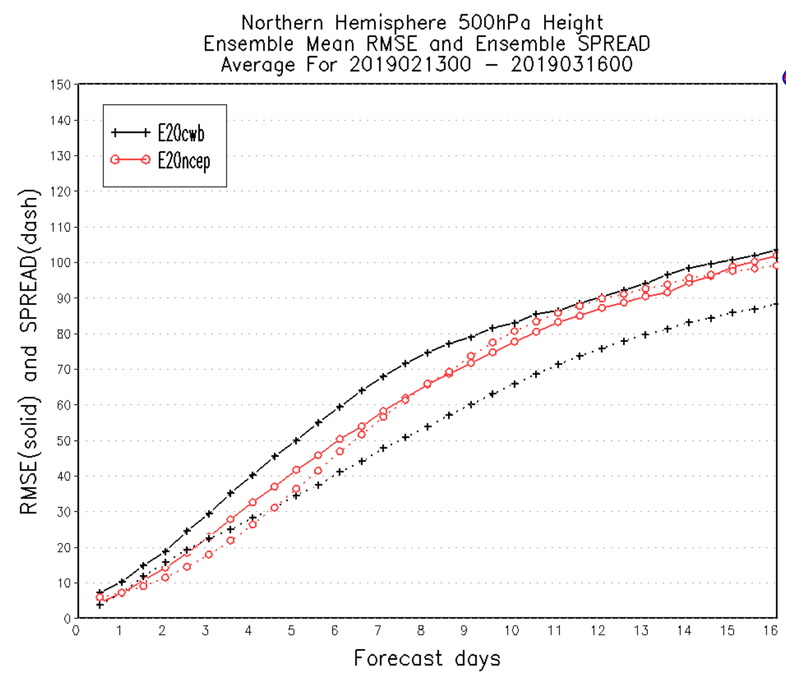

Figure 2. The RMSE and ensemble spread between CWBGEPS (black) and NCEP GEFS (dash).

means the performance of CWBGEPS is not so good. The preliminary view is that we should composite the initial perturbations subtly (enlarge the amplitude?) or fine tune our SPPT relative parameters. However, we probably should improve the CWBGFS itself. Because we find in the first three days the singular vector perturbations growing well and even the initial magnitude just half of NCEP GEFS (Fig. 2). We are not sure why after three days those perturbation cannot keep the growth rate like the first three days and have the enough spread.

Table 1. CWBGEPS Structure

CWBGFS Ensemble Prediction System			
resolution	deterministic model control run		T511L60 T319L60
	ensemble for 45 days		T319L60
	TLM and adjoint model		T42L60
initial perturbation, singular vector	north hemisphere		20^0N-80^0N, 0^0-360^0
	south hemisphere		20^0S-80^0S, 0^0-360^0
	Tropical area	Indian ocean	20^0S-20^0N, 45^0E-85^0E
		South China Sea	20^0S-20^0N, 90^0E-120^0E
		East Philippines	20^0S-20^0N, 120^0E-165^0E
optimization time	48 hrs		
ensemble size	20		
forecast length	45-day		

2. Downscaling

In Taiwan, for the fine agriculture, agricultural research institute needs to estimate which kind of crops should be cultivated in next season based on monthly, subseasonal or seasonally forecast. Furthermore, the long term temperature forecast is used for estimating the insect control or disease prevention in cultivation. For the agriculture needs, CWBGEPS is in charged of affording the two meter temperature 45 days forecast at some specific locations, the greenhouses or the fields. In the future, CWBGEPS is also going to provide rainfall, solar radiation, relative humidity, and wind speed/direction for farmers. The Fig. 3 is the example to show the T2m forecast in Taipei. The green line and red line stand for CWBGEPS control run and ensemble mean. The gray lines are from ensemble members. The blue line is high resolution CWBGFS for 16 days weather forecast. The gray bar marks one standard deviation of ensemble members, and it somehow estimates the highly probability range. The green, red and blue line are always stick together around 5-7 days and then they will separate and evolve to different curves pattern. The 5-7 days is the limitation of weather forecast (the period of time more accurate) and after that 10 days or longer forecast we need to use different view to explain the forecast. The pentad average forecast or same target time different initial time forecast average or some kind of tendency probability could be more appropriate than single specific value forecast for long term forecast.

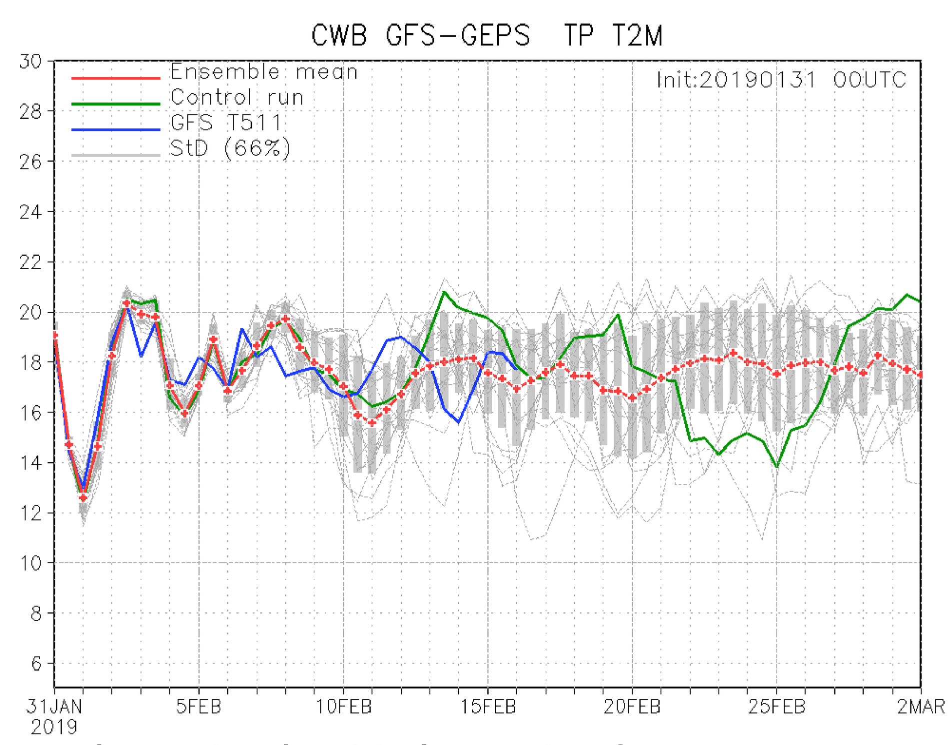

Figure 3. The 30 days T2m forecast at Taipei based on CWBGEPS initial time 20190131 00 UTC 45 days forecast.

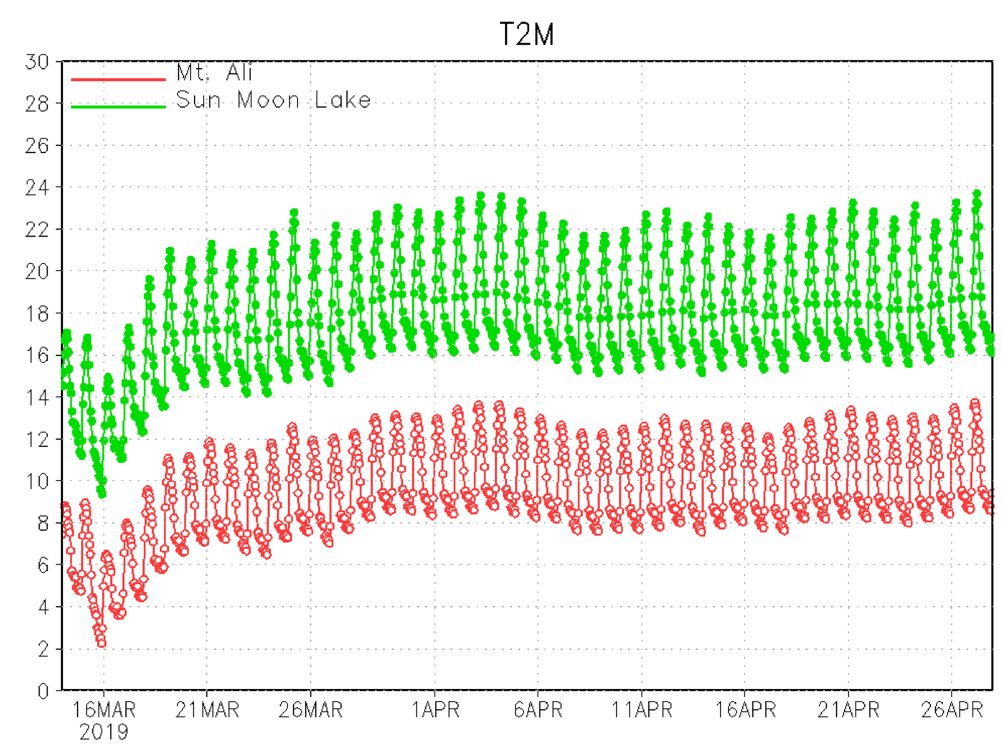

Figure 4. The 45mdays T2m hourly forecast in Taiwan Sunmoon lake (green) and Ali mountain (red).

At the same time, it is not so economic to output high time resolution, e.g., hourly forecast in 45 days forecast. On the other hand, the hourly forecast variation is probably not accurate at downscaling location. We can train, the ***machine learning method,*** the specific location real hourly observation and get the hourly variation regression equation then fit in the daily downscaling 45 days NWP forecast (Fig. 4). The regression we used is ***support vector regression***. Moreover, the support vector regression not only gives the hourly variation but also gives higher spatial resolution possible variation based on ensemble members. Next, we are going to build ***neural network learning*** for fine downscaling Taiwan area forecast by combining CWBGEPS and observation (Alpaydin, 2014; Hsieh, 2009).

3. Summary

The singular vector perturbations of CWBGEPS reflect the dynamic instability around midlatitude jet stream. However, CWBGEPS cannot have well spread especially around tropical area. The SPPT can cooperate well with singular vector perturbations and some parameters like vertical profile and the amplitude factor are still necessary fine tuning. Some interesting energy cascades in midlatitude or tropical areas are worth studying.

There are many jobs in the next stage. They include: for example, how to define the amplitude of singular vector in different time and area, how to select different singular vector calculation areas, how to compose singular vectors, how to extend the predictability of CWBGEPS, e.g. nudging SST, coupling ocean, etc.; should we try to use EnKF or others to form ensemble perturbations; what if we do not have TLM and ADJ in next new dynamic core model and what should we do; is it possible to interpolate current singular vectors into new model grid; what kind of probabilistic products can we contribute to forecast; what kind of downscaling method can we use in Taiwan; Machine learning or AI can help to raise the downscaling forecast quality or not.

We can image that the boundary between weather forecast and climate forecast will be getting blurred. The predictive products of NWP model will be extended to be prescriptive products which are the goal of CWBGEPS.

4. Reference

Alpaydin, E., 2014: *Introduction to Machine Learning*. MIT press, 640pp.

Buizza R., J. Tribbia, F. Molteni and T. Palmer, 1993: Computational of optimal unstable structures for a numerical weather prediction model. *Tellus*, **45A**, 388-407.

Hsieh, W. W., 2009: Machine learning methods in the environmental sciences: Neural networks and kernels. Cambridge Univ. press, 349pp.

Palmer, T. N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. J. Shutts, M. Steinheimer, A. Weisheimer, 2009: Stochastic parameterization and model uncertainty. *ECMWF Technical Memorandum*, **598**, 42pp.