Stochastic Representation of NCEP GEFS to Improve Sub-seasonal Forecast

Yuejian Zhu, Wei Li, Xiaqiong Zhou and Dingchen Hou

Abstract The National Centers for Environmental Prediction (NCEP) Global Ensemble Forecast System (GEFS) has been in daily operation to provide probabilistic guidance for public since December 1992. Since July 2017, the GEFS was extended from 16 days to 35 days forecast to support NCEP Climate Prediction Center (CPC)'s sub-seasonal forecast. The latest GEFS version was upgraded in three areas to improve sub-seasonal forecast: (1) introducing a new set of stochastic physical perturbations to improve model uncertainty representation for the tropics; (2) a 2-tiered SST approach to consider ocean impact; and (3) a new scale-aware convection scheme to improve model physics for tropical convection and MJO forecasts. The new set of stochastic physical perturbations include stochastic kinetic energy backscatter to make up subscale energy lost during model integration; stochastic physics perturbation tendency with five different spatial and temporal scales to perturb physical tendency; and stochastic perturbed humidity on the model lower level. After upgraded to new set of stochastic physical perturbations, the MJO forecast skill has been improved from 12.5 days of a 25-month period to nearly 22 days by combining all three modifications include stochastic physics. In the extratropics, the 500-hPa geopotential height; surface temperature and precipitation are improved for sub-seasonal timescale as well. However, the raw forecast skills of surface temperature and precipitation are extremely low, and the results imply that calibration may be important and necessary for surface temperature and precipitation forecast for the sub-seasonal timescale due to the large systematic model errors.

Keywords NCEP · GEFS · Stochastic representation · Sub-seasonal forecast

Y. Zhu (✉) · W. Li · X. Zhou · D. Hou
Environmental Modeling Center, NCEP/NWS/NOAA, 5830 University Research Ct., College Park, MD 20740, USA
e-mail: Yuejian.Zhu@noaa.gov

D. A. Randall et al. (eds.), *Current Trends in the Representation of Physical Processes in Weather and Climate Models*, Springer Atmospheric Sciences,
https://doi.org/10.1007/978-981-13-3396-5_15

1 Introduction

With the improvement of accuracy of weather forecasting and the increasing computational capacity, a seamless forecast that ranges from weather to seasonal timescale is in growing interest and demanding in general public and service sectors in order to protect life and properties. Extending the weather forecast to cover sub-seasonal timescale clearly has great socioeconomic significance. However, in scientific aspect, improving the forecast skill on this timescale is quite challenging. This gap in the forecast skill between weather and climate is partially due to the limitation of forecast predictability (Lorenz 1969) and less sensitivity to the initial condition which benefits the weather scale yet insufficient sensitivity to the boundary and external forcing which benefits the seasonal and longer lead time (Vitart 2014; Johnson et al. 2014; Liu et al. 2016; Troccoli 2010; Tian et al. 2017). Imperfectness of the representation of the model dynamics and physics, however, should be considered as the major source of uncertainties and errors for all lead time (Buizza et al. 1999). The approaches that aim to reasonably represent the model uncertainty thus become a practical method to reduce the model errors in recent years. The efforts in this regards include a multi-model ensemble method (Shin and Krishnamurti 2003; Palmer et al. 2004; Kirtman et al. 2014) that represents the overall uncertainty from different models; a stochastic total tendency perturbation method (STTP, Hou et al. 2008) that represents the uncertainty related to both dynamic and physics in single model; a stochastic physics perturbation tendency scheme (SPPT, Buizza et al. 1999; Palmer et al. 2009) that represents the uncertainty related to total model physical process; In addition to the stochastic perturbation on the tendency, Stochastic Kinetic Energy Backscatter (SKEB, Shutts and Palmer 2004; Shutts 2005; Berner et al. 2009; Shutts et al. 2015) is another way to present forecast uncertainty through considering the energy at non-resolved scales which cannot cascade to larger scales due to the model's finite resolution. All these methods have been used in operational centers and research community (Palmer et al. 2009).

Since it was implemented into operation in 1992, the NCEP GEFS has been widely used as probabilistic forecast guidance for the forecast within 2 weeks (Zhou et al. 2017). Regardless of the initial perturbation, the operational version of GEFS uses STTP to represent the model uncertainty. With the contribution of both initial uncertainties and perturbation in total tendency, the ensemble mean forecast of GEFS has outperformed the GFS deterministic forecast on the anomaly correlation of forecast lead at Day-8 of Northern Hemisphere 500 hPa geopotential height to represent mid-level general circulation for the past few years (Fig. 1).

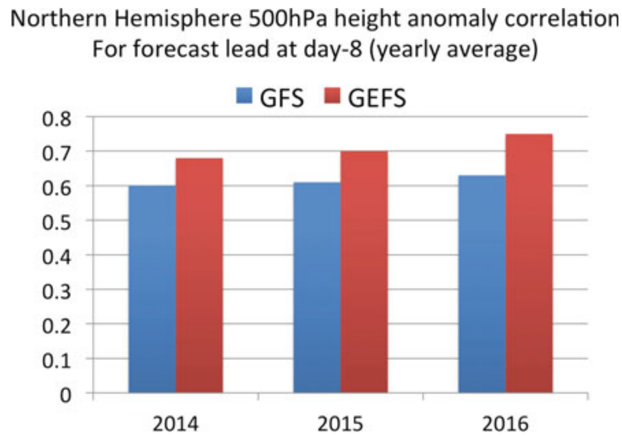


Fig. 1 Northern hemisphere 500 hPa geopotential height anomaly correlation for forecast lead at day-8 of GFS forecast (blue) and ensemble mean forecast (red) during years 2014–2016

2 Stochastic Physics Perturbation Schemes Tested in NCEP GEFS 35-Day Forecast

To be aligned with NOAA's mission of generating a unified coupled forecast system to cover the timescale from weather to seasonal, GEFS has carried out investigations on the strategy to potentially improve the forecast skill on week 3 and 4 time range (sub-seasonal timescale), and further to cover monthly forecast. A recent investigation is testing the impact of different stochastic perturbation schemes that represent the model uncertainty on the performance of sub-seasonal forecast (Zhu et al. 2017, 2018; Li et al. 2018). The motivation for this work came from the concerns of the under-dispersion (or overconfidence) of the current operational version of GEFS (GEFS v11 with EnKF initial perturbation + STTP) on medium range forecast especially over the tropics (Hou et al. 2008; Zhou et al. 2016, 2017).

Although STTP scheme compensates the less error growth from initial perturbations to some degree, the impact of the STTP is mainly over extratropics during boreal winter season with less impact on the spread over tropical region. It is well known that MJO is a major source of the predictability on sub-seasonal timescale. Therefore, improving the representation of the model uncertainty over tropics is a possible pathway to potentially improve this source of sub-seasonal predictability. A suite of three widely accepted stochastic perturbation methods (SPs hereafter) is thus applied to GEFS to represent the model uncertainties instead of STTP more efficiently (Table 1, second row). The scheme of SPs are: SKEB from expectation of making up subscale energy lost due to imperfect computation algorithms; SPPT with five different spatial and temporal scales (Figs. 2 and 3); Stochastic Perturbed Humidity (SHUM; Tompkins and Berner 2008) with single spatial-temporal scale, and near model surface layers. These schemes have already been implemented in the National Center Environmental Prediction (NCEP) Global Forecast System (GFS) model in the hybrid-EnKF data assimilation system, then basically available for use

Table 1 The configuration differences for four experiments

Experiments	Stochastic schemes	Boundary (SST)	Convection
CTL	STTP	Default	Default
SPs	SKEB + SPPT + SHUM	Default	Default
SPs + SST_bc	SKEB + SPPT + SHUM	2-Tiered SST	Default
SPs + SST_bc + SA_CV	SKEB + SPPT + SHUM	2-Tiered SST	Scale-aware convection

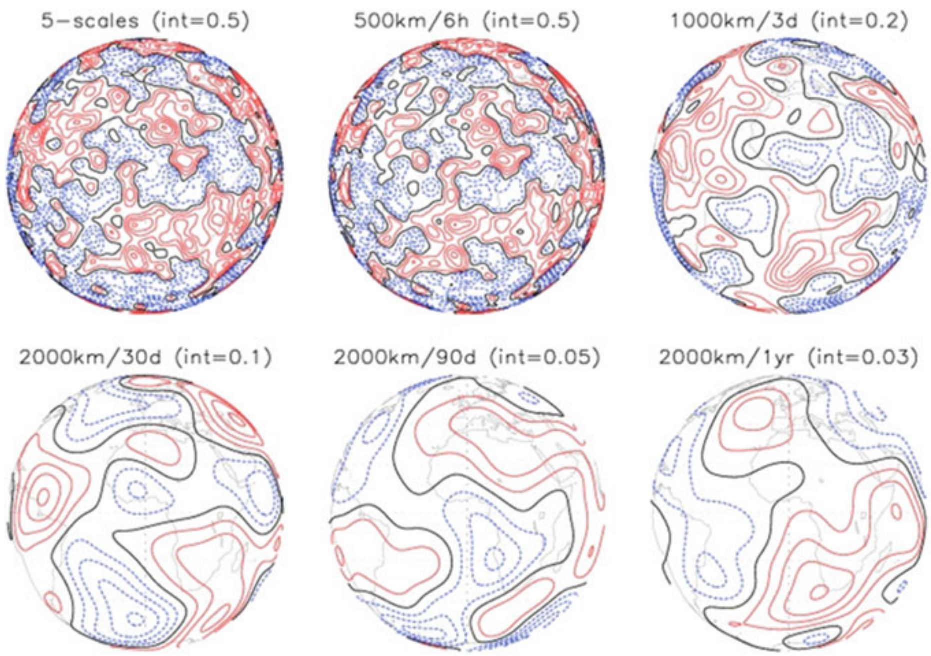


Fig. 2 5-scale random patterns used in stochastic perturbed physics tendencies (SPPT). On the top of each plot, the numbers (except for upper left) represent the scales of spatial and temporal perturbations with contour intervals in the bracket. The upper left is for combined total 5-scales

in the GEFS for testing and modification. Detailed descriptions of these schemes are as follows.

The SKEB scheme has been used to represent dynamical uncertainty through subgrid-scale processes that propagate upscale. A stream function forcing from the total dissipation has been applied to SKEB. Depending on numerical model design, the numerical dissipation (i.e., the diffusions) is only part to be considered in current GFS version. The generations of such perturbations on each vertical level are independently to provide some vertical coherence through vertical smoothing. Overall, the SKEB scheme should improve the global power spectrum and increase forecast spread.

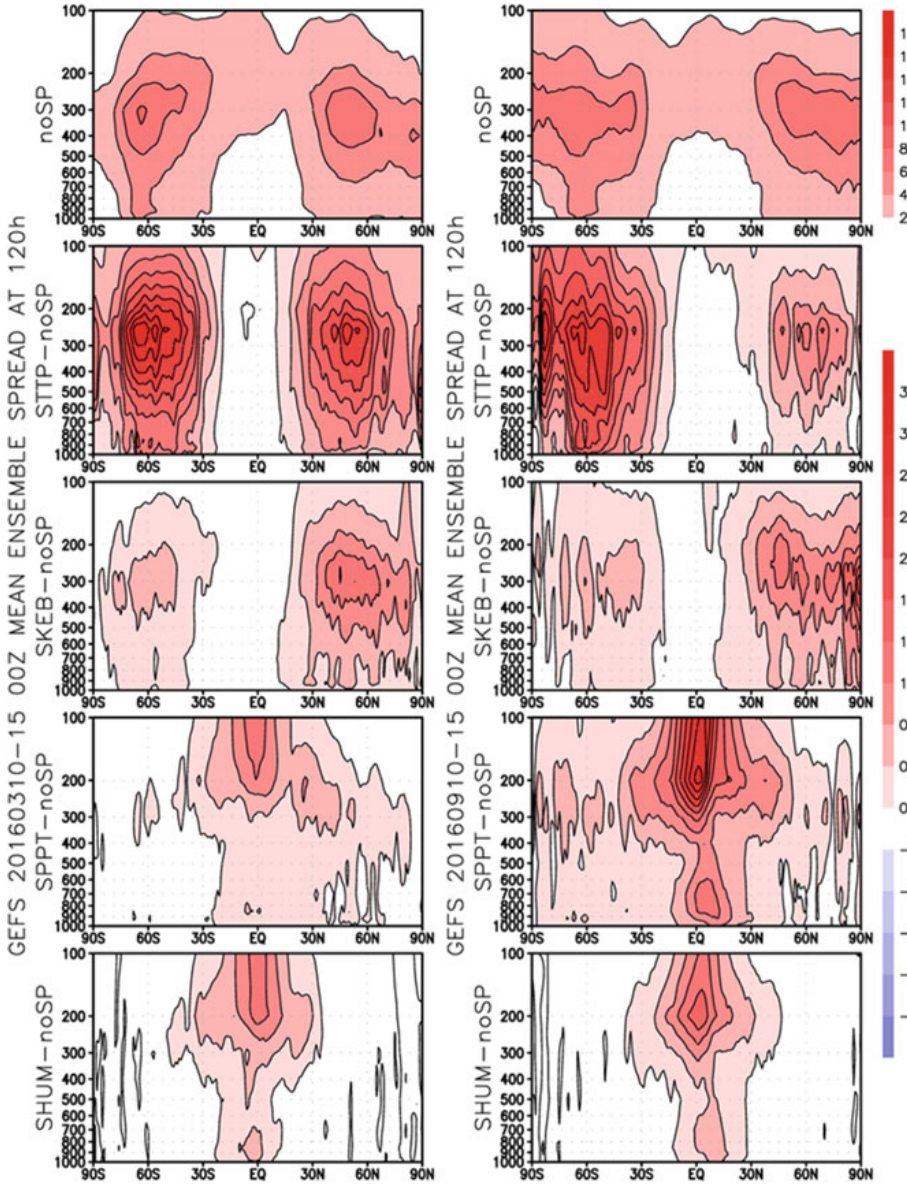


Fig. 3 Global meridional cross section showing the impact of stochastic perturbations for the atmosphere (cross section) for 120 h forecasts from six spring initializations (left) and six fall initializations (right). Paneled are the differences of zonal wind spread from CTL for (top) no stochastic physical perturbations, and the difference of STTP (upper middle); SKEB (middle); SPPT; (lower middle) and SHUM (bottom)

The SPPT scheme perturbs the total tendencies of temperature, wind, and water vapor during numerical integration generated by the GFS physics parameterizations (after all physics processes). The current version of SPPT implies five different random patterns with different timescales and correlation length scales to generate the tendency perturbations. The patterns of the stochastic perturbations, in general, are uniform in the vertical levels, except their magnitude are reduced and taped to zero gradually for both of near surface and above the tropopause. The maximum amplitudes of five scales are 0.8, 0.4, 0.2, 0.08, and 0.04 respectively. Figure 2 demonstrates the individual independent random scale patterns and combined 5-scale random pattern.

The SHUM scheme perturbs the near-surface humidity only; based on the concept that the uncertainty in humidity can have nonlinear impacts as thresholds in physical parameterizations are crossed (e.g., convective initiation). SHUM uses the same random pattern generator as SPPT scheme but only a single spatial-temporal scale is used with maximum amplitude of 0.006. The perturbation is a maximum in the lowest model level and decreases exponentially with height. However, its impact transports rapidly to upper level of troposphere.

Since the new schemes (SPs), which is the combination of three schemes are introduced to replace the current operational STTP scheme in GEFS, an averaged spread at 120-hr forecast in two seasons are demonstrated in Fig. 3 to indicate the relative contribution and effect of each individual stochastic scheme. Compared to the ensemble spread without considering stochastic perturbations (i.e., noSP, top row), the control experiment with STTP (operational GEFS, the second row) produced extra spread in the extratropics area without major impact in the tropics. For the package of new stochastic physics schemes (SPs), however, an additional spread is produced through SKEB (3rd row), which has similar spatial contribution of STTP (2nd row). Both of the SPPT (row 4) and SHUM (last row) generate additional spread for tropics, but the evolution characteristics are slightly different (not shown here) from SHUM which only perturbs humidity in the near boundary. An increasing spread over tropics greatly improves forecast uncertainty representation and also enhances the tropical forecast skill.

Apparently, the SPs (combined three stochastic schemes) present forecast spread globally, especial for tropical area when compared to NCEP operational GEFS solo stochastic scheme (STTP). The statistical scores for tropical zonal winds of 850 and 250 hPa show huge improvements from introduced SPs for 2 years experiment periods (Fig. 4). For both of upper and lower atmosphere levels, increased forecast spread does also result in the reduced forecast error (root mean square error—RMSE; left plots of Fig. 4); and higher continuous ranked probability skill scores (CRPSS; right plots of Fig. 4). Moreover, the spreads are more closed to forecast errors (left plots of Fig. 4) indicates the representation of forecast uncertainty is more realistic than current operational GEFS (STTP).

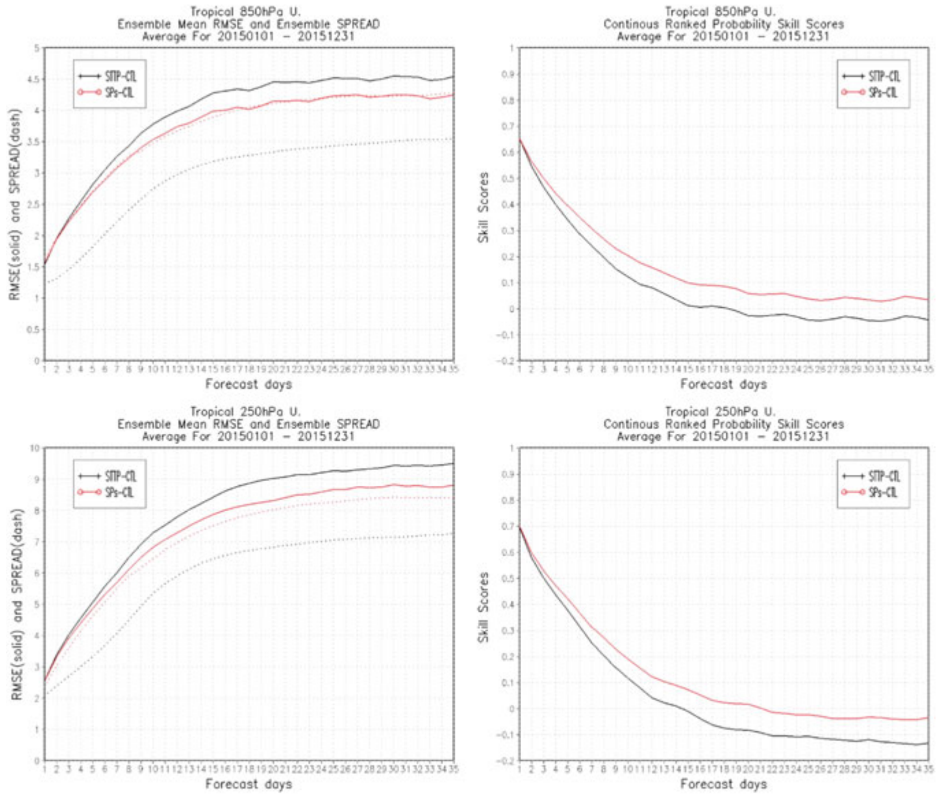


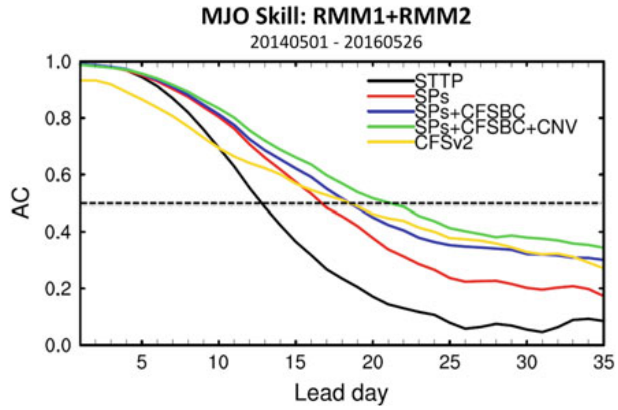
Fig. 4 RMS error of the ensemble mean (solid) and the ensemble spread (dash) (left), and CRPSS (right) are plotted every 24 h out to 35 days for 850-hPa (top) tropical (20°N–20°S) zonal wind during the Jan. 2015 to Dec. 2015 period comparing CTL (black) and SPs (red)

3 Other Strategies on Improving Ensemble Forecast on Sub-seasonal Timescale

The sub-seasonal forecast has different dependences from the short-term forecast. While the short-term forecast largely relies on the initial condition, the sub-seasonal forecast more and more relies on the boundary and external forcing. As such, for an uncoupled forecast system on sub-seasonal timescale, an accurate representation of the prescribed Sea Surface Temperature (SST) is of great importance (Li et al. 2001; Ling et al. 2015). The operational version of GEFS uses a prescribed SST that is initiated from analysis data and damps to climatology. Taking into account the day-to-day variability of the SST and as an intermediate stage between uncoupled and coupled forecast system, the underlying SST is updated using the bias-corrected SST from coupled model forecast (i.e., two-tiered SST, Table 1, third row).

As for the forecast system, an accurate representation of the physical process is critical to the forecast skill, the last strategy (or configuration) (Table 1, fourth row)

Fig. 5 MJO skills of the four different configurations of GEFS and CFSv2



that was tested is combining new SPs; two-tiered SST; and an upgraded Simplified Arakawa-Schubert (SAS) cumulus parameterization scheme that is both scale- and aerosol-aware (Han et al. 2017). The highlights of this upgraded convective parameterization scheme include: (1) the change of the rain conversion rate; (2) the modification of convective adjustment time in deep convection; (3) the cloud base mass flux in the shallow convection scheme becomes a function of mean updraft velocity; (4) convective inhibition (CIN) in the sub-cloud layer is an additional trigger condition to suppress unrealistic spotty rainfall; and (5) convective cloudiness is enhanced by suspended cloud condensate in an updraft.

The performance of the different GEFS configurations is demonstrated in Figs. 5 and 6. Since the Madden Julian Oscillation (MJO) is the dominant mode on the sub-seasonal predictability, MJO and its associated components are one of the emphases to evaluate the capability of the forecast system on sub-seasonal timescale. Compared to STTP scheme, the performance of the 850 hPa zonal wind over the tropics indicated a significant improvement associated with the increase of the spread in SPs (Fig. 4). The skill of the upper level zonal wind showed similar improvement (Figure not shown. please confirm), indicating a positive impact of the SPs on the MJO associated circulation. The RMM MJO skill increased from ~12.5 days in STTP scheme to 16.8 days in SPs. Combing SPs and updated SST further result in the increased MJO skill to 18.5 days. Combing SPs with updated SST and updated convection scheme lead to increase the MJO skill to 22 days (Fig. 5). The impact of the different configurations on the Northern Hemisphere large-scale circulation indicated the consistent result as the MJO (Fig. 6), with the improvement from STTP to SPs. The statistics, in terms of NH 500 hPa geopotential height anomaly correlations for average period of week-2 (days 8–14) and weeks 3 and 4 (days 15–28), include NCEP Climate Forecast System version 2 (CFSv2). The results indicate (1) All three new configurations show similar or better score than GEFS operation (ctl) for week-2, but much better than CFSv2; (2) All there new configurations demonstrate the very valuable skills for weeks 3 and 4 than GEFS operation (ctl), and much better than CFSv2.

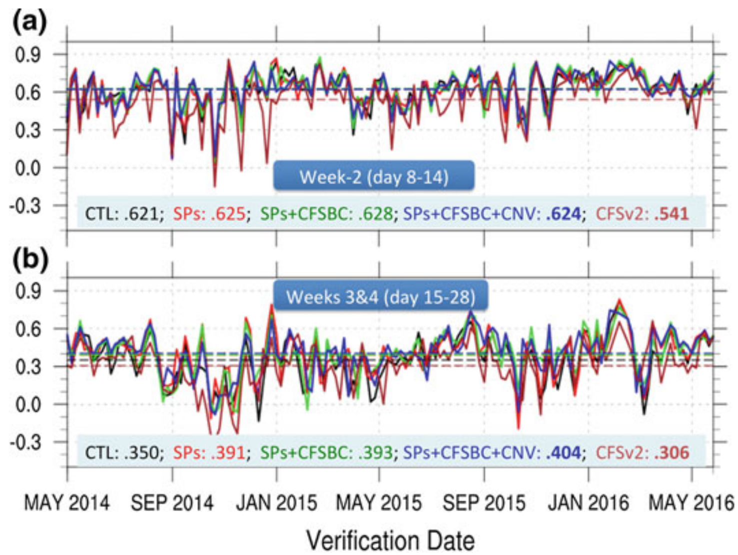


Fig. 6 The time series of ensemble mean anomaly correlation for Northern Hemisphere (20°N–80°N) 500-hPa geopotential height from May 2014 to May 2016 for different configurations (CTL-black; SPs-red; SPs + CFSBS-green and SPs + CFSBC + CNV-purple) and CFSv2 (orange) for lead week-2 (**a**) and weeks 3 and 4 (**b**). Days 15–28 (weeks 3 and 4 average). Average scores are shown in the bottom of each plot

4 Towards Physically Based Stochastic Parameterization

As we demonstrated in Sect. 2 for various stochastic perturbation schemes, most of them are in current operational ensemble forecast system that is still preliminary approach to assimilate model based uncertainties. There are many limitations in the application of the stochastic schemes. For example, the SKEB scheme highly depends on the accumulation of dissipation in the numerical integration, the horizontal and vertical diffusions scheme, gravity wave drag and mountain blocking parameterization et al. A contribution of SKEB will be greatly reduced when model resolutions are increased and when numerical schemes are improved. In addition, perturbation in SPPT varies with model physics thus varies with total physics tendency. The spatial and temporal de-correlation of the stochastic patterns thus does not really reflect uncertainty associated with individual physical process.

Figure 7 is a schematic diagram which demonstrates the current status of the stochastic perturbations, and the approach that represents the model uncertainties through realistic stochastic parameterization which is most possibly to be applied in the future. In the same time, two valuable studies have been done based on the operational ECMWF ensemble forecast system to apply (1) Independent random patterns to perturb different physical processes (or iSPPT). It is a similar procedure to current stochastic schemes (SPPT) but accounting stochastics for each individual physical process (Christensen et al. 2017); (2) Stochastic perturbed selected 20 phys-

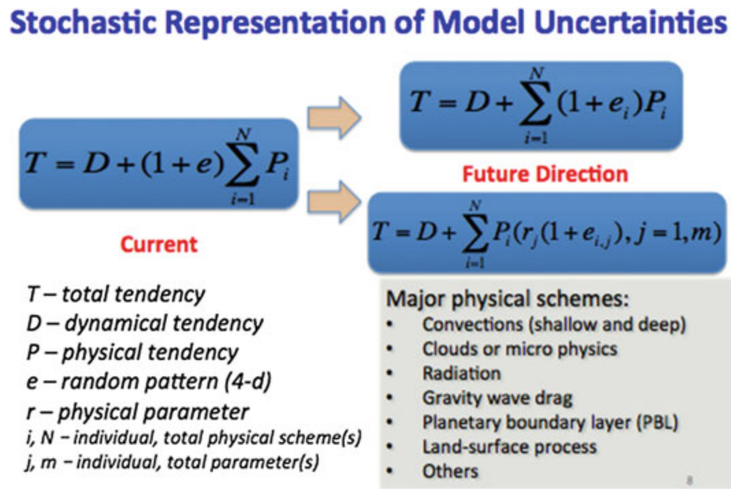


Fig. 7 Schematic diagram to present current status and future direction for stochastic representation of model uncertainties

ical parameters (SPP). Both of them increase ensemble spreads in general. The later one could change vertical distributions of forecast uncertainties significantly and thus may represent model uncertainties through the interaction of physical processes more realistically.

With the rapid progress in ensemble forecast system development and better understanding on the model physical process, the representation of forecast uncertainties from model dynamics and physics should be more approach to realistic atmosphere. Following this progress, many other sources of uncertainties, such as soil moisture and soil temperature from land model, sea surface temperature from ocean model, reflectivity of snow, and sea ice from sea ice model will be considered to improve weather forecast and sub-seasonal-climate prediction.

5 Summary

Stochastic perturbation is important processes that can help to improve sub-seasonal prediction after it succeeds for weather forecast. It advanced MJO skills significantly and associated tropical atmospheric circulation (850 and 200 hPa zonal winds, Li et al. 2018). It also enhanced extratropical prediction skills for weeks 3 and 4 average. In contrast to NCEP CFS v2 that is a coupling system with lower model resolution and older model physics, latest GEFS configuration has taken great advantage with new SPs, two-tiered SST and new convective parameterization in terms of tropical and extratropical, weather and sub-seasonal prediction. There are two areas we should focus on in near future: (1) Improve current physical tendency perturbation scheme to represent physical processes more realistically; (2) Consider other sources of uncertainties from land, sea, and other surface boundary.

Acknowledgements The authors would like to thank all of the helps from EMC ensemble team members, and Dr. Bing Fu helped to provide Figs. 2 and 3; Mr. Eric Sinsky provided Figs. 4 and 6 in particular. This study is partially supported through NWS's Office of Science and Technology Integration (OSTI) and NOAA's Climate Program Office (CPO)'s Modeling, Analysis, Predictions, and Projections (MAPP) program.

References

Berner, J., G.J. Shutts, M. Leutbecher, and T.N. Palmer. 2009. A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *Journal of the Atmospheric Sciences* 66 (3): 603–626.

Buizza, R., M. Miller, and T. Palmer. 1999. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal Royal Meteorological Society* 125 (560): 2887–2908.

Christensen, H.M., S.-J. Lock, I.M. Moroz, and T.M. Palmer. 2017. Introducing independent patterns into the stochastically perturbed parametrization tendencies (SPPT) scheme. *Quarterly Journal of the Royal Meteorological Society, Part A* 143 (706): 2168–2181.

Han, J., W. Wang, Y.C. Kwon, S.-Y. Hong, V. Tallapragada, and F. Yang. 2017. Updates in the NCEP GFS cumulus convection schemes with scale and aerosol awareness. *Weather and Forecasting*. https://doi.org/10.1175/WAF-D-17-0046.1.

Hou, D., Z. Toth, Y. Zhu, and W. Yang. 2008. Evaluation of the impact of the stochastic perturbation schemes on global ensemble forecast. In *Proceedings of the 19th conference on probability and statistics*, New Orleans, LA, American Meteor Society. https://ams.confex.com/ams/88Annual/webprogram/Paper134165.html.

Johnson, N.C., D. Collins, S. Feldstein, M. L'Heureux, and E. Riddle. 2014. Skillful wintertime North American temperature forecasts out to 4 weeks based on the state of ENSO and the MJO. *Weather and Forecasting* 29: 23–38. https://doi.org/10.1175/WAF-D-13-00102.1.

Kirtman, B.P., D. Min, and J.M. Infanti. 2014. The North American multimodel ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. *Bulletin of the American Meteorological Society* 95: 585–601.

Li, W., R. Yu, H. Liu, and Y. Yu. 2001. Impacts of diurnal cycle of SST on the intraseasonal variation of surface heat flux over the western Pacific warm pool. *Advances in Atmospheric Sciences* 18 (5): 793–806.

Li, W., Y. Zhu, X. Zhou, D. Hou, E. Sinsky, C. Melhauser, M. Pena, H. Guan, and R. Wobus. 2018. Evaluating the MJO prediction skill from different configurations of NCEP GEFS extended forecast. *Climate Dynamics*, https://doi.org/10.1007/s00382-018-4423-9.

Ling, T., M. Xu, X.-Z. Liang, J.X.L. Wang, and Y. Noh. 2015. A multilevel ocean mixed layer model resolving the diurnal cycle: Development and validation. *Journal of Advances in Modeling Earth Systems* 07. https://doi.org/10.1002/2015ms000476.

Liu, X., T. Wu, S. Yang, T. Li, W. Jie, L. Zhang, Z. Wang, X. Liang, Q. Li, Y. Cheng, H. Ren, Y. Fang, and S. Nie. 2016. MJO prediction using the sub-seasonal to seasonal forecast model of Beijing Climate Center. *Climate Dynamics*. https://doi.org/10.1007/s00382-016-3264-7.

Lorenz, E. 1969. The predictability of a flow which possesses many scales of motion. *Tellus* 21: 289–307. https://doi.org/10.1111/j.2153-3490.1969.tb00444.x.

Palmer, T.N., et al. 2004. Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER). *Bulletin of the American Meteorological Society* 85: 853–872.

Palmer, T.N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. Shutts, M. Steinheimer, and A. Weisheimer. 2009. Stochastic parametrization and model uncertainty. Technical Report ECMWF RD Tech. Memo. 598, 42 pp. http://www.ecmwf.int/publications/.

Shin, D.W., and T.N. Krishnamurti. 2003. Short- to medium-range superensemble precipitation forecasts using satellite products: 1. Deterministic forecasting. *Journal of Geophysical Research* 108 (D8): 8383. https://doi.org/10.1029/2001jd001511.

Shutts, G. 2005. A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal Royal Meteorological Society* 131: 3079–3102.

Shutts, G. 2015. A stochastic convective backscatter scheme for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society: Part A* 141 (692): 2602–2616.

Shutts, G., and T.N. Palmer. 2004. The use of high-resolution numerical simulations of tropical circulation to calibrate stochastic physics schemes. In *Proceedings of the ECMWF/CLIVAR simulation and prediction of intra-seasonal variability with emphasis on the MJO*, Reading, United Kingdom, European Centre for Medium-Range Weather Forecasts, 83–102.

Tian, D., Eric F. Wood, and X. Yuan. 2017. CFSv2-based sub-seasonal precipitation and temperature forecast skill over the contiguous United States. *Hydrology and Earth System Sciences* 21: 1477–1490.

Tompkins, A.M., and J. Berner. 2008. A stochastic convective approach to account for model uncertainty due to unresolved humidity variability. *Journal Geophysical Research* 113: D18101.

Troccoli, A. 2010. Seasonal climate forecasting. *Meteorological Applications* 17: 251–268. https://doi.org/10.1002/met.184.

Vitart, F. 2014. Evolution of ECMWF sub-seasonal forecast skill scores. *Quarterly Journal of the Royal Meteorological Society* 140: 1889–1899. https://doi.org/10.1002/qj.2256.

Zhou, X., Y. Zhu, D. Hou, and D. Kleist. 2016. Comparison of the ensemble transform and the ensemble Kalman filter in the NCEP global ensemble forecast system. *Weather and Forecasting* 31: 2058–2074.

Zhou, X., Y. Zhu, D. Hou, Y. Luo, J. Peng, and D. Wobus. 2017. The NCEP global ensemble forecast system with the EnKF initialization. *Weather and Forecasting* 32: 1989–2004.

Zhu, Y.X., M. Zhou, W. Pena, C.Melhauser Li, and D. Hou. 2017. Impact of sea surface temperature forcing on weeks 3 and 4 forecast skill in the NCEP global ensemble forecasting system. *Weather and Forecasting* 32: 2159–2173. https://doi.org/10.1175/WAF-D-17-0093.1.

Zhu, Y., X. Zhou, W. Li, D. Hou, C. Melhauser, E. Sinsky, M. Pena, B. Fu, H. Guan, W. Kolczynski, R. Wobus, and V. Tallapragada. 2018. An assessment of subseasonal forecast skill using an extended global ensemble forecast system (GEFS). *Journal of Geophysical Research* 6732–6745. https://doi.org/10.1029/2018JD028506.