Generation of IC Perturbation in GSI -GEFS System

V.S. Prasad vsprasad@ncmrwf.gov.in

Ensembles Technique at NCMRWF

- GFS based Global Data Assimilation and Forecasting (NGFS) is NCMRWF's legacy system and is being run on a routine basis since 1994.
- In this system NCMRWF employ Ensemble techniques in:
 a) for generation Ensemble forecasts –GEFS
 - b) for the improvement of IC's
- The work ensemble forecasting was initiated in mid 90's in collaboration with NCEP (Kalnay and others).
- Iyengar et al., 1996; Iyengar et al., 2004 and Kar et al., 2011 made initial attempts at NCMRWF to develop ensemble forecasting system using Breeding of Growing Modes (BGM) technique developed by Kalnay

Finally, a Global Ensemble Forecasting System (GEFS) at T190L28 resolution (about 70 km in horizontal resolution and 28 levels in the vertical) with 20 members has been implemented at NCMRWF for near real time medium range forecasting in 2012. GEFS is initialized by using then operational T574 deterministic analysis and forecast outputs (Prasad et al., 2011) by employing ETR (Wei et al., 2008) and STTP techniques.

	İ				
S. No	HPC (with Performance)	Period	GDAF System		
1	Cray – XMP (4 Mflops)	1988-1993	T80L18 with SSI		
2	Cray – YMP (16 Mflops)	1993-1998			
3	Dec-Alpha (9.6 Gflops)	1999-2011			
4	Cray SV1 (28.8 Gflops)	2001-2006			
5	PARAM (500 Gflops)	2006-2011	T254L64 GSI		
6	Cray – X1E (1.1 Tflops)	2006-2011	T382L64 GSI		
7	IBM – Power 6 (24 Tflops)	2010 onwards	T574L64 GSI		
8	IBM — Idataflux (350Tflop)	2015 onwards	Jan 2016	T574L64 3DHybrid T254(ENS),Eularian	
			Jun 2016	T1534L64,3D-Hybrid T574(ENS), semi-Lag	
9	Cray – XC40 (2.8 Pflops)	2018 onwards	18Jul 2018	T1534L64 4D-ENS-Var,SemiLag	

Schematic diagram of GEFS T192L28 - operational at NCMRWF in 2012

(Wei et al., 2008), BV-ETR

- 1. The 20 perturbed initial conditions are run for 6 forecast hours.
- 2. Differences of these forecasts wrt deterministic analysis are determined.
- The differences are then made "orthogonal" and then re-scaled to a size representative of the known observational and first guess errors.
- The cycle repeats every 6-hour, breeding a new 20-member set of independent perturbations as illustrated.

To emulate the effect of uncertainty in NWP models on the forecast variables, temperature, specific humidity, & winds, the Stochastic Total Tendency Perturbation (STTP) scheme is applied.



First Time generated probabilistic forecast products and many other ensemble post-processing products.



Day-5 forecast of 24-hr rainfall valid for 28th Jul 2012 (top left) over RSMC region. The probability of the rainfall in the 1-2cm/day range (top right), 2-5cm/day range (bottom left) and 5-10 cm/day range (bottom

Hybrid-3Dvar Analysis (Wang et al 2013)





Prasad & Johny 2016, JESS, 125, 521-538

Hybrid EnsVar assimilation system using 80 member EnKF ensembles



Prasad et al 2016, JESS, 125, 1509-1521

Percentage improvement in Wind RMSE (computed against radiosonde observations) over 3D Var in 24 hr forecasts of hybrid schemes (EnKF, ETR) at 850, 500 and 200 hPa Pressure levels

Pressure (hPa)	EnKF	ETR
850	6.5	1.03
500	8.3	1.36
200	10.85	1.90

Using these ensembles forecasts and the deterministic analysis, from Hybrid-ENKF system a Medium Range forecasting- GEFS system was devleoped.

IC for GEFS at NCMRWF

GFS Model	Eularian	Eularian	Semi- lagrangian	Semi- lagrangian
Physiscs	GFS-V12	GFSV-12	GFS-V13	GFS-14
Initial Perturbation	BV-ETR	ENKF-06	ENKF-06	ENKF-06
Model Uncertainity	STTP	STTP	STTP	STTP
Resolution	T192/L28	T254L64	T574	T1534
Period	Experimental	Jan2016- May2016	Jun2016-17Jul 2018	18Jul 2018 -

The skill of NWP forecasts from NCMRWF, improved substantially around 2011 and it is attributed to the increase in the resolution of the model, increase in the amount of data being assimilated, improvements in data assimilation techniques.

Verification of Day 03 Forecasts against Indian Radiosondes Root Mean Square Error (RMSE) of 850 hPa winds in m/s (2005-2015) The anomaly correlation values are comparatively higher in the T574 GFS with a gain of 1 day in the skill of the forecasts.



A reanalysis & Retrospective forecast for a 20-year period 1999-2018 was carried out with this setup. Prasad et al et al ,2017; Current Science, 112(2), 370-377

Ensembles from Retrospective Anal.

- 10 member Ensembles were also created using IC's from the retrospective analysis by using additive inflation technique proposed by Whitaker and Hamill (2012).
- Additive inflation, the additive noise was drawn from a full year's inventory of differences between 48- and 24-h forecasts valid at the same time. The additive perturbations were applied to the IC's
- These parameters were tuned so that the average
- background ensemble spread matched the averagebackground errors.
- Ankur and Surya et al used these ensembles in seasonal forecasting.

Methods of ensemble generation



- Burst Ensemble- Members of a large ensemble (e.g., 50 members are initialized from the same time)
- Lagged ensemble- Smaller set of model integrations are started from consecutive initial times

Srivastava, A., Rao, S.A., Pradhan, M. *et al.* Gain of one-month lead time in seasonal prediction of Indian summer monsoon prediction: comparison of initialization strategies. *Theor Appl Climatol* (2020). https://doi.org/10.1007/s00704-020-03470-3

Experiment design



Hindcast period: 2003-2017 CTL run (L)

- Lagged initialization
- Feb, Mar & Apr ICs
- 10-12 members

Sensitivity run (P)

- Perturbed initialization
- 10 members
- Initialized on 01st
 of Feb, Mar & Apr

AISMR Skill





SKILL	FEB (P)	MAR (P)	APR (P)	FEB (L)	MAR (L)	APR (L)
GPCP	0.31	0.49	0.37	0.23	0.63	0.55
IMD	0.47	0.65	0.49	0.37	0.64	0.44

- For all initializations, the lagged and perturbed ICs have comparable skill scores
- Perturbed Feb IC is somewhat better compared to lagged initialization approach.

Model Skill & Teleconnections

			Skill			Teleconn	ections	
	MEA N	SD	ISMR (GPCP)	ISMR (IMD)	Nino 3.4	IODE	ISMR vs Nino 3.4	ISMR vs IODE
GPCP	6.92	0.5 9	-	-	-	-	-0.73	-0.24
IMD	7.45	0.6 3	-	-	-	-	-0.67	-0.27
FEB (P)	3.95	0.3 7	0.31	0.47	0.50	0.70	-0.31	0.55
MAR (P)	4.06	0.3 2	0.49	0.65	0.66	0.63	-0.39	0.46
APR (P)	4.26	0.4 7	0.37	0.49	0.80	0.77	-0.55	0.71

Mean monsoon rainfall is similar across both lagged & perturbed model runs.

- FEB (L) & MAR (P) Nino 3.4 skill is slightly better but is similar for other ICs
- FEB (P) & MAR (L) IODE skill is better

ENSO teleconnections have become stronger in the recent period 2003-2017

Teleconnections with ENSO are surprisingly weak in the model for 2003-2017

IOD teleconnections are opposite to what is expected

Monthly Model Skill: Nino 3.4



- At longer leads (FEB IC), the Nino 3.4 skill is higher in the later months for lagged method
- Comparable monthly evolution of skills at shorter leads

Monthly Model Skill: IOD East pole



- At longer leads (FEB IC), the IOD skill is higher in the later months for perturbed method
- At shorter leads , the skill in the initial 2-3 months is higher for lagged method

Mean biases: SST & Rainfall



Mean biases in SST & rainfall are almost similar

Conclusions

- Both Burst and Lagged ensembles result in similar statistics for all India Summer Monsoon Rainfa II
- One advantage of the perturbed ICs is that one do not h ave to wait for the whole month observations (Initiate all ensemble run s on 1st of each month)
- Both ICs resulted in RMSE >> Spread for AISMR (i.e under dispersive). For SST indices RMSE and Spread are almost equal
- The question of how to improve spread in ensembles is s till eluding us for AISMR

Ensemble spread of Nino 3 SST forecasts



Ensemble spread of SST forecasts as a function of lead time: The daily evolution of the spread during the first month

Vialard et al.

LA – Lagged average	SP – Stochastic physics
WP – Wind	SWT – WP+TP+SP
perturbation	NDA – No data
TP - Temperature	assimilation

- WP & SP display a different behavior
- Strong relaxation to observed SST during the analysis prevents any significant spread in SST at initial time in WP.
- TP, SWT, and LA provide a better estimate of the uncertainties during the early range of the forecast than WP and SP.
- In TP and SWT, there is a slight decrease of the spread during the first week of the forecast that might correspond to the poisy

Ensemble spread of Nino 3 SST forecasts



Since El Niño predictability is generally believed to stem from the knowledge of oceanic initial conditions, one would normally expect uncertainties in oceanic initial conditions such as those generated in the WP experiment, to give rise to corresponding uncertainties in El Niño forecasts.

However, the spread in experiment WP (with uncertainties in initial conditions) is

Since "burst ensembles" allow a more timely delivery of the forecasts than LA, they should be preferred to LA in an operational system.