Stochastically Perturbed Parameterization Scheme for the Soil Temperature and Moisture with an Optimized Tuning Parameters within an Ensemble Data Assimilation System

Sujeong Lim[1], Seon Ki Park[1,2,3*], and Claudio Cassardo[4]

[1]Center for Climate/Environment Change Prediction Research, Ewha Womans University, Seoul, Republic of Korea
[2]Severe Storm Research Center, Ewha Womans University, Seoul, Republic of Korea
[3]Department of Climate and Energy System Engineering, Ewha Womans University, Seoul, Republic of Korea
[4]Department of Physics and NatRisk Centre, University of Torino, Torino, Italy

EWHA WOMANS UNIVERSITY

Motivation:

- In the ensemble data assimilation (EDA) system, the model uncertainties can be represented by the ensemble spread --- a standard deviation of background error covariance (BEC).
- The insufficient ensemble spread can cause the analysis to ignore the observation.
- This underestimation is also found in the coupled land-atmospheric modeling system, especially near the surface where the heat flux exchanges are crucial as the lower boundary conditions.

Research Objective:

I. **To develop** the stochastically perturbed parameterization (SPP) schemes for the Noah land surface model (Noah LSM) using soil temperature and soil moisture within the coupled WRF-Noah LSM system to represent the under-estimated ensemble spread in near-surface

II. **To optimize** the random forcing tuning parameters in the SPP-Noah LSM using a global optimization algorithm instead of enormous sensitivity experiments

III. **To inflate** the ensemble BEC using the optimized SPP-Noah LSM in the EDA system

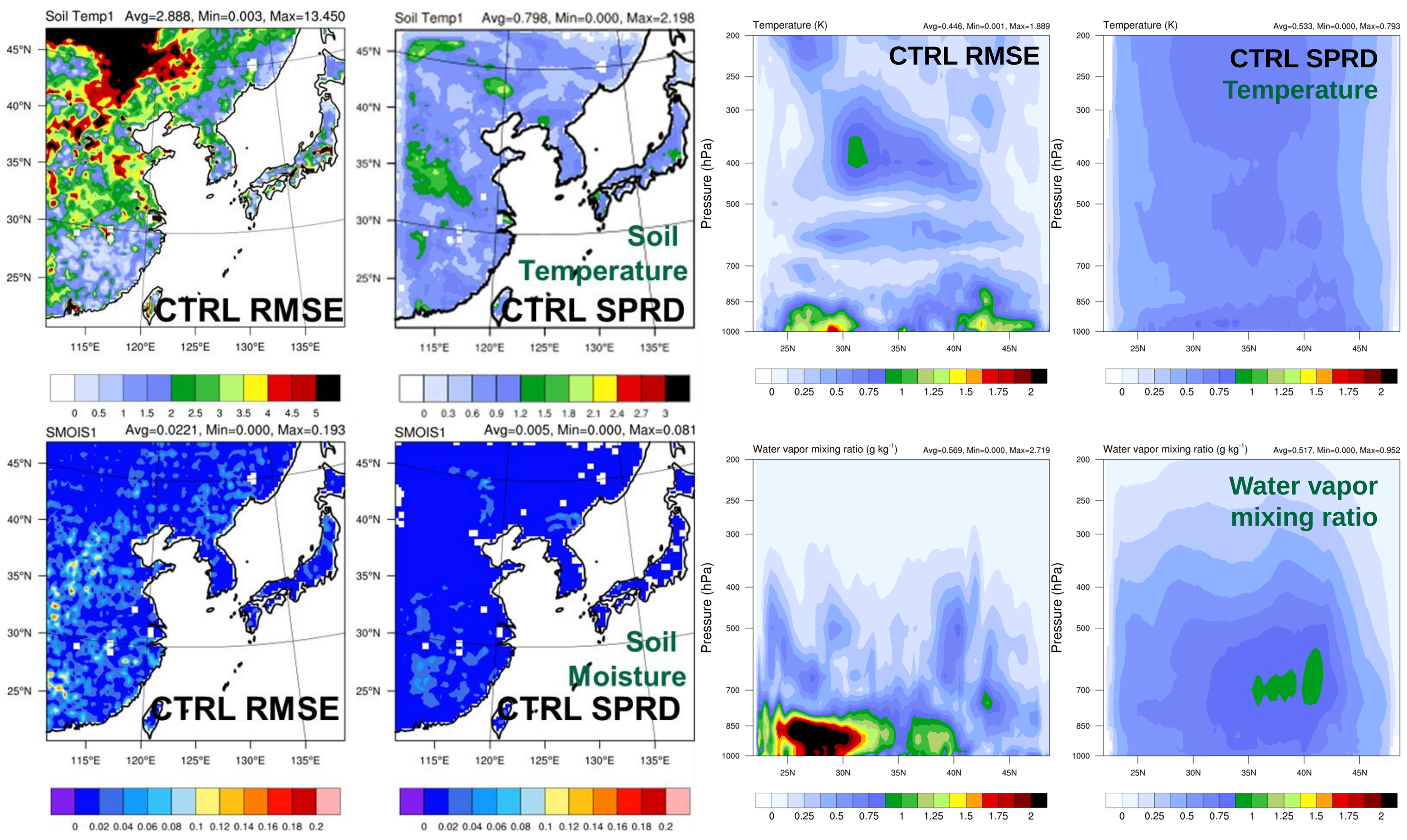

Fig 1. (Left) Ensemble mean RMSE and ensemble mean spread for soil temperature (top) and soil moisture (bottom) from the CTRL, (Right) ensemble mean RMSE and ensemble mean spread for zonal mean temperature (top) and water vapor mixing ratio (bottom) from the CTRL over the land.

SPP scheme for the Noah LSM (SPP-Noah LSM)

- It perturbs soil temperature (ST) or soil moisture (SM) using a spatially and temporally correlated random forcing at each grid point every time steps within the coupled WRF-Noah LSM system to represent the near-surface uncertainty.

$$x_{i,new}^1 = x_i^1 + r_i$$

x_i^1: ST or SM at the first soil layer (0 ~ 10 cm), i: ensemble member
r: Random forcing ($-\sigma \le r \le \sigma$) following Gaussian distribution with zero mean
(The perturbations sum up to zero so as not to introduce a systematic drift in the model)

- Additional prescription to the SM perturbing:
  1) New SM falls within the respective bounds sets by the wilting point and the saturation level
  2) The perturbations in areas under snow cover or with frozen soil are set to zero.

r is a function of tuning parameters; amplitude, decorrelation length and time scale

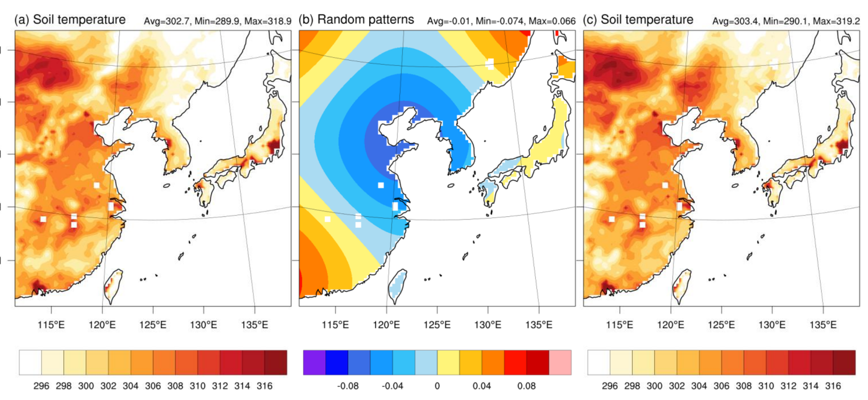

Fig 2. (a) original ST, (b) random forcing, and (c) new ST (K).

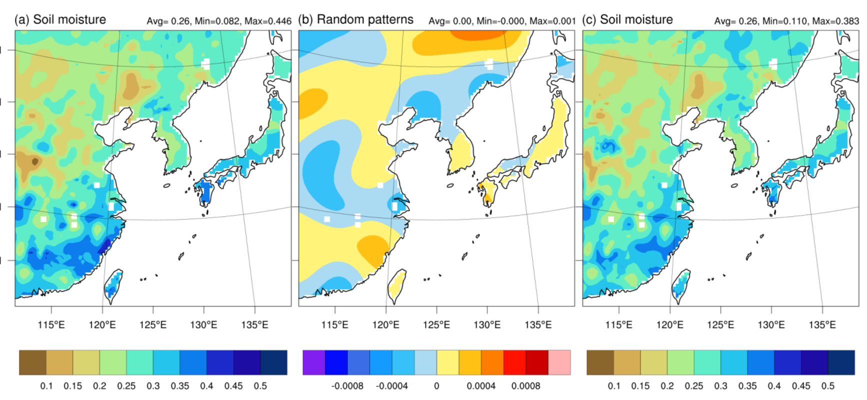

Fig 3. (a) original SM, (b) random forcing, and (c) new SM (m^3 m^{-3}).

Micro-Genetic Algorithm (μ-GA)

- μ-GA is based on the natural selection or survival of the fittest to evolve the best potential solution over several generations to the most-fit.
- We design a fitness function using the normalized mean squared errors (MSE) to explain the interaction between LSM and the planetary boundary layer (PBL) in terms of accuracy as:

$$\textit{fitness function} = \frac{MSE(\boldsymbol{x})}{\sigma(\boldsymbol{x})^2_{\text{reference}}} + \frac{1}{7}\sum_{k=1}^{z} \frac{MSE(\boldsymbol{y_k})}{\sigma(\boldsymbol{y_k})^2_{\text{reference}}}$$

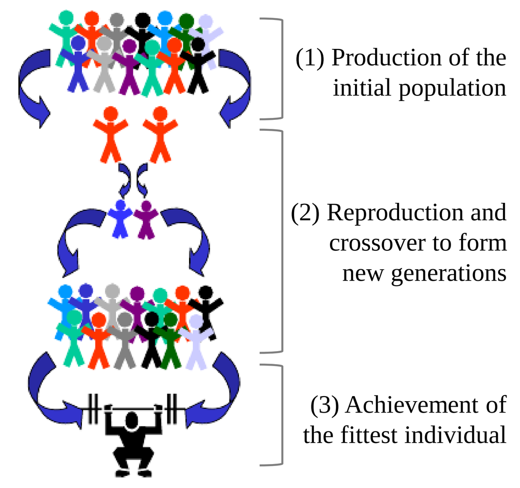

Fig. 4. Illustration of a simple GA.

where $\boldsymbol{x}$ is ST (K) or SM (m^3 m^{-3}), $\boldsymbol{y}$ is temperature (K) or water vapor mixing ratio (g kg^{-1}), $\boldsymbol{z}$ represents vertical layers from 850 to 1000 hPa with 7 levels, and $\sigma_{reference}$ is the standard deviation of reference data (GFS analysis).

- We consider diurnal variations of soil variables and optimize the each daytime and nighttime tuning parameters using 5 ensembles at 00/12 UTC 1 Aug 2018.

Table 1. Optimized ST parameters for daytime and nighttime (OSTP-D, OSTP-N). Same for the SM (OSMP-D, OSMP-N).

Tuning parameters (ranges)	OSTP-D	OSTP-N	OSMP-D	OSMP-N
Amplitude (0.01~0.64)	0.13 K	0.01 K	0.003 m^3 m^{-3}	0.003 m^3 m^{-3}
Length scale (50~3200 km)	2900 km	100 km	250 km	700 km
Time scale (0~900 s)	120 s	900 s	900 s	120 s

BEC Inflation in EDA System Using an Optimized SPP-Noah LSM

- GSI/Ensemble Kalman Filter (EnKF) assimilates the PREPBUFR (conventional data) using 27 ensembles
- Experimental period: 2018.08.01.06 UTC ~ 2018.08.07.00 UTC (*spin-up: ~ 08.03.18 UTC)
  - CTRL: Control DA cycles
  - STP1 (or SMP1): BEC inflation by perturbing the ST or SM using the daytime tuning parameters
  - STP2 (or SMP2): BEC inflation by perturbing ST or SM using the day/nighttime tuning parameters, considering the diurnal variations

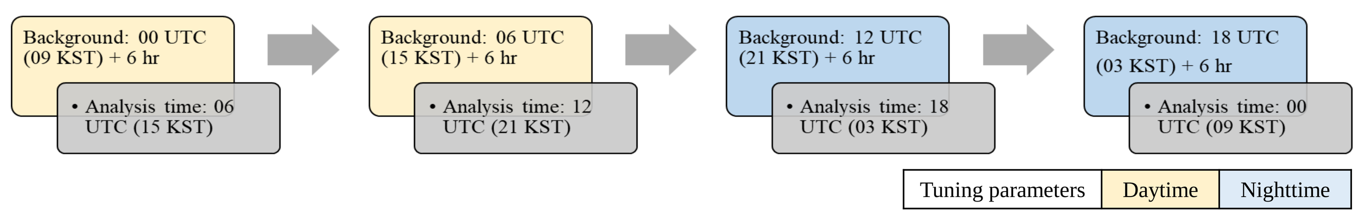

Fig 5. Example: schematic diagram of STP2 (or SMP2) in DA cycling system.

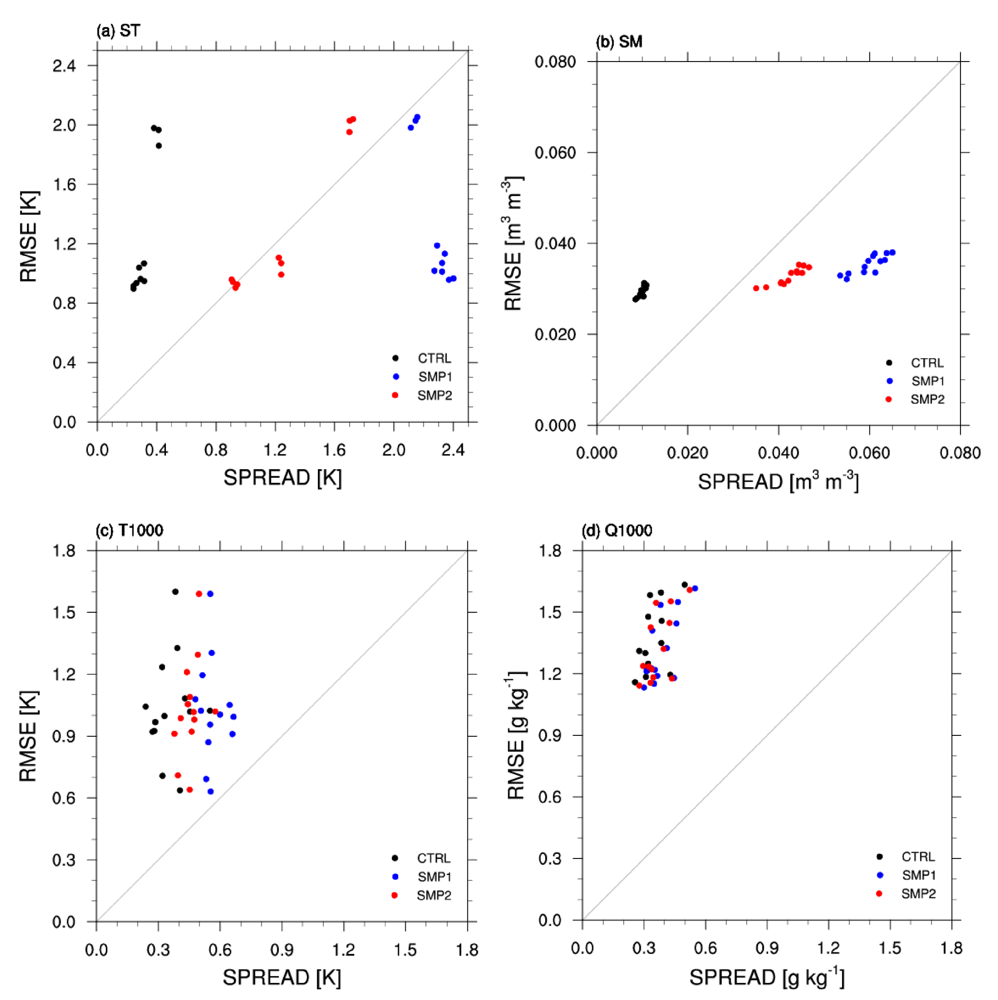

Fig 6. Scatter plots of the ensemble mean spread and RMSE for (a) ST, (b) SM, (c) temperature at 1000 hPa (T1000), and (d) water vapor mixing ratio at 1000 hPa (Q1000).

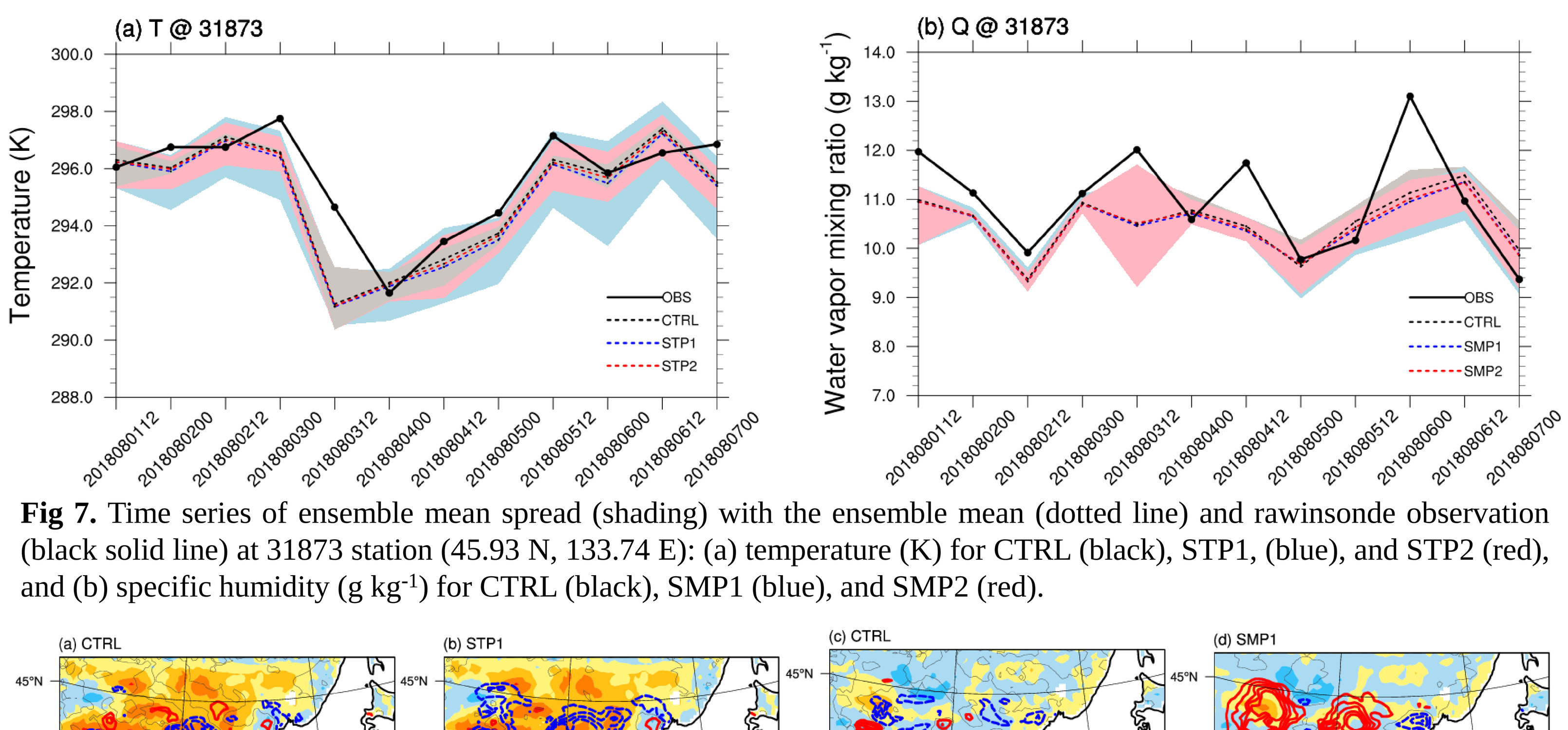

Fig 7. Time series of ensemble mean spread (shading) with the ensemble mean (dotted line) and rawinsonde observation (black solid line) at 31873 station (45.93 N, 133.74 E): (a) temperature (K) for CTRL (black), STP1, (blue), and STP2 (red), and (b) specific humidity (g kg^{-1}) for CTRL (black), SMP1 (blue), and SMP2 (red).

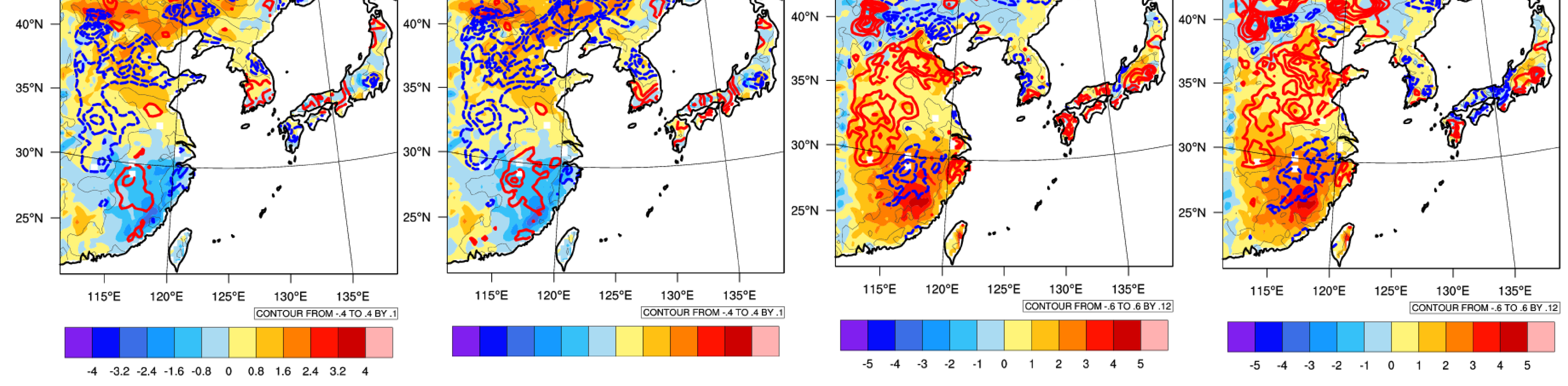

Fig 8. The analysis increment (colored contours; positive in red, negative in blue, and zero in gray) and the background error against GFS analysis (shaded) for temperature (in K) in (a) CTRL and (b) STP1, and for water vapor mixing ratio (in g kg^{-1}) in (c) CTRL and (d) SMP1. Results are averaged from 850 hPa to 1000 hPa.

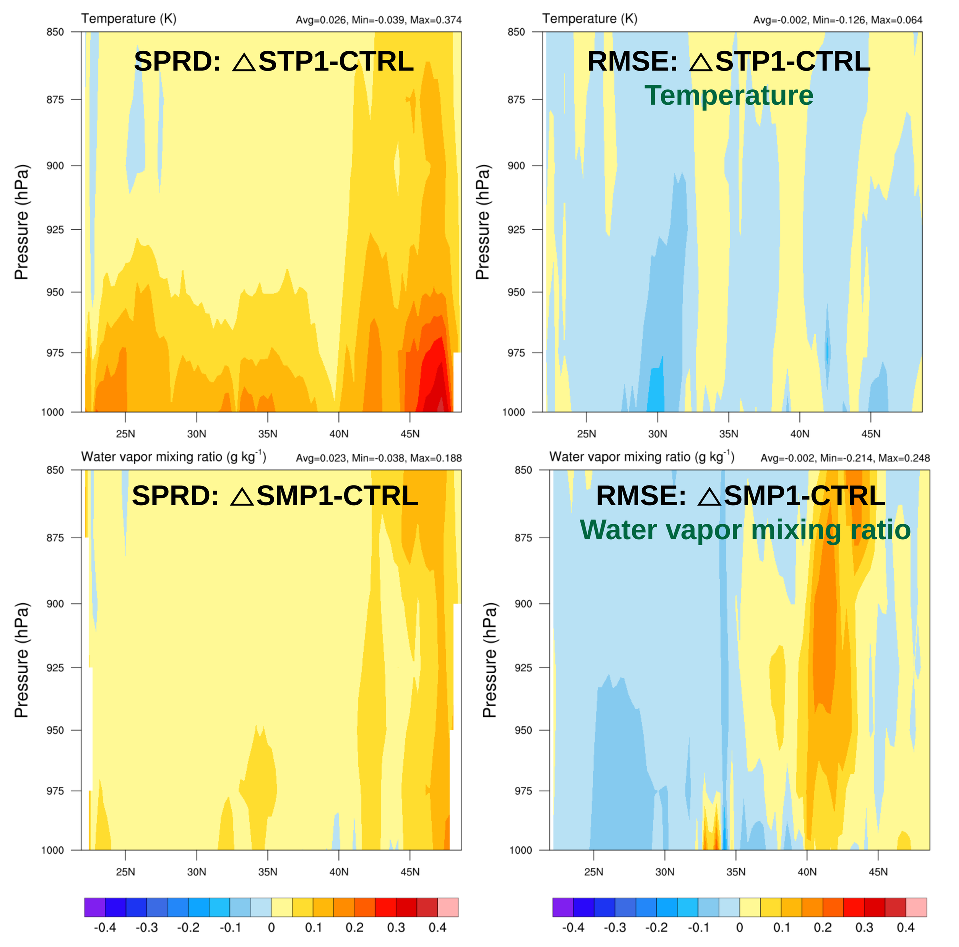

Fig 9. (top) Δ ensemble mean spread (STP1-CTRL) and Δ ensemble mean RMSE (STP1-CTRL) for temperature (K) over the land. (bottom) Δ ensemble mean spread (SMP1-CTRL) and Δ ensemble mean RMSE (SMP1-CTRL) for water vapor mixing ratio over the land.

Summary

- ❖ The ST and SM perturbations can indirectly inflate the ensemble BECs of temperature and water vapor mixing ratio in the PBL of the EDA system.
- ❖ The SPP-Noah LSM with diurnal variations depicts reasonable ensemble spreads for soil variables, but the ensemble spreads for atmospheric variables are less effective.
- ❖ The inflated ensemble spread helps to include more observations and to produce an adequate analysis increment reducing the background error in the PBL.
- ❖ The SM perturbation requires additional prescriptions to prevent inadequate analysis increments and RMSE increase.

Future Plans

- We need to include sufficient cases for optimization and various observations to evaluate the fitness as for the soil as well as atmospheric variables.
- We will investigate the impacts of SPP-Noah LSM in the coupled land-atmosphere data assimilation system.

Acknowledgement This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT) (No. NRF-2021R1A2C1095535). S. Lim was also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2020R1A6A3A13069223).

EWHA WOMANS UNIVERSITY CCCPR SSRC UNIVERSITÀ DEGLI STUDI DI TORINO

Workshop on model uncertainty, 9-12 May 2022, ECMWF, Reading