Federal Service for Hydrometeorology and Environmental Monitoring

HYDROMETEOROLOGICAL CENTRE OF RUSSIA

About weather—at first hand

Medium range ensemble prediction system of Hydrometcenter of Russia based on SL-AV model: First results

Vasily Mizyak [1], Mikhail A. Tolstykh[2,1], Vladimir Rogutov[1], Ksenia Alipova[2], Gordey Goyman[2]

[1] Hydrometeorological Research Centre of Russia, Moscow, Russia
[2] Institute of Numerical Mathematics RAS, Moscow, Russia

vmizyak@mecom.ru, m.tolstykh@inm.ras.ru. rogutovv@mail.ru , ksusha_ast@mail.ru, gordeygoyman@gmail.com

CONFERENCE ON COMPUTATIONAL INFORMATION TECHNOLOGIES FOR ENVIRONMENTAL SCIENCES IN COMMEMORATION OF RAS CORRESPONDING MEMBER VASILY LYKOSOV
CITES '2021
22 – 26 NOVEMBER 2021
MOSCOW, RUSSIA

Ensemble Weather Forecast

- *Allows taking into account the uncertainties in the representation of the initial state of the atmosphere and in the numerical model*
- *It requires much more computational resources compared to the deterministic forecast*
- *It is necessary to have a slightly different forecasts (different models, or different simulation parameters of the same model, or **<u>different initial data</u>** for the same model, etc)*

Components of the system

1. Global Atmosphere Model SL-AV

- *Operational use in the Hydrometcenter of Russia*
- *0.9° x 0.72° horizontal resolution and 96 vertical levels version*
- *Semi-Lagrangian finite-difference dynamical core*
- *Version with Stochastically Perturbed Parametrizations (SPP)*

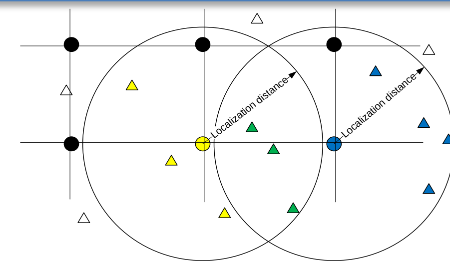
2. Data assimilation system based on LETKF

- *in-situ observations:*
  - ✓*SYNOP, SHIP – reports at observing stations over land and marine surface data from ships (Ps, T2m, RH2m, about 10000 observations)*
  - ✓*TEMP – radiosondes at 32 vertical levels (T, RH, U, V, about 2000 observations)*
  - ✓*AIREP – aircraft observations (T, U, V, about 10000 observations)*
- *some kinds of satellite observations (but not most useful):*
  - ✓*AMV – satellite cloud motion observations (U, V, 500000 observations)*
  - ✓*ASCAT – satellite sea surface winds speed observations (400000–700000 observations per day)*
- *multiplicative and additive inflation*

3. Ensemble verification system

- *Scores recommended by the Lead Centre on Verification of Ensemble Prediction (Brier Score, Relative Operating Characteristic, Continuous Ranked Probability Score, etc.)*

Data assimilation system – Local Ensemble Transform Kalman Filter



- Observations localization:
  - only observations from localization radius are used with, weights are depending on distance to grid point
  - high parallel efficiency because of independent performing at every grid point
- ensemble of forecasts is used for covariance matrix estimation

$$\mathbf{P}^b = \frac{1}{k-1}\mathbf{X}^b\mathbf{X}^{bT}$$

- transforming to k-dimensional ensemble space → solving the problem → inverse transforming to n-dimensional model space
- analysis ensemble can be used as initial data for the ensemble forecast

Satellite upper-air wind observations AMV

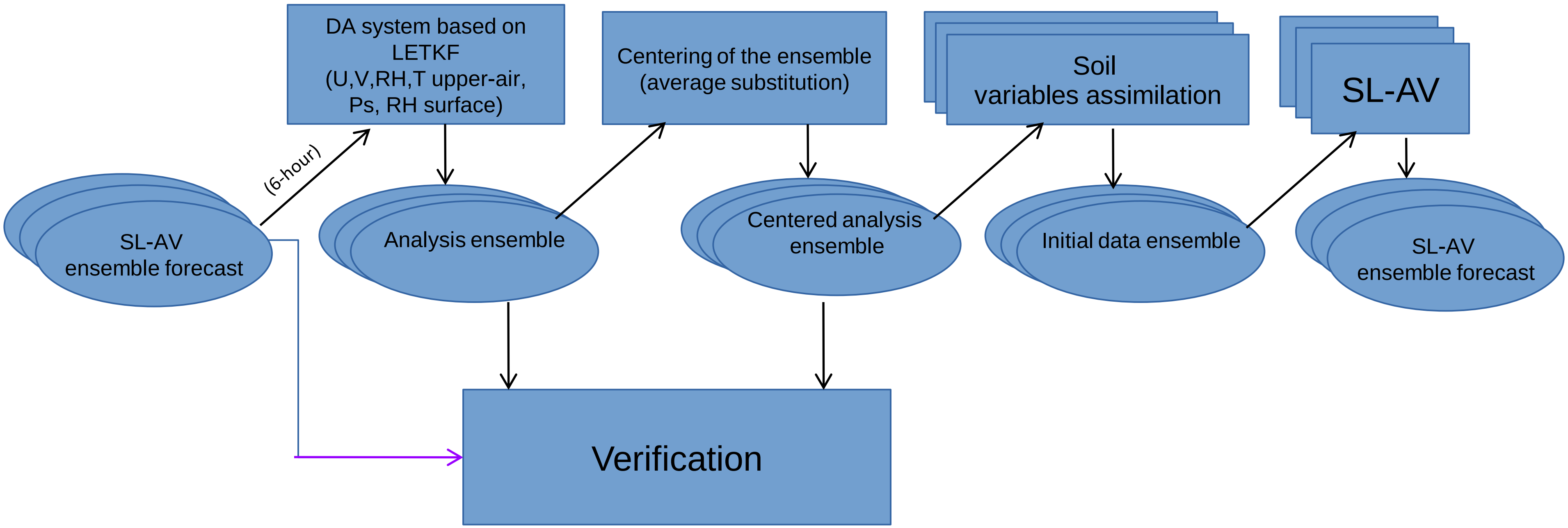
- *obtained by tracking motion of identical structures clouds or water vapor on two or more consecutive images*
- *global coverage*
- *measures the speed and wind direction in the upper-air*
- *the only source of upper-air wind information over the oceans, in the tropics and in the polar regions, have a global coverage*
- *it is difficult to obtain proper height of observation due to 2d image and therefore errors while height assignment are arises – height reassignment has performed*

Satellite sea-surface wind observations ASCAT

- *ASCAT (Advansed SCATerometer) data available from METOP-a and METOP-b polar orbiting satellites*
- *every device measures surface wind speed in 2 550-km swaths which are separated from the satellite ground track by about 336 km for the minimal orbital height*
- *data resolution – 25km*
- *ASCAT grouped into superobservations with grid scale comparable to the SLAV model resolution. Observations were rejected: in inner lakes, near coast line, with high model orography and low surface pressure*

Ensemble prediction system

- The ensemble deviation from the ensemble mean analysis is calculated
- Each deviation is added to HMC operational analysis
- The modified ensemble is used to obtain an ensemble of forecasts by model simulation
- The mean analysis coincides with the operational one and the spread is generated in the ensemble assimilation system



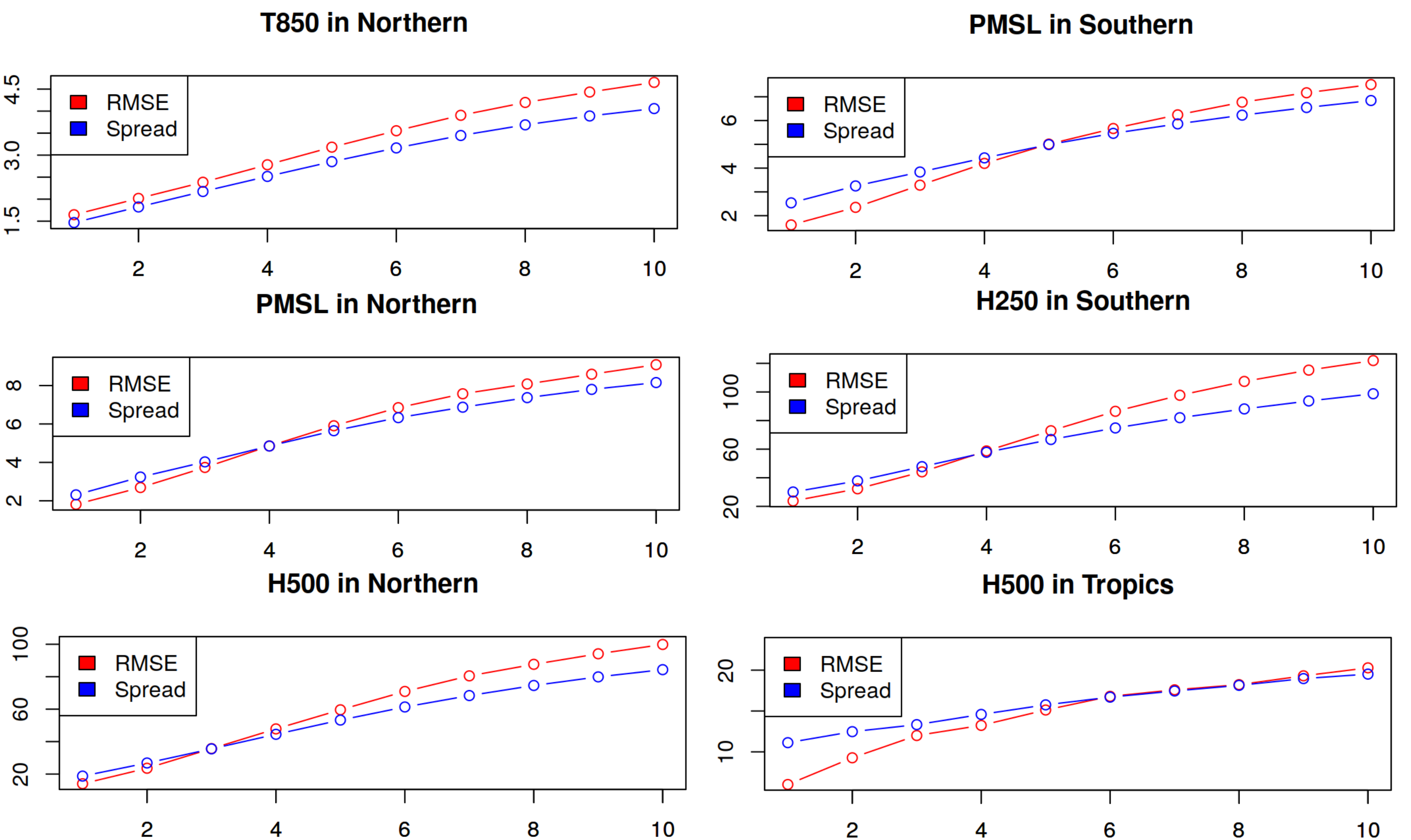
Practical implementation

- parallelization distributing the data across MPI-processors (by latitudes) + OpenMP-threads (by longitudes)
- Cray XC40-LC supercomputer installed at the Main Computer Center of Roshydromet
- 60 members of ensemble
- 992 cores is used:
  - 248 MPI x 4 OpenMP for the DA system
  - parallel launch of 31 models, using 4 MPI x 16 OpenMP
- 10-day forecasts at 00 UTC, background (6-hours forecasts) at 06, 12 и 18 UTC
- runtime of whole program complex is less then 1,5 hour for the 10 day forecast obtaining

Experiment results

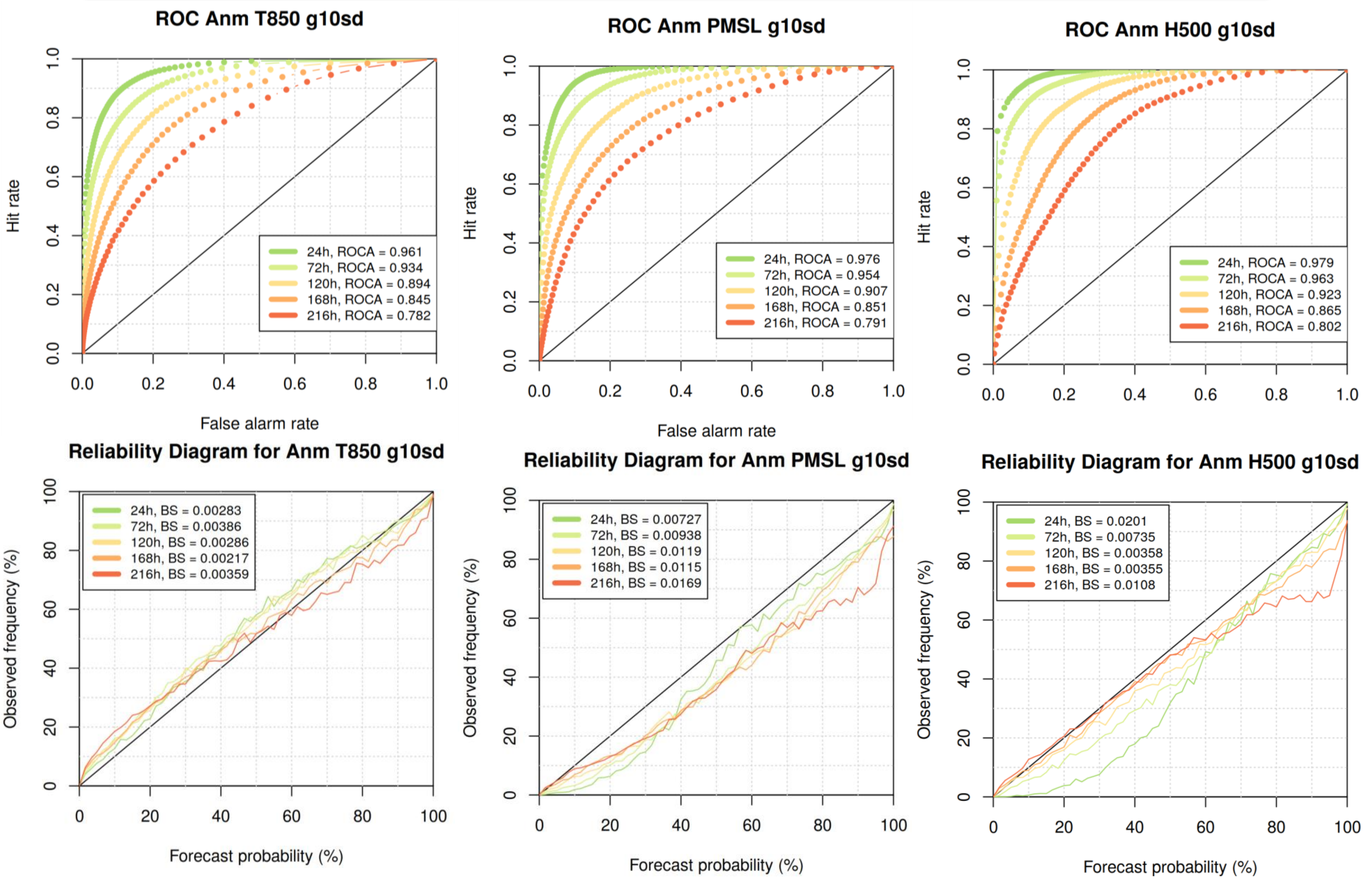
- Forecasts of January 2021 have been calculated
- Whole observation data set has been assimilated (AMV, SATOB, TEMP, SYNOP, SHIP, AIREP)
- Correlated errors of AMV observation has been used while assimilation
- Scores recommended by the Lead Centre on Verification of Ensemble Prediction System have been calculated and compared with results of the leading centers
- Ensemble spread has been compared with the RMSE of the control (ensemble mean) forecast - should approximately coincide
- CRPS (Continuous ranked probability score) has been compared with absolute error of control (mean) deterministic forecast

Ensemble mean vs spread of EPS forecast

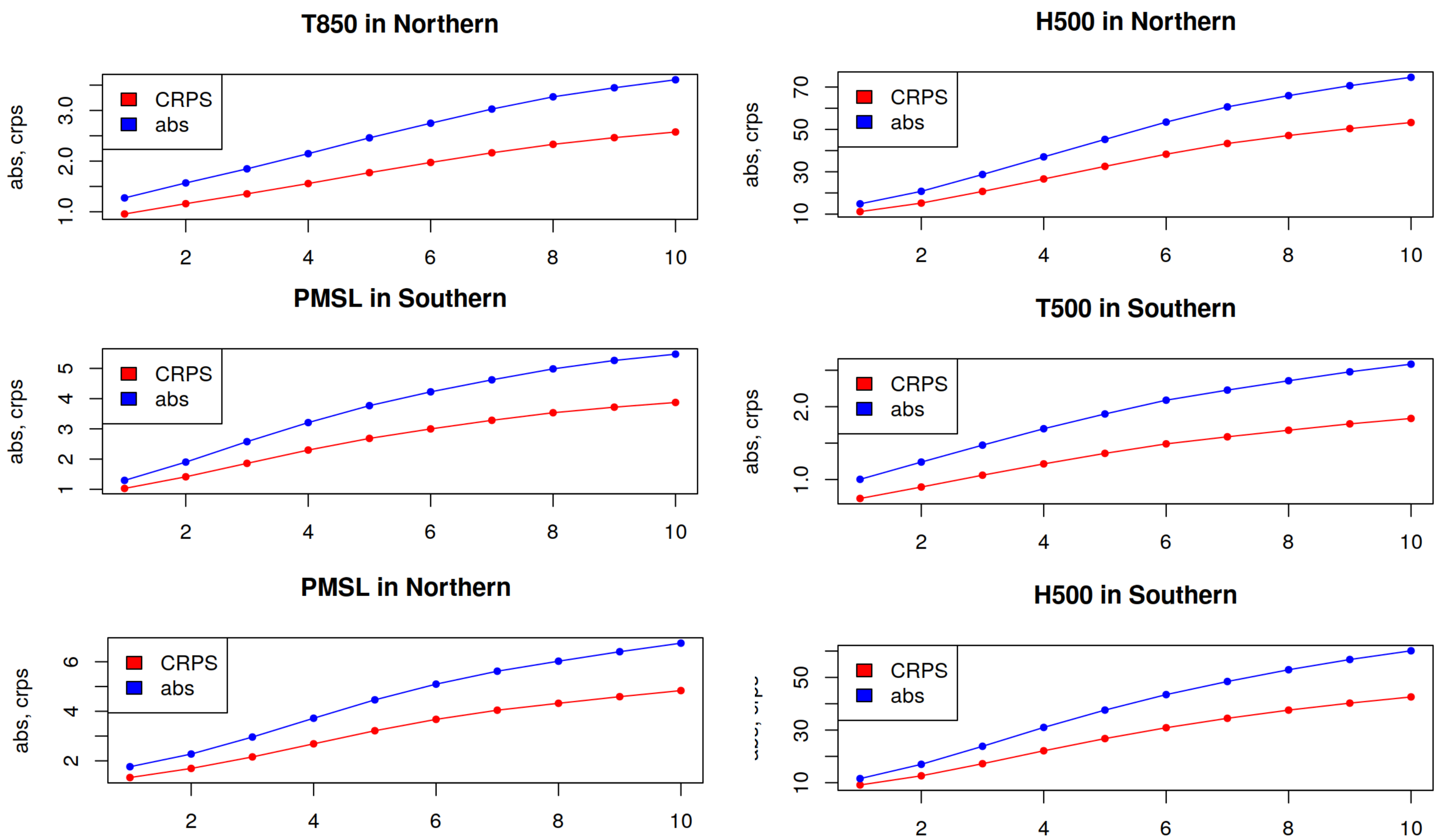


sprd – average standard deviation of ensemble averaged over region
rms -- root mean square error of mean forecast averaged over region
curves should be as close as possible

Reliability diagrams and Relative Operating Characteristic (ROC)



CRPS and absolute error of control (mean) deterministic forecast



CRPS – Continuous ranked probability score of ensemble forecast averaged over region
abs -- abs error of mean forecast averaged over region
curves should be as close as possible

Conclusions

- *current version already is capable of providing a satisfactory ensemble forecast*
- *improvement of data assimilation techniques could increase forecast accuracy*
- *using of correlated observation of AMVs slightly improves spread, but sufficiently increases time of analysis performing*
- *using of 60 ensemble members improves spread compared to 40 members*
- *we plan to investigate ensemble size influence on ensemble spread and other key features of ensemble*

Federal Service for Hydrometeorology
and Environmental Monitoring

HYDROMETEOROLOGICAL
CENTRE OF RUSSIA

About weather—at first hand

THANKS FOR YOUR ATTENTION!

Ensemble scores definitions

Brier score measures the mean squared probability error

$$BS = \frac{1}{N}\sum_{i=1}^{N}(f_i - o_i)^2$$

- N = number of points in the "domain" (spatio-temporal)
- o_i = 1 if the event occurs
  = 0 if the event does not occur
- f_i = is the probability of occurrence according to the forecast system (e.g. the fraction of ensemble members forecasting the event)

Continuous ranked probability score (CRPS) measures the difference between the forecast and observed CDFs

$$CRPS = \int_{-\infty}^{\infty}(P_{fcst}(x) - P_{obs}(x))^2 dx$$