Kaplunenko D.D., Ponomarev V.I., Krokhin V.V.

ASSESSING OF CLIMATE CHANGE TENDENCIES IN THE NORTHEAST ASIA AND NORTHWEST PACIFIC USING THE MULTIPLE IMPUTATION METHOD

V.I. Il'ichev Pacific Oceanological Institute, FEB RAS, Vladivostok
E-mail: dimkap@poi.dvo.ru

- The problem of assessing longtime regional climatic changes by the data of instrumental observations is complicated by the presence of missed data in the time series rows.
- Meteorological and oceanographic data in different WDC can not be in the same quality and scale for the same stations
- Data of climatic characteristics (SST, air temperature, precipitations) have a different statistical distribution. So it is necessary to find a set of methods or a method of assessing which will consider this fact.

- The goal of this work is assessing climatic tendencies on the Far East during 20^{th} century for SST, air temperature and precipitation using EOF-analysis and linear trend estimation.
- This aim would be provided by the applying of methods of Multiple Imputation, developed by Little, Rubin and Schafer for the data preparation (missed data augmenation)

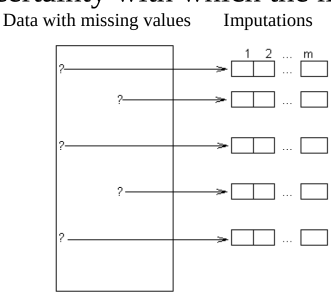
Multiple imputation method

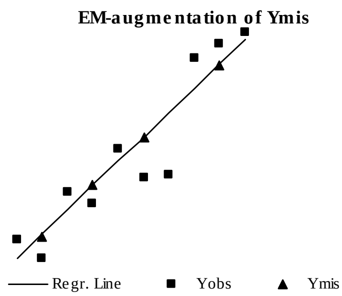
– Main point of method – augmentation of missed data values and assessing of statistical parameters by obtained samples using their distribution statistics (mean values, variances, covariances). Method is realized as successive applying of EM- and DA-algorithms which are based on assessing function of maximal likelihood.

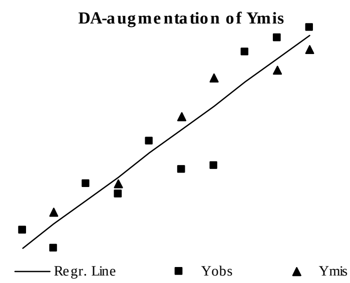
Little, R.J.A. & Rubin, D.B. (1987). ***Statistical Analysis with Missing Data*****. J. Wiley & Sons, New York.**

Schafer, J.L. (1997). ***Analysis of Incomplete Multivariate Data*****. Chapman & Hall, London.**

In multiple imputation (MI), each missing value is replaced by a set of m>1 plausible values drawn their predictive distribution. The variation among the m imputations reflects the uncertainty with which the missing values can be predicted from observed ones.







The **EM algorithm** is a general method for obtaining maximum-likelihood estimates of parameters from incomplete data. EM iterates between the following two steps:

· **E-step**: Replace missing sufficient statistics by their expected values given the observed data, using estimated values for the parameters; and

· **M-step**: Update the parameters by their maximum-likelihood estimates, given the sufficient statistics obtained from the E-step.

The convergence behavior of EM is related to the rates of missing information (how much information about the parameters is contained in the missing part of the data relative to the observed part). High rates of missing information can lead to slow convergence; low rates lead to rapid convergence.

Data augmentation (DA) is an iterative simulation technique, a special kind of Markov chain Monte Carlo (**MCMC**). In DA there are three types of quantities: observed data, missing data, and parameters. The missing data and parameters are unknown. DA alternately performs the following steps:

· **I-step**: Impute the missing data by drawing them from their conditional distribution given the observed data and assumed values for the parameters; and

· **P-step**: Simulate new values for the parameters by drawing them from a Bayesian posterior distribution given the observed data and the most recently imputed values for the missing data.

Alternating between these two steps sets up a Markov chain that converges to a stationary distribution, the joint distribution of the missing data and parameters given the observed data. DA bears a strong resemblance to the EM algorithm, and may be regarded as a stochastic version of EM. It is useful for multiple imputation of missing data. By running DA for a large number of cycles, and storing the results of a few I-steps along the way (with enough cycles in between to ensure independence), one obtains proper multiple imputations of the missing data.

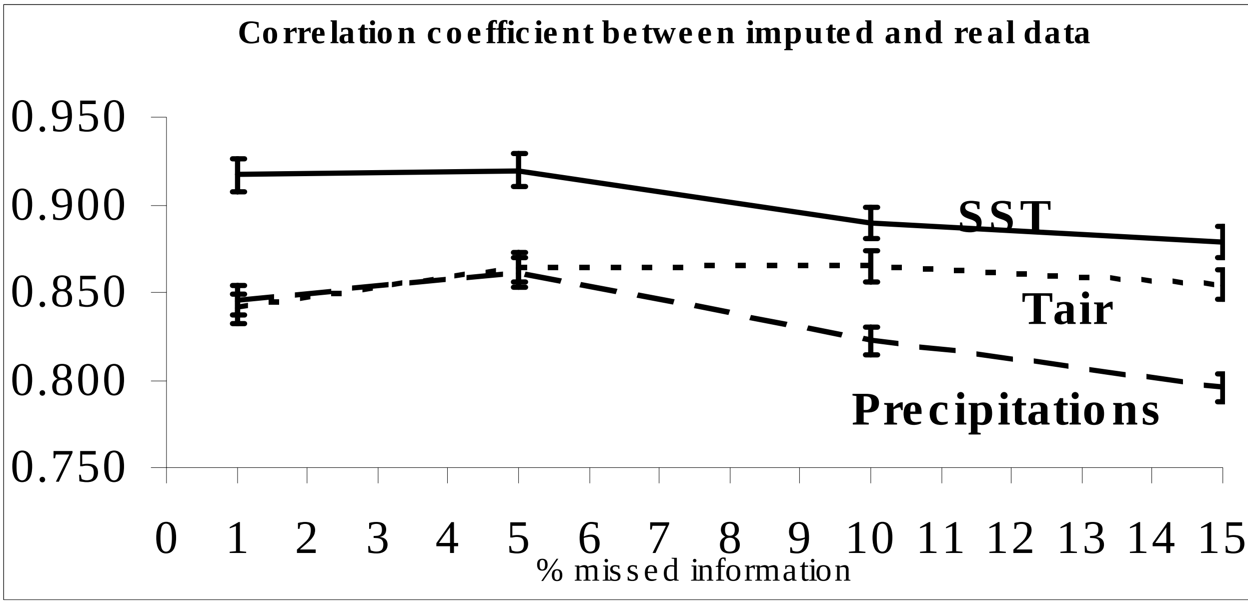
The method of Rubin (1987) for combining m sets of point estimates and standard errors has been extended to allow joint inferences about groups of parameters.

From each of the m imputed data sets, one must calculate and store estimates of the parameters in question along with an estimated covariance matrix.

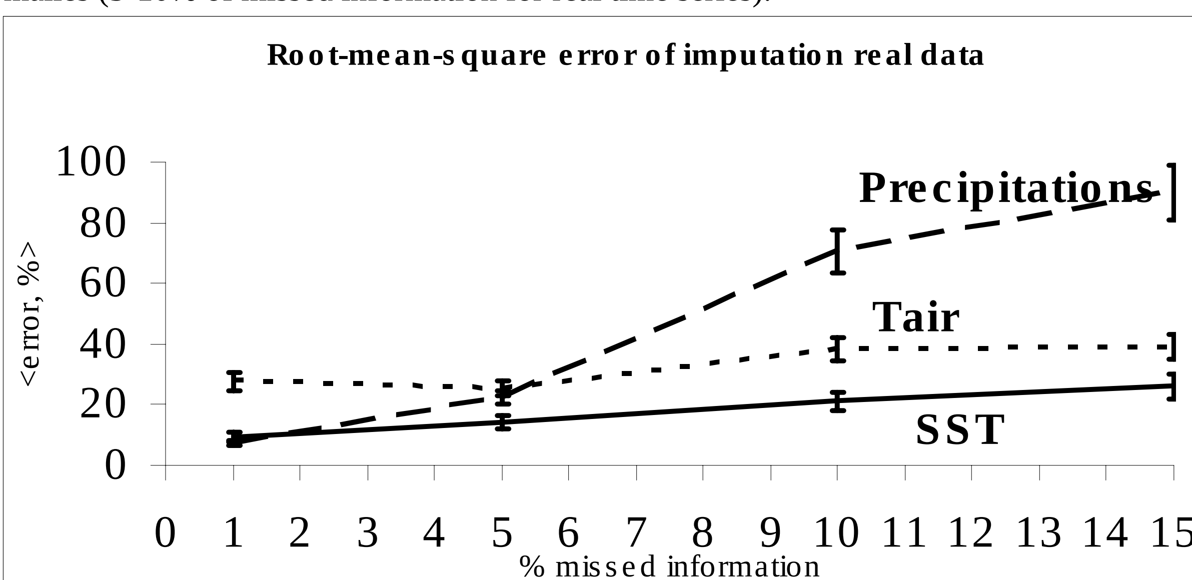
Data augmentation results

– Method approbation:

– empirical assessing of successfulness missed data imputation by artifical insertion of missed data in the time series and further assessing of root mean square error of data augmentation

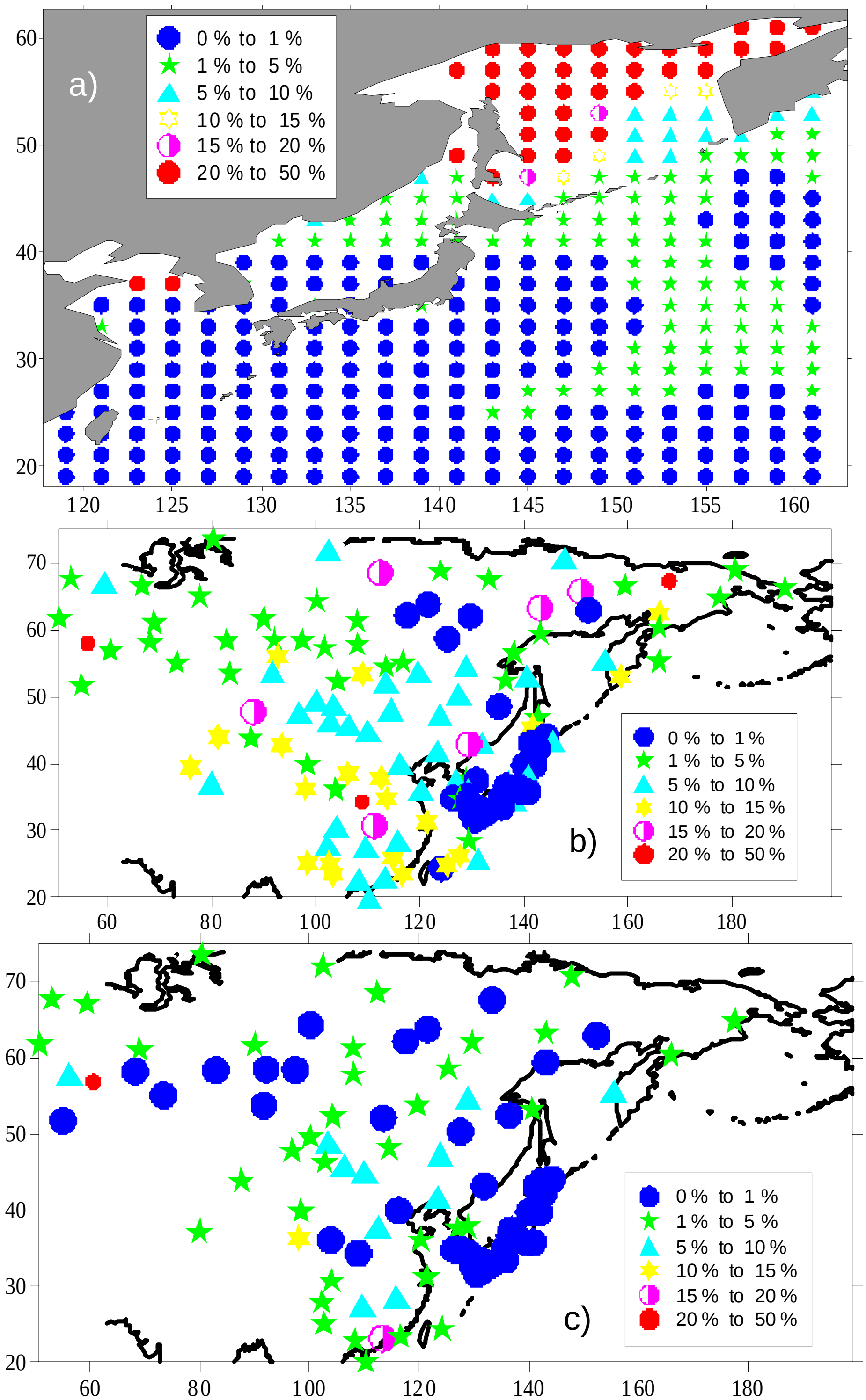


– The quality of augmentation is good enough for our purposes due to the small increase of mean square deviation and high value of correlation coefficient between real and imputed data anomalies (5-10% of missed information for real time series).



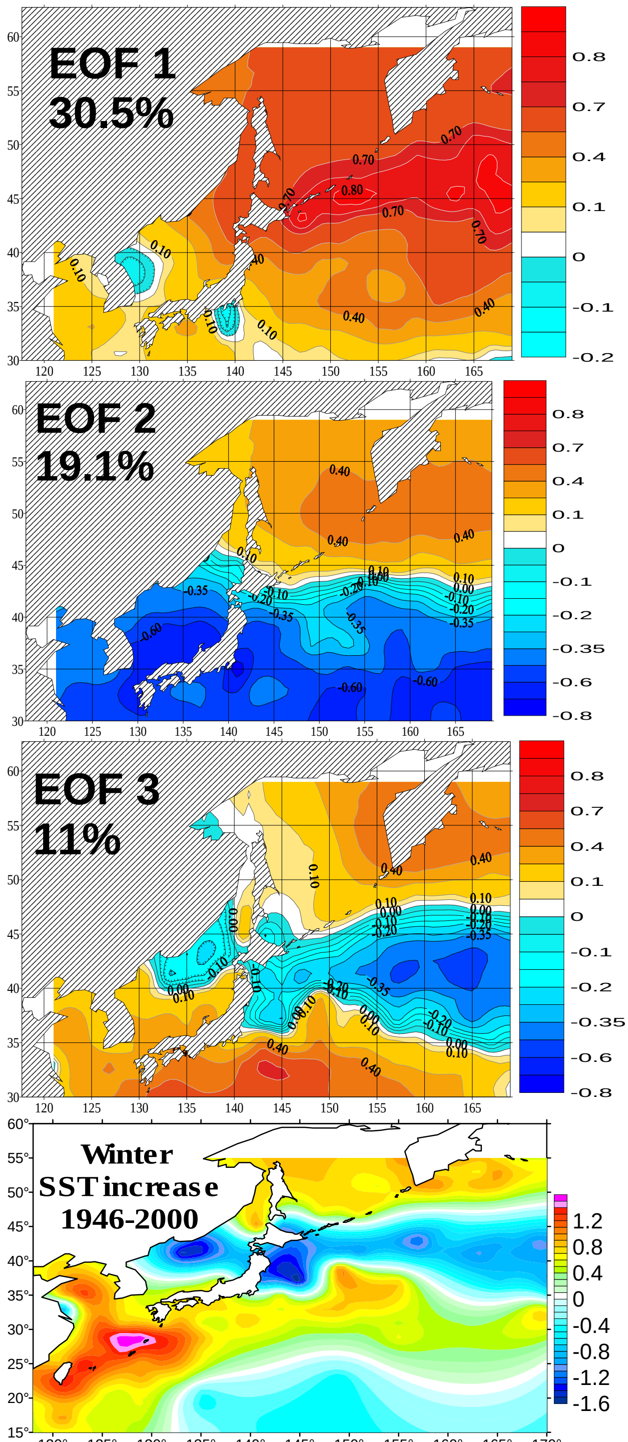
– Estimated parameters for augmented and real data is agreed with each other.

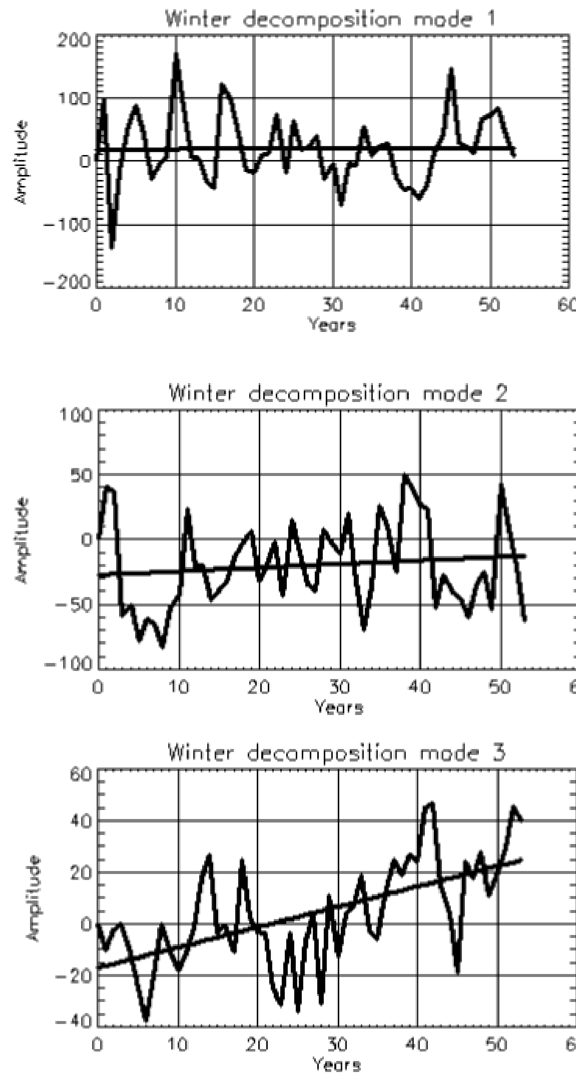
Northwest Pacific and Northeast Asia: missed data distribution.



Distribution of missed data for Northwest Pacific SST a), Northeast Asia air temperature b) and precipitations c).

SST: EOF-decomposition for winter period and linear trends estimation

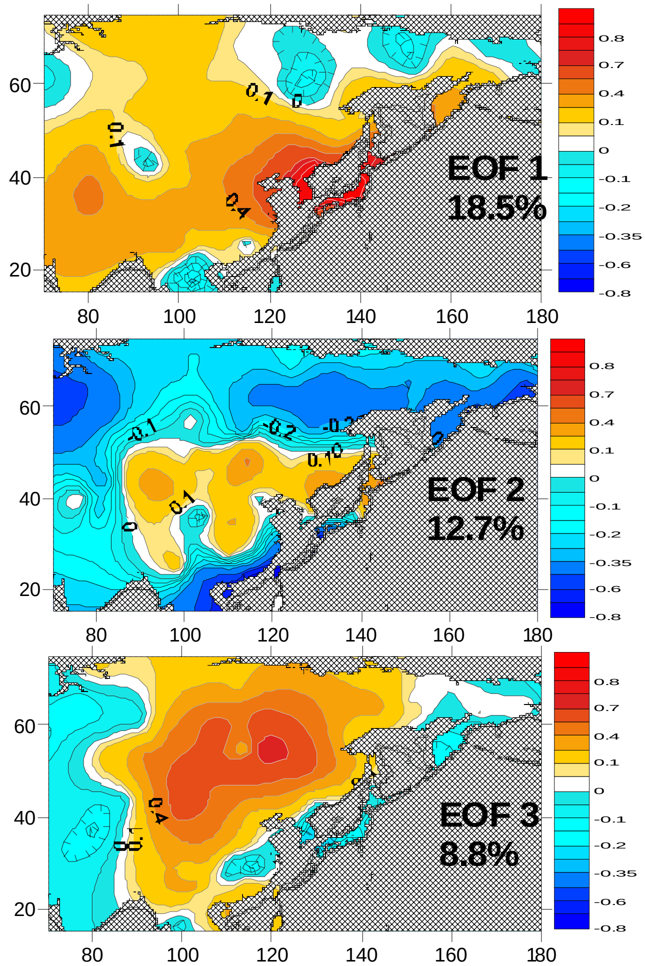


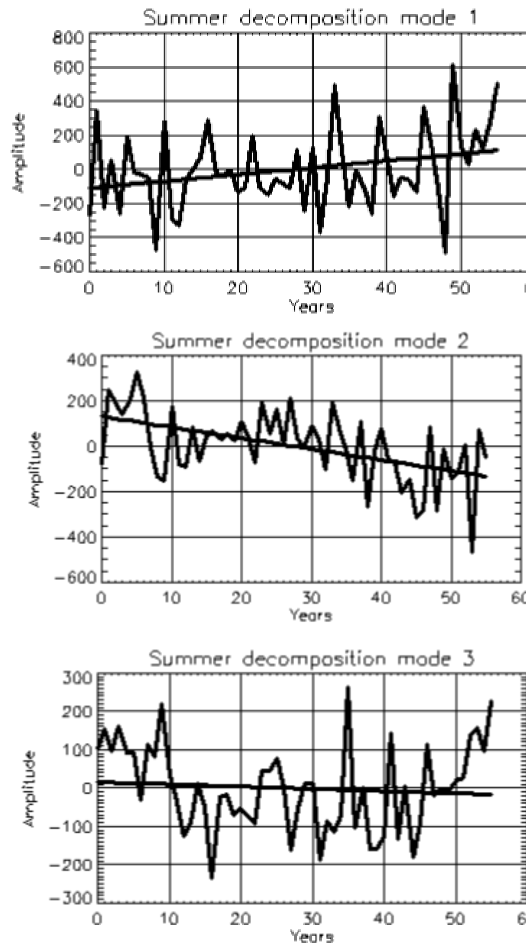


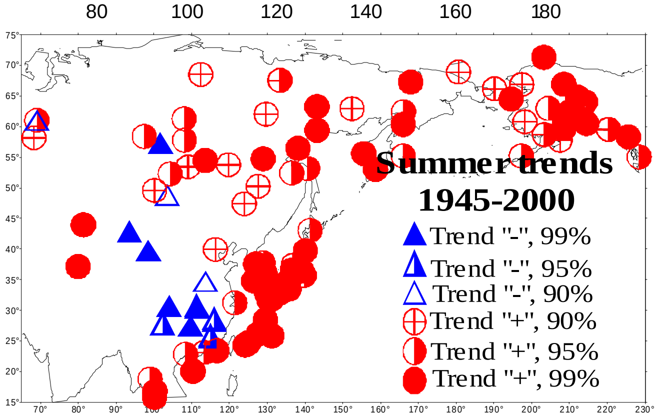
- Negative sign area for mode 3 EOF decomposition (11% of variance) has the same location with negative tendencies found for this region (Ponomarev et al, 2003)
- this might be possible evidence of strengthening processes in the subarctic frontal zone during considerated period .

– Statistically significant positive SST trend in Kuroshio region and in the northwest area of the Pacific subarctic gyre dominates in cold season of year
– Negative SST trend occurs in the Oyashio region and occupies south western area of subarctic Pacific gyre.

Air temperature: EOF-decomposition for winter period and linear trends estimation





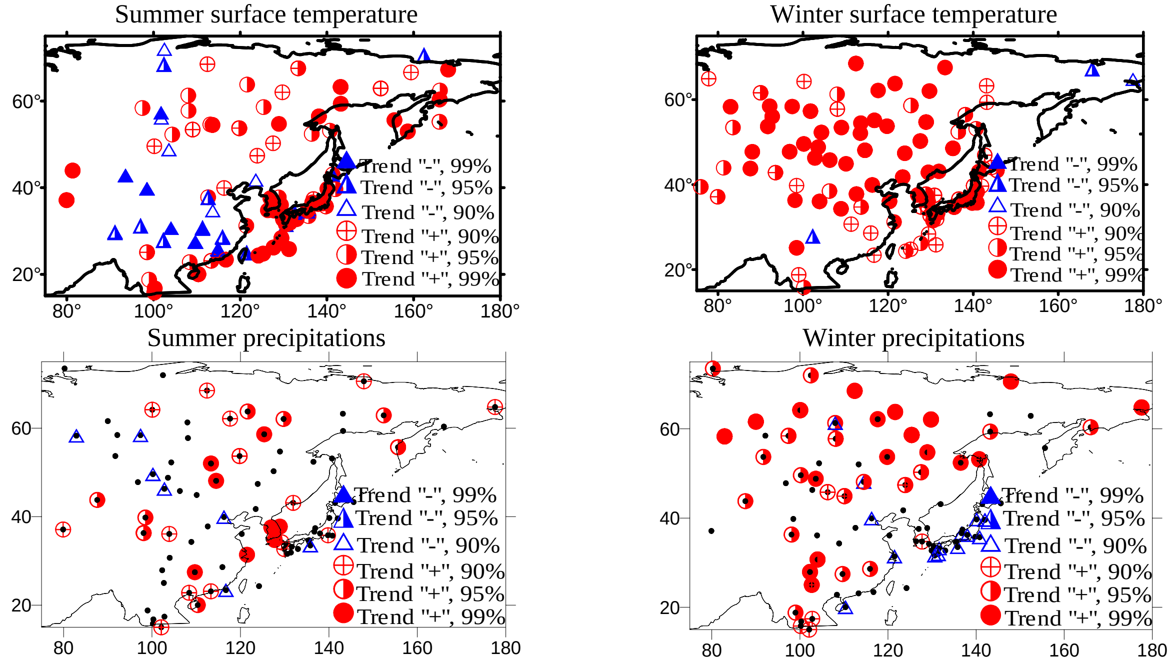


- **Negative sign area for mode 2 EOF decomposition (12.7% of variance) has the same location with negative tendencies found for this region (Ponomarev et al, 2001)**
- **This facts indicates nonuniform warming for considered regions**

Also have been found that warming tendency in winter accompanies precipitation decrease in Japan and Russian Primorye Region adjacent to the Northwest East/Japan Sea. Significant precipitation reduce in Japan takes place in October, December and January, with exception of subarctic area (Hokkaido Island) where precipitation slightly increases in December-March and in August but decreases in May-July.

06a

Air temperature and precipitations : linear trends estimation



The warming tendency in air temperature accompanies precipitation increase in this are of moderate latitudes. The tendency sign depends on the season.

CONCLUSIONS

- Multiple Imputation method is valid for augmentation data of SST, air temperature and precipitation values.
- MI is good enough for assessing of climate change tendencies in the Northeast Asia and Northwest Pacific by mentioned characteristics and might be a perspective for common statistical analysis of different characteristics with missed values with using of Rubin rules of calculation
- Results of climate change tendencies analysis showed the presence of significant positive SST trend in Kuroshio region and negative trend for Oyashio region. The warming tendency in air temperature accompanies precipitation increase in this area of moderate latitudes of Northeast Asia.
- EOF-analysis of SST and Air temperature showed the presence of relation between observed modes variability and significant trends locations in the mentioned region.