Seasonal climate prediction for North Eurasia

Vladimir N Kryjov

North Eurasia Climate Center, Moscow, Russia
E-mail: kryjov@yahoo.co.uk

Received 18 October 2011
Accepted for publication 1 February 2012
Published 2 March 2012
Online at stacks.iop.org/ERL/7/015203

Abstract
An overview of the current status of the operational seasonal climate prediction for North Eurasia is presented. It is shown that the performance of existing climate models is rather poor in seasonal prediction for North Eurasia. Multi-model ensemble forecasts are more reliable than single-model ones; however, for North Eurasia they tend to be close to climatological ones. Application of downscaling methods may improve predictions for some locations (or regions). However, general improvement of the reliability of seasonal forecasts for North Eurasia requires improvement of the climate prediction models.

Keywords: operational seasonal forecast, North Eurasia, climate models, multi-model ensemble, downscaling

1. Introduction
The socio-economic development of the countries located in North Eurasia is strongly climate and weather dependent. The importance of seasonal climate prediction has been realized by the world meteorological community. The World Meteorological Organization (WMO) initiated the establishing of the Lead Center for Long Range Forecast Multi-Model Ensembles (WMO LC LRF-MME) in Seoul, Korea, and the Lead Center for the Long Range Forecast Verification System (WMO LC SVS-LRF) in Melbourne, Australia. Twelve operational centers issuing global seasonal model predictions have achieved the status of WMO Global Producing Centers (GPC). In March 2012, WMO plans to start issuing a global seasonal climate update: a climate monitoring and seasonal climate forecasting bulletin with global coverage. The forecasting section of the bulletin will be based on the results produced by the Global Producing Centers and both WMO Lead Centers. This activity essentially improves the provision of the world’s societies and economies with seasonal forecasts.

A brief overview of the current status of seasonal forecasting for North Eurasia and possible ways of improving it is given in this work.

2. Existing capabilities
2.1. Single-model forecasts
In order to improve the regional interpretation of the global long range forecasts produced by the Global Producing Centers, WMO established a set of WMO Regional Climate Centers, with the North Eurasia Climate Center (NEACC) being in charge of the operational long range forecasts for North Eurasia. NEACC issues a set of long range forecasts based on the outputs from the single model of the Hydrometeorological Research Center of Russia (Tolstykh 2010). The one-month-lead seasonal climate prediction is based on the outputs from the single model of the Hydrometeorological Research Center of Russia (Tolstykh 2010). It should be noted that according to the WMO regulations (WMO 2002; see their figure 1), a one-month-lead seasonal forecast is a forecast of the seasonal mean values issued one month in advance of the beginning of the forecast period. For example, the seasonal forecast for wintertime (December through February, DJF) temperature implies the issuing of the forecast of the seasonal (DJF) mean temperature values by the beginning of November, with model integrations being initialized using October observations. In
Figure 1. RPSS of the seasonal forecasts of wintertime (upper panel) and summer (lower panel) North Eurasia station temperature interpolated from the GPC-Moscow model grid-point forecasts. RPSS values exceeding approximately 0.08 (orange and red points) are significant at the 2.5% significance level in one-tailed tests as revealed from 1000 Monte Carlo runs. The verification was performed in cross-validation mode with five years withheld.

particular, NEACC issues the forecast of wintertime seasonal mean temperature values in late October. The operational forecasting system consists of a ten-member ensemble produced by the global atmospheric model with a resolution of 1.1° latitude by 1.4° longitude and 28 vertical levels. No ocean model is used, with sea surface temperature (SST) being specified by the persistence of initial observed SST anomalies for the whole forecast period atop the SST annual cycle. Hindcasts cover the 25 yr period from 1979–2004. The Hydrometeorological Research Center of Russia is one of the WMO Global Producing Centers, GPC-Moscow, and model outputs match all the WMO requirements listed in the Manual on the Global Data-Processing and Forecasting System (WMO 2010).

Although being successful in the prediction of the extremely low Arctic Oscillation index and the extremely cold North Eurasia winter of 2009/10, in general, NEACC's seasonal climate predictions, similarly to those from all the seasonal climate prediction models, are quite poor in the middle and high latitudes. Figure 1 shows the rank probability skill score (RPSS; Wilks 1995) of the seasonal forecasts of tercile probabilities interpolated from the GPC-Moscow model outputs as regards climatological forecasts for 70 weather stations more or less uniformly distributed throughout North Eurasia. These 70 stations comprise a standard set of the stations used at the Hydrometeorological Research Center of Russia and NEACC for verification of the long range forecasts. The postprocessing of the model forecasts includes correction of the model bias in accordance with the WMO requirements (WMO 2010), that is, the systematic errors in model climatology are corrected on the basis of the climatology of the historical forecasts and that of the observations.

The verification assessments shown in figure 1 have been performed on the cross-validated forecast series with five years withheld. WMO recommends (WMO 2002) that any influence of the data from the years which follow the year of forecast should be prevented in a cross-validation procedure. An analysis of the autocorrelation structure of the station temperature series has revealed that for almost all of them an autocorrelation function becomes negative at the time shift of one–two years. So, the verification assessments performed on the cross-validated series, with the data for five years (the year of forecast and the following four years) being withheld, match the WMO recommendations.

The RPSS values exceeding approximately 0.08 (orange and red points) are statistically significant at the 2.5% significance level in the one-tailed test as revealed from 1000 Monte Carlo runs. The threshold value has been exceeded for several stations. Since the total number of the stations is 70, there is a probability that these significant RPSS values have been obtained occasionally (Wilks 1995, chapter 5.4) and field significance tests accounting for spatial dependences between the stations have been conducted. Indeed, the probability of
obtaining of such numbers (11 for winter and 8 for summer) of 2.5% significant RPSS values occasionally exceeds 5% and 8% for winter and summer, respectively, as revealed from 1000 Monte Carlo tests, with spatial dependence between the stations being accounted for by scrambling in the time domain of the whole yearly sets of station forecasts. Unfortunately, as was mentioned above, a poor performance of seasonal climate predictions for the middle and high latitudes is a common shortcoming of all the seasonal climate prediction models, even the state-of-the-art ones, both two-tier and one-tier (Kharin and Zwiers 2001, 2003, Yun et al 2003, Kryjov et al 2006, Bundel et al 2011). The conclusion of the poor performance, particularly for North Eurasia, is supported by the verification assessments of the seasonal forecasts from the twelve WMO Global Producing Centers which are posted on the Web-site of the WMO LC SVS-LRF (www.bom.gov.au/wmo/lrfvs/index.html).

2.2. Multi-model forecasts

The tendency in the past decade has been to use multi-model ensembles in seasonal climate prediction. Nowadays, a number of (mainly international) climate centers issue multi-model seasonal climate predictions. For North Eurasia, multi-model ensemble predictions issued by Asia Pacific Economic Cooperation Climate Center (APCC, Busan, Korea), the International Research Institute for Climate and Society (IRI, Palisades, New York) and WMO LC LRF-MME are available. As has been shown by numerous studies (Kharin and Zwiers 2002, Peng et al 2002, Barnston et al 2003, Hagedorn et al 2005, Min et al 2009), a multi-model ensemble forecast which combines several single-model forecasts outperforms each of these single-model forecasts taken separately. The rationality of the multi-model ensemble has been discussed in a number of studies (Vislocky and Fritsch 1995, Fritsch et al 2000, Hagedorn et al 2005, Weigel et al 2008, Min et al 2009). In a multi-model ensemble forecast, the errors of the (bias corrected) individual model forecasts mutually offset each other and increase the signal to error ratio. This can be illustrated with a decomposition of an individual forecast (a single-model ensemble member forecast) into the signal and error components (Kryjov 2010):

\[ X = y + \text{err}_c + \text{err}_m + \text{err}_r. \]

In this formula X is a single-model ensemble member forecast; y is a signal—a fraction of the forecast positively correlated with observations Y; \( \text{err}_c \) is a (constant) bias of the model climatology with respect to observed climatology caused by the model formulation; \( \text{err}_m \) is a varying model systematic error caused by the model formulation; \( \text{err}_r \) is a ‘weather noise’—normally distributed random error. \( \text{err}_c \) is considered constant and it is easily eliminated by adjustment of the model climatology to the observed one on the basis of historical forecasts. The last term, the random error, is reduced by the use of an ensemble of model integrations started from different initial conditions, with this random error converging to zero with increase of the ensemble size. However, as is obvious from the equation, no single-model ensemble size increase leads to a reduction of \( \text{err}_m \). The purpose of the use of a multi-model ensemble is just the reduction of this error. This is based on the hypothesis that these errors from individual models are independent of each other. Under these conditions, since signal \( y \) is common to all the models, the averaging of the model forecasts leads to an increase of the signal to error ratio because the errors from individual models tend to mutually offset each other.

Indeed, in the regions for which predictability is high, e.g., the equatorial Pacific, such offset of the model errors leads to extraction of a useful signal common to all the participating models and improvement of the seasonal forecast. Meanwhile, for the regions for which atmospheric predictability is low and the errors are larger than the useful signal, mutual offset of the errors results, on the one hand, in increasing of the reliability of forecasts, but, on the other hand, in the tendency to lead to forecasts close to climatology ones (Min et al 2009). In particular, on the basis of seven-model ensemble historical forecasts for 21 yr, Min et al (2009) have shown that about 70% of the multi-model forecasts of the tercile probabilities of temperature for the northern extratropics are within the range of 0.20–0.40, which is close to the climatological tercile probability value of 0.33. In comparison, the fraction of such forecasts for the Tropics, for which predictability of temperature is high, does not exceed 40%. North Eurasia is an area of low predictability (see, e.g., Robertson et al 2004, Min et al 2009, Bundel et al 2011), with the fraction of the forecasts close to climatological ones being about 80%. In the global multi-model ensemble seasonal forecasts issued by the climate centers which operate large multi-model ensembles, APCC (www.apcc21.org), IRI (portal.iri.columbia.edu) and WMO LC LRF-MME (wмолс.org), forecasts for North Eurasia tend, as a rule, to be close to climatology ones, that is, uncertain. The multi-model ensemble approach in seasonal forecasting is a great step forward as compared with a single-model one. For the regions with appropriate predictability, the use of multi-model ensemble forecasts provides a more reliable basis for decision-making than single-model forecasts and the advantage of the multi-model approach is entirely obvious for these regions. Meanwhile, for the regions with low predictability, particularly for North Eurasia, the multi-model ensemble forecasts prevent us from making wrong decisions, which is a definite advantage; however, support for correct decisions based on them is quite limited.

3. Future development

A reasonable way to improve seasonal climate prediction for North Eurasia is to apply a downscaling. Developed for quantification of the local effects of global change processes (von Storch 1995, Zorita and von Storch 1999), nowadays statistical downscaling is becoming a standard and routine technique in long range forecasting for regional/local scales (see, e.g., Feddersen and Andersen 2005, Kang et al 2009, Juneng et al 2010, Min et al 2011).

The downscaling approach is based on the view that the regional climate is conditioned by two factors: the large
scale climate state and regional (local) physiographic features. However, the use of the downscaling technique in seasonal climate prediction differs from that in climate projection for which it was initially developed. Whereas in climate projections the relationships are established between the observed local variables and large scale fields derived from observations (e.g., by reanalysis), in seasonal predictions the relationships are established between the observed local variables and model simulated forecast fields in which the large scale atmospheric patterns are usually spatially shifted as compared with observations. Therefore, downscaling in seasonal prediction yields model forecast correction rather than localization.

Correcting the raw model forecasts with the downscaling technique, it is possible to improve to some degree the forecasts and forecast performance. In particular, this is shown by the results from the tests with the downscaling method suggested by Min et al. (2011) modified for a single-model version. This method is based on regression with a probabilistic interpretation of the forecasts based on the estimation of forecast uncertainty associated with both regression random errors and the ensemble spread. Another peculiarity of the method is the selection of the model predictor (model grid-point series), the predictand being the station temperature, with the use of the cross-validation cycle, which allows one to improve the stability of the selection of predictors. This ‘inner’ cross-validation cycle, which is used for the selection of the model predictors, should not be confused with an ordinary verification cross-validation cycle, which in this case is an ‘outer’ one. Forecasts are formulated in terms of tercile probabilities based on the Gaussian probability distribution function (PDF). The parameters of the forecast PDF are estimated as an arithmetic mean of the mean values and standard deviations of the Gaussian PDFs predicted from the simple linear regressions (with estimated uncertainty) of the station temperature on the selected model grid-point series which feature a congruence coefficient significant at the 5% level (depending on the effective number of degrees of freedom, the grid-point threshold coefficient value slightly varies around 0.45). It should be noted that the use of the arithmetic mean of the standard deviations implies that the grid-point series are not considered independent. The details of the downscaling method applied are described by Min et al. (2011).

The verification assessments of the downscaled forecasts have been performed on the cross-validated forecast series (the ‘outer’ cross-validation) with five years withheld, that is, similarly to the verification assessments of the interpolated forecasts shown in figure 1.

Results from the tests are shown in figure 2. The basic model forecast is the same as for figure 1—it is the single-model GPC-Moscow global forecast—and the assessment procedures are absolutely the same so this figure can be compared with figure 1. The number of stations for
which RPSS exceeds 0.08 is 29 for winter and 21 for summer. The probability of obtaining such figures occasionally is less than 0.1%.

In general, the number of stations (out of 70) for which downscaled forecasts outperform climatological ones is 44 for winter and 43 for summer forecasts. For interpolated forecasts those figures are 28 and 27, respectively. So, downsampling improves to some degree the performance of the seasonal predictions. However, it cannot be assessed as reliable yet. Even improvement by downsampling based on the multi-model ensemble (Min et al 2011) does not achieve very much.

4. Conclusion

Reliable seasonal climate prediction for middle and high latitudes, particularly for North Eurasia, is a great challenge for modern climate science. The work goes in two directions: improvement of the climate models and improvement of the methods of postprocessing the model output, with the former being the basic one. Both simple and sophisticated postprocessing methods are aimed at the correction of the varying systematic model errors. It is required that the model outputs at least contain the useful signal and that this signal is sufficiently large to be extracted from the raw model forecast. Nowadays, the seasonal climate prediction models are governed by the boundary conditions, mainly confined to sea surface temperature in the eastern equatorial Pacific (Shukla et al 2000). The area of reliable seasonal prediction is also mainly confined to the eastern equatorial Pacific. However, many empirical studies (see, e.g., Mosedale et al 2006, Sinha and Toppel 2006, Pan 2007) reveal that some seasonal predictability may arise from the heat and mass exchange between the ocean and atmosphere in other regions, including the extratropics. In addition, inclusion of the highly resolved stratosphere may also improve the long range simulation of the atmospheric processes in the extratropics (see, e.g., Baldwin and Dunkerton 2001, Perlwitz and Harnik 2004, Kryjov and Park 2007). General improvement of the reliability of seasonal forecasts for the Earth regions with low atmospheric predictability requires improvement of the climate prediction models. Meanwhile, reliable seasonal climate prediction for North Eurasia remains a great challenge.

Acknowledgments

Development of the downscaling technique was funded by Asia Pacific Economic Cooperation Climate Center (Busan, Korea) and the Central Research Institute for Electric Power Industry (Tsukuba, Japan).

References

Baldwin M P and Dunkerton T J 2001 Stratospheric harbingers of anomalous weather regimes Science 294 581–4
Barnston A G, Mason S, Goddard L, DeWitt D G and Zebiak S E 2003 Increased automation and use of multimodel ensembles in seasonal climate forecasting at the IRI Bull. Am. Meteorol. Soc. 84 1783–96
Bundel A Yu, Kryzhow V N, Min Y-M, Khan V M, Vifand R M and Tishchenko V A 2011 Assessment of probability multimodel seasonal forecast based on the APCC model data Russ. Meteorol. Hydrol. 36 135–54
Feddersen H and Andersen U 2005 A method for statistical downscaling of seasonal ensemble predictions Tellus A 57 398–408
Fritsch J, Hilliker M, Ross J and Vislocky R L 2000 Model consensus Weather Forecast. 15 571–82
Hagedorn R, Doblas-Reyes F J and Palmer T N 2005 The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept Tellus A 57 219–33
Juneng L, Tangang F T, Kang H, Lee W-J and Seng Y K 2010 Statistical downsampling forecasts for winter monsoon precipitation in Malaysia using multimodel output variables J. Clim. 23 17–27
Kang H, Park C-K, Saji N H and Ashok K 2009 Statistical downscaling of precipitation in Korea using multimodel output variables as predictors Mon. Weather Rev. 137 1928–38
Kharin V V and Zwiers F W 2001 Skill as function of time scale in ensemble of seasonal hindcast Clim. Dyn. 17 127–41
Kharin V V and Zwiers F W 2002 Climate predictions with multi-model ensemble J. Clim. 15 793–9
Kharin V V and Zwiers F W 2003 Improved seasonal probability forecast J. Clim. 16 1684–701
Kryjov V N 2010 Postprocessing of the model long range forecasts at the North Eurasia Climate Centre 80 Years of Hydrometeorcentre of Russia (Moscow: Hydrometcentre of Russia) pp 247–53
Kryjov V N and Park C-K 2007 Solar modulation of the El-Niño/Southern oscillation impact on the Northern Hemisphere annular mode Geophys. Res. Lett. 34 L10701
Kryjov V N et al 2006 Assessment of the climate forecasts produced by individual models and MME methods APCC Technical Report 2006, No.1 vol 1, p 534
Min Y-M, Kryjov V N and Oh J-H 2011 Probabilistic interpretation of regression-based downscaled seasonal ensemble predictions with the estimation of uncertainty J. Geophys. Res. 116 D08101
Min Y-M, Kryjov V N and Park C-K 2009 A probabilistic multimodel ensemble approach to seasonal prediction Weather Forecast. 24 812–28
Mirvis V M, Meleshko V P, Gavrilina V M, Matyugin V A and L’vyova T Yu 2006a Monthly meteorological forecasting with the MGO hydrodynamic-statistical method. I. Deterministic forecast; results from 26-year series of testing Russ. Meteorol. Hydrol. 2006 1–10
Mirvis V M, Meleshko V P, Gavrilina V M, Matyugin V A and L’vyova T Yu 2006b Monthly meteorological forecasting with the MGO hydrodynamic-statistical method. II. Probabilistic forecast: analysis and interpretation of ensemble distribution, forecast calculation and verification Russ. Meteorol. Hydrol. 2006 1–10
Mosedale T J, Stephenson D B, Collins M and Mills T C 2006 Granger causality of coupled climate processes: ocean feedback on the North Atlantic oscillation J. Clim. 19 1192–4
Pan L L 2007 Synoptic eddy feedback and air–sea interaction in the North Atlantic region J. Clim. 20 647–59
Peng P, Kumar A and van den Dool H 2002 Analysis of multimodel ensemble predictions for seasonal climate anomalies J. Geophys. Res. 107 4710–21
Perlwitz J and Harnik N 2004 Downward coupling between the stratosphere and troposphere: the relative roles of wave and zonal mean processes J. Clim. 17 4902–9
Robertson A W, Lal U, Zebiak S E and Goddard L 2004 Improved combination of multiple atmospheric GCM ensembles for seasonal prediction Mon. Weather Rev. 132 2723–44
Shukla J et al 2000 Dynamical seasonal prediction Bull. Am. Meteorol. Soc. 81 2593–606
Sinha B and Topliss B 2006 A description of interdecadal time-scale propagating North Atlantic sea surface temperature anomalies and their effect on winter European climate, 1948–2002 J. Clim. 19 1067–79
Tolstykh M A 2010 Global Semilagrangian Model For Numeric Weather Prediction (Moscow: Hydrometcentre of Russia)
Vislocky R L. and Fritsch J M 1995 Improved model output statistics forecast through model consensus Bull. Am. Meteorol. Soc. 76 1157–64
von Storch H 1995 Inconsistencies at the interface of climate impact studies and global climate research Meteorol. Z. 4 72–80
Weigel A P, Liniger M A and Appenzeller C 2008 Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? Q. J. R. Meteorol. Soc. 134 241–60
Wilks D S 1995 Statistical Methods in the Atmospheric Sciences (New York: Academic) p 467
WMO 2002 Standardised Verification System (SVS) for Long-Range Forecasts (LRF), New Attachment II-9 to The Manual on The GDPS (WMO-No. 485 24) vol 1 (Geneva: WMO)
WMO 2010 Manual on the Global Data-Processing and Forecasting System. Volume I—Global Aspects (WMO-No. 485) (Geneva: WMO)
Yun W-T, Stefanova L and Krishnamurti T N 2003 Improvement of the multimodel superensemble technique for seasonal forecast J. Clim. 16 3834–40
Zorita E and von Storch H 1999 The analogue method as a simple statistical downscaling technique: comparison with more complicated methods J. Clim. 12 2474–89