**PAPER • OPEN ACCESS**

# Mesoscale atmospheric modelling technology as a tool for creating a long-term meteorological dataset

To cite this article: Vladimir Platonov *et al* 2017 *IOP Conf. Ser.: Earth Environ. Sci.* **96** 012004

View the article online for updates and enhancements.

## Related content

- Mesoscale modelling of the summer climate response of Moscow metropolitan area to urban expansion
M I Varentsov, P I Konstantinov and T E Samsonov

- Mesoscale high-resolution modeling of extreme wind speeds over western water areas of the Russian Arctic
Vladimir S Platonov and Alexander V Kislov

- Viability of Controlling Prosthetic Hand Utilizing Electroencephalograph (EEG) Dataset Signal
Azizi Miskon, Suresh A/L Thanakodi, Mohd Raihan Mazlan et al.

# Mesoscale atmospheric modelling technology as a tool for creating a long-term meteorological dataset

**Vladimir Platonov, Alexander Kislov, Gdaly Rivin, Mikhail Varentsov, Inna Rozinkina, Mikhail Nikitin and Mikhail Chumakov**

Lomonosov Moscow State University, Department of Geography, Section of Meteorology and Climatology. 119991, Moscow, Leninskie Gory MGU, 1, GSP-1, phone 8(495)939-29-42
Hydrometcentre of Russia, Moscow, 123242, B. Predtechensky per., 11-13, phone 8(495)255-13-09
A.M.Obukhov Institute of Atmospheric Physics, Russian Academy of Sciences. 119017, Moscow, Russian Federation, Pyzhyovskiy pereulok, 3, phone 8 (495) 951-85-49.
Gazprom VNIIGAZ LLC, Moscow region, 115583, box 130, Razvilka village, 15-1.

vplatonov86@gmail.com, avkislov@mail.ru

**Abstract.** A detailed hydrodynamic simulation of major meteorological parameters for the last 30 years (1985 – 2014) has been performed for the Sea of Okhotsk and the Sakhalin Island. The regional non-hydrostatic atmospheric model COSMO-CLM was used for this long-term simulation with horizontal resolutions of ~13.2, ~6.6 and ~2.2 km. This dataset was created to help in the investigation of statistical characteristics and physical mechanisms of formation of extreme weather events (primarily wind speed extremes) on small spatio-temporal scales. The detailed meteorological information thus obtained could be used to take into account the coast configuration, mountain systems, and other important mesoscale features of the terrain. This paper describes a proposed downscaling technology for long-term simulations with three "nested domains". The results of verification of the dataset and estimation of extreme wind velocities are presented.

## 1. Introduction

The hydrometeorological information about the Arctic and Far East regions is presented, apart from poor station data, by reanalysis data (NCEP/NCAR [1], NCEP-CFSR [2], ERA-Interim [3], MERRA [4], NARR [5], ASR [6,7], etc.) and global climate modeling data with a spatial resolution of about tens of kilometers. This resolution may be insufficient for many goals and applications. For example, extreme weather events, especially genesis of extreme wind speeds, are evidently associated with local-scale and mesoscale processes including non-hydrostatic effects. In Kislov et al. [8] it was shown that a global model could not reproduce a significant part of wind extremes distribution, so-called 'dragons' [9]. It is good reason to use more detailed resolution and more precise models and

simulations, namely, non-hydrostatic mesoscale regional atmospheric models. There is an additional advantage for use of mesoscale models, specifically for taking into account complex terrain characteristics, such as coastline, mountain topography, detailed surface features, etc.

However, an investigation of extreme weather events and wind speeds could be considered using the long-term simulations and robust statistical estimates of outputs. The combination of high resolution, detailing of mesoscale processes, a long period for analysis and assessments is a good challenge for new research, explanations of physical mechanisms and possible trends of extreme wind speeds. Solution of the corresponding task requires an appropriate model tool, computational resources; many test runs for model configuration optimization, selection of model domains, and development of a general technology, validation and verification of model results.

## 2. Data and methods

### 2.1 Model description
The COSMO-CLM model (ver. 5.0) is used as the main tool for the creation of this long-term meteorological archive. COSMO-CLM is the climate version of the well-known mesoscale COSMO model, including some modifications and extensions adapted to the long-term numerical experiments. It is developed by German Weather Service (DWD) and CLM-Community (see CLM Community site http://www.clm-community.eu/, [10, 11]). The COSMO-CLM model is based on the primitive Navier-Stokes equations describing the dynamics of compressible fluid in the moist atmosphere. The model equations are solved on the rotational grid 'latitude-longitude' ($\lambda$, $\varphi$) with a pole tilt. It helps to minimize the problem of meridians convergence over the pole. The numerical scheme is realized on Arakawa C-grid, and the vertical coordinate is a hybrid Gal-Chen coordinate.

The standard configuration of the COSMO-CLM model was applied with the Runge-Kutta integration scheme with $5^{th}$ advection order. The height-based hybrid Gal-Chen coordinate [12] is given as an analogue of the $\sigma$-coordinate from the surface up to $Z_0$ level and as the Z-coordinate above the $Z_0$ level. The Ritter and Geleyn radiation scheme [13] is based on the $\delta$ two-stream version of the radiation transfer equation. The precipitation formation described by a bulk microphysics parameterization, Tiedtke mass-flux schemes with equilibrium closure based on moisture convergence are used for moist and shallow convection [14]. Turbulence is described by a prognostic TKE-based scheme, with 2.5 order closure [15]; the spectral nudging technique [16] and Smagorinsky diffusion included. A full description of the COSMO Model physics, dynamics and parameterizations is available on (http://www.cosmo-model.org/content/model/documentation/core/default.htm).

### 2.2 Experiments and downscaling technology description
The COSMO-CLM model runs were executed using ERA-Interim reanalysis [3] data (~$0.75^0$ resolution) as driving conditions over the base domain (~13.2 km resolution) for the 1985 – 2014 period (i.e. 30 years). Additionally, many external parameters data (e.g., land/sea mask, roughness length, leaf area index, etc.) came from EXTPAR v.3.0 tool provided by CLM-Community (www.clm-community.eu/index.php?menuid=174&reporeid=260) over the base domain (data from EXTPAR are provided in Table 1). All these data went into preprocessing routine int2lm and were interpolated onto the domain grid.

Generally, the 'nested domains' technique was realized through three nested computational areas, with 13.2, 6.6, and 2.2 km resolutions. As a result, more than 20 detailed meteorological fields (see Appendix) within the 30-year period with 1-hour temporal resolution over the Okhotsk Sea and Sakhalin region were obtained. Initially, the model run was executed over the starting domain ('base domain') with a raw 13.2-km resolution through the whole period. After that, these outputs were used as initial and boundary conditions for interpolation and model run over the next domain, with 6.6 km resolution. And, finally, in the same way, we have executed model runs over the 2.2 km resolution domain with the driving conditions from the previous domain's outputs. However, some shortcomings of the downscaling method should be noted. The first one is associated with occurring artifacts and

spurious effects at the domain boundaries. The second one is the bias associated with coarse resolution, which the regional climate model also adds by its own errors to the output data. Nevertheless, using the 'nested domains' scheme with finer resolution and reasonable choice of boundaries could allow us to reduce these errors and biases.

**Table 1.** External parameters list provided by EXTPAR tool and required for the COSMO-CLM model run initialization.

| Parameter | Name | Units |
|-----------|------|-------|
| HSURF | Geometrical height | m |
| Z0 | Roughness length | m |
| LAI_V | Leaf area index, vegetation period | |
| LAI_R | Leaf area index, non-vegetation period | |
| FR_LAND | Land cover fraction | 1 |
| PLCOV_V | ground fraction covered by plants, vegetation period | 1 |
| PLCOV_R | ground fraction covered by plants, non-vegetation period | 1 |
| ROOTDP | Root depth | m |
| SOILTYP | Soil type | - |
| T_CL | Soil temperature | K |
| FR_LAKE | Lakes fraction cover | 1 |

The 'spectral nudging' technique was additionally applied over the base domain in order to control the model behavior and link to the real atmospheric dynamics. This technique assimilates the large-scale components of atmospheric circulation from reanalysis data not only on the boundaries, but also within the whole model domain. It limits the possible model retreat from the real conditions. It is based on the two-dimensional Fourier decomposition of reanalysis and regional model fields and the succeeding adjustment of simulation results. The temperature, wind speed, geopotential, and pressure were assimilated using this spectral nudging technique. Taking into account that the global control of atmospheric circulation is executed by the large-scale systems in the middle and upper troposphere, we have assumed 500 km and more as the spatial scale, and 850-hPa pressure level as the lowest for assimilation.

Practically, this computational scheme was realized with model runs for a period of several months (from 3 – 4 up to a year). Such duration was used because of restrictions of computational resources and data storage volumes, as well as a risk of technical crashes of experiments during the continuous runs for longer periods. All simulations were executed on 288 cores of the RSC "Tornado" supercomputer system in the Main Computer Center of Roshydromet with a peak performance of 35 Tflops. 30-year runs over 13.2 and 6.6 km domains together used ~2,400 hours of the CPU time. Additionally, many hours consumed test runs, debugging experiments, and simulations over the 2.2-km domain (for the most extreme wind speed events), approximately 200 – 300 CPU hours.

*2.3. Selection of domains' boundaries*
The selection of domains for simulations is one of the most important stages of preparation to simulations. There are many competing considerations to take into account. First, it is necessary to find a compromise between the limitations of computational resources and the size of the territory. Second, it is not recommended to draw the domain boundary near large mountain ranges, because it affects gravitational boundary waves propagation. Third, the boundaries of the next nested domain should be shifted inside from the previous ones by 4 – 10 model grids, i.e. by 20 – 50 km. Ultimately,

trying to take into account all these considerations, we have selected the following boundaries, as shown in Figure 1. The main characteristics of these domains are listed in Table 2.
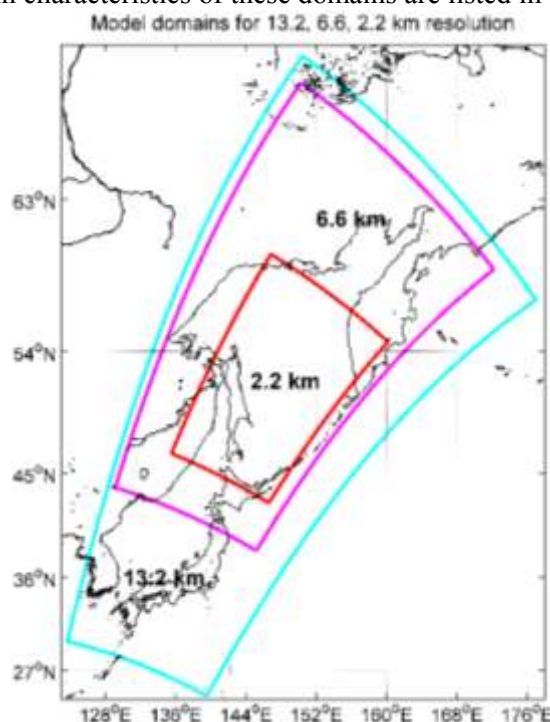


**Figure 1.** Map of boundaries of model domains with 13.2, 6.6 and 2.2 km horizontal resolutions.

**Table 2.** Main characteristics of model domains.

| Characteristics | 13.2 km domain | 6.6 km domain | 2.2 km domain |
|---|---|---|---|
| Longitude of tilted pole | $110^0$ E | $110^0$ E | $110^0$ E |
| Latitude of tilted pole | $60^0$ N | $60^0$ N | $60^0$ N |
| Total number of grid points | 145*355 = 51475 | 228*525 = 119700 | 300*500 = 150000 |
| Horizontal resolution, grad (km) | $0.12^0$ (~13.2 km) | $0.06^0$ (~6.6 km) | $0.02^0$ (~2.2 km) |
| Time step, sec | 120 | 60 | 20 |
| Number of model levels in atmosphere | 40 | 40 | 50 |
| Number of model levels in soil | 9 | 9 | 9 |
| Source of initial and boundary conditions | ERA-Interim ($\sim0.75^0$) | COSMO-CLM 13.2 km | COSMO-CLM 6.6 km |

An important reason to select boundaries for the 'base' domain (13.2 km resolution) was accounting for the synoptic-scale features of atmospheric circulation over the given region, including the monsoon system and the associated cyclonic activity, both in summer typhoons and winter polar fronts. Therefore, the eastern boundaries were extended farther to the Pacific Ocean. It allowed to catch out severe cyclones at early stages of its formation and reproduce its evolution better. A realistic simulation over the 'base' domain was important to get adequate driving conditions for the 6.6 km

domain runs. The boundaries of the 6.6 km domain are only slightly smaller than the 'base' one, contouring the Okhotsk Sea. This boundary seems to be sufficient for a good reproduction of the wind speed climatology over the region. The 2.2 km domain covers the Sakhalin Island, parts of the Okhotsk Sea, the Kamchatka peninsula, and the Japan Sea. Since the main goal of this domain was to simulate the extreme wind speeds near the Sakhalin Island, this domain was enlarged over the Okhotsk Sea in order to resolve and reproduce many mesoscale cyclonic features contributing to the formation of extreme winds.

## 3. Results and discussion

### 3.1. Verification of dataset

The verification of the obtained dataset was performed with observation data obtained from the "Hydrometcenter database" and www.rp5.ru archive on temperature, wind speed, and wind gusts ($U_g$ > 10 m/s). The stations located in less than 100 km from the coast and 500 m of absolute height were sorted out. Estimates of the model quality were obtained over three verification areas (see Figure 2): 1) Okhotsk sea, adjacent water and coastal areas (124 stations), 2) Sakhalin Island, strait of Tartary, the southern part of Okhotsk Sea (50 stations), and 3) the northern part of the eastern Sakhalin coast (6 stations). The primary comparison was made for 2014 because of the most thorough observational data archive. We have considered four periods for assessment – January-March (JFM), April-June (AMJ), July-September (JAS), and October-December (OND) of 2014.
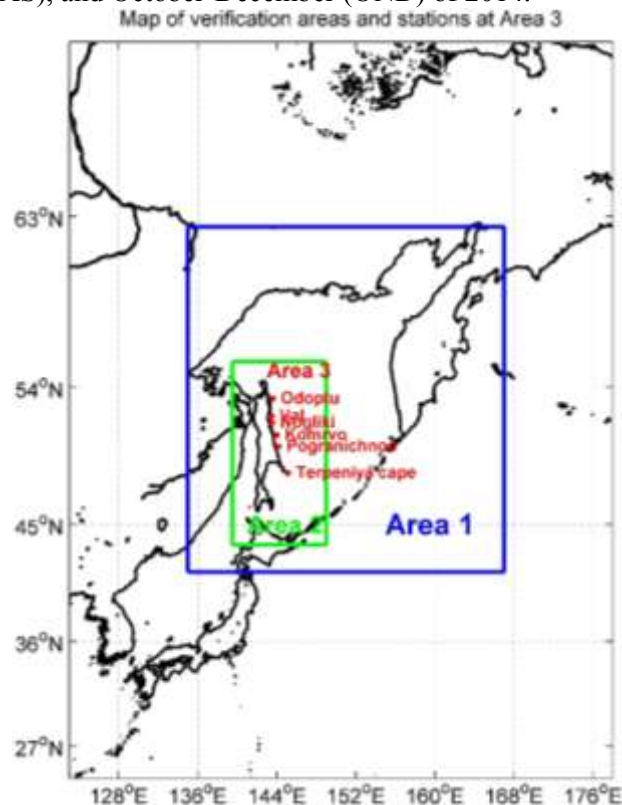


**Figure 2.** Map of verification areas 1, 2 and stations in area 3.

We applied a specific technique for verification of the modelled temperature and wind speed. For each station, that model grid for comparison was defined which had the minimal RMSE with observations. The corresponding model grid was found over the 25 x 25 km square around the station, i.e. 3 x 3 model grids with a 13.2-km resolution and 5 x 5 model grids with a 6.6 km resolution. Moreover, stations having more than 5% observation lacunas were rejected.

Summarizing the verification results, the mean temperature errors are about 0.5 $^0$C, while the RMSE reaches 2 - 3 $^0$C. The wind speed is overestimated by the model over inlands (RMSE is up to 2 m/s). It's noteworthy that wind gusts were reproduced by the model rather well (the ME was up to 1 m/s, the RMSE was 2 – 3.5 m/s, and the correlation coefficients were 0.8 and more), despite a fairly simple algorithm [17]. Additionally, the most noticeable errors are for Area 1, 2 the errors are less for Area 2, and much less for Area 3.

For several stations, errors may be high, but it could be explained by errors in the model land-mask and its inconsistences with the real surface and terrain in these cases. Seasonal courses of the parameters and error were revealed. RMSE for temperature over Area 1 is maximal in winters and minimal in summers. This is associated with overestimation of temperature by the model during strong freezings and, consequently, stable stratifications near the surface. For the Sakhalin's eastern coast, the situation is vice versa. The correlation coefficients for temperature are more than 0.9 during the entire year; therefore, the model reproduces its synoptic and daily variability correctly. The largest RMSEs for wind speed in the OND period, i.e. the 'storm season', are associated with maxima absolute values in this time. Thus, the lower correlation coefficients during the summer season could be explained by the frequency of convective movements. The transition from the resolution of 13.2 km to 6.6 km does not provide any significant improvement, but one should note a slight decrease of the errors and increase of the correlation coefficients. Also, a decrease of the errors range for parts of the seasons, parameters and areas was observed.

As for the spatial distribution of errors, maximal temperature errors are observed at the coastal stations (mostly in summer) or in inner lands (mostly in winter). The summer errors could be explained by a strong contrast between the sea and land and, hence, by the complexity in reproduction of the local land-sea interactions. The winter errors are linked to the well-known overestimation of turbulent mixing in the boundary layer by the COSMO-CLM model in stable stratification conditions that leads to rising of the model temperature.

The spatial distribution of wind speed errors is simpler. They are larger over the inlands, because of inconsistency between the real and model-defined roughness length in the land-sea mask for several stations. The errors for the coastal stations are smaller due to a less number of factors determining the wind speed.

Verification for the 2.2-km domain was performed only for extreme wind situations listed in Table 3. Generally, comparison characteristics for these short periods of extremes are not worse than analogous ones for seasons. Therefore, the model could reproduce both the background seasonal variability and its dynamics during the extreme winds with the same quality. Using finer resolution slightly improves the model results, especially by transition to 2.2 km.

*3.2. Synoptic analysis of extreme wind situations*
Extreme wind speed situations were sorted out from the given archive (15 cases listed in Table 3) and analyzed according to the synoptic processes. Two main types of synoptic situations were identified as favorable for the genesis of extreme winds. All selected storms were observed during the cold season. The synoptic features leading to extreme wind speeds were associated with a strong thermal gradient between the cold continental air mass over Eastern Eurasia and the marine polar air masses over the Pacific and Okhotsk Sea during winter. It leads to strengthening of the cyclonic activity at the polar and arctic fronts over the Okhotsk Sea and Sakhalin Island. All selected extreme winds were of northern directions and caused by cyclones.

The first type (most frequent) was associated with developing of cyclones over the Primorsky Kray or Japan Sea (example for 28.01.1989, 07 UTC, Figure 3a). Then it intensified, moved to the west or north coast of the Sakhalin Island, crossed it, and came to the Okhotsk Sea. Here, in the rear of the cyclone, the wind speed usually intensified again. The intensity and duration of extreme winds cases varied from a few hours to a day and 25 to 35 m/s.

**Table 3.** Dates of extreme synoptic situations over the Okhotsk Sea region and maximal wind speed values (m/s) reproduced by COSMO-CLM model.

| № | Date | Wind speed, m/s | | |
|---|---|---|---|---|
| | | 13.2 km | 6.6 km | 2.2 km |
| 1 | 25.03.1987 | 29.7 | 31.6 | 31.1 |
| 2 | 28.01.1989 | 32.3 | 34.3 | 34.3 |
| 3 | 19.12.1989 | 28.6 | 30.0 | 31.8 |
| 4 | 8.11.1995 | 27.0 | 26.7 | 27.0 |
| 5 | 1.02.1996 | 28.3 | 29.2 | 28.9 |
| 6 | 14.02.1996 | 32.8 | 34.2 | 34.0 |
| 7 | 28.02.1999 | 33.3 | 35.2 | 35.5 |
| 8 | 14.11.1999 | 27.4 | 25.3 | 28.7 |
| 9 | 11.01.2000 | 28.3 | 29.3 | 29.6 |
| 10 | 12.01.2001 | 30.8 | 30.3 | 30.7 |
| 11 | 5.12.2002 | 27.7 | 30.0 | 29.0 |
| 12 | 8.01.2005 | 29.8 | 28.7 | 30.6 |
| 13 | 27.02.2006 | 28.9 | 29.0 | 30.6 |
| 14 | 3.02.2014 | 27.7 | 25.5 | 26.7 |
| 15 | 6.03.2014 | 24.9 | 24.8 | 26.8 |

The second type was characterized by the penetration of intensified cyclones from the Pacific Ocean through the Kamchatka peninsula, Kuril or Japan Islands, especially Hokkaido (example for 19.12.1989, 11 UTC, Figure 3b). Next, the deepening cyclones crossed the Okhotsk Sea and got close to the Sakhalin Island, causing extreme winds offshore. This group was characterized by less deep but larger cyclones.

A comparison of the reproduction of these extreme situations at different model resolutions showed that on 2.2 km the wind speed maximums are slightly higher (the added value is 2 – 3 m/s) and its spatial distribution is more spotted compared to rawer resolutions. In most cases using a finer resolution leads to decreasing of the mean speed and gusts errors. At the same time, these features are not so evident by transition from 13.2 to 6.6 km, in contrast to the transition from 6.6 km to 2.2 km. It follows that 6.6 km resolution does not provide a significant improvement of the model results due to a bad resolving of the 'grey zone' processes, obviously. However, the 2.2 km resolution improves the reproduction of extreme winds in the surface layer because of explicit resolving of deep convection, detailed terrain, coastline, and turbulent motions description.
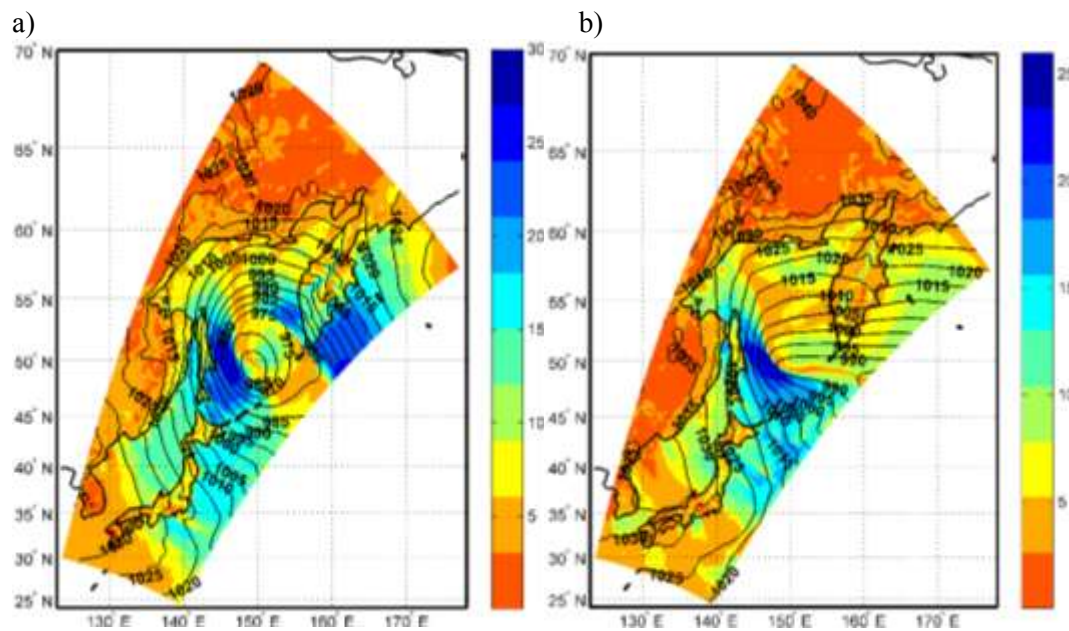
**Figure 3.** Sea level pressure (hPa) and wind speed (m/s) 28.01.1989, 07 UTC (a) and 19.12.1989, 11 UTC (b) reproduced by COSMO-CLM model on the 13.2 km domain.

## 4. Conclusions and recommendations for future research

It is for the first time in the Russian Federation that a technology of detailed regional modelling of atmospheric processes was realized on long-term timescales. Overall, the analysis showed that the used downscaling technique with applying the COSMO-CLM model reproduced the meteorological conditions, spatial distribution, seasonal and synoptic variability of temperature and extreme wind speed for the study area with approximately the same adequate quality. Some relations between the reproduction quality of mesoscale atmospheric circulation features and the horizontal resolution of the model were revealed. In particular, it was shown that ~6.6 km resolution does not provide any significant improvement in comparison to ~13.2 km resolution, whereas ~2.2 km resolution provides an appreciable improvement of the quality.

The obtained high-resolution dataset will be used for a full and comprehensive analysis of the physical mechanisms of extreme weather events, the reproduction quality of hydrometeorological fields, their statistical estimates, climatological trends, spatial distribution of a large variety of hydrometeorological parameters on diurnal, seasonal, and interannual timescales, using this information for a detailed environment state assessment and many other objectives.

## Appendix

List of main output variables:
- Sea level pressure (hPa)
- Air temperature 2 and 10 m ($^0$C)

- Potential air temperature 2 and 10 m ($^0$C)
- Surface temperature ($^0$C)
- Specific humidity 2 and 10 m (g/kg)
- 10 m zonal and meridional wind (m/s)
- 10 m wind speed maxima during 1 hour (m/s)
- 10 m wind gust during 1 hour (m/s)
- Hourly snow and rain precipitation (kg/m$^2$, both)
- Radiation (direct and diffuse) shorter and longer than 700 nm  (W/m$^2$, both)
- Downward shortwave radiation (W/m$^2$)
- Upward shortwave radiation (W/m$^2$)
- Downward and upward longwave radiation (W/m$^2$)
- Sensible and latent heat flux (W/m$^2$)

### References

[1]  Kalnay E *et al.* 1996 The NCEP/NCAR 40-year reanalysis project  *BAMS* **77** 3 437 – 71
[2]  Saha S *et al.* 2010 The NCEP climate forecast system reanalysis  *BAMS* **91** 1015
[3]  Dee D P *et al.* 2011The ERA-Interim reanalysis: Configuration and performance of the data assimilation system  *Quart. J. RMS* **137** 656 553 – 97
[4]  Rienecker M M *et al.* 2011MERRA: NASA's modern-era retrospective analysis for research and applications *J. of Climate* **24** 14 3624 – 48
[5]  Mesinger F *et al.* 2006North American Regional Reanalysis  *BAMS* **3** 343 – 60
[6]  Bromwich D H *et al.* 2009 Data assimilation challenges for high-resolution reanalysis in the Polar regions: The Arctic System Reanlaysis *10th Annual WRF Users' Workshop* 1 – 4
[7]  Bromwich D H, Wilson A B, Bai L-S, Moore G W K, Bauer P A 2016 Comparison of the regional Arctic System Reanalysis and the global ERA-Interim Reanalysis for the Arctic *Quart. J. of RMS* **142** 644 – 58
[8]  Kislov A, Matveeva T, Platonov V 2015 Wind speed extremes in Arctic area 2015 *Fund. and Appl. Clim.* **2** 63–80 (In Russian)
[9]  Kislov A, Matveeva T 2016 An extreme value analysis of wind speed over the European and Siberian parts of Arctic region *Atm. and Clim. Sci.* **6** 205 – 23
[10] Böhm U, Kuecken M, Ahrens W, Block A, Hauffe D, Keuler K, Rockel B, Will A 2006 CLM – the climate version of LM: brief description and long-term applications *COSMO Newsletters* **6** 225–35
[11] Rockel B, Will A,  Hense A 2008 The regional climate model COSMO-CLM (CCLM) *Met. Zeit.* **17** 4 347–48
[12] Gal-Chen T, Somerville R C J 1975 On the use of a coordinate transformation for the solution of the Navier-Stokes equations *J. of Comput. Physics* **17** 2 209 – 28
[13] Ritter B, Geleyn J F 1992 A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations *Mon. W. Rev.* **120** 2 303 – 25
[14] Tiedtke M 1989 A comprehensive mass flux scheme for cumulus parameterization in large-scale models *Mon. Wea. Rev.* **117** 8 1779 – 800
[15] Herzog H-J, Vogel G and Schubert U 2002 LLM – a nonhydrostatic model applied to high-resolving simulations of turbulent fluxes over heterogeneous terrain *Theor. Appl. Clim.* **73** 67 – 86
[16] Von Storch H, Langenberg H, Feser F 2000 A spectral nudging technique for dynamical downscaling purposes *Mon. Wea. Rev.* **128** 10 3664 – 73
[17] Schulz J-P, Heise E 2003 A new scheme for diagnosing near-surface convective gusts *COSMO Newsletter* **3** 221 – 5