

CORDEX data visualization using QGIS

Issues and solutions on the example of Kazakhstan

Kalamkas YESSIMKHANOVA – Mátyás GEDE

DOI: 10.30921/GK.74.2022.4.4

Absztrakt: Számos tanulmány foglalkozik az éghajlatváltozás elemzésével és az éghajlati szimulációk vizsgálatával. Hiányoznak azonban az éghajlati adatok vizualizációjáról szóló tanulmányok, különös tekintettel a CORDEX- (Coordinated Regional Downscaling Experiment) adatokra. Ennek a munkának a célja a NetCDF (Network Common Data Form) fájlformátumban tárolt éghajlati adatok vizualizációs szempontjainak áttekintése, különös figyelmet fordítva a CORDEX adatok térbeli referenciaproblémájára. Ezenkívül nagyon kevés tanulmány foglalkozik Kazahsztán területével, míg egyes régiókban sokféle éghajlati adat és modell áll rendelkezésre. Ebben a tekintetben is érdekes a régió elemzése. A projekt eredményeit Kazahsztán különböző időszakokban előrejelzett hőmérséklet-változását ábrázoló térképek mutatják be.

Abstract: Many studies are devoted to climate change analysis and climate simulations examination. However, there is a lack of studies about climatic data visualization, specifically of CORDEX (Coordinated Regional Downscaling Experiment) data. Therefore, this work is aimed at reviewing visualization aspects of climate data stored in NetCDF (Network Common Data Form) file format. Particular attention was given to spatial reference matter of CORDEX data. Additionally, very few studies are available for the area of Kazakhstan while there is multiplicity of climate data and models for some regions. In this regard, carrying out analysis for this region is valuable. Results of the project are performed as maps visualizing temperature change across Kazakhstan over different time periods.

Kulcsszavak: klimatikus adatok ábrázolása, georeferálás, klímaváltozás, CORDEX, NetCDF, QGIS

Keywords: climate data visualization, georeferencing, climate change, CORDEX, NetCDF, QGIS

Introduction

Like any process, climate change has consequences both globally and locally. In this regard, it has an importance to visualize climate changes, because visualization facilitates insight and interpretation of data and processes. One of the effective visualization tools is a map. It is feasible to demonstrate the scale of the problem through maps. Such powerful tool should be used for society awareness of its social responsibility.

Modelling climate change is important due to its application urgency. Models based on climate data simulations provided by Coordinated Regional Downscaling Experiment (CORDEX) are reliable and extensive. CORDEX initiative was created to provide climate data for different regions and in several resolutions (Giorgi Filippo – Gutowski William J., 2015). Particularly in this work, the Regional Climate Model (RCM) for Central Asia is going to be considered (WRCP CORDEX 2022). Kazakhstan has a huge territory, precisely 2,724,900 km². So, when we speak about climate and climate changing, it should be considered that climate is dissimilar at different parts

of the country. Consequently, climate change policy should be reckoned accordingly. That is why local models are useful and reliable for risk assessment, decision-making processes and allow undertaking measures for preventing and retaining risk susceptible areas. Climate simulations are helpful in rendering images of future models. Based on the results of future changes in temperature it can be figured out what effects it can cause in climate system in future perspectives.

Most works related to climate data are not focused on visualization part, only on the analysis of the data itself. Since climate data is divided by large regions within CORDEX like Central Asia, Europe, and Central America and so on, data analysis, usually, is presented for whole regions. Moreover, plotted data were generated by software not designed for cartographic visualization. So, from cartographic point of view it has visualization issues, such as illegibility of small areas, lack of labels and other insufficiencies. Additionally, modification of the data is limited in software for creating plots (for instance, Panoply) and that is why GIS (Geographic Information System) applications are relevant.

The first step of data handling in GIS applications is georeferencing the data. However, georeferencing of CORDEX data, which is stored in NetCDF format, may collapse due to misinterpretation which may occur in QGIS application. For this reason, the present study was elaborated to get insights how to cope with these difficulties. Two ways of georeferencing data in NetCDF format are considered. Results of this research may contribute to spreading local climate models for the territory of Kazakhstan due to its deficient volume. As the objective was to visually represent climate data, results are provided with sample maps of Kazakhstan revealing temperature change over time periods.

Methodology

The research methodology is built on a comprehensive approach of data collection, analysis and compilation of raw data, and comparison of the results. Visual interpretation of spatial climate data can be accomplished with GIS technology products. Powerful GIS software, such as QGIS/ArcGIS, allow handling processes of any complexity. QGIS was deliberately used due to its

free availability. QGIS can process a great deal of file formats with no pre-processing. However, some of them require to be prepared.

Visual interpretation of spatial data is going to be based on climatic data stored in NetCDF file format. These data are one of the comprehensive base sources and easy to process. Source data were obtained from the official website of the Earth System Grid Federation (ESGF) in NetCDF file format. Climate information for Central Asia domain is accessible in two spatial resolutions: 0.220 and 0.440 (25 and 50 km). Since Kazakhstan is located in Central Asia, temperature change is analysed for 25 km model resolution, which is the best for this region. Near surface air temperature change assessment was performed based on monthly values for optimistic scenario in RCM ALARO-0 with driving model of CNRM-CERFACS-CNRM-CM5. CNRM-CM5 is a model of CNRM-CERFACS (National Centre for Meteorological Research and Centre Européen de Recherche et de Formation Avancée) (Voltaire et al. 2013).

Climate Data Operators (CDO) is an assembly of command operation lines to handle and analyse climate data. The operators include starting from simple statistics and mathematics functions to complex calculation like spatial interpolation (Max Planck Institut 2022).

CDO commands were compiled in Cygwin environment which runs on Microsoft Windows. Codes created for Unix-like OS (Operating System) can be compiled in Cygwin tool, which is the

case with CDO. All data are processed with CDO commands and in QGIS. All CDO commands were taken from CDO guidebook (Schulzweida, 2021).

First of all, preliminary operations are required in order to open climate data stored in NetCDF file format in GIS software. Data manipulation, specifically, merging datasets, operate statistical values (unit conversion), arithmetic operations (subtracting, computing mean values), and re-projecting were handled with CDO. For instance, to compute mean value for a particular time period (2040–2070), the CDO command below was applied.

```
cdo timmean-selyear,2040/2070
input.nc output.nc
```

Next stage is opening raster image in QGIS. Although initial data is georeferenced, sometimes it will not be placed correctly in QGIS without coordinate system transformation. This is caused by missing a line of code with projection parameters, which QGIS needs for data referencing. Projection parameters are already encoded in all data, and applications devoted for generating a plot are able to interpret that information. However, QGIS has a special codification of projection parameters, and it is not able to read out projection information if it does not match unified code. Figure 1 shows demonstration of the data with no projection conversion in QGIS environment. Figure 1 illustrates world countries boundaries, where the data should cover the region of Central Asia. However, it is displaced and displayed incorrectly.

One of the solutions is re-projecting data to longitude/latitude coordinate system, which is the CORDEX standard. It has relevance, because the native coordinate system of RCM ALARO-0 uses Lambert Conformal Conic projection (Top, S. et al. 2020). In order to make process of model georeferencing more convenient it is practical to apply interpolation with CDO. Firstly, interpolation with CDO takes only single-line code which is shown below.

```
cdo remapbil,n256 input.nc output.nc
```

Secondly, GIS applications handle Coordinate Reference System (CRS) of the data with no errors. In the current study the grid was interpolated to Gaussian N256. Gaussian grid is a rectangular horizontal coordinate system with an assigned number of coordinates. It is used for scientific modelling on a sphere (Hortal et al. 1991). Bilinear interpolation method was used for re-projecting. Bilinear interpolation method only works on quadrilateral curvilinear source grids. The selection of interpolation method is not stipulated by any criteria, so it was simply random choice to visualize the data for Kazakhstan territory. There are several methods of interpolation available with CDO (Max Planck Institut 2022). Details about spatial reference can be extracted with CDO command line below.

```
cdo griddes input.nc
```

Main spatial reference attributes of the original projection and interpolated grid are provided in Table 1.

Table 1. Grid attributes

| | Original projection | Interpolated grid |
|-------------------------------|----------------------------|-------------------|
| Grid mapping | Lambert Conformal | Gaussian 256 |
| Grid type | projection | Gaussian |
| Grid size | 74259 | 524288 |
| X size | 333 | 1024 |
| Y size | 223 | 512 |
| X name | x | lon |
| X long name | x coordinate of projection | longitude |
| X units | km | degrees east |
| Y name | y | lat |
| Y long name | y coordinate of projection | latitude |
| Y units | km | degrees north |
| X first | 0 | 0.3515625 |
| Xinc | 25 | |
| Y first | 0 | |
| Yinc | 25 | |
| Longitude of central meridian | 74.64 | |
| Latitude of projection origin | 47.82 | |
| Standard parallel | 47.82 | |

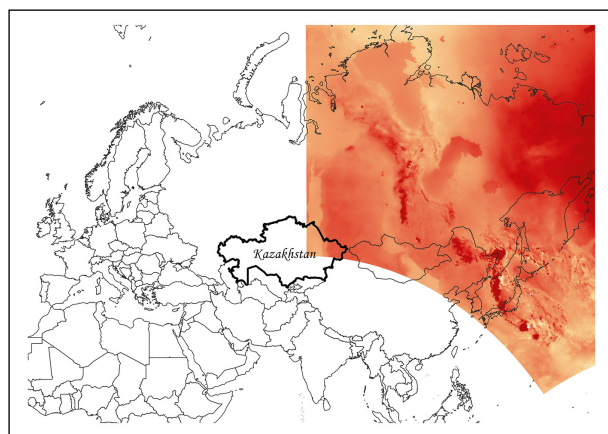


Figure 1. Raster data is displaced in QGIS due to the lack of correct georeferenced definition

Once the interpolation is completed, georeferencing settings in QGIS are not required and the data is placed correctly automatically. Figure 2 reveals the proper alignment of interested area after interpolation. Here, the data is placed into the area where it belongs to in Central Asia.

This result can be reached out with applying another approach as well. As mentioned above, the code line with projection parameters is missing in the original file. Therefore, adding missing command would be logically suitable. In order to re-encode projection information in original data, an external *.txt file should be created with required features of projection. New grid information file can be embedded into initial file by following CDO command.

```
cdo remapbil,grid.txt input.nc output.nc
```

QGIS decodes projection information in the following form that consists of name of projection, longitude of central meridian, latitude of projection origin, standard parallel, false easting, false northing, ellipsoid, and its semi-major axis. All essential parameters are encrypted in the way given below.

```
proj4_params = "+proj=lcc
+lat_1=47.82 +lat_2=47.82
+lat_0=47.82 +lon_0=74.64
+x_0=-12500. +y_0=-12500.
+ellps=sphere +a=6371229.
+b=6371229. +units=m +no_defs"
```

Values of settings are presented in the Table 1 and can be derived from initial data file. Moreover, take notice of X and Y of first coordinates. In this case, Table 1 shows that margins of X first and Y first are equal to 0. However, these parameters should be modified by simple calculation of them. This re-evaluation is critical due to its importance because they are referencing points of projection. These two parameters differ, and values depend on raster size. Recalculation is simple and requires elementary math operations. Half sizes of raster multiplied by resolution give starting point values of X and Y axes. So, while embedding projection information to initial file values of X first and Y first was changed from 0 to X first = -4162500, Y first = -2787500.

The coordinate system of QGIS environment was set up to WGS-84 (World Geodetic System 1984) and the world country outlines shapefile has the same referencing. Therefore, Figures 1 and 2 are demonstrated in WGS-84 coordinate system. However, the coordinate system was changed to Pseudo-Mercator after data sampling (Figure 3).

The following stage is data selection for the interested area. Sampling data for the region of Kazakhstan is one of the critical steps as it gives clear vision of what data is supposed to

be processed. Since CORDEX data is distributed by regions, deriving data by country outlines is impossible. Thus, selecting the area of interest was proceeded in QGIS by clipping the region with the polygon of Kazakhstan.

There are other applications for generating plots and animations apart from mentioned programs, for instance Panoply (Schmunk, 2020). However, this application is not user-friendly enough for data management. In Panoply, modification of data visualization is extremely limited compared to QGIS. In Panoply, it is not possible to upload shapefiles and operate with them. If there is a necessity of demonstrating large regions, for instance, whole Europe or Central America, the use of these applications is justified. If sampling area by country is essential, then GIS applications are handier. Apart from this, applying proper symbology and labelling is more convenient in GIS software. So, the choice of software depends on your demands. If the goal is simply visualizing the data for presentation purposes, then it is worth using Panoply. However, if the purpose is to analyse data and do some complex manipulation with data, it is rational to choose QGIS or any other GIS desktop applications.

Work outcomes are shown as georeferenced maps of Kazakhstan with temperature change statistics. Maps indicate temperature change between the historic reference and the rest of the 21st century period in Kazakhstan. Timeframe from 1977 to 2005 is the historical reference and future time is divided by 30 years. Standard time to witness climate change is 30 years and it is identified by the World Meteorological Organization (WMO) (Arguez&Vose, 2011). So, in this project three 30-year time periods 2010–2040, 2040–2070 and 2070–2100 were set up. Figure 3 illustrates temperature changes in Kazakhstan and it shows that all changes are positive, which assumes that mean temperature will increase over time. It is notable that these simulated changes correspond

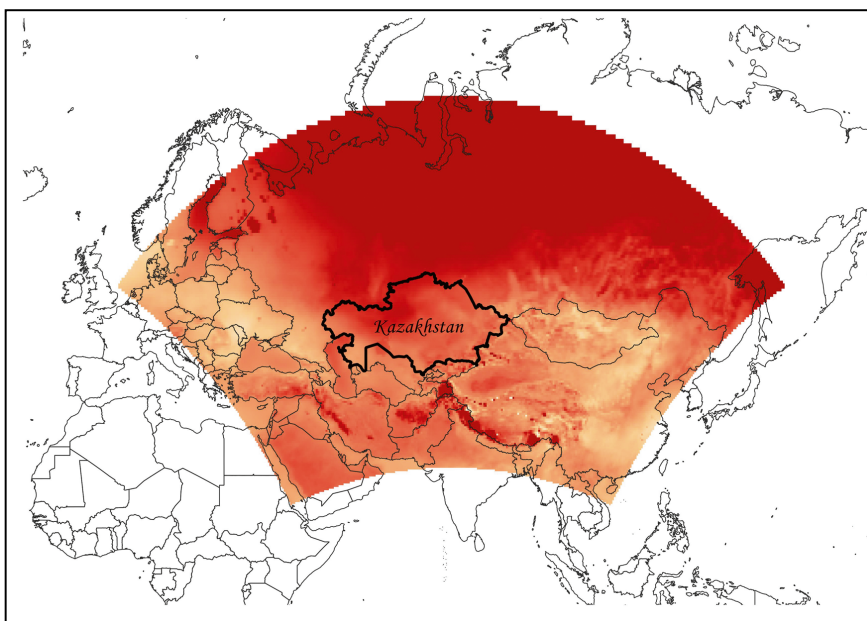


Figure 2. Raster data is at the correct location after re-projection

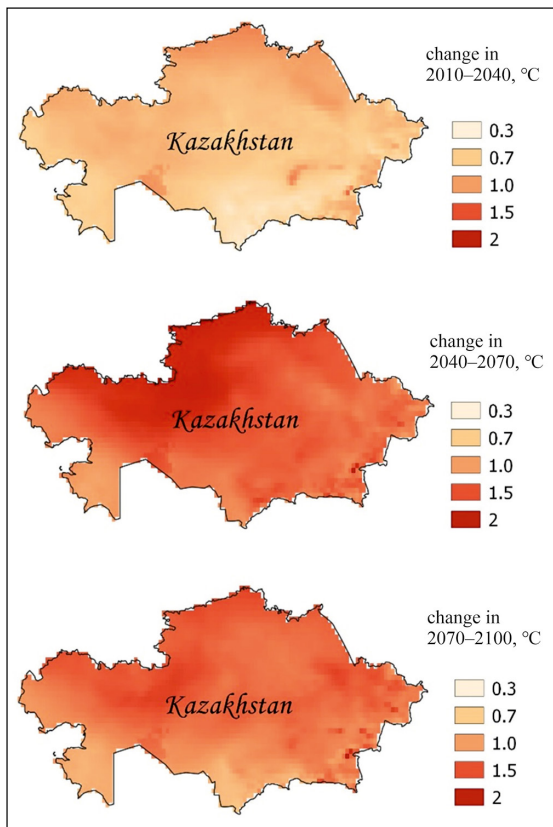


Figure 3. Expected changes in mean temperature in 30 years in Kazakhstan

to Representative Concentration Pathway (RCP) 2.6. This is one of the four pathways which means forcing level will decrease to 2.6 W/m² by 2100, which is the best scenario of probable future emission pathways (van Vuuren et al. 2011)

Mean temperature of time period 2010–2040 will be higher than 1977–2005 for 0.3–1.0 Celsius degrees. Temperature will alter considerably between 2040 and 2070, so compared to 1977–2005, the difference is drastically higher up to 2 Celsius degrees. Specifically, northwestern part of the territory will be warmer for 2 centigrade and the rest of the country is going to be warmer for about 1.5 Celsius degrees. Interval 2070–2100 represents cooling down after 2040–2070 approximately for 0.5 centigrade, however regardless, compared to historical reference point temperature will be higher, and shifts lie between 0.7 and 1.5 centigrade. Overall, graphics demonstrate warming process and show impact of global warming on the study area.

This study depicts that mean temperature will increase for 1.5–2 Celsius degrees within the 21st century even with the best possible emission scenario. It corroborates scientific research showing temperature rises globally. So, according to this calculations Kazakhstan should take into action climate change strategy. Since Kazakhstan ratified the Paris Agreement on climate change it also works on implementing of Sustainable Development Goals by United Nations General Assembly. This work may help in monitoring the situation on temperature change in the country within the 13th sustainable goal (UNDP 2022).

Summary

There are some regions where there is shortage of visualized climate data like Kazakhstan. Disseminating maps and process of data visualization has value for students, junior professionals, government workers, and all who work with such kind of data. Data visualization perspectives of CORDEX data are shown in this study. From cartographic point, attention was given to data georeferencing. Two approaches of spatial referencing are covered in the research. In addition, other visualization issues of CORDEX data are considered. Moreover, outcomes of this work can help local government in decision-making process and elaboration of climate change policy.



Kalamkas Yessimkhanova
PhD student

ELTE Eötvös Loránd University,
Institute of Cartography and
Geoinformatics
kalamkasyessimkhanova@gmail.com



Mátyás Gede
Associate
professor, PhD

ELTE Eötvös Loránd University,
Institute of Cartography and
Geoinformatics
saman@inf.elte.hu

References

- Arguez, A. – Vose, R. S. (2011). The Definition of the Standard WMO Climate Normal: The Key to Deriving Alternative Climate Normals, *Bulletin of the American Meteorological Society*, pp. 699–704. DOI: 10.1175/2010BAMS2955.1
- Giorgi Filippo - Gutowski William J. (2015) Regional Dynamical Downscaling and the CORDEX Initiative, *Annu. Rev. Environ. Resour., Annual Reviews*, pp. 467–490. DOI: 10.1146/annurev-environ-102014-021217
- Hortal - Mariano -Simmons, A. J. (1991). Use of reduced Gaussian grids in spectral models. *Monthly Weather Review* 119.4: 1057–1074. DOI: 10.21957/v413vy4fg
- Max Planck Institut (2022) Climate Data Operators. <https://code.mpimet.mpg.de/projects/cdo>
- Schmunk, R. (2020). NASA Goddard Institute for Space Studies: User's Guide to Panoply.
- Schulzweida, U. (2021). MPI for Meteorology: CDO User Guide.
- Top, S. – Kotova, L. – De Cruz, L. – Aniskevich, S. – Bobylev, L. – De Troch, R. – Gnatiuk, N. – Gobin, A. – Hamdi, R. – Kriegsmann, A. – Remedio, A.R. – Sakalli, A. – Van De Vyver, H. – Van Schaeuybroeck, B. – Zandersons, V. – De Maeyer, P. – Termonia, P. – Caluwaerts, S. (2021). Evaluation of regional climate models ALARO-0 and REMO2015 at 0.22° resolution over the CORDEX Central Asia domain. *Geosci. Model Dev.*, 14, 1267–1293. DOI: 10.5194/gmd-14-1267-2021
- van Vuuren, D.P. – Edmonds, J. –Kainuma, M. – Riahi, K. – Thomson, A. – Hibbard, K. – Hurtt, G.C. – Kram, T. – Krey, V. – Lamarque, J.F. – Masui, T. – Meinshausen, M. – Nakicenovic, N. – Smith, S.J. – Rose, S.K. (2011). The representative concentration pathways: an overview. *Climatic Change* 109, 5. DOI: 10.1007/s10584-011-0148-z
- Voldoire, A. – Sanchez-Gomez, E. Salas y Mélia, D. – Decharme, B. – Cassou, C. et al. (2013). The CNRM-CM5.1 global climate model: description and basic evaluation. *Climate Dynamics*, Springer Verlag, 40, 2091–2121. DOI: 10.1007/S00382-011-1259-Y
- UNDP (2022) Sustainable Development Goals. 13 – Climate Action. <https://www.undp.org/sustainable-development-goals#climate-action>
- WRCP CORDEX (2022) Region 8: Central Asia. <https://cordex.org/domains/region-8-central-asia>