

README concerning the dataset COSMO-REA6 available via DKRZ /pool/data resources

Date: 19.01.2024

T I T L E O F T H E D A T A S E T COSMO-REA6 (COSMO-REA6 Regional Reanalysis)

P A T H T O T H E D A T A S E T /pool/data/COSMO-REA

O W N E R / P R O D U C E R O F T H E D A T A S E T -Contact details of the person responsible for the dataset (also in the long run) COSMO-REA Support Contact Address, DWD, cdc.rea@dwd.de Martin Schupfner, DKRZ, schupfner@dkrz.de Michael Borsche, DWD, michael.borsche@dwd.de

D A T A U S A G E L I C E N S E - Are there any restrictions on the data reuse? If yes, please provide details. - Which license are the data associated with? COSMO-REA model data produced by DWD is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0; <https://creativecommons.org/licenses/>). Consult '<https://www.wdc-climate.de/ui/project?acronym=COSMO-REA6>' for terms of use governing COSMO-REA output, including citation requirements and proper acknowledgment. The data producers and data providers make no warranty, either express or implied, including, but not limited to, warranties of merchantability and fitness for a particular purpose. All liabilities arising from the supply of the information (including any liability arising in negligence) are excluded to the fullest extent permitted by law.

C O N T E N T O F T H E D A T A S E T - What kind of data is it? This dataset comprises reanalysis data for Europe created with the regional high-resolution reanalysis system COSMO-REA6 using global reanalysis ERA-Interim as forcing. The reanalysis dataset covers the period 1995 to August 2019 and has a spatial resolution of 0.055° (6 km). For more details, see below: Methods used for data creation. - Which variables are contained and how are they described (naming conventions, data standards)? The data has been standardised according to a slight variation of the CORDEX standard within the NFDI4Earth pilot project OcMOD. Deviations are an additional DRS element - the modeling realm. The CV incl. CMOR tables has been set up in the following gitlab repository: <https://gitlab.dkrz.de/dicad-pp/ocmod> This dataset comprises a subset of the COSMO-REA6 output (see below: Methods used for data creation), consisting of atmospheric and soil/land variables of the COSMO model with frequencies: hourly, 6-hourly, daily and monthly.

D A T A U S A G E S C E N A R I O S - What can the data be used for? - Who are expected users at DKRZ? In the survey, conducted by DWD and DKRZ to determine the scientific parameters to provide and the use cases, the following use cases were named: - Reference Data - Energy Application - Climate Dynamics - Extreme Events - Evaluation - Forcing Data - Impact Modeling The survey results are available here: [https://c6dreq.dkrz.de/files/ocmod\\_dreq.php](https://c6dreq.dkrz.de/files/ocmod_dreq.php)

M E T H O D S U S E D F O R D A T A C R E A T I O N - How was the data created? Who created the data? Do existing publications describe the data? The high-resolution reanalysis system COSMO-REA6 has been developed based on the NWP model COSMO in a cooperation between Hans-Ertel-Centre for Weather Research (HErZ) and the German Meteorological Service (DWD). Initially, COSMO-REA6 was produced for 1995-2014 at University of Bonn, Hans-Ertel-Centre for Weather Research (HErZ), see <https://www.herz.uni-bonn.de/wordpress/>. It was extended from 2015 onwards by DWD. The DWD provides selected COSMO-REA6 output under [https://opendata.dwd.de/climate\\_environment/REA/](https://opendata.dwd.de/climate_environment/REA/). The data provided via ESGF and WDCC underwent an additional CMIP/CORDEX-like standardisation (“CMORisation”), conducted in a cooperation between DKRZ and DWD within the OcMOD (Observations closer to Model Data) NFDI4Earth Pilot, and also comprises a different selection of parameters that was determined through a community survey. The regional reanalysis system for Continental Europe matches the domain of the CORDEX EURO-11 specifications, albeit at a higher spatial resolution, i.e.,  $0.055^\circ$  (6 km) instead of  $0.11^\circ$  (12 km). It comprises the assimilation of observational data using the existing nudging scheme of COSMO complemented by snow, SST and soil moisture analysis modules and uses ERA-Interim data as lateral boundary conditions. The reanalysis data set covers the period 1995 to August 2019. Due to the end of production of the forcing global reanalysis ERA-Interim, COSMO-REA6 has also been discontinued. A second version of COSMO-REA6, labeled “R6G2”, that uses the data of the global reanalysis ERA5 as lateral boundary conditions, is currently under development at DWD.

References: Bollmeyer, C.; Keller, J. D.; Ohlwein, C.; Wahl, S.; Crewell, S.; Friederichs, P.; Hense, A.; Keune, J.; Kneifel, S.; Pscheidt, I.; Redl, S.; Steinke, S. (2015). Towards a high-resolution regional reanalysis for the European CORDEX domain. doi:10.1002/qj.2486 Kaspar, Frank; Niermann, Deborah; Borsche, Michael; Fiedler, Stephanie; Keller, Jan; Potthast, Roland; Rösch, Thomas; Spanghehl, Thomas; Tinz, Birger. (2020). Regional atmospheric reanalysis activities at Deutscher Wetterdienst: review of evaluation results and application examples with a focus on renewable energy. doi:10.5194/asr-17-115-2020

Weblinks: [https://www.dwd.de/DE/klimaumwelt/klimaueberwachung/reanalyse/reanalyse\\_node.html](https://www.dwd.de/DE/klimaumwelt/klimaueberwachung/reanalyse/reanalyse_node.html)  
<https://reanalysis.meteo.uni-bonn.de/?COSMO-REA6> <https://nfdi4earth.de/2participate/pilots>  
<https://nfdi4earth.de/images/22KasparObservationscloserToModelDataOcMOD.pdf>

I S S U E S - Are there any peculiarities associated with the dataset which a data re-user should know about? Owing to the experimental nature of this reanalysis, the surface altitude (orographie) was not held constant throughout the model run, but two slightly varying versions of surface altitude were used: one for the simulation years 2007-2012 (that was published as orog(fx)), one for the rest of the model run (i.e. 1995-2006, 2013-2019). Instead of a time invariant variable “orog(fx)”, therefore a monthly mean variant “orog(mon)” has been published. The data files for atmospheric 3D variables (eg. ta, hus) that require orog as

auxiliary coordinate variable for the 'height based hybrid Gal-Chen coordinate' have the correct version of the two surface altitude fields attached.

**VOLUME OF THE DATASET ( AND POSSIBLE CHANGE S T H E R E O F )** - What is the volume of the dataset? - Will the volume of the dataset in- or decrease with time? If so, by how much? The full dataset consists of 56 Terabytes available via /pool/data/COSMO-REA . The dataset is also available via the ESGF project COSMO-REA and via the World Data Centre for Climate (<https://www.wdc-climate.de/ui/project?acronym=COSMO-REA6>). The 3D variables on atmospheric model levels available via /pool/data/COSMO-REA, ESGF and WDCC are only provided on the 6 lowest model levels. The full vertical set of model levels is provided in the DKRZ DOKU long term archive only (LINK TBA).

**T I M E H O R I Z O N O F T H E D A T A S E T O N / P O O L / D A T A** - How long is the data planned to be made available via /pool/data? 10 years, unless this version becomes obsolete (eg. by the users only relying on the to-be-provided successor R6G2 in the future). In this case the COSMO-REA6 data will still be available via the WDCC ([https://doi.org/10.26050/WDCC/CR6\\_EU6](https://doi.org/10.26050/WDCC/CR6_EU6)) and DOKU long term archives while R6G2 would be provided on disk and via the DKRZ and DWD ESGF nodes. - Storage resources have currently been granted until 12/2025.