Štěpánek, P., 2026. GriSt: a high-resolution gridded database of meteorological elements for the Czech Republic. In: Tolasz, R., Šustková, V. (Ed.), Proceedings of the 2nd PERUN Project Conference (TA ČR, SS02030040). CHMI, Praha, pp. 171–177, ISBN 978-80-7653-095-9. https://doi.org/10.59984/978-80-7653-095-9.21

GriSt: a high-resolution gridded database of meteorological elements for the Czech Republic

(GriSt: gridovaná databáze meteorologických prvků ve vysokém rozlišení pro Českou republiku)

Štěpánek, P., Zahradníček, P., Farda, A.

Petr Štěpánek, CzechGlobe – Global Change Research Institute CAS; Czech Hydrometeorological Institute, Brno Regional Office, stepanek.p@czechglobe.cz

Pavel Zahradníček, CzechGlobe – Global Change Research Institute CAS; Czech Hydrometeorological Institute, Brno Regional Office, pavel.zahradnicek@chmi.cz

Aleš Farda, CzechGlobe – Global Change Research Institute CAS, farda.a@czechglobe.cz

https://doi.org/10.59984/978-80-7653-095-9.21

Abstract: GriSt is a high-resolution gridded dataset of meteorological variables based on station observations measured by the Czech Hydrometeorological Institute (CHMI). It includes mean, maximum and minimum air temperature, precipitation, relative humidity, wind speed and sunshine duration. The input observations undergo comprehensive quality control using the MetQC system. For long-term climatological applications, the quality-controlled series are additionally homogenized and completed where data are missing, providing temporally consistent input data.

The station records are interpolated on a daily basis to a regular grid with a spatial resolution of 500 m using regression kriging. Geographic coordinates, elevation and other terrain characteristics are used as predictors. An additional interpolated residual field is incorporated so that the resulting grids reproduce the observed values at station locations while providing spatially continuous information across the entire domain.

The dataset covers the period from 1961 to the present and is continuously updated. Daily and hourly products are available, together with temporally aggregated layers. Two main spatial variants are produced: GriSt, focused on the Czech Republic, and GriSt Extended, covering the Czech Republic and surrounding regions. Owing to its long temporal coverage, high spatial resolution and consistency across meteorological variables, GriSt is particularly suitable for climate analyses, impact studies and climate-service applications developed within the PERUN project. It can also serve as a reference dataset for the downscaling and bias correction of climate-model simulations.

Keywords: gridded climate dataset – station observations – quality control – homogenization – regression kriging – climate services

Abstrakt: GriSt je gridovaný dataset meteorologických prvků ve vysokém prostorovém rozlišení, který vychází ze staničních pozorování měřených Českým hydrometeorologickým ústavem (ČHMÚ). Zahrnuje průměrnou, maximální a minimální teplotu vzduchu, srážkové úhrny, relativní vlhkost, rychlost větru a délku slunečního svitu. Vstupní pozorování procházejí komplexní kontrolou kvality pomocí systému MetQC. Pro dlouhodobé klimatologické aplikace jsou časové řady navíc homogenizovány a doplněny v místech chybějících údajů, čímž vzniká časově konzistentní vstupní datový soubor.

Staniční záznamy jsou v denním kroku interpolovány do pravidelné sítě s prostorovým rozlišením 500 m metodou regresního krigingu. Jako prediktory jsou využívány zeměpisné souřadnice, nadmořská výška a další charakteristiky reliéfu. Do výsledku je zahrnuto také interpolované pole reziduí, díky němuž gridy reprodukují pozorované hodnoty v polohách stanic a současně poskytují prostorově spojitou informaci pro celé zájmové území.

Dataset pokrývá období od roku 1961 do současnosti a je průběžně aktualizován. K dispozici jsou produkty v denním a hodinovém kroku i časově agregované vrstvy. Vytvářeny jsou dvě hlavní prostorové varianty: GriSt, zaměřený na území České republiky, a GriSt Extended, pokrývající Českou republiku a okolní regiony. Díky dlouhému časovému pokrytí, vysokému prostorovému rozlišení a vzájemné konzistenci meteorologických prvků je GriSt zvláště vhodný pro klimatologické analýzy, impaktové studie a klimatické služby rozvíjené v rámci projektu PERUN. Dataset může rovněž sloužit jako referenční podklad pro zpřesňování a korekci systematických chyb klimatických simulací.

Klíčová slova: gridovaný klimatický dataset – staniční pozorování – kontrola kvality – homogenizace – regresní kriging – klimatické služby

Introduction

Meteorological observations are irregularly distributed in space, whereas many environmental and impact applications require spatially continuous fields on a regular grid. Gridded datasets make it possible to estimate conditions away from observing stations, calculate area means, provide complete inputs for impact models and compare observations with climate-model output. Their value, however, depends not only on spatial resolution but also on the consistency of the station series and on a transparent description of how observations are transformed into raster fields.

GriSt was developed as a common observational basis for climate analyses and climate services in the Czech Republic. It links three processing levels used in operational and research practice: quality control of station observations, preparation of temporally consistent series where required, and spatial interpolation at a daily or hourly time step. The set of variables – mean, maximum and minimum air temperature, precipitation, relative humidity, wind speed and sunshine duration – was selected so that the fields can be used jointly, including for calculation of reference evapotranspiration according to the FAO Penman–Monteith approach (Allen et al. 1998). This type of internally consistent multivariable input is used, for example, in operational agrometeorological applications and as input to soil-water-balance models such as SoilClim (Hlavinka et al. 2011).

The purpose of this proceedings contribution is practical documentation for users of the dataset. It summarizes the established production chain and the available product variants, explains their different purposes, and illustrates their use in climate services and impact modelling. GriSt is well suited to applications developed within PERUN, but PERUN represents only one of several contexts in which the dataset is used. Detailed formulation and validation of newer interpolation developments are outside the scope of this contribution and will be reported separately.

Input data and dataset variants

The principal inputs are station observations measured by the Czech Hydrometeorological Institute and, for the wider-domain product, by meteorological services in surrounding countries. The core variables are mean, maximum and minimum air temperature, precipitation totals, relative humidity, wind speed and sunshine duration. Together they provide the meteorological inputs required for calculation of reference evapotranspiration and for coupled impact models. The standard temporal basis is daily; operational products are also prepared at an hourly step for the principal meteorological variables and derived evapotranspiration products. The main characteristics of the two product concepts are summarized in Tab. 1.

Two complementary raster concepts are produced. The earlier Czech-domain GriSt product is based primarily on technical series: quality-controlled, homogenized and completed records with long and temporally consistent coverage. Because homogenization requires sufficiently long records and some stations must be excluded during preparation of the technical series, this variant uses a smaller station network. Its principal advantage is temporal consistency, which is essential for climatology, trend analysis and historical comparison. By contrast, GriSt Extended uses all available quality-controlled measurements, including shorter records, and covers the Czech Republic together with surrounding regions. The denser input network generally provides better spatial representation, especially in mountains, near national borders and across transboundary catchments. This is the currently maintained concept used for operational products and spatially demanding applications, particularly in hydrology. The distinction and recommended uses are listed in Tab. 1.

Table 1. Main characteristics of the GriSt products described in this contribution.

Characteristic	**GriSt**	**GriSt Extended**
Spatial domain	Czech Republic (operational calculation may use a wider Czech–Slovak domain with a Czech cut-out)	Czech Republic and surrounding regions; wider variants also available
Spatial resolution	500 m	500 m; a 3 km version covering a wider Central European domain is also produced
Time coverage	1961–present; selected longer reconstructions exist	1961–present
Time step	Daily; operational daily and hourly products for the principal meteorological variables and evapotranspiration	Daily; operational updates according to data availability
Primary purpose	Long-term climatology, trend assessment, impact studies, climate services and scenario reference fields	High spatial representativeness, border regions, mountainous terrain and complete transboundary catchments
Input concept	Technical series: quality controlled, homogenized and completed; fewer stations with sufficiently long records	All available quality-controlled observations, including shorter series; denser station network

Quality control and homogenization

Raw station data may contain erroneous, repeated, duplicated or missing values. The operational quality-control procedure is based on the ProClimDB software (Štěpánek 2010), later reimplemented in R as MetQC. Because no single test reliably identifies all types of problems, MetQC combines temporal and spatial information. Tested values are compared with neighbouring stations, first-difference series and expected values estimated from the surrounding network. Neighbours can be selected by distance and correlation, and their values can be standardized to the elevation of the tested station before comparison.

The automated procedure assigns flags and produces tables, graphs and maps for expert review. In the operational convention, values can be classified as valid, erroneous, suspect, repeated, duplicated or missing. This combination of automated detection and visual diagnostics is important because station density, local topography and the behaviour of individual meteorological elements differ substantially.

Long climatological series used for the technical-series variant are additionally tested for non-climatic discontinuities caused, for example, by station relocation, changes in instruments or observation practice. Detection is performed on monthly, seasonal and annual series using relative comparisons with neighbouring stations and several statistical tests, including the Standard Normal Homogeneity Test, the Maronna–Yohai bivariate test, and the t-test, among others (Štěpánek et al. 2011a). Confirmed breaks are corrected at the daily scale. The Distribution Adjusted by Percentiles method (Squintu et al. 2020) compares the empirical distributions of candidate and reference series before and after a break and applies percentile-dependent adjustments. Missing values are filled after homogenization so that

estimates are based on series not affected by artificial shifts (Štěpánek et al. 2013). In the all-station GriSt Extended concept, observations undergo quality control but shorter records are retained to maximize spatial information.

Spatial interpolation

Each day and meteorological variable are processed separately. The established GriSt workflow uses regression kriging, in which the large-scale spatial pattern is described by a regression model and the remaining local component is interpolated from station residuals (Hengl et al. 2007). Geographical coordinates, elevation and other terrain characteristics can be used as predictors; only predictors that contribute to the specific interpolation are retained. Particular attention is paid to the semivariogram used for the residual component.

The resulting field is calculated on a regular grid with a nominal spatial resolution of 500 m. A final residual layer is added to the regression-kriging estimate. This step is operationally important: it makes the raster reproduce the observed station values at station locations, while the regression and geostatistical components provide a continuous and topographically structured field between stations. The approach follows the general interpolation framework used for technical series and quality control described by Štěpánek et al. (2011b), extended from point estimates to high-resolution raster production. Examples of daily minimum-temperature fields and the locations of the input stations are shown in Fig. 1.

The interpolation is performed separately for each day, both for the historical period and for operational updates. This allows the relationships between the meteorological variable and terrain to vary from day to day, rather than imposing a single climatological lapse rate. In this respect, the approach differs from methods such as MISH, developed at the Hungarian Meteorological Service, in which interpolation parameters are derived in advance from long-term homogenized series and modelled climate-statistical relationships (Szentimrey and Bihari 2007). The method is suitable for producing both long climatological records and operational maps. Spatial interpolation uncertainty nevertheless depends on station density, data availability, terrain complexity and the variable considered; this is one reason why the all-station GriSt Extended product is advantageous in mountainous regions and for hydrological applications.

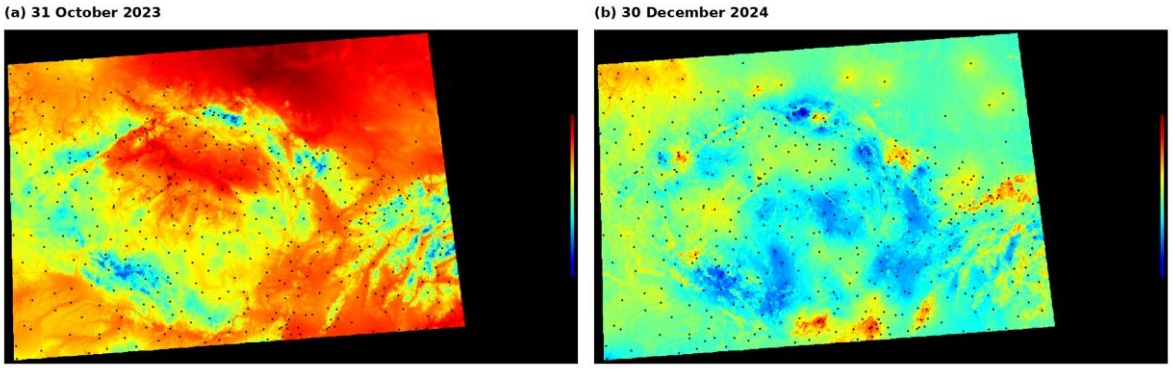


Fig. 1. Examples of daily minimum-air-temperature fields from GriSt Extended for the Czech Republic and surrounding areas. Black dots indicate the input stations. The displayed domain corresponds to the area used for the validation of ALADIN-Climate/CZ. The examples illustrate the topographically structured spatial detail and the continuous coverage between stations.

Products and applications

Daily raster layers can be aggregated to monthly, seasonal, annual and multi-annual products or sampled at arbitrary points and polygons. The GriSt processing chain provides the observational basis for several established climate-service and decision-support systems. In InterSucho.cz, the historical GriSt record is used to calibrate the SoilClim soil-water-balance model and to place current soil-moisture conditions

in a long-term climatological context. GriSt data are also supplied to the HAMR drought and water-availability assessment system, where they support the linked soil-water-balance and hydrological modelling chain. The same GriSt framework also supports FireRisk.cz and AgroRisk.cz, where historical data are used for calibration and reference climatology, while newly generated daily or hourly fields are supplied continuously for operational public services. These applications demonstrate the practical value of a long, internally consistent dataset that links historical calibration with routine real-time operation and allows current conditions to be expressed as anomalies or percentiles relative to the historical climate.

GriSt is also used as the observational reference for transforming climate-model simulations to a high-resolution grid. In this role, the dataset provides the baseline spatial structure and observed distributions to which modelled climate-change information can be added or against which model output can be bias-corrected. This principle is applied in CzechGlobe climate-information products, including ClimRisk, and in outputs prepared for PERUN. The jointly developed ClimateProofing.cz framework combines climate information originating from ClimRisk and PERUN into a user-oriented product for climate-risk and resilience assessment. Further details of the PERUN scenario preparation and its link to climate proofing are given in the companion contribution by Štěpánek et al. (2026).

PERUN is therefore one important application context rather than the sole purpose of the dataset. Where GriSt has been incorporated into PERUN impact studies, it provides a consistent observational reference and a common set of meteorological inputs for evapotranspiration, soil-water balance, drought, agricultural, forestry and hydrological calculations. The distinction between GriSt and GriSt Extended remains important: the technical-series product prioritizes temporal homogeneity, whereas the all-station extended product prioritizes the densest possible spatial representation of the wider region (Tab. 1).

Limitations and interpretation

A 500 m grid should not be interpreted as an observation every 500 m. It is a spatial estimate constrained by the available station network, the selected predictors and the interpolation model. Fine-scale patterns visible in the raster mainly express terrain-related structure represented by the predictors; local effects not observed by the station network cannot be reconstructed reliably. Accuracy is generally higher where stations are dense and where the relationship between the variable and terrain is stable. The denser all-station network of GriSt Extended improves spatial representation but does not replace the temporal advantages of homogenized technical series.

For each daily or hourly raster, a set of diagnostic outputs is generated alongside the final field. These include summary statistics such as RMSE and other measures of agreement with station observations, as well as graphical outputs showing the stations used in the interpolation and the spatial structure of the resulting field. Additional processing metadata and intermediate information are also stored. Together, these diagnostics support retrospective quality assessment, comparison among individual maps and variables, and a more informed evaluation of spatial uncertainty.

The exact network and processing choices differ among variables. Precipitation is observed at many more sites than the full set of climatological variables, but it is also spatially more intermittent. Wind, humidity and sunshine duration have fewer long-term records and are more sensitive to changes in instrumentation or exposure. Metadata and expert evaluation therefore remain important even in an extensively automated workflow.

The present contribution describes the GriSt production workflow that has been developed and used operationally for approximately 15 years. During the past year, the methodology has undergone substantial further development, particularly in the treatment of residual interpolation and agreement with station observations. These recent developments are outside the scope of this proceedings

contribution and will be presented in a follow-up paper in an international journal. The resulting methodology is also planned for implementation in an R package.

Conclusion

GriSt converts quality-controlled station observations into spatially continuous daily and operational hourly fields suitable for climate analysis, impact modelling and climate services. Its main strengths are a long historical record, a 500 m grid, day-specific regression-kriging models, agreement with observations at station locations and the availability of a mutually consistent set of meteorological variables needed for reference evapotranspiration and soil-water-balance calculations.

The two raster concepts serve different but complementary needs. GriSt based on technical series provides temporal consistency for climatology and historical comparison, while GriSt Extended retains all available quality-controlled observations to improve spatial representation, especially in mountains and transboundary catchments. Both concepts support operational services and climate-impact applications; within PERUN they provide suitable reference fields and model inputs, while the same data chain also underlies broader products such as InterSucho.cz, FireRisk.cz, AgroRisk.cz, ClimRisk and ClimateProofing.cz.

Acknowledgments

The research was partially funded by the Technology Agency of the Czech Republic (grant no. SS02030040, PERUN). The preparation and processing of the gridded meteorological datasets for the present climate also build on methodological and data outputs developed within the SustES project (CZ.02.1.01/0.0/0.0/16_019/0000797), the AdAgriF project (Advanced Methods of Greenhouse Gas Emission Reduction and Sequestration in Agriculture and Forest Landscapes for Climate Change Mitigation; CZ.02.01.01/00/22_008/0004635), and the ACECE project (Czech Science Foundation, grant no. 24-14581L).

References

Allen, R. G., Pereira, L. S., Raes, D., Smith, M., 1998. Crop evapotranspiration: Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. Food and Agriculture Organization of the United Nations, Rome, 300 pp. ISBN 92-5-104219-5.

Hengl, T., Heuvelink, G. B. M., Rossiter, D. G., 2007. About regression-kriging: From equations to case studies. *Computers & Geosciences*, **33**, pp. 1301–1315. https://doi.org/10.1016/j.cageo.2007.05.001

Hlavinka, P., Trnka, M., Balek, J., Semerádová, D., Hayes, M., et al., 2011. Development and evaluation of the SoilClim model for water balance and soil climate estimates. *Agricultural Water Management*, **98**, pp. 1249–1261. https://doi.org/10.1016/j.agwat.2011.03.011

Squintu, A. A., Schrier, G., Štěpánek, P., Zahradníček, P., Klein Tank, A., 2020. Comparison of homogenization methods for daily temperature series against an observation-based benchmark dataset. *Theoretical and Applied Climatology*, **140**, pp. 285–301, https://doi.org/10.1007/s00704-019-03018-0

Szentimrey, T., Bihari, Z., 2007. Mathematical background of the spatial interpolation methods and the software MISH (Meteorological Interpolation based on Surface Homogenized Data Basis). In: Szalai, S., Bihari, Z., Szentimrey, T., Lakatos, M. (Eds.), Proceedings of the Conference on Spatial Interpolation in Climatology and Meteorology, COST Action 719, COST Office, Luxembourg, pp. 17–28. ISBN 92-898-0033-X.

Štěpánek, P., 2010. ProClimDB – Software for Processing Climatological Datasets. Czech Hydrometeorological Institute, Regional Office Brno. Online documentation: climahom.eu.

Štěpánek, P., Zahradníček, P., Brázdil, R., Tolasz, R., 2011a. Metodologie kontroly a homogenizace časových řad v klimatologii. Czech Hydrometeorological Institute, Prague, 118 pp., ISBN 978-80-86690-97-1.

Štěpánek, P., Zahradníček, P., Huth, R., 2011b. Interpolation techniques used for data quality control and calculation of technical series: an example of a Central European daily time series. *Időjárás*, **115**, pp. 87–98.

Štěpánek, P., Zahradníček, P., Farda, A., 2013. Experiences with data quality control and homogenization of daily records of various meteorological elements in the Czech Republic in the period 1961–2010. *Időjárás*, **117**, pp. 123–141.

Štěpánek, P., Farda, A., Meitner, J., Zahradníček, P., Fischer, M., et al., 2026. Preparation of climate scenarios for impact studies in the PERUN project: combining selected GCMs and the regional ALADIN model. In: Tolasz, R., Šustková, V. (Eds.), Proceedings of the 2nd PERUN Project Conference (TA ČR, SS02030040). CHMI, Praha, pp. 75–82, ISBN 978-80-7653-095-9. https://doi.org/10.59984/978-80-7653-095-9.10