Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

ZAMG

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

DACH recommendations on uncertainties and interpretation of grid point values of station-based grid data

Based on analyses and considerations by Christoph Frei (MeteoSwiss) on the occasion of the DACH - Workshop of 25 August 2021

Station-based grid data are based on station data that have been spatially interpolated using (geo-) statistical methods. These are used wherever information is needed between stations with meteorological measurements. Grid data are also needed as a reference for the evaluation and bias adjustment of climate projections. However, users need to be aware of the limitations of such grid data in order to apply them correctly. In the following, general limitations in station-based grid data are described that arise largely independently of the methodology used and must be taken into account in a professional application of these data. Grid data from satellite and radar measurements or from reanalyses are not discussed here, as they contain other limitations.

Measurement uncertainty

One source of uncertainty in climate grid data sets is the measurement errors of the collected station data. These errors usually contain a systematic and a random component. They vary in magnitude depending on the parameter, measurement system, meteorological and geographical conditions. Precipitation measurements with conventional pluviometers, for example, systematically underestimate the real precipitation. In the DACH region, this systematic bias ranges from about 5% in the lowlands in summer to more than 50% at wind-exposed locations above 1500 m in winter (Sevruk 1985; Richter 1995, Kochendorfer et al. 2017). In addition, there are random errors, which are very large in relative terms for small amounts of precipitation. Measurement errors, systematic and random, potentially significantly affect the accuracy of grid datasets. The contributions of random measurement errors can be quantified via cross-validation, and their magnitude is usually described in documentations of the datasets. Users are encouraged to consult the relevant literature to familiarize themselves with the magnitudes and characteristics of measurement errors, and to assess their effects on the specific application. The spatial propagation of measurement errors is currently not considered in the grid data sets. It would have to be modelled geo-statistically analogous to the propagation of the measured values themselves. For the spatial estimation of the systematic precipitation deficit, for example, temperature and wind fields would have to be explicitly taken into account.

Conditional Bias

The spatial interpolation of measurements at stations is inevitably associated with uncertainties, because the spatial distribution of a climate parameter is only described incompletely from the available measurements. On the one hand, this uncertainty manifests itself in random errors of the individual estimates at grid points. In addition, it manifests itself in "conditional biases", systematic errors depending on the location in the corresponding distribution. In general, the low extremes are overestimated and the high extremes underestimated. The interpolation uncertainty, although random in single cases, causes systematic errors (biases) in the climatological distribution and thus biases in common climate indices. In a daily precipitation data set, the "wet-day frequency" is overestimated, as well as the "wet-day intensity" and, to a particular extent, the high quantiles (heavy precipitation) are underestimated. The effects of interpolation uncertainty can also affect the mean values: When creating a background field, for example for seasonal temperature averages, the estimates are underestimated at abnormally warm locations (e.g. in cities in summer) and overestimated at abnormally cold locations (e.g. in deep Alpine valleys in winter). More generally, the actual spatial variability (roughness) of a parameter in a (deterministic) grid data set is underestimated.

The phenomenon of conditional bias (often referred to as "smoothing effect" or "representativity error") is related to interpolation uncertainty in combination with the goal of minimizing the error of interpolation (optimal prediction). Error minimization is a fundamental principle of statistical prediction, and inherent in all methods of producing grid datasets, either explicitly in statistical methods (e.g. regression, kriging, general additive models) or implicitly in heuristic methods (e.g. in calibration via cross-validation, weighting methods). Conditional biases are therefore a very general limitation of today's grid datasets, independent of the method used for production. Avoiding them requires a fundamental paradigm shift from classical deterministic estimation to possible realizations based on stochastic simulation (e.g. Cornes et al. 2018, Frei & Isotta 2019). However, more methodological research and time for development are needed before this paradigm shift can be fully implemented.

The magnitude of conditional biases is directly dependent on the magnitude of interpolation uncertainty. That is, they are
(a) smaller in precipitation in winter than in summer (spatial scale of precipitation systems),
(b) smaller in regions and in time periods with dense observation networks (amount of available information),
(c) smaller for mean values over larger areas than for point values at one location,
(d) smaller for monthly mean values than for daily values,
(e) smaller in the middle of the distribution than for extremes.

These dependencies can also cause inconsistencies in the data sets. For example, if the station density varies over time, artificial trends arise because the conditional bias changes over time.

Unfortunately, the conditional biases present in a grid dataset are usually poorly documented, which is related to the complex dependencies, but also to the lack of awareness in the grid data community itself. This makes it difficult for a user to get an informed picture of this constraint on their specific application. A very valuable and general recommendation to avoid the effects of conditional bias is to interpret data sets on larger spatial scales. There, the uncertainties and thus the conditional biases are smaller. Information on the critical space scale (effective resolution) can be found in data set documentations (see also details in the following section).

<u>Effective resolution</u>

The mesh size of modern regional grid datasets is in the range of 1 km. The typical spacing of measuring stations, for current measuring networks, is 15 km for precipitation and 30 km for temperature. For measurement networks prior to 1960, this distance is even significantly wider. The spatial distribution in a 1-km grid dataset is therefore largely based on the relationship of the parameter to topographic characteristics. If these relationships are unclear, e.g. generally for precipitation or temperature during complex weather events (“Föhn”, cold-pools in winter, topographically influenced boundary layers during high pressure in summer), then estimates at this scale are associated with considerable uncertainty. This manifests itself in excessive (unrealistic) smoothness and in systematic errors in the statistical characteristics (conditional bias). The estimated values at a grid point do not have the climate (statistical characteristics) of point values or pixel means at this scale, but the climate of an area mean over a larger area (e.g. Hiebl and Frei, 2017). In most grid datasets available today, the scale of the resolved features is much coarser than the mesh size, often a multiple of it. In that case, the mesh size used in the gridding is just a technical quantity but bares little information about the space scales actually resolved. The misunderstanding that "finer mesh size equals better resolution" often leads to unjustified interpretations of the data sets by users.

The "effective resolution" is the scale limit beyond which a user can expect the statistics calculated from the grid data (indices, quantiles, spatial variance) to be realistic, i.e. largely free of conditional biases. For variables with poor relationships to topographic predictors (e.g. precipitation), the effective resolution is of the order of the typical station spacing. For small-scale summer precipitation even a multiple of this. For variables where there are auxiliary quantities providing a high degree of additional information for the interpolation (e.g. the topography for temperature), the effective resolution is also finer than the station spacing. This is also true for flat regions and for variables that vary little in space. The km-scale mesh size of most grid datasets in the DACH area is the lowest common denominator for most applications. It simplifies their technical implementation, e.g. in hydrological applications that often run on the km-scale or for the validation of regional climate models that currently exist on 2-20 km mesh sizes.

It is important that users consider the limited effective resolution of the datasets and assess what effects this may have on the application beforehand. Spatial modelling of non-linear processes at fine spatial scales, for example, is highly susceptible to errors in spatial variance and to underestimation of extremes, which are common in today's climate datasets. Similarly, there is a risk that the validation of extremes (e.g. high precipitation quantiles) in km-scale climate models will be affected by the conditional biases in the reference datasets. Here, a prior spatial aggregation to at least the mean station spacing is recommended. The station spacing is parameter-dependent in most measurement networks. Aggregation to sub-catchments is also useful and facilitates e.g. the water balance analysis. Finally, applications with high requirements on long-term consistency are at risk from temporal changes in the density of station data used in the datasets. Again, what can help further here is spatial aggregation to scales that can be reliably mapped from the available measurement networks over the entire period.

Observational datasets (and applications based on them, such as bias-adjusted high-resolution climate projections) are often seen as a worry-free interface between climatology and climate impact research. It is tempting to look at these datasets as if there is a measurement at every km point. This notion is wrong. Although such datasets offer significant advantages in application, their proper use requires the user to address details of their construction, especially the density of the measurement networks used, and to assess the risks of constraints on their application. We encourage users to evaluate their use case together with the creators of the grid datasets.

References

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., & Jones, P. D. (2018). An ensemble version of the E-OBS Temperature and Precipitation Datasets. Journal of Geophysical Research: Atmospheres, 9391-9409. https://doi.org/10.1029/2017JD028200

Frei, C., and Isotta, F. A. (2019). Ensemble spatial precipitation analysis from rain-gauge data: Methodology and application in the European Alps. J. Geophys. Res. Atmos, 124. https://doi.org/10.1029/2018JD030004

Hiebl, J., and Frei, C. (2018). Daily precipitation grids for Austria since 1961 - development and evaluation of spatial dataset for hydroclimatic monitoring and modelling. Theor. Appl. Climatol., 132, 327-345. https://doi.org/10.1007/s00704-017-2093-x

Kochendorfer, J., Nitu, R., Wolff, M., Mekis, E., Rasmussen, R., Baker, B., Earle, M. E., Reverdin, A., Wong, K., Smith, C. D., Yang, D., Roulet, Y.-.A., Buisan, S., Laine, T., Lee, G., Aceituno, J. L. C., Alastrué, J., Isaksen, K., Meyers, T., Brækkan, R., Landolt, S., Jachcik, A., and Poikonen A. (2017). Analysis of single-age-shielded and unshielded measurements of mixed and solid precipitation from WMO-SPICE, Hydrol. Earth Syst. Sci., 21, 3525-3542, https://doi.org/10.5194/hess-21-3525-2017

Richter, D. (1995). Results of methodical investigations on the correction of the systematic measurement error of the Hellmann precipitation gauge. Report of the German Weather Service, 194, 93 pp.

Sevruk, B. (1985). Systematic precipitation measurement error in Switzerland. In: Precipitation in Switzerland. (Ed. Sevruk B.), Beiträge zur Geologie der Schweiz - Hydrologie, 31.