Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-51-AC1, 2016

Geoscientific Model Development Discussions

Interactive comment on "An open-source MEteoroLOgical observation time series DISaggregation Tool (MELODIST v0.1.0)" *by* Kristian Förster et al.

Kristian Förster et al.

kristian.foerster@uibk.ac.at

Received and published: 11 May 2016

We would like to thank Dr. Ina Pohle for her detailed review of our manuscript and for her constructive comments and suggestions. The points raised in this review are highly appreciated and will help us to improve our manuscript. Please find our detailed response below.

(Reviewer's comments are in italics)

General comments:

The manuscript by Förster et al. presents the software package

MELODIST, a frame- work of state of the art methods for disaggregating meteorological time series. The methods included comprise deterministic and stochastic approaches with several options to choose for the individual meteorological variables. The disaggregation methods are described concisely with adequate reference to the relevant literature. The general applicability of the disaggregation methods is assessed by comparisons between observed hourly data and disaggregated hourly data based on daily variables. Therefore, five stations in contrasting climates have been chosen. The model code itself is well documented, the software package is easy to apply and modify and thus has high potential on being used e.g. by hydrologist who require hourly input data for models. The manuscript is well structured and written. The methods are clearly documented and critically assessed both with reference to the literature and by own analyses of the authors. The conclusions are well supported by the results. I recommend the article for publication after minor revision for the following issues:

- Introduction: motivate the need of a disaggregation to hourly data more directly - for which purposes are data in hourly resolution needed (give examples)

Based on available literature, we will add some examples and applications for which disaggregation methods are required. We will add one additional paragraph in the introduction:

"In contrast, hourly meteorological time series are required for numerous applications in geoscientific modelling. Typical applications in hydrology include both derived flood frequency analyses (e.g., Haberlandt and Radtke, 2014) and water balance simulations (Waichler and Wigmosta, 2003). In ecological modelling, sub-daily meteorological data are required for, e.g., the estimation of epidemic dynamics of plant fungal pathogens (Bregaglio et al., 2010)"

Thank you for pointing us in this direction.

> *- Introduction: while the relevant literature concerning disaggregation methods is addressed, reference to other tools / software packages for disaggregation of single meteorological variables (e.g. HyetosR) is missing*

We will refer to the HyetosR package which we were not aware of. We appreciate this hint! The revised version of the manuscript will include a reference to this software (page two, first bullet point):

"For instance, the rainfall disaggregation package "HyetosR" (Kossieris et al, 2012, ITIA, 2016) provides an extensive parameter estimation methodology which is based on observed time series."

> *- Results: It is of interest, whether the distributions of the hourly data are preserved. Table 2 gives only mean values and standard deviations. Do the parameters of the distribution functions differ between observed and disaggregated hourly values?"*

We agree that the comparison of mean values and standard deviations only gives a simplified review of the distribution of these values. This is a valid point which we have discussed intensively. The variables addressed in the manuscript have different distributions which is why it is not possible to fit one single type of distribution function. For instance, temperature might be represented by a normal distribution for many sites, whereas precipitation is characterised by a lower limit of zero and asymmetry. To best possibly address the need for distributions and to keep the manuscript concise without extensive additions regarding theoretical distribution functions and parameter estimation, we decided to add an additional figure to the revised version of the manuscript including histograms of both observed and disaggregated values for each variable and each station.

> *- Results: On which basis have the times and locations for the result figures been chosen? Are these the times locations where the disaggregation results fit the observations best? It might be helpful to add performance measures also for the time periods displayed.*

This question seems to refer to the example figures (Fig. 3 to Fig. 7) since only these figures include results of disaggregation methods for selected times and locations. In fact, the example figures for each variable have been randomly selected. They have been designed to show exemplarily how each of the methods work. You are right to say that this information needs to be clarified. In the revised version of the manuscript, we will explain in section 3.1 that the times and locations have been randomly selected. Adding performance measures for each method is a good point as this information would prove helpful. In principle, this is not a problem at all. However, this would require one additional table for each example plot. In our opinion, these additional tables would go beyond the scope of the exemplary type of figures. Therefore, we suggest adding the RMS error for each method to the legend in order to give an idea of model performance for each method for the times displayed in each figure (except for precipitation).

Minor comments:

> *Page 1, line 1: Maybe specify: "Observations of hourly time series" / "Monitoring data in hourly resolution"*

We will rewrite this sentence accordingly: "Meteorological time series with one-hour time step are required in many applications in geoscientific modelling."

> *Page 3, line 24-26: Can be deleted, the reader should be familiar with the difference between deterministic and stochastic approaches.*

We agree that most of the readers should be familiar with these terms. However, since the evaluation of stochastic methods requires multiple runs to perform statistical analyses, we believe that some introductory remarks might improve comprehensibility regarding the study design.

Page 6, line 4: replace "small scale" with "sub-hourly"

Done.

Page 6, line 5: sentence unclear

We agree that this sentence should be improved. We will revise this statement in the following way: "This idea best corresponds to averages of wind speed for a given increment of time (e.g., one hour) rather than instantaneous measurements."

Page 6 line 25: specify distribution (uniform)

This information was missing: "The function rnd is a random number generator which draws random numbers between 0 and 1 from a uniform distribution."

General language comment: check when to use "a" and "an"

We will review and correct the document with respect to the usage of "a" and "an". Thank you.

Page 10 line 18: why is this approach not referred to as "inverse distance weighting"

At present, this method simply transforms the mass curve from one station to another. Distance measures, which might be relevant if more than one highly resolved station is considered, are not considered in this method since the focus of the methods presented is on single sites only. However, a distance-related weighting considering more than one station can be easily applied to this method. This feature is implemented in the already cited IDWP program.

Page 13 line 2: replace "are not reproducable" with "is not reproducible" or: "cannot be reproduced"

Done.

Page 14 line 2 & line 29: these lines are redundant.

Yes. We removed the redundant sentence in line 29.

Page 15 line 3: can you give a ballpark figure on computational costs, e.g. disaggregation of 10 years of temperature data?

Thank you for this suggestion. The following information will be added to the revised version: "disaggregating 5 years of daily precipitation recordings using the cascade model takes less than 4 seconds on a notebook with a 2 GHz i7 CPU)"

Page 15 line 5: give examples here (or in introduction)

As pointed out earlier, we will add some examples in the introduction.

Table 1: Please state whether "data availability" refers to hourly data

Yes, “data availability” refers to hourly data. We will add this information to the caption.

Figure 2: scale of the points – hard to perceive differences

We slightly increased the dot size in order to improve perception. However, the difference between the two stations in Central Europe is small (De Bilt: 803 mm, Obergurgl: 851 mm).

References

Bregaglio, S., Donatelli, M., Confalonieri, R., Acutis, M., and Orlandini, S.: An integrated evaluation of thirteen modelling solutions for the generation of hourly values of air relative humidity, Theor. Appl. Climatology, 102, 429–438, doi:10.1007/s00704-010-0274-y, 2010.

Haberlandt, U. and Radtke, I.: Hydrological model calibration for derived flood frequency analysis using stochastic rainfall and probability distributions of peak flows, Hydrol. Earth Syst. Sci., 18, 353–365, doi:10.5194/hess-18-353-2014, 2014.

ITIA: HyetosR: An R package for temporal stochastic simulation of rainfall at fine time scales, http://www.itia.ntua.gr/en/docinfo/1200/, accessed on 09 Mai 2016, 2016.

Kossieris, P., Koutsoyiannis, D., Onof, C., Tyralis, H., and Efstratiadis, A.: HyetosR: An R package for temporal stochastic simulation of rainfall at fine time scales, European Geosciences Union General Assembly, Geophysical Research Abstracts, 14, 11 788, 2012.

Waichler, S. R. and Wigmosta, M. S.: Development of Hourly Meteorological Values From Daily Data and Significance to Hydrological Modeling at H. J. Andrews Experimental Forest, J. Hydrometeorol., 4, 251–263, doi:10.1175/1525-7541(2003)4<251:dohmvf>2.0.co;2, 2003.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-51, 2016.