Interactive comment on "The design, deployment and testing of Kriging models in GEOframe" by Marialaura Bancheri et al.

General Comments

Overall quality of the paper

Firstly, thanks the authors for this manuscript that delivering a new software product to science world is important for diversity and productiveness. In addition, I want to remark that the review is to improve the paper.

The paper presents an interpolation package for climatic and environmental variables. The objectives of the paper are to bring an alternative package which is applicable with good performance. By using Kriging techniques, the modeling spatial distribution of hourly temperature and rainfall in an Alpine catchment located in north of Italy and leave-one-out cross validation of the model are the scope of study. And it is suitable in terms of themes of GMD journal.

The paper is generally clear and well-written; the methods, assumptions and results are rationally structured and presented, but I think the language has some issues in the sense of fluency. In the manuscript, the structure of parenthetical sentence is overutilized. Instead of this structure, simple and short sentences with proper conjunctions must be preferred for fluency and explicity of text. Clarity of the manuscript can be enhanced a bit more by using clear, fluent and open expressions. In addition, I do not approve the separation of word with a hyphen at the end of the line.

In my opinion, the paper is at good level in terms of significance, quality and reproducibility of scientific and presentation quality has a fair level. I think it could be accepted after revision of the manuscript.

<u>Advantages of the manuscript:</u>

- A new open source software product
- Comparison of 4 types of Kriging and 11 semivariogram models
- Comparison with R gstat in terms of multi-model approach

<u>Disadvantages of the manuscript:</u>

- The lack of multi-site (which have different geo-climatic characteristics) comparison
- Independence of validation dataset
- The insufficiency of rainfall data used (the modeling of a single storm event does not provide insight into the good performance)

Specific Comments

Major-minor issues (criticisms), review based sections, clarification of figures and tables, evaluation of references, questions

Against the possibility of future modifications, the title must include version number (v. 0.9.8) of the model.

In the abstract, time scale (hourly) of used data could be mentioned. Also, "2008 year hourly temperature", "rainfall storm event (11 h) in 2008" and "spatio-temporal distribution of climatic and environmental variables" expressions could be included in the abstract. The importance of the study for hydro-meteorological studies can be processed in some more detail in the abstract and the introduction.

A bit more of recent references on Kriging as itemized below can be given in introduction section. Moreover, references on Kriging open source algorithms can be added and it is provided the differences of between this paper and the current literature.

- Saghafian and Bondarabadi, 2008. "Validity of regional rainfall spatial distribution methods in mountainous areas".
- Di Piazza et al., 2011. "Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy".
- Bostan et al., 2012. "Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey".
- Wang et al., 2014. "Comparison of interpolation methods for estimating spatial distribution of precipitation in Ontario, Canada".
- Adhikary et al., 2016. "Genetic programming-based ordinary kriging for spatial interpolation of rainfall".
- Jin et al., 2016. "Estimating Loess Plateau average annual precipitation with multiple lineer regression kriging and geographically weighted regression kriging".

As a reference, the study of titled "A statistical approach to some mine valuations and allied problems at the Witwatersrand" by Krige (1951) which is the originator of Kriging method, can be quoted within the framework of respect for labor. As you know, Kriging method is named after him.

Kriging methods are geostatistical methods that consider the spatial autocorrelation of sample data. Due to based on stochastic process, Kriging assumes that sample data is stationary and normally distributed. And the data needs the de-trending and de-clustering processes. In this context, fit of normal distribution and autocorrelation function are analyzed. Normal distribution and spatial autocorrelation of residuals are even investigated. Within this context, detailed information about the processes of outlier treatment, normalization, detrending, declustering, optimization of semivariogram model and neighborhood search must be provided.

In introduction, some open source geostatistical tools are given. In addition to them, software products such as;

- GRASS (grass.osgeo.org)
- GSLIB (www.gslib.com)
- Surfpack (dakota.sandia.gov/content/surfpack)
- jk3d (sourceforge.net/projects/jk3d)
- Map Window (mapwindow.org)
- SGeMS (sgems.sourceforge.net)
- GitHub Py Krige (github.com/bsmurphy/PyKrige)
- QGIS (qgis.org)
- uDig (udig.refractions.net)
- GeoDa (spatial.uchicago.edu/software)
- ArcGIS (arcgis.com) -even it can be counted from the point of view of integrating open source codes-

can be indicated.

The study contains continuous modeling of temperature and event based modeling of rainfall by SIK package. Temperature outputs are also compared with R gstat. Although temperature is more stabile climatic variable in according to the spatial distribution, precipitation is more changeable that has more the spatial variance. Thus, I think continuous modeling and comparison with different package results of precipitation make greater contribution to this paper.

It is stated that plug-in to a hydrological model of the SIK model is easy. Entering more details about this will be good. Furthermore, the SIK results can be inputted into a hydrological model and the performance of hydrological model can be assessed.

In some parts of the paper, it is specified that the data of 97 meteorological stations is used, but there are 93 stations in Appendix D. More than this, there are 81 stations in Figure 4; 57 stations in Figure 8. In short, an inconsistency in number of stations is in existence. When examining Figure 8, I have doubts about a manipulation for high performance. Please enlighten me about this matter.

To get to the clarification of figures, I think Figure 3 and 5 must be made smaller, due to they are nonsensically very big. In Figure 4, basin boundary and river network should be shown for more understanding of the study area. In the legend of Fig. 4, the ranges of elevation must be like "210.1-850, 850-1500, 1500-2100, 2100-2750, 2750-3388" with round numbers as an example. For Figure 5 and 9, please give an explanation of X and Y axis in the relevant places of the text. In the legend of Figure 5, please illustrate experimental semivariance as only black dot without line and the others as only line without dot. In Figure 6, size of dots can be smaller in terms of clearly view of each semivariogram model. A scale bar (km) must be added on Figure 7 where the intervals of temperature must be as "-6-0, 0-5, 5-10, 10-15, 15-19" with round values. Likewise, A scale bar and north arrow must be placed in Figure 8 which has two same legends of elevation, one of them must be removed. In comparison of spatial data, same color scale must be used for spatial layers. In this respect, Figure 7 (temperature scale) and 8 (elevation scale) are good examples. However, sizes of RMSE plots

are not same in Figure 8. In despite of the authors allege the visualization reasons as excuse, I do not accept this situation. For a reasonable comparison, same size scale of RMSE must be used. I suggest the followed intervals for RMSE in both OK and DK output layers of June 2008. "0-1, 1-2, 2-3, 3-6, 6-9, >9 with max value". A conspicuous another issue about Figure 8 is that max RMSE value is 9 in spite of given as 11.95 in the text. The measured rainfall is not shown as black dashed line in the legend of Figure 10. In Figure 11, the color scale of precipitation is changed from white to dark blue. But the colors of precipitation classes are very close to each other, because of this situation the transition of precipitation surface is not distinguished exactly. To remove this problem, selection of a more colorful scale will be suitable. For instance; red, yellow, sea green, light and dark blue, purple. It changes respectively from dry and wet. Two raster layers (DEM and precipitation), both of which contain white color, are overlapped using transparency feature of the layers in the same figure; but this is caused the confusion with regard to understanding of the figure. If the presentation of elevation layer is as contour lines, the confusion dies. The intervals of the precipitation and DEM must be respectively as "2.21-2.7, 2.7-3.2, 3.2, 3.7, 3.7-4.2, 4.2-4.7, 4.7-5.2, 5.2-5.7, 5.7-6.02" and "0-300, 300-2900, 2900-3388". Figure 11 also needs a scale bar. Together with NSE; R^2, RMSE and PBIAS values of DK and similarly all the performance criteria of OK methods can be given in Figure 12. And all, the captions of the figures can be shortened as follows and the required explanations relating to the figures should be stated in the relevant places of the text. In all the figures, writings and titles of chart and axis must be had standard font size.

- Figure 1. Flow chart of modeling process.
- Figure 2. Flow chart of validation process.
- Figure 1. Flow chart of modeling and validation processes. (Figure 1 and 2 can be combined into one)
- Figure 4. Geo-location of study area and position of meteorological stations.
- Figure 5. Fitting of the experimental semivariogram for 15th June 2008 12:00 CET.
- Figure 6. Monthly variation of the NSE index over entire hourly temperature dataset.
- Figure 7. Maps and histograms of spatialized temperature for 15th February 2008 and 15th June 2008.
- Figure 8. Bubble plots of the RMSE obtained using OK and DK.
- Figure 9. Boxplots of the semivariograms of the precipitation event of 29th and 30th June 2008.
- Figure 10. Comparison between the four types of Kriging and the measured rainfall.
- Figure 11. Spatial interpolation of the precipitation applying OK and the linear semivariogram model.

A little more information such as slope, land cover, vegetation, hypsometric elevation and river discharge about the study area can be given. Basin characteristics are important in understanding and evaluation of hydro-meteorological modeling in particular type and distribution of precipitation for this study. Info of DEM is not presented in the paper. Please provide source of DEM.

Please give definitions, explanations and formulations of the performance criteria used (NSE, R^2, RMSE, PBIAS) in the paper. For NSE, it is referred to Nash and Sutcliffe (1970).

The headers of the tables can be changed as below because of they are also overlong. The statements about the tables can be presented in the relevant parts of the text not in the table header. It is emphasized that Bessel semivariogram model is the best in according to Table 1. How was this information reached? How was the integration of 4 criteria (NSE, R^2, RMSE, PBIAS) achieved? Logarithmic model results also look pretty good, why was not the logarithmic model in the best 5 of the authors? How were the best 5 determined? In my opinion, the logarithmic model is candidate for the best and Bessel, logarithmic, exponential, linear and spherical models are the best 5 semivariograms. Please clarify this matter. In Table 2, mean of performance criteria values of all the stations can be given for 4 types of Kriging.

- Table 1. Performance results of semivariogram models used
- Table 2. Results in terms of goodness of fit indices between the measured and interpolated rainfall values for two stations

In the manuscript, it is stated that the overall performances of both SIK and gstat tools are good. I don't think gstat has good performance in according to Figure 12. Please research the success limits of NSE values in the literature. In this context, please rewrite the related parts and the last sentence of the abstract. In plain words, I did not expect it to be so different. SIK has almost overwhelming superiority. I have some suspicions in this regard. Please inform about the optimization of gstat, is it performed sensitively? How do you describe this failure and explain it?

In conclusion section, resolution of interpolation grid data (100 m) can be mentioned because of the resolution is an important issue for the hydrological model results.

The following abbreviations and symbols must be added to Appendix B:

- PSO: particle swarm optimization
- OK: ordinary kriging
- LOK: local ordinary kriging
- DK: detrended kriging
- LDK: local detrended kriging
- SIK: spatial interpolation kriging
- OMS3: object modeling system v.3
- MS: modeling solutions
- DP: design patterns
- LOO: leave-one-out
- GIS: geographical information systems
- a.s.l. : above sea level
- CET: central European time
- DEM: digital elevation model
- DOI: digital object identifier

- Γ: matrix of the two point variograms
- Λ: N-uple of the unknown weights
- B: N-uple containing the variograms among the ungauged site and the measured sites

In most of the equations in Appendix C, multiplication is indicated by a dot. I think there is no need to use these dots.

In Appendix D, sorting of the stations by elevation not station ID becomes more meaningful due to the strong relationship between the elevation and temperature or precipitation. Giving of mean and median elevation values of the all stations would be better.

In appendix of code availability, the authors specifies that the documentation of SIK components are available. Is a detailed user manual provided?

The sentence of "Lastly, the authors declare that they have no conflict of interest." should be placed at the end of author contribution appendix.

Technical Corrections

Corrections, typing errors

In some parts of the paper, it is specified that the data of 97 meteorological stations and 10 semivariograms are used, but there are 93 stations in Appendix D and 11 semivariogram models in Figure 5/Table 1.

P1 L7-P8 L14- P11 L8- P20 L18: "97" ---> "93"

P1 L5-P20 L20-P22 L2: "ten" ---> "eleven"

P10 L6-P10 L14-P11 in Fig. 5 caption-P11 L2-P12 in Table 1 header-P20 L2: "10" ---> "11"

P1 L3: "in a hydrological model" ---> "to a hydrological model"

P1 L8: "compared to the results" ---> "compared to the results of temperature"

P1 L11+: "These data, besides being important by themselves, are the natural input to distributed and semi-distributed hydrological models, where their quality and precision affect the accuracy of results." ---> Please give a few references.

P1 L15+-P1 L19-P1 L22-P3 L5+-P5 L5-P7 L17: The references aren't chronologically presented, please chronologize.

P1 L18: "anong" ---> "among"

P1 L20: "shown" ---> "showed"

P2 L25: "(Formetta et al., 2014).In order" ---> "(Formetta et al., 2014). In order". No space after point.

P3 L13-P3 L15-P3 L27: The ordering of in-equation parenthesis is not appropriate. The correct ordering is "{ [( )] }".

P3 L9-P3 L22-P 3 L23: "N-ple" ---> "N-uple"

P3 L26-P5 L21-P7 L11: "e.g." ---> "e.g.,"

P3 L27: ":=" ---> "="

P4 L2-P4 L3: "i.e." ---> "i.e.,"

P4 L19: "are the local" ---> "are respectively the local"

P4 L28: "interpolate a theoretical model for the semivariogram" ---> "fit a theoretical model to the semivariogram"

P4 L30: "produce maps" ---> "produce continuous surface maps"

P5 L7: "Efron and Efron" ---> "Efron"

P5 L10: "produced" ---> "performed"

P5 L16: "granularity whose rational" ---> "detail whose logic"

P5 L17: "that, for the reader convenience are presented in A" ---> "presented in Appendix A for the reader convenience"

P5 L19: "Object Modelling System v3 (OMS3)" ---> "OMS3". Previously presented as open.

P5 L20: "here" ---> "are"

P5 L22: "Modelling Solutions (MS)" ---> "MS". Previously presented as open.

P5 L22: "complicate" ---> "complicated"

P5 L25: "particular cares from the researcher who programs them" ---> "any special effort of the computer programmer of them"

P5 L29+: "in various components, and inserted in MS as alternatives, thus opening the way to compare, inside the same chain of tools" ---> "various components inserted in MS as alternatives. Thus inside the same chain of tools, it enables to compare"

P5 L31: "in Formetta et al. (2014) and in Bancheri (2017)" ---> "Formetta et al. (2014) and Bancheri (2017)"

P6 L12: "Particle swarm" ---> "Particle swarm optimization (PSO) technique"

P6 L16-P11 L7-P14 L2-P17 L2: "show" ---> "shows"

P7 L10: "addition(" ---> " addition (". No space before parenthesis.

P7 L11: "Design Patterns (DPs," ---> "DP". Previously presented as open.

P7 L13-P7 L14-P7 L15: "DPs" ---> "DP"

P8 L13-P20 L11: "modeling solutions" ---> "MS". Previously presented as open.

P9 in Fig. 4-P18 in Fig. 11: "Meteo_stations" ---> "Meteorological stations"

P9 in Fig. 4: "DEM" ---> "DEM (m)"

P18 in Fig. 11: "DEM" ---> "Contour lines (m)". In Fig. 11, contour lines must be used instead of DEM.

P9 L9: "out-layers" ---> "outliers"

P9 L10-P10 L3: "asses" ---> "assess"

P10 L1+: "nor too low or too high" ---> "neither too low nor too high"

P10 L10: "their performances" ---> "performances"

P11 L1: "the correlation coefficient" ---> "the coefficient of determination"

P11 L5: "4 typer of Kriging, ~~OK, LOK, DK, LDK,~~" ---> "4 types of Kriging"

P11 L9: It is specified that the two local cases were performed using the ten closest stations to the interpolation point. Why 10 stations? Has neighborhood analysis ever carried out?

P11 L10: "meaning" ---> "it means"

P11 L12: "exists" ---> "exist"

P12 in Table 1: Please sort the semivariograms by the performance.

P13 in Fig. 6: Please remove expression of "hourly temperature" from the chart titles as it is expressed in the caption.

P14 L2: "7" ---> "Figure 7"

P14 L3: "during" ---> "in"

P15 L1: "alms" ---> "m a.s.l."

P15 L2: "1.83 ^{o}C" ---> "1.83 ^{o}C in the DK method"

P15 in Fig. 8: Please remove expression of "monthly mean temperature" from the map titles. The altitude legend scale has not min and max values, please add these.

P16 in Fig. 9 chart title: "30 h" ---> "11 h"

P16 L2: "953 m" ---> "943 m". According to Appendix D.

P16 L6: "local ordinary kriging" ---> "LOK". Previously presented as open.

P16 L7: "ID 2170" ---> "ID 1270"

P16 L6-P16 L8: It is stated that the LOK computed using the 5 closer stations. Is it determined with manual calibration of the neighborhood? If search of neighbor numbers was performed, it must be specified in the text. If not, explain that how it was determined. In addition to, the number of closest neighbor stations is 10 for temperature and 5 for rainfall. Please interpret this.

P17 L1: "DEM (100 m resolution)" ---> "DEM"

P17 L3: "[mm/h]" ---> "(mm/h)"

P19 L4: "design pattern" ---> "DS". Previously presented as open.

P19 L4: "figure 3" ---> "Figure 3"

P19 L7: "opertion" ---> "operation"

P19 L10: "offers" ---> "offer"

P19 L11+: "the same number of bins and the same cutoff distance" ---> "the same number of bins and cutoff distance"

P19 in Fig. 12 legend: "SI component" ---> "SIK component"

P19 in Fig. 12 caption: "package" ---> "packages"

P19 in Fig. 12: NSE value 1.0 is not found in the figure.

P20 L13: "design patterns" ---> "DS". Previously presented as open.

P20 L18+: "Both the interpolation of the temperature and the interpolation of the rainfall" ---> "Both the interpolation of the temperature and the rainfall"

P20 L22: "detrended Kriging" ---> "DK". Previously presented as open.

P20 L22: "Single event rainfall, on the contrary do not" ---> "On the contrary, single event rainfall did not"

P20 L26: "GPL v3" ---> "GPL v3 (www.gnu.org/licenses/gpl-3.0.en.html)"

P20 L27: "a a-personal" ---> "a-personal"

P20 L28: "~~thus~~"

P21 L1: "instead, for instance, of Mercurial," ---> "instead of for instance Mercurial"

P21 L1: "found at" ---> "found at followed link"

P21 L5: "includes" ---> "include"

P21 L7: "Maven and Gradle, in particular" ---> "in particular Maven and Gradle"

P21 L7+: "Our favor, among the possibilities, was for" ---> "Our favor among the possibilities was"

P21 L12: "instead than another All of them" ---> "than another. All of them"

P21 L13: "programmers' community, but are" ---> "community of programmers, but they are"

P21 L13+: "scientists making increasingly difficult for them maintain their own code" ---> "scientists who are increasingly struggling against difficulties to maintain their own code"

P21 L18: "and run all the software" ---> "and all the software"

P21 L20: "GPL v3 license (www.gnu.org/licenses/gpl-3.0.en.html)" ---> "GPL v3 license"

P21 L22: "alternatives are, among others," ---> "alternatives among others are"

P21 L25: "our" ---> "ours"

P22 in Appendix B: "Percent bias" ---> "percent bias"

P22 in Appendix B: "Coefficient of determination" ---> "coefficient of determination"

P22 in Appendix B: "Root mean square error" ---> "root mean square error"

P25 in Appendix D: The elevation of station 90145 is 3399 m as given in Appendix D and P15 L1. However the max elevation value of study area is 3388 m in Fig. 4. Here, there is a mistake (3399>3388), please correct it.

P29 L5: "Fisheries research" ---> "Fisheries Research"

P29 L25-P29 L37-P30 L10: "Oxford university press" ---> "Oxford University Press"

P29 L29: "Efron, B.~~ and Efron, B.~~"

P29 L36: "Springer Publishing" must be added.

P30 L2: "Journal of hydrology" ---> "Journal of Hydrology"

P30 L7: "Journal of applied meteorology" ---> "Journal of Applied Meteorology"

P30 L8: "International journal of geographical information systems" ---> "International Journal of Geographical Information Systems"

P30 L21: "Biotechnol. Agron. Soc. Environ., vol. 17(2), 392-406" must be added.

P30 L27: "Mathematical geology" ---> "Mathematical Geology"

P30 L31: "Publ. Climatol., 46" must be added.

P30 L35: "JAWRA Journal of the American Water Resources Association, 21(3), 365–380" must be added.

P30 L36: "Monthly weather review" ---> "Monthly Weather Review"

The references in **P30 L22** and **P30 L24** must be replaced due to the chronological order.

P31 L3: "Agricultural systems" ---> "Agricultural Systems"

P31 L6: "Data Acquisition And Processing, Analysis, Forecasting And Other Applications, (WMO No.168)" must be added.

Recommendations

- The validation of SIK component can be performed with remote sensed data (radar and satellite products) besides of ground based data (earth observation stations).
- The climatic and environmental variables can be analyzed at different time scales as daily, monthly, seasonal and annual.
- The interpolation performance can be evaluated in different areas in point of multi-site approach.
- Co-Kriging using secondary data as elevation, vegetation and temperature can be embodied into the SIK.
- Kling-Gupta Efficiency (KGE) criteria whose advantages over NSE are stated in Gupta et al., 2009, can be added as performance benchmark.
- Interpolated temperature and rainfall data can be inputted into a hydrological model. And accuracy and reliability of interpolation can be tested in terms of performance of hydrological model.
- In assessment of interpolation performance, jack knife and k-fold cross validation can be utilized for independent data set of validation procedure.
- Fit of normal distribution of residuals can be tested due to Kriging modeling is a stochastic process.
- Uncertainty analysis can be carried out.
- Comparison based on pixels of raster layer can be implemented.
- Comparison in terms of lowland-highland, underestimate-overestimate, density of stations and representation of low-high rainfall.

References

Adhikary, S. K., Muttil, N. and Yilmaz, A. G., 2016. Genetic programming-based ordinary kriging for spatial interpolation of rainfall. *J. Hydrol. Eng.*, 21 (2), -1--1.

Bostan, P. A., Heuvelink, G. B. M. and Akyurek, S. Z., 2012. Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey. *Int. J. of Applied Earth Observation and Geoinformation*, 19, 115-126.

Di Piazza, A., Lo Conti, F., Noto, L. V., Viola, F. and La Loggia, G., 2011. Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy. *Int. J. of Applied Earth Observation and Geoinformation*, 13, 396-408.

Gupta, H. V., Kling, H., Yilmaz, K. K. and Martinez, G. F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377, 80-91.

Jin Q., Zhang J., Shi M. and Huang J., 2016, Estimating Loess Plateau average annual precipitation with multiple lineer regression kriging and geographically weighted regression kriging. *Water*, 8, 266, 1-20.

Krige, D. G., 1951. A statistical approach to some mine valuations and allied problems at the Witwatersrand. *MSc Thesis of the University of Witwatersrand*.

Nash, J. E., and Sutcliffe, J. V., 1970. River flow forecasting through conceptual models. Part I-A discussion of principles. *Journal of Hydrology*, 10(3), 282-290.

Saghafian, B. and Bondarabadi, S. R., 2008. Validity of regional rainfall spatial distribution methods in mountainous areas. *J. of Hydrol. Eng.*, 13 (7), 531-540.

Wang, S., Huang, G. H., Lin, Q. G., Li, Z., Zhang, H. and Fan, Y. R., 2014. Comparison of interpolation methods for estimating spatial distribution of precipitation in Ontario, Canada. *Int. J. Climatol.*, 34: 3745-3751.