Interactive comment on
"Optimal Design of Hydrometric Station Networks Based on Complex Network Analysis" by
Ankit Agarwal et al.

Anonymous Referee #1

We thank the reviewer for investing his/her valuable time in our manuscript. We think that the reviewer has misunderstood some of the material, and we apologize if our presentation has not been totally clear in all instances. We understand that conciseness is particularly important for manuscripts like this which builds on emerging ideas in the very fast-evolving field of complex network theory, as well as on new ideas around similarity measures, such as event synchronization, which are rather new in hydrology.

We have responded (in black) to each reviewer comment (in red).

General comments

The presented work is based on an interesting and attractive idea, namely the transposition of complex network analysis methods to evaluate and support the optimal design of hydrometric networks. A new metrics is proposed to weight and rank the relative importance of the nodes of the network: the weighted-degree-betweenness (WDB). Two nodes of the hydrometric network are considered as connected if the occurrence of heavy events is sufficiently synchronized at the two stations. If I understand, correctly stations with strong similarities with other stations will have a large number of connections and hence a high WDB value and conversely. The approach is tested against a large and rich data set composed of 1229 German raingauges. Two criterions are used to compare different strategies to remove 10% of the stations of the network: the so-called network efficiency (average value of the inverse of path lengths between two nodes of the network) and the interpolation (i.e., kriging) error. According to the results, removing the lowest ranking stations (stations with the lowest WDB values) has the lowest impact on both criterions, i.e., the proposed ranking measure helps apparently identifying the less influential stations, the station that can be removed from the networks with the most limited consequences on the measurements.

We thank the reviewer for a constructive summary of our manuscript and also for his/her critical and supportive suggestions. Your feedback is vitally important to increase the readability of the work.

Major comments

If I understand correctly stations with strong similarities with other stations will have a large number of connections and hence a high WDB value and conversely.

This is a misunderstanding. The ranking measure that we propose (WDB) is not simply related to the number of links. In fact, the issue that is raised by the reviewer is one of the limitations of the traditional node ranking measure called degree: high number of links → high degree → higher rank. This limitations was one of the reasons why we developed the new measure WBD.

The difference between degree and our new measure WBD is discussed through the artificial network example in Fig. 2 and Table 2. For instance, node 5 has the highest WDB score but the lowest decree score. This means that the most important node (according to our measure) has the lowest number of links in the network!

The difference between degree and WBD and the limitations of the traditional measure decree are also explained in the text, e.g. on page 4, line 9 *("…The degree can explain the importance of nodes to some extent, but nodes that own the same degree may not play the same role in a network. For instance, a bridging node connecting two important nodes might be very relevant though its degree could be much lower than the value of less important nodes…")* or on page 6, line 7 *("… Degree is limited as node ranking measure since it cannot distinguish between different roles in the network. For example, nodes 5, 7, and 8 have the same degree (k_i=2), but node 5 serves as bridge node linking the two parts of the network…")*.

However, to make it even more explicit, we propose to add the following statement (marked in yellow) in section 2.2: *"The degree k of a node in a network counts the number of connections linked to the node directly. For example, the degree of nodes 1, 2 and 4 in network N1 (Fig. 1a) is 1 and for node 3 is 3. In the network N2 (Fig. 1b),* *all nodes have degree 3. The degree can explain the importance of nodes to some extent, but nodes that own the same degree may not play the same role in a network. For instance, a bridging node connecting two important nodes might be very relevant though its degree could be much lower than the value of less important nodes. Moreover, stations with strong similarities with other stations will have a large number of connections resulting in a high degree and hence a high rank and conversely. "*

This being said, the article appears to draw an extremely counter-intuitive if not absurd conclusion: the stations with the lowest correlation with the other stations of the network, station that a priori provide important additional information, should be removed first. This conclusion is highly questionable and may be explained by the selection of inadequate ranking and efficiency evaluation methods. **At least, some further analyses should be conducted before the publication of the manuscript can be considered.** The ranking method is selected without considering the final objective and is probably inadequate.
My feeling is that the proposed approach leads to attribute the highest ranks to the stations with the lowest relative information content which is exactly the opposite of what is meant.

This statement of the reviewer seems to be a consequence of his/her misunderstanding of the rank measure. We have already mentioned that our proposed measure (WDB) identifies stations which provide additional information, for example, on page 14, line 15: *"... it awards stations which provide unique information which cannot be generated from other stations in the network ...*" However, to be absolutely clear, we will add a discussion on how our measure differs from a correlation analysis and accounts for the information content of stations (see response to next reviewer comment).

There is one situation where our method would require additional care: Let's imagine a node that is unrelated to other nodes (no links). Physically, one might imagine this scenario in a meteorological sub-region characterized by fine-scale convective thunderstorms with sparse rain gauge coverage. Hence, precipitation event synchronization across rain gauges in that sub-region would be poor. In that case, indeed, this station would not be the part of constructed network, and would not be ranked. This station should be treated carefully as it provides unique information. We will add a remark on this particular situation in the revision.

An explanation is clearly missing at the beginning of the manuscript to explain why the network construction method and the proposed WDB are suited to rate the relative information content of the stations of the network.

We thank the reviewer for this proposal and propose to add two paragraphs in the introduction:

"… The study aims to identify influential and redundant stations based on the relative information content. In the past several measures, such as zero-lag correlation or time-delayed correlation, have been used. However, these measures are limited by the underlying assumptions, e.g. measuring linear relations. Further, they give equal weight to high and low rainfall values, whereas the main information content in a rainfall time series is embedded in the larger values. In this study, we use event synchronization (ES) as a similarity measure for the network construction. ES is a suitable measure for event like, non-Gaussian data such as precipitation (Stolbova et al., 2014; Tass et al., 1998). It has advantages over other time-delayed correlation techniques (e.g., Pearson lag correlation), as it allows us to define the event time series by determining the threshold, and as it uses a dynamic time delay. The latter refers to a time delay that is adjusted according to the two time series being compared, which allows for better adaptability to the region of interest…."

"… Identifying key nodes in complex networks has attracted increasing attention in recent years (Chen et al., 2012; Hou et al., 2012; Jensen et al., 2016; Kitsak et al., 2010; Zhang et al., 2013). There are several methods to evaluate the importance of nodes (Hu et al., 2013). Degree (k), betweenness centrality (B) and closeness centrality (CC) are the methods commonly used in complex networks (Gao et al., 2013). Studies in different disciplines have shown that degree and betweenness centrality often outperform other node-ranking measures (Gao et al., 2013; Liu et al., 2016). In this study, we propose a novel measure called weighted degree-betweenness (WDB), combining degree *(k)* and betweenness centrality (*B*), which combines the advantages of both. Our case studies show that the proposed measure WDB has an even higher discrimination power compared to betweenness centrality and that it effectively ranks the nodes in the network. WDB is more sensitive to the different roles of nodes, such as global connecting nodes, hybrid nodes, and local centers, and provides a more informative ranking than the existing node ranking measures.

Moreover, the validation based on kriging necessitates a more in-depth analysis and probably further tests to be conducted. The authors consider a so-called kriging error which definition has first to be clarified. It seems to be a theoretical kriging error standard deviation provided by the ArcGIS software geostatistical extension. In fact, this standard deviation depends on the location. What is provided is certainly an average value over the whole considered area – this, of course, has to be clarified by the authors. This standard deviation depends on the network density, on the variance of the rainfall fields and on the characteristics of the variogram. At least the variogram and variance of fields have to be provided as a support to the analysis for the various tested networks.

We thank the reviewer for highlighting the need for such more detailed information. We use the kriging standard error (KSE) which is the square root of the kriging variance (Adhikary et al., 2015; Xu et al., 2018). We estimate the kriging standard error across the entire study area. We will add the following information in the revision:

The kriging variance $\sigma_z^2(x_o)$, in the ordinary kriging can be computed as

$$\sigma_z^2 = \mu_z + \sum_{i=1}^{n} w_i \gamma(h_{oi}) \quad for \sum_{i=1}^{n} w_i = 1$$

where γ(h) is the variogram value for the distance h; h_{0i} is the distance between observed data points xi and xj; μz is the Lagrangian multiplier in the Z scale; h_{0j} is the distance between the unsampled location x_0 (where estimation is desired) and sample locations x_i; and n is the number of sample locations.
The square root of the kriging variance, also named as kriging standard error (KSE), is used as a gauge network evaluation factor which can reflect the performance of optimal gauge combination".
In addition, we will add the variance of the rainfall fields and the variogram, in the revised version.

Removing typical rain gauges can easily have a tricky impact on the average theoretical kriging error standard deviation: the lower density of the network may be partly compensated by the higher homogeneity of the measured rainfall fields (higher decorrelation distances and lower field variances). This compensation effect could explain the modest influence or even the positive effect in table 4 for case 2. In fact, the theoretical error standard deviations are two much dependent on the network itself to enable comparisons between network structures.

We thank the reviewer for his/her suggestion. However, we would like to bring to your attention that the main motivation was to see, how removing the low ranking and high ranking stations impacts the kriging error variance and thereby verifying the efficiency of the method. We would like to point out that the proposed WDB method is independent of the distances between the nodes and considers only the similarity. We agree that when there is strong homogeneity in the rainfall field then a smaller number of stations is required to capture the variance. We believe that 3 different networks, whereas each network is reduced by 10% of its stations, can be compared to the full network by the kriging approach, as for example done in Adhikary et al., 2015; Kassim and Kottegoda, 1991; Xu et al., 2018; Yeh et al., 2006. However, to better analyze the effect of the proposed node ranking in reducing a station network, we will perform additional analyses for inclusion in the revised version.

More classical comparison methods, based for instance on observed interpolation errors, should absolutely be selected and tested by the authors. The distance between interpolated fields obtained with the complete (reference) and reduced networks could, for instance, be evaluated. Interpolation errors could also be computed based on a leave-one-out sampling method providing more realistic estimates of real interpolation errors. Of course, the leave-one-out test station should be selected before the network reduction methods are applied. These verification methods are computationally probably expensive but absolutely necessary.

We will perform additional comparisons by analyzing the interpolation errors, based on a leave-one-out sampling method.

According to these doubts concerning the adequacy of the proposed method and the soundness of the conclusions, I do not recommend the publication of the manuscript and the real-world application of the suggested ranking method unless the proposed additional verifications are conducted.

Again, we would like to thank the reviewer for his/her comments. The doubts raised are on the one hand a consequence of a misunderstanding by the reviewer. This confusion between the two measures (degree and WDB) can be easily rectified by being even more explicit in the description of the method. On the other hand, we feel that the reviewer raised important issues about the comparison using the kriging method. We will provide further details about the kriging and we will add further comparison methods to better evaluate our method.

Reference

Adhikary, S. K., Yilmaz, A. G. and Muttil, N.: Optimal design of rain gauge network in the Middle Yarra River catchment, Australia:, Hydrol. Process., 29(11), 2582–2599, doi:10.1002/hyp.10389, 2015.

Chen, D., Lü, L., Shang, M.-S., Zhang, Y.-C. and Zhou, T.: Identifying influential nodes in complex networks, Phys. Stat. Mech. Its Appl., 391(4), 1777–1787, doi:10.1016/j.physa.2011.09.017, 2012.

Gao, C., Wei, D., Hu, Y., Mahadevan, S. and Deng, Y.: A modified evidential methodology of identifying influential nodes in weighted networks, Phys. Stat. Mech. Its Appl., 392(21), 5490–5500, doi:10.1016/j.physa.2013.06.059, 2013.

Hou, B., Yao, Y. and Liao, D.: Identifying all-around nodes for spreading dynamics in complex networks, Phys. Stat. Mech. Its Appl., 391(15), 4012–4017, doi:10.1016/j.physa.2012.02.033, 2012.

Jensen, P., Morini, M., Karsai, M., Venturini, T., Vespignani, A., Jacomy, M., Cointet, J.-P., Mercklé, P. and Fleury, E.: Detecting global bridges in networks, J. Complex Netw., 4(3), 319–329, doi:10.1093/comnet/cnv022, 2016.

Kassim, A. H. M. and Kottegoda, N. T.: Rainfall network design through comparative kriging methods, Hydrol. Sci. J., 36(3), 223–240, doi:10.1080/02626669109492505, 1991.

Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E. and Makse, H. A.: Identification of influential spreaders in complex networks, Nat. Phys., 6(11), 888–893, doi:10.1038/nphys1746, 2010.

Liu, J., Xiong, Q., Shi, W., Shi, X. and Wang, K.: Evaluating the importance of nodes in complex networks, Phys. Stat. Mech. Its Appl., 452, 209–219, doi:10.1016/j.physa.2016.02.049, 2016.

Stolbova, V., Martin, P., Bookhagen, B., Marwan, N. and Kurths, J.: Topology and seasonal evolution of the network of extreme precipitation over the Indian subcontinent and Sri Lanka, Nonlinear Process. Geophys., 21(4), 901–917, doi:10.5194/npg-21-901-2014, 2014.

Tass, P., Rosenblum, M. G., Weule, J., Kurths, J., Pikovsky, A., Volkmann, J., Schnitzler, A. and Freund, H.-J.: Detection of n : m Phase Locking from Noisy Data: Application to Magnetoencephalography, Phys. Rev. Lett., 81(15), 3291–3294, doi:10.1103/PhysRevLett.81.3291, 1998.

Xu, P., Wang, D., Singh, V. P., Wang, Y., Wu, J., Wang, L., Zou, X., Liu, J., Zou, Y. and He, R.: A kriging and entropy-based approach to raingauge network design, Environ. Res., 161, 61–75, doi:10.1016/j.envres.2017.10.038, 2018.

Yeh, M.-S., Lin, Y.-P. and Chang, L.-C.: Designing an optimal multivariate geostatistical groundwater quality monitoring network using factorial kriging and genetic algorithms, Environ. Geol., 50(1), 101–121, doi:10.1007/s00254-006-0190-8, 2006.

Zhang, X., Zhu, J., Wang, Q. and Zhao, H.: Identifying influential nodes in complex networks with community structure, Knowl.-Based Syst., 42, 74–84, doi:10.1016/j.knosys.2013.01.017, 2013.