Interactive comment on "Investigation of hydrological time series using copulas for detecting catchment characteristics and anthropogenic impacts" by

T. Sugimoto, A. Bárdossy, G. S. S. Pegram and J. Cullmann

General remarks of authors

First of all, we deeply appreciate the care and effort taken by the reviewers in examining this paper.

We checked the literature and the papers concerning asymmetry once more. We agree with the comments of anonymous referee #1 that our definition of asymmetry is similar to the definition of Joe H. (2014) in the sense that it compares the asymmetry along both diagonals of a bivariate copula. Although advanced modeling with new asymmetry function is intriguing, it is beyond the scope of this study.

It seems that the explanation about the relation between asymmetry and hydrograph was not clear, leading to questions and comments from both referees. An effort has been made to substitute more comprehensive figures and text for a better explanation of the material.

The main modifications and improvements in the manuscript are

- Expansion of asymmetry definition from expectation notation to integration notation
- Further comprehensive illustration for the relation between hydrograph and asymmetry.
- In Figure 6 (Asymmetry and Catchment), x-coordinate is log scaled.
- In Section 3.1 (deseasonalization of data) redundant equations and explanations are deleted
- The mistakes in the equations and English have been corrected

Point-to-Point Response to Anonymous Referee #1 – *referee comments in italics*

Nice study that appears to be the first dealing with select asymmetrical properties and interpretations of copula models in a context of daily streamflow statistics for which asymmetry is known to exists. The asymmetry is related to the generalized hydrograph shape. Much of the authoritative text literature (e.g. Nelsen, 2006; Joe, 2014; Durante and Sempi, 2015) do not comprehensively tackle the asymmetry problem of a copula.

Nelsen (2006) is basically devoid of "skewness" (asymmetry) computations— understandably so. Joe (2014, p.66) discusses skewness of a copula and the orientation of the skewness appears conceptually similar (not necessarily numerically equal) to the A1 definition (primary diagonal) of eq. 9. A unique contribution by the paper is the A2 definition (secondary diagonal) of eq. 10. This reviewer has seen many bivariate plots of hydrologic phenomena (such as daily streamflow) and notes the secondary diagonal asymmetry. This asymmetry means a fair share of copula families seen in the literature arguably are in applicable because they have symmetry on the secondary diagonal. This reviewer would like A1 and A2 to also be expressed in direct terms of integration of the copula formula or its density. For example, a Joe (2014) definition for the primary diagonal is: \int\int_{[0,1]} (v-u)C(u,v) du dv from which a secondary asymmetry definition (not identified by Joe) can result \int\int_{[0,1]} (v+u-1)C(u,v) du dv - (1/2) Can the authors of the paper expand the definitions of A1 and A2 beyond the "expectation" notation?

Author's Response (Definition of Asymmetry1 and Asymmetry2)**:**

The “expectation notation” was conventionally used in this research, so there is no reason not to express the equation in integration form beyond the expectation notation as follows:

$$A_1(k) = E\left[(U_t - 0.5)(U_{t+k} - 0.5)\left((U_t - 0.5) + (U_{t+k} - 0.5)\right)\right]$$
$$= \int_0^1\int_0^1 (u_t - 0.5)(u_{t+k} - 0.5)(u_t + u_{t+k} - 1)\, c(u_t, u_{t+k})\, du_t du_{t+k} \tag{9}$$

$$A_2(k) = E\left[-(U_t - 0.5)(U_{t+k} - 0.5)\left((U_t - 0.5) - (U_{t+k} - 0.5)\right)\right]$$
$$= \int_0^1\int_0^1 -(u_t - 0.5)(u_{t+k} - 0.5)(u_t - u_{t+k})\, c(u_t, u_{t+k})\, du_t du_{t+k} \tag{10}$$

It seems sensible, because the terms such as $(u_t + u_{t+k} - 1)$ and $(u_t - u_{t+k})$ appear in this notation, which is comparable to the asymmetry definition by Joe (2014) and anonymous referee #1 :

Asymmetry1 : $\int\int (v-u)c(u,v)dudv$

Asymmetry2 : $\int\int (v+u-1)c(u,v)dudv - 0.5$

In general, there seem other ways to define and apply the asymmetry. L-comoments (L-coskew) suggested by anonymous referee#1 can be one of them.

Have the authors considered the ***L-comoments*** *(Serfling and Xiao, 2007)? But more importantly, the very recent "break through" of L-comoment (bivariate L-moment, bivariate L-skew) definition (Brahimi et al. [2015]) directly in terms of a copula. L-coskew (bivariate skew) \delta^{[12]}_{3;\mathbf{C}} = \int\int_{[0,1]} (60v^2 - 60v + 12) * C(u,v) du dv - (1/2) \delta^{[21]}_{3;\mathbf{C}} = \int\int_{[0,1]} (60u^2 - 60u + 12) * C(u,v) du dv - (1/2)*

Author’s Response (Suggestion for using L-comoments and L-coskew)**:**

L-comoments or L-coskew (Serfling and Xiao, 2007) were not really known to our group. So, we quickly checked the theory in the papers and summarize the main features below.

- L-moments are defined as linear combinations of order statistics.
- The advantage of using order statistics is that, it is not necessary to assume the existence of second order statistics or the statistics of higher order. This can be suitable for heavy-tail distributions.
- L-comoments or L-coskew are extensions of L-moments to the multivariate case.

These functions are theoretically interesting and can be regarded as an advanced definition of copula asymmetry.

The authors generally think that the use of such sophisticated functions enable us to tackle with problems of hydrology and earth system sciences in different ways. For example, the application of such functions for asymmetry1 might be interesting, although their application is beyond the main focus in this research.

These integrals can readily by numerically approximated or integrated by Monte Carlo methods enhanced by low-discrepancy sequence methods. Some final thoughts. A similar study as this does not really appear to have been done. Whereas, this review generally thinks that the physical interpretations of the watershed *and climatology are mechanism producing asymmetry, care is suggested to avoid over interpretations until a great suite of similar studies can be conducted. For example, 9164, line 24 "... A1 ... asymmetry can be related to temporal distribution of precipitation" (what scale of time?) or "... A2 ... more related to catchment and rainfall characteristics ... or ... interseasonal characteristics of climate".*

These are deeply important properties and suggest that copulas are an avenue forward in watershed/climate stochastic modeling. Intuition seems to be correct, but expansion of the authors' thoughts and statements to interpretation of A1 and A2 or other skewness measures or bivariate moment (L-moment) would be informative.

Also, given that we know typical storm water hydrographs are asymmetrical and are inherently formed by a cascade of processes (e.g. water parcel survival from input to output — Markov of sorts), is there a connection between A1/A2 and storm water hydrographs (e.g. unit hydrographs)?

Author's Response (Relation between asymmetry and hydrograph):

We note that anonymous referee#1 gave some positive comments but also the warning about the necessity of careful thought and expressions for asymmetry. It seems important, because this can influence the decent usage of copulas and its fruitful results in the future.

In retrospect, our explanation about the relation between hydrograph and asymmetry seemed to be not good enough in this manuscript, which raised several questions or remarks.

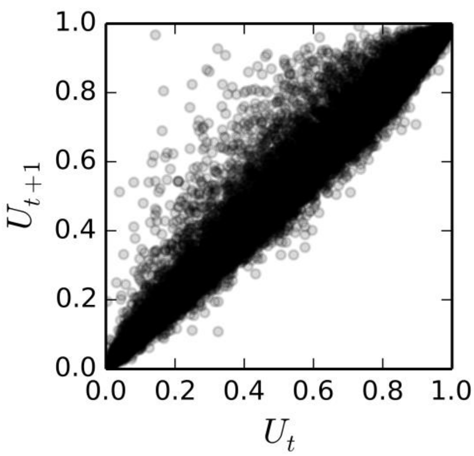
from interactive comment of anonymous referee #1:

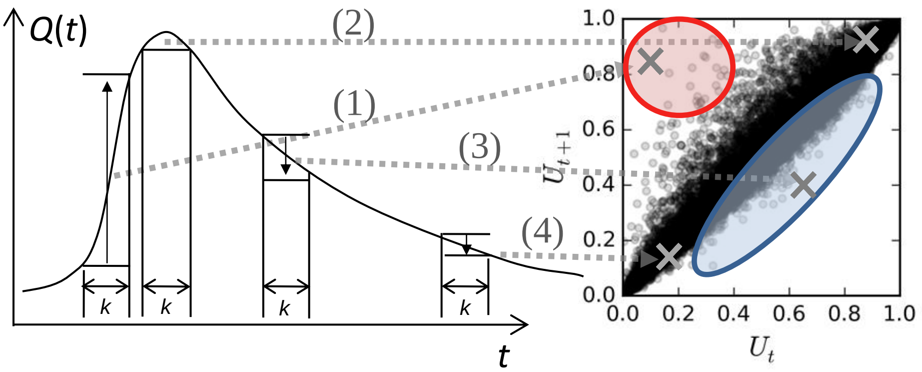
- *... is there a connection between A1/A2 and storm water hydrographs (e.g. unit hydrographs)?*
- *A1 ... asymmetry can be related to temporal distribution of precipitation" (what scale of time?)*
- *expansion of the authors' thoughts and statements to interpretation of A1 and A2*

from interactive comment of anonymous referee #2:

- *Section 3. I would give more practical explanation about Copula asymmetry. It is not fully clear.*

In order to answer these questions, Figure3 has been modified as shown below





Previous version of figure 3 (top) and New version of figure 3 (bottom)

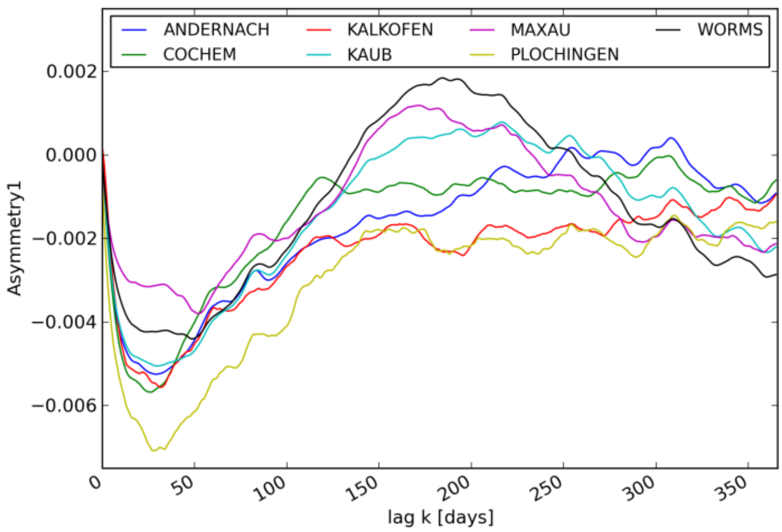
Sketch of the transformation of the values from sample hydrograph (left) to the points on scatterplot of ranks (right): empirical copula calculated from two values separated by time lag $k=1$ [days] in a discharge time series of Andernach where $rank\,correlation = 0.9870$, $A_1(k=1) = -0.0002398$ and $A_2(k=1) = -0.00011037$. The possible combinations of high and low values, which has large impacts on asymmetry, are numbered: (1) low to high, (2) high to high, (3) high to low, (4) low to low. Negative contribution to asymmetry2 is drawn with red circle and positive contribution with blue oval.

This figure illustrates where each pair of values on hydrograph can be plotted on empirical copula. For example, it can be seen there are more points in upper left corner, which demonstrates how the shape of hydrograph can be related to the asymmetry of these empirical copulas. This figure and additional explanation will replace the current figure3 and explanation.

A1 ... asymmetry can be related to temporal distribution of precipitation" (what scale of time?)

Author's Response (Further explanation about asymmetry1):

The asymmetry1 would change depending on the lag k [days] similar to the case of asymmetry 2 (please see the figure below) but based on different reasons. The answer to the question is that the asymmetry1 is significantly small (-0.002 ~ -0.006) for small time scale (lag k = 1~100 [days]) . This is important because this asymmetry can be potentially related to the precipitation of the region. Some basic investigation for asymmetry1 was conducted in the original study (Sugimoto, T., 2014. Copula based stochastic analysis of discharge time series. PhD Thesis. Nr. 232. University of Stuttgart, Germany). It is copied below, but finally not included for the organization of this paper.



In this sense, no concrete conclusion or over interpretation should be given, but it still may make sense to mention the possible mechanism behind it so that it can be the hint for the possible future research works.

or other skewness measures or bivariate moment (L-moment) would be informative.

Hopefully, the new Figure3 and some additional explanation about asymmetry1 will carry the message, so the following sentences were slightly corrected:

(original text at 9164 Line 25 in discussion paper)
This asymmetry can be related to the intrinsic temporal distribution of precipitation.

(improved text)
This implies that the intrinsic temporal distribution of precipitation can be investigated based on this asymmetry, possibly with advanced asymmetry functions such as bivariate moments based on L-moments (Brahimi et al., 2015).

(original text at 9165 Line 2 in discussion paper)
This asymmetry can be related to the characteristics of the runoff and catchment.

(improved text)
This asymmetry can be related to the shape of the hydrograph, therefore the characteristics of the runoff and catchment.

References

Brahimi, Chebana, and Necir (2015) Copula representation of bivariate L-moments: A new estimation method for multi-parameter two-dimensional copula models, Statistics, 49(3)[497–521].

Durante and Sempi (2015) Principles of copula theory, CRC Press.
Nelsen, RB (2006) An introduction to copulas, Springer.
Joe, H. (2014) Dependence modeling with copulas, CRC Press.
Serfling and Xiao (2007) A contribution to multivariate L-moments: L-comoment matrices, Journal of Multivariate Analysis, 98[1765-1781].

9160, Lines 25 and 30: There is confusion in the technical writing aspect of mentioning ARIMA and then evidently switching conceptually to "Fourier analysis". This review suggests that a proof reading would resolve potential confusion.

Author's Response (technical proof reading about ARIMA and Fourier Analysis)

We checked again the literature (Huang et al., 1998) . For the Fourier Analysis, the system must be periodic or stationary and EMD methods have been developed to overcome the restriction. ARIMA is designed originally for stationary process, assuming the no change of the background system. In this sense ARIMA and Fourier analysis is related, but maybe the technical description was not clear, so the text at 9160 Line1 in discussion paper was improved.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. A Math. Phys. Eng. Sci. doi:10.1098/rspa.1998.0193

9162, Line 9: "this statistics" —> "these statistics"

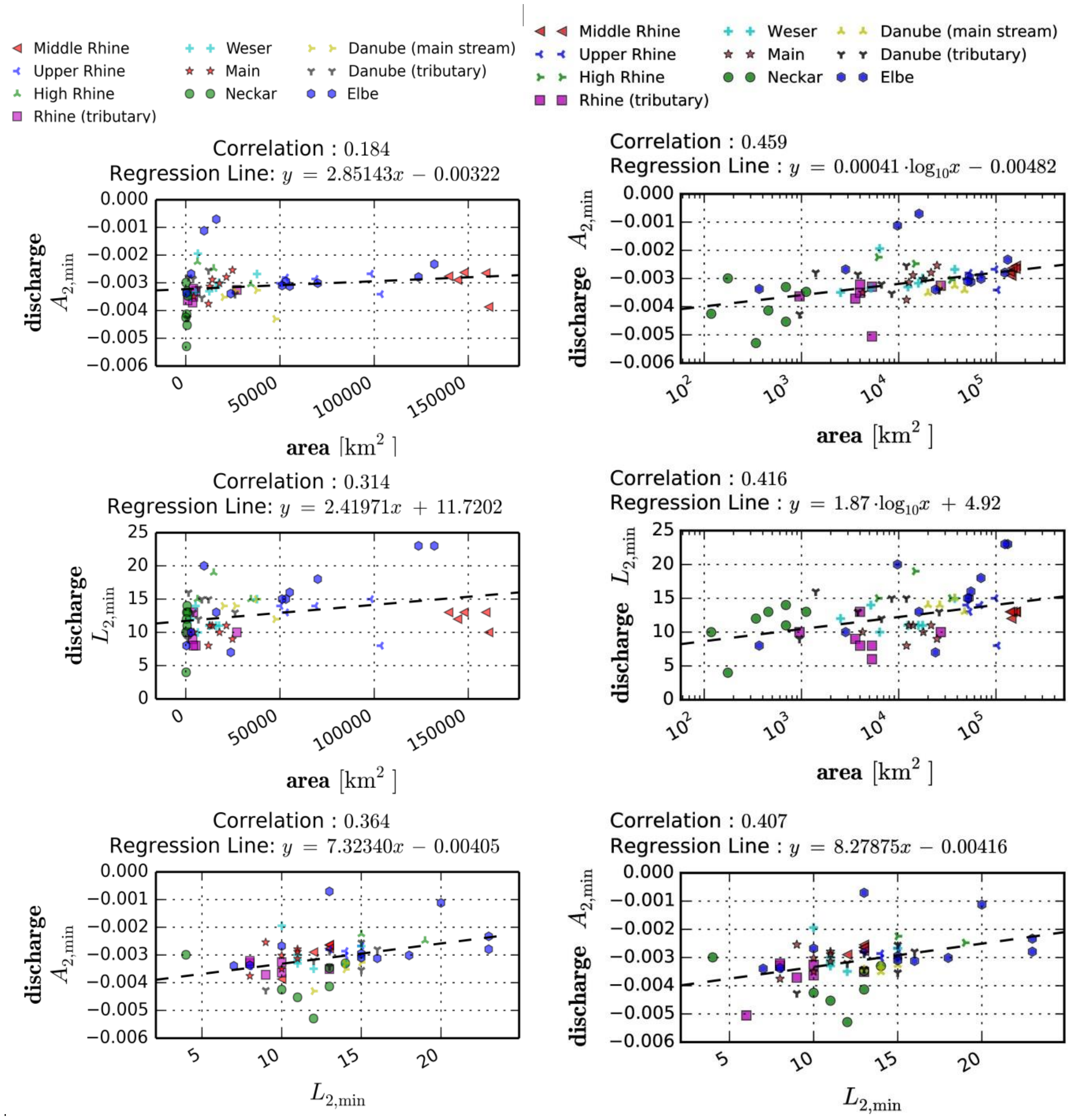
Author's Response:
Thank you very much for pointing out the mistakes. This will be corrected in the revised version of manuscript.

9168, Line 14: missing minus sign in definition of A2(k,t)?

Author's Response:
Yes, this is again a mistake. We thank you for pointing out this error.

Figure 6: Shouldn't the horizontal axis be cast in logarithms?

Author's Response:
For the figure 6 (figure6 old), the same result was plotted on the graph with log-scaled x-coordinate (figure6 new). The correlation and regression line were also calculated based on the log-scaled catchment area. ($x' = \log_{10} x$). Now, it is more clear that there are linear relationships between area and asymmetry measures (A2min, L2min). Thank you very much for pointing this out.



Old version of Figure6 (left) and new version of Figure6 (right):
Relation between Asymmetry and catchment characteristics: minimum of asymmetry2 of discharge and catchment area (top), lag at minimum of asymmetry2 of discharge and catchment area (middle), minimum of asymmetry2 of discharge and lag at minimum of asymmetry2 of discharge (bottom)

Point-to-Point Response to Anonymous Referee #2 – *referee comments in italics*

Received and published: 26 October 2015
The manuscript provides an interesting set of tools based on copula function for investigating discharge time series dynamic.
The topic is particularly interesting since it is in line with the recent and innovative use of copula. Up today copula was applied mainly to perform multivariate frequency analysis while it is potentially useful for detecting and interpreting observed data. This paper is a clear example. The manuscript is easy and pleasant to read, <u>however it includes many analyses and methods that</u>, maybe, it could be worth to split it in two papers.

In the following minor and major concerns are listed.

1) In the abstract API acronym should be defined.

Author's Response:
Thank you for pointing out this. It will be corrected.

2) In the Introduction line 20-22. If the aim is to investigate on the catchment status and the anthropogenic impact, I do not think it is obvious that the solution is to analyze the discharge time series, the reader could expect to see the analysis of the crosscorrelation between rainfall and runoff time series.

Author's Response:
we agree that cross correlation between rainfall and runoff can be the first choice. There are several studies about them, but in our opinion, not enough to explain the causality. The corresponding expressions in abstract will be reconsidered.

3) Section 3. I would give more practical explanation about Copula asymmetry. It is not fully clear.

Author's Response:
Please see "Author's Response (Relation between asymmetry and hydrograph)" in the previous section in this document

4) Section 3.1 line 15. "and instead of "und" 5) Section 3.1 line 25. related "to" temporal distribution

Author's Response:
Thank you very much for pointing out the mistakes

6) Section 3.1 page 9165-9166. The de-seasonalization approach is well known
(Grimaldi, S. Linear parametric models applied to daily hydrological series (2004) Journal of Hydrologic Engineering, 9 (5), pp. 383-391), maybe you can remove the equations
in order to make easier the text.

Author's Response:
Thank you for pointing out this. Section3.1 , the several equation and redundant explanation were removed, instead reference to the study of Grimaldi (Grimaldi, 2004) was added.

7) Section 3.1 pag 9166. I am not surprised to have a residual periodicity since you have removed the annual one. Maybe a weekly periodicity could be still detected.

Author's Response:
Yes, the weekly periodicity might still exist. The important argument here is that the asymmetry remains after certain normalizing treatment of original. This asymmetry is now more reasonable to explain catchment characteristics, because the influence of annual cycle is eliminated. (Not that asymmetry itself is different from month to month. In this sense, the seasonality cannot be fully removed).

*8) Section 4.1.In general this section is very interesting. I would suggest to better explain if the distance D is based on empirical copula and why this is important; and **the uncertainty** of the estimated distance. Maybe these notions are already included in the text but it should be better clarified.*

Author's Response:

I would suggest to better explain if the distance D is based on empirical copula and *why this is important*

Yes, it is based on empirical copula. This study started with the analyzing the asymmetry of empirical copula. After that distance D was examined as an extension to it. It is not necessarily important to use empirical copula, but seems sensible to use it for the purpose of this study.

and the uncertainty of the estimated distance.

There seem two aspects about uncertainty:

1. Uncertainty of Model

From the definition, copula variance can be related to the model uncertainty; how much the natural system is varying. This can be related to the potential calibration difficulty of hydrological model or any parameter estimation of global circulation model.

(the following text is added at 9179 line 15 in discussion paper, original text)
This asymmetry can be related to the intrinsic temporal distribution of precipitation.

The copula based measures introduced in this study can be related to the potential model uncertainty, that is, how much the natural system is varying.

2. Uncertainty of the statistic

Estimating uncertainty of copula distances might be interesting, but seems complicated. It is possible to calculate copula distances for 77 discharge data from different gauging stations, but these data from the same river or same regions should be interrelated and not independent. Thus, it seems not to be simple to estimate the uncertainty of copula distances, therefore this matter is not really discussed in this paper. Copula distances are just calculated for the independent stationary Gaussian processes in order to provide some impression.

These arguments are not clear in the manuscript, so some correction has been done so that these are clearer.