Reply to Referee #2

In the following please find the corrections and comments to the referee’s response. For clarity, the comments of the referee were copied in black and our comments are in blue.

This paper presents an extensive data set from an agricultural catchment in central Germany, where various parts of the water cycle, streams, groundwater, precipitation and soil water are sampled for about 2 years. This data set is used to investigate runoff generation and connectivity between the water cycle components and builds up on a hydrometric paper published in 2014. The paper states that the groundwater system controls streamflow, and no strong influence of precipitation on groundwater and streamflow was observable. The paper is very well written, very informative in background, and the field effort and the data set are great. Nevertheless, there are also several shortcomings in the paper that need attention. I think the analysis needs some more rigor and attempts to a better quantification should be made.

We thank Referee #2 for the comments which were very useful to improve the paper and prepare an improved version of the manuscript. We answer below to each comment in a point-to-point reply.

At first, the paper often claims statistical differences etc., without presenting p-values etc. These p-values need to be reported, this would be absolutely necessary for the reader to be convinced. I understand that lacks of variability in stable isotopes restrict the use of the classical tools, but I do not fully agree about the lack of variability in this work. In my opinion, there is a variability of δ2H in Figure 4. Just plotting a moving average in the stream flow stable isotopes should present some variability. In the lowest panel you clearly see the heaviest stable isotopes values around 7/1 or 8/1, while the values in later winter early Spring are lightest (not sure how important snow in the area would be). Reporting this, e.g. temperature would be good, I don’t think this area (with this elevation) has a long term snow cover over the winter. But that said, the variation seem to be 5-6‰ (the figures are not that easily visible). This variation is a factor ten above measurement precision, and comparable (or even higher) than the differences in McGuire et al. (2005, WRR) (their figure 4). They used 1‰ (and clearly less for 18O) for mean transit time estimations in the HJA.

We made substantive changes to the previous version of the manuscript. In general, we improved the quality of the manuscript by including additional statistical analyses and a hydrological model for the Vollnkirchner Bach subcatchment to bring together hydrometric observations and isotopic based process understanding. We further estimated mean transit times (MTT) for the Vollnkirchner Bach (sites 13, 18 and 94) and the Schwingbach (sites 11, 19 and 64) using FlowPC (Version 3.1, Małoszewski and Zuber (1996)). Different models (dispersion model, exponential model, exponential-piston-flow model, linear model, and linear-piston-flow model) were compared for their results (sigma as goodness of fit) as well as statistical comparisons for site differences were run (bootstrapping for cross-validation). However, the calculated output data did not fit the observed values in terms of the quality criterion sigma and model efficiency. This was mainly due to the small seasonal variations in stream water isotopic signatures. We also bias-corrected the precipitation input data with two different approaches: the mean precipitation value was subtracted from every single precipitation value and then divided by the standard deviation of precipitation isotopic signatures. Afterwards, this value was subtracted from the weekly precipitation values (bias1). For the second approach, the difference of the mean stream water isotopic value and the mean precipitation value was calculated and also subtracted from the weekly precipitation values (bias2).
However, even these bias corrections of the input data did not improve the model outputs (see Table 1), further down).

We plotted the moving average through the streamflow as well as through the groundwater isotope data as recommended by the reviewer (now Fig.7 and 8).

Through the application of a new data analysis tool, topology inference network mapping (Kolaczyk, 2014) in combination with a principal component analysis (Jolliffe, 2002), we further showed the δ¹⁸O isotope relationships between surface and groundwater sampling points based on significant correlations (p<0.05).

Further, the previous paper (Orlowski et al., 2014, Water) presents one event that showed reaction of stable isotopes on incoming precipitation. I am not aware of the number of sampled events, but I think the authors should clearly put more effort on presenting individual events, perform Isotope hydrograph separation on them, and present when and when we do not have some precipitation influence on the response and why. This would be a very good link with the hypothesis 1 (where you should clarify the meaning of “strong temporal” (page 1814), because every reader will have a different perception of such a subjective term), and the importance of the switch between different sources in a catchment and the link to catchment stage are important.

Unfortunately, the event data is rather limited and especially not based on sequential sampling. Thus, we could not present further data showing the influence of event/pre-event water on the isotopic streamflow dynamics.

Here I would suggest the TRANSIT approach (Weiler et al., 2003, WRR), since it could yield comparable results to the presented response times in Orlowski et al (2014). Further you can present what the difference between the two streams is, if they are individually sampled.

We estimated MTT for the two studied streams (Schwingbach and Vollnkirchener Bach). However, we did not obtain any meaningful results. Alternatively, we set up a hydrological model for the Vollnkirchener Bach subcatchment using CMF (Catchment Modelling Framework) by Kraft et al. (2011). CMF is a modular framework for hydrological modelling based on the finite volume approach by Qu and Duffy (2007). We used CMF to simulate water fluxes and advective transport of stable water isotopes (¹⁸O and ²H) to study the origin (Windhorst et al., 2014) and age of water. The generated model is a highly simplified representation of the Vollnkirchener Bach subcatchment’s groundwater body. For estimating the age of groundwater and its flow direction in the subcatchment, a virtual tracer with a given concentration and a fixed decay rate was used in CMF. This approach allowed overcoming the lack of temporal variation in surface and groundwater isotopic signatures. Our model further provides the opportunity to make use of stable water isotope information along with climate, land use, and soil type data, in combination with a digital elevation map to estimate residence times >5 years. We think that this CMF-based method is superior in obtaining catchment functioning insights as compared to classical MTT estimations that empirically try to match observed stable water isotopes.

Beyond, the limitations of the methods/or result presentation I think the merger of results and discussion was not that efficient for the paper. This takes out some clarity, and leaves the reader with some wondering what’s new. And I don’t think this actually accounts for the information wealth the paper delivers. Thus I would like to see a separation of results and discussion in the revised
version, and a discussion that also outlines what we learnt new compared to the current understanding in runoff generation in developed/agricultural catchments.

As recommended by the reviewer, we separated the results and discussion section and restructured the manuscript.

The discussion and the cited references need some stronger focus on the research question of the paper, e.g. I do not think that a general discussion of how the precipitation compares to other precipitation stable isotope studies is necessary. Same holds for discussing stable isotopes in soil water. All valid and interesting points, but please present in the results only data in detail that is necessary for the research question, and shorten the presentation of precipitation data. This will help the paper to get more focussed and will eventually help a clear discussion of the generated understanding.

Following the recommendations of the reviewer, we shortened the section on precipitation isotopes. However, we do not agree that the soil water isotope section should thoroughly be reduced, since most of the process understanding could be gained through soil water isotope data. In addition, this would also conflict with suggestions of other reviewers.

In the introduction you outline general effects of fractionation, precipitation behavior of stable isotopes in detail, while this is not the focus of the work. Please present towards the end of the introduction (Page 1811 and 1812 are really nice) why this work was performed, and how these hypotheses are based on the current research need. I had the feeling this was not convincingly presented, this will also help to present the novelty in the discussion. In summary, add event based result section, were you can present response (also hydrometric) for several events, e.g. using TRANSEP (if sequential precipitation stable isotope data is available for events), and explain differences between the events. Second, separate results and discussion, and focus in the discussion points that are clearly related to the research questions, and how that compares to other work.

Since we did not sample precipitation sequentially, we did not add an additional results section on that topic. However, different rainfall-runoff event types could be detected using the lag-to-peak-time approach in a previous study by Orlowski et al. (2014) and a hydrograph separation was presented likewise in that earlier study in the same catchment.

General comments: The manuscript seems to have some problems with “ff”.

P1810L2: I think the abstract would need an introduction sentence that sets the research field and reasoning.

We edited the abstract. It now reads as follows: “A dual stable water isotope ($\delta^2$H and $\delta^{18}$O) study was conducted in the developed (managed) landscape of the Schwingbach catchment (Germany). The two-year weekly to biweekly measurements of precipitation, stream, and groundwater isotopes revealed that surface and groundwater are decoupled from the annual precipitation cycle but showed bidirectional interactions between each other. Apparently, snowmelt played a fundamental role for groundwater recharge explaining the observed differences to precipitation $\delta$-values.

A spatially distributed snapshot sampling of soil water isotopes in two soil depths at 52 sampling points across different land uses (arable land, forest, and grassland) revealed that top soil isotopic signatures were similar to the precipitation input signal. Preferential water flow paths occurred under
forested soils explaining the isotopic similarities between top and subsoil isotopic signatures. Due to human-impacted agricultural land use (tilling and compression) of arable and grassland soils, water delivery to the deeper soil layers was reduced, resulting in significant different isotopic signatures. However, the land use influence smoothed out with depth and soil water approached groundwater δ-values. Seasonally tracing stable water isotopes through soil profiles showed that the influence of new percolating soil water decreased with depth as no remarkable seasonality in soil isotopic signatures was obvious at depth >0.9 m and constant values were observed through space and time.

Since classic isotope evaluation methods such as mean transit time calculation failed, we established a hydrological model to estimate groundwater ages and flow directions within the Vollnkirchener Bach subcatchment. Our model revealed that complex age dynamics exist within the subcatchments and that much of the runoff must has been stored in the catchment for much longer than event water.

Tracing stable water isotopes through the water cycle in combination with a hydrological model was valuable for determining interactions between different water cycle components and unravelling age dynamics within the study area. This knowledge can further improve catchment specific process understanding of developed, human-impacted landscapes.”

P1811L13 “Garvelmann et al., 2012”, I do think this is a wrong citation, there is not transit time work involved there. Please re-check, otherwise this would be confusing.

We deleted this reference here.

Same line: “transit”, sometimes you are using “transit” sometimes “residence” time in the manuscript, please unify where it makes sense.

We unified the terminology throughout the manuscript.

P1811L17: You have quite some substantial elevation difference in the catchment, more than 100m, and call the catchment low-mountainious, I don’t think this is comparable to real low angle catchments.

There does not exist an international standard definition of a mountain (Goudie and Goudie, 2013) and thus, a low-mountainous terrain. The definition of mountain regions is largely arbitrary because multiple criteria can be used to define such areas, e.g. relative relief, threshold altitude (1000 m) etc. (Perry and Taylor, 2009). Following Perry and Taylor (2009) a hilly terrain, which we equate with a low-mountainous terrain, has an altitudinal difference of 50-100 m (over 5 km distance). In the Schwingbach catchment, 28% of the area exhibits slopes with a gradient >10%. Over a 5 km longitudinal section an altitudinal difference of 92 m could be observed, however, highly depend on the transect. In general, elevation in the Schwingbach catchment ranges between 233–415 m a.s.l.. We conclude that the catchment belongs to the low-mountainous region according to these classifications.

Further, you need to better support this claim of poorly understood, I am not sure about that. What exactly? Further, I think here and in the following lines you need to better describe what was done and understood in developed catchments (also make the difference developed in sense of urbanisation or agriculture or both) and what is still a question of research. Please cite the necessary references to lay out the claim.
P1811L22-24: You need to support this claim better with (more) references, and clearly state why it is limited. The why helps to focus the paper, the “that” is not so important.

We have edited this section as follows: “Moreover, due to human-induced alterations of river systems (e.g. channelisation of streambeds or draining) (O’Driscoll et al., 2010), water fluxes in developed (managed) landscapes can be especially diverse. Almost all European river systems were already substantially modified by humans before river ecology research developed (Klapper, 1990; Allan, 2004). Through changes in land use, land cover and irrigation, agriculture has substantially modified the hydrological cycle in terms of both water quality and quantity (Gordon et al., 2010) as well as altered the functioning of aquatic ecosystem processes (Pierce et al., 2012; Rockström et al., 2014). This complex character of developed, agricultural dominated catchments is often disregarded and established research approaches often failed to fully capture agro-ecosystem functioning at multiple scales (Orlowski et al., 2014). Since agricultural land use (arable land, permanent crops, and grassland) is the most dominant land use in Europe (UNEP, 2002), there exists a pressing need to understand biogeochemical fluxes (e.g. nitrogen compounds or pesticides) coupled with water fluxes in these managed landscapes (Orlowski et al., 2014) and to figure out a way to embed this landscape heterogeneity or the consequence of the heterogeneity into models (McDonnell et al., 2007).

One way to better understand the relationship between precipitation, stream, soil, and groundwater, is a detailed knowledge about the isotopic composition of the various water sources (surface, subsurface, and groundwater) and their variation in space and time."

P1811L24: “residence time” unify terminology.

We unified the terminology throughout the manuscript.

P1811L28: “Kendall and McDonnell, 1998” The book consists of individual chapters from various authors, please cite the author of the relevant chapter, and have the Editor of the book in the reference list. Please improve throughout the manuscript.

We corrected it throughout the manuscript.
“water lines” in general a catchment should have one LMWL, since it describes the precipitation, it reads confusing here. Since water sources other than precipitation cannot have a LMWL.

We changed the sentence as follows: “To compare different water sources on the catchment-scale, a local meteoric water (LMWL) line is developed and evaporation water lines (EWLs) are used.”

“strong temporal” Please avoid subjective descriptions. You also could relate this hypothesis to new/old water paradox work from Kirchner and other authors.

We moved away from stating hypotheses and changed the paragraphs as follows: “...This interaction between groundwater and surface water remains poorly understood in many catchments throughout the world but process understanding is fundamental to effectively manage the quantity and quality of water resources (Ivovic, 2009). Sklash and Farvolden (1979) showed very early, that groundwater plays an important role as a generating factor for storm and snowmelt runoff processes. In many catchments, streamflow responds promptly to rainfall inputs but variations in passive tracers such as water isotopes are often strongly damped (Kirchner, 2003). This indicates that storm runoff in these catchments is dominated mostly by “old water” (Buttle, 1994; Neal and Rosier, 1990; Sklash, 1990). However, not all “old water” is the same (Kirchner, 2003). This catchment behaviour was described by Kirchner (2003) as the old water paradox. Thus, there is evidence of complex age dynamics within catchments and that much of the runoff is stored in the catchment for much longer than event water (Rinaldo et al., 2015). Still, some of the physical processes controlling the release of “old water” from catchments are poorly understood, roughly modelled, and the observed data do not suggest a common catchment behaviour (Botter et al., 2010).”

“...To capture spatial landscape heterogeneity, but to keep data acquisition simple, stable water isotope data were coupled with hydrodynamic data from a previous study by Orlowski et al. (2014) in the developed Schwingbach catchment (Germany) to unravel water flow paths and interactions between different water cycle components. Results obtained through this earlier study imply that the Schwingbach catchment is highly responsive indicated by fast runoff responses to precipitation inputs (Orlowski et al., 2014). Moreover, groundwater reacted almost as quickly as streamflow to precipitation events with raising head levels. Thus, the catchment showed “old water” paradox like behaviour (Kirchner, 2003). We further showed that streamflow was predominantly generated in the catchment headwater area and that gaining and losing stream reaches occurred in parallel along the studied stream affected by the underlying geology.

Thus, stable water isotopes in combination with hydrodynamic data of a two-year monitoring period (July 2011 to July 2013) were utilised to explore spatio-temporal isotopic variations, unravel linkages between the different water cycle components, investigate the transformations from precipitation to soil and groundwater, and analyse the effect of small-scale landscape characteristics (i.e. soil physical properties, topographic wetness index (TWI), distance to stream, and vegetation cover) on soil water isotopic composition. Further, stable water isotope data was utilized to estimate groundwater ages and flow directions in the Vollnkirchener Bach subcatchment via an hydrological model setup based on the findings of Orlowski et al. (2014).”
We thoroughly edited the Introduction (see above) and included the following: “To compare different water sources on the catchment-scale, a local meteoric water (LMWL) line is developed and evaporation water lines (EWLs) are used. They represent the linear relationship between $\delta^2$H and $\delta^{18}$O of meteoric waters (Ingraham, 1998) in contrast to the global meteoric water line (GMWL), which describes the world-wide average stable isotopic composition in precipitation (Craig, 1961a). Thus, the comparison of stable isotope data for stream, soil, or groundwater samples relative to the global or local meteoric water lines can provide general understandings on water cycle processes at specific research sites (Song et al., 2011).”

P1815L10ff: You should make clearer when you write about the one stream and about the other. Present the individual catchment boundaries in fig1. In the text it needs to be more clear what information is linked to what (sub-)catchment.

We edited the respective paragraphs in the manuscript to make the description of the two streams clearer. Moreover, we modified Figure 1 and included the Vollnkirchener Bach subcatchment boundaries as well as the names of the streams in Figure 1c. In addition, Figure 1 now encompasses a map showing the location of the study area in Germany (Fig. 1a) and a map of the soil sampling points along the Vollnkirchener Bach (Fig. 1d).

P1816L6ff.: Do you have long term data for the area, would be better than presenting precip sum for only one year. Catchment outlet refers to the Schwingbach Catchment?

For comparisons, we included the following section in the manuscript: “The climate in the study area is classified as temperate with a mean annual temperature of 8.2°C. An annual precipitation sum of 633 mm (for the hydrological year 1 November 2012 to 31 October 2013) was measured at the catchment’s climate station (site 13, Fig. 1b). The year 2012 to 2013 was an average hydrometeorological year. For comparison, the climate station Giessen/Wettenberg (25 km N of the catchment) operated by the German Meteorological Service (DWD, 2014) records a mean annual temperature of 9.6 °C and a mean annual precipitation sum of 666 ± 103 mm for the period 1980 to 2010.”

P1816L19 “Significant” should only be used in relation to statistical tests. Please avoid otherwise.

The sentence now reads as follows: “Substantial snowmelt peaks were observed during December 2012 and February 2013.”

P1816L9ff. “>114 L s$^{-1}$. Why “>” and not the exact value? I think the unit might be better in depth (mm)/area to have comparability if this is a lot or not.

We included the following information in the sentence: “Discharge peaks from December to April (measured by the use of RBC-flumes with maximum peak flow of 114 L s$^{-1}$, Eijkelkamp Agrisearch Equipment, Giesbeek, NL) and low flows occur from July until November.”

Moreover, we changed the discharge unit to mm d$^{-1}$.

P1818L9-12: I think such information should be in the introduction, not in the method section.

We edited the Introduction and decided to move the sentence to the Results section: “Distances to the stream are linked to water flow path lengths and were therefore supposed to be a controlling
factor. However, no impact of different distances to the stream on soil water isotopic signatures could be observed.”

P1820, Chapter 3.1. Here a decision regarding the focus is necessary; the whole section of precipitation is very detailed, while not really linked to the hypotheses/research question. In general this could be shortened. Further I think the introduction of the LMWL (3.5) should be together with chapter 3.1.

We decided not to shorten the section on “Variations of precipitation isotopes and d-excess” but to strengthen it and to include results of Figure 10 on the LMWL and EWLs (now Fig. 3) in this section.

P1824L19ff: As outlined earlier, there are variations, with heavier values in summer/fall and lighter values in spring. It is not even close to a straight line.

As described above the MTT calculations with FlowPC did not provide meaningful results, not even after bias-correcting the precipitation input data (Table 2). The failure of the MTT estimations is mainly attributed to the little variation in stream water isotopic signatures, which have to be weekly averages for the calculations with FlowPC. Thus, potential variations in stream water isotopes are averaged out.

P1825L8: Here again, standard deviation is already 3 times higher than analytical precision.

See argumentation above.

P1825L21: Did you try some mixing calculations?

We included the following sentence in the Discussion section: “Just by comparing mean precipitation ($\delta^{18}O = -6.2 \pm 3.1$), stream (e.g. $\delta^{18}O = -8.4 \pm 0.4$ for the Vollnkircher Bach), and groundwater isotopic signatures ($\delta^{18}O = -8.2 \pm 0.4$ for the meadow) (Tab. 1), it is obvious that simple mixing calculations do not work either.”

P1826L4/5 “Statistically similar”, please provide p-value (t-test?)

We applied a new data analysis tool, network mapping (Kolaczyk, 2014) in combination with a principal component analysis (Jolliffe, 2002) to determine the $\delta^{18}O$ isotope relationships between surface and groundwater sampling points based on significant correlations ($p<0.05$). These new results are included in the the Section on “Isotopes in groundwater”.

P1826L14: Please also report p-value here and throughout the manuscript.

See reply above.

P1826L23-27: These lines are not needed. This should be clear from intro and method, here you should just report the results.

We deleted this sentence from the manuscript.

P1827L12-13: “. . .(Fig. 8).”, this is difficult to see, please test statistically incl. p-value. Same for lines 16 and 20.

We generally included p-values in the revised manuscript where necessary.
P1827L22-24: What did you expect to see? So how should rainfall influence the groundwater signal? Maybe the volume of preferential flow is low compared to the volume of GW and no (strong) effect can be seen.

The following paragraph was already given in the manuscript as explanation: “Subsoil isotopic values were statistically equal to stream and groundwater isotopic values (Fig. 4) implying that the catchment was under baseflow conditions during the sampling campaign and that capillary rise of groundwater occurred. Nevertheless, the rainfall isotopic signal was not directly transferred through the soil to the groundwater body, even so prompt groundwater head level raises as a result of rainfall-runoff events occurred. This likewise supports the assumption of double paradox-like catchment behaviour.”

P1827-1828L28-2: It would be important to work out the site differences and related the differences in measurement to this.

We edited this paragraph as follows: “Garvelmann et al. (2012) obtained high resolution δ2H vertical depth profiles of pore water at various points along two fall lines of a pasture hillslope in the southern Black Forest (Germany) by applying the H2O(liquid)–H2O(vapor) equilibration laser spectroscopy method. The authors showed that groundwater was flowing through the soil in the riparian zone (downslope profiles) and dominated streamflow during baseflow conditions. Their comparison indicated that the percentage of pore water soil samples with a very similar stream water δ2H signature is increasing towards the stream channel (Garvelmann et al., 2012). In contrast, we found no relationship between the distance to stream and soil isotopic values in the Vollnkirchener Bach subcatchment over various heights above sea level (235–294 m a.s.l.).”

P1830L10ff. Lots of details, but it needs to be clearer what the message is. I think with separating results/discussion this will automatically become clear, if the discussion is related to the hypotheses and general progress in the field.

We separated the Results and Discussion and changed the order of the sections in the revised manuscript.

P1831L5: If the focus is on the LMWL please add this to the precipitation isotope section. If you want to compare the water lines of individual water cycle components make that clear, and the analysis statistically more sound.

We moved the paragraph on the LMWL and EWLs to the Section on “Variations of precipitation isotopes and d-excess” and shortened it.

P1831L21: How do you exclude moisture recycling?

We edited the respective paragraphs to make it clearer: “Considering d-excess values, it is well-known that precipitation events originating from oceanic moisture show d-excess values close to +10‰ (Craig, 1961a; Dansgaard, 1964; Wu et al., 2012) and one of the main sources for precipitation in Germany is moisture from the Atlantic Ocean (Stumpp et al., 2014)...Continental precipitation events originating from oceanic moisture can approach d-excess values of +10‰ (Wu et al., 2012) (Fig. 2, solid line). Air mass trajectories at intercontinental, southern and eastern regions are suggested to be more stable with less variable moisture sources in these regions compared to sites...
near the coast (Stumpp et al., 2014). Therefore, rainout histories on the continent itself are more stable (Stumpp et al., 2014).”

P1832L5-6: From the elevation range they seem not to be comparable.

See argumentation above on low-mountainous regions.

P1832L17: Here event based hydrograph separation would be nice, they should present how much % groundwater contributions are during events.

See argumentation above.

P1832L20: “a larger contribution. . .” You did not present calculations.

We edited the sentence as follows: “Even so streamflow and groundwater head levels promptly responded to precipitation inputs, there was no obvious change in their isotopic composition due to rain events (old water paradox behaviour).”

Figure 1: Please present sub-catchments. Fonts are too small to decipher.

Changed as recommended.

Figure 2: What is the relation to hypotheses/research questions? I think this figure could be removed.

Deleted Figure 2 from the manuscript.

Figure 4: Too small. For trying, plot a moving average, I guess you start seen some seasonal patterns.

Changed as recommended.

Figure 5: Larger fonts needed.

Changed as recommended.

Figure 8: font size

Changed as recommended.

Figure 10: please report equations of water lines, here.

We included the respective equations in Figure 10 (now Fig. 3).