Reply to Referee #1

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 study presents stable water isotopic data for a catchment in Germany. While there is much variability in the precipitation isotopes, there was little in the groundwater and stream water. The manuscript is well written and easy to read, however, lacks rigor and has limited quantifications backing up the main conclusions. This makes it difficult to see novelty and assess what the actually contribution of the work is toward advancing understanding of rainfall-runoff processes in a general sense (i.e., beyond the empirical sense of this specific location). The challenge to the authors, which will require significant revision and additional analysis, is to take their data and develop a procedure/approach overcoming the lack of variability. While I can appreciate that the lack of variability "restricted the use of classical isotope hydrology techniques" (P1810L25), it leads to two clear questions: (1) Why not develop a new method to leverage the data you have or (2) Why not measure other tracers (geochemicals or electrical conductivity come to mind) that better map the flow domain? The authors opt towards choice (1) which I can agree with since it is difficult to add analysis and sampling after the fact. This makes the central question of this research how to turn lemons into lemonade? Personally, I think this is a pretty important central question given the ubiquity with which stable water isotopic data is currently being collected. However, the approach developed is extremely qualitative so it leaves the reader wondering what we have learned here and how to learn the same thing at a different location. The authors do a good job highlighting the need for such knowledge (see around P1812L2-6) but the study leaves something to be desired on how to achieve such knowledge.

Of course, I do not have any great suggestions on what that method/approach could be since that is the core of the research. What I am looking for here is a methodology that can allow for comparisons of this site to others. For example, perhaps the connections to land cover could be leveraged to develop a basic mixing model allowing for characterization of impacts of spatial vegetation pattern heterogeneity on hydrological response in these types of landscapes? Then you could do some cross validation estimates (bootstrap or leave-on-out type) on the error and uncertainties incurred? Development of a modeling extension (in whatever form) is recommended here to help provide a clear hypothesis testing/quantification framework. This would help demonstrate exactly how "conducting a stable water isotope study in Schwingbach catchment helped to identify relationships between precipitation, stream, soil, and groundwater in a developed catchment" (P1832L13). That would constitute a significant contribution and move this study beyond its heavy empirical tilt (which is needed).

We gratefully acknowledge the comments of the reviewer, which helped us to improve the manuscript. In general, we performed additional statistical analyses and set up a new hydrological model for the Vollnkirchener Bach subcatchment to explain mixing processes.

1) As a new data analysis tool, we used a network map (Kolaczyk, 2014) in combination with a principal component analysis to demonstrate the isotopic relationships between surface and groundwater sampling points.

2) We further utilized the isotope data to calculate the mean transit times (MTT) for the Vollnkirchner Bach (sites 13, 18 and 94) and the Schwingbach (sites 11, 19 and 64) using FlowPC. We bias corrected the input data (precipitation) to improve model results and

statistically compared these results with the initial non-corrected model results. 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, even after bias-correcting the input data. This was mainly due to the small seasonal variations in stream water isotopic signatures. Therefore, we conclude that the application of MTT estimation methods based on stable water isotopes failed in the Schwingbach catchment and developed a new data-driven groundwater model to simulate observed stable water isotope data.

3) To further overcome the criticized qualitative data analyses approach and to verify and validate the hydrological processes, the Catchment Modelling Framework (CMF) by Kraft et al. (2011) was used to setup a hydrological model for the Vollnkirchner Bach subcatchment. Thereby, we were able to estimate spatially distributed groundwater ages. The flexible setup of CMF and the variety of available flow-accounting equations allows customizing the setup making this modeling framework especially suitable to be used in our isotope tracer study (Windhorst et al., 2014). The additional data analyses now allow a better comparison of our study area to other sites.

We refrained from using other tracers, since intense measurement and sampling campaigns have already been conducted in previous studies (Lauer et al., 2013, Orlowski et al., 2014).

Good that the hypotheses are clearly stated. However, there are some ambiguous words in there that reinforce the qualitative nature of the study and it is questionable how testable these really are. For example in hypothesis (1), what do you really mean by strong? Does it mean high in amplitude or quick changes (steep slopes in time)? It would be good to put this in the context of something measureable or quantifiable. For hypothesis (2), there is an inherent assumption of instantaneous mixing throughout the groundwater. Early work from Sarah Dunn [Dunn SM, McDonnell JJ, Vache KB. 2007. Factors influencing the residence time of catchment waters: A virtual experiment approach. Water Resources Research 43: W06408] and more recent work by others (e.g., Markus Hrachowitz [Hrachowitz M, Savenije H, Bogaard TA, Tetzlaff D, Soulsby C. 2013. What can flux tracking teach us about water age distribution patterns and their temporal dynamics? Hydrology and Earth System Sciences 17: 533–564]; Ype van der Velde [Van der Velde Y, Torfs PJJF, van der Zee SEATM, Uijlenhoet R. 2012. Quantifying catchment-scale mixing and its effects on time-varying travel time distributions. Water Resources Research 48: W06536]) have really questioned such complete mixing. So, this might actually be a rather poorly constructed hypothesis. Lastly, hypothesis (3) would require measuring physical distributions of flow pathways in the subsurface at a scale not really achieved here, would it not? Why not focus in on a clear and testable hypothesis to better streamline the presentation of a central key finding? I think this will be achieved when considering the recommendations of the previous paragraph.

We have revised the Introduction completely. The section now reads as follows: *"The application of stable water isotopes as natural tracers in combination with hydrodynamic methods has been proven to be a valuable tool for studying the origin, formation, and interrelationship between surface water and groundwater (Blasch and Bryson, 2007; Goni, 2006), partitioning evaporation and transpiration (Phillips and Gregg, 2003; Rothfuss et al., 2010, 2012; Wang and Yakir, 2000), and further mixing processes between various water sources (Aggarwal et al., 2007; Clark and Fritz, 1997c; Kendall and*

Coplen, 2001; Wu et al., 2012). Particularly in catchment hydrology, stable water isotopes play a major role since they can be utilised for hydrograph separations (Buttle, 2006; Hoeg et al., 2000; Ladouche et al., 2001; Munyaneza et al., 2012), to calculate the mean transit time (McGuire et al., 2002, 2005; Rodgers et al., 2005b), to investigate water flow paths (Barthold et al., 2011; Goller et al., 2005; Rodgers et al., 2005a), or to improve hydrological model simulations (Birkel et al., 2010; Koivusalo et al., 1999; Liebminger et al., 2007; Rodgers et al., 2005b). However, spatio-temporal sources of stream water in low angle, developed catchments are still poorly understood. This is partly caused by damped stream water isotopic signatures excluding traditional hydrograph separations (Klaus et al., 2015). Unlike the distinct watershed components found in steeper headwater counterparts, lowland areas often exhibit a complex groundwater–surface water interaction (Klaus et al., 2015). 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 (Ivkovic, 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).

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. In principal, isotopic signatures of precipitation are altered by temperature, amount (or rainout), continental, altitudinal, and seasonal effects. They are mainly influenced by prevailing atmospheric conditions during rainfall and snowfall causing a depletion of isotopes (Araguás-Araguás et al., 2000; Blasch and Bryson, 2007; Clark and Fritz, 1997c; Gat, 1996; Rohde, 1998). The input signal becomes more pronounced in snow-

dominated systems where snowfall and snowmelt are depleted in heavy stable water isotopes relative to rainfall (Maule et al., 1994; O'Driscoll et al., 2005). Stream water isotopic signatures can reflect precipitation isotopic composition and moreover, depend on discharge variations affected by seasonally variable contributions of different water sources such as bidirectional water exchange with the groundwater body during baseflow, or high event-water contributions during stormflow (Genereux and Hooper, 1998; Koeniger et al., 2009). Following the way of precipitation over the unsaturated zone to the groundwater, the process of infiltration in itself is known to be a non-fractionating process (Gonfiantini et al., 1998), except for mixing between different water pools (e.g. moving and standing water) (Gat, 1996). However, precipitation falling on vegetated areas is intercepted by plants and re-evaporated thus isotopically fractionated. The remaining throughfall infiltrates slower and can be affected by evaporation resulting in an enrichment of heavy isotopes, particularly in the upper soil layers (Gonfiantini et al., 1998; Kendall and Caldwell, 1998). In the soil, specific isotopic profiles develop, characterized by an evaporative layer near the surface especially under arid and semi-arid climate. This decreases exponentially with depth (Zimmermann et al., 1968), representing a balance between the upward convective flux and the downward diffusion of the evaporative signature (Barnes and Allison, 1988). In humid and semi-humid areas, this exponential decrease is generally interrupted by the precipitation isotopic signal. Hence, the combination of the evaporation effect and the precipitation isotopic signature determine the isotope profile in the soil (Song et al., 2011). Once soil water reaches the saturated zone, this isotope information is finally transferred to the groundwater (Song et al., 2011). Soil water can therefore be seen as a link between precipitation and groundwater, and the dynamics of isotopic composition in soil water are indicative of the processes of precipitation infiltration, evaporation of soil water, and recharge to groundwater (Blasch and Bryson, 2007; Song et al., 2011).

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 δ^2H 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).

Identifying the origin of water vapour sources and moisture recycling (Gat et al., 2001; Lai and Ehleringer, 2011), the deuterium-excess (d-excess), defined by Dansgaard (1964) as $d = \delta^2H - 8 \times \delta^{18}O$ can be used, since the d-excess mainly depends on the mean relative humidity of the air masses formed above the ocean surface (Zhang et al., 2013). In addition, the d-excess reflects the prevailing conditions during evolution, interaction, or mixing of air masses en route to the precipitation site (Froehlich et al., 2002).

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)."

Lastly, the results and discussion should be separated. Having them combined contributes to the qualitative feel of the presentation. It makes the results read more like a story than a presentation of substantial findings.

We have separated the results and discussion section.

Further, the findings seem to echo much of what is already seen in the literature. This potentially points to a lack of novelty. With this, I think the study has a bit farther to come before it can be considered ready for publication in HESS. As it reads now, it is more suitable for a regional journal or a journal with a more empirical focus (which I think the authors can move beyond).

We have revised and improved the whole manuscript and e.g. included a hydrological model to estimate groundwater ages and flow directions in the Vollnkirchener Bach subcatchment. Our model 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. Such long residence times could previously only be determined via other tracers such as Tritium (e.g. Michel (1992)). If stable water isotope information is used alone, it is known to cause a truncation of stream residence time distributions (Stewart et al., 2010).

Thus, we are convinced that the manuscript is now ready for publication in HESS.

Minor/Editorial Comments

The title does not seem grammatically correct. Should be something like: "Exploring water cycle dynamics by sampling a multitude of stable water isotope pools: : :" or "Exploring water cycle dynamics through sampling multiple stable water isotope pools: : :"

The title now reads as follows: "*Exploring water cycle dynamics by sampling multiple stable water isotope pools in a small developed landscape of Germany*"

The last paragraph of the introduction is awkward. Since this is the paragraph that sets the tone for the presentation, it is fairly important. What was the "former" study? Are you referencing previous work that already used these data? Please improve this (see also general comments).

When referring to the "former" study we are talking about the findings of a previous study by Orlowski et al. (2014) conducted in the same catchment. The findings from the previous study were useful and, in fact, necessary to successfully complete the current research. As such, the present analysis complements the one published in the previous paper (Orlowski et al., 2014), but has a

distinct different focus. However, the previous study does not make use of isotope data presented in this manuscript.

We have generally revised the Introduction section (see reply above).