Reply to reviewer 2

We would like to thank you for reviewing our manuscript and for providing constructive and useful comments. Below is our responses to your specific comments.

The line number referenced in this reply are from the revised manuscript (attached as Manuscript_revised_2.pdf).

Comment 1: The manuscript contains an excessive number of abbreviations (e.g., PWSS, SIW, SLV, BIMM, JVD, WTP, MGD, SEPT, SWTP etc.) that sometimes make the reading difficult. Please, use acronyms sparingly, only when strictly necessary.

- Thank you for the advice. We have tried to reduce the usage of acronyms throughout the manuscript. Some examples are provided below. Acronyms replaced by full word are highlighted in yellow.

  L 24: "Our work highlights the ability of stable isotopes in water to analyze PWSS and document aspects…."
  L 122: "Jordan Valley District operates three water treatment plants (WTP). The Jordan Valley Water Treatment Plant…"
  L 139: "The Jordan Valley District water distribution system consists of one primary (Fig. 1), several secondary (line 2 through 6, Fig. 1) and numerous tertiary …"
  L 142 -144: "Water from Jordan Valley Treatment Plant is pumped directly into transmission line 1 and water from South East Treatment Plant water is pumped into transmission lines 2 and 3 (Fig. 1)."
  L 152- 154: "Each month from May to October 2015, we sampled water at sources contributing to the Jordan Valley District service area and at numerous locations ("distribution sites" or simply "sites", Figure 1) on the Jordan Valley District transmission lines."
  L 290: "The resulting values were summed across all distribution sites $(\sum A_I X f_{I,J})$ and divided by the total area of Jordan Valley District supply region to obtain the areal coverage (*A*) of that source."
  L 349: "For sites on the western portion of the Jordan Valley District, the model-inferred mean JVTP contributions were…."
  L523: "….comparing the observed and predicted stable isotopes in water (or other conservative geochemical tracers) at several…."

Comment 2: Introduction: although the results are clear, the Introduction misses a clear definition of the working hypothesis upon which to establish specific objectives. The objective should stem from the analysis of the identification of research gaps in the current literature and/or from a practical management issue in the study PWSS. Pease, re-structure the introduction taking this into consideration.

- Thank you so much for pointing it out. We have rewritten the introduction section that clearly states the problem and its importance. We have identified the research gaps and have pointed out the ways in which and why BIMM can be a useful tool to study urban water systems.

The restructured introduction is included below. (L 29 – 105 in the revised manuscript).

1. Introduction

Public water supply systems (PWSS) are an important component of the critical infrastructure supporting human development across the globe. The complexity of PWSS can vary widely, ranging from linear, single-source distribution systems to branched distribution networks using multiple water sources and complex storage systems. To understand the stability of water supplies, conduct risk evaluation, and develop effective and efficient responses for particular threats (supply contamination, infrastructure failure, etc.), it is critical to understand the physical and spatial structure of the distribution network, connectivity within the system, and the links between the point-of-use and environmental water sources.

The physical structure of the distribution system and basic information on water sources are generally well documented for most first-world metropolises. In these settings, water managers traditionally rely on network analyses to study different aspects of water distribution systems, including pressure gradients, flow rates, water losses from the supply system, identification of vulnerable sections, and tracking of disinfectants and contaminants (*Boryczko and Tchórzewska-Cieślak*, 2014; *Pietrucha-Urbanik*, 2015; *Yoo et al.*, 2015). These analyses are generally robust; however, they are seldom validated using observational data and can suffer from shortcomings including the absence of unique solutions in underdetermined systems, assumption of invariant flow rates, uncomprehensive or non-inclusiveness of uncertainty in the analysis (*Waldrip et al.*, 2016), and outdated/incorrect information on infrastructure (*Liggett and Chen*, 1994). Beyond statistical and computational issues, hydrodynamic modelling requires extensive and detailed information about the PWSS, including node elevation, pipe length and diameter, and pump operating data. For many communities in the developing world, where distribution networks are commonly unregulated and decentralized, even basic information on supply system structure and source contributions may be incorrect, incomplete, undocumented, or difficult to obtain. Hydrodynamic modeling of PWSS in such cases can be challenging and prone to significant errors.

With growing water security challenges due to climate change (*Arnell*, 1999; *Vörösmarty et al.*, 2010), expanding complexity and dynamicity of urban water systems and increasing detrimental effects of aging water infrastructure in many countries (*Hanna-Attisha et al.*, 2016; *Kaushal*, 2016; *Larsen et al.*, 2016; *Schnoor*, 2016) it is important to develop new techniques and methods to study PWSS that requires minimal information on the physical structure and connectivity within the supply system. In this regard, new techniques and methods are being developed to understand (1) failure in the water distribution system with limited, imprecise and ambiguous information on the supply structure (*Najjaran et al.*, 2005; *Ismail et al.*, 2011; *Bolar et al.*, 2013; *Kabir et al.*, 2015) and (2) analyze the water distribution system in a probabilistic framework (*Waldrip et al.*, 2016; *Waldrip et al.*, 2018).

Stable isotopes in water (SIW) can serve as an important tool to study water management within complex PWSS. SIW are naturally occurring tracers of the terrestrial hydrological cycle and significant isotopic differences between water sources can exists at catchment, regional, and global scales due to seasonal biases in recharge, differences in meteoric water composition, altitude, and meteorological factors such as temperature, humidity and wind speed (*Dansgaard*, 1964). Varying isotopic signatures among the water sources (precipitation, river, lakes, reservoirs, shallow and deep groundwater, etc.) makes SIW an effective tracer to understand and investigate natural and human-natural coupled systems, for examples please refer to (*Dawson and Ehleringer*, 1991; *Rozanski et al.*, 1992; *Gat*, 1995;

von Grafenstein et al., 1999; *Kennedy et al.*, 2011; *Klaus and McDonnell*, 2013; *Gabor et al.*, 2017; *Matheny et al.*, 2017; *Jameel et al.,* 2018).

In urban settings, stable isotopes and other geochemical tracers have been used successfully to understand effects of stormwater control measures on urban stream (*Jefferson et al.*, 2015), detect infiltration rates in urban sewers (*De Bénédittis et al.,* 2005; *Kracht et al.,* 2007), partition waste water and groundwater in urban sewers (*De Bondt et al.,* 2018) and determine the age of drinking water in a PWSS (*Waples et al.,* 2015). Recent studies have also shown that stable isotopes of tap water in urban areas can be used to characterize active water management practices, identify linkages between socioeconomic factors and water management practices, and quantify the effects of climate variability on water resources (*Ehleringer et al.,* 2016; *Jameel et al.*, 2016; *Tipple et al.*, 2017).

Here, we collaborated with the Jordan Valley Water Conservancy District (JVD, also referred as Jordan Valley District) to conduct an isotopic survey of waters from their service area within Salt Lake Valley metropolitan area (SLV) of northern Utah, USA (Fig. 1). JVD is a multi-source public water distribution network (Fig. 1) and we attempt to understand mixing between water sources at various sites (subsequently referred to as distribution sites) distributed on the transmission lines using SIW. This work extends the earlier work of *Jameel et al.,* (2016) and *Tipple et al.,* (2017) beyond identifying broad water management patterns of a PWSS and explores the capacity of SIW to provide quantitative, spatially- and temporally-resolved estimates of source contributions within a well-defined and characterized PWSS.

We conducted our study during a 6-month period (May 2015 – October 2015), and using information on the production volume from the different sources, we analyze the stable isotope data at a monthly resolution within a Bayesian framework to generate quantitative estimates (with uncertainty) of the contribution of individual sources at the distribution sites. These analyses reveal basic information on supply and transport dynamics within the system, reflecting the physical structure of the supply system and the geographic distribution of sources. Finally, we combine the monthly analyses to characterize the spatial structure of the system in terms of contribution areas for the different sources across the supply network. Our results suggest that SIW-based Bayesian isotope mixing models (BIMM) could be a powerful and useful tool to interrogate PWSS, provide observational validation to hydrodynamic models, track contaminants and disinfectants within the supply system, and monitor water rights in PWSS managed by or for multiple stakeholders. This technique can be particularly useful in understanding water management practices of urban centers in the developing world which are undergoing rapid expansions and are generally decentralized, which makes conventional hydrodynamic techniques difficult to apply.

Comment 3: L8. Figure 2 is introduced quite abruptly in the text, before the M&M section, even though it seems to me to present some results of this study. I suggest redefining its position in the structure of the manuscript.

- We have not referenced figure 2 in the introduction section and have introduced it in the result sections. It is now Fig. 4 in the updated manuscript.

Comment 4: In the M&M section, please specify in a clearer way the concept of prior and posterior fractional contributions (see also captions of Fig. 3, 4 and 5).

- We have added an additional paragraph in the method sections detailing the concepts of the prior and posterior distributions and how they were implemented in our paper. We have rectified the typos in the captions of figure 3, 4 and 5.

The edited paragraph in the revised manuscript (L 226 – 243 and L 264 – 267) is included below:

L 226 – 243:

In Bayesian analysis, the parameters to be estimated are initially assigned values (or distributions) that are believed to be the best estimate of the parameters. These initial values are referred to as prior values. After observing the data, the initial (or prior) values of the parameters are updated, and posterior estimates of the parameter values are obtained. Please refer to (*Parnell et al.*, 2010) and (*Hoff*, 2009) for more detailed information on Bayesian isotope mixing model and Bayesian Statistical methods, respectively. In our analysis, the parameters to be estimated were the fractional contribution of the different sources (*f's*) at the distribution sites. We therefore assign prior values to *f's* that are then updated using the observed isotope ratios at the distribution sites ($\mu\delta^2\mathrm{H}_I$, $\delta\mu^{18}O_I$) to obtain the posterior f estimates.

In general, if no information exists about the parameters, they are assigned a non-informative prior value. The default non-informative prior assigned to the Dirichlet distribution is the Jeffreys prior, where each element of the vector α is assigned a value of 1/*K* (with *K* being the number of sources) (*Parnell et al.*, 2010). However, if pre-existing information about the parameters exists, they are assigned informed prior values. In our case, we had existing information on the volumetric contribution of each source to the water system as a whole (obtained from JVD, Table 1) as well as the distance between the sources and the distribution sites, each of which, we assert, should affect the probability that a given source supplied water to a given site. Therefore, we assigned prior values for each supply site based on these information.

L 264 – 267:

Based upon the above-described method, the prior contributions of selected sources at the distribution sites for June 2015, in spatial and isotope space are shown in Fig. 2 and Fig. 3, respectively. The prior values were then updated using the observed isotope values at the distribution sites ($\mu\delta^2\mathrm{H}_I$, $\delta\mu^{18}O_I$) to obtain the posterior values of *f's*.

Comment 4: L208. These assumptions should be discussed.

- By "assumed" we mean that the isotope data was modeled as a bivariate distribution. The two assumptions in this are a) that the individual isotope data are normally distributed and b) that they are correlated. The histogram of the isotope data (δ^2H and δ^{18}O) suggests that they are mostly normally distributed. If we don't assume a joint (bivariate) distribution, we ignore the correlation between the isotopes. δ^2H and δ^{18}O are strongly correlated ($r > 0.8$ for all the months). By using a joint distribution, and its associated variance-covariance matrix, we capture both the individual variability and the co-variation. An actual assumption here will be to assume that δ^2H and δ^{18}O are independent.

- Further, most isotope mixing models assume isotope data are normally distributed and are modeled as univariate or bivariate normal distribution (for example: *Parnell et al.,* 2010).

- In this regard, we have added couple of sentences (see below) in the manuscript discussing the above points (L 205– 207 in the revised manuscript).

"The bivariate normal distribution accounts for the potential correlation between δ^2H and δ^{18}O. By using a joint distribution, and its associated variance-covariance matrix, we capture both the individual variability and the co-variation. "

Comment 5: The Hotteling multivariate t-test assumes a multivariate normal distribution: is the same distribution assumed at L209?

- Yes.

Comment 6: I suggest reporting d-excess values as indicators of evaporative effect, to corroborate these statements.

We have rephrased the sentence that suggests that deuterium excess values decreased from spring to fall. We have included the table showing the Deuterium Excess values as a supplementary table and have referenced it in the text. (L 308– 312 in the updated manuscript).

"JVTP isotope ratios increased from June to October, 2015, as did SETP isotope ratios from July to October, 2015, which can be due to evaporative enrichment of the heavy isotopes in upstream reservoirs of the Provo River system from spring to fall (mean d-excess for JVTP decreased from 5.1‰ in June 2015 to 3.9‰ in October 2015, Table S1)."

Month	Deuterium Excess values (per mil)
June 2015	5.1
July 2015	4.4
August 2015	4.2
September 2015	3.7
October 2015	3.9

Comment 7: L314. Please, be more specific: in what sense this work goes beyond previous work? This is part of the discussion on the originality and novelty of this research.

- We have rephrased the sentence to include the ways in which the work goes beyond the previous work.

Below is the excerpt from the updated manuscript (L 324 - 328).

"Our model builds upon the work of *Jameel et al.,* (2016) and *Tipple et al.,* (2017), but goes beyond their analyses of identifying district level water management patterns by providing quantitative, spatially- and temporally-resolved estimates of source contributions at specific locations throughout the SLV supply system."

Comment 8: L419. I suggest moving this paragraph to section 3.5.

- We have kept this paragraph in this section as we feel it is more relevant to this section. A couple of sentences in section 3.5 discusses the same concept.

Below is the excerpt from the revised manuscript (L 496 – 500).

"In our analysis, the isotopic compositions of major sources were distinct, allowing our model to quantify the contribution from the major sources at the distribution sites with robust estimates of uncertainty across the supply system. However, for distribution sites with isotope values intermediate to candidate sources (group 3, Fig. 6) the specificity of our result was limited by non-unique solutions."

Technical corrections and minor comments

Comment 1: L45. Please, specify what these particular threats are.

- We have included the possible threats in the paper (supply contamination, infrastructure failure, etc.). L 34 in the revised manuscript.

Comment 2: L91. The reference to Gorski et al., 2015 applies to vapour. Is still relevant?

- We have removed it from the reference list.

Comment 3: L122. Please, specify that the study was carried out in the USA for the benefit of the international readers.

- We have included the information (Line 109 in the revised manuscript).

Comment 4: Why? To avoid overexploitation? Please, specify.

- Over exploitation could be one of the possible reasons for well rotation but there could be various other reasons such as contractual agreements between the JVD and other water districts. We have rewritten the sentence (included below, L 120 - 122 in the updated manuscript)

"Not all wells operate simultaneously, rather only 2-5 wells operate at any given time and the operating wells are rotated every few months due to contractual obligations and to avoid overexploitation."

Comment 5: L129. Is this especially true for irrigation purposes?

- The increased demand is due to irrigation as well as increase in municipal water demand. Municipal water consumption in summer months are more than double of winter consumptions. We have rewritten the sentence (included below, L 115 in the updated manuscript) and included the word "municipal"

“….however, during the summer season (June – August) an additional 5-7 sources are often used to meet increased municipal water demand (personal communication, JVD operations manager).”

Comment 6: L138. Please use SI units.

- We have used SI units in the revised manuscript.

Comment 7: L146. Why do these well have so high salt concentrations?

- Prior contamination due to mining. The southwestern part of Salt Lake Valley is very close to the Kennecott Copper Mine (one of the largest open pit mines in the world).

Comment 8: L163: Combine with L174.

- We have restructured the paragraph (from L 160 in the updated manuscript)

“Distribution sites and surface water sources (Provo River and Wasatch creeks) were sampled 1-3 times per month. Groundwater wells were sampled 1-5 times respectively, during the entire study period (May 2015 to October 2015). When a given well was not sampled in its month of operation, the mean value observed for the same well during other month(s) of our study period was used to characterize water supplied from that well.”

Comment 9: L184. Specify the sample size

- We have included the information in the updated manuscript (n = 9).

Comment 10: L199. The acronym has been already defined at L115

- We have rewritten the manuscript with only the acronym.

Comment 11: L201. Information, such as?

- We have added a list of possible prior information. Line 195 in the revised manuscript.

“….inclusion of prior information into the statistical analysis (such as those from previous studies, ancillary data and subjective knowledge)…”

Comment 12: L345. Typo: to our model. L355. Typo: lines.

- We have corrected the typos.

Comment 13: L380. Can you explain why you used credible intervals and not confidence intervals?

- Since Bayesian statistics is fundamentally different from frequentist statistics, credible intervals are the Bayesian equivalent of the confidence interval. A credible interval is defined as an interval in which the parameter to be determined has a given probability.

Comment 14: L396. Replace “credible intervals” with “CI”.

- We have replaced credible intervals with CI.

Comment 15: L513. I doubt cities in developing countries have to fund to perform costly isotope analysis over a long time span or large areas. Perhaps add a comment here.

- With the advent of laser spectrometers stable isotope analyses have become relatively easy, robust, fast and cheap. We analyzed more than 1000 samples on Picarro laser spectrometers for this study and the total cost of sample analysis was around $7000 ($7 per sample). Preliminary analysis and pilot study on a medium sized city can be conducted with a budget of less than $2000. Continuous long term measurements will indeed be costly but careful planning and sample optimization can lead to reduction of costs. The last sentence in the conclusion now discusses this issue.

Below is the excerpt from the manuscript (L 532 -534).

> Considering that stable isotope analysis of most water samples is now rapid (minutes) and inexpensive, geochemically-based BIMMs offer an attractive tool for studying and monitoring PWSS in support of management and water security.

Comment 16: Fig. 1. In the caption define WTP and SWTP.

- We have defined them in the figure caption.

Comment 17: Fig. 7, L780. Replace panel a with (a).

- Done.