Examining the premise of the "LakeFlow" discharge algorithm in the contiguous United States

by

Aote Xin

B.Eng., Nanjing Tech University, 2014

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Geography and Geospatial Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Approved by:
Major Professor
Dr. Jida Wang

Copyright

Abstract

Lakes and reservoirs are essential storage units of the global drainage system. These stores influence downstream river discharge through detaining and releasing water mass at their interface. LakeFlow is a proposed algorithm to estimate discharge at the lake-river interface for the forthcoming Surface Water and Ocean Topography satellite mission (SWOT). With the direct measurements provided by SWOT, LakeFlow bridges the storage change to an improved parameterization of the inflow and outflow channel properties that are not observable to SWOT but necessary for discharge calculations. The success of LakeFlow relies on a wide existence of the premise, which is that the discharge difference between the primary inflow and outflow river reaches is the first-order control on storage variation. However, additional hydrologic controls and anthropogenic factors, which are not directly observable to SWOT, could also be substantial and therefore undermine the LakeFlow's premise. For this reason, we have examined the premise conditions for 26 reservoirs in different hydroclimate conditions across the contiguous United States (CONUS), and further constrained the uncertainties by accounting for lateral flow and surface evaporation through a synergy of multi-source remote sensing and modeling datasets. When comparing reservoir storage change with discharge difference between the reservoir's inflow and outflow river reaches for a two full year, 18 out of 26 reservoirs have a Nash-Sutcliffe efficiency (NSE) greater than 0, exceeding the mean flow benchmark. 12 out of the 26 reservoirs have an NSE greater than 0.5, indicating a satisfactory premise. The majority of these reservoirs with a satisfactory premise are located in less humid regions. By comparing reservoirs that were and

were not designed for water supply, our results suggested that human water withdrawal is probably the most significant contributor to a less satisfactory premise. As we expect an overall lesser extent of human water consumption from natural lakes, our results based on reservoirs may serve as a minimum baseline for the actual feasibility of LakeFlow. Simulated lateral flow from the Global Reach-scale A Priori Discharge Estimates for SWOT (GRADES) demonstrates a high effectiveness in improving the premise condition, particularly in humid regions. The inclusion of surface evaporation estimates from the CONUS Reservoir Evaporation Dataset (CRED) dataset appears to be most effective in drier regions. Coupled with both surface evaporation and lateral flow datasets, the LakeFlow premise conditions are improved by a median value of 27.4% for reservoirs in more humid region, and 74.5% in reservoirs that were not built specifically for water supply. These results support the overall feasibility of LakeFlow and motivate continued testing on the accuracy of reservoir inflow/outflow estimations.

Table of Contents

List of Figures

List of Tables

List of Equations

Acknowledgements

For this master's project, I would like to devote my highest appreciation to Dr. Jida Wang for being a crucial part to this project from the grand idea initiation, method design, supervision, manuscript review and quality control. Dr. Wang is an exceptional research scientist, and I want to thank him for all the help he has offered to the completion of this project and to my learning at Kansas State University. It has been a rewarding experience having Dr. Wang as both my mentor and advisor. Looking ahead, I am confident that the things I learn from Dr. Wang, both as an individual researcher and as a colleague, will be carried with me in the long haul.

I would like to thank my committee members, Dr. Abigail Langston for the crucial suggestions and comments on the project, and Dr. George Allen for processing and providing lateral flow simulations from GRADES, making critical suggestions on method improvement and commenting on the thesis draft. I am very glad and grateful for the opportunity to have Abby and George as my committee members.

I would like to thank Dr. Colin Gleason in the Department of Civil and Environmental Engineering at University of Massachusetts Amherst for the critical comments on this research project during the American Geophysical Union fall meeting.

I want to acknowledge the Steve Kale Summer Research Fellowship and NASA's SWOT Science Team grant to support my research activities during the summers of 2020 and 2021, respectively. I am also grateful for the department's Teaching Assistantship during most of the time in my M.S. pursuit.

I want to thank My girlfriend Wang Jiaqian for being the cornerstone of my life, and my cat Lulu for the invaluable companionship during quarantine. My American family (Leon and

Jennie Brown and their wonderful family) raised me up, welcomed me to their family and gave me a home. Outside of the world of research and study, their support is the most important element in helping me to complete the thesis.

It was an absolute pleasure to work as a Research Assistant for Dr. Angela Y.S. Park in the Department of Political Science and her amazing research team. I want to thank Angela for her guidance and leadership both in life and research.

I am truly happy for being a part of the Geography community at Kansas State University. I am proud to be a K-State Wildcat and grateful for having my master's study completed at the Department of Geography and Geospatial Sciences. There are people among the graduate students and faculty members from the department, who I became best friends with, who I constantly align myself to as my life model, and who went out of their ways to offer me with the emotional support that I needed the most. Without a question, and I firmly believe that my study at Kansas State University has prepared me well to become a confident physical geographer and a GIS professional in either academia or industry.

My experience is graduate school is awesome and life-changing through the amazing people I have the pleasure to work with and meet with. Most importantly, it has been a humbling experience that, at times, pushed and motivated me to become more resilient. If I ever come back to read my thesis again, I would like to remind myself of the most important things in every single day from what I have learnt from the completion of my master's thesis: 1. The love and passion for the work, 2. The right working ethics and spirit, 3. The positive attitude, and 4. Always having the right people next to me.

Chapter 1 - Introduction

1.1 Background

1.1.1 Human Impacts on Rivers

Freshwater is an important natural resource for terrestrial life forms. Human civilization has long relied on freshwater resources to secure food and energy supply, maintain environmental health, and sustain societal, economic and industrial development for a fast-growing population (Lehner et al., 2011; Oki and Kanae, 2006; Rodell et al., 2018; Siebert et al., 2010). Being an essential hydrologic component in the water cycle, rivers route approximately 40% of the global precipitation over the continents back to the ocean (Oki and Kanae, 2006). Despite covering only 0.58% of the Earth's glacial-free land surface (0.35% for rivers wider than 30m), rivers play an essential role in controlling the land-atmosphere fluxes (Allen and Pavelsky, 2018). For instance, rivers are considered to be a hotspot for carbon exchange, having a substantial contribution to the global carbon budget despite their limited surface area (Raymond et al., 2013).

The natural flows of rivers have long been harnessed and regulated by humans for flood control, irrigation, navigation, domestic and industrial water supply and energy generation. Globally, 63% of the rivers that are wider than 1000km are classified as non-free-flowing rivers, mainly due to an ongoing increase in human interference (Grill et al., 2019). Around rivers across global basins, human interference can be categorized as five pressure factors: (1) river fragmentation; (2) flow regulation; (3) sediment trapping; (4) water consumption; and (5) infrastructure development (Grill et al., 2019). Dams and reservoirs as the end product of human hydraulic engineering can be found across the world's densely-populated regions. Most developed countries, notably the United States, have witnessed a boom in the number of large dams throughout the 20th century (Graf, 1999; 2006). The developing countries or regions, notably India, China, Southeast Asia, South America and Africa, are currently undergoing a rapid increase in the total number of dams (Grill et al., 2019; Lehner et al., 2011; Wada et al., 2017; Zarfl et al., 2014).

Dams alleviate the stress from a growing domestic and irrigation consumption of freshwater resources, sustain energy supply for nearby settlements, and provide services like flood control, navigation and recreation. Despite these benefits to the human societies, dams as in-stream barriers, pose an uncertain threat to the livelihoods of the local aquatic species by stopping the nutrients from moving downstream, reshaping the geomorphic conditions of aquatic habitats, and keeping them from migrating upwards or downwards to feed and to reproduce (Graf, 2006; Wellmeyer et al., 2005). The impact of dams on regional hydrology was documented in previous research: the dam construction decreased the natural lake inundation area across the basin, lowered the levels of downstream lakes and surrounding groundwater tables, and resulted in a compensation of sea level rise (Chao et al., 2008; Wada et al., 2017; Wang et al., 2014; Wang et al., 2013). Though one study has suggested that compared to dam construction, climate variability could be the dominant driver of the observed decadal lake decline across Yangzte Plain, China from 2000 to 2011 (Wang et al., 2017). Dams would keep on leaving even larger footprints at a rapid pace due to an increase in dam construction in the developing world (Grill et al., 2019; Wada et al., 2017). Dams also place an ecological impact and affect water quality in rivers and lakes (Winton et al., 2019). Its impact on increasing surface evaporation loss has also been implied or documented in previous studies (Wang et al., 2018; Woolway et al., 2020; Zhao and Gao, 2019; Zhao et al., 2022).

Hydrologists and fluvial geomorphologists have greatly advanced our understanding of the various impacts of dams on rivers. Streamflow monitoring at the interface between reservoirs and river reaches provides the necessary measuring data to reveal changes on river conditions through time. By comparing the pre- and post-dam flow conditions based on in situ measurements, Wellmeyer et al. (2005) argued that dams alter the downstream flow regime mainly by changing the magnitude and timing of low and peak flows through detaining and releasing water mass at their interface. Specifically, dams reduce levels of peak flows and increase levels of low flows. For 36 large dams, the reduction rate for instantaneous maximum flows could reach 67% while 30-day minimum flows could increase by 52% by comparing gauge records between regulated and unregulated reaches. The timing of low and peak flows is highly affected by dam operation for flood control and water supply (Graf, 2006). Lajoie et al. (2007) later argued that the larger the watershed, the greater inter-annual variability of the timing of maximum monthly discharge. The main geomorphic impact from dams on downstream rivers/channels usually includes the increased downcutting from a response to a “hungry water” condition and planform changes where channels become more stabilized and less degree of lateral migration might occur (Graf, 2006; Shields Jr et al., 2002).

1.1.2 River Discharge

Discharge is defined as the volume of water moving down a channel at a given point per unit time, allowing the water resources in rivers to be measured as volumetric fluxes (Anderson and Anderson, 2010; Mueller et al., 2013; Turnipseed and Sauer, 2010). Streamflow is highly subject to the intensity, duration and frequency of local rainfall events, and is also related to specific geomorphic conditions, human activities and land cover types of the surrounding watershed. Watershed geomorphology considers the physical characteristics of a watershed as a factor that could affect streamflow. Basin features in watershed geomorphology like watershed area, length, shape and drainage density are related to the amount of rainfall that the watershed could receive, the arrival time for precipitation to become streamflow, and the time of concentration (i.e., the length of time it takes for runoff in a catchment to reach the outlet). Land cover conditions affect the streamflow by changing the peak flow level and the time for peak flow to occur. For instance, increasing the impervious surface in urban areas would increase the intensity of flood events, adding more strain to the city infrastructure (Anderson and Anderson, 2010; Chin, 2012). Existing studies (Chai et al., 2019; Li and Jin, 2017; Oki and Kanae, 2006; Syed et al., 2010) also suggest that climate change and climate variability are also driving the global discharge to become more variable.

Reservoir inflow estimation and inflow forecast are critical to water resource management and hydropower operation (Gragne et al., 2015). Reservoir outflow estimation is critical to human and ecosystem health, reservoir operations and water resources management in downstream regions. Flow records from upstream reservoirs inform downstream water agencies as to the timing and quantity of water that can be available (Biancamaria et al., 2015; Bonnema et al., 2016a). In transboundary basins, sharing of these data is, however, not always viable due to political concerns or poor institutional capacity. With thousands of new dams planned for construction in the near future, there is an imperative need to monitor the transboundary reservoir outflow in these regions. Reservoirs in exorheic basins detain streamflow from draining to the ocean, substantially increasing the residence time of terrestrial water (Hosen et al., 2021; Vörösmarty and Sahagian, 2000; Vörösmarty et al., 1997). An improved understanding of the roles of dams and reservoirs in the global water cycle requires an improved estimation of reservoir outflow at continental and regional scales.

1.2 Literature Review

We hereby provide a short review of the methods and tools that have been applied to measure river discharge. These methods include the traditional field-based methods, hydrological modeling and several discharge algorithms which incorporate direct measurements from satellite remote sensing. Followed by the introduction, we will introduce some of the discharge algorithms that specifically apply the direct measurements from the forthcoming Surface Water and Ocean Topography (SWOT) satellite mission. A comprehensive and systematic review of methods and tools applied for discharge estimation is available in Gleason and Durand (2020).

1.2.1 Estimating River Discharge – Field-Based Methods

Field-based methods and technologies yield the most accurate estimates of river discharge. Because of their high accuracy, these in situ estimates are often seen as "measurements" and are used as the baseline data for model calibration or other applications that require highly accurate discharge data. But intrinsically they are still approximations of true discharge values.

Acoustic Doppler Current Profilers (ADCP) and weir equations are often used to measure river discharge in the field setting (Gleason and Durand, 2020). The United States Geological Survey (USGS) has relied principally on ADCPs to attain discharge measurements (Turnipseed and Sauer, 2010). Usually stationed on a moving boat, the ADCP measures the Doppler shift of acoustic signals to derive the flow velocity and then the discharge (Mueller et al., 2013;

Turnipseed and Sauer, 2010). As river discharge can also be written as a product of cross-section area and flow velocity, in a method similar to ADCP proposed by Costa et al. (2000) and Costa et al. (2006), the river discharge is computed by measuring the channel cross-section area with ground-penetrating radar, and the flow velocity with a pulsed Doppler radar.

Stream gauges transform the stage of the river surface into an estimation of discharge, often via an empirically derived rating curve (Leopold and Maddock Jr, 1953). These site-specific rating curves were carefully calibrated before application with a high-precision field instrument, usually ADCP. The ADCP discharge measurements were used to estimate discharge with stage measurements according to an observed stage-discharge relation (i.e., a site-specific rating curve). The flow and channel data from the USGS stream-gaging network are publicly available on the USGS National Water Information System (Turnipseed and Sauer, 2010). Despite being seen as more accurate estimates, in situ measurements have several disadvantages, making them inappropriate or insufficient in some scientific studies. The most notable disadvantage is its spatial coverage. Essentially, in situ measurements are spatially discrete point estimates of discharge. The cost of instrumenting and then maintaining these stream gauges is high, which has resulted in an uneven spatial distribution of stream gauges where more can be found in more populated and more developed regions. This has left ungauged basins like the Tibetan Plateau or some other transboundary basins to be the gray area for discharge monitoring. In addition, there has been an ongoing demise of stream gauge networks on a global scale (Gleason et al., 2018; Gleason and Wang, 2015). These disadvantages of gauge measurements have challenged the scientific community to come up with approaches in estimating discharge that could potentially complement the measurement gaps within the field-based methods.

1.2.2 Estimating River Discharge – Water Balance and Hydrological Modeling

Water balance lays the foundation of many hydrological models to simulate streamflow, in which these models parse the hydrologic cycle into components where the major fluxes occur (Arnold et al., 1998; Beven and Kirkby, 1979; Burnash et al., 1973; Gleason and Durand, 2020; Liang et al., 1994; Samaniego et al., 2010). In situ measurements and remote sensing data can be used to tune the model parameters for calibration. For a drainage basin, oftentimes in hydrological modeling, the total streamflow (i.e., the amount of water that leaves a basin through rivers and streams) is balanced out by the precipitation, evapotranspiration and basin storage change. In some modeling methods, before routing out to an entire river network, the streamflow is derived from three components: precipitation excess (surface runoff), interflow (subsurface runoff) and baseflow. During a rainfall event, part of the precipitation (i.e., precipitation excess) flows above the surface, reaches the channel and eventually becomes the streamflow. While the rest of the precipitation, after being intercepted by vegetation surfaces (e.g., tree leaves or bushes), enters the soil through infiltration, and reaches the channels by becoming interflow and baseflow (i.e., groundwater discharge). Interflow occurs in the unsaturated zone (i.e., above the water table surface but below the land surface). Baseflow is the movement of groundwater (i.e., below the water table surface) towards the channels. For interflow, it might take days to years to become streamflow, and weeks to millennia for baseflow to become streamflow.

With advances in computational capacity and improvement on baseline topological dataset, Lin et al. (2019) used hydrological modeling to estimate the global river discharge at 2.94 million reaches for a total period of 35 years. The Global Reach-level A priori Discharge Estimates for Surface Water and Ocean Topography (GRADES) dataset was made available with the top-of-the-line global raster hydrographic data, MERIT Hydro (a novel hydrography map based on the Multi-Error-Removed Improved-Terrain DEM (MERIT-DEM)), at 3-arcsec resolution (~90m) (Yamazaki et al., 2019). The Variable Infiltration Capacity (VIC) model was used in GRADES to model the land surface processes (Liang et al., 1994). Routing Application for Parallel Computation of Discharge (RAPID) was used for river network routing (Lin et al., 2019). RAPID is based on the traditional Muskingum method for river routing, and can be used to calculate the flow at any reaches in a river network for hydrological modeling (Eijkhout et al., 2011).

Bonnema et al. (2016a) used a model-based mass balance approach with remotely-sensed storage variations to estimate reservoir outflow, and assessed the sensitivity of each hydrologic control. The reservoir outflow was balanced out with reservoir storage variation, surface evaporation and reservoir inflow. They derived the reservoir inflow by using the curve number (CN) method and the Variable Infiltration Capacity (VIC) model and surface evaporation from an energy-balance model (Liang et al., 1994; Te Chow, 2010; Woodward et al., 2003). Reservoir storage change was obtained from water level measurements from remote sensing and reservoir elevation-area relationship from Shuttle Radar Topography Mission (SRTM) DEM (Bonnema et al., 2016a; Liang et al., 1994). This method was tested on two reservoirs, with the most accurate having a rRMSE of 46.8%. There are many other hydrological models but they are not the focus of this study.

1.2.3 Estimating River Discharge – Remote Sensing Methods

Enabled by satellite remote sensing, the improved spatial and temporal coverages in our ability to observe the Earth have advanced our understanding of the Earth's systems, which made it possible to utilize remotely-sensed hydraulic information to estimate river discharge. Several methods are based on the basic concept that some remotely-sensed hydraulic variables (e.g., river stage, river with and inundation area, etc) correlate with river discharge (Bjerklie et al., 2005b; Kouraev et al., 2004; Smith et al., 1996; Smith et al., 1995; Smith and Pavelsky, 2008). As a response to an increase in discharge, rivers of various conditions (i.e., vegetation types, sediment types, bank material, channel cross-sectional shape, etc) might undergo a mix of channel processes which might result in a change in inundation area, river stage and river width. The performance of these methods depends on the sensitivity of this hydraulic information to changing discharge, and the availability of stream gauge data for constructing and calibrating the rating curves. (Alsdorf et al., 2007; Kouraev et al., 2004; Smith et al., 1996; Smith et al., 1995; Smith and Pavelsky, 2008). Similar to how stream gauges work, a site-specific rating curve that characterizes the width-discharge relationship or stage-discharge relationship was carefully calibrated, and then used with remotely sensed hydraulic variables to estimate discharge. For example, Smith et al. (1995) and Smith et al. (1996) used Synthetic-Aperture Radar (SAR) to measure the widths of the braided glacially fed rivers in Alaska, which yielded satisfactory estimates of river discharge. In these braided rivers with easily eroded banks, the effective river width (the total inundation area in a given reach divided by the reach length) was shown to be sensitive to the changes in river discharge, thus making it possible to estimate discharge for braided rivers. In summary, these methods are intrinsically empirical and rely heavily on high-quality in situ data to parameterize such relationships between hydraulic variables and discharge, making it impractical in remote regions where such in situ data are not available. The signal quality of remote sensing takes a strong hold on the final discharge estimates. The curves are also specific to channel conditions. It remains uncertain if the curve for one channel cross-section could be applied to that for another. In addition to the rating-curve method, instead of directly measuring the channel surface area, Brakenridge et al. (2007) provided an alternative to estimate discharge via a remote sensing estimator from satellite passive microwave measurements. The general concept of this method is based on the difference in emissivity between water and land. This estimator is the radiance ratio of two parcels: a “measurement” parcel comprised of water pixels centered over a river, and a “calibration” parcel comprised of land pixels adjacent to but unaffected by the local river. An increase in river discharge results in an expansion of water area in a “water parcel”, thus increasing the remote sensing estimator. The remote sensing estimator was shown to correlate well with the river discharge in several sampled rivers (Brakenridge et al., 2007).

Using a C and X band interferometry, the SRTM satellite mission provides the world’s first image-based global Digital Elevation Model (DEM) of about 80% of Earth’s land surface, all land between 60° north latitude and 56° south latitude at 3 arcsec resolution (~90m) (Farr et al., 2007). The incidence angle of the X band radar payload is a fixed 54° and the angle for the C band radar payload is in the range of 30° to 58°. At these angles, much of the backscattered energy from the water surface is lost due to specular reflection, leading to poor measurements of surface water. SRTM lays the foundation for global satellite interferometry, but many existing missions like itself are not suitable for observing surface water dynamics or flow. Under this reality, the forthcoming Surface Water and Ocean Topography (SWOT) satellite mission was proposed, jointly developed by the National Aeronautics and Space Administration (NASA) and Centre national d'études spatiales (CNES), the French space agency, in partnership with the Canadian Space Agency (CSA) and UK Space Agency (UKSA). A Ka band Radar interferometer (KaRin) was chosen as SWOT’s main payload (Biancamaria et al., 2015). A Ka band comes with several advantages: it has a higher frequency in the microwave spectrum, allowing for finer spatial resolutions than low-frequency bands in the SWOT’s design

(Biancamaria et al., 2015). KaRin has a near-nadir incidence angle (0.6°-3.9°) which makes it particularly suited to monitoring surface water. Almost as a specular reflector, water reflects much of the backscattered energy emitted from SWOT back to SWOT, while land, as a rougher surface, backscatters it to all directions. This distinction is used to extract water masks from satellite images. In addition, SWOT can observe all continental surfaces between 78°S and 78°N latitude, capable of covering some mid-high latitudes where surface water is most abundant, while SRTM comparatively has a limited spatial coverage which is between 60°S and 60°N.

There are two antennas on KaRin, each at an end of a 10-meter baseline. During operation, one antenna emits electromagnetic signals and two antennas receive the backscattered signals. Therefore, each point on the ground surface on Earth will be observed from two different positions. By comparing the phase difference, the elevation of the point is resolved.

SWOT has immense hydrologic values and provides new possibilities for method innovation in river discharge estimation with remotely sensed hydraulic variables. SWOT will enable the first ever direct and simultaneous observation of inundation extent (W), height (H) and slope (S) at unprecedented spatial and temporal resolutions (Biancamaria et al., 2015). It has a 15% relative errors on its width measurements, a less-than 10 cm height accuracy and a centimeter-level slope accuracy. For its temporal coverage, SWOT has an orbit repeat period of about 21 days but could provide more than one overpass within a 21-day repeat period over continents between 78°S latitude and 78°N latitude. The number of revisit times ranges from a maximum of 2 at the equator to more than 10 above 70°N or 70°S, making it suitable for observing sub-monthly surface water dynamics. For its spatial coverage, SWOT could observe water bodies greater than $(250m)^2$ and rivers wider than 50 to 100m.

Several discharge algorithms have been developed over the past decade based on SWOT's measuring capabilities (Bonnema et al., 2016b; Brinkerhoff et al., 2020; Durand et al., 2014; Gleason and Smith, 2014; Gleason et al., 2014; Hagemann et al., 2017). As neither channel cross-section area nor flow velocity is directly observable to SWOT, these algorithms share the same basic methodology: to transform the remotely sensed hydraulic variables (W, H, S) to discharge through flow laws. Gleason and Smith (2014) applied a river-specific At-Many-station Hydraulic Geometry (AMHG) relationship to estimate discharge with satellite remote sensing. As a geomorphic scaling phenomenon fundamental to natural rivers, AMHG describes a correlative relationship between At-a-station Hydraulic Geometry (AHG) parameters from multiple cross sections along a reach, reducing the number of site-specific unknown parameters in a geomorphic relationship. Gleason and Wang (2015) further complemented AMHG by providing its mathematical basis and suggested using the percentage of intersected rating curves as the indicator of the existence of a strong AMHG relationship. Hagemann et al. (2017) presented the Bayesian AMHG-Manning (BAM) algorithm, which combines Manning's equation and the AMHG in a synergistic framework. This approach is "statistically defensible" under a Bayesian mathematical setting, quantifying the uncertainty in the estimates like discharge. BAM algorithm showed improved overall performance over 19 rivers with synthetic and model-derived data (Hagemann et al., 2017).

Similarly, Durand et al. (2014) proposed the Metropolis Manning (MetroMan) algorithm which uses a derivative of Manning's equation to estimate river discharge. MetroMan has been tested on a variety of rivers, and has proven to be reliable among several discharge algorithms (Durand et al., 2016; Durand et al., 2014; Tuozzolo et al., 2019). Manning's equation is an empirically-derived hydraulic relationship to estimate flow velocity and river discharge. The equation consists of a cross-section component and a flow velocity component, which is written as follows:

$$Q = A\frac{1}{n}H^{2/3}\,S^{1/2}$$

Equation 1 Manning's Equation

Where: A is the cross-section area of the river; n is the Manning's roughness coefficient or a flow resistance term, reflecting whatever the water is interacting with; H is the river stage or depth; and *S* is the river surface slope or gradient. n, *H* and *S* make up the mean flow velocity of the channel. SWOT is only able to observe changes in cross-section area A and river stage *H* and is not able to observe the cross-section area below the lowest observed river elevation. For convenience, they write the manning's equation and divide the cross-section area into two components, one that is observable to SWOT, and one that is not:

$$Q = \frac{1}{n}(A_0 + \delta A)^{5/3}\,W^{-2/3}S^{1/2}$$

Equation 2 Manning's Equation in "MetroMan"

Where: A_0 is the cross-sectional area at the lowest observed river elevation (i.e., a product of river stage and bathymetry), δA is the change in the cross-sectional area relative to A_0, and W is the river width. All these quantities are averaged across a reach length. In this equation, δA, W and *S* can be either directly observed by SWOT or calculated from the directly observed terms. These three quantities are therefore taken as the known variables in the equation. Baseflow cross-section A_0 and roughness coefficient *n* are 2 channel properties not observable to SWOT, therefore taken as the unknown variables in the equation. It was also assumed that A_0 and *n* do not vary in time.

Naturally, the 2 unknown variables (i.e., A_0 and *n*), also the channel properties not observable to SWOT, need to be parameterized before they can be used with SWOT observations to calculate discharge. MetroMan parameterizes the channel properties by estimating the unknown variables (A_0 and n) from the known variables (δA, W, S) via a flow law mass-conservation equation. A Metropolis algorithm was used in a Bayesian approach to calculate the best estimate of 2 unknown variables with their associated error distributions. This technique requires a prior estimate or namely the first guess of A_0 and n to initialize the iteration. Once n and A_0 are estimated, the discharge could be calculated in a Manning's equation with δA, W, and S directly observed by SWOT. MetroMan has later been updated to use a prior estimate of mean annual flow and n. Its feasibility to estimate discharge was demonstrated in two comparison studies (Durand et al., 2016; Tuozzolo et al., 2019).

Previous studies suggested that the final discharge estimate product from MetroMan and other SWOT discharge algorithms were highly sensitive to the selection of the prior estimates (Durand et al., 2016; Durand et al., 2014; Hagemann et al., 2017; Tuozzolo et al., 2019). More specifically, the parameterization of river properties that are not observable to SWOT would largely determine the accuracy of the final discharge estimates. These studies have shown that the performance of the algorithms substantially decreased without the integration of a priori data from in situ measurements or hydrologic modeling. This has posed a challenge to these discharge algorithms in ungauged basins were limited to no a priori knowledge is available.

Due to human water management, the interface discharge in vicinity of large water stores, such as lakes and reservoirs, would be very different from the free-flowing open-channel discharge. Like MetroMan and BAM, most core algorithms were not designed specifically for the interface between river reaches and water stores, nor did they take advantage of the river-store mass conservation. These limitations call for future improvements on discharge algorithms to account for the river-store interactions. There are also many other remote sensing discharge algorithms.

1.2.4 Integrating River-Store Mass Conservation to Discharge Estimation

Wang et al. (2019) recently proposed "LakeFlow", which utilizes the mass conservation of lakes or reservoirs and their inflow and outflow river reaches to estimate discharge. It is a generic concept that could be applied to other discharge estimation frameworks, like MetroMan, to potentially improve estimation accuracy. Without using additional a priori information, this algorithm aims to improve the discharge estimation around global lakes and reservoirs by bridging their storage change, which is easy for SWOT to accurately observe, to an improved parameterization of the channel properties of the inflow and outflow river reaches. The two key elements in this algorithm are Manning's equation and the mass conservation between the water stores and their major connecting rivers (hereafter referred to as "the river-store mass conservation"). The derivative of Manning's equation used in MetroMan is adopted for this novel algorithm.

The amount of water that a surface store (i.e., a lake or a reservoir) receives as inflow subtracted by the amount of water it loses from its outflow reaches should approximately equal storage change in a given period time. The remaining residual includes some hydrologic controls like lateral inflows, surface evaporation and, etc. For artificial water stores like large reservoirs, water consumption (e.g., water supply, irrigation) could be a major hydrologic output. The anthropogenic impact on artificial impoundment mainly comes from our growing needs for water supply, and such impact could be paramount or trivial for reservoirs serving their purposes of design. For a natural lake where such human impact is absent, the storage change could be more closely balanced out by the flow difference between a lake's major inflow and outflow reaches multiplied by time. A river-store mass conservation relationship can be described as follows:

$$dV = \int_{t=0}^{p} (Q_I - Q_O)_t + \varepsilon$$

Equation 3 River-Store Mass Conservation

Where: dV is the net storage change of a surface store in a time period p. As a product of lake surface area changes and level changes,dV could be easily calculated with SWOT's direct measurements of surface water; Q_I and Q_O notate the instantaneous flow for the store's SWOT-observable inflow and outflow river reaches during p. As SWOT can observe rivers wider than 50-100m, any river reaches that meets the measurement requirement would be considered in representing Q_I or Q_O; ε is a residual term reflecting the total influence from other hydrologic controls (e.g., surface evaporation and lateral flows) and anthropogenic controls (i.e., direct human water withdrawal).

To bridge the equations from Manning's and river-store mass conservation together, we substitute the flow terms (one for inflow and one for outflow) in the river-store mass conservation with Manning's equation. Assuming there is only one SWOT-observable inflow reach and one SWOT-observable outflow reach, this substitution of flow terms would yield:

$$dV = \int_{t=0}^{p} \left(n_I^{-1} \left(A_{0,I} + \delta A_I \right)^{5/3} W_I^{-2/3} S_I^{1/2} - n_O^{-1} \left(A_{0,O} + \delta A_O \right)^{5/3} W_O^{-2/3} S_O^{1/2} + \varepsilon \right)_t$$

Equation 4 River-Store Mass Conservation in the form of Manning's Equation

Where: the storage change dV and the residual term ε stay the same; n_I, $A_{0,I}$, δA_I, W_I and S_I notate the roughness coefficient, cross-section area at the baseflow level, change in cross-section area, surface water width and surface water slope of the store's inflow river reach. n_O, $A_{0,O}$, δA_O, W_O and S_O notate the same set of hydraulic variables and channel properties of the store's outflow river reach. With each of the overpasses, SWOT could observe dV, δA, W and S for the inflow and outflow reaches once. This leaves the unknown variables in the equation to be each reach's channel properties that are not observable to SWOT (n_I, n_O, δA_I, δA_O) and a residual term ε. These channel properties are considered to not vary in time. Supposing that ε which reflects the total influence of other hydrologic and anthropogenic controls is significantly smaller than the discharge difference between inflow and outflow reaches, we would be able to solve the equation of 4 unknown variables with the same number of SWOT overpasses given ideal observation conditions. Once the channel properties of inflow and outflow reaches have been solved, the algorithm could provide discharge estimates for inflow and outflow reaches with Manning's equation at each SWOT's overpass.

In this novel algorithm, Manning's equation is modified to account for the river-store mass conservation. By providing an observation-based solution to the channel properties that are not observable to SWOT, this algorithm, which does not necessarily rely on but allows for the incorporation of a priori knowledge in channel parameterization to constrain the estimates of parameters, has great potential in improving discharge estimation around global water stores in both gauged and ungauged basins.

1.2.5 The Premise of LakeFlow

The river-store mass conservation takes a crucial part in the channel parameterization for LakeFlow (Wang et al., 2019). While existing algorithms depend on the SWOT's observation of hydraulic variables of the channels, this method takes better use of SWOT's hydrologic capabilities in capturing surface water dynamics. Ideally, the SWOT-observable inflow and outflow reaches are in close vicinity of water stores to allow the discharge difference between these reaches (i.e., dQ or $Q_I - Q_O$) to be the first-order control on storage change. Only when the scale of ε is significantly smaller than SWOT-observable dV, can the magnitude and variation of dV be significantly captured by dQ.

Surface evaporation could largely influence the scale of ε. Following current trends of global warming, in the future, evaporative demand could contribute to a global drought expansion independent of precipitation changes (Cook et al., 2014; Dai, 2012). By estimating the evaporation loss from Landsat-derived reservoir areas and pan-derived evaporative rates, Zhang et al. (2017) suggested that the annual evaporation loss in 200 major reservoirs in Texas was equivalent to 20% of their total active storage. Through remote sensing and modeling, Zhao and Gao (2019) implemented the heat storage change into the Penman equation and provided improved estimates of the monthly evaporation loss on 721 reservoirs from 1985 to 2014 across the Contiguous United States (CONUS), which account for 90.2% of the large reservoir storage capacity in the CONUS. They proposed a Reservoir Evaporative Losses Index (RELI) to evaluate the role of evaporation in the total water losses (i.e., the sum of evaporation, water withdrawal and other water losses) of reservoirs. The 30-year annual average RELIs for Lake Mead in Nevada/Arizona and other 4 reservoirs in Texas ranged from 35% in Wright Patman Lake (Texas) to 65% in Lake Mead (Arizona and Nevada). They found out that evaporation could account for a large portion of the overall water loss in arid and semiarid regions.

The impact of lateral inflows from tributaries (in comparison to the major river reaches) that are not observable to SWOT in the river-store mass conservation could not be ignored, especially in humid and subhumid regions where a greater level of precipitation-induced lateral inflows is likely. The scales of lateral flows are largely related to the hydroclimate conditions and watershed geomorphic factors (e.g., the drainage area) on the water stores, as they together determine the total volume of water that becomes lateral inflows. By comparing results based on different assumptions regarding lateral flows, Durand et al. (2014) suggested that the lateral inflows that contributed to the SWOT-observable rivers could place a large impact upon the discharge estimation accuracy in the MetroMan discharge estimation algorithm. They saw a large increase in uncertainty when lateral flows are not accounted. The rRMSE, an error performance indicator which will be discussed more in the Methodology section, was 10% when the lateral inflows were available, and 36% when they were not.

According to a USGS technical report on Lake Seminole in Florida, while surface water accounts for 81% of the reservoir inflow, groundwater accounts for 18% of the reservoir inflow, (Dalton et al., 2004). The reservoir inflow composition in Lake Seminole indicates that groundwater inflow could be large enough to place an impact on the algorithm's premise. The control of groundwater-surface water exchange on reservoir water balance depends on several factors such as local soil characteristics, water table gradient, and effluent or influent surface conditions (Gurrieri and Furniss, 2004). Since observations or estimations of these factors are relatively difficult to obtain, this thesis does not explicitly quantify the impact of groundwater exchange in reservoirs.

In comparison to natural lakes, anthropogenic factors could have a non-negligible role on the water budget of a regulated reservoir (Lehner et al., 2011). We might see dV responds to dQ differently among reservoirs designed for different purposes. Water withdrawal, including irrigation use and domestic water supply, hydropower generation, navigation, and flood prevention are among the common purposes of design for large dams and reservoirs. In dams built for water withdrawal, direct water pumpage could dominate the water budget in reservoirs where their designed capacity greatly exceeds the average annual runoff in the region.

1.3 Statement of Problem

It is very crucial before the application of LakeFlow to assess and compare the scales of each hydrologic and anthropogenic controls in the river-store mass conservation among lakes and reservoirs. The scale of ε determines if LakeFlow could produce accurate river discharge estimates by directly influencing the conditions of the algorithm's premise. Through a premise test, we would have a better sense of where and when this algorithm could yield a higher estimation accuracy in the context of reservoir purposes and hydroclimate conditions.

Aforementioned studies have highlighted the importance of evaporation loss, lateral flows, water withdrawal and groundwater exchange in the lake water balance and discharge estimation (Dalton et al., 2004; Durand et al., 2014; Zhang et al., 2017; Zhao and Gao, 2019). However, an intercomparison of the scales and magnitude of these hydrologic and anthropogenic controls on the river-store mass conservation remains lacking. Bonnema et al. (2016a) conducted a sensitivity test among hydrologic controls (reservoir inflow, storage change and surface evaporation) that were used to estimate reservoir outflow through a mass-balance approach in two hydropower reservoirs. By comparing the error results from including a different set of hydrologic controls, they revealed a minimal impact of surface evaporation on the outflow estimation accuracy. The fact that only two reservoirs were tested in their study calls for a need to include more reservoirs from a wide variety of climate and non-climate conditions, for a more comprehensive investigation of the roles and scales of secondary controls represented by ε in the river-store mass conservation.

The premise of the LakeFlow algorithm and scales of ε have not been tested on a broader geographical scale. A comprehensive investigation of premise scenarios across a variety of natural and human conditions is necessary prior to LakeFlow's application as the feasibility of the algorithm relies highly on it. Such efforts would help to answer questions like: where and on what conditions can the premise of the algorithm be observed; what are the secondary controls that have a significant impact on the river-store mass conservation, and how do their scales change with different natural and human conditions; could simulations and estimates of these secondary controls help to improve the premise conditions?

SWOT has a mean revisit period of eleven days, meaning that SWOT could only provide its remotely-sensed measurements on surface water about three times a month on average (Biancamaria et al., 2015). On a global scale, the number of overpasses could range from a maximum of two to ten. With these temporal sampling gaps between revisit days, it remains unknown whether SWOT would be able to capture the dynamics described in the river-store mass conservation for the majority of the water stores. If it could, are there any patterns as to in what geographical or climatic conditions, could SWOT better capture the mass conservation? Additional study is needed to further investigate the impact of SWOT's temporal sampling gaps on the premise of LakeFlow.

1.4 Research Objectives

This thesis aims to break the key barriers in understanding the feasibility of LakeFlow in its large application by widely testing the premise of LakeFlow across various geographic regions. The testing of the premise improves our understanding of how well the reservoir storage change and the discharge difference between the reservoir's main inflow and outflow river reaches can characterize the river-store mass conservation on a large spatial scale, a nice indicator of the future LakeFlow's performance when incorporating with SWOT's measurements. Therefore, this study has real-world implications. In summary, this research has the following objectives:

I. To test whether the algorithm's premise (that is, dV is the first-order control on dQ or expressed as $\mathrm{dV} \approx \mathrm{dQ}$) widely exists across the CONUS.

II. To assess the scales of the secondary controls (ε), and provide a comprehensive analysis of how their scales vary with hydroclimate conditions and reservoirs' purposes of design.

III. To find out if the inclusion of the secondary controls using model simulations (lateral inflow) and/or estimated dataset (surface evaporation) helps close the mass budget and thus improve LakeFlow's premise condition, and if it does, how we can use these datasets on reservoirs to potentially improve the accuracy of river discharge estimates from LakeFlow.

IV. To investigate if SWOT can capture the dynamics in the river-store mass conservation under the impact of SWOT's temporal sampling gaps.

1.5 Thesis Outline

This study introduces a discharge estimation algorithm that aims to break the key barriers to improving the accuracy of discharge estimation. The performance of LakeFlow relies heavily on a theoretical premise. As it remains uncertain how the premise of the algorithm could be affected by hydroclimate and human interference, this study sheds light on the applicability of the algorithm on a large scale by testing the premise in over 20 reservoirs built for various purposes and across various hydroclimate conditions. This study further aims to improve the premise conditions by adding model simulations of lateral inflow and a multi-source dataset for reservoir surface evaporation. The results of the addition are also presented and discussed in the following sections. Finally, this study looks at the SWOT's temporal sampling and investigates the impact of SWOT's temporal sampling on the premise conditions. Chapter 2 is the Methodology section of this thesis, and will focus on introducing our selection and the processing of data, the methods that address the proposed research questions individually, and how they would be able to facilitate the analysis of these questions and error metrics that evaluate the error performance. Chapter 3 will focus on presenting the statistical results by applying the proposed methods to processed data, and how these results facilitate further discussion on our research topics. Chapter 4 is the conclusion of the thesis.

Chapter 2 - Methodology

2.1 Data Acquisition

2.1.1 Sample Selection

The sensor payload on the SWOT satellite will be able to observe all lakes and reservoirs with an area greater than 100m by 100m (Biancamaria et al., 2015). We followed the mission guideline to sample studied reservoirs accordingly from a global dam inventory. The Global Reservoir and Dam (GRanD) database (version 1.3) contains 7,320 records of dams and their associated reservoirs with a cumulative storage capacity of over 6,800 km^3 (Lehner et al., 2011). GRanD includes all dams whose reservoirs have a storage capacity of over $0.1km^3$. The dams are georeferenced, and the spatial boundaries of the associated reservoirs are specified as polygons at a high spatial resolution. Each dam is associated with multiple non-spatial attributes, including the name of the dams and reservoirs, construction year, primary use, storage capacity and surface water extent. As nearly all reservoirs included in GRanD have a surface area over $0.1km^2$, we used the database to help us select reservoirs that are observable to SWOT.

Under the current design of swath coverage, rivers that are wider than 50m can be observable to SWOT (Biancamaria et al., 2015). To test our algorithm premise on reservoirs and rivers that could be observed by SWOT once it is launched, we need to make sure that not only the reservoirs selected with GRanD are large enough, but also their surrounding river reaches are wide enough so SWOT could provide useful measurements. Any river reach, either at the inflow or outflow point of a reservoir, would be considered in the river-store mass conservation as the primary river reaches if they are observable to SWOT. For our study, we used the SWOT River Database (SWORD) (Altenau et al., 2021) to further filter out the reservoirs that connect rivers that are not observable to SWOT but were previously selected using GRanD. The Global River

Widths from Landsat (GRWL) database which includes rivers wider than 30 m (Allen and Pavelsky, 2018) served as the primary basis for the development of SWORD (Altenau et al., 2021). Therefore, for a reservoir that was pre-selected using GRanD, at least one inflow river reach and one outflow river reach need to be included in SWORD, so that the river-store mass conservation could be characterized by SWOT.

The reservoir selection was also determined by the availability of gauge data and reservoir evaporation data. We further filtered out the reservoirs where either gauge data or reservoir surface evaporation data were not available to us. The detailed description of the source of gauge measurements for streamflow and reservoir storage is covered in Section 2.1.2. A detailed description of the reservoir surface evaporation dataset is covered in Section 2.1.3.

To test the algorithm premise as thoroughly as possible, we selected 26 dams across the contiguous United States (CONUS) with both the associated reservoirs and their connected inflow and outflow river reaches observable to SWOT (Figure 2-1). Among the 26 sampled reservoirs, four are located in arid regions, ten in semiarid regions, three in subhumid regions, and nine in humid regions (Figure 2-2). The Hydrologic Landscape Regions (HLR) of the United States dataset (Wolock, 2003) helped us to identify the hydroclimate conditions at which the selected reservoirs are located. This dataset contains watersheds grouped into five classes according to their similarities in landscape and climate characteristics (Wolock, 2003).

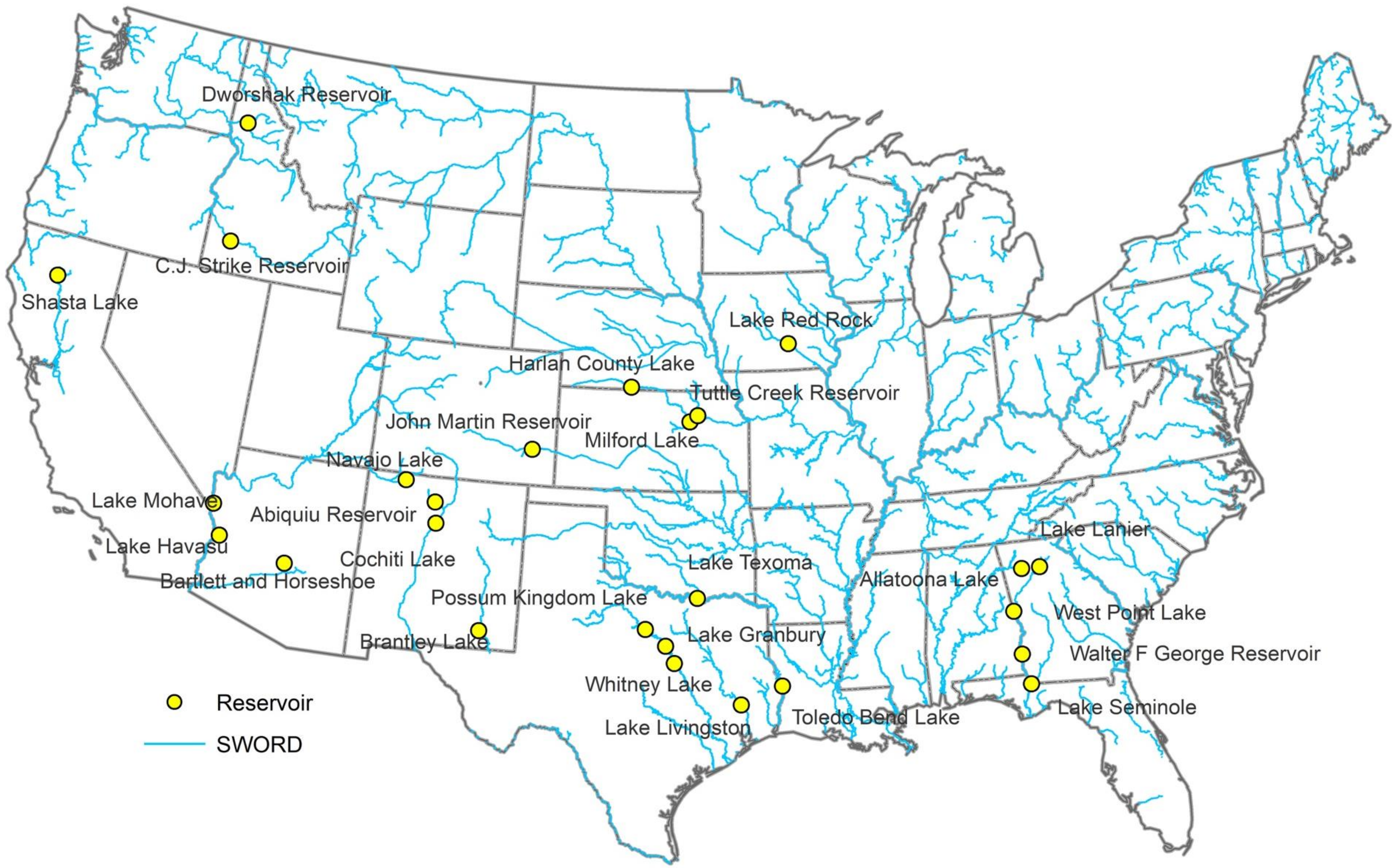


Figure 2-1 The 26 reservoirs selected with Global Reservoir and Dam (GRanD) database across the contiguous United States. The line features are SWOT-observable rivers included in the SWOT River Database (SWORD).

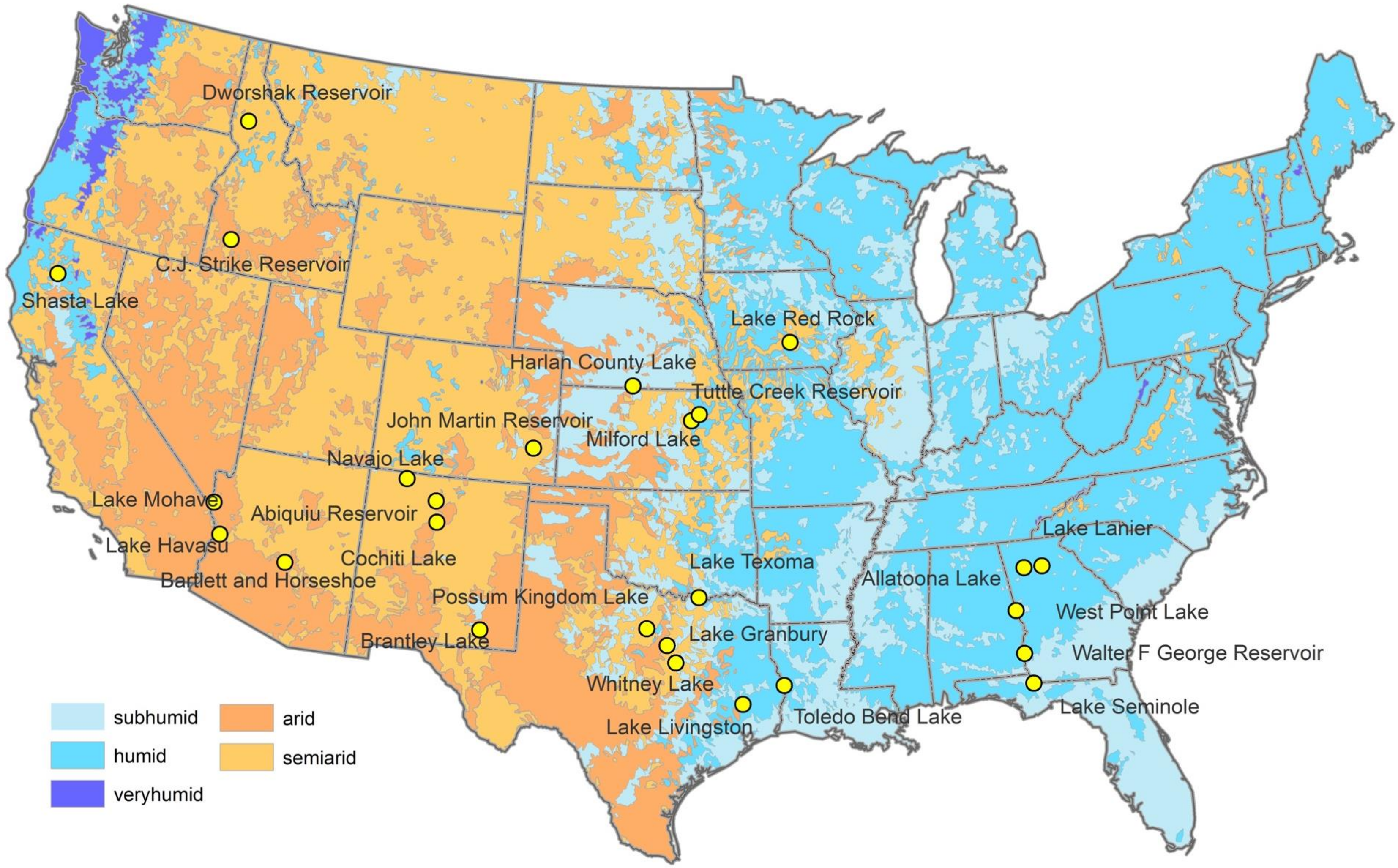


Figure 2-2 The 26 reservoirs selected for this study and the hydrologic landscape regions of which they belong to. The Hydrologic Landscape Regions (HLR) of the United States dataset contains watersheds grouped into five classes according to their similarities in landscape and climate characteristics (Wolock, 2003).

2.1.2 Flow Records and Reservoir Storage

To test the algorithm premise and to characterize the river-store mass conservation, we need the flow data for both the primary inflow and outflow river reaches, and the reservoir storage variation data. Then the flow data can be used to calculate the discharge difference between the primary inflow and outflow reaches, and eventually be compared to reservoir storage change. The United States Geological Survey (USGS) provides daily measurements of river discharge, reservoir storage, and gauge height collected from approximately 1.9 million sites in all 50 states (https://waterdata.usgs.gov/nwis). The current and historical data for these measurements are maintained and released to the public by the USGS web server (https://maps.waterdata.usgs.gov/mapper/index.html). As considered to be close approximations of the true "discharge" of the river reaches and "storage" of the reservoir, these gauge measurements from USGS have widely been used for validation as baseline reference within the hydrologic community. Therefore, these gauge measurements would be sufficient to represent the actual river-store mass balance of these reservoirs. Daily observations of reservoir storage and stream discharge are available for all 26 reservoirs selected for our study and their surrounding primary inflow and outflow river reaches. Though the maintenance of stream or lake gauges may have been discontinued for some reservoirs, the historic data available on the USGS database are still a valuable source for investigating conditions of the river-store mass conservation. As a result, we collected daily in situ measurements of reservoir storage and stream discharge of their primary inflow and outflow river reaches for 2 consecutive water years, albeit the start years are not always the same among the reservoirs. According to USGS explanations (https://water.usgs.gov/nwc/explain_data.html), a water year starts on the first day of October of the first year, and ends on the last day of September of the second year .

On the USGS database, the stream discharge is documented with a unit of cubic feet per second. Reservoir storage is often documented with a unit of acre-feet except for several large reservoirs with a more sizable unit of thousand-acre feet. For stream discharge, different USGS gauge stations might give their readings at different times of the day, with the most common being 6:00 am, 8:00 am and midnight. When gauge readings of discharge are available more than once a day, a mean value will be generated as the daily summary data. There was not a general pattern for a reservoir as to when an observation at a given time would be available to users. For our study, we would use the daily mean observation for both stream discharge and reservoir storage whenever it is available. There would be cases when daily mean observation is not available. When this happens, we would use the discharge and reservoir storage readings taken at a single time of a day when available, based on the assumption that the stream discharge and reservoir storage would not vary greatly during a day. The assumption holds for most days in a water year when the climate is less variable. As the final step of the data preparation, the unit used in the discharge daily time series was converted from cubic meter per second to million cubic meters per day, and the unit used in reservoir storage daily time series was converted from either cubic foot per second or thousand cubic meters per second to million cubic meters.

Using the daily time series of reservoir storage from USGS, the daily time series of storage changes were generated from measurements every two consecutive days. Differently, the discharge of the primary outflow river reach was subtracted from the discharge of the primary inflow river reach on the same day to generate the daily time series of discharge difference between the primary inflow and outflow reaches.

2.1.3 Secondary Controls

This study also aims at improving the river-store mass conservation that was solely characterized by reservoir storage and stream discharge of the primary inflow and outflow river reaches by adding model simulations of lateral flows and multi-source datasets of reservoir surface evaporation on a daily time scale.

The CONUS Reservoir Evaporation Dataset (CRED) (Zhao and Gao, 2019) is a multi-source dataset that provides monthly time series of evaporation rate and amount (as the product of evaporation rate and lake surface area) for 721 reservoirs in the United States from March 1984 to October 2015. The lake heat storage term in the Penman Equation was considered when calculating the evaporation rates in CRED. The lake surface area dynamics were extracted from the Landsat-based water surface occurrences (Pekel et al., 2016) and improved by better leveraging the contaminated pixels from the cloud, shadows and more (Zhao and Gao, 2018). The evaporation rates were documented in a unit of millimeters per day. The evaporation amount was calculated as the product of evaporation rates and lake surface area in a unit of thousand cubic meters per month. The CRED dataset contains evaporations rates and amounts derived from 3 sources of monthly meteorological data, TerraClimate (Abatzoglou et al., 2018), North American Land Data Assimilation System (NLDAS) (Xia et al., 2012) and Global Land Data Assimilation System (GLDAS) (Rodell et al., 2004). The mean of the monthly evaporation amounts from these 3 data sources was available. For the scope of this study, we will use the daily time series of the mean of the monthly evaporation amounts for 26 studied reservoirs.

The Global Reach-level A priori Discharge Estimates for SWOT (GRADES) dataset simulates global discharge values of river flows at 2.94 million river reaches (Lin et al., 2019). Daily streamflow was simulated on a 90m high accuracy DEM and later routed out using a

Routing Application for Parallel computation of Discharge river routing model (RAPID). Excluding the primary inflow and outflow reaches that were previously selected by using the SWORD database, the river reaches that directly touch the boundary of the selected reservoirs would be regarded as the tributaries that are not observable to SWOT. Yet, these river reaches drain to the water stores, making a likely substantial contribution to the reservoir storage variation. By using the ArcGIS platform, these river reaches included in the GRADES dataset were manually selected for flow calculation. The daily discharge values were then calculated for a continuous 2-year period that corresponds to the duration length of in situ gauge measurements for each reservoir. A sum of river flows for all tributaries that drain to the reservoir according to GRADES was generated to represent the total impact of lateral flows.

2.2 Observed River-Store Mass Conservation

2.2.1 SWOT's Temporal Sampling Gaps

At 890.6 km with an inclination of 77.6 degrees, SWOT has an orbital cycle of 21 days but may be able to provide more than one overpass in a span of one orbital cycle (Biancamaria et al., 2015). As documented, during one SWOT's orbital period, the number of revisits ranges from a maximum of 2 in regions close to the equator to more than 10 above 70 degrees in both north and south hemispheres. Therefore, the number of overpasses during a SWOT orbital period may vary greatly depending on the specific locations of selected reservoirs. The French government space agency, Centre National D'Etudes Spatiales (CNES), maintains an online service named AVISO+ (i.e., Archiving, Validation and Interpretation of Satellite Oceanographic data) that provides data products from satellite altimetry (https://www.aviso.altimetry.fr/en/home.html). The ground track and orbit ephemeris in the shapefile format for SWOT are available for download on AVISO+'s website (https://www.aviso.altimetry.fr/en/missions/future-missions/swot/orbit.html). With the SWOT swath information and the spatial locations of selected reservoirs, the maximum number of overpasses for each reservoir was calculated. As our algorithm would need observations on reservoir's main surface body and its main inflow and outflow reaches, the SWOT overpass needs to cover the majority of the reservoir body along with its primary inflow and outflow reaches. In this study, the centroid of the lake body was used to calculate the number of overpasses. The ArcGIS provides a spatial join geoprocessing tool that joins attributes from one feature to another based on the spatial relationship. If the centroid of a lake body was within three SWOT overpasses, the number of the overpasses would be three and the revisit time for this lake would be one SWOT repeat period (i.e., 21 days) divided by three, yielding seven days. By using the ArcGIS platform, we calculated the number of overpasses during a SWOT orbital period (i.e., 21 days) for each reservoir by spatial joining the centroids of the reservoir polygons from GRanD with the SWOT swath polygons. With the number of overpasses, the SWOT revisit time for each reservoir was then calculated.

2.2.2 Temporal Rescaling of Gauge Measurements

In situ measurements of reservoir storage and streamflow from the USGS National Water Information System were collected once or more times per day. We then aggregated the daily measurements of each reservoir by their corresponding SWOT revisit period. SWOT's direct measurements of surface water are temporally discrete, therefore before testing the algorithm's premise, an aggregation of daily measurements is necessary to reflect the impact of SWOT's temporal sampling on the discharge estimates.

For example, a reservoir has a SWOT revisit time of 5 days. Daily time series of discharge difference and storage change would need to be aggregated every 5 days. Our USGS daily time series all begin their observation on the first day of October. Assuming that October the first is the beginning of a SWOT revisit period, then October the fifth would be the last day of the first SWOT revisit period. Measurements of storage change and discharge difference over the first SWOT revisit period (i.e., from October the first to the fifth) would then be aggregated. The output values are the cumulative storage change or discharge difference over a span of a SWOT revisit period. The aggregation will continue for the subsequent revisit periods, until the last available revisit period occurs.

The beginning of a revisit period could also be the second, third, fourth or fifth of October assuming the length of a SWOT revisit period is 5 days. Therefore, followed by the example, five sequences of discharge difference or storage change having different starting days would be created from an aggregation of single original daily time series.

2.2.3 Comparison between Reservoir Storage Change and Discharge Difference

Our study looks into a theoretical hydrological relationship (i.e., river-store mass conservation of lakes and reservoirs). We aim at revealing how well this relationship, across reservoirs, can be characterized by using observations on reservoirs and their primary inflow and outflow river reaches, and making implications on how LakeFlow can be better applied moving forward. To analyze the conditions of the river-store mass conservation, we compared the observations of reservoir storage change to the observations of discharge difference, on the time scale of one SWOT repeat period. Error performance metrics and their associated performance criteria are crucial components of this study, so we can evaluate the hydrological relationship for each reservoir and conduct a fair intercomparison among them all.

However, the characteristics of our data (i.e., reservoir storage change and discharge difference between inflow and outflow river reaches) are different from flow data. It remains unknown whether the commonly-used error performance metrics in flow simulation/estimation would be adequate to allow a fair evaluation and intercomparison in our study. In fact, Knoben et al. (2019) argued that there has not been a single perfect modeling performance metric that is suitable for every study occasion and objective in the hydrological modeling community. We will start with a short review of the error performance metrics we decided to use in this study, and continue with a discussion on their possible suitability according to limitations and constraints of our error metrics regarding the characteristics of our data. In addition, it is important to achieve a better understanding of the error performance criteria and how these criteria are interrelated, as such efforts help to understand what is causing a hydrological relationship to appear stronger in a reservoir.

Statistical results were reported according to six error performance metrics: Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970), Kling Gupta Efficiency (KGE) (Gupta et al., 2009), relative Root Mean Square Error (rRMSE) (Bjerklie et al., 2005a), Mean of Relative Residual (MRR) (Bjerklie et al., 2005a), the Standard Deviation of Relative Residuals (SDRR) (Bjerklie et al., 2005a) and a distribution-free overlapping index (Pastore and Calcagni, 2019). The NSE and KGE performance metrics are two traditionally-used measures of model performance used in hydrology (Gupta et al., 2009; Knoben et al., 2019; Lamontagne et al., 2020). MRR, SDRR and rRMSE have been widely used in previous research where the estimated flow product was evaluated against the observations (Durand et al., 2016; Durand et al., 2014; Gleason and Durand, 2020; Hagemann et al., 2017; Nickles et al., 2020a; Nickles et al., 2020b; Nickles et al., 2019; Tuozzolo et al., 2019). The MRR, SDRR and rRMSE altogether allow for a discussion over how much of total error (represented by rRMSE) was caused by bias (represented by MRR) and how much was caused by variability, enabling a fair intercomparison of five discharge estimation algorithms over their estimation performance across 19 rivers (Durand et al., 2016). rRMSE, rBias (a different name for MRR) and NSE error metrics were used to evaluate the performance of Bayesian AMHG-Manning (BAM) algorithms across the same set of 19 rivers (Hagemann et al., 2017).

NSE and KGE are commonly used for calibrating hydrological models, optimizing the values of their parameters and evaluating model performance (Gupta et al., 2009; Knoben et al., 2019; Moriasi et al., 2015). In a single number, the NSE and KGE reveal the similarity of observed and simulated flow records (Gupta et al., 2009; Moriasi et al., 2015; Nash and Sutcliffe, 1970). The equations for calculating NSE and KGE are as follows:

$$NSE = 1 - \frac{\sum_{t=1}^{t=T}\left(Q_{sim}(t) - Q_{obs}(t)\right)^2}{\sum_{t=1}^{t=T}\left(Q_{obs}(t) - \overline{Q_{obs}}\right)^2}$$

Equation 5 Nash-Sutcliffe Efficiency (NSE)

$$KGE = 1 - \sqrt{(r-1)^2 + (\frac{\sigma_{sim}}{\sigma_{obs}} - 1)^2 + (\frac{\mu_{sim}}{\mu_{obs}} - 1)^2}$$

Equation 6 Kling Gupta Efficiency (KGE)

Where T is the total number of time steps, indicating the length of the flow records. $Q_{sim}(t)$ and $Q_{obs}(t)$ are the discharge values from model simulations and observation at a certain time step t. $\overline{Q_{obs}}$ is the mean of all flow observations. r is the linear correlation between all observations and all simulations, σ_{obs} and σ_{sim} are standard deviations of all observations and simulations of discharge. μ_{obs} and μ_{sim} are the mean of all observations and simulations of discharge. μ_{obs} is equivalent to $\overline{Q_{obs}}$.

NSE is dimensionless and falls into the range of [-inf to 1.0]. Its value indicates the portion of the variance in the observations that can be explained by the simulations. NSE uses the mean flow as an inherent benchmark predictor when it equals zero. NSE = 0 has been popularly used as the separator for "good" and "bad" model performance, although its predictive skill in a more complex setting is questionable (Knoben et al., 2019). NSE = 1 indicates that there is a perfect correspondence between simulated and observed flow, and that the model simulations have reached maximum (ideal) performance. NSE < 0 indicates that the model simulations are a worse predictor than the mean of all observations. Without accounting for different hydrologic processes and physical conditions, this evaluation criterion (i.e., NSE > 0 indicates good model performance; NSE < 0 indicates bad model performance) is considered to be a low-level evaluation benchmark, especially in cases where seasonal conditions and flow regimes are different (Gupta et al., 2009; Knoben et al., 2019).

By design, KGE was created for model performance evaluation by treating correlation, bias error and variability error as three separate error components, different from NSE (Gupta et al., 2009). As previously explained, r represents the linear correlation, $\frac{\sigma_{sim}}{\sigma_{obs}}$ is the quotient of the standard deviations of the simulations and observations, indicating the variability error, $\frac{\mu_{sim}}{\mu_{obs}}$ is the quotient of the means of the simulations and observations, indicating the bias error. It is important to point out that NSE values (or scores) and KGE values cannot be treated as equivalent as there is no unique relationship between the two error performance metrics, and there is no inherent benchmark for KGE, while for NSE, there is (Gupta et al., 2009; Knoben et al., 2019). An interpretation of NSE values on model performance should not guide the interpretation of KGE values (Knoben et al., 2019).

Similar to NSE, KGE falls in the range of [-inf to 1.0]. KGE = 1 indicates a maximum performance that a model could achieve for a perfect correspondence between simulations and observations of streamflow. While Koskinen et al. (2017) used KGE = 0 as an indicator of model simulation providing no better estimate as the mean of all observations, a technical report done by Knoben et al. (2019) argued that KGE does not have an inherent benchmark in the form of mean flow as NSE, and KGE < 0 does not necessarily mean that the model simulation is a worse indicator than mean flow. In other words, KGE = 0 does not share the same implication of achieving the same explanatory power as the mean of the observations as NSE = 0, and there is no theoretical basis for treating KGE = 0 as a general benchmark indicator for model performance. By definition, KGE calculates the distance from an ideal model performance according to three dimensions (linear correlation, variability and bias). In a case where all simulations equal the mean of observations ($\sigma_{sim} = 0,\ \mu_{sim} = \mu_{obs}$, r = 0), the KGE value equals $1 - \sqrt{2} \approx$ -0.41, where a mean flow benchmark can be defined (Knoben et al., 2019). These findings were supported in mathematical terms and also in synthetic experiments in their study. Therefore, a KGE value of -0.41 could be used as an indicator where model simulations have the same explanatory power as the mean of all flow observations.

The relative statistics, namely rRMSE, MRR and SDRR have provided a fair comparison for existing discharge algorithms among various river reaches (Durand et al., 2016). The relative residuals on which the calculation of three relative metrics is based, are written for one time step as:

$$RR(Relative\ Residuals) = \frac{(observation - estimate)}{observation}$$

Equation 7 Relative Residuals

The mean, standard deviation of Relative Residuals (RR) and root of the mean square RR at all time steps are calculated and referred to as MRR, SDRR and rRMSE. A large MRR value and a low SDRR value indicate a possible large bias from the estimates to the observations, but the overall variability is well captured. MRR shows agreement in value magnitude and direction, and SDRR shows agreement in value variation and spread. rRMSE can be interpreted as the total errors from MRR and SDRR, indicating an overall error performance. The three metrics have a mathematical relationship which can be written as:

$$rRMSE^2 = MRR^2 + SDRR^2$$

Equation 8 Decomposition of rRMSE into MRR and SDRR

The last error metric we used for evaluating the observed river-store mass conservation is a distribution-free overlapping index. The index is relevant to the area intersected by the probability density functions (Pastore and Calcagni, 2019), and therefore provides a useful insight into the similarity between our reservoir storage change and discharge difference. The detailed mathematical calculation behind this distribution-free overlapping index could be found in (Pastore and Calcagni, 2019). An R package was developed for calculating this overlapping index. For further details on this R package and the useful functions included in the package, see the manual at https://cran.r-project.org/web/packages/overlapping/overlapping.pdf.

Our motivation is not to find the best overall performance metric for the large hydrological field (as the variety of application scenarios, and the complexities in hydrological processes and their interactions make it unlikely for such metric to even exist), but use performance metrics and their associated criterion that are most adequate and can be best utilized to allow for a fair comparison of a hydrological relationship characterized by observations.

As previously described, NSE is a useful error performance metric that has been traditionally used to evaluate and optimize model performance. Gupta et al. (2009) conducted an in-depth decomposition analysis where they mathematically broke down NSE into three components (the bias, the correlation and a measure of relative variability), and identified the two major limitations. First, as the bias term in NSE is normalized by the standard deviation of the observations, the importance of the bias error will vary among our sampled reservoirs. When the variability in the observation data for one reservoir is high, the bias component will have a smaller contribution to the final value of NSE. In other words, NSE will not be as sensitive to the bias error in the simulations when the variability in the observations is high, leading to a large volume balance error. Second, as NSE is being maximized, the variability in the flow data tends to be systematically underestimated, leading to a smaller peak flow value. In addition to two findings, it was also acknowledged that because the baseflow errors are typically smaller than peak flow errors, NSE is generally found to be highly sensitive to peak flow errors (Gupta et al., 2009).

As evident from the above equations, NSE and KGE are different in their calculation, though both measures can be represented in terms of three components which measure the linear correlation, the bias and the variability (Gupta et al., 2009). Though KGE was created based on the decomposition of NSE by the bias, the correlation and a measure of relative variability, KGE is not meant for being a better/improved error performance metric than NSE, simply less prone to the systematical limitations in NSE (Gupta et al., 2009).

The relative statistics, namely rRMSE, MRR and SDRR are three error performance metrics used in Durand et al. (2016) to characterize relative discharge errors. rRMSE is highly sensitive to baseflow errors that are generally small relative to peak flow errors, leading to an inflation of rRMSE value (Hagemann et al., 2017). In a relative residual on which these relative statistics are based, the difference between the observation and simulation is normalized by the observation at the current time step. Further inflation on values of the three relative statistics could be induced by having a large number of small values in the observations.

In our analysis, the reservoir storage change and discharge difference between major inflow and outflow river reaches at an approximately weekly time scale are used as the observations and simulations when evaluating error performance. Although three of the common characteristics of streamflow data is a strong seasonal periodicity, variability and skewness (i.e., a rapid increase in late winter and spring followed by a gradual decrease until the end of autumn) (Lamontagne et al., 2020; Livina et al., 2003), no body of literature is relevant to the data characteristics of reservoir storage change and discharge difference between a reservoir's major inflow and outflow river reaches.   It is therefore uncertain whether the variability in our observation data and simulation data would be high. But we are certain that the scales of both reservoir storage and discharge difference will be largely impacted by damming operations and hydroclimate variability, and are also related to the size, the geometry and the capacity of the reservoir and its connected major river channels.

Followed by our discussion on NSE, if the variability in the observation data (i.e., reservoir storage change at approximately a weekly time scale) is high, the bias error will contribute less to the final value of NSE, deemphasizing the bias component. The second finding by Gupta et al. (2009) on NSE is less relevant to our study's setting as we do not intend to maximize NSE values for optimization purposes. Moriasi et al. (2015) provide guidelines for evaluating model performance through a meta-analysis of performance data for four hydrological models. They suggest that for watershed-scale models, NSE > 0.50 indicates a 'satisfactory' model performance. To reaffirm, our main objective of this study is not to establish new performance evaluation criteria that is specific to river-store mass conservation, but rather to make useful interpretations of our results from error performance metrics that allow a fair comparison among sampled reservoirs. Therefore, in our study, we will use the performance evaluation criteria recommended by Moriasi et al. (2015) and find out how the results would be by considering the premise condition of a reservoir to be "satisfactory" or better if the NSE for discharge difference as an estimate of storage change exceeds 0.50. An NSE value over 0.70 and less than 0.80 indicates a "Good" performance while an NSE over 0.8 indicates a "Very Good" performance.

All six error metrics were calculated for our sampled reservoirs to evaluate the similarities between discharge difference and reservoir storage change. In the calculation of these metrics, the reservoir storage change was regarded as the benchmark observations while the discharge difference between the primary inflow and outflow river reaches were regarded as the estimates of the observations. NSE, KGE, rRMSE and an overlapping index are measures of similarity between the reservoir storage change and discharge difference in a single value. There have been over 23,500 Google Scholar citations to Nash and Sutcliffe (1970) (02 May 2022) and NSE was regarded as "by far the most utilized index in hydrological applications" (Biondi and Todini, 2018). Advantages exist behind the widespread usage and the popularity of NSE, including an extensive body of literature on its applications, discussions over the error behaviors of NSE and the impact of the data characteristics in hydrology. That is why we decide to use NSE as the central error performance metric in the study. KGE, rRMSE and an overlapping index are supplementary to the use of NSE, providing additional insights and references to the error performance results and facilitating the discussion on the river-store mass conservation on sampled reservoirs. MRR and SDRR will be primarily used in our second research question, for assessing the secondary controls within river-store mass conservation characterized by reservoir storage change and discharge difference between the reservoir's major inflow and outflow river reaches.

Outliers are common in gauge measurements. The causes of outliers could be attributed to systematic errors or extreme events in our data. Outliers, small or large, could have a huge impact on the statistical results. Yet the exact impact that outliers have on our statistical results is currently unknown. Furthermore, error performance metrics behave to outliers very differently followed by our previous discussions on these metrics (e.x., rRMSEs are sensitive to baseflow errors; NSE is less sensitive to bias when the variability is high in observations but sensitive to peak flow errors). Therefore, it is necessary to conduct a systematical analysis of the impact of outliers on our error performance metrics.

Followed by this motivation, we applied an outlier-removal technique to potentially filter out the measurements that exceed the bounds within which the majority of the dataset is located in (i.e., between 2.5% and 97.5% quantiles). This method takes into consideration the overall distribution of the measurements within two years, and sees records that are outside of the range of 2.5%and 97.5% quantiles as out-of-bounds outliers. For comparison, we calculated each error metric, except for the overlapping index, first with measurements that outliers are removed and then with measurements that outliers are left untouched. We believed that outliers could have little impact on the overlapping index, as this index is more sensitive to any discrepancy between the probability distributions of the compared datasets, therefore unlikely to be biased due to outliers.

To summarize how outliers were removed in our study in the example of NSE, first, we calculated the 0.025 and 0.975 quantiles for the time series of reservoir storage change. The storage changes that are not within the 0.025 and 0.975 quantiles would be marked as out of bounds, and the days on which storage changes were marked as out of bounds would not be included in the calculation of NSE. Then, the same measures were completed for the time series of discharge difference. And as the result, a list of days on which either storage change or discharge difference was out of bounds was generated. As previously written, the measurements on these days would not be included in the calculation of "quantile" NSE. With this list, the "quantile" NSE was calculated by excluding the measurements of both discharge difference and storage change on the days on the list. This NSE was referred to as the "quantile" or "outlier-removed" version of NSE.

While for NSE, outliers were removed from the measurements of discharge difference and storage change, outliers were handled at a different level for the three "relative" statistics. This indirect approach was necessary. As explained in previous sections, MRR, SDRR and rRMSE were all derived from a base error metric, relative residuals. As the reservoir storage or discharge difference level might not vary substantially for days (for example, during the dry season), minimal reservoir storage change and discharge difference were left likely to result in an unrealistically huge relative residual. Due to this reason, the same quantile outlier-removal method was applied to remove the outliers in the relative residuals after they were calculated.

Overall, calculating each error metric except for overlapping index in two versions would allow us to find out how outliers in the hydrologic records influence the statistical results. This would also help us to constrain the values of relative residuals when the denominator (for instance, the reservoir storage oftentimes might stay very stable and therefore gives a very small reservoir storage change) is extremely small which would result in an extremely large relative residual. As three metrics in our study, MRR, SDRR and rRMSE, are all derivative of the relative residuals. We expect to see how two versions of these relative statistics differ from each other and how effective the "quantile" measure is in constraining the values of relative statistics.

Accordingly, we made scatterplots of reservoir storage change against discharge difference at the time scale of one SWOT revisit period (the length of the SWOT revisit period varies from a reservoir to reservoir). In addition to scatterplots, hydrographs showing monthly time series of reservoir storage change and discharge difference in two water years were made for selected reservoirs. The daily time series of discharge difference and storage change were aggregated by month to produce the monthly time series. The monthly time series generated here was titled "Observed dQ" and "Observed dV". There were other time series on the hydrograph besides "Observed dQ" and "Observed dV": "dQ + q"and "dQ - E".

2.3 Assessment of Secondary Controls

Such efforts helped us answer: where and on what hydroclimate conditions would the premise of the algorithm better be met; what are the overall scales and variabilities of the impact of secondary controls on the characterized river-store mass conservation, and how do their scales and variabilities change by different hydroclimate conditions and reservoir purposes? Therefore, based on the results which we intend to use to analyze the observed conditions of river-store mass conservation characterized by reservoir storage change and discharge difference from the first research question, we continue to compare the results of the studied reservoirs by hydroclimate conditions and the designed reservoir purpose, respectively.

As suggested, the evaluation of model performance should be treated as a multi-objective problem, and the decomposition of a single error performance criterion (in their case, NSE and KGE) into different components allows a deeper understanding of the causalities of model performance (Gupta et al., 2009; Knoben et al., 2019). Followed by Section 2.2.3, the decomposition of rRMSE gives two components: MRR and SDRR. The main advantage of using relative statistics is that, they allow us to form a more in-depth understanding of the error compositions, as MRR and SDRR characterize errors that are, respectively, bias-based and time-variability-based. Therefore, a set of two error performance metrics, namely MRR and SDRR are used, in addition to the overlapping index, NSE, KGE and rRMSE, to assess the impact of secondary controls within the river-store mass conservation characterized by reservoir storage change and discharge difference between the major reservoir inflow and outflow river reaches. MRR helps to understand the scales of secondary controls, and SDRR helps to understand the time-variabilities of secondary controls.

Additionally, datasets of hydroclimate conditions and the designed reservoir purpose are used respectively to help analyze how the impact on the secondary controls might relate to different hydroclimate conditions and human water withdrawal.

2.3.1 Hydroclimate Conditions

The Hydrologic Landscape Regions (HLR) of the United States dataset divides the contiguous United States into five large regions according to their similarities in landscape and climate characteristics from humid to arid (Wolock, 2003). According to HLR, 12 reservoirs are located in humid or subhumid regions with the remaining 14 in arid or semiarid regions. The overall scales of secondary control were analyzed by comparing the conditions of river-store mass conservation observed in reservoirs that are located in different hydroclimate regions. In more arid regions, the scale of surface evaporation could be larger than in humid regions. In humid regions, the scale of lateral flows could be larger than in the arid regions. By comparing the statistical results from arid and humid regions, we will have a better understanding of how hydroclimate conditions influence the conditions of river-store mass conservation and the overall scales of secondary control.

2.3.2 Reservoir Purposes

The GRanD dataset contains various non-spatial attribute information on each reservoir from construction year, descriptive administrative locations, dam height and reservoir purposes. The main reservoir purposes documented in the GRanD dataset include irrigation, hydroelectricity, water supply, flood control and navigation. Irrigation and water supply are referred to as human water withdrawal/consumption. According to the GRanD dataset, 18 (69.23%) of the 26 sampled reservoirs were built for human water withdrawal. We analyzed the conditions of river-store mass conservation of these water-supply reservoirs, and compared them to reservoirs that were not specifically designed for human water withdrawal.

2.4 Reservoir Surface Evaporation and Lateral flow

We will apply the discharge estimates of lateral flows from the GRADES database and surface evaporation from the CRED database to assess how the inclusion of these secondary control could influence the conditions of the river-store mass conservation characterized solely by water stores and their primary inflow and outflow river reaches, and whether the inclusion could improve the premise conditions.

Statistical results are reported using the same five error metrics as mentioned in Section 2.2.3 (i.e., Nash-Sutcliffe Efficiency (NSE), Kling Gupta Efficiency (KGE), relative Root Mean Square Error (rRMSE), Mean of Relative Residual (MRR), the Standard Deviation of Relative Residuals (SDRR)) to assess the effectiveness in adding model simulations of lateral inflow and estimated reservoir surface evaporation. They have been widely used as the benchmark error performance for discharge estimates in previous studies (Biancamaria et al., 2011; Bonnema et al., 2016b; Durand et al., 2016; Durand et al., 2014; Hagemann et al., 2017). Detailed descriptions of these error metrics can be found in Section 2.2.3.

2.4.1 Reservoir Surface Evaporation

The original monthly reservoir surface evaporation amounts made available from the CRED dataset are in a unit of thousand cubic meters per month. Before adding it to our analysis, the time scale of the original monthly dataset was then adjusted to a daily time scale. The process of the time scale adjustment acknowledges the number of days in each month and divided the monthly evaporation amounts by the number of days in each corresponding month. Once a daily time series of reservoir surface evaporation was created in a unit of a million cubic meters, this daily time series of surface evaporation were combined with the daily time series of discharge difference before undergoing an aggregation for a SWOT repeat period time scale.

The combined daily time series of discharge difference and surface evaporation were also aggregated by month to generate the monthly time series. This monthly time series titled “dQ - E” was to reflect the impact of reservoir surface evaporation on the river-store mass conservation that was characterized solely by discharge difference and storage change.

2.4.2 Lateral Flows

The original time series of lateral flows from the GRADES database was in a unit of cubic meters per second. Unit conversion to a million cubic meters per day was completed before it could be added to the daily time series of discharge difference and compared to the reservoir storage change. The combined daily time series of discharge difference and lateral flows was aggregated by a SWOT revisit period and by month.

A monthly time series titled "dQ + q" was added to the monthly hydrograph to reflect the impact of lateral flows on the river-store mass conservation that was characterized solely by discharge difference and storage change.

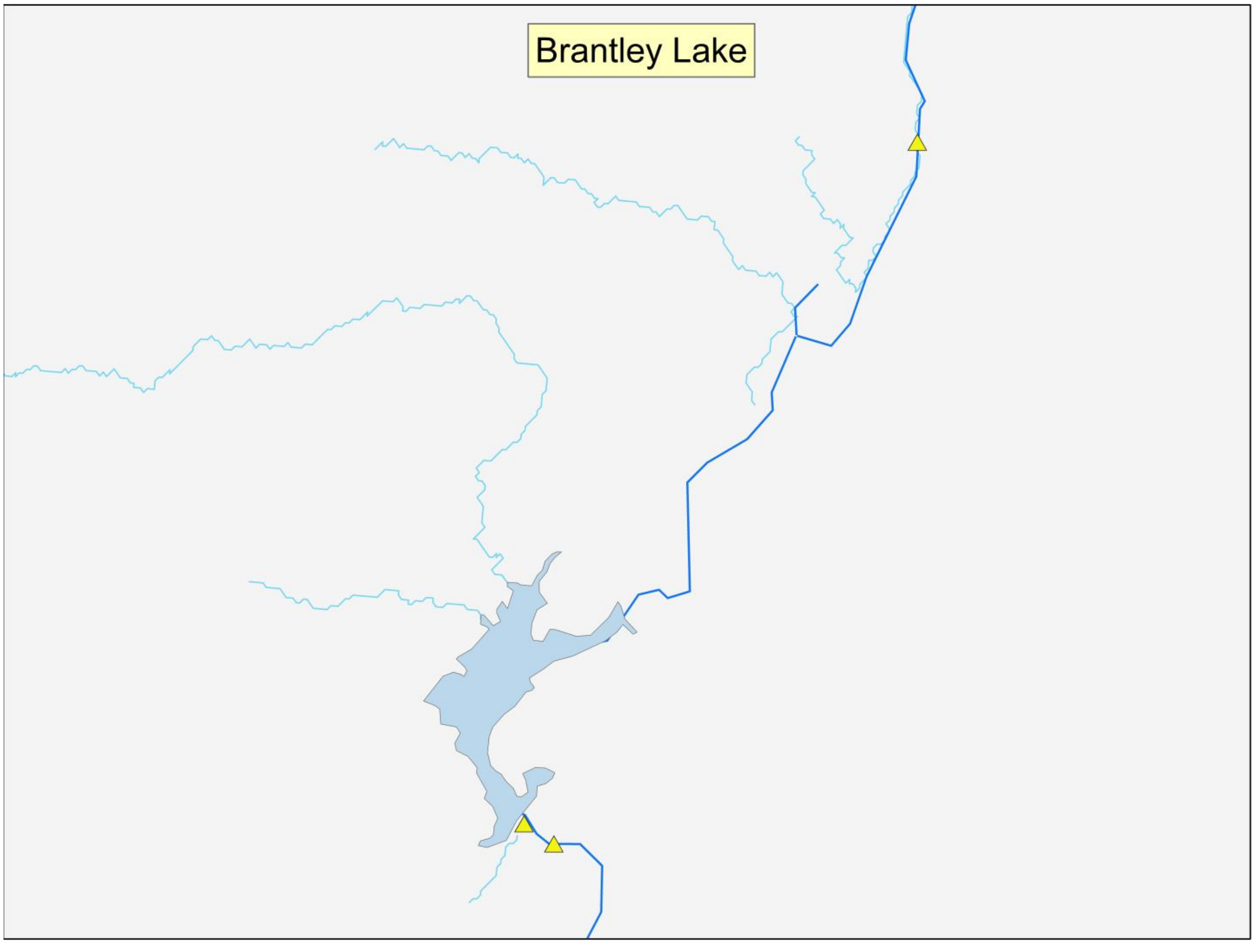


Figure 2-3 Vector representation of Brantley Lake, located in New Mexico, USA. Main water body of Brantley Lake (centered polygon), primary inflow (east) and outflow (west)

river reaches (highlighted in navy blue), tributaries that contribute to lateral flow (light blue) and gauge stations (triangles in yellow) are included in the figure. The swath (shaded areas) and ground track (red lines) approximate SWOT's coverage at this region.

2.5 SWOT's Temporal Sampling

2.5.1 Data Processing

In Section 2.2.2, a SWOT temporal timescale was integrated into the original SWOT measurements of discharge difference, by aggregating the daily time series by a SWOT revisit period. The aggregated time series will be considered as the referenced benchmark and is named as "Observed time series of discharge difference/Observed dQ" to avoid confusion.

As previously written, SWOT has an orbital cycle of roughly 21 days but could provide more than one overpass above most regions. SWOT measurements are temporally discrete, meaning that SWOT could only provide a suite of measurements every revisit period. This characteristic draws questions on how well SWOT could capture the river-store mass conservation.

Two assumptions were made when we simulated the impact of SWOT discrete temporal sampling on daily measurements. First, measurements of discharge difference from the USGS daily time series would only be available on days when SWOT revisited. Second, variations of discharge difference within each SWOT revisit period would follow a linear pattern, gradually increasing or decreasing from the discharge difference on the first revisit day to the discharge difference on the following day. We also considered the different scenarios where the first complete SWOT revisit period would occur differently.

With these two assumptions, estimations of discharge difference on days when SWOT observation was not available were linearly interpolated from measurements of discharge difference on days when SWOT observation was available. As this time series contains estimations of measurements when SWOT observation was not available, it was also referred to as the “estimated daily time series of discharge difference”. Then we aggregated the estimated daily time series by a SWOT revisit period. Sequences of estimated time series will then be compared to sequences of observed time series to investigate the impact of SWOT’s temporal sampling.

For example, for a reservoir that has a SWOT revisit period of five days, an observed time series of discharge difference would be derived through an aggregation of the original daily time series every five days while an estimated time series was interpolated from the original in situ measurements every five days. The observed time series would be the reference when analyzing the impact of SWOT’s temporal sampling on discharge difference and conditions of river-store mass conservation.

In addition, we added the estimated time series to the hydrograph. These estimated time series demonstrated the monthly uncertainties of discharge difference from SWOT’s temporal sampling. The number of estimated time series was determined by the length of SWOT revisit period of a specific reservoir. For example, a five-day revisit period will result in five sequences of estimated daily time series, with each sequence having a different starting day for the first complete SWOT revisit period.

2.5.2 Assessing SWOT’s Temporal Sampling

In two studies that specifically analyzed the impact of SWOT sampling characteristics on discharge, KGE and rRMSE were used as the central error performance metrics (Nickles et al., 2020b; Nickles et al., 2019). We use a set of five error performance metrics including NSE, KGE, rRMSE, MRR, SDRR and a Wilcoxon rank sum test (Cuzick, 1985) to assess the impact of SWOT's temporal sampling in this question. Detailed descriptions of these error metrics and how they were calculated can be found in Section 2.2.3. NSE, KGE and rRMSE are error performance metrics that assess the impact of SWOT's temporal sampling in a single value. Supplementarily, the use of MRR and SDRR improves our understanding on scales and the variabilities of the SWOT's temporal sampling on the time series of discharge difference. The rank sum test aims to determine whether two independent groups of data come from the same population, used as a measure of similarity between the observed and estimated time series of discharge difference between a reservoir's major inflow and outflow river reaches.

Chapter 3 - Results and Discussion

The following chapter reports the results that address the objectives of the study. Based on the results, I will provide the discussions regarding our research questions.

3.1 The Observed River-Store Mass Conservation

3.1.1 The Overall LakeFlow Premise Conditions

To test whether LakeFlow's premise widely exists across the CONtiguous United States (CONUS), we applied a suite of four error metrics, including NSE, KGE, rRMSE, and overlapping index (Section 2.2.3), to 26 reservoirs sampled in our study. We used the Nash-Sutcliffe Efficiency (NSE) as the central error performance metric in discovering whether reservoir storage change is the first-order control on discharge difference between the reservoir's major inflow and outflow river reaches. KGE, rRMSE and an overlapping index are supplementary to the use of NSE for assessing the river-store mass conservation characterized by reservoir storage change and discharge difference.

With daily gauge measurements of streamflow discharge and reservoir storage capacity, we were able to test the premise conditions for 26 reservoirs across the CONUS. Tables 3.1 and 3.2 summarizes the observed conditions of river-store mass conservation for 26 reservoirs according to 6 error metrics. Each error metric, except the overlapping index, was calculated in two versions, with full record length and with outliers removed by a quantile measure. As described in Section 2.2.3, NSE is by far the most widely-used error performance metric in hydrology to evaluate model performance (Lamontagne et al., 2020). Therefore, we will use NSE as the primary performance evaluation indicator for assessing the observed conditions of river-store mass conservation among our sampled reservoirs. The “All” NSE uses the all the time steps within the observations. The mean, standard deviation and median of this version of NSEs among 26 reservoirs are -8.0, 22.52 and 0.42. NSE > 0 indicates that the discharge difference between the reservoir’s major inflow and outflow river reaches (when NSE is used for model performance evaluation, this would be the model simulations) has higher explanatory skill than the mean of the reservoir storage change (when NSE is used for model performance evaluation, this would be the observations). Our results show that 18 (69.2%) out of 26 reservoirs have an NSE value greater than zero. By using the error performance criteria for NSE suggested by Moriasi et al. (2015), 12 reservoirs (46.15%) show an NSE value greater than 0.5 (categorized as ‘Satisfactory’ or above), among which one reservoir (3.85%) has an NSE value between 0.5 and 0.7 (categorized as ‘Satisfactory’), two reservoirs (7.70%) have an NSE value between 0.7 and 0.8 (categorized as ‘Good’), nine reservoirs (34.62%) have an NSE value greater than 0.8 (categorized as ‘Very Good’; the highest ranked category). For a “quantile” NSE, the mean, standard deviation and median among 26 reservoirs are -36.72, 157.66 and 0.41. The results according to the “quantile” NSE on each performance evaluation category are similar to those according to an “All” NSE: In a total of 15 reservoirs (57.69%) have an NSE value greater than zero, indicating a higher predictive skill than the mean of observations. Among these 15 reservoirs, 12 reservoirs (46.15%) show an NSE value greater than 0.5 (categorized as ‘Satisfactory’ or above), one reservoir (3.85%) has an NSE value between 0.5 and 0.7 (categorized as ‘Satisfactory’), three reservoirs (11.54%) have an NSE value between 0.7 and 0.8 (categorized as ‘Good’), eight reservoirs (30.77%) have an NSE value greater than 0.8 (categorized as ‘Very Good’; the highest ranked category). Both versions of NSE show that 12 (46.15%) out of 26 reservoirs have a satisfactory or better premise condition (i.e., NSE > 0.5). Figure 3-1 summarizes the observed conditions of river-store mass conservation of 26 sampled reservoirs across different hydroclimate regions according to a “quantile” NSE.

As evidenced by Knoben et al. (2019), a KGE value greater than -0.41 indicates that a model is a better predictor than the mean of the observations. Like NSE, two versions of KGE were calculated. According to results based on “All” KGE, only six (23.1%) out of 26 sampled reservoirs have a KGE value greater than -0.41, five reservoirs (19.23%) greater than 0 and only one reservoir (3.85%) over 0.41. The mean, standard deviation and median of “All” KGE of 26 sampled reservoirs are -97.05, 270.31 and -5.89. For a “quantile” KGE, eight reservoirs (30.8%) exceed the -0.41 benchmark, three reservoirs (11.54%) have a KGE greater than 0 and one reservoir (3.85%) over 0.41. The mean, standard deviation and median of this version of KGE among 26 sampled reservoirs are -34.33, 65.88 and -2.38.

Gupta et al. (2009) and Knoben et al. (2019) both argued that understandings of model performance according to NSE values should not guide our understandings according to KGE values as there is no unique relationship between these two metrics. Through examples, Knoben et al. (2019) also showed that high NSE values do not necessarily translate to high KGE values.

We fully acknowledge this argument and discuss the possible distinct behaviors of KGE and NSE, and what possibly causes the results from NSE and KGE to be so different.

The fact that 18 reservoirs meet the mean flow benchmark (NSE > 0; Gupta et al. (2009); Nash and Sutcliffe (1970)) according to NSE (15 for a “quantile” NSE) and only six reservoirs meet the mean flow benchmark (KGE > -0.41; Knoben et al. (2019)) according to KGE (eight for a “quantile” KGE) suggests a difference between the two metrics’ behaviors given that the same datasets were used in the calculations.

The variability in the reservoir storage change dataset (perceived as observations) plays a large role in affecting the error behaviors in NSE. Explicitly, NSE is less sensitive to bias when the variability is high in the observations. The degree of variability can be indicated by the coefficient of variations (CV; a ratio of the standard deviation and mean of the sample). The decomposition of NSE suggested that an interplay exists between the bias component and the variability component (Gupta et al., 2009). When the variability in the observation dataset (storage change) is large (i.e., a high CV value), the bias component would be underestimated, resulting in a small contribution to the final value of NSE. Unlike NSE, linear correlation, bias and variability are separate components of KGE. KGE would be more subjective to bias errors than NSE even when the observations have a large CV score, resulting in worse KGE results (Gupta et al., 2009).

It is assumed that one of the most prominent impacts of calculating a “quantile” version of these error metrics (i.e., in the discharge difference or reservoir storage change datasets, filtering out the time steps where the values exceed an interval of 2.5% and 97.5% quantiles) is a decrease of variability in both observation and simulation datasets, in addition to the possible impact on correlation and bias. By comparing NSE and KGE respectively between the “All” and “quantile” values, our results show that six reservoirs have a higher NSE value (i.e., improved NSE result) after the time steps were filtered out by quantiles, while 12 reservoirs have a higher KGE value (i.e., improved KGE result). On the other hand, 20 reservoirs have a lower NSE value while 14 reservoirs have a lower KGE value.

It is difficult to have clear conclusions about the behaviors of NSE or KGE without a decomposition of these error performance metrics into different error components. According to our results, we also see that the mean and standard deviation of the NSE values change substantially between “All” and “Quantile” groups of NSE. From “All” to “Quantile”, the mean and standard deviation of all NSE values changes from -8.04 to -36.72, and from 22.52 to 157.66. Six reservoirs still see an increase in NSE values. This indicates an overall trend of decrease in NSE values with even larger individual differences. This could be explained by having a decrease in the variability in the datasets: as the degree of variability in the observations becomes smaller, NSE is more sensitive to the bias-related errors, overcompensating the possible improvement from having smaller variability-related errors in the datasets for 20 reservoirs. This trend of decrease in NSE values could also be related to a change in linear correlation or bias errors between the observations and simulations after the time steps where measurements that are out of bounds got removed. Meanwhile, the mean and standard deviation of the KGE values change substantially between “All” and “Quantile” groups of KGE. From “All” to “Quantile”, the mean and standard deviation of all KGE values changes from -97.05 to -34.33, and from 270.31 to 65.88. An increase in KGE values is seen in 12 reservoirs with smaller individual differences. This could be explained that in these 12 reservoirs, removing the time steps where large variability errors occur in the datasets to create the quantile version of the KGE results in a trend of improvement in KGE values for the 12 reservoirs. The remaining 14 reservoirs still see a decrease in KGE values, suggesting an also important impact of linear correlation and bias errors.

To summarize, by comparing the results between “All” and “quantile” groups, we have seen a more prominent trend in NSE where the NSE values decrease in 20 reservoirs (76.92%) when the time steps where the values in observations and simulations were out of bounds were removed. But using the same treatment, we have seen a more complicated case for KGE where 12 reservoirs have increased their KGE values while 14 reservoirs decreased. Further analysis on the behaviors of the error metrics is needed where we can have a decomposition of error performance metrics based on different aspects of error components., so we see what aspect of errors is causing a particular error performance metric (e.g., NSE or KGE) to become larger or smaller.

The statistical results according to rRMSE, MRR and SDRR show overall large values of three error performance metrics, enormously exceeding the benchmark set for the evaluation of discharge estimates in Durand et al. (2016), where an rRMSE of 0.35 was used as a threshold benchmark. Even the smallest value of rRMSE ("All”), indicating the best overall rRMSE result found in Lake Mohave with a value of 1.10 was substantially larger than the “0.35” benchmark. This could be mainly attributed to the small denominator impact (i.e., having small values of reservoir storage change) in the relative residuals, described in the previous section. The inflation of rRMSE we see in our results could also be caused by the metric’s large sensitivity to errors where the discharge difference is small (Hagemann et al., 2017).

According to the descriptive statistics on rRMSE results in our sample, overall, we see a trend of improvement in rRMSE results (or, decrease in rRMSE values) after using a “quantile” technique. The mean, standard deviation and median for the “All” rRMSE values in our sample give 19.91, 33.25 and 5.06. For “quantile” rRMSE values, the mean, standard deviation and median are 7.25, 17.33 and 1.57. In addition, according to a “quantile” rRMSE, all our reservoirs, 26 in total, have an improved rRMSE result indicating a decrease in total errors.

The mean, standard deviation and median for the “All” SDRR values are 19.89, 33.25 and 5.05, for “quantile” SDRR values are 7.20, 17.33 and 1.56. In addition, according to a “quantile” SDRR, all our reservoirs, 26 in total, have an improved SDRR result indicating a decrease in time-varying errors.

The mean, standard deviation and median for the “All” MRR values are 0.57, 0.85 and 0.37, and for “quantile” MRR values are 0.38, 0.64 and 0.33. According to a “quantile” MRR, 19 reservoirs (73.08%) out of 26 have an improved MRR result, indicating a smaller bias error, while seven reservoirs (26.92%) have a worse MRR result, indicating an increase in bias errors.

Despite an increase in bias errors (MRR) for seven reservoirs after a “quantile” technique was applied, the rRMSE values end up being lower for these seven reservoirs with also lower SDRR values.

The distribution-free overlapping index has shown a mean of 0.59, a standard deviation of 0.21 and a median of 0.61 among 26 reservoirs. Figure 3-2 summarizes the observed conditions of river-store mass conservation for 26 studied reservoirs according to six error metrics (calculated in two versions).

Our results according to NSE have shown that 18 reservoirs (69.23%) have shown that the discharge difference time series has a greater explanatory power than the mean of reservoir storage changes (NSE > 0), among which 12 reservoirs (46.15%) have met the “satisfactory” or above criteria according to a performance evaluation criterion suggested by Moriasi et al. (2015). Our results indicate that for nearly half reservoirs in our sample, the river-store mass conservation can indeed be characterized by reservoir storage change and discharge difference between reservoir's major inflow and outflow reaches.

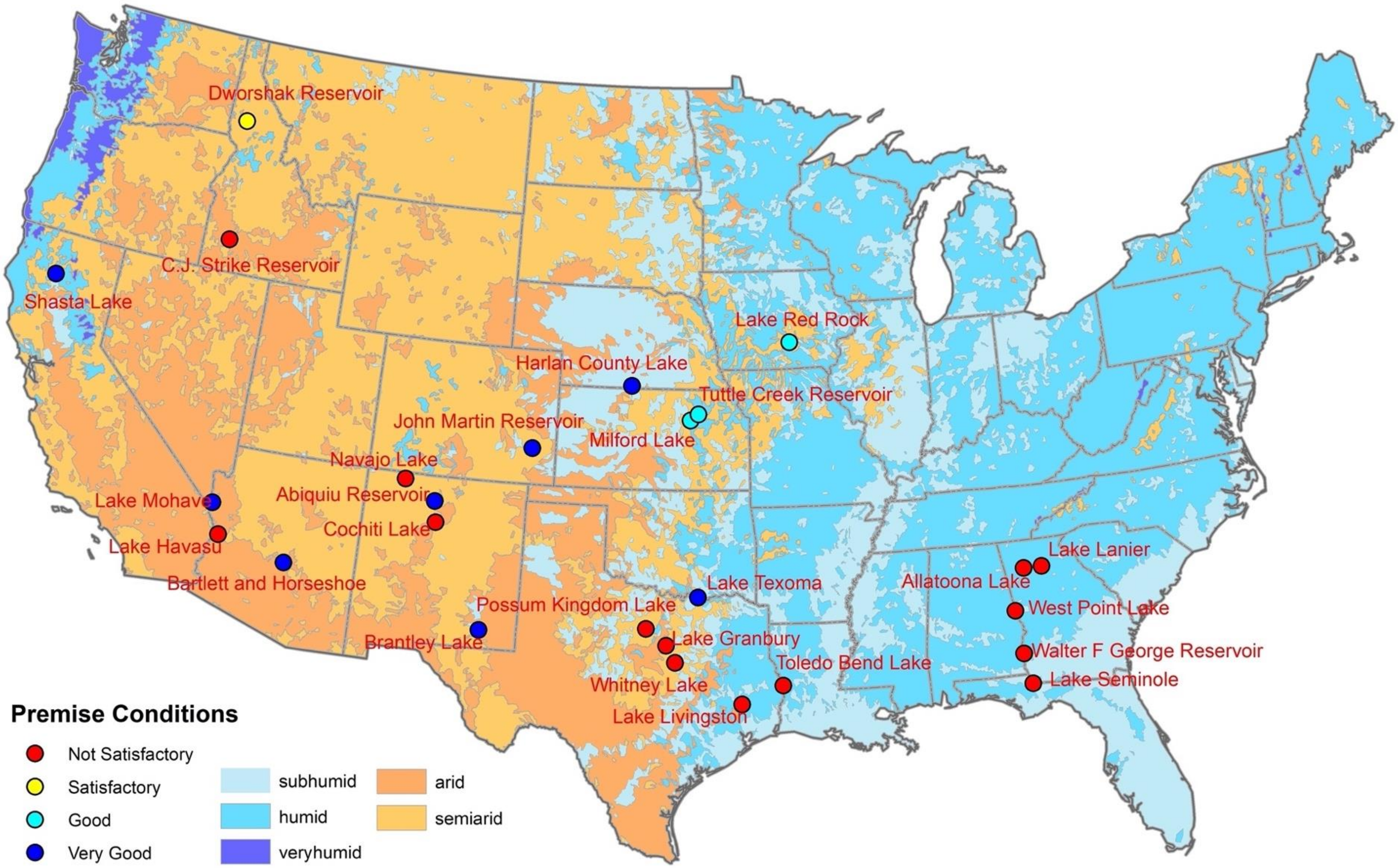


Figure 3-1 Premise conditions of reservoirs across different hydroclimate regions in contiguous United States, showing as four classes in an ascending order: "Not Satisfactory" (NSE < 0.5), "Satisfactory" (0.5 < NSE < 0.7), "Good" (0.7 < NSE < 0.8) and "Very Good" (NSE > 0.8).

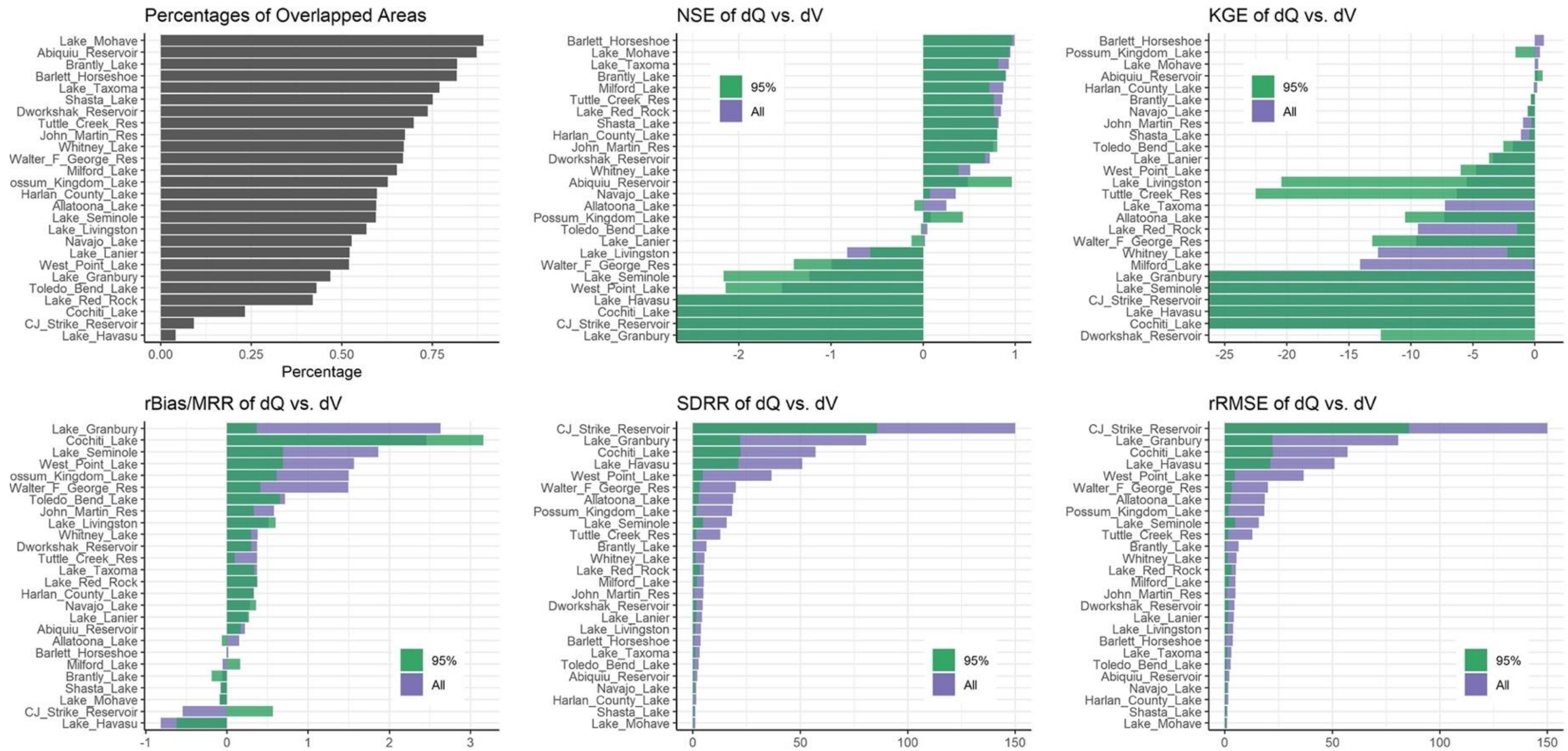


Figure 3-2 Observed conditions of river-store mass conservation for 26 reservoirs according to six error metrics (outliners removed using percentiles)

Table 3.1 Observed conditions of river-store mass conservation for 26 reservoirs according to an overlapping index, NSE and KGE ("ALL" are results from all observations, and "Quan" are results after removing outliners)

Reservoirs	Overlap	NSE		KGE	
		All	Quan.	All	Quan.
Red Rock	0.42	0.85	0.77	-9.43	-1.43
Barlett Horseshoe	0.82	0.99	0.97	0.72	0.01
Dworkshak	0.74	0.73	0.67		-12.43
Livingston	0.57	-0.82	-0.57	-5.5	-20.44
Toledo Bend	0.43	0.05	-0.02	-1.77	-2.55
Seminole	0.59	-1.23	-2.17	-93.29	-130.44
Shasta	0.75	0.82	0.81	-1.12	-0.43
John Martin	0.67	0.76	0.81	-0.95	-0.25
Taxoma	0.77	0.93	0.82	-7.25	-0.1
Tuttle Creek	0.7	0.86	0.77	-6.29	-22.5
Walter F George	0.67	-0.99	-1.4	-9.57	-13.11
Abiquiu	0.87	0.48	0.96	0.25	0.62
Allatoona	0.59	0.25	-0.09	-7.3	-10.46
Brantly	0.82	0.9	0.89	-0.29	-0.34
CJ Strike	0.09	-51.89	-805.22	-312.24	-219.09
Cochiti	0.23	-36.01	-60.82	-1214.62	-177.05
Harlan County	0.6	0.8	0.81	0.19	-0.09
Granbury	0.47	-97.08	-50.04	-91.73	-82.58
Havasu	0.04	-30.56	-43.03	-631.2	-185.89
Lanier	0.52	0.02	-0.13	-3.39	-3.69
Mohave	0.89	0.95	0.93	0.27	0.02
Milford	0.65	0.87	0.72	-14.1	-0.13
Navajo	0.53	0.35	0.08	-0.54	-0.57
Possum Kingdom	0.63	0.08	0.43	0.4	-1.56
West Point	0.52	-1.53	-2.14	-4.76	-5.98
Whitney	0.67	0.51	0.38	-12.63	-2.21
Mean	0.59	-8.04	-36.72	-97.05	-34.33
Median	0.61	0.42	0.41	-5.89	-2.38
Standard Deviation	0.21	22.52	157.66	270.31	65.88

Table 3.2 Observed conditions of river-store mass conservation for 26 reservoirs according to rRMSE, MRR and SDRR ("ALL" are results from all observations, and "Quan" are results after removing outliners)

Reservoirs	rRMSE		MRR		SDRR	
	All	Quan.	All	Quan.	All	Quan.
Red Rock	5.09	3.25	0.37	0.37	5.08	3.22
Barlett Horseshoe	3.66	0.59	0.02	0	3.66	0.59
Dworkshak	4.54	1.83	0.37	0.3	4.53	1.8
Livingston	3.83	1.32	0.51	0.6	3.79	1.17
Toledo Bend	2.77	0.94	0.72	0.65	2.68	0.68
Seminole	15.87	4.85	1.86	0.69	15.76	4.8
Shasta	1.19	0.6	-0.08	-0.07	1.19	0.6
John Martin	4.92	0.8	0.58	0.33	4.89	0.73
Taxoma	3.15	1.11	0.37	0.34	3.13	1.06
Tuttle Creek	12.81	1.59	0.37	0.1	12.81	1.59
Walter F George	20.03	3.23	1.49	0.41	19.98	3.2
Abiquiu	2.13	0.69	0.22	0.17	2.12	0.67
Allatoona	18.69	2.74	0.15	-0.06	18.69	2.74
Brantly	6.36	0.5	-0.06	-0.19	6.36	0.47
CJ Strike	149.97	85.66	-0.55	0.57	149.97	85.65
Cochiti	57.07	22.35	2.45	3.16	57.02	22.13
Harlan County	1.66	0.52	0.33	0.32	1.63	0.41
Granbury	80.69	22.11	2.63	0.37	80.65	22.1
Havasu	50.95	21.18	-0.81	-0.62	50.94	21.17
Lanier	4.23	1.56	0.25	0.27	4.22	1.53
Mohave	1.1	0.5	-0.09	-0.08	1.09	0.49
Milford	5.02	1.92	-0.05	0.16	5.02	1.91
Navajo	1.67	0.92	0.28	0.36	1.65	0.85
Possum Kingdom	18.35	1.79	1.5	0.61	18.29	1.68
West Point	36.66	4.64	1.56	0.69	36.63	4.59
Whitney	5.36	1.36	0.38	0.3	5.34	1.33
Mean	19.91	7.25	0.57	0.38	19.89	7.2
Median	5.06	1.57	0.37	0.33	5.05	1.56
Standard Deviation	33.25	17.33	0.85	0.64	33.25	17.33

3.2 Secondary Controls

To assess the overall scales and variabilities of the impact of secondary controls, we use MRR and SDRR as they allow for a more in-depth understanding of the error compositions, as MRR and SDRR characterize errors that are, respectively, bias-based and time-variability-based.

3.2.1 Individual Patterns of Secondary Controls

Lake Red Rock is located at the interface of semiarid and humid regions of Iowa. This lake is classified as being in semiarid regions in this study. Its main use according to GRanD is flood control. An “All” NSE of 0.85 exceeds “Very Good” criterion suggested by Moriasi et al. (2015). Discharge difference overestimates reservoir storage change in most months during the studied period. However, every June, the discharge difference substantially underestimates reservoir storage change. It might be attributed to seasonal flash flood events in these regions.

Bartlett and Horseshoe are two cascade reservoirs located at the interface of semiarid and arid regions of Arizona. This lake is classified as being in semiarid regions in this study. Their main use is water supply. The premise condition stands firm with an “All” NSE of 0.99. Discharge difference and reservoir storage change align very closely throughout the duration of interest, indicating a strong observed condition of river-store mass conservation.

Dworshak Reservoir is located in the semiarid region of Idaho. Its main use is flood control with second as hydroelectricity and water supply. Discharge difference underestimates reservoir storage change overall with substantial volume gaps during late spring and early summer. An “All” NSE for Dworshak is 0.73.

Lake Livingston is located in the humid region of Texas. Its main use is flood control. With an “All” NSE of -0.82, bias can be observed between discharge difference and reservoir storage change, and discharge difference underestimates reservoir storage change overall. During some months, the bias can be quite substantial.

Toledo Bend Lake is located in the humid region of Louisiana. Its main use is hydroelectricity with a second as water supply. The observed condition of river-store mass conservation is complicated, likely subjective to both human withdrawal and tributary flows. With an “All” NSE of 0.05, discharge difference underestimates reservoir storage change in most months. The greatest gaps between discharge difference and reservoir storage change occur during winter, spring and early summer.

Lake Seminole is located in the subhumid region of Florida. Its main use is navigation, with second use of hydroelectricity and flood control. With an “All” NSE of -1.23, the observed condition of river-store mass conservation is not strong. Discharge difference underestimates reservoir storage change in most months. Some greatest gaps occur during spring, fall and winter.

Shasta Lake is located in the interface of humid and subhumid regions of California. This lake is classified as being in humid regions. Its main use is irrigation, and its second use are hydroelectricity, navigation and water supply. The premise condition stands firm with an “All” NSE of 0.82. The discharge difference underestimates reservoir storage change generally only in a very small extent. The deviation of discharge difference from reservoir storage changes mostly occurs during winter and summer.

John Martin Reservoir is located in the semiarid region of Colorado. Its main use is flood control with a second use as irrigation. The premise condition stands firm in this reservoir with an “All” NSE of 0.76. Discharge difference and reservoir storage change align well in most month while discharge difference overestimates reservoir storage change during summer and fall.

Lake Taxoma is located at the interface of humid and semiarid regions of Oklahoma. This lake is classified as being in humid regions. Its main use is flood control, with a second use as water supply. With an “All” NSE of 0.93, discharge difference aligns with reservoir storage change except for several months in summer and fall when the discharge difference underestimates reservoir storage change. This might be caused by the water withdrawal directly from the lake body lowering down reservoir storage change.

Tuttle Creek Reservoir is located at the interface of humid and subhumid regions of Kansas. This lake is classified as being in humid regions. Its main use is flood control with a second use as navigation. The premise condition stands firm with an “All” NSE of 0.86. The discharge difference slightly underestimates reservoir storage change.

Walter F George reservoir is located in the humid region of Alabama. Its main use is navigation with 2 second use as hydroelectricity and flood control. With an “All” NSE of -0.99, discharge difference underestimates reservoir storage change in general.

Abiquiu Reservoir is located in the semiarid region of New Mexico. Its main use is flood control with two second use as irrigation and water supply. The premise stands firm with an NSE of 0.48. discharge difference slightly underestimates reservoir storage change in Spring. In other months they align well.

Allatoona Lake is located in the humid region of Georgia. Its main use is flood control, and its second use are water supply and hydroelectricity. With an “All” NSE of 0.25, discharge difference and reservoir storage change match in trends but with substantial gaps in between. discharge difference underestimates reservoir storage changes generally.

Brantley Lake is located in the arid region of New Mexico. Its main use is flood control with the second use being irrigation. The premise stands firm with an “All” NSE of 0.90. discharge difference slightly overestimates reservoir storage change during late spring and early summer.

CJ Strike reservoir is located in the arid region of Idaho. Its main use is hydroelectricity with second use as irrigation. With an “All” NSE of -51.89, discharge difference and reservoir storage change have very weak relationship except in March and April. Part of the reason might be attributed to the distance between inflow gauge to lake body (~50km). Summer flash floods in the region where inflow gauge is located might also be responsible. Many canals used for water supply was also found around the lake body.

Cochiti Lake is located in the semiarid region of New Mexico. Its main use is flood control with a second use as irrigation. With an “All” NSE of -36.01, discharge difference and reservoir storage change have weak relationship except for winter. discharge difference overestimates reservoir storage changes generally. This could be attributed to the distance between inflow gauge and lake (16km). The main reason behind a weak observed premise condition might be water withdrawal from the lake body.

Harlan County Lake is located in the subhumid region of Nebraska. Its main use is flood control with a second use as irrigation. The general trends of discharge difference and reservoir storage change match well with an “All” NSE of 0.80. discharge difference underestimates reservoir storage change during spring and overestimates it in other seasons, which might be attributed to precipitation change during the seasons, intensified by irrigation use.

Lake Granbury is located at the interface of humid and semiarid regions of Texas. This lake is classified as being in humid regions. Its main use is irrigation. With an “All” NSE of -

97.08, there is a relatively poor discharge difference reservoir storage change relationship. discharge difference underestimates reservoir storage change in most seasons.

Lake Havasu is in the semiarid region of Arizona. Its main use is water supply with two second use as irrigation and hydroelectricity. The trends of discharge difference and reservoir storage change matches well. But the bias between reservoir storage change and discharge difference is great with an “All” NSE of -30.56. The main reason behind this deviation might be attributed to the pumping plants and canals that are withdrawing water from the main lake, like Colorado River Aqueduct.

Lake Lanier is located in the humid region of Georgia. Its main use is flood control with a second use as hydroelectricity. The trends of discharge difference and reservoir storage change match well. discharge difference underestimates reservoir storage changes substantially in most seasons with an “All” NSE of 0.02.

Lake Mohave is located in the arid region of Arizona. Its main use is hydroelectricity. The trends match ideally showing a strong discharge difference-reservoir storage change relationship. discharge difference overestimates reservoir storage changes slightly in summer with an “All” NSE of 0.95.

Milford Lake is located at the interface of semiarid and subhumid regions of Kansas. This lake is classified as being in semiarid regions. Its main use is flood control, and its second use is water supply. The relationship of discharge difference and reservoir storage change is very strong in general with an “All” NSE of 0.87.

Navajo Reservoir is located in the semiarid region of New Mexico. Its main use is irrigation. discharge difference underestimates reservoir storage change in some seasons with an “All” NSE of 0.35. This could be attributed to the uneven precipitation between seasons.

Possum Kingdom Lake is located in the semiarid region of Texas. Its main use is water supply. The relationship of reservoir storage change and discharge difference is good for most of the months with an “All” NSE of 0.08.

Whitney Lake is located in the semiarid region of Texas. Its main use is water supply with two second use as flood control and hydroelectricity. With an “All” NSE of 0.51, discharge difference and reservoir storage change have an overall good relationship with discharge difference slightly underestimates reservoir storage change.

3.2.2 General Patterns of Secondary Controls

It is then important to understand the impacts of secondary controls and what might affect the scales of their impacts. The hydroclimate regions where reservoirs are located in were identified using the Hydrologic Landscape Region (HLR) dataset. Among the 26 reservoirs sampled in our study, four are located arid regions, ten in semiarid regions, three in subhumid regions, and nine in humid regions. The main and second uses of reservoirs and dams are available on the Global Reservoir and Dam dataset (GRanD), a “geographically explicit and reliable” database of over 7000 dams worldwide (Lehner et al., 2011). 18 reservoirs (69.23%) fall into the category of water withdrawal dams (i.e., having a main or second use as water supply or irrigation) according to GRanD.

With the reservoir purpose information provided by GRanD, we calculated the mean, median and the standard deviation of error performance values from water-supply dams and non-water-supply dams, respectively, according to an overlapping index, NSE, KGE, rRMSE, MRR and SDRR. The term “Water-supply dams/reservoirs” refers to dams/reservoirs that have a main or second use of water supply or irrigation documented in GRanD. Table 3.3 and Table 3.4

summarize the statistical results according to different error performance metrics for water-supply reservoirs and non-water-supply reservoirs, respectively, using descriptive statistics (mean, standard deviation and median).

By comparing water-supply reservoirs to non-water-supply reservoirs, the average NSE value ("All" version; so as below) increases from -11.50 to -0.24, the average KGE value increases from -134.95 to -16.49, the average rRMSE value decreases from 23.23 to 12.45 the average SDRR value decreases from 23.21 to 12.42, the average MRR value increases from -0.47 to 0.79 and the overlapping score increases from 0.58 to 0.61. The average overlapping index value shows an increase from 0.58 to 0.61 by comparing water-supply reservoirs to non-water-supply reservoirs, not as substantial as what occurs with NSE, KGE and rRMSE. The results according to the "quantile" versions of error performance metrics show the same trends of variations. NSE, KGE, rRMSE and the overlapping index are all measures of overall error performance. Our results show that, the average error performance for non-water-supply reservoirs according to NSE, KGE, rRMSE and SDRR are all largely better than those for water-supply reservoirs, with the exception of MRR showing a decline in error performance. The average MRR ("All") value for water-supply reservoirs is 0.47, for non-water-supply dams it is 0.79. The average MRR ("quantile") for water-supply reservoirs is 0.37, lower, but barely, than the average value of 0.38 for non-water-supply reservoirs. In addition to a comparison on the mean values, the comparison on standard deviations between the two reservoir purpose categories shows a consistent result where the standard deviations of all error performance metrics are substantially smaller for non-water-supply reservoirs than for water-supply reservoirs.

Unlike other error performance metrics, having a larger mean of MRRs in non-water-supply reservoirs suggests that human water withdrawal likely reduces the bias-related errors in water-supply reservoirs. Having a smaller mean of SDRRs in non-water-supply reservoirs suggests that human water withdrawal likely increases the time-varying errors in water-supply reservoirs. Therefore, based on our results, we strongly argue that human water withdrawal largely affects the ability of which gauge-measured reservoir storage change and discharge difference characterize the river-store mass conservation, and human water withdrawal on reservoirs is a large cause of a less satisfactory premise condition for LakeFlow, by largely increasing more time-varying errors.

The descriptive statistics of our results using different error performance metrics (an overlapping index, NSE, KGE, rRMSE, MRR and SDRR) were calculated in four hydroclimate conditions (arid, semiarid, subhumid, humid), respectively. The descriptive statistics are summarized in Table 3.5, 3.6, 3.7 and 3.8, each corresponding to one hydroclimate condition: arid, semiarid, subhumid and humid.

The overlapping index, NSE, KGE and rRMSE are measures of similarity between two datasets in a single value. The overlapping index for reservoirs shows similar means in four hydroclimate conditions but higher standard deviations in arid and semiarid regions than in subhumid and humid regions (means: 0.62, 0.56, 0.57, 0.61; standard deviations: 0.31, 0.25, 0.04, 0.11; in an order of arid, semiarid, subhumid, humid regions). This suggests that though a similar mean value exists in all hydroclimate regions, the actual overlapping index values for reservoirs are closer to this mean in subhumid and humid regions than in arid and semiarid regions. Between the four hydroclimate regions, the best NSE result with the largest mean and the smallest standard deviation of NSE values occur together in subhumid regions (both in “All” and “quantile” versions). The best KGE result with the largest mean and the smallest standard deviation of KGE values occur together in humid region (both ins “All” and “quantile” versions). According to the “All” version of rRMSEs, the smallest average rRMSE for a hydroclimate region occurs in the semiarid regions, while the smallest standard deviation of rRMSE for a hydroclimate occurs in the subhumid regions. But it is important to point out that the average rRMSE value in semiarid regions (15.38) is not substantially lower than that of humid regions (16.38) or subhumid regions (18.07), as it otherwise appears in the NSE or KGE. The smallest average MRR value occurs in, surprisingly, arid region while the smallest average SDRR value occurs in the semiarid regions, though its average SDRR value is not smaller than that of humid or subhumid, by much.

According to the “quantile” version of rRMSEs, the smallest average rRMSE for a hydroclimate region occurs in the subhumid region, while the smallest standard deviation of rRMSE for a hydroclimate region also occurs in the subhumid regions. As suggested in the previous section, a “quantile” version of rRMSE needs to be used with discretion due to the complicated nature of its impact on MRR and SDRR. Therefore, for the discussion here, we will only defer to an “All” version of rRMSE.

Our results have suggested a discrepancy among results from different error performance metrics in four hydroclimate regions. While the overlapping index results are similar among four regions, subhumid regions have the overall best NSE results, humid regions has the overall best KGE results, and semiarid regions has the overall best rRMSE results (though not by much). As two products from a decomposition of rRMSE, the best MRR results can be found in arid regions while the best SDRR can be found in semiarid regions. Therefore, we asked for alternative ways to descriptive statistics in order to analyze the possible impact of hydroclimate conditions on secondary controls.

The monthly hydrographs are generated for each reservoir in this study that demonstrates, at a monthly timescale, the variations of reservoir storage change and discharge difference between the reservoir's major inflow and outflow for two years. The scatterplots are also generated for each reservoir. The disadvantage of using such hydrographs and scatterplots alone for analysis is that the findings are not quantitative based (like descriptive statistics), and therefore might lack enough accuracy. But these plots remain a nice tool that allows us to discern the general trends or patterns in the data analysis process. This is indeed one limitation of this study which calls for future studies to conduct quantitative analysis that perhaps addresses the seasonal changes in reservoir water balance. Figures 3-3 and 3-4 are the monthly hydrograph and scatterplot for West Point Lake. We will first give an example of how the interpretation is conducted on one reservoir, West Point Lake, then summarize the overall patterns and trends we have identified from interpreting the plots for all reservoirs in this study, lastly followed by interpretations for each of the remaining reservoirs. The scatterplots and monthly hydrographs for all reservoirs in our study can be found in the Appendix.

West Point Lake is located in Georgia and belongs to the subhumid region according to the HLR dataset. Its main use, according to GRanD dataset, is flood control with a second use as hydroelectricity. The discharge difference (blue) and reservoir storage change (red) lines share similar trends of variations, but substantial bias errors can be seen with the largest variance occurring in the late winter and spring season on the hydrograph. The scatterplot also shows that discharge difference substantially underestimates reservoir storage change. Our plots imply a substantial impact of lateral inflow from the tributaries that are too small for SWOT to observe on the river-store mass conservation characterized by reservoir storage change and discharge difference.

According to our interpretations on each reservoir in this study, we have found that, at large, discharge difference tends to underestimate reservoir storage change (meaning reservoir gaining water from other sources) in more humid regions, implying a substantial impact of lateral inflow from the tributaries that are too small for SWOT to observe. In arid and semiarid regions, discharge difference overestimates reservoir storage change only during certain months (notably, late spring), suggesting a non-negligible impact of reservoir surface evaporation, though an underestimation of discharge difference on reservoir storage change still occurs widely.

Our findings based on visual interpretations of plots are consistent with the results according to MRR. A negative value of MRR indicates an overestimation of discharge difference on reservoir storage change. A positive value of MRR indicates an underestimation of discharge difference on reservoir storage change. Our results have shown six reservoirs with a negative MRR value indicating an overestimation, with the remaining 20 with a positive MRR value indicating an underestimation. Among the six reservoirs that have a negative MRR value, three are located in arid regions, two in semiarid regions and one in humid regions. According to our results based on MRR, we can argue that: first, an underestimation of discharge difference on reservoir storage change is more prevalent than an overestimation among the reservoirs in our study. Second, if an overestimation occurs, it is more common to see it in drier regions.

These findings lead us into thinking what are the possible explanations behind these phenomena (i.e., a widespread underestimation of discharge difference; an overestimation more easily to occur in dry regions). And what roles might the secondary controls (specifically focusing on human water withdrawals, lateral inflows and reservoir surface evaporation) play.

As discussed in Section 2.4, reservoir surface evaporation, lateral inflow and human water withdrawal (for irrigation or water supply) are several secondary controls that might have a large impact on the river-store mass conservation characterized by reservoir storage change and discharge difference. Among the three major secondary controls, lateral inflow is the only water flux considered as an input in a reservoir's water balance, while reservoir surface evaporation and human water withdrawal (directly from lake storage) are all fluxes that leave the system as water outfluxes. The underestimation of discharge difference on reservoir storage change in our reservoirs indicates the existence of additional water influx in the water balance. Therefore, we argue that lateral inflow, being the only water influx among the three major secondary controls, is the main cause of the first observed phenomenon, the reason that such underestimation commonly occurs in our sample. Further, the scales of this influx from lateral inflow might be even larger than the water outfluxes from other secondary controls, namely surface evaporation and human water withdrawal. This finding exceeds our previous expectation on the impact of lateral inflow being more prevalent in humid regions, and implicates that one possible way to improve the observed river-store mass conservation for a better premise condition for LakeFlow, is to account for the contribution of the lateral inflow in the studied reservoirs, no matter in which hydroclimate region they are located. This is precisely what the focus of this study will be in Section 3.3.

Our results suggest that the overestimation of discharge difference on reservoir storage change is a lot less common in our sample at large (6 out of 26 reservoirs) and is more likely to occur in dry regions (three in arid regions, two in semiarid regions and one in humid region). We argue that this could be attributed to both reservoir surface evaporation and human water withdrawal. It is likely to have high intensity of water withdrawal from the lake for irrigation use and domestic water supply in the arid regions. A dry environment and possibly higher average annual temperature could also incur a higher level of surface evaporation. We might be able to improve the observed river-store mass conservation for a better premise condition for LakeFlow, by accounting for reservoir surface evaporation in drier regions, though the improvement might not be effective and could even incur more errors.

Figure 3-3 Monthly hydrograph showing reservoir storage change (Observed dV; blue line), discharge difference (Observed dQ; red), discharge difference subtracted by estimated reservoir surface evaporation (Observed dQ – E; green), discharge difference added by modeled lateral inflow (Observed dQ + q; orange), discharge difference subtracted by estimated reservoir surface evaporation and added by modeled lateral inflow (Observed dQ – E + q; purple), estimated discharge difference with SWOT's temporal sampling (Estimated dQ; gray)

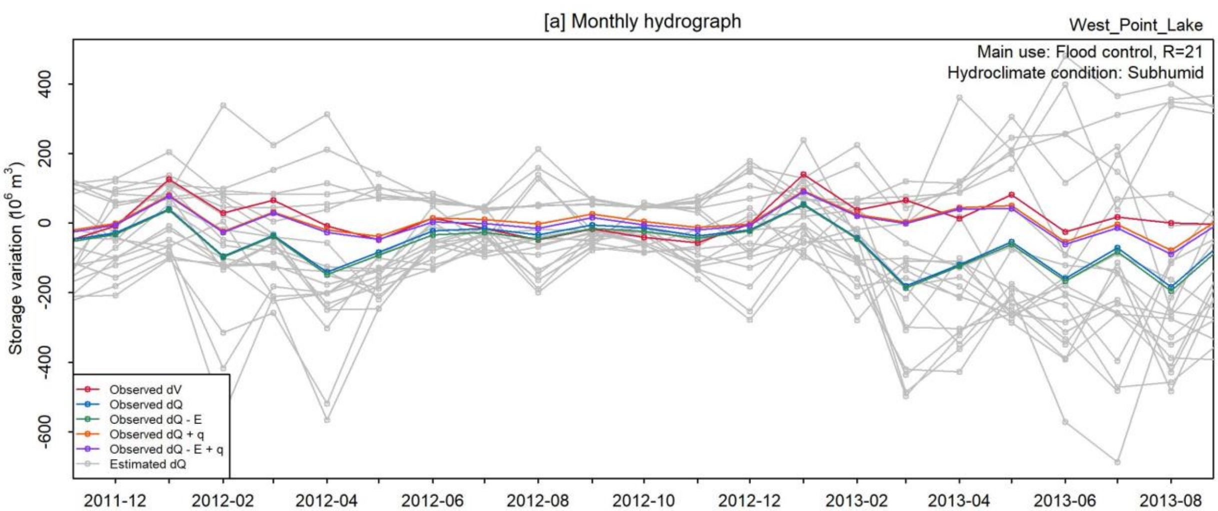


Figure 3-4 A scatterplot showing reservoir storage change against discharge difference in a unit of one million cubic meter at a timescale of the SWOT revisit time (R) in West Point Lake

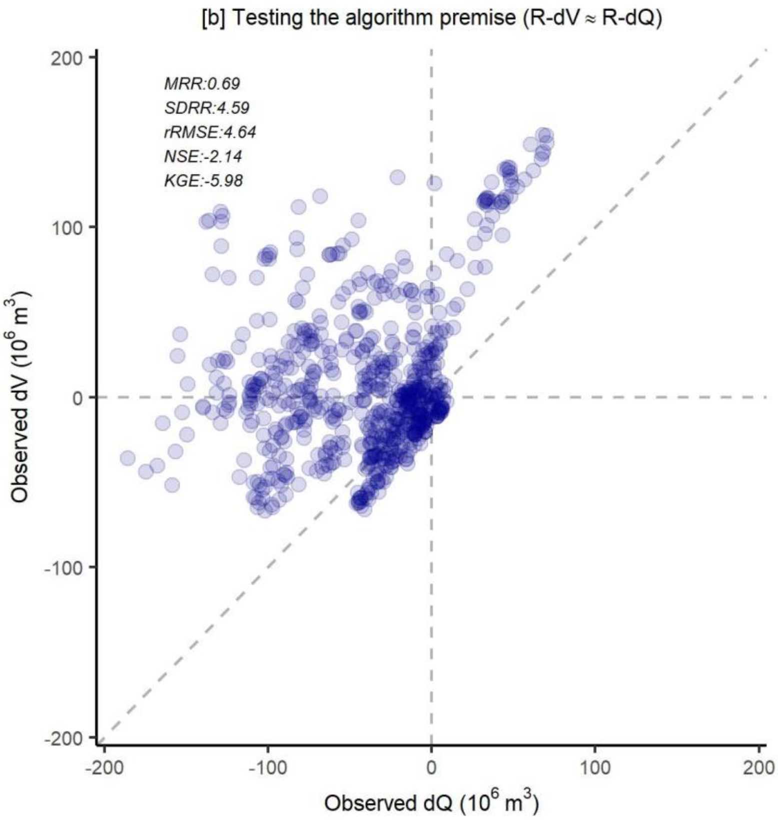


Table 3.3 Summary of an overlapping index, NSE, KGE, rRMSE, MRR and SDRR for water-supply dams.

Descriptive Statistics	Overlap. Index	NSE		KGE		rRMSE		MRR		SDRR	
		All	Quan.	All	Quan.	All	Quan.	All	Quan.	All	Quan.
MEAN	0.58	-11.50	-52.83	-134.95	-38.62	23.23	9.31	0.47	0.37	23.21	9.25
STD. DEV.	0.24	25.79	183.52	313.47	72.32	38.01	20.05	0.87	0.74	38.01	20.06
MEDIAN	0.64	0.50	0.55	-1.77	-1.06	4.97	1.24	0.31	0.31	4.96	1.20

Table 3.4 Summary of an overlapping index, NSE, KGE, rRMSE, MRR and SDRR for non-water-supply dams.

Descriptive Statistics	Overlap. Index	NSE		KGE		rRMSE		MRR		SDRR	
		All	Quan.	All	Quan.	All	Quan.	All	Quan.	All	Quan.
MEAN	0.61	-0.24	-0.49	-16.49	-24.70	12.45	2.61	0.79	0.38	12.42	2.57
STD. DEV.	0.13	0.96	1.21	29.18	40.75	11.07	1.51	0.68	0.26	11.06	1.51
MEDIAN	0.58	-0.40	-0.35	-5.89	-9.55	8.95	2.41	0.44	0.39	8.94	2.39

Table 3.5 Summary of observed conditions of river-store mass conservation on “arid” hydroclimate conditions according to an overlapping index, NSE, KGE, rRMSE, MRR and SDRR.

Descriptive Statistics	Overlap. Index	NSE		KGE		rRMSE		MRR		SDRR	
		All	Quan.	All	Quan.	All	Quan.	All	Quan.	All	Quan.
MEAN	0.62	-12.32	-200.65	-78.30	-54.91	40.59	21.86	-0.03	0.16	40.58	21.83
STD. DEV.	0.31	22.84	349.05	135.06	94.79	63.18	36.83	0.40	0.31	63.19	36.85
MEDIAN	0.75	0.83	0.85	-0.62	-0.29	5.64	0.65	-0.08	0.13	5.62	0.61

Table 3.6 Summary of observed conditions of river-store mass conservation on "semiarid" hydroclimate conditions according to an overlapping index, NSE, KGE, rRMSE, MRR and SDRR.

Descriptive Statistics	Overlap. Index	NSE		KGE		rRMSE		MRR		SDRR	
		All	Quan.	All	Quan.	All	Quan.	All	Quan.	All	Quan.
MEAN	0.56	-6.17	-9.89	-209.02	-38.06	15.38	5.59	0.47	0.48	15.36	5.54
STD. DEV.	0.25	13.62	21.39	405.65	71.82	19.86	8.13	0.85	0.94	19.85	8.09
MEDIAN	0.64	0.50	0.55	-9.43	-1.49	5.06	1.81	0.33	0.30	5.05	1.74

Table 3.7 Summary of observed conditions of river-store mass conservation on "subhumid" hydroclimate conditions according to an overlapping index, NSE, KGE, rRMSE, MRR and SDRR.

Descriptive Statistics	Overlap. Index	NSE		KGE		rRMSE		MRR		SDRR	
		All	Quan.	All	Quan.	All	Quan.	All	Quan.	All	Quan.
MEAN	0.57	-0.66	-1.17	-32.62	-45.51	18.07	3.34	1.25	0.57	18.01	3.26
STD. DEV.	0.04	1.04	1.40	42.95	60.10	14.37	1.99	0.66	0.17	14.38	2.02
MEDIAN	0.59	-1.23	-2.14	-4.76	-5.98	15.87	4.64	1.56	0.69	15.76	4.59

Table 3.8 Summary of observed conditions of river-store mass conservation on "humid" hydroclimate conditions according to an overlapping index, NSE, KGE, rRMSE, MRR and SDRR.

Descriptive Statistics	Overlap. Index	NSE		KGE		rRMSE		MRR		SDRR	
		All	Quan.	All	Quan.	All	Quan.	All	Quan.	All	Quan.
MEAN	0.61	-10.66	-5.54	-14.88	-17.32	16.38	3.91	0.71	0.29	16.35	3.85
STD. DEV.	0.11	30.56	15.75	27.30	24.37	23.72	6.48	0.80	0.25	23.72	6.51
MEDIAN	0.59	0.05	-0.09	-6.29	-10.46	4.23	1.56	0.37	0.34	4.22	1.53

3.3 Effectiveness of Adding Major Secondary Control

In Section 3.2.2 in which we suggest that accounting for the contribution of lateral inflow could improve the observed river-store mass conservation for a better premise condition for LakeFlow, no matter in which hydroclimate region the reservoir is located. We also suggest that using the same idea on surface evaporation could achieve a similar result for reservoirs located in drier regions but the improvement might not be as effective. Followed by these suggestions, we applied the model simulations of lateral flows (q) and estimations of surface evaporation (E) into our analysis and hope to find out how the inclusion of these secondary controls can compensate or on the other hand, enlarge the residuals between reservoir storage change and discharge difference, and thus improve observed river-store mass conservation for a better premise condition for LakeFlow. For the testing purpose, daily lateral flow time series were acquired from the Global Reach-level A priori Discharge Estimates for SWOT (GRADES) (Lin et al., 2019), and monthly E time series were acquired from the Contiguous United States Reservoir Evaporation Dataset (CRED) (Zhao and Gao, 2019).

We calculated a set of five error performance metrics, including NSE, KGE, rRMSE, MRR and SDRR. Detailed descriptions of these error performance metrics can be found in Section 2.2.3. Figures 3.5, 3.6, 3.7, 3.8 and 3.9 show the observed premise conditions for river-store mass conservation, added with E and q both separately and collectively, according to NSE, KGE, rRMSE, MRR and SDRR. An outliner removal process was completed for each calculation of NSE by using quantiles (shown as “95%” in the following figures).

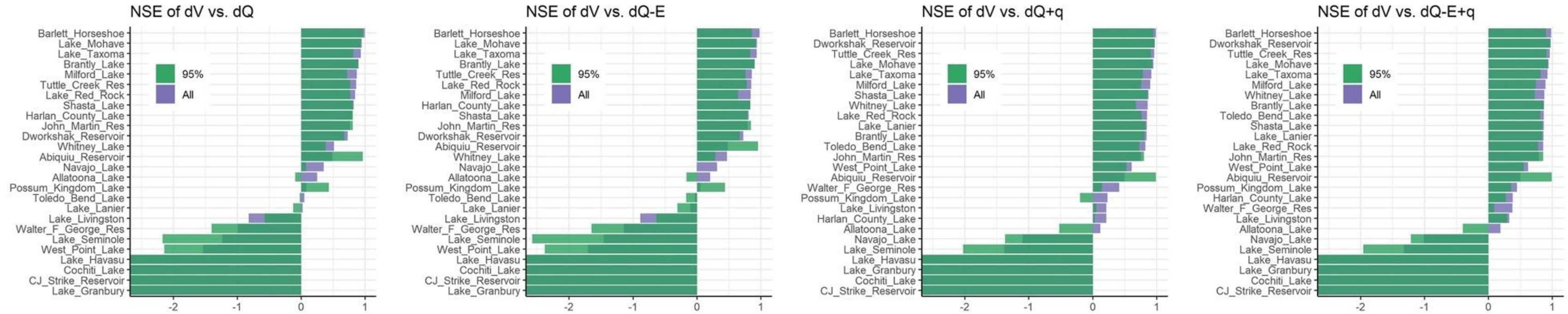


Figure 3-5 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to NSE ("ALL" are results from all observations, and "95%" are results after removing outliners).

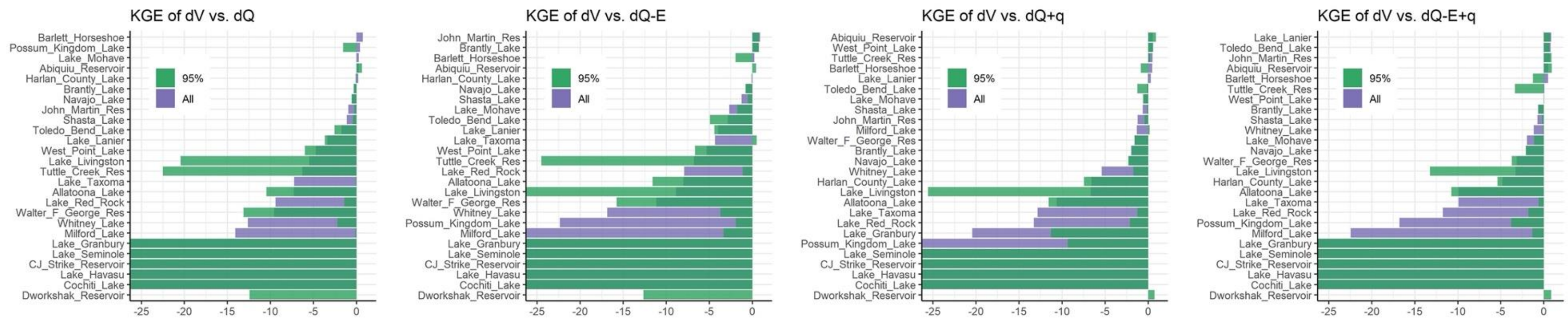


Figure 3-6 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to KGE ("ALL" are results from all observations, and "95%" are results after removing outliners).

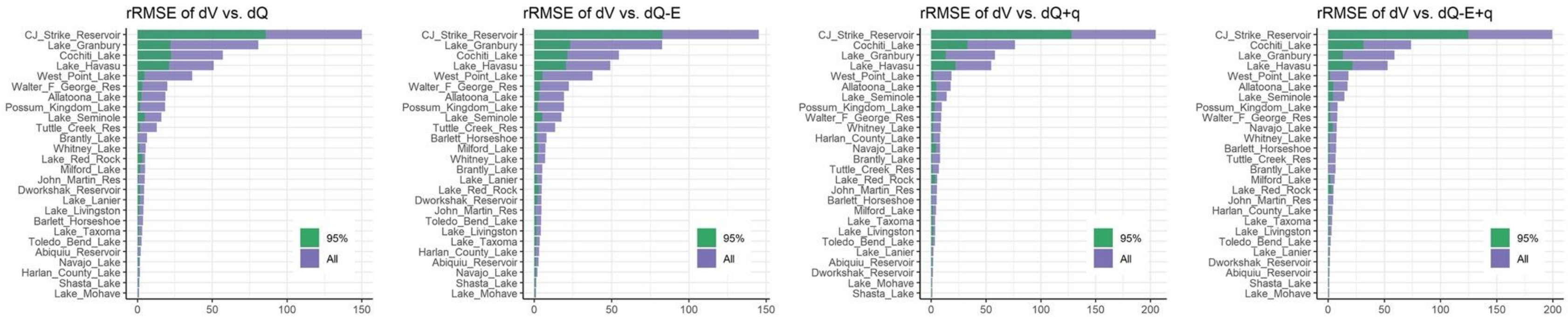


Figure 3-7 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to rRMSE ("ALL" are results from all observations, and "95%" are results after removing outliners).

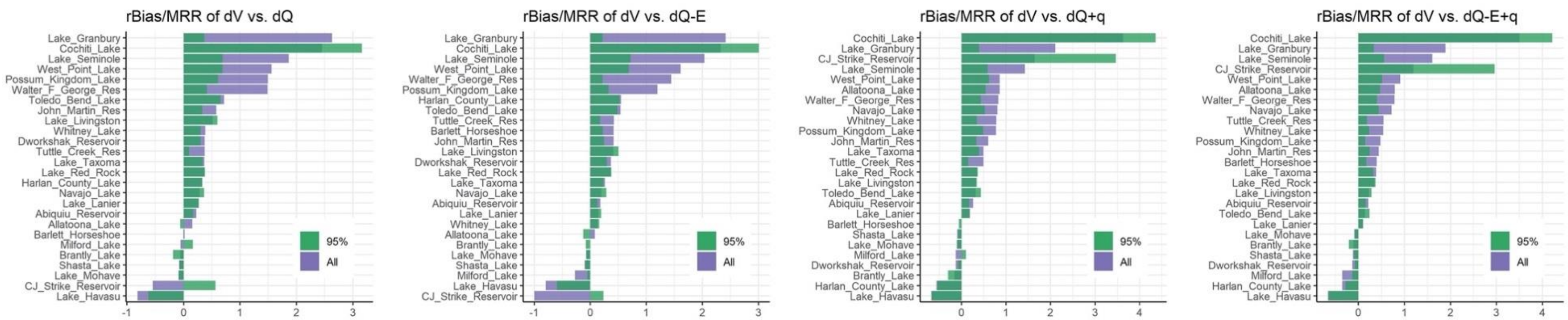


Figure 3-8 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to MRR ("ALL" are results from all observations, and "95%" are results after removing outliners).

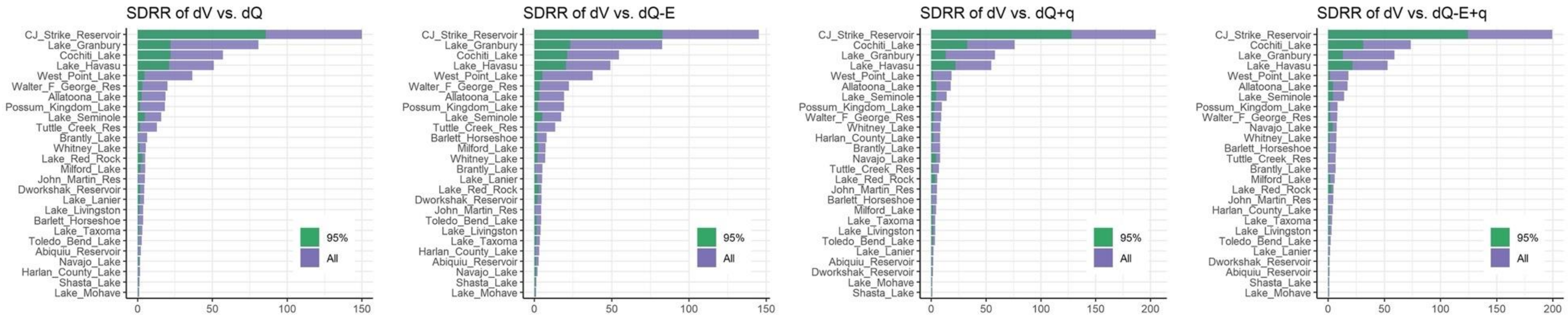


Figure 3-9 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to SDRR ("ALL" are results from all observations, and "95%" are results after removing outliners).

Table 3.9 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to NSE results calculated from all observations.

Reservoirs	NSE			
	dQ	dQ-E	dQ+q	dQ-E+q
Red Rock	0.845	0.849	0.849	**0.854**
Barlett Horseshoe	**0.993**	0.978	0.986	0.981
Dworkshak	0.725	0.723	0.969	**0.969**
Livingston	-0.823	-0.889	0.214	**0.324**
Toledo Bend	0.047	-0.031	0.825	**0.865**
Seminole	**-1.234**	-1.460	-1.387	-1.326
Shasta	0.820	0.809	**0.872**	0.862
John Martin	0.762	**0.790**	0.757	0.789
Taxoma	0.930	**0.934**	0.917	0.923
Tuttle Creek	0.861	0.858	**0.960**	0.959
Walter F George	-0.991	-1.146	**0.413**	0.372
Abiquiu	0.485	0.483	0.496	**0.496**
Allatoona	**0.250**	0.208	0.118	0.184
Brantly	0.899	**0.904**	0.840	0.867
CJ Strike	-51.888	**-48.665**	-102.057	-97.287
Cochiti	-36.012	**-33.581**	-61.407	-57.901
Harlan County	0.800	**0.830**	0.211	0.378
Granbury	-97.079	-100.647	**-53.778**	-54.927
Havasu	-30.563	**-28.527**	-34.476	-32.286
Lanier	0.022	-0.106	0.845	**0.858**
Mohave	**0.948**	0.936	0.948	0.941
Milford	0.869	0.838	**0.901**	0.890
Navajo	**0.354**	0.312	-1.101	-1.015
Possum Kingdom	0.082	0.057	0.228	**0.446**
West Point	-1.534	-1.708	0.608	**0.618**
Whitney	0.511	0.469	0.857	**0.873**
Mean	-8.04	-7.91	-9.25	-8.86
Median	0.42	0.39	0.68	0.70
Standard Deviation	22.52	22.65	25.34	24.37

Table 3.10 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to NSE results with outliers removed from observations.

Reservoirs	NSE (Quan.)			
	dQ	dQ-E	dQ+q	dQ-E+q
Red Rock	0.769	**0.776**	0.767	0.775
Barlett Horseshoe	**0.968**	0.864	0.947	0.900
Dworkshak	0.673	0.670	0.966	**0.966**
Livingston	-0.573	-0.636	0.055	**0.287**
Toledo Bend	-0.024	-0.167	0.730	**0.811**
Seminole	-2.167	-2.579	-2.027	**-1.959**
Shasta	0.806	0.790	**0.854**	0.841
John Martin	0.806	0.845	0.802	**0.855**
Taxoma	**0.816**	0.834	0.784	0.815
Tuttle Creek	0.766	0.759	**0.911**	0.910
Walter F George	-1.403	-1.651	**0.144**	0.089
Abiquiu	0.961	0.955	0.988	**0.989**
Allatoona	**-0.092**	-0.166	-0.524	-0.405
Brantly	0.892	**0.904**	0.811	0.859
CJ Strike	-805.223	**-752.988**	-1628.974	-1550.597
Cochiti	-60.815	**-56.619**	-103.416	-97.812
Harlan County	0.805	**0.837**	0.031	0.273
Granbury	-50.037	-53.794	**-25.032**	-25.542
Havasu	-43.026	**-40.124**	-48.701	-45.598
Lanier	-0.127	-0.306	0.821	**0.836**
Mohave	**0.933**	0.922	0.932	0.926
Milford	0.719	0.645	**0.757**	0.740
Navajo	**0.075**	0.002	-1.373	-1.214
Possum Kingdom	0.431	**0.440**	-0.197	0.356
West Point	-2.143	-2.381	0.529	**0.553**
Whitney	0.383	0.288	0.676	**0.723**
Mean	-36.724	-34.649	-69.144	-65.755
Median	0.407	0.364	0.703	0.732
Standard Deviation	157.657	147.446	318.933	303.595

Table 3.11 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to KGE results calculated from all observations.

Reservoirs	KGE			
	dQ	dQ-E	dQ+q	dQ-E+q
Red Rock	-9.426	**-7.903**	-13.271	-11.748
Barlett Horseshoe	**0.718**	0.245	0.475	0.481
Dworkshak	-inf	-inf	-inf	-inf
Livingston	-5.499	-8.879	-6.672	**-3.287**
Toledo Bend	-1.766	-2.848	-0.157	**0.775**
Seminole	-93.286	-179.674	-142.403	**-56.018**
Shasta	-1.119	-1.246	**-0.620**	-0.747
John Martin	-0.955	**0.895**	-1.197	0.728
Taxoma	-7.249	**-4.337**	-12.843	-9.931
Tuttle Creek	-6.290	-6.770	**0.524**	0.048
Walter F George	-9.574	-11.164	**-1.577**	-3.164
Abiquiu	0.250	0.052	**0.538**	0.529
Allatoona	**-7.299**	-7.998	-10.630	-9.931
Brantly	-0.287	**0.742**	-1.969	-0.665
CJ Strike	-312.243	**-294.841**	-606.845	-589.442
Cochiti	-1214.622	**-1149.507**	-1768.639	-1703.523
Harlan County	**0.193**	-0.105	-6.568	-4.823
Granbury	-91.734	-111.646	**-20.423**	-39.913
Havasu	-631.202	**-608.695**	-676.494	-653.986
Lanier	-3.390	-3.961	0.310	**0.867**
Mohave	**0.265**	-2.688	-0.469	-1.950
Milford	-14.098	-39.870	**-1.318**	-22.458
Navajo	**-0.536**	-0.727	-2.221	-2.022
Possum Kingdom	**0.397**	-22.376	-40.571	-16.768
West Point	-4.756	-5.333	**0.535**	0.033
Whitney	-12.632	-16.856	-5.388	**-1.163**
Mean	-97.046	-99.420	-132.716	-125.123
Median	-5.895	-7.337	-3.804	-3.225
Standard Deviation	270.312	256.522	383.872	370.340

Table 3.12 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to KGE results with outliers removed from observations.

Reservoirs	KGE (Quan.)			
	dQ	dQ-E	dQ+q	dQ-E+q
Red Rock	-1.430	**-1.118**	-2.119	-1.805
Barlett Horseshoe	**0.006**	-1.938	-0.838	-1.268
Dworkshak	-12.430	-12.648	0.729	**0.852**
Livingston	-20.435	-43.159	-25.577	**-13.237**
Toledo Bend	-2.552	-4.914	-1.242	**0.655**
Seminole	**-130.438**	-956.154	-1294.061	-277.055
Shasta	-0.433	-0.527	**-0.100**	-0.196
John Martin	-0.252	0.772	-0.409	**0.874**
Taxoma	-0.099	**0.519**	-1.229	-0.607
Tuttle Creek	-22.504	-24.502	**0.243**	-3.369
Walter F George	-13.111	-15.750	**-1.563**	-3.719
Abiquiu	0.616	0.444	0.904	**0.915**
Allatoona	**-10.462**	-11.581	-11.538	-10.749
Brantly	-0.337	**0.784**	-1.901	-0.607
CJ Strike	-219.088	**-209.282**	-411.731	-404.773
Cochiti	-177.051	**-167.474**	-244.637	-233.246
Harlan County	-0.094	**-0.030**	-7.446	-5.364
Granbury	-82.580	-131.382	**-11.268**	-31.939
Havasu	-185.886	-192.007	-180.024	**-174.000**
Lanier	-3.686	-4.387	0.018	**0.762**
Mohave	**0.022**	-1.733	-0.571	-1.139
Milford	-0.130	-3.340	**0.197**	-1.325
Navajo	**-0.568**	-0.776	-2.310	-2.084
Possum Kingdom	**-1.557**	-1.905	-9.335	-3.792
West Point	-5.984	-6.646	**0.573**	0.034
Whitney	-2.212	-3.697	-1.706	**-0.110**
Mean	-34.334	-68.940	-84.882	-44.858
Median	-2.382	-4.042	-1.635	-1.565
Standard Deviation	65.882	191.809	264.584	104.759

Table 3.13 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to MRR results calculated from all observations.

Reservoirs	MRR			
	dQ	dQ-E	dQ+q	dQ-E+q
Red Rock	0.37	0.36	0.35	**0.35**
Barlett Horseshoe	0.02	0.42	**0**	0.4
Dworkshak	0.37	0.36	**-0.12**	-0.13
Livingston	0.51	0.42	0.33	**0.23**
Toledo Bend	0.72	0.54	0.32	**0.14**
Seminole	1.86	2.03	**1.43**	1.6
Shasta	**-0.08**	-0.1	-0.1	-0.12
John Martin	0.58	**0.42**	0.6	0.44
Taxoma	0.37	**0.26**	0.5	0.39
Tuttle Creek	**0.37**	0.42	0.49	0.55
Walter F George	1.49	1.44	0.83	**0.78**
Abiquiu	0.22	**0.18**	0.27	0.22
Allatoona	0.15	**0.08**	0.86	0.79
Brantly	**-0.06**	-0.01	-0.16	-0.1
CJ Strike	**-0.55**	-1	1.65	1.19
Cochiti	2.45	**2.33**	3.63	3.5
Harlan County	**0.33**	0.55	-0.56	-0.35
Granbury	2.63	2.41	2.11	**1.89**
Havasu	-0.81	-0.8	-0.68	**-0.66**
Lanier	0.25	0.15	0.2	**0.09**
Mohave	-0.09	**-0.08**	-0.1	-0.09
Milford	**-0.05**	-0.28	-0.12	-0.35
Navajo	0.28	**0.2**	0.81	0.72
Possum Kingdom	1.5	1.2	0.78	**0.48**
West Point	1.56	1.61	**0.86**	0.91
Whitney	0.38	**0.13**	0.79	0.54
Mean	0.57	0.51	0.57	0.52
Median	0.37	0.36	0.42	0.39
Standard Deviation	0.85	0.85	0.89	0.84

Table 3.14 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to MRR results with outliers removed from observations.

Reservoirs	MRR (Quan.)			
	dQ	dQ-E	dQ+q	dQ-E+q
Red Rock	0.37	0.37	**0.37**	0.37
Barlett Horseshoe	**0**	0.23	-0.06	0.17
Dworkshak	0.3	0.29	**-0.07**	-0.08
Livingston	0.6	0.51	0.35	**0.28**
Toledo Bend	0.65	0.49	0.44	**0.24**
Seminole	0.69	0.72	0.59	**0.56**
Shasta	**-0.07**	-0.09	-0.07	-0.09
John Martin	0.33	**0.24**	0.34	0.25
Taxoma	0.34	**0.23**	0.4	0.32
Tuttle Creek	**0.1**	0.17	0.15	0.19
Walter F George	0.41	**0.22**	0.44	0.41
Abiquiu	0.17	**0.13**	0.18	0.16
Allatoona	**-0.06**	-0.13	0.55	0.47
Brantly	-0.19	**-0.08**	-0.3	-0.2
CJ Strike	0.57	**0.23**	3.46	2.96
Cochiti	3.16	**3.01**	4.36	4.21
Harlan County	0.32	0.52	-0.54	**-0.28**
Granbury	0.37	**0.22**	0.4	0.33
Havasu	-0.62	**-0.6**	-0.67	-0.65
Lanier	0.27	0.19	0.18	**0.1**
Mohave	-0.08	**-0.07**	-0.09	-0.08
Milford	0.16	**-0.06**	0.1	-0.13
Navajo	0.36	**0.28**	0.53	0.44
Possum Kingdom	0.61	0.33	0.49	**0.15**
West Point	0.69	0.69	0.62	**0.52**
Whitney	0.3	**0.16**	0.35	0.24
Mean	0.38	0.32	0.48	0.42
Median	0.33	0.23	0.35	0.24
Standard Deviation	0.64	0.61	1.07	0.99

Table 3.15 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to SDRR results calculated from all observations.

Reservoirs	SDRR			
	dQ	dQ-E	dQ+q	dQ-E+q
Red Rock	5.08	**4.65**	5.44	4.98
Barlett Horseshoe	**3.66**	7.91	4.87	7.16
Dworkshak	4.53	4.62	1.5	**1.43**
Livingston	3.79	4.09	3.62	**3.03**
Toledo Bend	2.68	4.39	3.3	**2.25**
Seminole	15.76	17.34	**14.08**	14.37
Shasta	1.19	1.29	**1**	1.09
John Martin	4.89	**4.46**	5.15	4.7
Taxoma	**3.13**	3.37	3.7	3.58
Tuttle Creek	12.81	13.38	6.91	**6.55**
Walter F George	19.98	22.35	9	**8.08**
Abiquiu	2.12	2.65	1.9	**1.41**
Allatoona	18.69	19.24	17.79	**17.26**
Brantly	6.36	**5.04**	8.03	6.55
CJ Strike	149.97	**145.14**	204.59	199.48
Cochiti	57.02	**54.7**	76.24	73.61
Harlan County	1.63	**3.08**	8.11	3.98
Granbury	80.65	82.59	**58.13**	59.08
Havasu	50.94	**49.2**	54.63	52.87
Lanier	4.22	4.86	1.98	**1.72**
Mohave	1.09	1.07	1.13	**1.05**
Milford	5.02	7.14	**4.46**	5.9
Navajo	**1.65**	2.01	8	7.72
Possum Kingdom	18.29	19.1	9.5	**8.54**
West Point	36.63	37.7	18.43	**18.19**
Whitney	5.34	**7.04**	8.44	7.29
Mean	19.89	20.32	20.77	20.07
Median	5.05	6.04	7.45	6.55
Standard Deviation	33.25	32.27	42.1	41.17

Table 3.16 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to SDRR results with outliers removed from observations.

Reservoirs	SDRR (Quan.)			
	dQ	dQ-E	dQ+q	dQ-E+q
Red Rock	3.22	**2.92**	3.5	3.13
Barlett Horseshoe	**0.59**	1.58	0.71	1.39
Dworkshak	1.8	1.88	0.85	**0.81**
Livingston	**1.17**	1.38	1.76	1.36
Toledo Bend	0.68	1.34	1.16	**0.75**
Seminole	4.8	5.13	4.42	**4.26**
Shasta	0.6	0.64	**0.5**	0.54
John Martin	0.73	**0.54**	0.76	0.55
Taxoma	**1.06**	1.24	1.26	1.12
Tuttle Creek	1.59	1.84	1.39	**1.08**
Walter F George	3.2	3.54	2.22	**2.15**
Abiquiu	0.67	0.99	0.74	**0.53**
Allatoona	**2.74**	3.02	4.73	4.46
Brantly	0.47	**0.4**	0.76	0.55
CJ Strike	85.65	**82.88**	127.93	124.48
Cochiti	22.13	**21.18**	32.79	31.32
Harlan County	**0.41**	0.59	2.04	1.36
Granbury	22.1	23.26	**13.23**	13.25
Havasu	21.17	**20.49**	22.24	21.55
Lanier	1.53	1.75	0.65	**0.54**
Mohave	0.49	0.46	0.51	**0.45**
Milford	1.91	2.58	**1.66**	2.1
Navajo	**0.85**	1.1	4.33	4.08
Possum Kingdom	**1.68**	2.22	2.95	1.8
West Point	4.59	5.19	1.91	**1.74**
Whitney	1.33	1.9	1.42	**1.29**
Mean	7.2	7.31	9.09	8.72
Median	1.56	1.86	1.71	1.37
Standard Deviation	17.33	16.74	25.35	24.68

Table 3.17 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to rRMSE results calculated from all observations.

Reservoirs	rRMSE			
	dQ	dQ-E	dQ+q	dQ-E+q
Red Rock	5.09	**4.66**	5.45	4.99
Barlett Horseshoe	**3.66**	7.92	4.87	7.17
Dworkshak	4.54	4.64	1.51	**1.44**
Livingston	3.83	4.11	3.64	**3.04**
Toledo Bend	2.77	4.43	3.31	**2.25**
Seminole	15.87	17.45	**14.15**	14.45
Shasta	1.19	1.29	**1**	1.09
John Martin	4.92	**4.48**	5.19	4.72
Taxoma	**3.15**	3.38	3.74	3.6
Tuttle Creek	12.81	13.39	6.92	**6.58**
Walter F George	20.03	22.4	9.03	**8.12**
Abiquiu	2.13	2.66	1.92	**1.43**
Allatoona	18.69	19.24	17.81	**17.28**
Brantly	6.36	**5.04**	8.03	6.55
CJ Strike	149.97	**145.15**	204.59	199.48
Cochiti	57.07	**54.75**	76.32	73.69
Harlan County	**1.66**	3.13	8.13	3.99
Granbury	80.69	82.62	**58.17**	59.11
Havasu	50.95	**49.21**	54.63	52.88
Lanier	4.23	4.86	1.99	**1.72**
Mohave	1.1	1.08	1.14	**1.06**
Milford	**5.02**	7.15	4.47	5.91
Navajo	**1.67**	2.02	8.04	7.76
Possum Kingdom	18.35	19.14	9.53	**8.55**
West Point	36.66	37.73	18.45	**18.21**
Whitney	**5.36**	7.04	8.48	7.31
Mean	19.91	20.34	20.79	20.09
Median	5.06	6.04	7.48	6.56
Standard Deviation	33.25	32.28	42.1	41.17

Table 3.18 Observed conditions of river-store mass conservation added with secondary factors both individually and collectively according to rRMSE results with outliers removed from observations.

Reservoirs	rRMSE (Quan.)			
	dQ	dQ-E	dQ+q	dQ-E+q
Red Rock	3.25	**2.94**	3.52	3.16
Barlett Horseshoe	**0.59**	1.6	0.72	1.4
Dworkshak	1.83	1.91	0.86	**0.81**
Livingston	**1.32**	1.47	1.8	1.39
Toledo Bend	0.94	1.42	1.24	**0.79**
Seminole	4.85	5.18	4.46	**4.3**
Shasta	0.6	0.65	**0.51**	0.55
John Martin	0.8	**0.6**	0.83	0.6
Taxoma	**1.11**	1.26	1.32	1.16
Tuttle Creek	1.59	1.85	1.4	**1.1**
Walter F George	3.23	3.55	2.26	**2.19**
Abiquiu	0.69	1	0.76	**0.55**
Allatoona	**2.74**	3.03	4.76	4.49
Brantly	0.5	**0.41**	0.82	0.59
CJ Strike	85.66	**82.88**	127.97	124.52
Cochiti	22.35	**21.39**	33.08	31.61
Harlan County	**0.52**	0.79	2.11	1.39
Granbury	22.11	23.26	**13.23**	13.25
Havasu	21.18	**20.5**	22.25	21.56
Lanier	1.56	1.76	0.68	**0.55**
Mohave	0.5	0.47	0.52	**0.46**
Milford	1.92	2.58	**1.66**	2.11
Navajo	**0.92**	1.13	4.36	4.11
Possum Kingdom	**1.79**	2.25	2.99	1.81
West Point	4.64	5.24	2.01	**1.82**
Whitney	1.36	1.91	1.46	**1.31**
Mean	7.25	7.35	9.14	8.75
Median	1.57	1.88	1.73	1.39
Standard Deviation	17.33	16.74	25.37	24.7

3.3.1 Individual Patterns

Lake Red Rock is located in the semiarid region. Before adding any datasets of secondary controls, the NSE value of this lake is 0.845, indicating a "Very Good" error performance. Our results also suggest an inclusion of both E and q according to NSE, while an inclusion of just E according to KGE and rRMSE.

Barlett and Horseshoe are two cascade reservoirs located in the semiarid region. Results according to NSE, KGE and rRMSE suggest no significant improvement is seen by adding any combination of datasets of secondary controls.

Dworshak reservoir is located in the semiarid region. Both rRMSE and NSE results suggest an inclusion of both E and q is necessary.

Lake Livingston is located in the humid region. rRMSE, NSE and KGE results all suggest an inclusion of both E and q is necessary.

Toledo Bend Lake is located in the humid regions. rRMSE, NSE and KGE results all suggest an inclusion of both E and q is necessary with a large improvement on the premise condition.

Seminole Lake is located in the subhumid region. NSE results do not suggest a use of datasets for secondary controls is necessary, while rRMSE results suggest the inclusion of q and KGE suggest the inclusion of both E and q.

Shasta Lake is located in the humid region. A "Very Good" NSE error performance is seen in Shasta Lake before any additional dataset is used. rRMSE, NSE and KGE results all suggest an inclusion of q is necessary.

John Martin is located in the arid region. RMSE, NSE and KGE results all suggest an inclusion of E is necessary with a small improvement on the premise condition.

Taxoma Lake is located in the humid region. Starting with a “Very Good” NSE error performance, both rRMSE and KGE results suggest a small improvement on premise condition by including just E.

Tuttle Creek Reservoir is located in humid region. NSE and KGE results suggest a use of q is necessary, while rRMSE results suggest the inclusion of q and E.

Walter F George reservoir is located in humid region. NSE and KGE results suggest a use of q is necessary, while rRMSE results suggest the inclusion of q and E.

Abiquiu Reservoir is located in semiarid region. NSE and rRMSE results suggest a use of E and q is necessary, while KGE results suggest the inclusion of q only.

Allatoona Lake is located in humid region. Both KGE and NSE results do not suggest a use of additional dataset, while rRMSE results suggest an inclusion of E and q.

Brantly Lake is located in arid region. All rRMSE, NSE and KGE results suggest that an inclusion of E only would improve the premise condition.

CJ Strike Reservoir is located in arid region. All rRMSE, NSE and KGE results suggest that an inclusion of E only would improve the premise condition.

Cochoti Reservoir is located in semiarid region. All rRMSE, NSE and KGE results suggest that an inclusion of E only would improve the premise condition.

Harlan County Reservoir is located in subhumid region. NSE results suggest an inclusion of E only while KGE and rRMSE do not suggest a use of any additional dataset.

Lake Granbury is located in humid region. All rRMSE, NSE and KGE results suggest that an inclusion of q only would improve the premise condition.

Lake Havasu is located in semiarid region. All rRMSE, NSE and KGE results suggest that an inclusion of E only would improve the premise condition.

Lake Lanier is located in humid region. All rRMSE, NSE and KGE results suggest that an inclusion of E and q would largely improve the premise condition.

Lake Mohave is located in the arid region. Both KGE and NSE results do not suggest a use of additional dataset, while rRMSE results suggest an inclusion of E and q.

Milford Lake is located in semiarid region. All rRMSE, NSE and KGE results suggest that an inclusion of q only would improve the premise condition.

Navajo Reservoir is located in semiarid region. Neither rRMSE, KGE and NSE results suggest a use of additional dataset.

Possum Kingdom Lake is located in semiarid region. KGE results do not suggest a use of additional dataset, while rRMSE and NSE results suggest an inclusion of E and q.

West Point Lake is located in subhumid region. KGE results suggest a use of q is necessary, while rRMSE and NSE results all suggest an inclusion of q and E.

Whitney Lake is located in semiarid region. rRMSE results do not suggest a use of additional dataset is necessary, while KGE and NSE results all suggest an inclusion of q and E.

3.3.2 General Patterns

Table 3.9 and 3.10 demonstrate the effectiveness of including estimated reservoir surface evaporation (E; CRED dataset) and simulated lateral inflow (q; GRADES dataset) and summarize the error performance results when using (1) the original discharge difference (dQ), (2) discharge difference subtracted by estimated reservoir surface evaporation (dQ - E), (3) discharge difference added by simulated lateral inflow and (4) discharge difference subtracted by estimated reservoir surface evaporation and added by simulated lateral inflow (dQ -E + q), according to NSE in two versions of 26 reservoirs in this study. The highest value of NSE, indicating the best error performance, that each reservoir receives, in four scenarios is bolded on the tables. Same tables are also generated for KGE (Table 3.11, 3.12), rRMSE (Table 3.17, 3.18), MRR (Table 3.13, 3.14) and SDRR (Table 3.15, 3.16). NSE, KGE and rRMSE are error performance metrics that indicate an overall error performance. A decomposition of rRMSE yields MRR and SDRR, indicating a bias-related error and time-varying errors, respectively.

Our results show that (1) according to NSE, five reservoirs (19.23%) have the highest NSE values using dQ, seven reservoirs (26.92%) using dQ – E, five reservoirs (19.23%) using dQ + q, nine reservoirs (34.62%) using dQ – E + q; (2) according to KGE, six reservoirs (23.08%) have the highest KGE values using dQ, seven reservoirs (26.92%) using dQ – E, seven reservoirs (26.92%) using dQ + q, five reservoirs (19.23%) using dQ – E + q; (3) according to rRMSE, six reservoirs (23.08%) have the highest rRMSE values using dQ, six reservoirs (23.08%) using dQ – E, three reservoirs (11.54%) using dQ + q, eleven reservoirs (42.31%) using dQ – E + q; To summarize, our results indicate that by including datasets that simulate or estimate secondary controls, 80.77% of reservoirs in our sample have shown an improved premise condition according to NSE (76.92% according to KGE and rRMSE). Including both datasets of lateral inflow and surface evaporation (dQ – E + q) has improved the river-store mass conservation for the largest number of reservoirs according to NSE (nine reservoirs; 34.62%) and rRMSE (eleven reservoirs; 42.31%), while according to KGE, including just surface evaporation (dQ – E) or just lateral inflow (dQ + q) ties the number of reservoirs with an improved river-store mass conservation (seven reservoirs; 26.92%).

These findings suggest that: first, the inclusion of datasets of secondary controls has a great potential in improving the river-store mass conservation for a better LakeFlow premise condition in most reservoirs; second, as we do not see in our results that one particular scenario of adding secondary controls being the most effective or more effective at large, it is necessary that we analyze our results by different hydroclimate conditions, and reservoir purposes to find out if there is any pattern there. Followed by this motivation, we will use NSE statistical results for the reservoirs in this study categorized by different hydroclimate regions from the HLR dataset and reservoir purposes from the GRanD dataset.

Table 3.19 summarizes the improvement of the premise conditions by integrating E and q individually and collectively in different hydroclimate regions. Among the 26 sampled reservoirs, 4 are located in arid regions, 10 in semiarid regions, 3 in subhumid regions, and 9 in humid regions. To allow a fair comparison of reservoirs in different hydroclimate regions, the humid and subhumid regions are grouped together as a single wet region, so as arid and semiarid regions as a single dry region, so a similar number of reservoirs can be achieved for the two regions (Dry: 14 reservoirs; Wet: 12 reservoirs). In the dry region, results show that the largest average (0.9%), maximum (443.9%) and median (0.1%) improvement is achieved by including both factors together (Q - E + q), although the standard deviation (161.2%) reaches the largest among all scenarios, indicating a large variability. Including q in reservoirs located in dry regions adds enormous uncertainty to the river-store mass conservation, as seen with a standard deviation of 123.5% when only q is included, and 161.2% when both q and E are included, while a standard deviation of 9.4% is seen when just E is included. In the wet region, the largest average (498.6%) and the maximum (3868.9%) improvement is seen after including both datasets together (Q - E + q), though the standard deviation of the improvement in this scenario is 1118.9%, the largest and the median is 27.4%, the second largest. It has also been seen that including q in reservoirs located in wet regions adds more uncertainty to the river-store mass conservation, as seen with a standard deviation of 1097.6% when only q is included, and 1118.9% when both q and E are included, while a standard deviation of 163.5% is seen when just E is included.

It would not be accurate to summarize our findings solely using the average improvement percentages due to the degree of variability in our sample indicated by having large standard deviations. It is perhaps the best to use the median when the distribution of data values is largely skewed and when there are clear “outliers”. In this case, the median is a better indication of where the center of our sample is located. In dry regions, the highest median improvement percentage is seen associated with an inclusion of E and q, while in wet regions, the highest median (28.0%) is seen with an inclusion of just q, slightly higher than the median (27.4%) with an inclusion of E and q though. Therefore, we suggest an inclusion of E and q might be most effective in improving the river-store mass conservation for a better LakeFlow premise condition in both dry regions and wet regions, and accounting for the contribution of lateral flows is necessary for reservoirs in wet regions.

Table 3.20 summarizes the possible improvement on the premise conditions of water-supply reservoirs and non-water-supply reservoirs. According to GRanD, 18 reservoirs were built for water-supply purposes while the rest were not. Our results show that the highest average, standard deviation, maximum and median improvement percentages are all associated with an inclusion of both E and q, in both water-supply reservoirs and non-water-supply reservoirs. In addition, the mean, standard deviation, maximum and median improvement percentages are a lot higher in non-water-supply reservoirs than in water-supply reservoirs. This finding suggests that an inclusion of E and q together is more effective in non-water-supply reservoirs than water-supply reservoirs.

Table 3.19 Summary of improvement from E and q on different hydroclimate conditions according to NSE calculated from the full length of records ("All").

Region	Count	Control	Mean	Standard deviation	Minimum	Maximum	Median
Dry	14	E	-2.4%	**9.4%**	**-30.9%**	6.7%	-0.3%
		q	-22.4%	123.5%	-410.8%	177.8%	-0.3%
		E & q	**0.9%**	161.2%	-386.4%	**443.9%**	**0.1%**
Wet	12	E	-69.0%	**163.5%**	-591.0%	3.7%	-9.7%
		q	481.3%	1097.6%	-73.6%	3805.0%	**28.0%**
		E & q	**498.6%**	1118.9%	**-52.8%**	**3868.9%**	27.4%

Table 3.20 Summary of improvement from E and q on reservoirs with different reservoir purposes according to NSE calculated from the full length of records ("All").

Reservoir Purpose	Count	Control	Mean	Standard deviation	Minimum	Maximum	Median
Supply	18	E	-12.0%	**38.3%**	**-165.6%**	6.7%	-0.9%
		q	69.5%	396.9%	-410.8%	1640.8%	-0.6%
		E & q	**94.8%**	420.5%	-386.4%	**1725.0%**	**0.9%**
Non-Supply	8	E	-80.7%	**193.0%**	-591.0%	0.4%	-9.7%
		q	526.5%	1240.8%	-12.4%	3805.0%	68.7%
		E & q	**536.3%**	1261.2%	**-7.5%**	**3868.9%**	**74.5%**

3.4 Impact of SWOT's Temporal Sampling

3.4.1 Number of SWOT overpasses

Table 3.21 and Figure 3-10 below summarizes the number of SWOT overpasses and the corresponding revisit time for the 26 reservoirs sampled in this study. The average number of SWOT overpasses for these reservoirs are 2 with an averaged revisit time of 11 days.

3.4.2 Error Results

We used a set of five error metrics and a Wilcoxon rank sum test, to evaluate the error performance in this question. Detailed descriptions of these error metrics and how they were calculated can be found in Section 2.2.3 and 2.5.

A Wilcoxon rank sum test (or Mann-Whitney U test) was added for error performance evaluation (Cuzick, 1985). The rank sum test aims to determine whether two groups of data come from the same population and requires no assumption for data distributions. Figure 3.11, Tables 3.22 and 3.23 summarize the results of 26 sampled reservoirs after simulating a SWOT sampling scheme on the observed river flow records according to 6 error metrics. For the rank sum test, 24 reservoirs have a *p*-value over 0.05, which means that in these reservoirs, there is not enough evidence to reject the null hypothesis that the observed discharge difference and the estimated discharge difference are from the same population group in our study. The mean of the p values from a rank sum test is 0.56. Such results indicate that there is strong evidence that estimated and observed discharge difference are from the same population.

11 reservoirs have an "All" NSE greater than 0.5, and 12 reservoirs have a "quantile" NSE greater than 0.5. The mean, standard deviation and median of an "All" NSE among all 26 reservoirs are -0.11, 0.35 and 1.38, and of a "quantile" NSE are -0.01, 1.26 and 0.42. Such

results indicate that nearly half of the reservoirs have shown satisfactory results when a SWOT sampling gap is applied. The KGE has suggested better performance in applying a SWOT sampling gap. 17 reservoirs have an “All” KGE greater than a recommended benchmark of -0.41 and 16 reservoirs have a “quantile” KGE greater than -0.41. The mean, standard deviation and median of an “All” KGE among all 26 reservoirs are 0.54, 0.40 and 0.65, and of a “quantile” KGE are -0.34, 0.61 and 1.01. Such results in KGE have shown that a widespread strong relationship between the estimated and observed discharge difference time series.

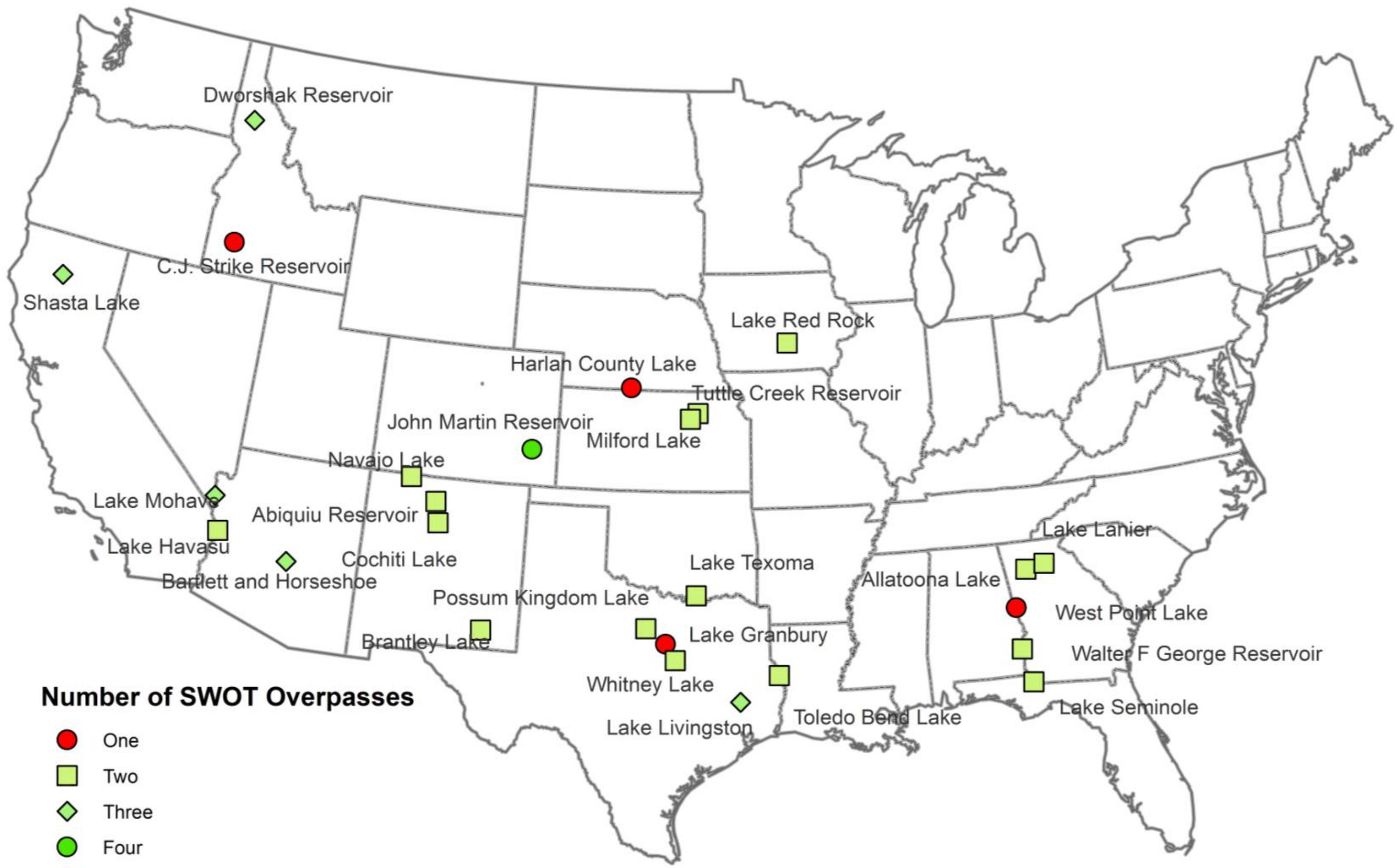


Figure 3-10 Summary of the numbers of SWOT overpasses in one orbital period for reservoirs in this study. Four reservoirs have one overpass, sixteen reservoirs have two overpasses, five reservoirs have three overpasses and one reservoir have four overpasses.

Table 3.21 The number of overpasses and revisit days of the studied reservoirs under the SWOT's design.

Reservoirs	Number of overpasses	Revisit days
John Martin	4	5
Barlett Horseshoe	3	7
Dworkshak	3	7
Livingston	3	7
Shasta	3	7
Mohave	3	7
Red Rock	2	11
Toledo Bend	2	11
Seminole	2	11
Taxoma	2	11
Tuttle Creek	2	11
Walter F George	2	11
Abiquiu	2	11
Allatoona	2	11
Brantly	2	11
Cochiti	2	11
Havasu	2	11
Lanier	2	11
Milford	2	11
Navajo	2	11
Possum Kingdom	2	11
Whitney	2	11
CJ Strike	1	21
Harlan County	1	21
Granbury	1	21
West Point	1	21
MEAN	2	11

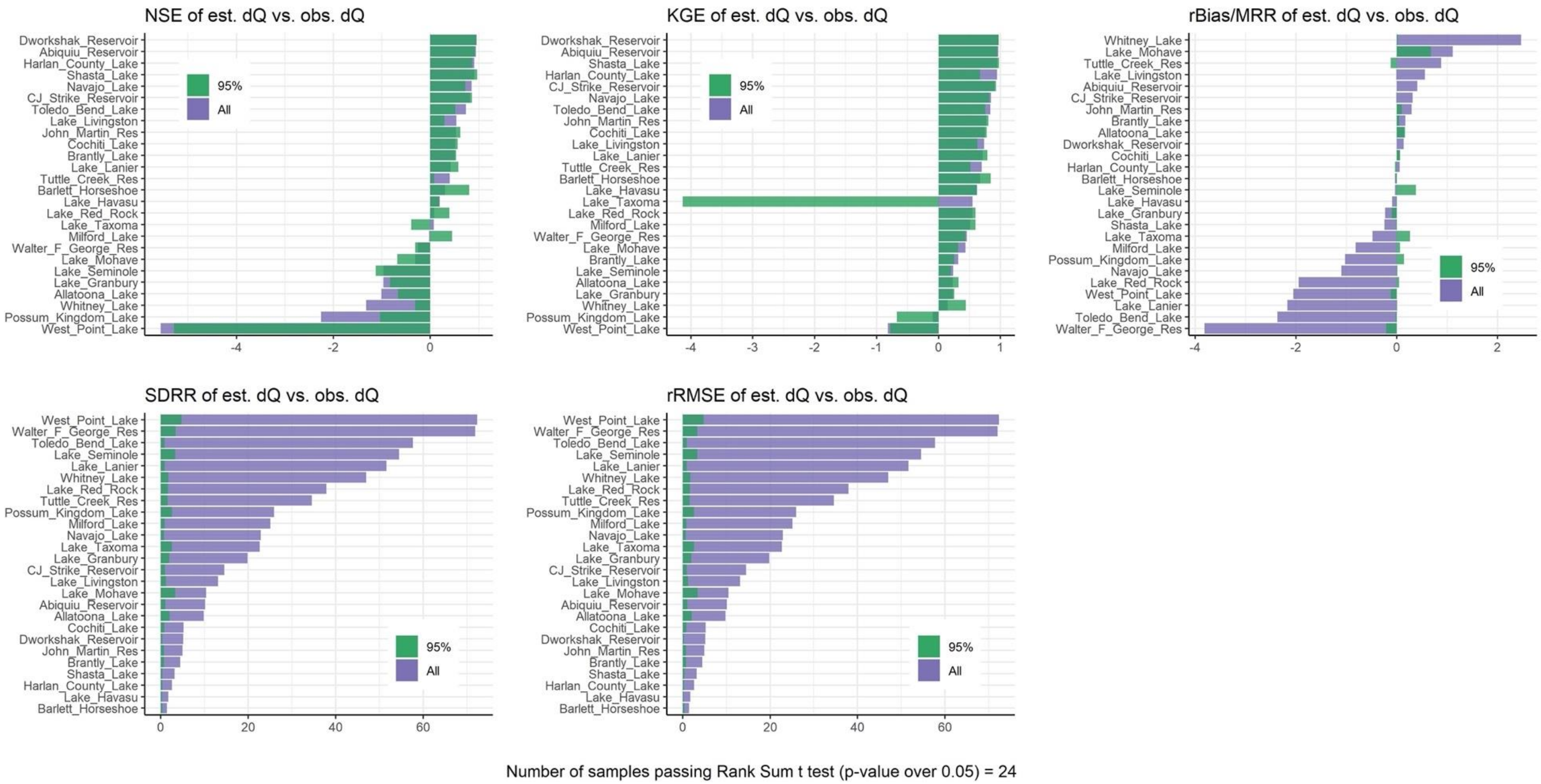


Figure 3-11 Comparison between the estimated and observed discharge difference with ("95%") and without removing the outliers ("ALL")

Table 3.22 Comparison between the estimated and observed discharge difference according to rank sum test, NSE and KGE ("95%") with and without removing the outliers ("ALL")

Reservoirs	Rank Sum	NSE		KGE	
		All	Quan.	All	Quan.
Red Rock	0.81	0.08	0.39	0.55	0.6
Barlett Horseshoe	0.05	0.3	0.8	0.67	0.85
Dworkshak	0.25	0.95	0.94	0.97	0.97
Livingston	0.73	0.54	0.29	0.73	0.63
Toledo Bend	0.44	0.73	0.51	0.84	0.76
Seminole	0.96	-0.96	-1.12	0.24	0.2
Shasta	0.92	0.91	0.97	0.95	0.98
John Martin	0.62	0.53	0.62	0.77	0.8
Taxoma	0.87	0.07	-0.39	0.55	-4.13
Tuttle Creek	0.32	0.4	0.08	0.7	0.52
Walter F George	0.79	-0.26	-0.31	0.46	0.44
Abiquiu	0.8	0.94	0.92	0.96	0.95
Allatoona	0.71	-1.01	-0.67	0.23	0.33
Brantly	0.13	0.51	0.53	0.32	0.25
CJ Strike	0.78	0.82	0.86	0.91	0.93
Cochiti	0.93	0.52	0.56	0.75	0.78
Harlan County	0.22	0.91	0.86	0.95	0.67
Granbury	0.26	-0.96	-0.83	0.23	0.26
Havasu	0.88	0.19	0.18	0.63	0.62
Lanier	0.79	0.42	0.58	0.72	0.79
Mohave	0.59	-0.31	-0.68	0.44	0.32
Milford	0.33	-0.02	0.45	0.51	0.6
Navajo	0.31	0.85	0.72	0.85	0.82
Possum Kingdom	0	-2.25	-1.04	-0.09	-0.67
West Point	0.23	-5.56	-5.29	-0.82	-0.78
Whitney	0.78	-1.32	-0.31	0.15	0.44
Mean	0.56	-0.11	-0.01	0.54	0.34
Median	0.66	0.35	0.42	0.65	0.61
Standard Deviation	0.31	1.38	1.26	0.4	1.01

Table 3.23 Comparison between the estimated and observed discharge difference according to rRMSE, MRR and SDRR ("95%") with and without removing the outliers ("ALL")

Reservoirs	rRMSE		MRR		SDRR	
	All	Quan.	All	Quan.	All	Quan.
Red Rock	37.94	1.65	-1.95	0.05	37.89	1.65
Barlett Horseshoe	1.42	0.39	-0.03	-0.02	1.42	0.39
Dworkshak	5.18	0.35	0.14	0	5.18	0.35
Livingston	13.15	1.21	0.56	-0.01	13.14	1.21
Toledo Bend	57.69	0.92	-2.37	-0.02	57.64	0.91
Seminole	54.51	3.34	-0.03	0.38	54.51	3.32
Shasta	3.18	0.39	-0.24	-0.01	3.18	0.39
John Martin	4.98	0.75	0.3	0.1	4.97	0.75
Taxoma	22.68	2.61	-0.48	0.26	22.67	2.6
Tuttle Creek	34.58	1.58	0.88	-0.12	34.57	1.58
Walter F George	72.01	3.38	-3.81	-0.21	71.91	3.37
Abiquiu	10.13	1.05	0.41	0	10.12	1.05
Allatoona	9.84	2.03	0.14	0.17	9.83	2.02
Brantly	4.49	0.76	0.17	0.05	4.49	0.76
CJ Strike	14.53	0.96	0.31	-0.01	14.53	0.96
Cochiti	5.22	0.83	0.06	0.07	5.22	0.82
Harlan County	2.58	0.34	0.06	-0.03	2.58	0.33
Granbury	19.83	1.98	-0.23	-0.1	19.83	1.98
Havasu	1.77	0.28	-0.09	-0.01	1.77	0.28
Lanier	51.65	0.94	-2.17	0.01	51.6	0.94
Mohave	10.47	3.4	1.11	0.68	10.41	3.34
Milford	25.11	0.88	-0.82	0.07	25.1	0.87
Navajo	22.9	0.73	-1.1	0.01	22.87	0.73
Possum Kingdom	25.96	2.58	-1.03	0.15	25.94	2.57
West Point	72.37	4.77	-2.05	-0.12	72.34	4.77
Whitney	47.03	1.75	2.47	0.02	46.97	1.75
Mean	24.28	1.53	-0.38	0.05	24.26	1.53
Median	17.18	1.01	-0.03	0	17.18	1.01
Standard Deviation	22.33	1.18	1.29	0.17	22.31	1.17

Chapter 4 - Conclusions

4.1 LakeFlow Premise Testing: Can reservoir storage change and discharge difference characterize the river-store mass conservation?

The algorithm's premise ($\delta Q \approx \delta V$) was tested in 26 reservoirs across large geographical areas in the contiguous United States. We used the Nash-Sutcliffe Efficiency (NSE) as the central error performance metric, and KGE, rRMSE and an overlapping index are supplementary to the use of NSE for assessing the river-store mass conservation characterized by reservoir storage change and discharge difference. Our results show that 18 (69.2%) out of 26 reservoirs have an NSE value greater than zero, exceeding the mean flow benchmark for NSE. By using the error performance criteria for NSE suggested by Moriasi et al. (2015), 12 reservoirs (46.15%) show an NSE value greater than 0.5 (categorized as 'Satisfactory' or above).

Our results also indicate differences in how some error performance metrics would behave given the same datasets used in the calculations. We suggest future studies could focus on the behaviors of the error metrics by conducting a decomposition of error performance metrics based on different aspects of error components, so we see what aspect of errors is causing a particular error performance metric (e.g., NSE or KGE) to become larger or smaller. The statistical results according to rRMSE, MRR and SDRR show overall very large values of the three error performance metrics.

By comparing the results before and after applying a 5% outlier-removal method, we find the following messages. First, removing the outliers in our discharge difference and reservoir storage change datasets did reduce the time-varying errors (observed from SDRR results). Second, in cases where we see an increase in bias errors (MRR), a decrease in time-varying errors (SDRR) could compensate the impact of bias errors (MRR), resulting in an overall lower rRMSE values. Third, the outlier-removal technique could increase or lower the bias errors, and this impact on bias is more complicated than on time-varying errors where we see 26 reservoirs all showing a lower SDRR value. Finally, considering the complicated nature of the impact of removing outliners in datasets on time-varying errors and bias errors, we urge that the "quantile" version for relative statistics results (rRMSE, SDRR, MRR) needs to be used with discretion.

4.2 Secondary Controls that Affect LakeFlow Premise Condition

We summarize the statistical results according to NSE, KGE and rRMSE for reservoirs that are water-supply and non-water-supply and calculated the descriptive statistics for the error performance results for each group of reservoirs. We strongly argue that human water withdrawal largely affects the ability of which gauge-measured reservoir storage change and discharge difference characterize the river-store mass conservation, and human water withdrawal on reservoirs is an important cause of a less satisfactory premise condition for LakeFlow, by largely increasing more time-varying errors.

Our descriptive statistical summary of error performance among NSE, KGE and rRMSE have indicated a discrepancy in four hydroclimate regions. Therefore, we used alternative ways (i.e., hydrographs and scatterplots) to descriptive statistics in order to analyze the possible impact of hydroclimate conditions on secondary controls.

According to our interpretations of each reservoir in this study, we have found that, at large, discharge difference tends to underestimate reservoir storage change (meaning reservoir gaining water from other sources) in more humid regions, implying a substantial impact of lateral inflow from the tributaries that are too small for SWOT to observe. In arid and semiarid regions, discharge difference overestimates reservoir storage change only during certain months (notably, late spring), suggesting a non-negligible impact of reservoir surface evaporation, though an underestimation of discharge difference on reservoir storage change still occurs widely. These findings based on visual interpretations of plots are consistent with our interpretations of MRR results, where a negative value indicates an overestimation and a positive value indicate an underestimation of discharge difference on storage change.

Therefore, we argue that lateral inflow, being the only water influx among the three major secondary controls, is the main reason that such underestimation commonly occurs in our sample. Further, the scales of this influx from lateral tributaries might be even larger than the water outfluxes from other secondary controls, namely surface evaporation and human water withdrawal. This finding exceeds our previous expectation on the impact of lateral inflow being only significant in humid regions, and implicates that one possible way to improve the observed river-store mass conservation for a better premise condition for LakeFlow, is to account for the contribution of the lateral inflow in the studied reservoirs, no matter in which hydroclimate region they are located.

Our results suggest that the overestimation of discharge difference on reservoir storage change is a lot less common in our sample at large (6 out of 26 reservoirs) and is more likely to occur in dry regions. We might be able to improve the observed river-store mass conservation for a better premise condition for LakeFlow, by accounting for reservoir surface evaporation in drier regions, though the improvement might not be effective and could even incur more errors.

Finally, we provide a summary of the premise condition for each reservoir in this study.

4.3 Feasibility of Ancillary Data on Improving the Observed River-Store Mass Conservation for a Better LakeFlow Premise Condition

We evaluated error performance by adding different combinations of estimated reservoir surface evaporation and simulated lateral inflows according to NSE, KGE and rRMSE. Additionally, we provided suggestions on how the datasets of secondary controls can be used for each reservoir in this study.

Our results have suggested the inclusion of datasets of secondary controls has a great potential in improving the river-store mass conservation for a better LakeFlow premise condition in most reservoirs.

NSE statistical results were summarized for reservoirs by different hydroclimate regions from the HLR dataset and reservoir purposes from the GRanD dataset. Descriptive statistics were generated to describe the data characteristics in each hydroclimate condition or reservoir purpose. By comparing the median improvement in each hydroclimate region, we suggest an inclusion of E and q might be most effective in improving the river-store mass conservation for a better LakeFlow premise condition in both dry regions and wet regions, and accounting for the contribution of lateral flows is necessary for reservoirs in wet regions. Additionally, our results also suggest that an inclusion of E and q together is more effective in non-water-supply reservoirs than in water-supply reservoirs.

4.4 SWOT's Temporal Sampling

SWOT temporal sampling gaps have a relatively limited influence on the discharge difference.

4.5 Limitations

This study will help to assess the validity of using river-store mass conservation to improve the parameterization of channel properties and thus improve the accuracy of discharge estimation, and provide guidance to the applications of LakeFlow.

Some limitations exist in this study. First, when comparing our results by reservoirs in different hydroclimate regions, the numbers of the reservoirs in each hydroclimate region are not closely equal. The smallest number is three from subhumid regions and the largest number is ten from the semiarid regions. This bias in sample size may lead to a certain degree of uncertainties in our findings.

Second, as previously suggested, a decomposition of error performance metrics like NSE and KGE is necessary for helping to better understand what aspect of errors is causing a particular error performance to be good or poor. This study was not able to provide enough insights on this matter.

Third, the HLR dataset allows us to summarize our reservoirs by different hydroclimate conditions, but there is one important assumption: reservoirs that are located in the same hydroclimate region defined in HLR dataset would be impacted by similar climate patterns (specifically related to precipitation). However, some reservoirs are located right at the interface of different hydroclimate regions. This would have a direct impact on some of the conclusions that were based upon a more absolute classification of hydroclimate conditions.

Fourth, the in situ gauges where our streamflow data were collected are not always in the vicinity of the reservoir, as they are not specifically established for measuring streamflow around the reservoir. This is often the case for the gauges that were set up upstream of a reservoir. The farther away an inflow gauge is from the main reservoir inlet, the more likely the hydrological behaviors (including the scale and the timing) of river inflow at this cross-section are different from those of an ideal reservoir inlet cross-section or the river section which SWOT will be able to observe. A major consequence is that we might have misjudged the LakeFlow premise condition for some reservoirs, potentially resulting in an underestimation or an overestimation.

Fifth, the thesis mostly focused land surface hydrological processes. The impacts of infiltration and groundwater recharge in the reservoirs are not directly quantified although they may be indirectly inferred by the unexplained residuals.

References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A. and Hegewisch, K.C. 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Scientific Data 5(1), 170191.

Allen, G.H. and Pavelsky, T.M. 2018. Global extent of rivers and streams. Science 361(6402), 585-588.

Alsdorf, D.E., Rodríguez, E. and Lettenmaier, D.P. 2007. Measuring surface water from space. Reviews of Geophysics 45(2).

Altenau, E.H., Pavelsky, T.M., Durand, M.T., Yang, X., Frasson, R.P.d.M. and Bendezu, L. 2021. The Surface Water and Ocean Topography (SWOT) Mission River Database (SWORD): A Global River Network for Satellite Data Products. Water Resources Research 57(7), e2021WR030054.

Anderson, R.S. and Anderson, S.P. (2010) Geomorphology: the mechanics and chemistry of landscapes, Cambridge University Press.

Arnold, J.G., Srinivasan, R., Muttiah, R.S. and Williams, J.R. 1998. Large Area Hydrologic Modeling and Assessment Part I: Model Development. Journal of the American Water Resources Association 34(1), 73-89.

Beven, K.J. and Kirkby, M.J. 1979. A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrological Sciences Bulletin 24(1), 43-69.

Biancamaria, S., Durand, M., Andreadis, K.M., Bates, P.D., Boone, A., Mognard, N.M., Rodríguez, E., Alsdorf, D.E., Lettenmaier, D.P. and Clark, E.A. 2011. Assimilation of virtual wide swath altimetry to improve Arctic river modeling. Remote Sensing of Environment 115(2), 373-381.

Biancamaria, S., Lettenmaier, D.P. and Pavelsky, T.M. 2015. The SWOT Mission and Its Capabilities for Land Hydrology. Surveys in Geophysics 37(2), 307-337.

Biondi, D. and Todini, E. 2018. Comparing Hydrological Postprocessors Including Ensemble Predictions Into Full Predictive Probability Distribution of Streamflow. Water Resources Research 54(12), 9860-9882.

Bjerklie, D.M., Dingman, S.L. and Bolster, C.H. 2005a. Comparison of constitutive flow resistance equations based on the Manning and Chezy equations applied to natural rivers. Water Resources Research 41(11).

Bjerklie, D.M., Moller, D., Smith, L.C. and Dingman, S.L. 2005b. Estimating discharge in rivers using remotely sensed hydraulic information. Journal of Hydrology 309(1-4), 191-209.

Bonnema, M., Sikder, S., Miao, Y., Chen, X., Hossain, F., Ara Pervin, I., Mahbubur Rahman, S.M. and Lee, H. 2016a. Understanding satellite-based monthly-to-seasonal reservoir outflow estimation as a function of hydrologic controls. Water Resources Research 52(5), 4095-4115.

Bonnema, M.G., Sikder, S., Hossain, F., Durand, M., Gleason, C.J. and Bjerklie, D.M. 2016b. Benchmarking wide swath altimetry-based river discharge estimation algorithms for the Ganges river system. Water Resources Research 52(4), 2439-2461.

Brakenridge, G.R., Nghiem, S.V., Anderson, E. and Mic, R. 2007. Orbital microwave measurement of river discharge and ice status. Water Resources Research 43(4).

Brinkerhoff, C.B., Gleason, C.J., Feng, D. and Lin, P. 2020. Constraining Remote River Discharge Estimation Using Reach‐Scale Geomorphology. Water Resources Research 56(11).

Burnash, R.J., Ferral, R.L. and McGuire, R.A. (1973) A generalized streamflow simulation system: Conceptual modeling for digital computers, US Department of Commerce, National Weather Service, and State of California.

Chai, Y., Li, Y., Yang, Y., Zhu, B., Li, S., Xu, C. and Liu, C. 2019. Influence of Climate Variability and Reservoir Operation on Streamflow in the Yangtze River. Scientific Reports 9(1), 5060.

Chao, B.F., Wu, Y.H. and Li, Y.S. 2008. Impact of artificial reservoir water impoundment on global sea level. Science 320(5873), 212-214.
Chin, D.A. (2012) Water-Resources Engineering, Pearson Education.
Cook, B.I., Smerdon, J.E., Seager, R. and Coats, S. 2014. Global warming and 21st century drying. Climate Dynamics 43(9-10), 2607-2627.
Costa, J.E., Cheng, R.T., Haeni, F.P., Melcher, N., Spicer, K.R., Hayes, E., Plant, W., Hayes, K., Teague, C. and Barrick, D. 2006. Use of radars to monitor stream discharge by noncontact methods. Water Resources Research 42(7).
Costa, J.E., Spicer, K.R., Cheng, R.T., Haeni, F.P., Melcher, N.B., Thurman, E.M., Plant, W.J. and Keller, W.C. 2000. measuring stream discharge by non-contact methods: A Proof-of-Concept Experiment. Geophysical Research Letters 27(4), 553-556.
Cuzick, J. 1985. A Wilcoxon-type test for trend. Statistics in Medicine 4(1), 87-90.
Dai, A. 2012. Increasing drought under global warming in observations and models. Nature Climate Change 3(1), 52-58.
Dalton, M.S., Aulenbach, B.T. and Torak, L.J. 2004 Ground-water and surface-water flow and estimated water budget for Lake Seminole, southwestern Georgia and northwestern Florida.
Durand, M., Gleason, C.J., Garambois, P.A., Bjerklie, D., Smith, L.C., Roux, H., Rodriguez, E., Bates, P.D., Pavelsky, T.M., Monnier, J., Chen, X., Di Baldassarre, G., Fiset, J.M., Flipo, N., Frasson, R.P.d.M., Fulton, J., Goutal, N., Hossain, F., Humphries, E., Minear, J.T., Mukolwe, M.M., Neal, J.C., Ricci, S., Sanders, B.F., Schumann, G., Schubert, J.E. and Vilmin, L. 2016. An intercomparison of remote sensing river discharge estimation algorithms from measurements of river height, width, and slope. Water Resources Research 52(6), 4527-4549.
Durand, M., Neal, J., Rodríguez, E., Andreadis, K.M., Smith, L.C. and Yoon, Y. 2014. Estimating reach-averaged discharge for the River Severn from measurements of river water surface elevation and slope. Journal of Hydrology 511, 92-104.
Eijkhout, V., Habets, F., Yang, Z.-L., Niu, G.-Y., Maidment, D.R. and David, C.H. 2011. River Network Routing on the NHDPlus Dataset. Journal of Hydrometeorology 12(5), 913-934.
Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D. and Alsdorf, D. 2007. The Shuttle Radar Topography Mission. Reviews of Geophysics 45(2).
Gleason, C. and Durand, M. 2020. Remote Sensing of River Discharge: A Review and a Framing for the Discipline. Remote Sensing 12(7), 1107.
Gleason, C.J. and Smith, L.C. 2014. Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry. Proceedings of the National Academy of Sciences 111(13), 4788-4791.
Gleason, C.J., Smith, L.C. and Lee, J. 2014. Retrieval of river discharge solely from satellite imagery and at-many-stations hydraulic geometry: Sensitivity to river form and optimization parameters. Water Resources Research 50(12), 9604-9619.
Gleason, C.J., Wada, Y. and Wang, J. 2018. A Hybrid of Optical Remote Sensing and Hydrological Modeling Improves Water Balance Estimation. Journal of Advances in Modeling Earth Systems 10(1), 2-17.

Gleason, C.J. and Wang, J. 2015. Theoretical basis for at-many-stations hydraulic geometry. Geophysical Research Letters 42(17), 7107-7114.
Graf, W.L. 1999. Dam nation: A geographic census of American dams and their large-scale hydrologic impacts. Water Resources Research 35(4), 1305-1311.
Graf, W.L. 2006. Downstream hydrologic and geomorphic effects of large dams on American rivers. Geomorphology 79(3-4), 336-360.
Gragne, A.S., Sharma, A., Mehrotra, R. and Alfredsen, K. 2015. Improving real-time inflow forecasting into hydropower reservoirs through a complementary modelling framework. Hydrology and Earth System Sciences 19(8), 3695-3714.
Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., Babu, S., Borrelli, P., Cheng, L., Crochetiere, H., Ehalt Macedo, H., Filgueiras, R., Goichot, M., Higgins, J., Hogan, Z., Lip, B., McClain, M.E., Meng, J., Mulligan, M., Nilsson, C., Olden, J.D., Opperman, J.J., Petry, P., Reidy Liermann, C., Saenz, L., Salinas-Rodriguez, S., Schelle, P., Schmitt, R.J.P., Snider, J., Tan, F., Tockner, K., Valdujo, P.H., van Soesbergen, A. and Zarfl, C. 2019. Mapping the world's free-flowing rivers. Nature 569(7755), 215-221.
Gupta, H.V., Kling, H., Yilmaz, K.K. and Martinez, G.F. 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology 377(1-2), 80-91.
Gurrieri, J.T. and Furniss, G. 2004. Estimation of groundwater exchange in alpine lakes using non-steady mass-balance methods. Journal of Hydrology 297(1-4), 187-208.
Hagemann, M.W., Gleason, C.J. and Durand, M.T. 2017. BAM: Bayesian AMHG-Manning Inference of Discharge Using Remotely Sensed Stream Width, Slope, and Height. Water Resources Research 53(11), 9692-9707.
Hosen, J.D., Allen, G.H., Amatulli, G., Breitmeyer, S., Cohen, M.J., Crump, B.C., Lu, Y., Payet, J.P., Poulin, B.A., Stubbins, A., Yoon, B. and Raymond, P.A. 2021. River network travel time is correlated with dissolved organic matter composition in rivers of the contiguous United States. Hydrological Processes 35(5), e14124.
Knoben, W.J.M., Freer, J.E. and Woods, R.A. 2019. Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. Hydrology and Earth System Sciences 23(10), 4323-4331.
Koskinen, M., Tahvanainen, T., Sarkkola, S., Menberu, M.W., Lauren, A., Sallantaus, T., Marttila, H., Ronkanen, A.K., Parviainen, M., Tolvanen, A., Koivusalo, H. and Nieminen, M. 2017. Restoration of nutrient-rich forestry-drained peatlands poses a risk for high exports of dissolved organic carbon, nitrogen, and phosphorus. Science of the Total Environment 586, 858-869.
Kouraev, A.V., Zakharova, E.A., Samain, O., Mognard, N.M. and Cazenave, A. 2004. Ob' river discharge from TOPEX/Poseidon satellite altimetry (1992–2002). Remote Sensing of Environment 93(1-2), 238-245.
Lajoie, F., Assani, A.A., Roy, A.G. and Mesfioui, M. 2007. Impacts of dams on monthly flow characteristics. The influence of watershed size and seasons. Journal of Hydrology 334(3-4), 423-439.
Lamontagne, J.R., Barber, C.A. and Vogel, R.M. 2020. Improved Estimators of Model Performance Efficiency for Skewed Hydrologic Data. Water Resources Research 56(9), e2020WR027101.

Lehner, B., Liermann, C.R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J.C., Rödel, R., Sindorf, N. and Wisser, D. 2011. High‐resolution mapping of the world's reservoirs and dams for sustainable river‐flow management. Frontiers in Ecology and the Environment 9(9), 494-502.
Leopold, L.B. and Maddock Jr, T. 1953 The hydraulic geometry of stream channels and some physiographic implications, p. 64, Washington, D.C.
Li, Z. and Jin, J. 2017. Evaluating climate change impacts on streamflow variability based on a multisite multivariate GCM downscaling method in the Jing River of China. Hydrology and Earth System Sciences 21(11), 5531-5546.
Liang, X., Lettenmaier, D.P., Wood, E.F. and Burges, S.J. 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research 99(D7), 14415-14428.
Lin, P., Pan, M., Beck, H.E., Yang, Y., Yamazaki, D., Frasson, R., David, C.H., Durand, M., Pavelsky, T.M. and Allen, G.H. 2019. Global reconstruction of naturalized river flows at 2.94 million reaches. Water Resources Research 55(8), 6499-6516.
Livina, V., Ashkenazy, Y., Kizner, Z., Strygin, V., Bunde, A. and Havlin, S. 2003. A stochastic model of river discharge fluctuations. Physica A: Statistical Mechanics and its Applications 330(1-2), 283-290.
Moriasi, D., Gitau, M., Pai, N. and Daggupati, P. 2015. Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria. Transactions of the ASABE 58(6), 1763-1785.
Mueller, D.S., Wagner, C.R., Rehmel, M.S., Oberg, K.A. and Rainville, F. 2013 Measuring discharge with acoustic Doppler current profilers from a moving boat, p. 116, Reston, VA.
Nash, J.E. and Sutcliffe, J.V. 1970. River flow forecasting through conceptual models part I — A discussion of principles. Journal of Hydrology 10(3), 282-290.
Nickles, C., Beighley, E., Durand, M. and Prata de Moraes Frasson, R. 2020a. Integrating Lateral Inflows Into a SWOT Mission River Discharge Algorithm. Water Resources Research 56(10), e2019WR026589.
Nickles, C., Beighley, E. and Feng, D. 2020b. The Applicability of SWOT's Non-Uniform Space–Time Sampling in Hydrologic Model Calibration. Remote Sensing 12(19), 3241.
Nickles, C., Beighley, E., Zhao, Y., Durand, M., David, C. and Lee, H. 2019. How Does the Unique Space‐Time Sampling of the SWOT Mission Influence River Discharge Series Characteristics? Geophysical Research Letters 46(14), 8154-8161.
Oki, T. and Kanae, S. 2006. Global hydrological cycles and world water resources. Science 313(5790), 1068-1072.
Pastore, M. and Calcagni, A. 2019. Measuring Distribution Similarities Between Samples: A Distribution-Free Overlapping Index. Frontiers in Psychology 10(1089), 1089.
Pekel, J.F., Cottam, A., Gorelick, N. and Belward, A.S. 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540(7633), 418-422.
Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Durr, H., Meybeck, M., Ciais, P. and Guth, P. 2013. Global carbon dioxide emissions from inland waters. Nature 503(7476), 355-359.

Rodell, M., Famiglietti, J.S., Wiese, D.N., Reager, J.T., Beaudoing, H.K., Landerer, F.W. and Lo, M.H. 2018. Emerging trends in global freshwater availability. Nature 557(7707), 651-659.
Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J.K., Walker, J.P., Lohmann, D. and Toll, D. 2004. The global land data assimilation system. Bulletin of the American Meteorological Society 85(3), 381-394.
Samaniego, L., Kumar, R. and Attinger, S. 2010. Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. Water Resources Research 46(5).
Shields Jr, F.D., Simon, A. and Steffen, L.J. 2002. Reservoir effects on downstream river channel migration. Environmental Conservation 27(1), 54-66.
Siebert, S., Burke, J., Faures, J.M., Frenken, K., Hoogeveen, J., Döll, P. and Portmann, F.T. 2010. Groundwater use for irrigation – a global inventory. Hydrology and Earth System Sciences 14(10), 1863-1880.
Smith, L.C., Isacks, B.L., Bloom, A.L. and Murray, A.B. 1996. Estimation of Discharge From Three Braided Rivers Using Synthetic Aperture Radar Satellite Imagery: Potential Application to Ungaged Basins. Water Resources Research 32(7), 2021-2034.
Smith, L.C., Isacks, B.L., Forster, R.R., Bloom, A.L. and Preuss, I. 1995. Estimation of Discharge From Braided Glacial Rivers Using ERS 1 Synthetic Aperture Radar: First Results. Water Resources Research 31(5), 1325-1329.
Smith, L.C. and Pavelsky, T.M. 2008. Estimation of river discharge, propagation speed, and hydraulic geometry from space: Lena River, Siberia. Water Resources Research 44(3).
Syed, T.H., Famiglietti, J.S., Chambers, D.P., Willis, J.K. and Hilburn, K. 2010. Satellite-based global-ocean mass balance estimates of interannual variability and emerging trends in continental freshwater discharge. Proceedings of the National Academy of Sciences 107(42), 17916-17921.
Te Chow, V. (2010) Applied hydrology, Tata McGraw-Hill Education.
Tuozzolo, S., Lind, G., Overstreet, B., Mangano, J., Fonstad, M., Hagemann, M., Frasson, R.P.M., Larnier, K., Garambois, P.A., Monnier, J. and Durand, M. 2019. Estimating River Discharge With Swath Altimetry: A Proof of Concept Using AirSWOT Observations. Geophysical Research Letters 46(3), 1459-1466.
Turnipseed, D.P. and Sauer, V.B. 2010 Discharge measurements at gaging stations, Reston, VA.
Vörösmarty, C.J. and Sahagian, D. 2000. Anthropogenic Disturbance of the Terrestrial Water Cycle. BioScience 50(9), 753-765.
Vörösmarty, C.J., Sharma, K.P., Fekete, B.M., Copeland, A.H., Holden, J., Marble, J. and Lough, J.A. 1997. The storage and aging of continental runoff in large reservoir systems of the world. Ambio 26(4) 210-219.
Wada, Y., Reager, J.T., Chao, B.F., Wang, J., Lo, M.H., Song, C., Li, Y. and Gardner, A.S. 2017. Recent Changes in Land Water Storage and its Contribution to Sea Level Variations. Surveys in Geophysics 38(1), 131-152.
Wang, J., Sheng, Y. and Tong, T.S.D. 2014. Monitoring decadal lake dynamics across the Yangtze Basin downstream of Three Gorges Dam. Remote Sensing of Environment 152, 251-269.

Wang, J., Sheng, Y. and Wada, Y. 2017. Little impact of Three Gorges Dam on recent decadal lake decline across China's Yangtze Plain. Water Resources Research 53(5), 3854-3877.

Wang, J. (PI), Sheng, Y. (co-PI), and Allen, G. (co-PI). Integrating reservoirs into SWOT's global surface water storage and discharge monitoring. Surface Water and Ocean Topography (SWOT) Science Team, National Aeronautics and Space Administration (NASA), Grant# 80NSSC20K1143, 2020–2024.

Wang, W., Lee, X., Xiao, W., Liu, S., Schultz, N., Wang, Y., Zhang, M. and Zhao, L. 2018. Global lake evaporation accelerated by changes in surface energy allocation in a warmer climate. Nature Geoscience 11(6), 410-414.

Wang, X., Gong, P., Zhao, Y., Xu, Y., Cheng, X., Niu, Z., Luo, Z., Huang, H., Sun, F. and Li, X. 2013. Water-level changes in China's large lakes determined from ICESat/GLAS data. Remote Sensing of Environment 132, 131-144.

Wellmeyer, J.L., Slattery, M.C. and Phillips, J.D. 2005. Quantifying downstream impacts of impoundment on flow regime and channel planform, lower Trinity River, Texas. Geomorphology 69(1-4), 1-13.

Winton, R.S., Calamita, E. and Wehrli, B. 2019. Reviews and syntheses: Dams, water quality and tropical reservoir stratification. Biogeosciences 16(8), 1657-1671.

Wolock, D. 2003 Hydrologic landscape regions of the United States. U.S. Geological Survey, Reston, Virginia.

Woodward, D.E., Hawkins, R.H., Jiang, R., Allen T. Hjelmfelt, J., Mullem, J.A.V. and Quan, Q.D. (2003) World Water and Environmental Resources Congress 2003, pp. 1-10.

Woolway, R.I., Kraemer, B.M., Lenters, J.D., Merchant, C.J., O'Reilly, C.M. and Sharma, S. 2020. Global lake responses to climate change. Nature Reviews Earth & Environment 1(8), 388-403.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y. and Mocko, D. 2012. Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. Journal of Geophysical Research: Atmospheres 117(D3).

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P.D., Allen, G.H. and Pavelsky, T.M. 2019. MERIT Hydro: A High‐Resolution Global Hydrography Map Based on Latest Topography Dataset. Water Resources Research 55(6), 5053-5073.

Zarfl, C., Lumsdon, A.E., Berlekamp, J., Tydecks, L. and Tockner, K. 2014. A global boom in hydropower dam construction. Aquatic Sciences 77(1), 161-170.

Zhang, H., Gorelick, S.M., Zimba, P.V. and Zhang, X. 2017. A remote sensing method for estimating regional reservoir area and evaporative loss. Journal of Hydrology 555, 213-227.

Zhao, G. and Gao, H. 2019. Estimating reservoir evaporation losses for the United States: Fusing remote sensing and modeling approaches. Remote Sensing of Environment 226, 109-124.

Zhao, G. and Gao, H.L. 2018. Automatic Correction of Contaminated Images for Assessment of Reservoir Surface Area Dynamics. Geophysical Research Letters 45(12), 6092-6099.

Zhao, G., Li, Y., Zhou, L. and Gao, H. 2022. Evaporative water loss of 1.42 million global lakes. Nature Communications 13(1), 3686.

Appendix A - Sampled Reservoirs – Vector Representations

Figure A.1 Abiquiu Reservoir

Figure A.2 Allatoona Reservoir

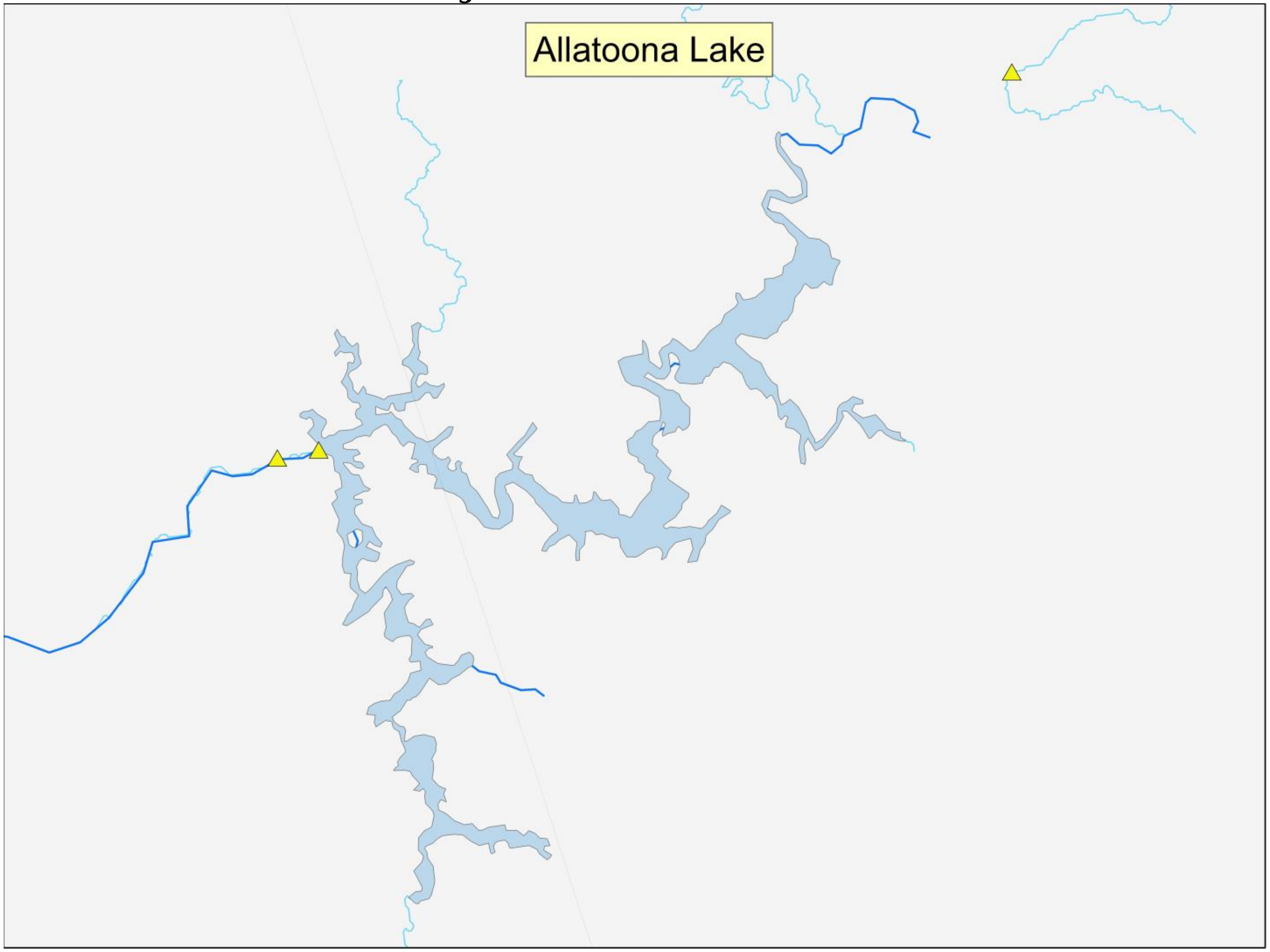


Figure A.3 Barlett and Horseshoe Reservoir

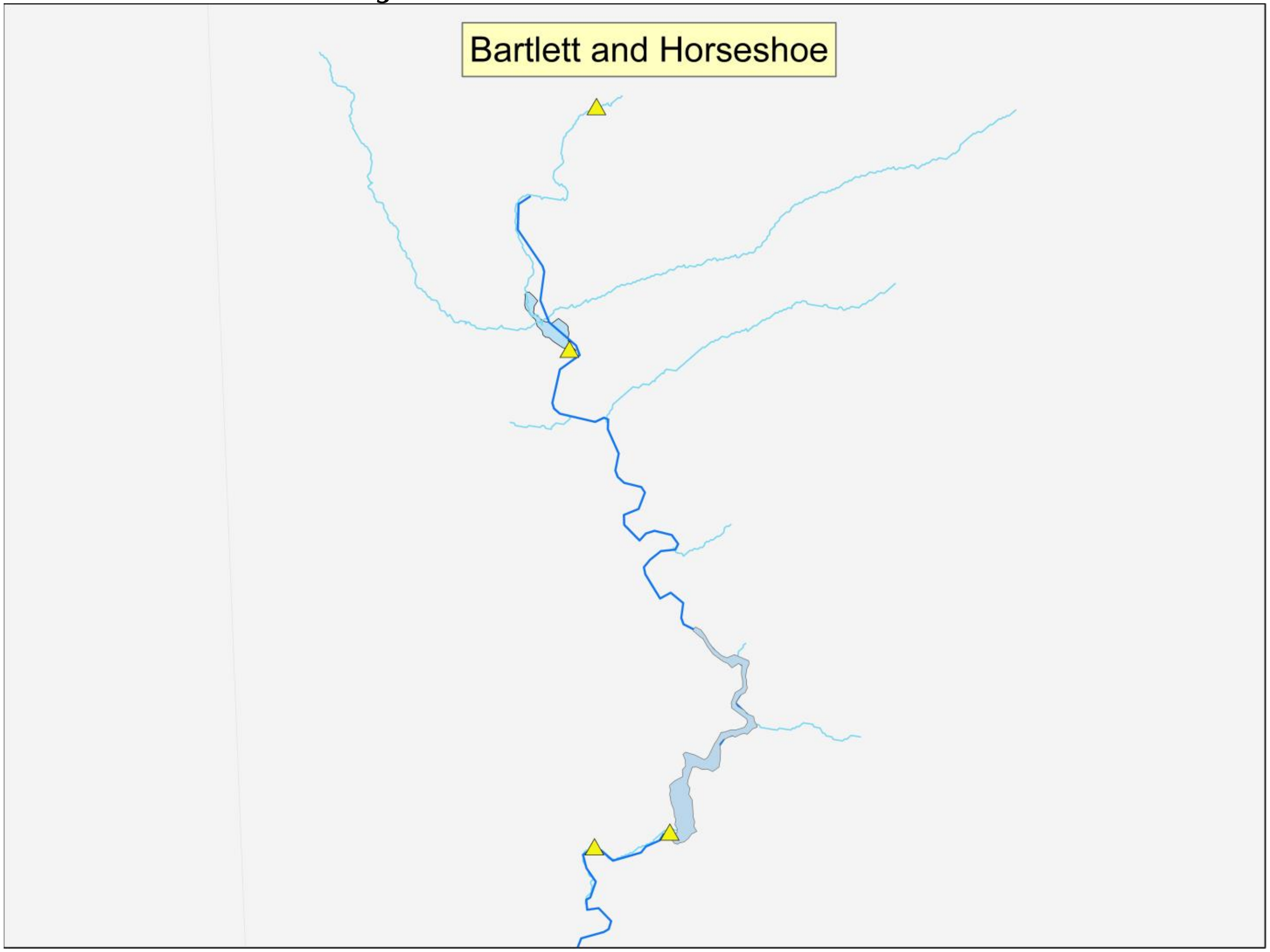


Figure A.4 Brantly Lake

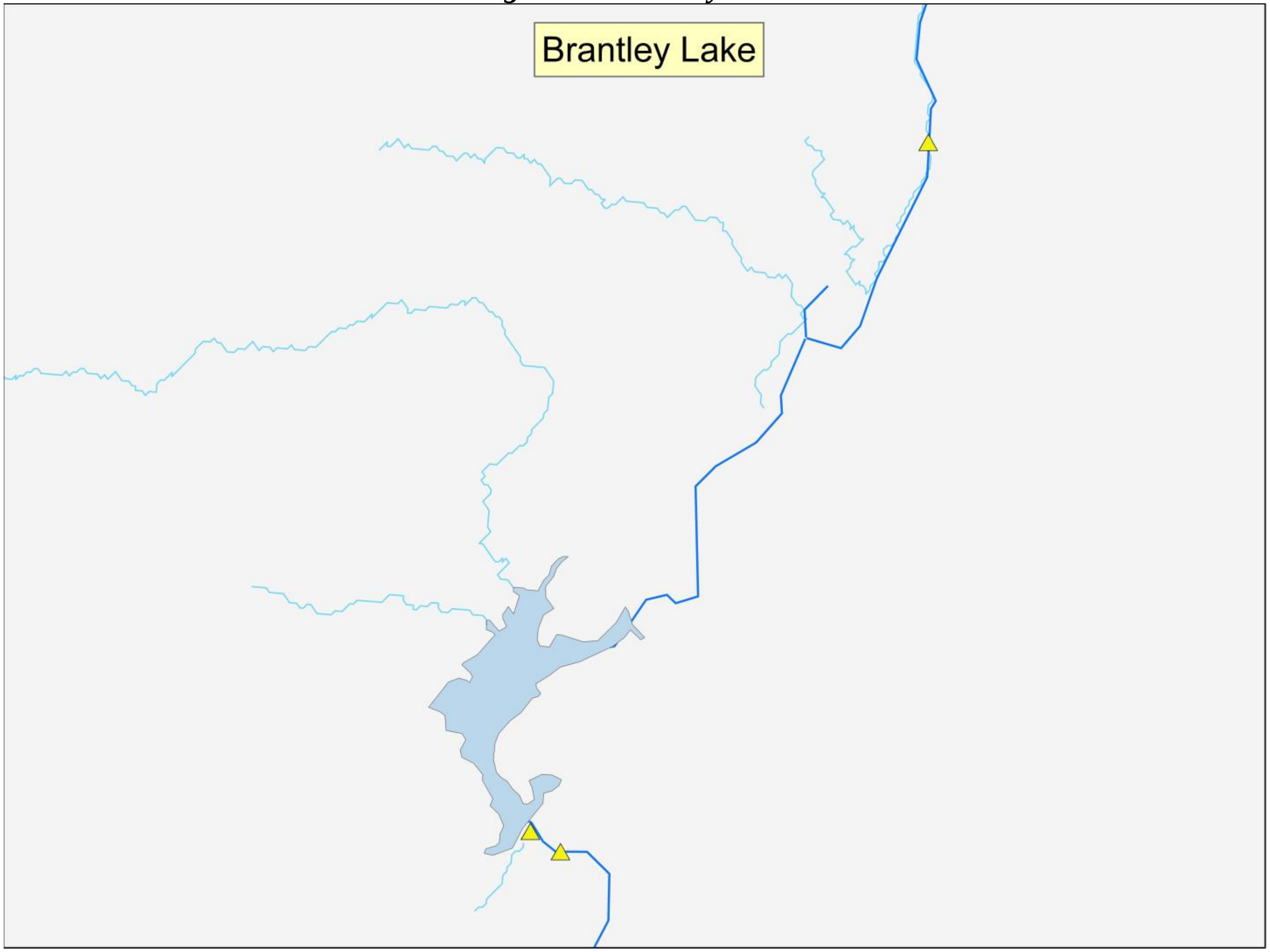


Figure A.5 CJ Strike Reservoir

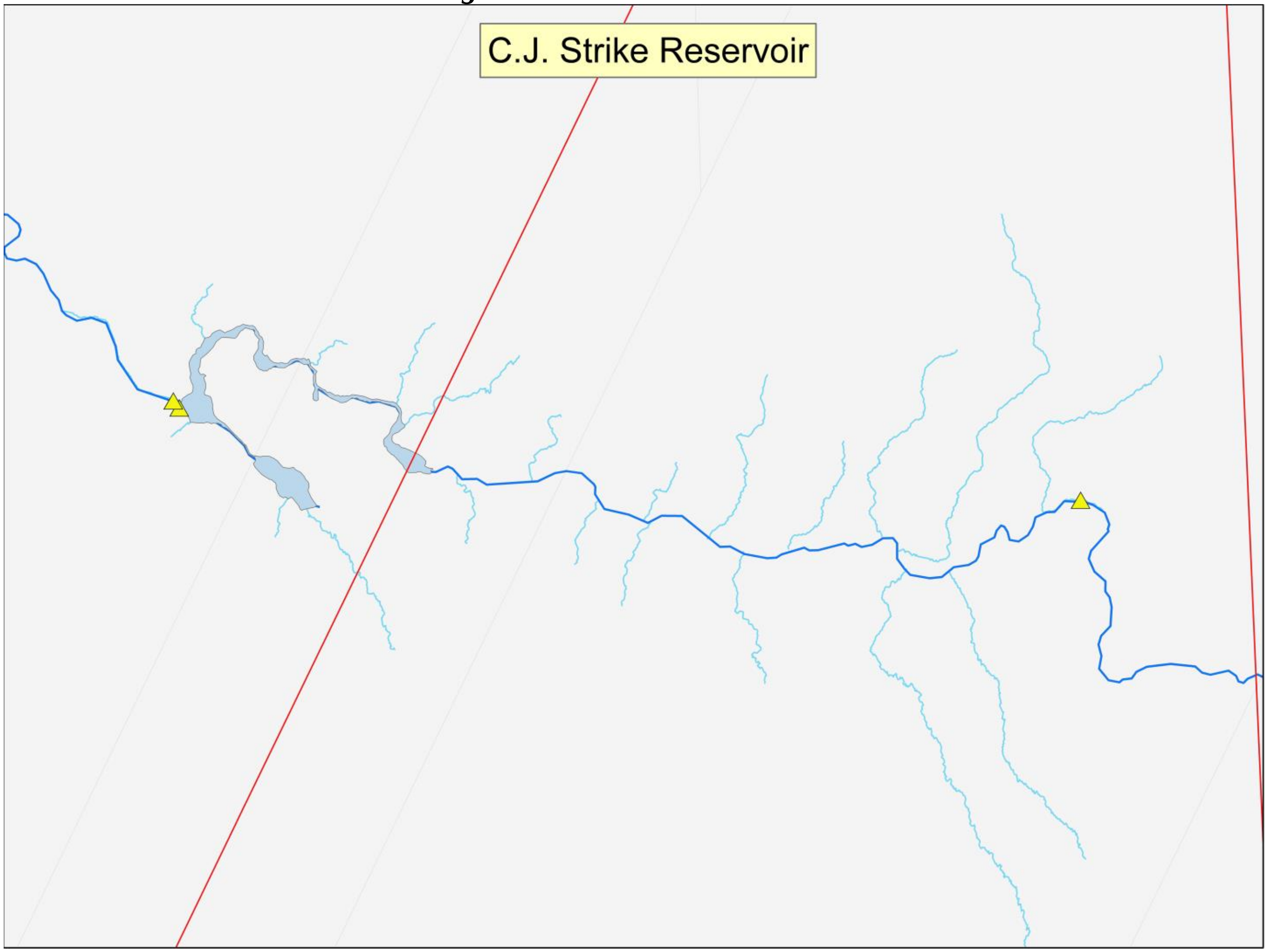


Figure A.6 Cochiti Lake

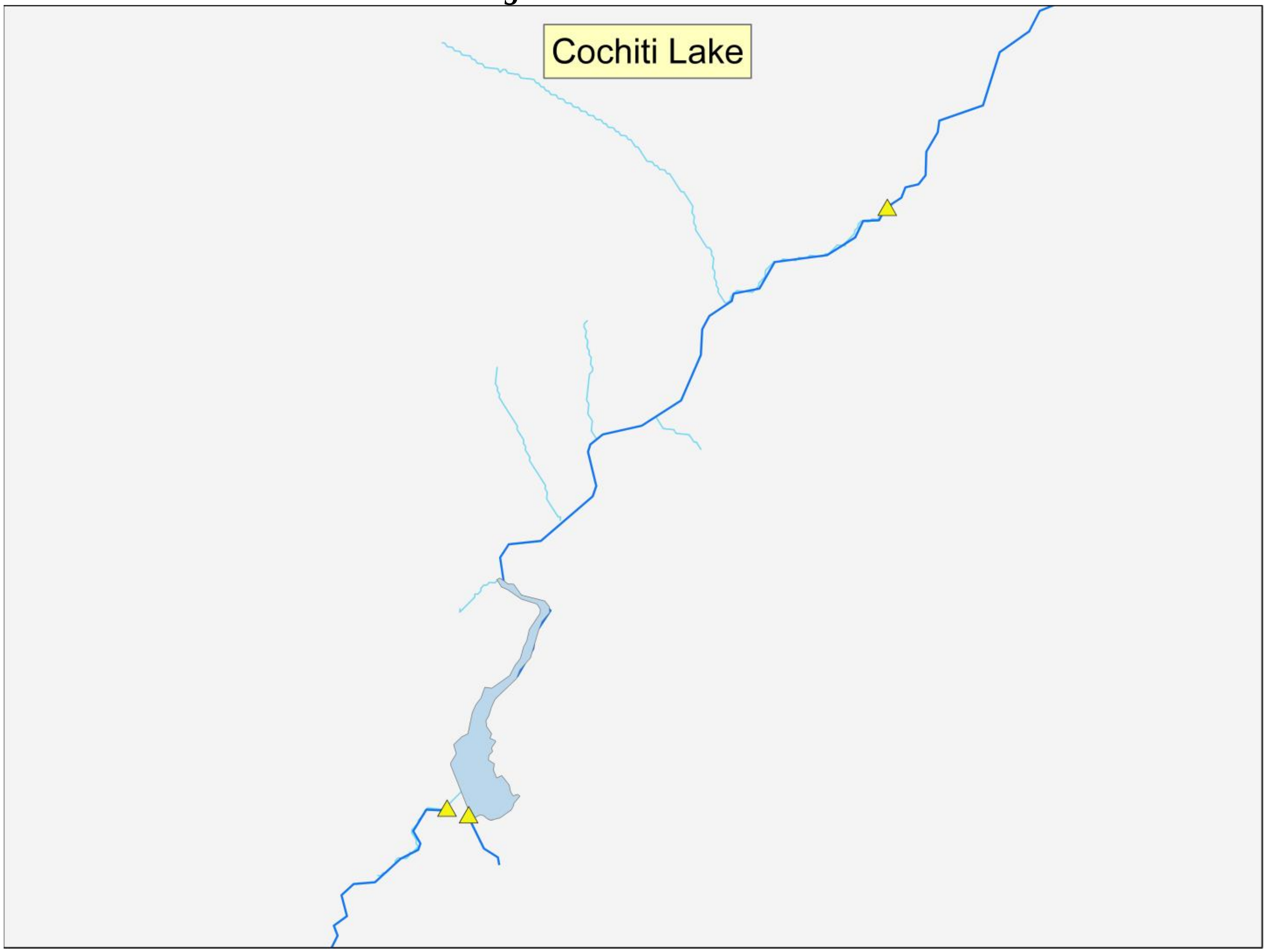


Figure A.7 Dworshak Reservoir

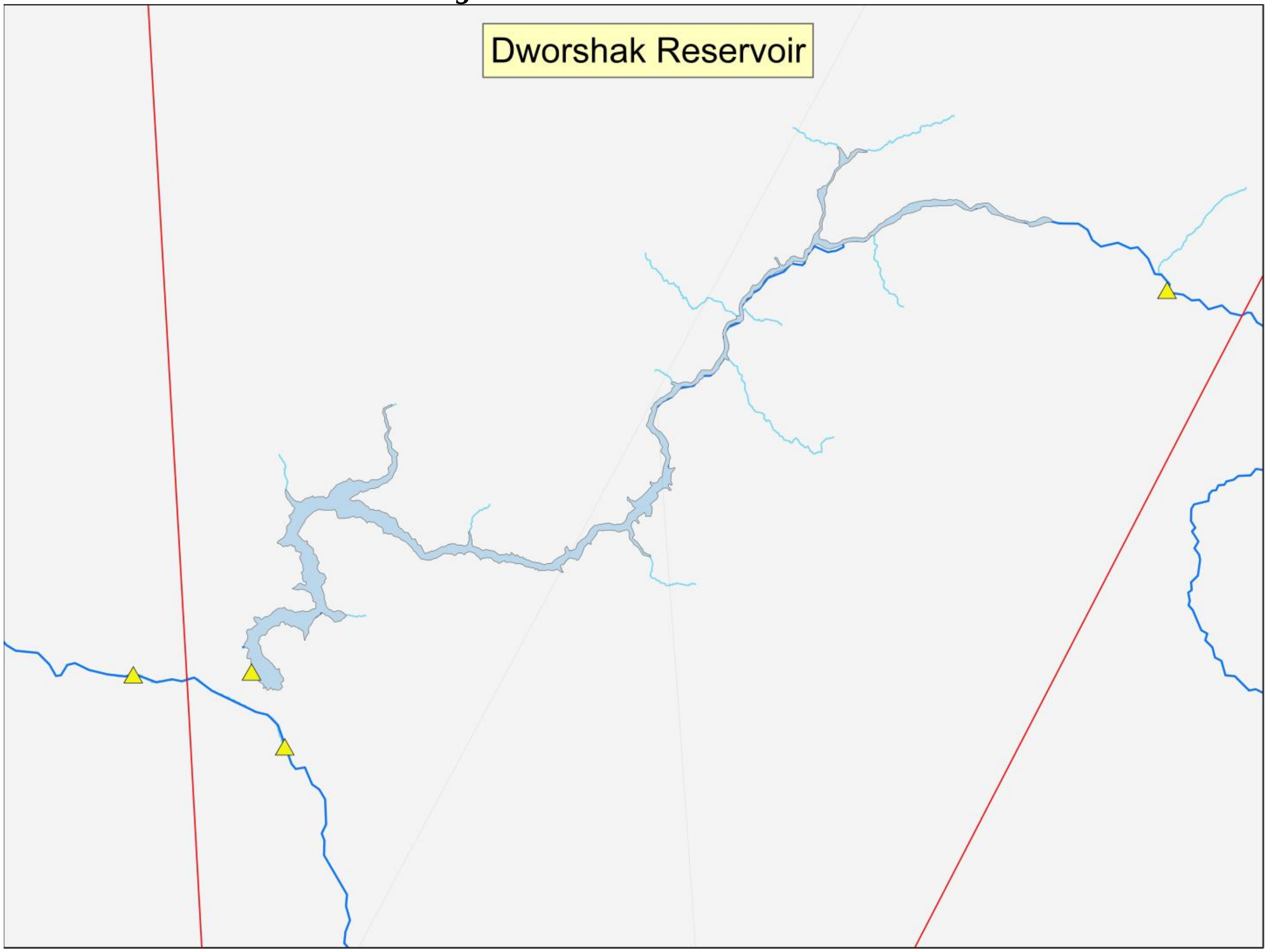


Figure A.8 Harlan County Lake

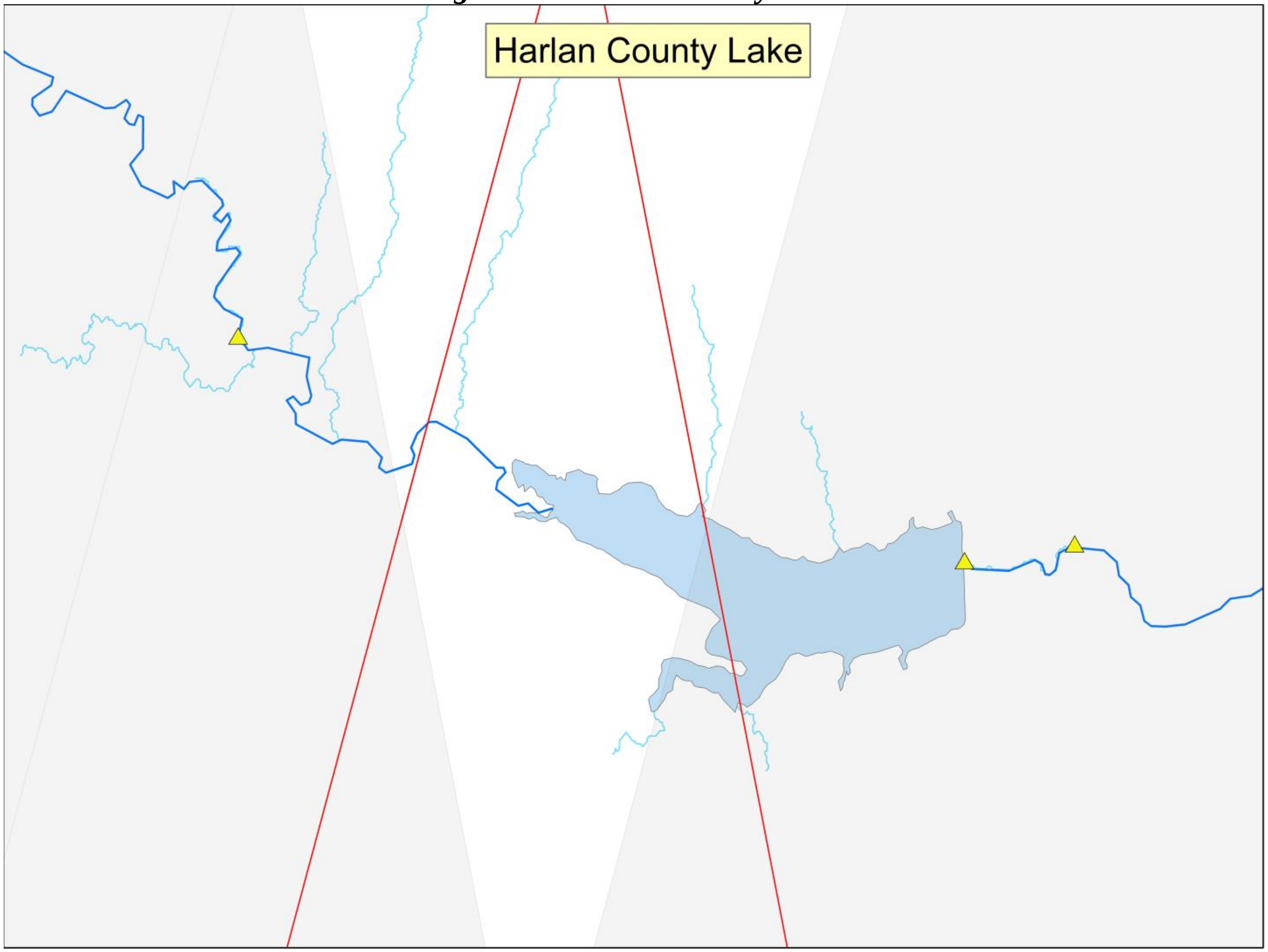


Figure A.9 John Martin Reservoir

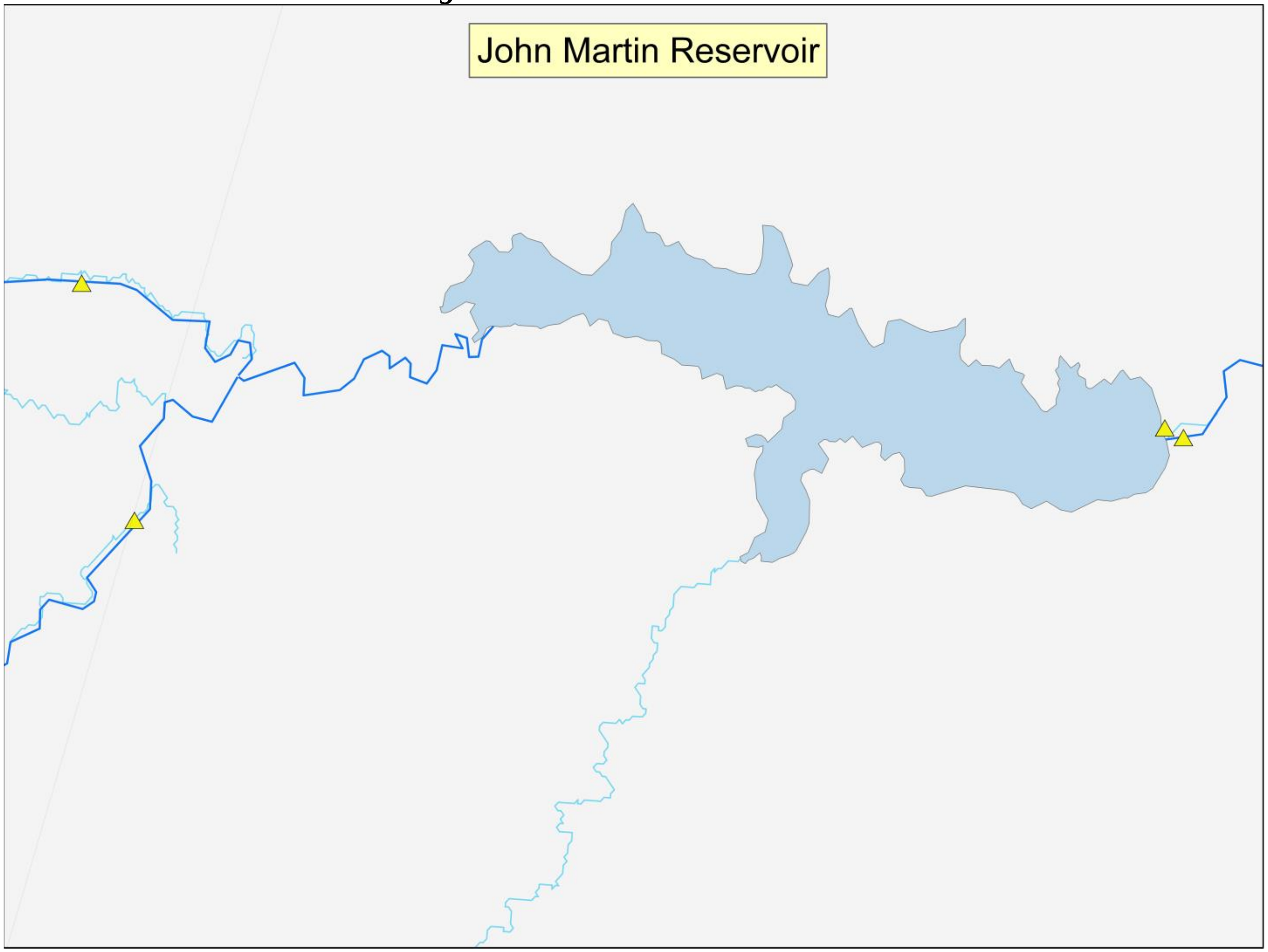


Figure A.10 Lake Granbury

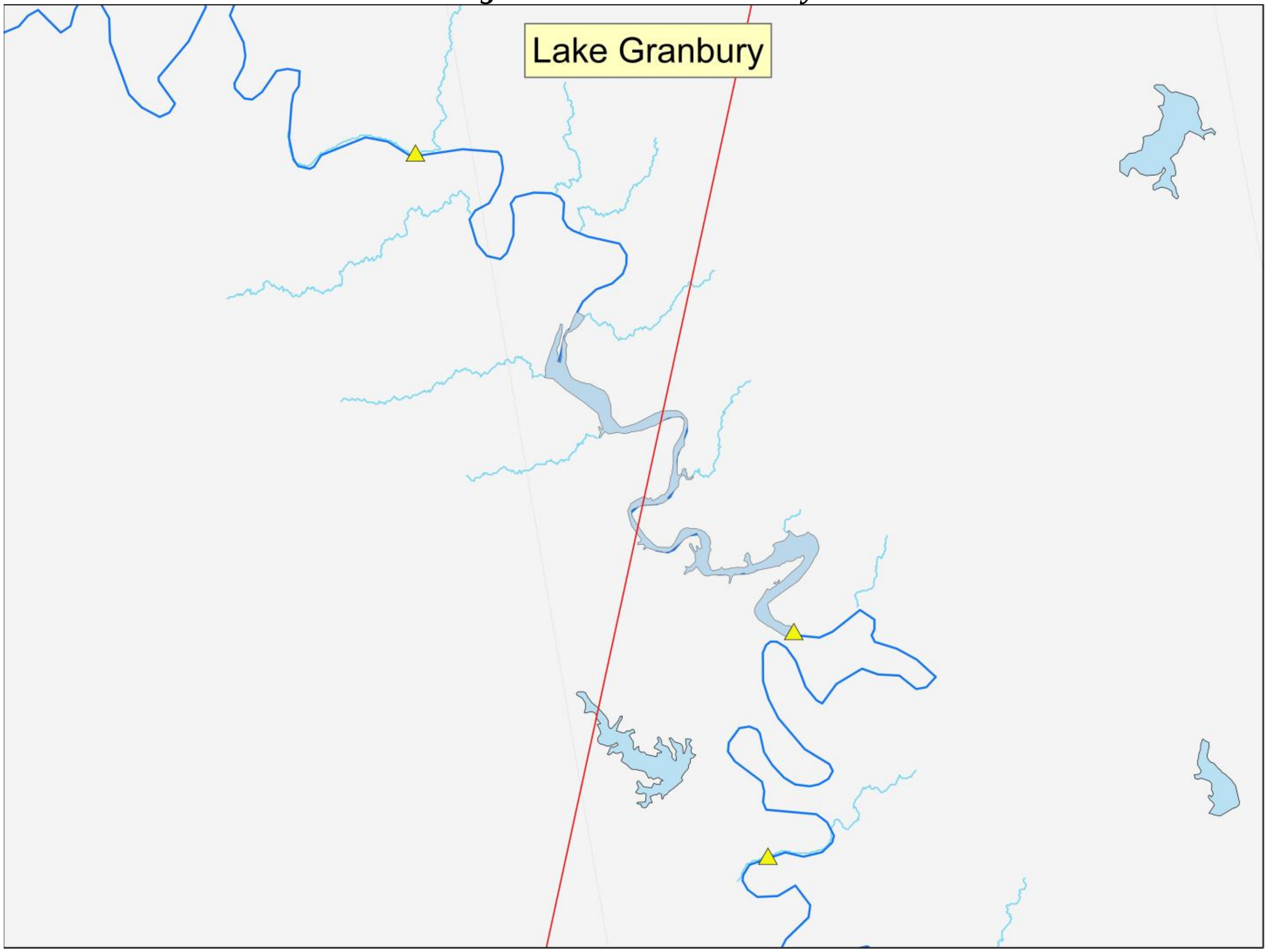


Figure A.11 Lake Havasu

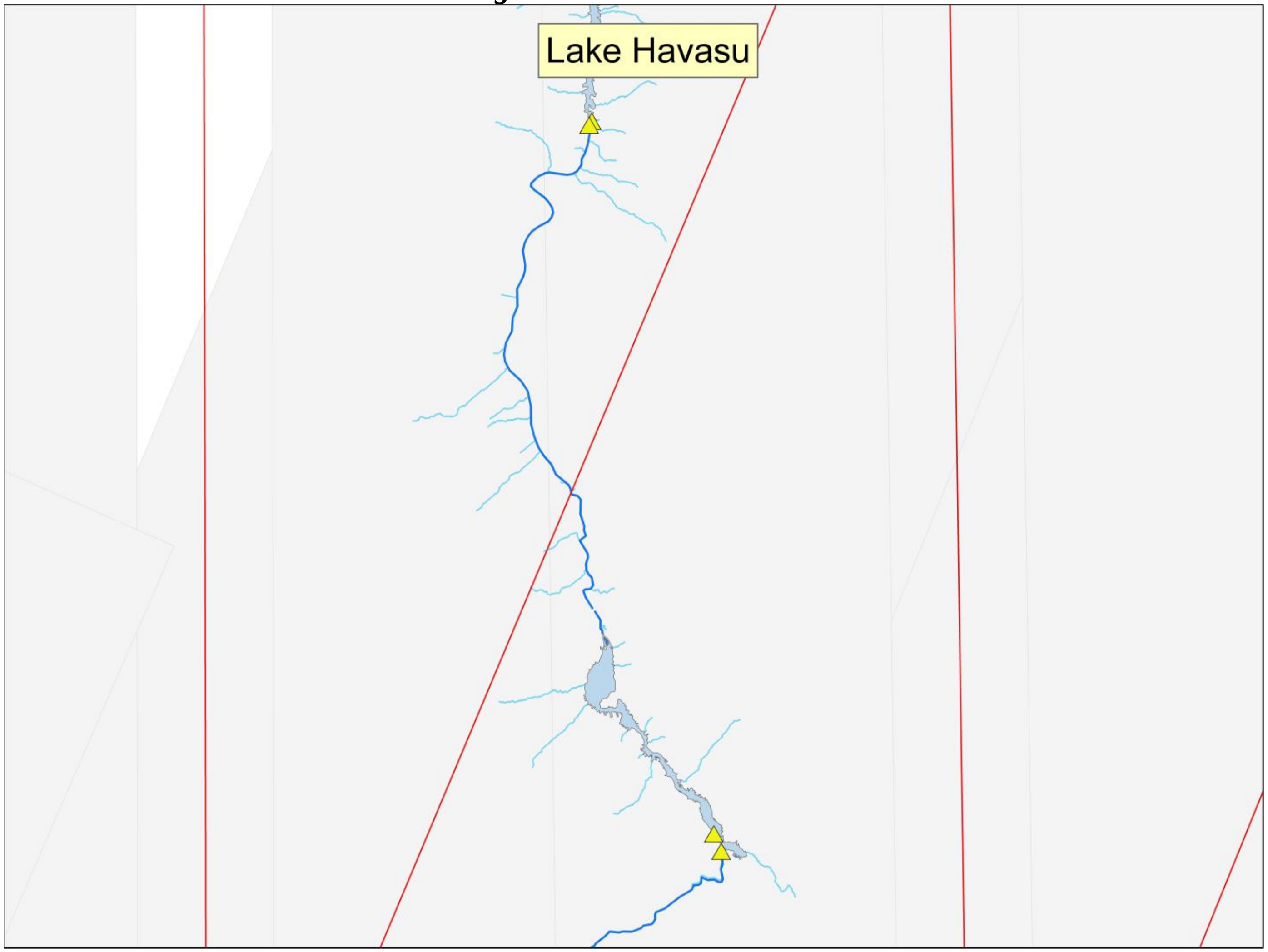


Figure A.12 Lake Lanier

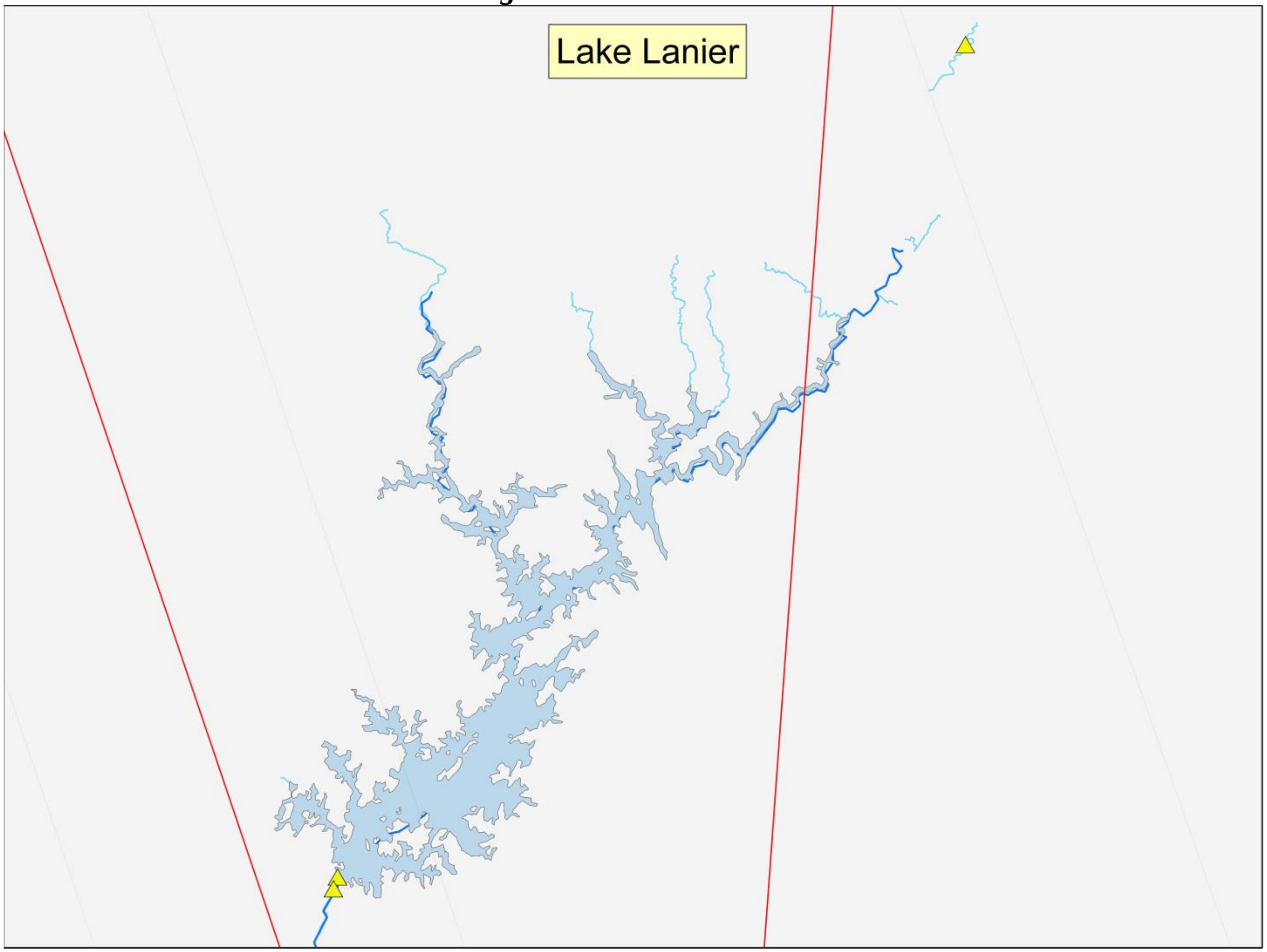


Figure A.13 Lake Livingston

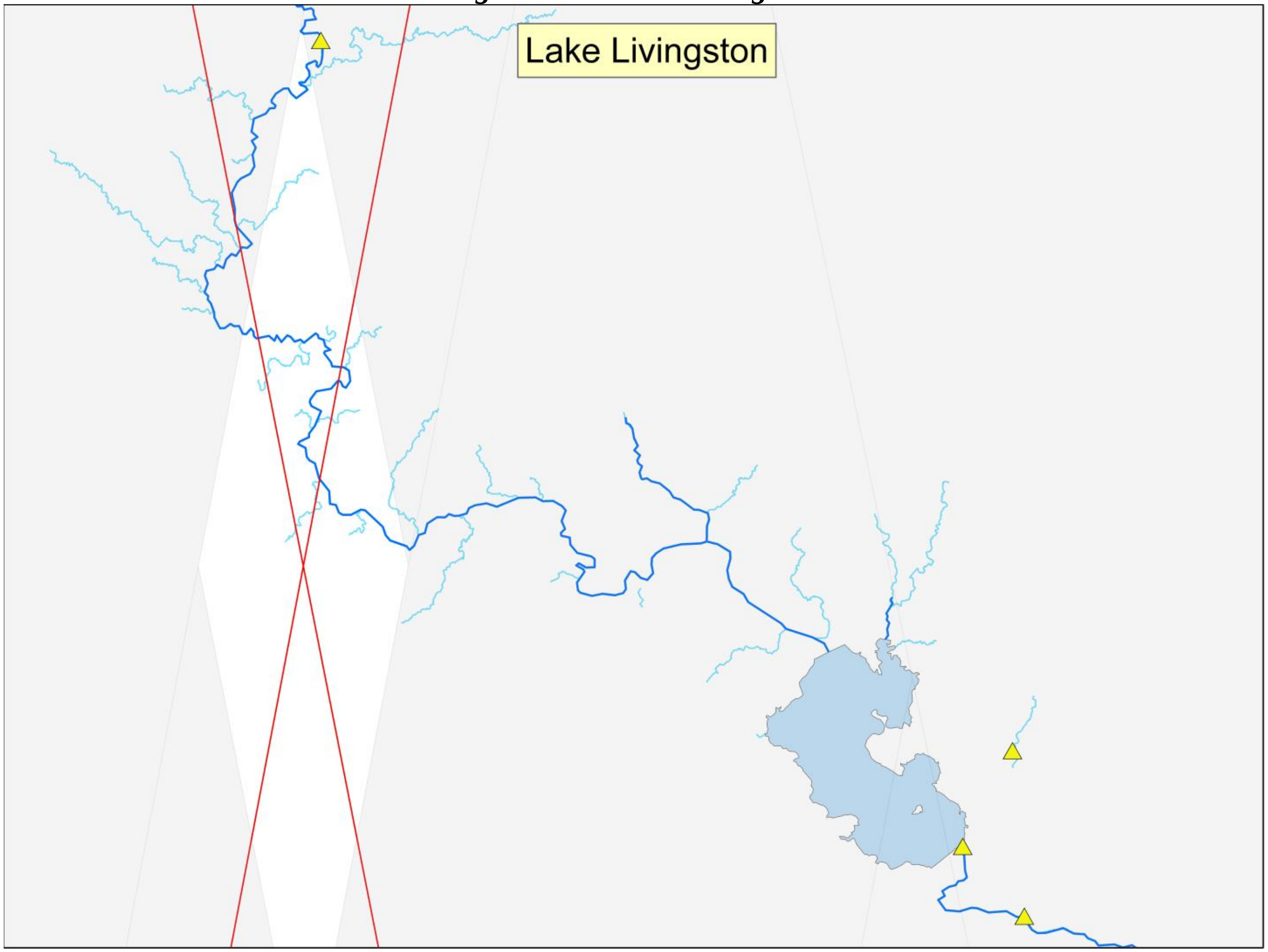


Figure A.14 Lake Mohave

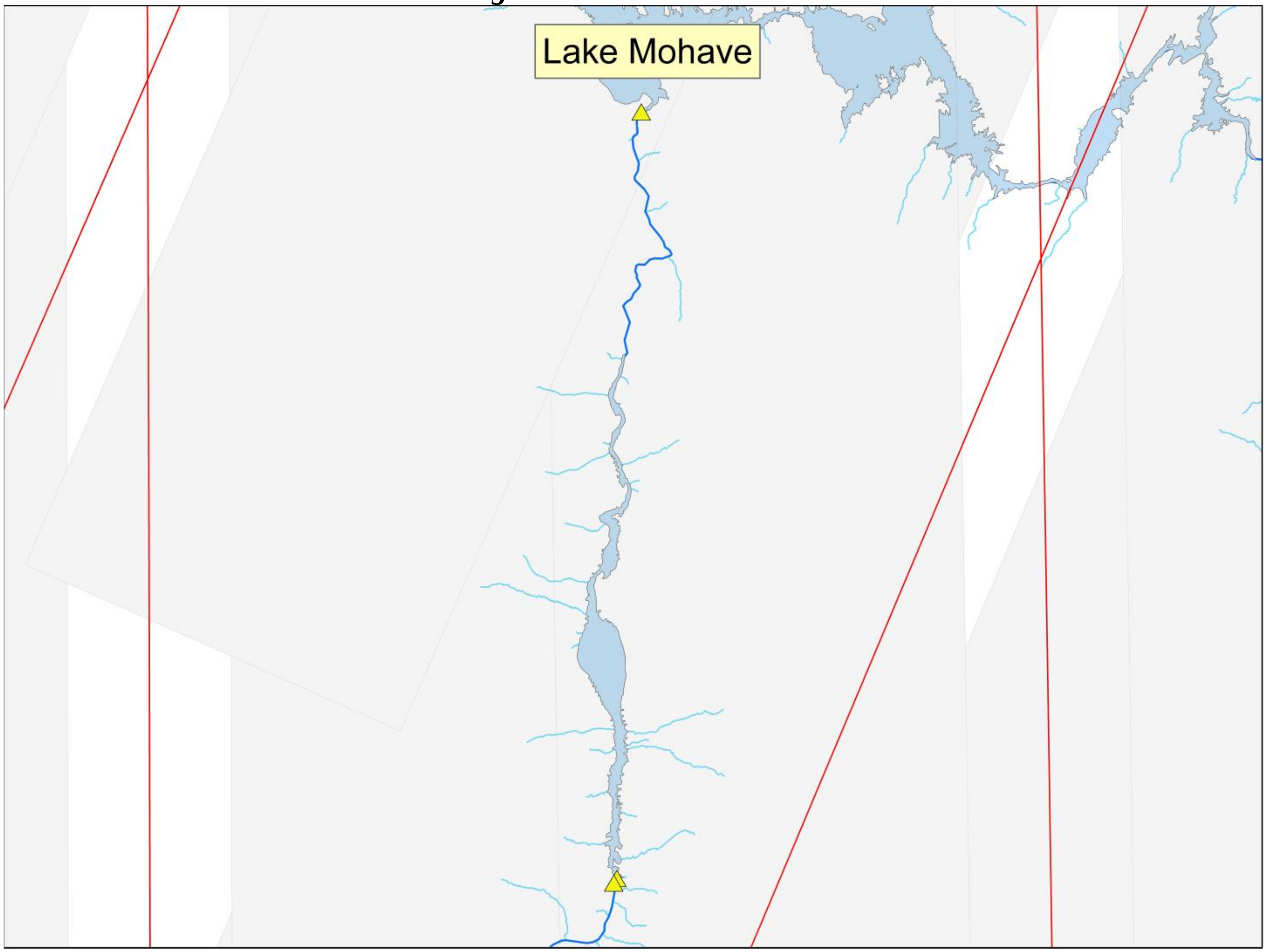


Figure A.15 Lake Red Rock

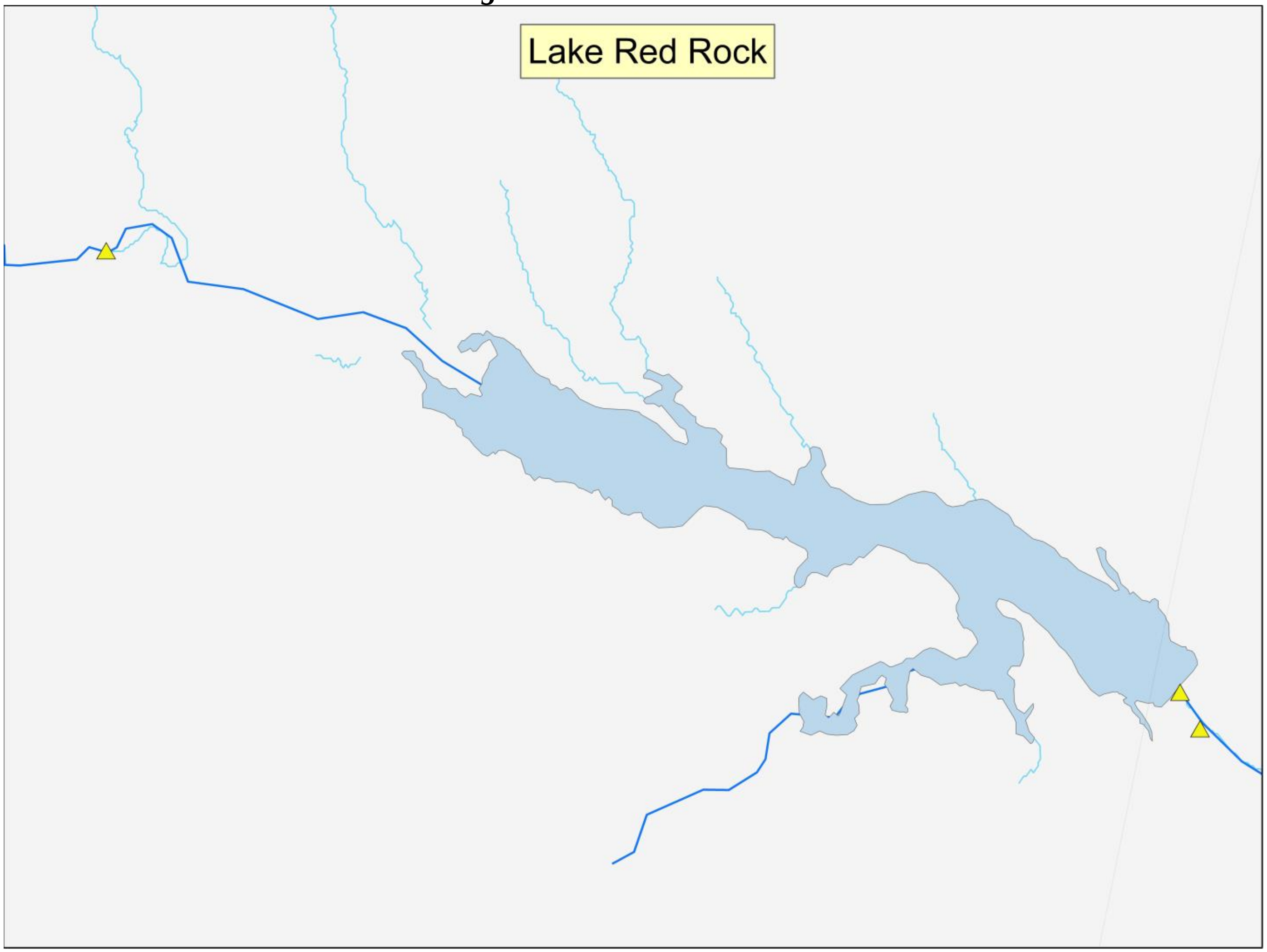


Figure A.16 Lake Seminole

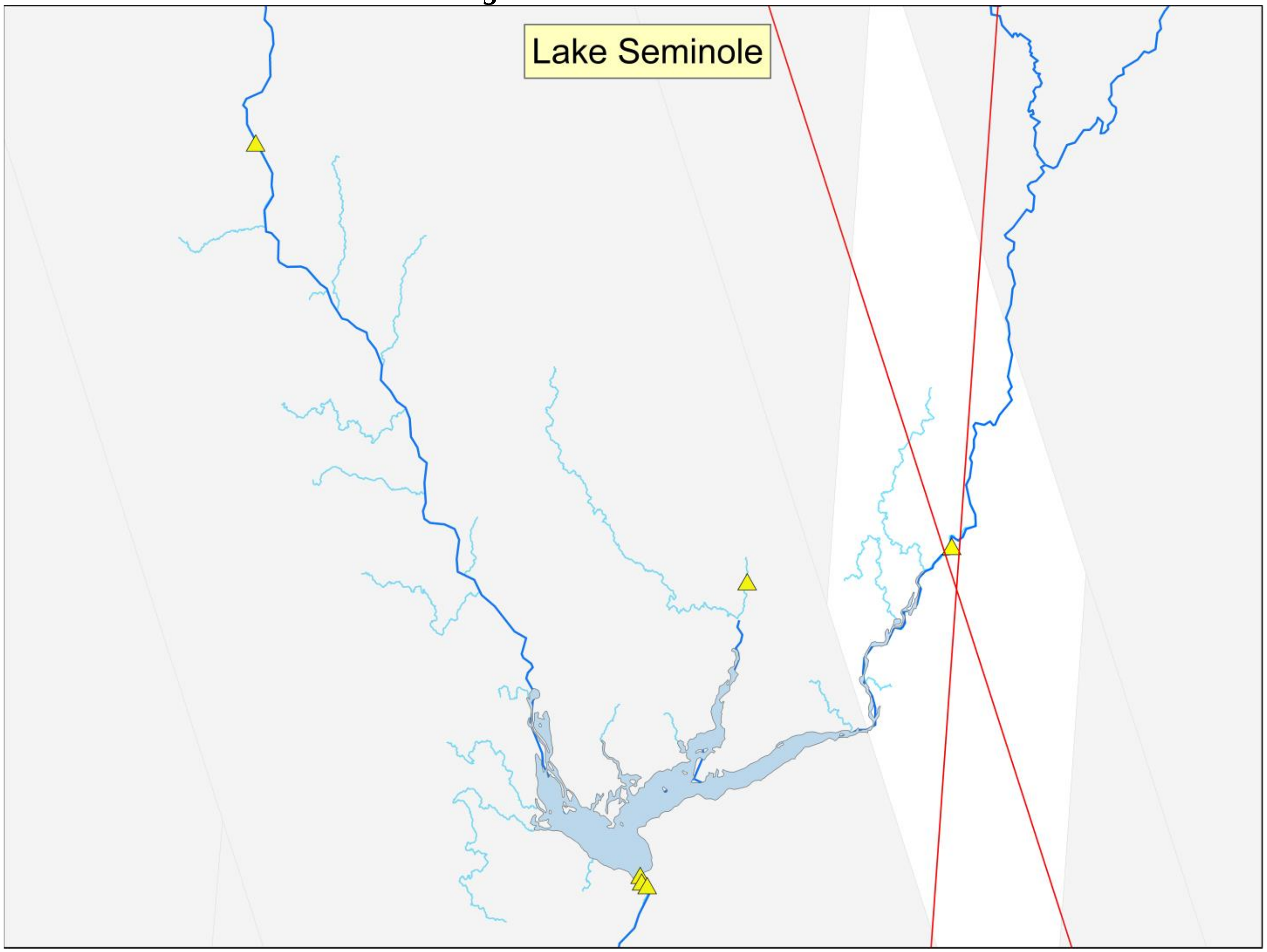


Figure A.17 Lake Taxoma

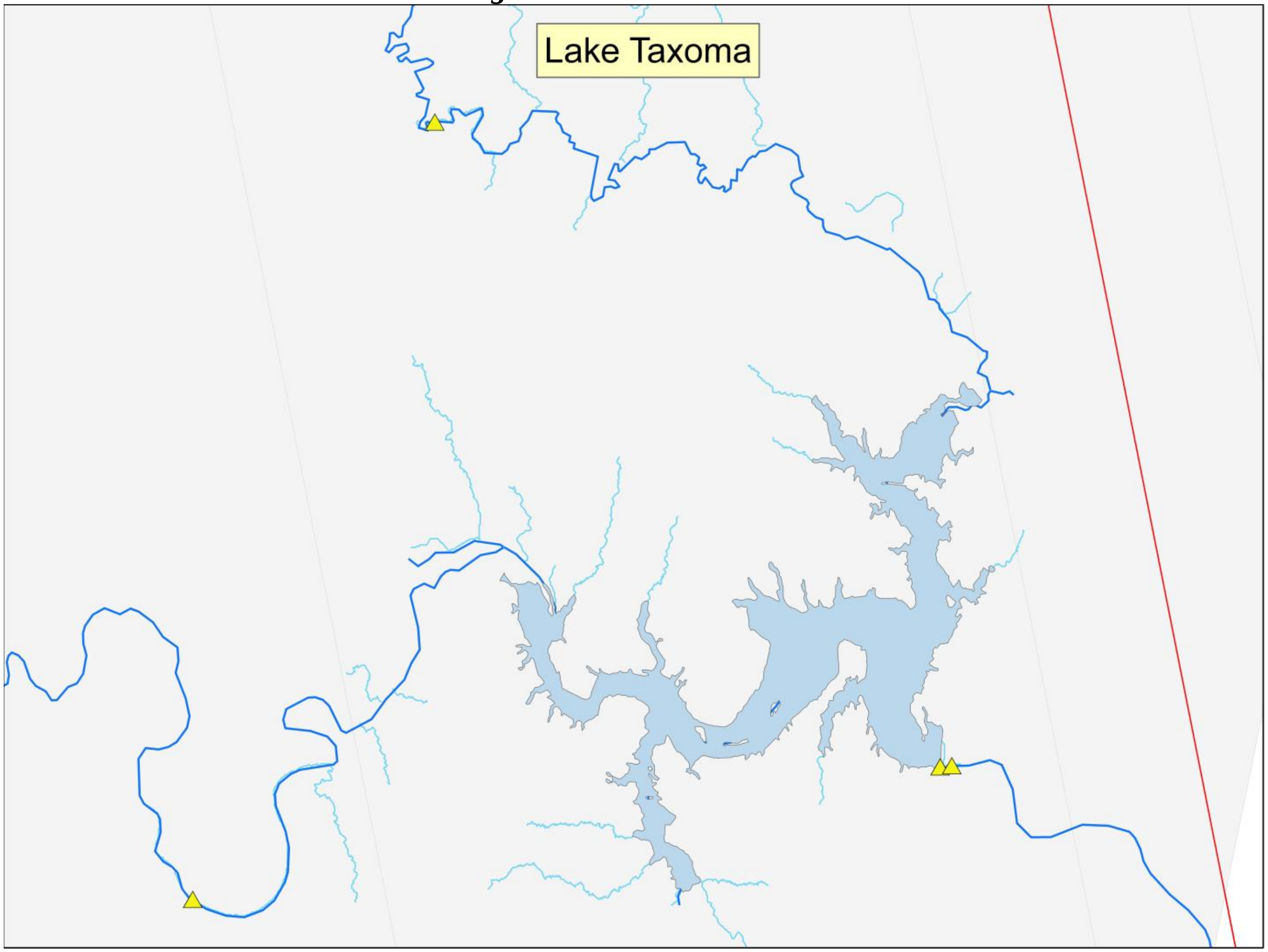


Figure A.18 Milford Lake

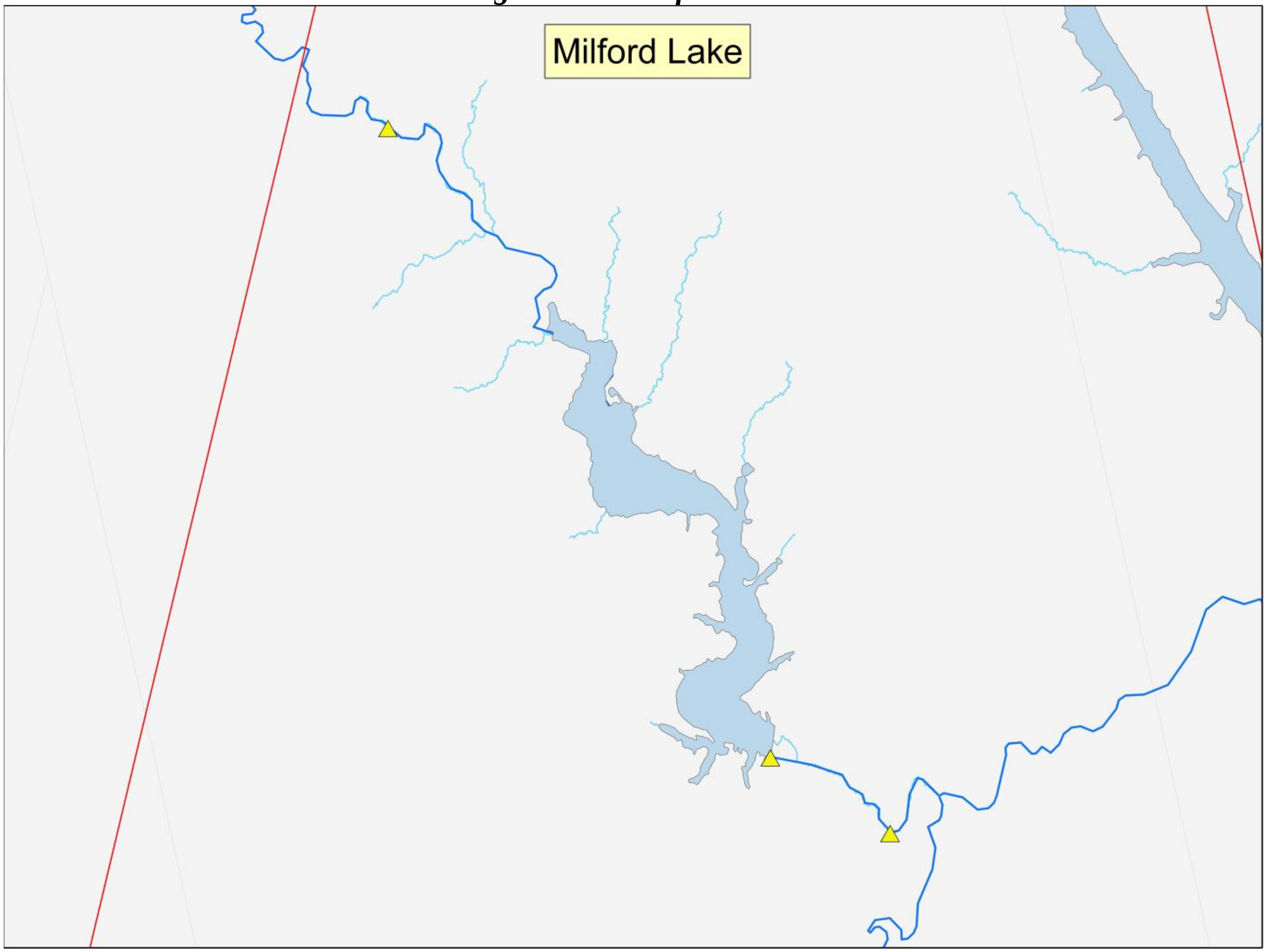


Figure A.19 Navajo Lake

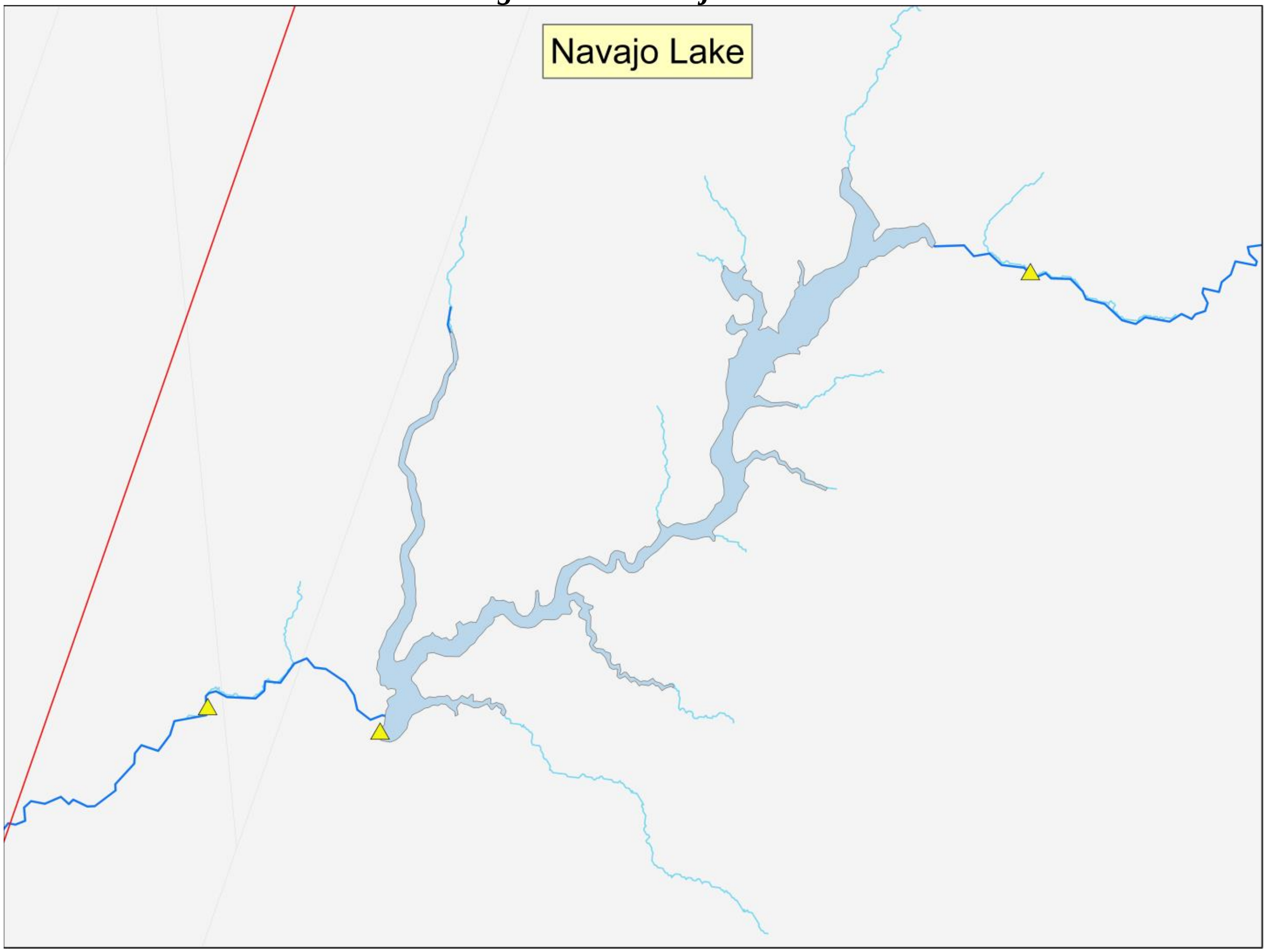


Figure A.20 Possum Kingdom Lake

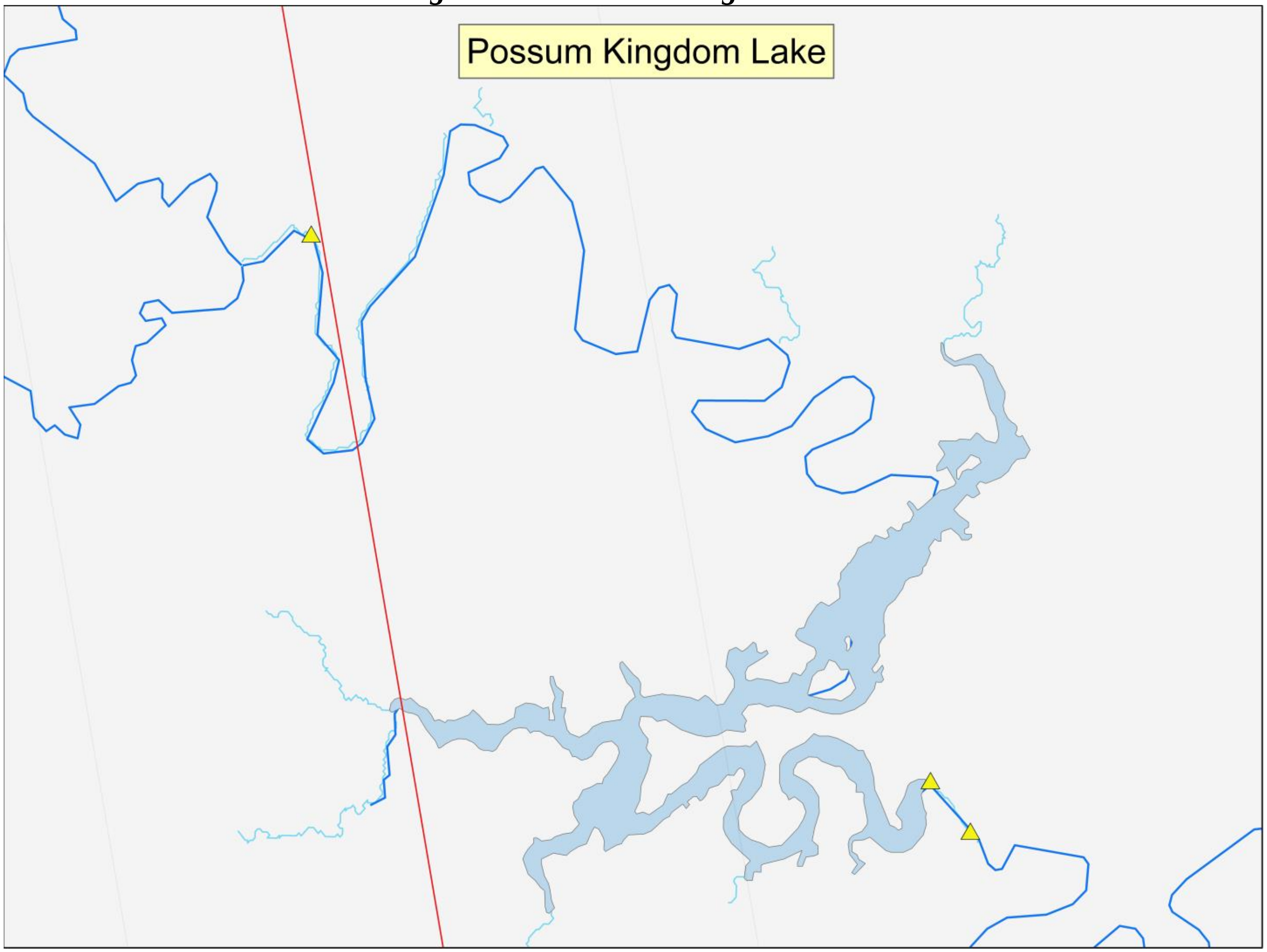


Figure A.21 Shasta Lake

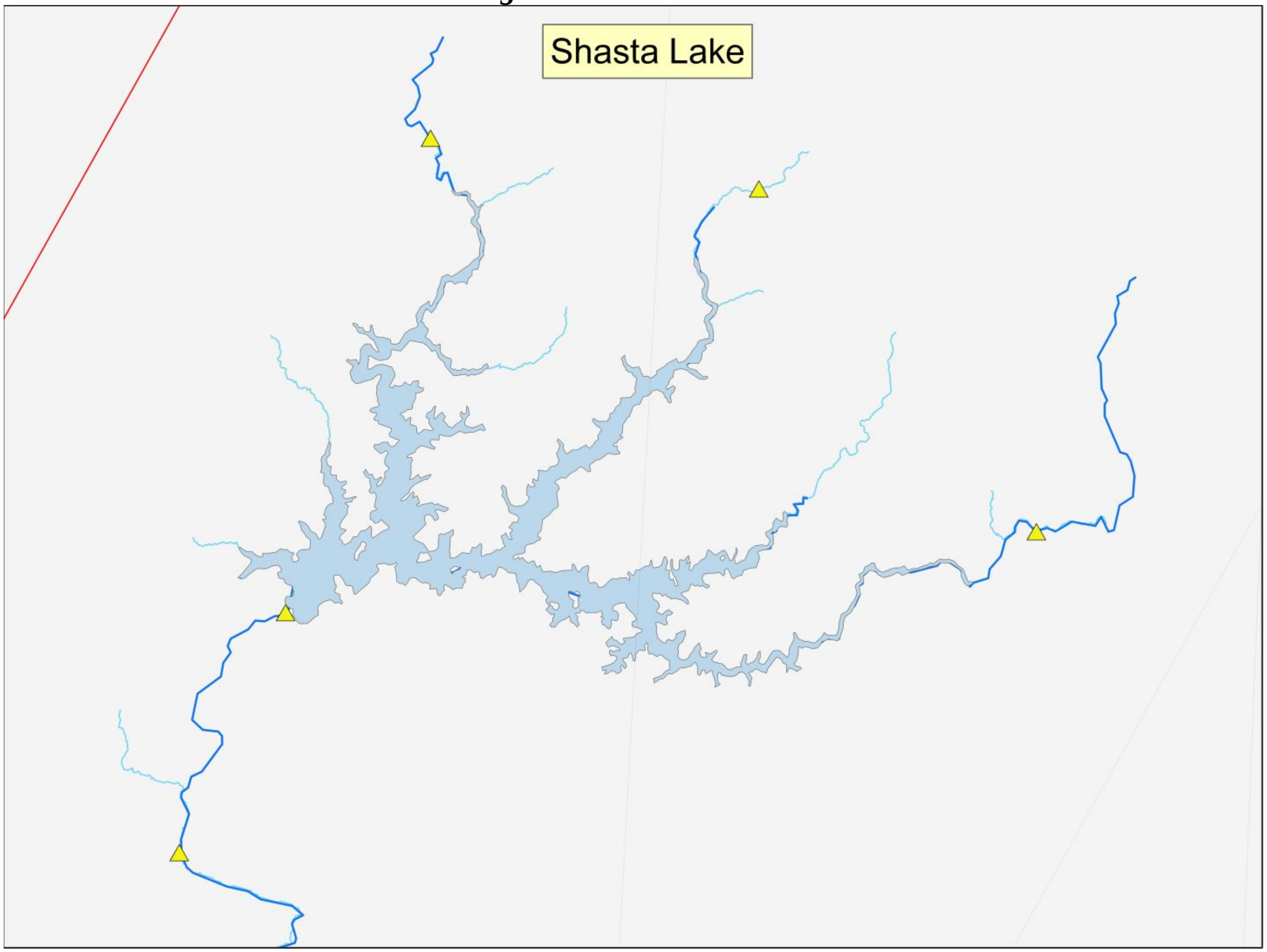


Figure A.22 Toledo Bend Lake

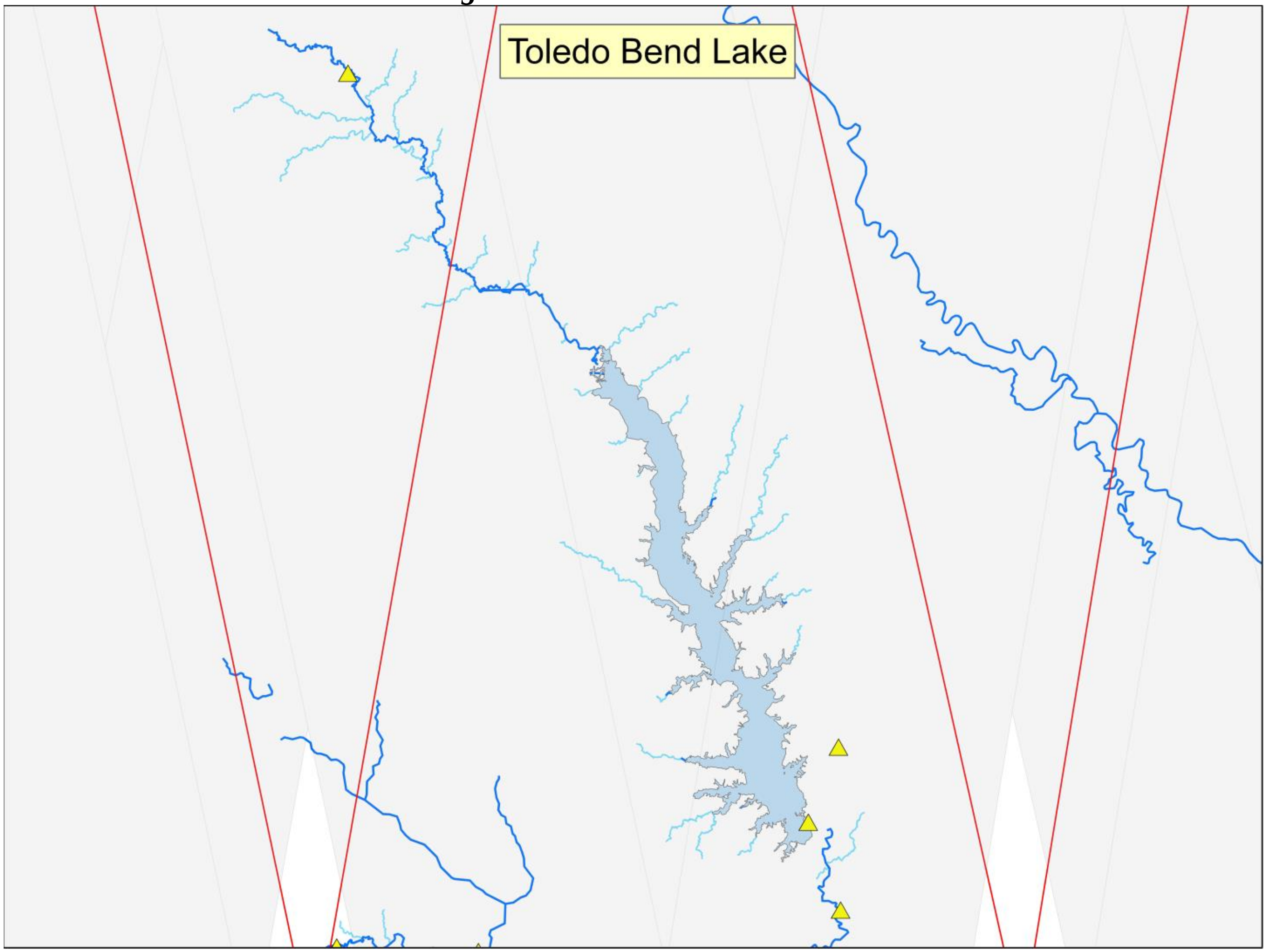


Figure A.23 Tuttle Creek Reservoir

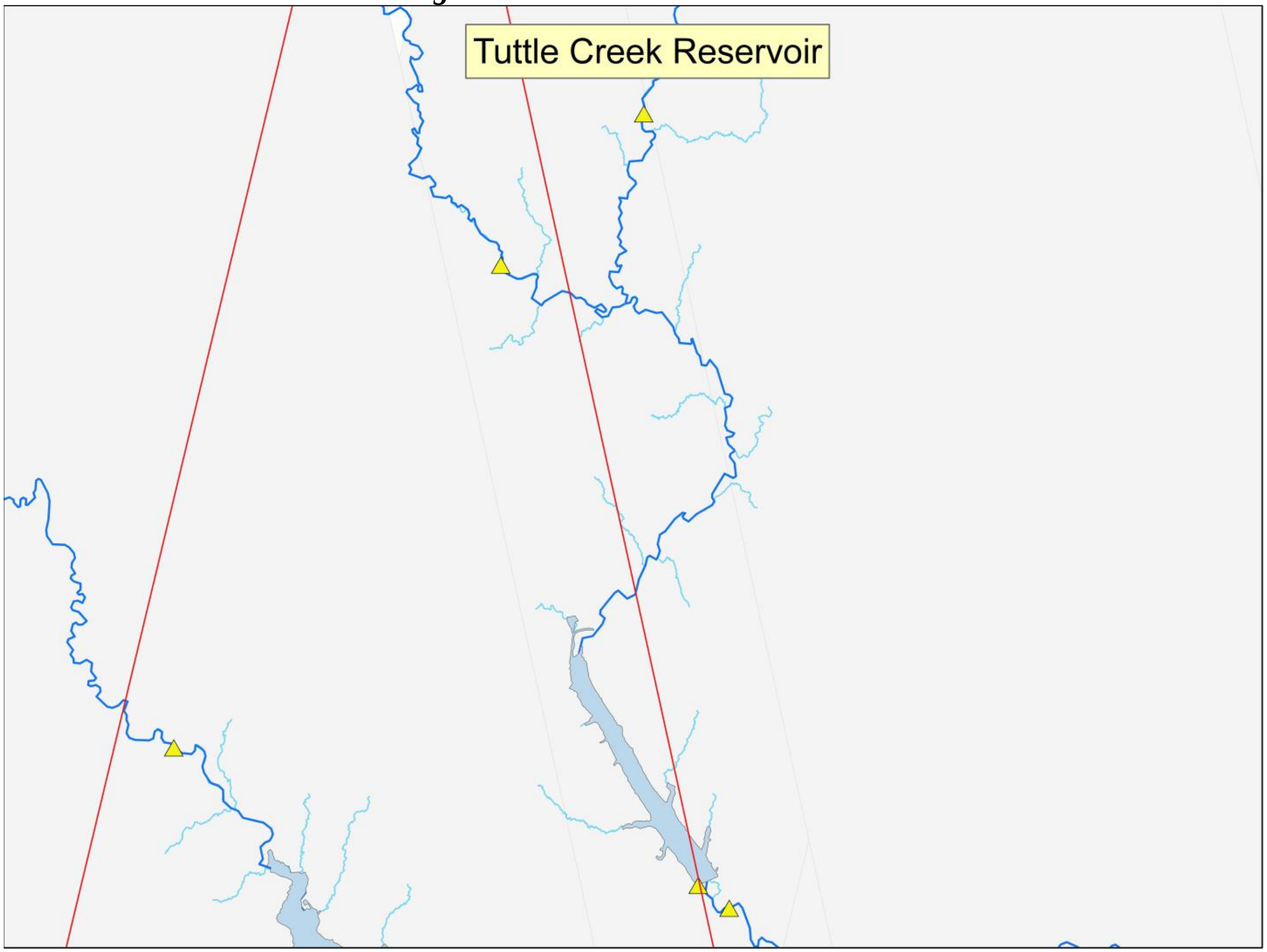


Figure A.24 Walter F George Reservoir

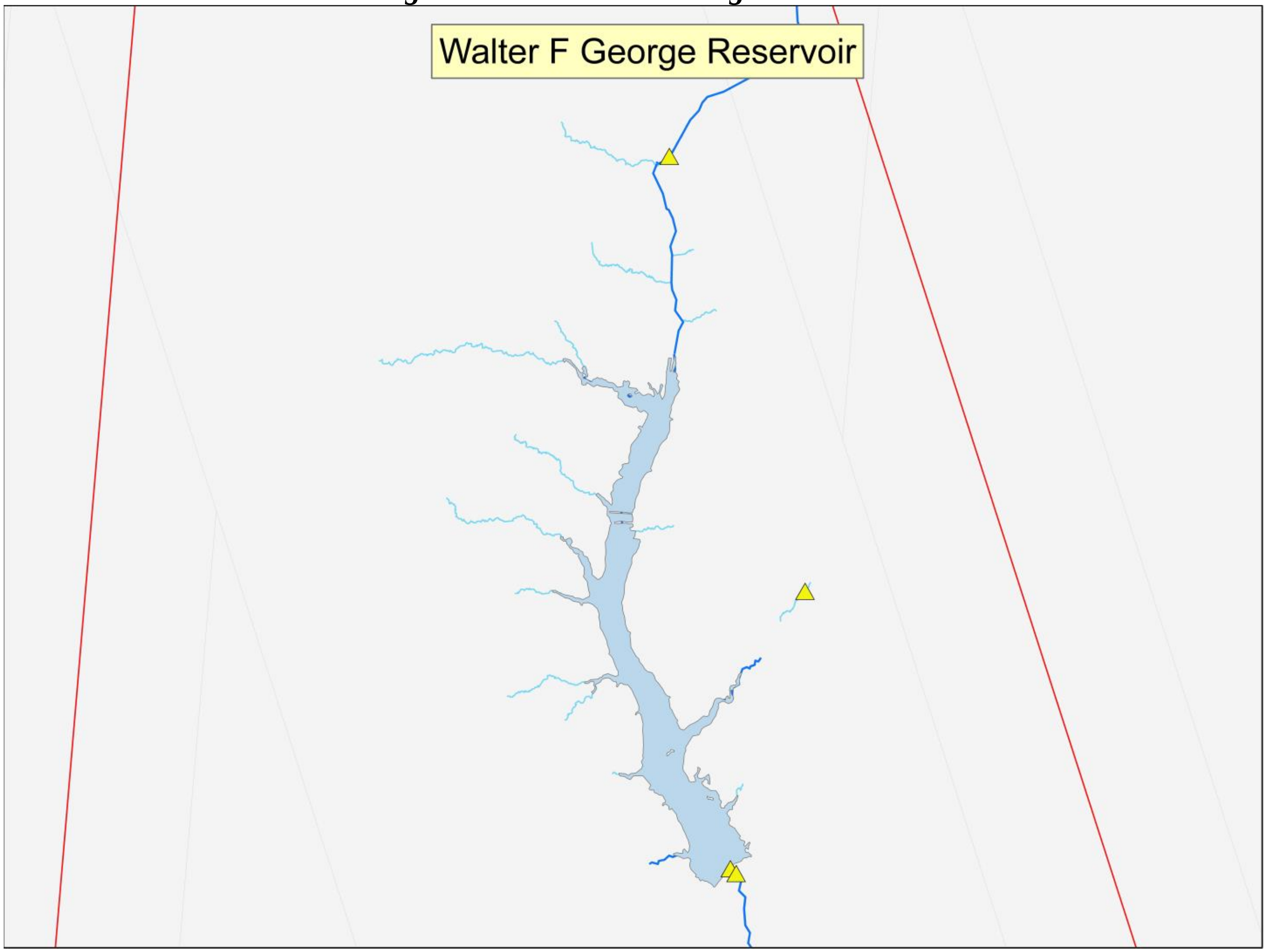


Figure A.25 West Point Lake

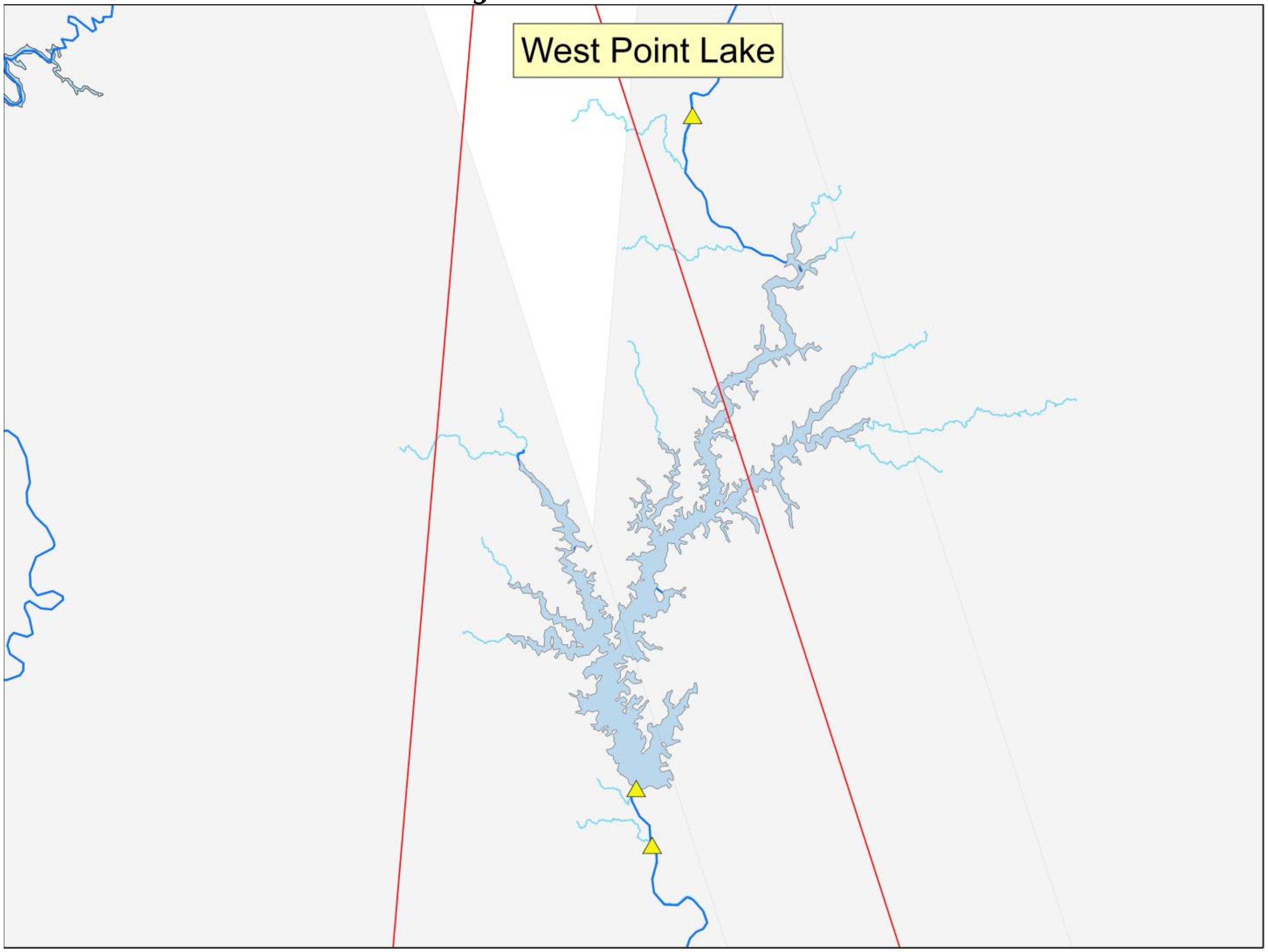
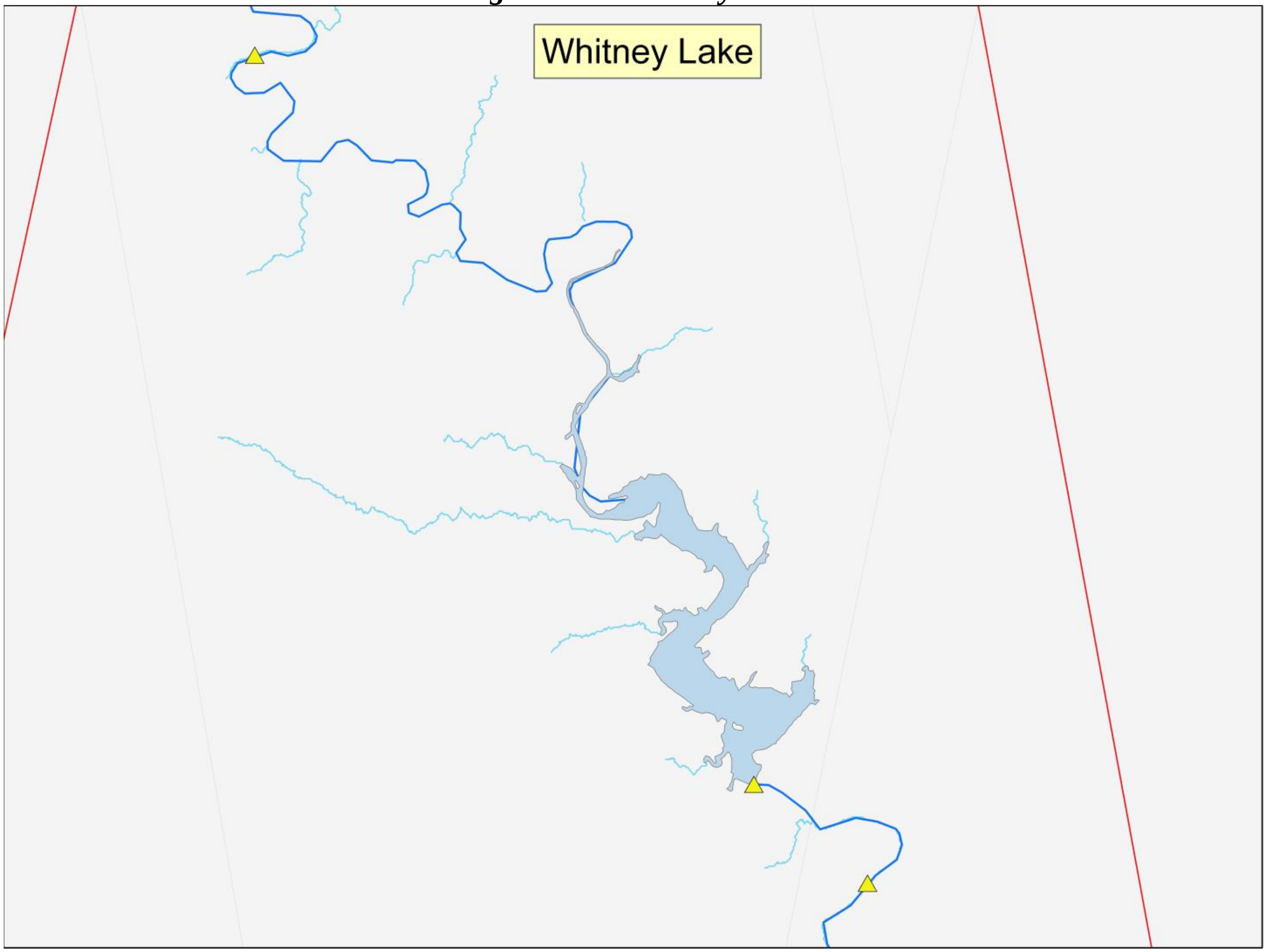


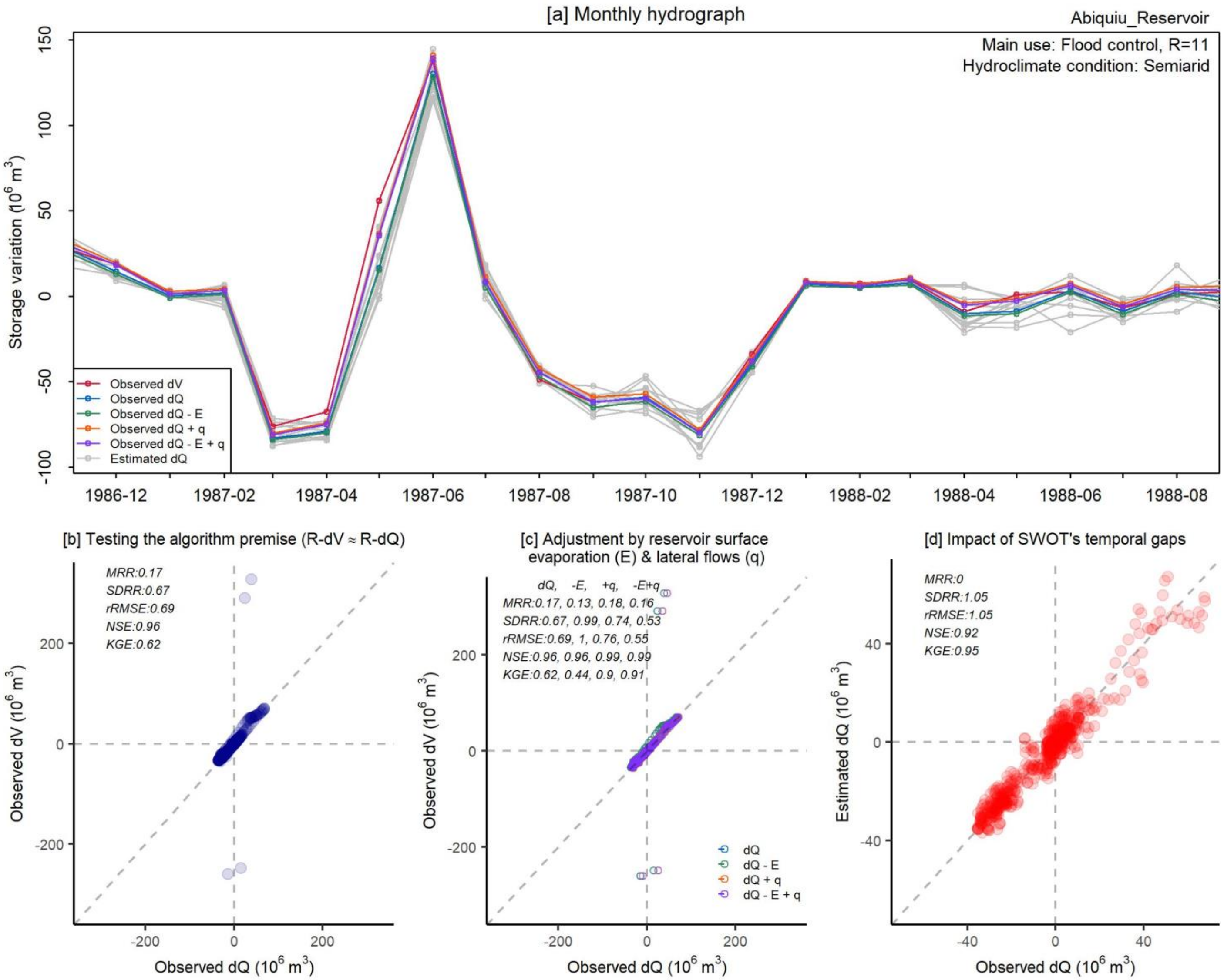
Figure A.26 Whitney Lake



Appendix B - Graphical Error Performance Measures

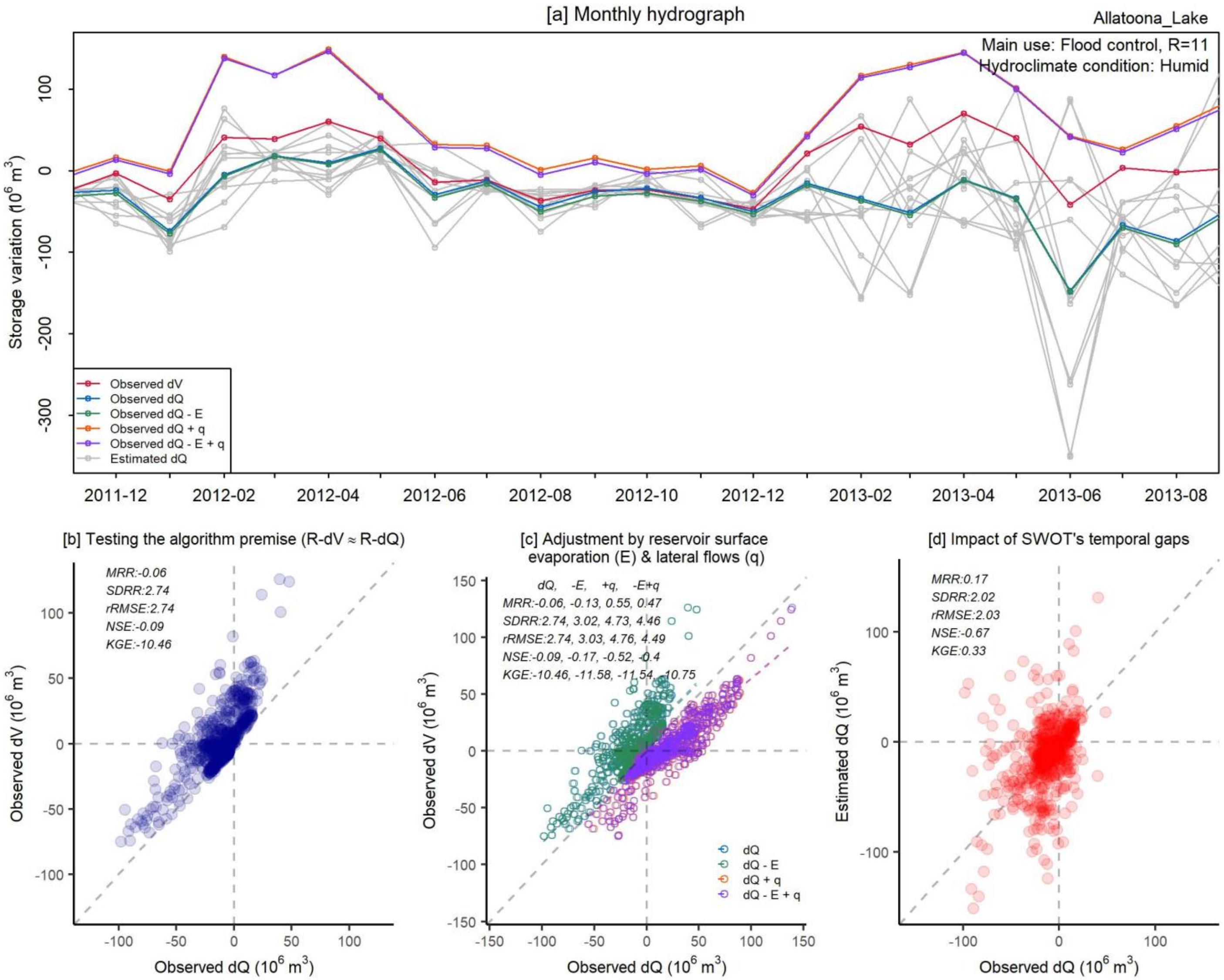
Appendix B.1 Monthly Hydrographs and Scatterplots

Figure B.1.1 Abiquiu Reservoir



* dQ: discharge difference between SWOT-observable inflow and outflow reaches, dV: reservoir storage variation
** E: reservoir surface evaporation, q: lateral flows. Significance level=0.05

Figure B.1.2 Allatoona Reservoir



* dQ: discharge difference between SWOT-observable inflow and outflow reaches, dV: reservoir storage variation
** E: reservoir surface evaporation, q: lateral flows. Significance level=0.05

Figure B.1.3 Barlett and Horseshoe Reservoir

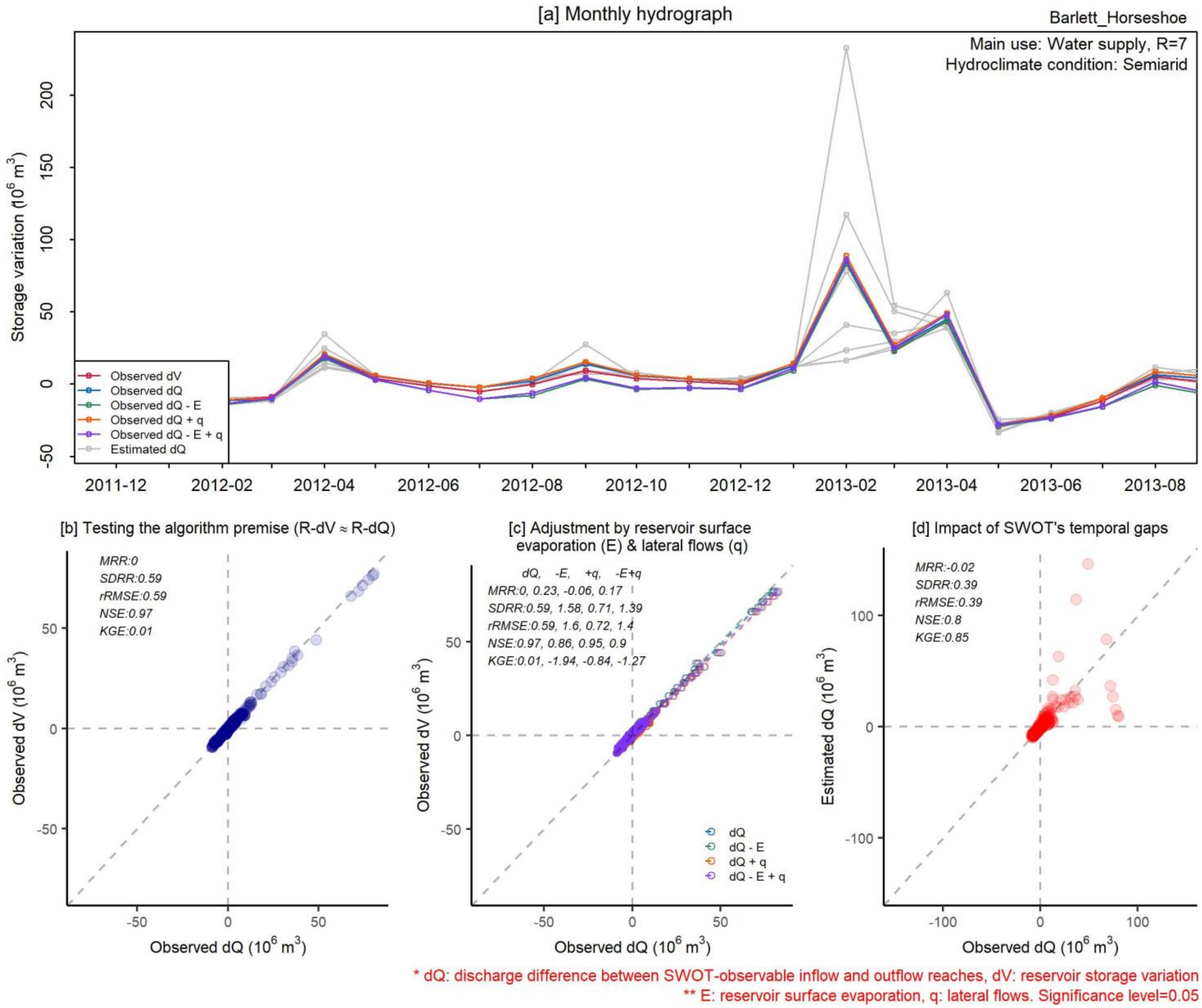


Figure B.1.4 Brantly Lake

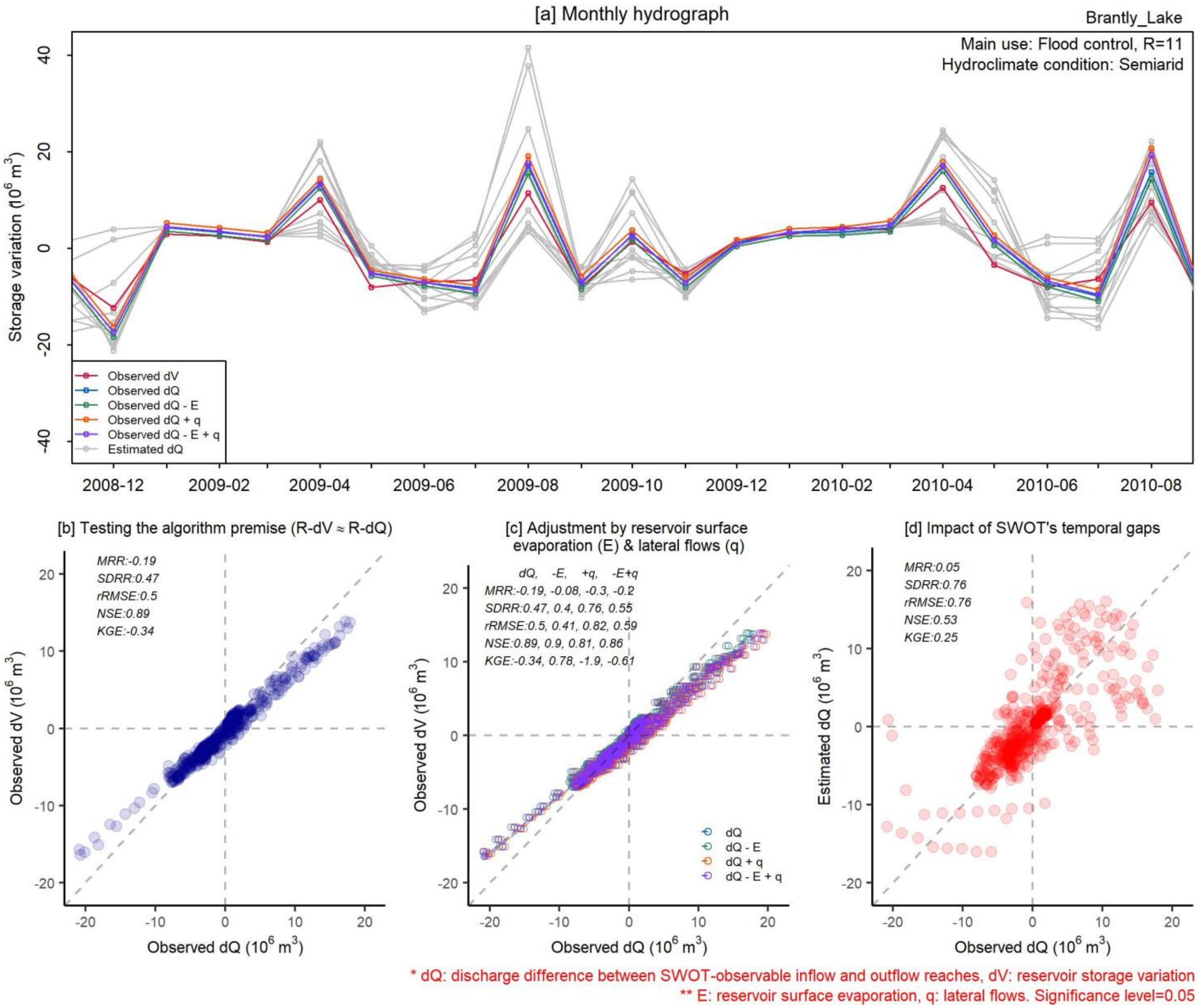


Figure B.1.5 CJ Strike Reservoir

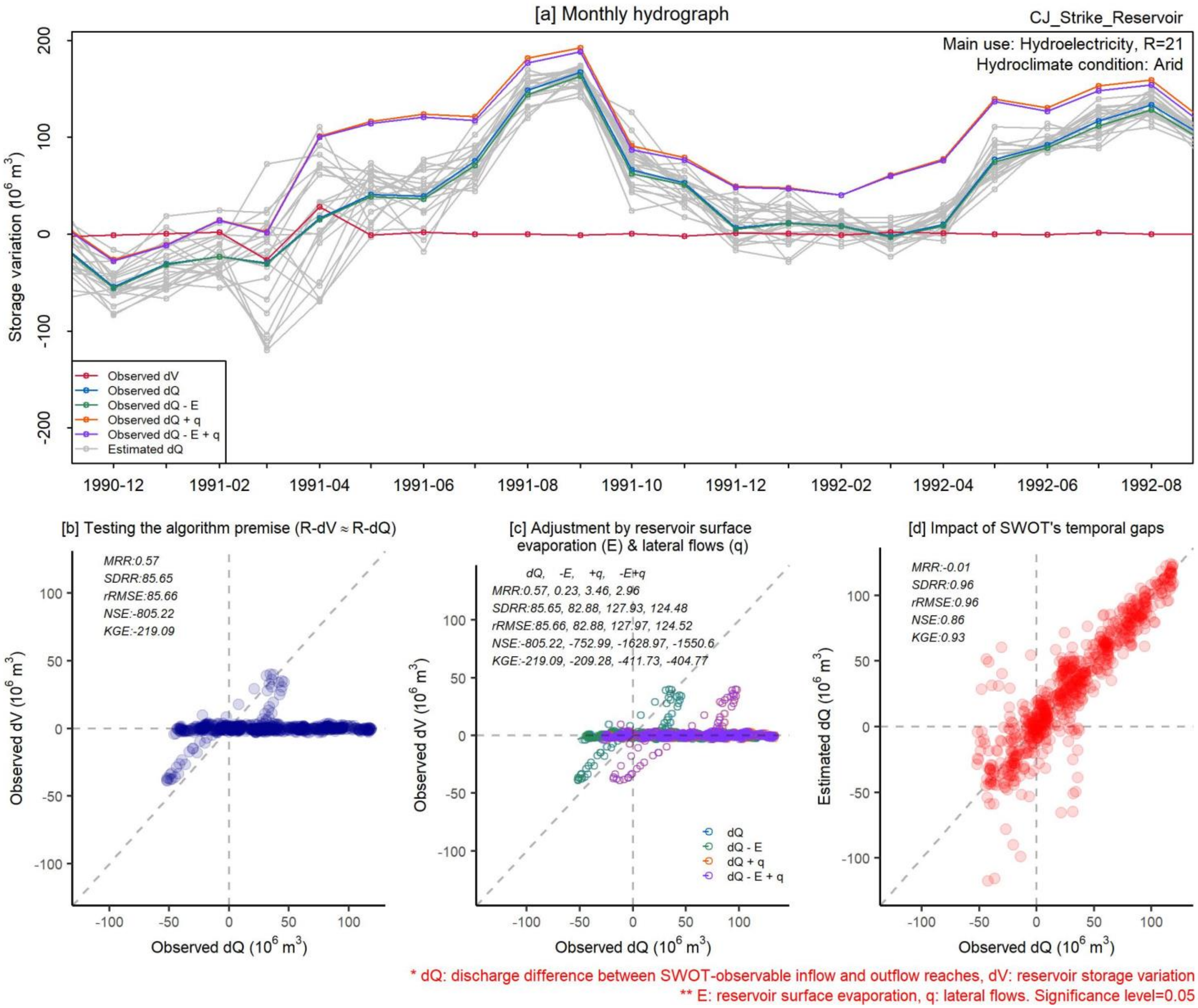


Figure B.1.6 Cochiti Lake

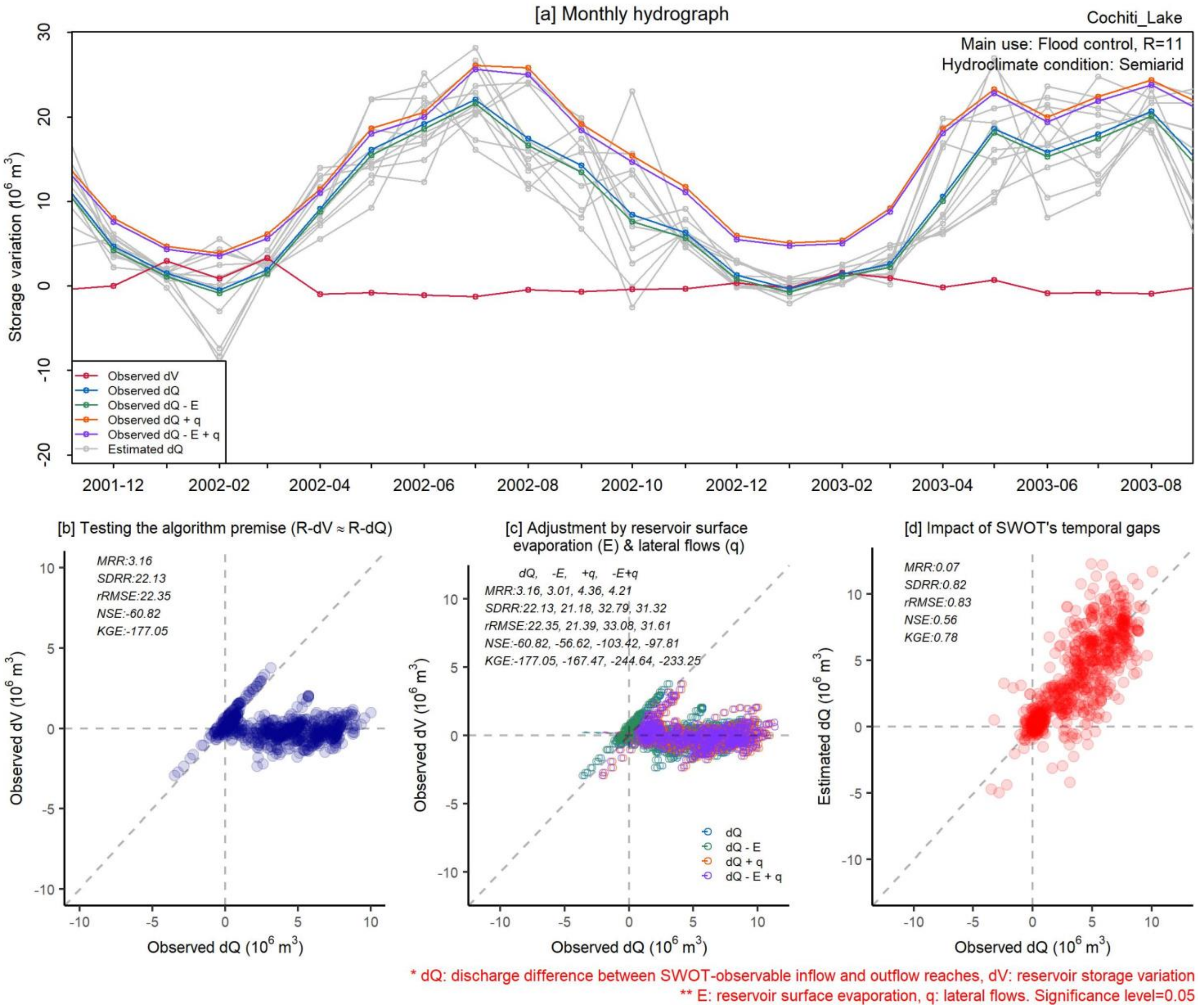
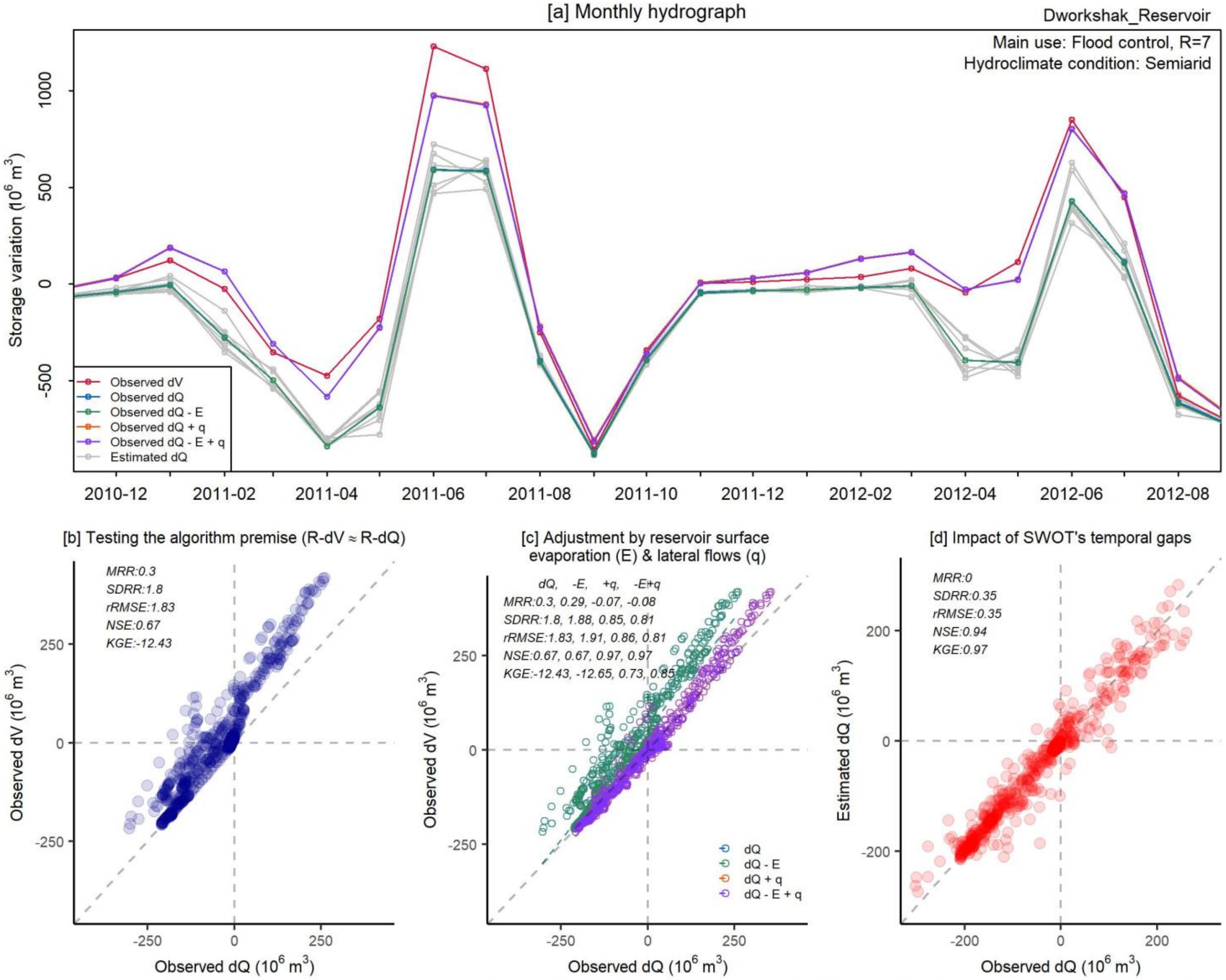
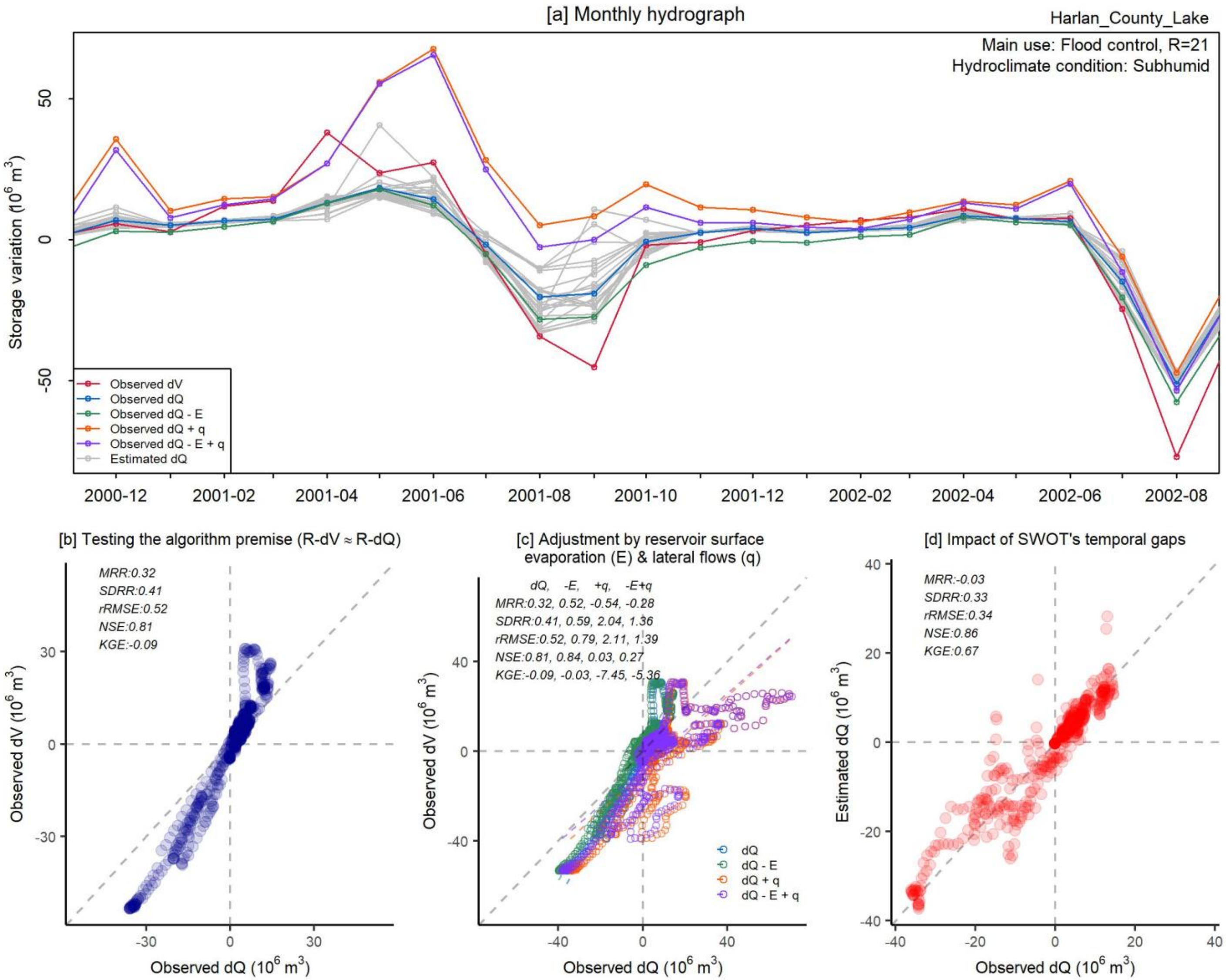


Figure B.1.7 Dworshak Reservoir



* dQ: discharge difference between SWOT-observable inflow and outflow reaches, dV: reservoir storage variation
** E: reservoir surface evaporation, q: lateral flows. Significance level=0.05

Figure B.1.8 Harlan County Lake



* dQ: discharge difference between SWOT-observable inflow and outflow reaches, dV: reservoir storage variation
** E: reservoir surface evaporation, q: lateral flows. Significance level=0.05

Figure B.1.9 John Martin Reservoir

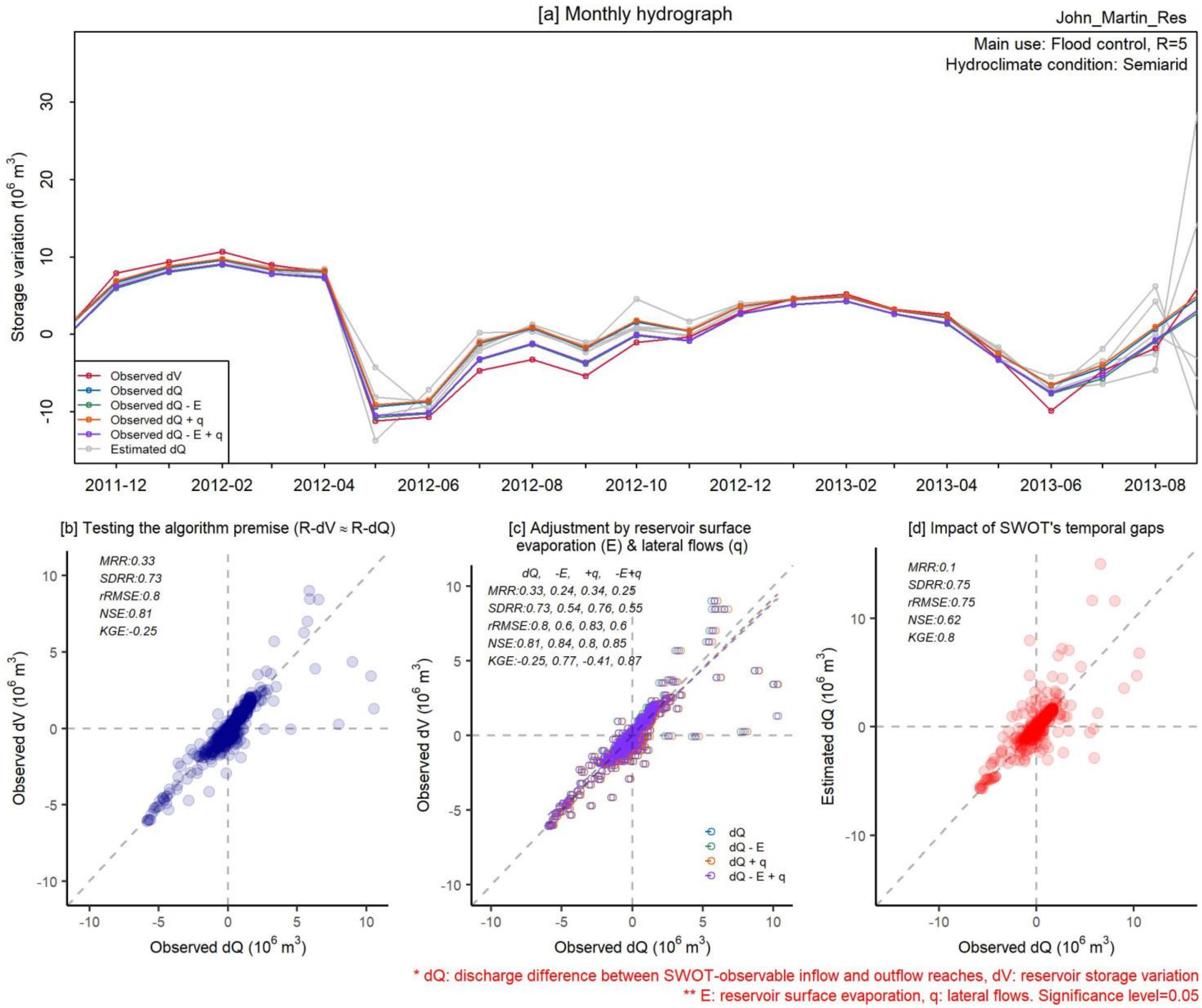


Figure B.1.10 Lake Granbury

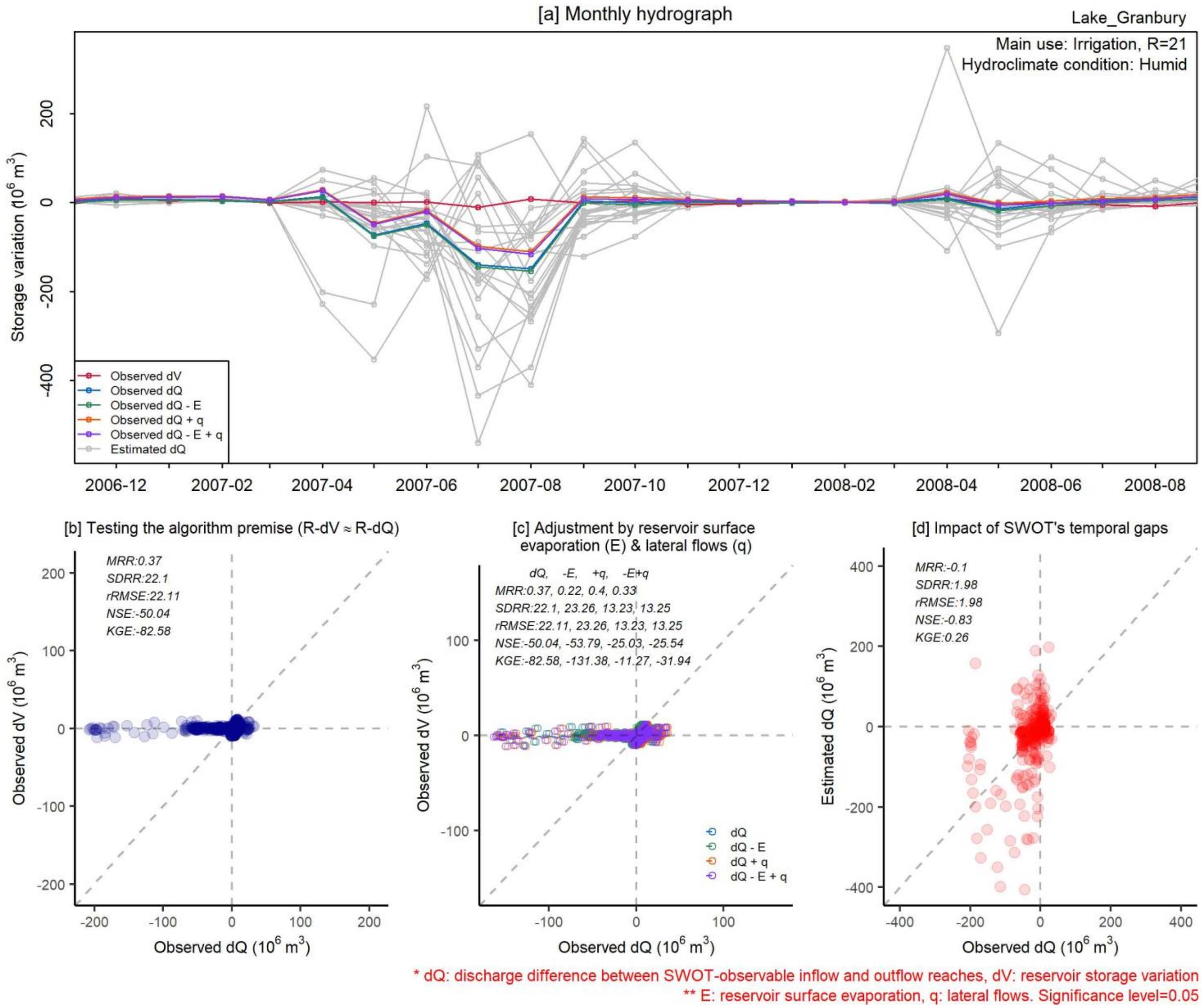


Figure B.1.11 Lake Havasu

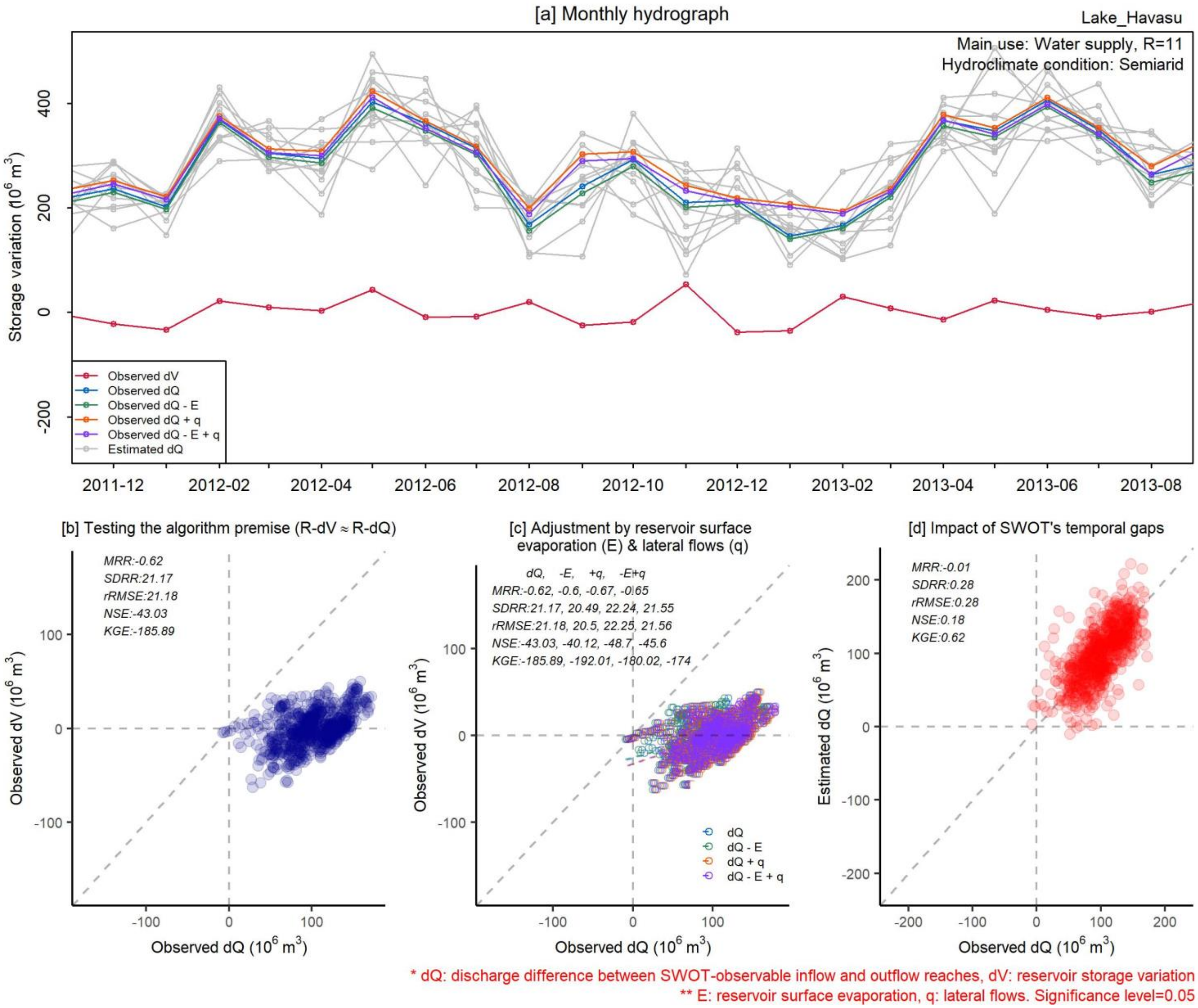
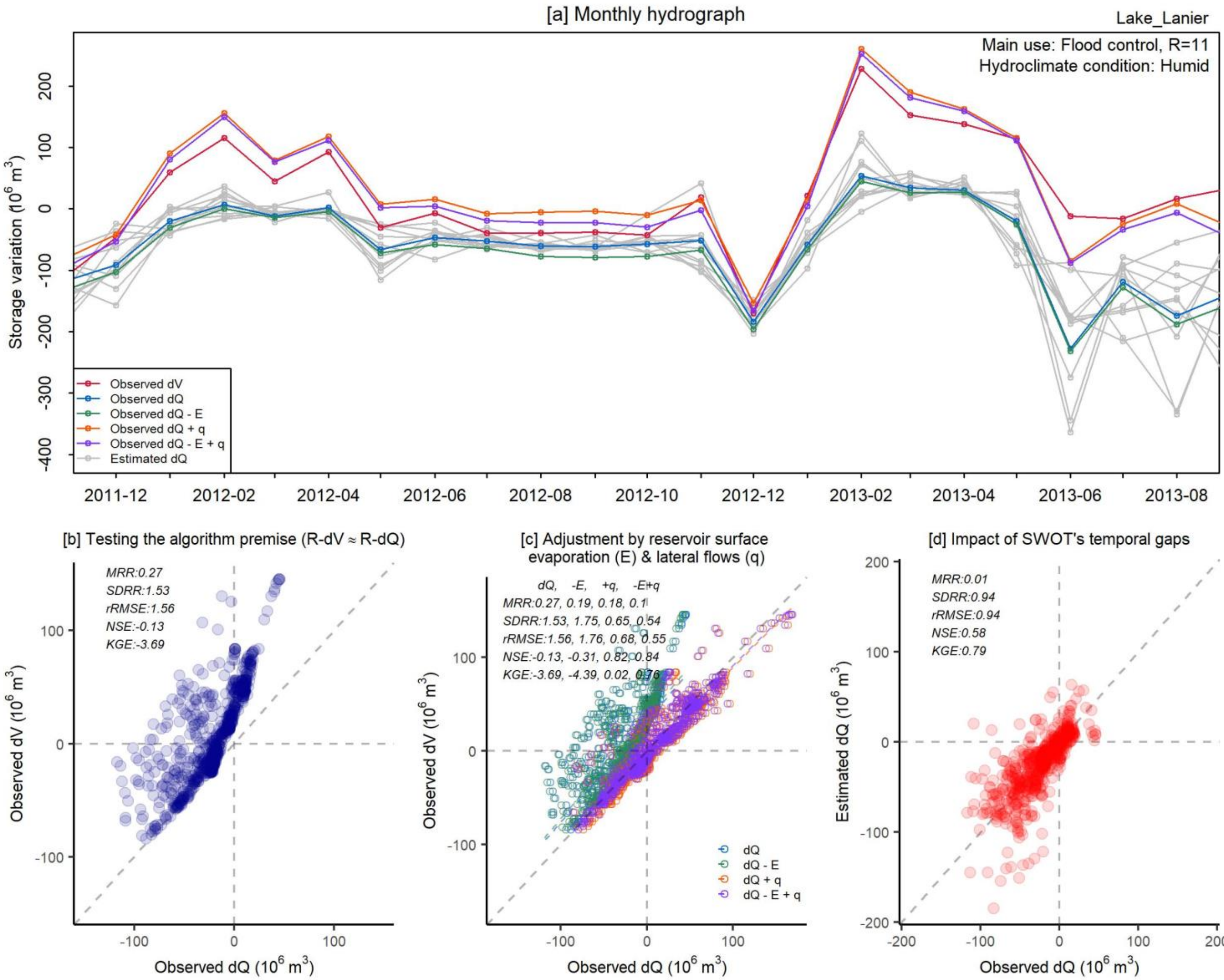


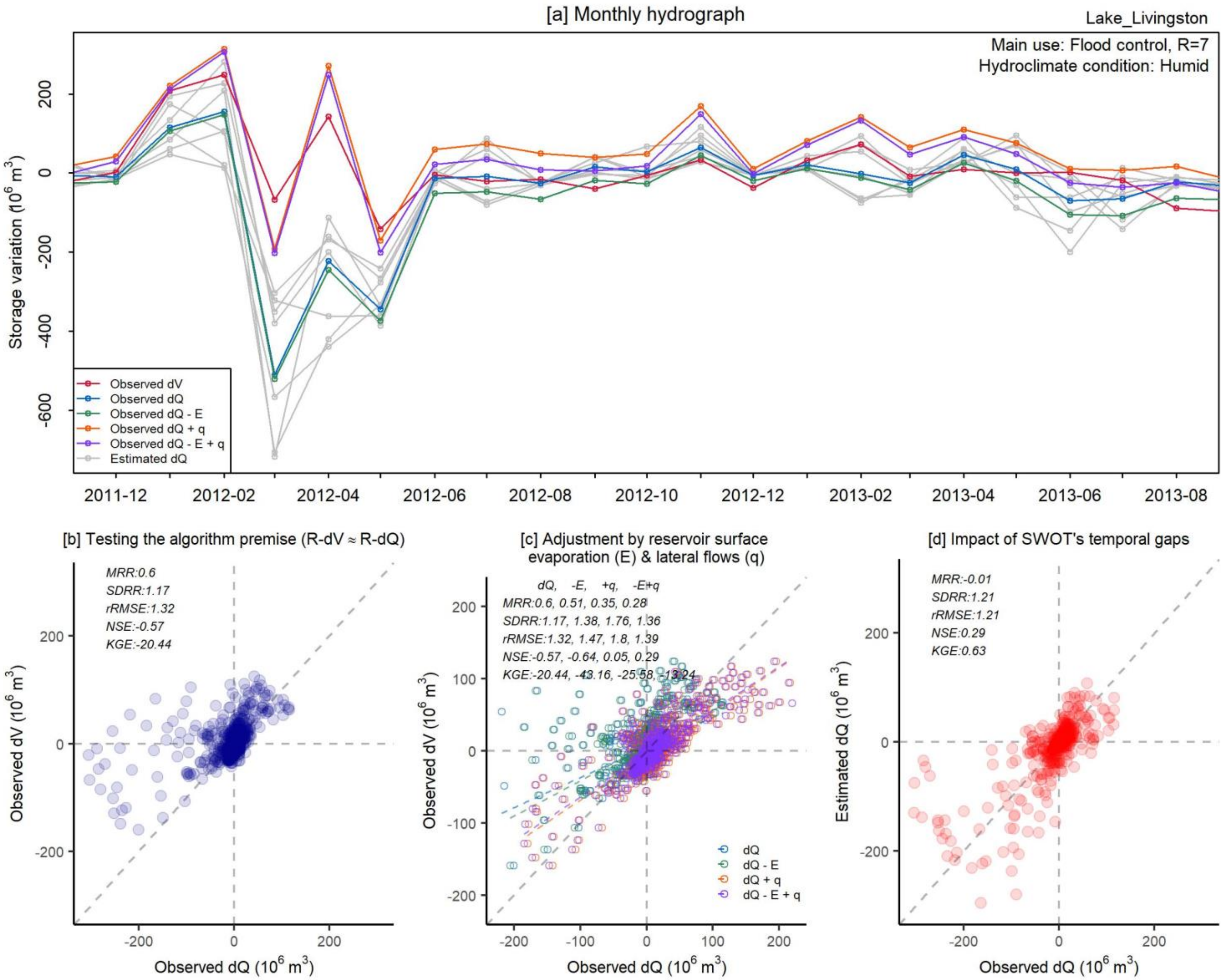
Figure B.1.12 Lake Lanier



* dQ: discharge difference between SWOT-observable inflow and outflow reaches, dV: reservoir storage variation

** E: reservoir surface evaporation, q: lateral flows. Significance level=0.05

Figure B.1.13 Lake Livingston



* dQ: discharge difference between SWOT-observable inflow and outflow reaches, dV: reservoir storage variation
** E: reservoir surface evaporation, q: lateral flows. Significance level=0.05

Figure B.1.14 Lake Mohave

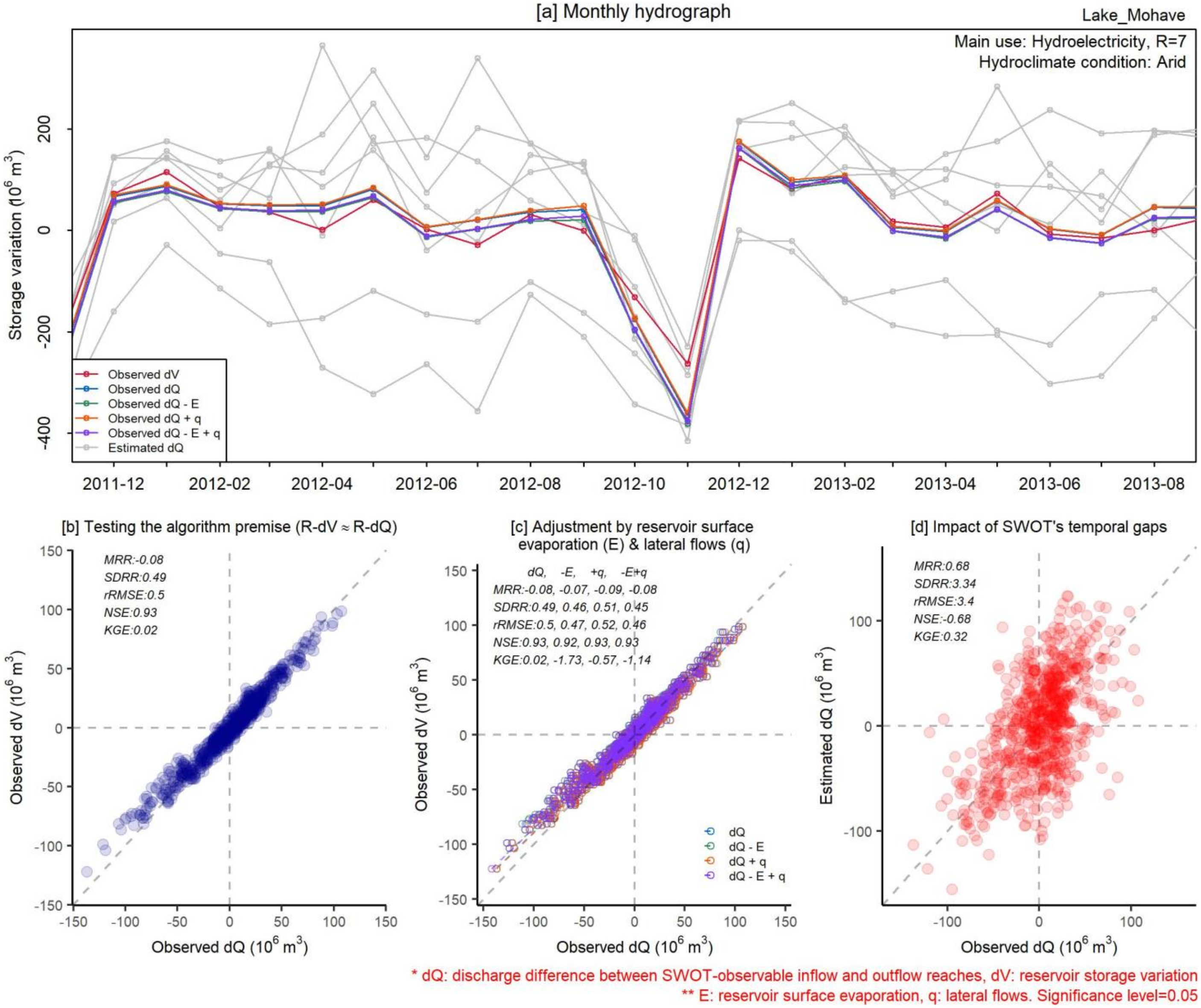


Figure B.1.15 Lake Red Rock

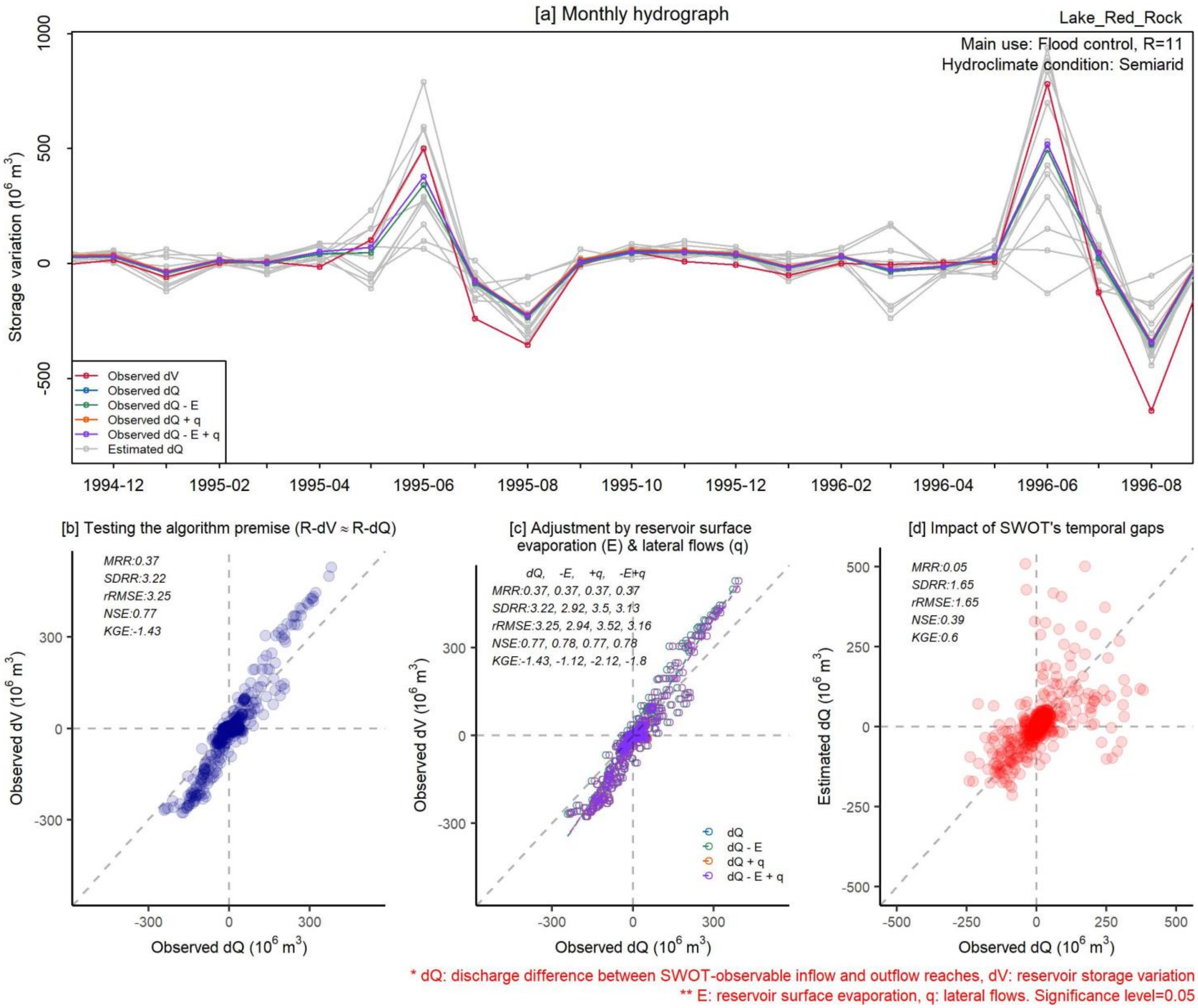


Figure B.1.16 Lake Seminole

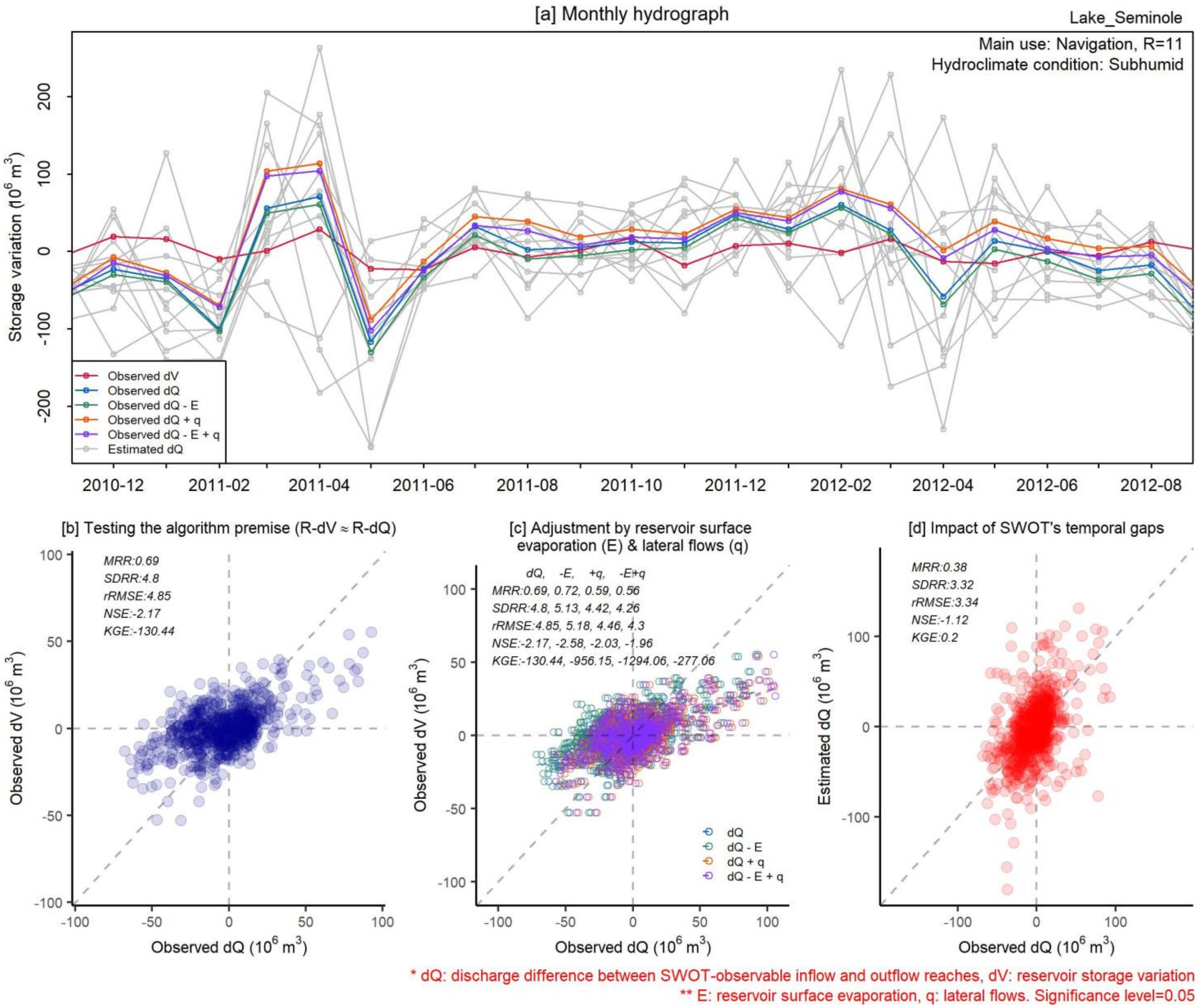


Figure B.1.17 Lake Taxoma

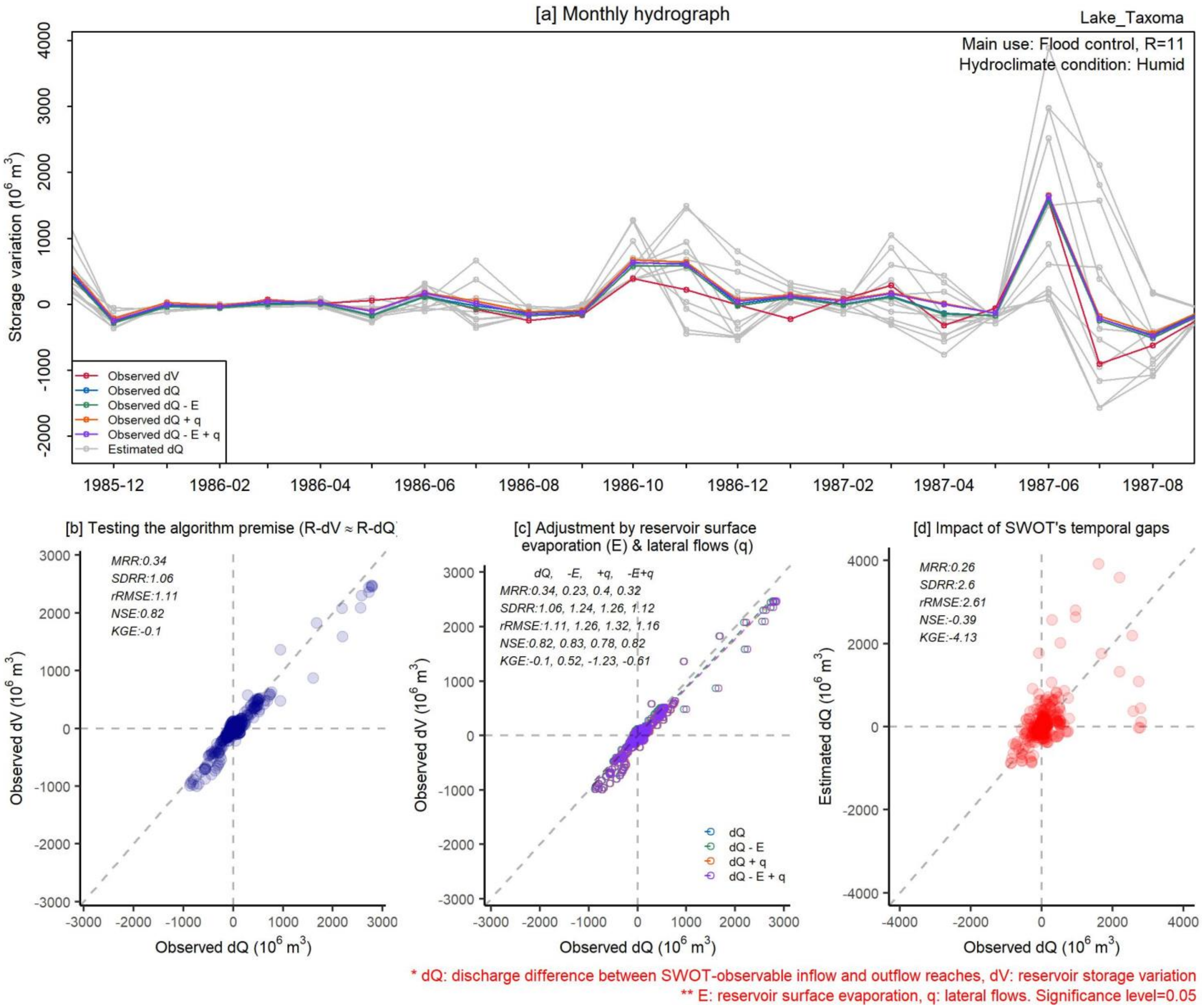


Figure B.1.18 Milford Lake

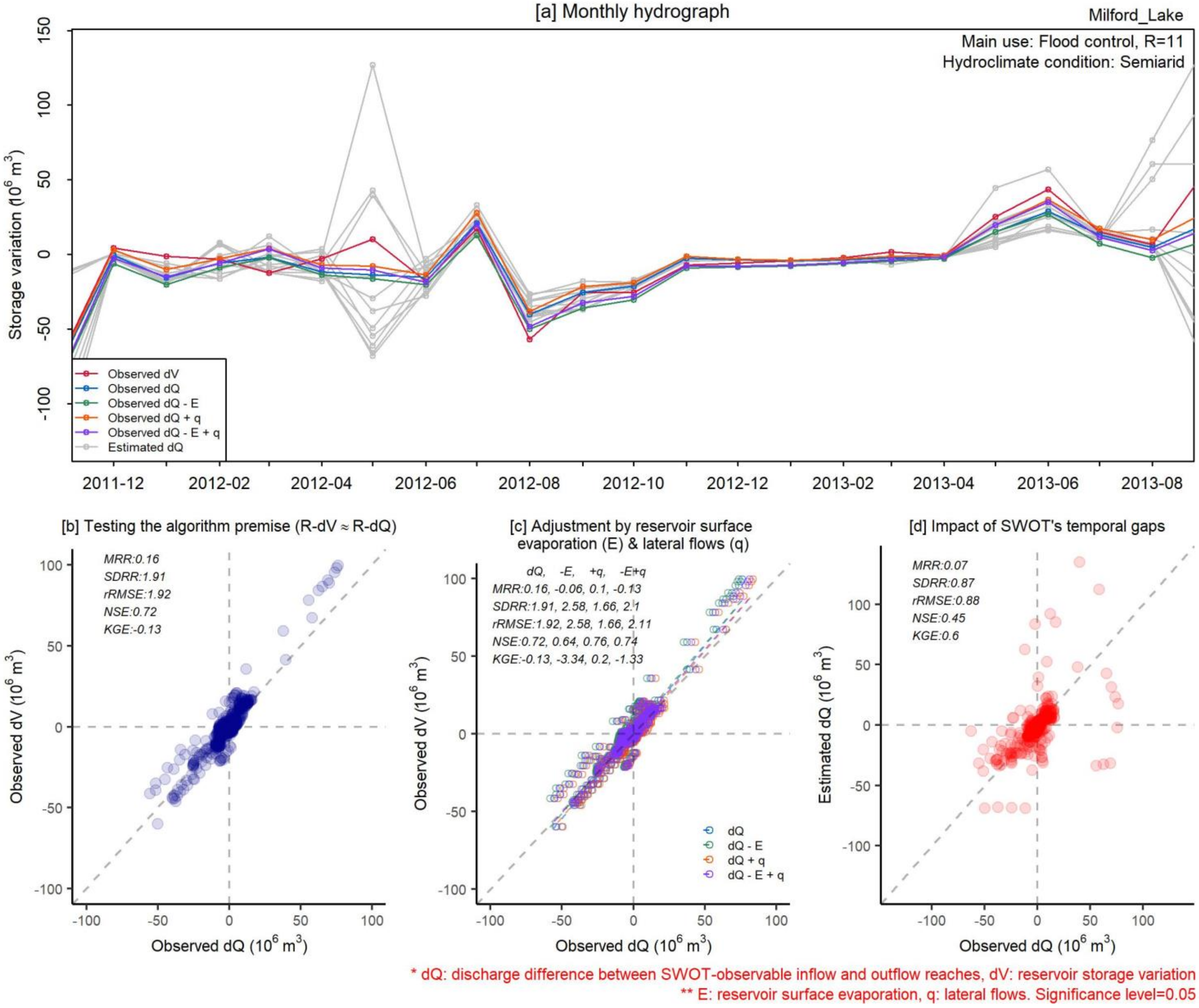
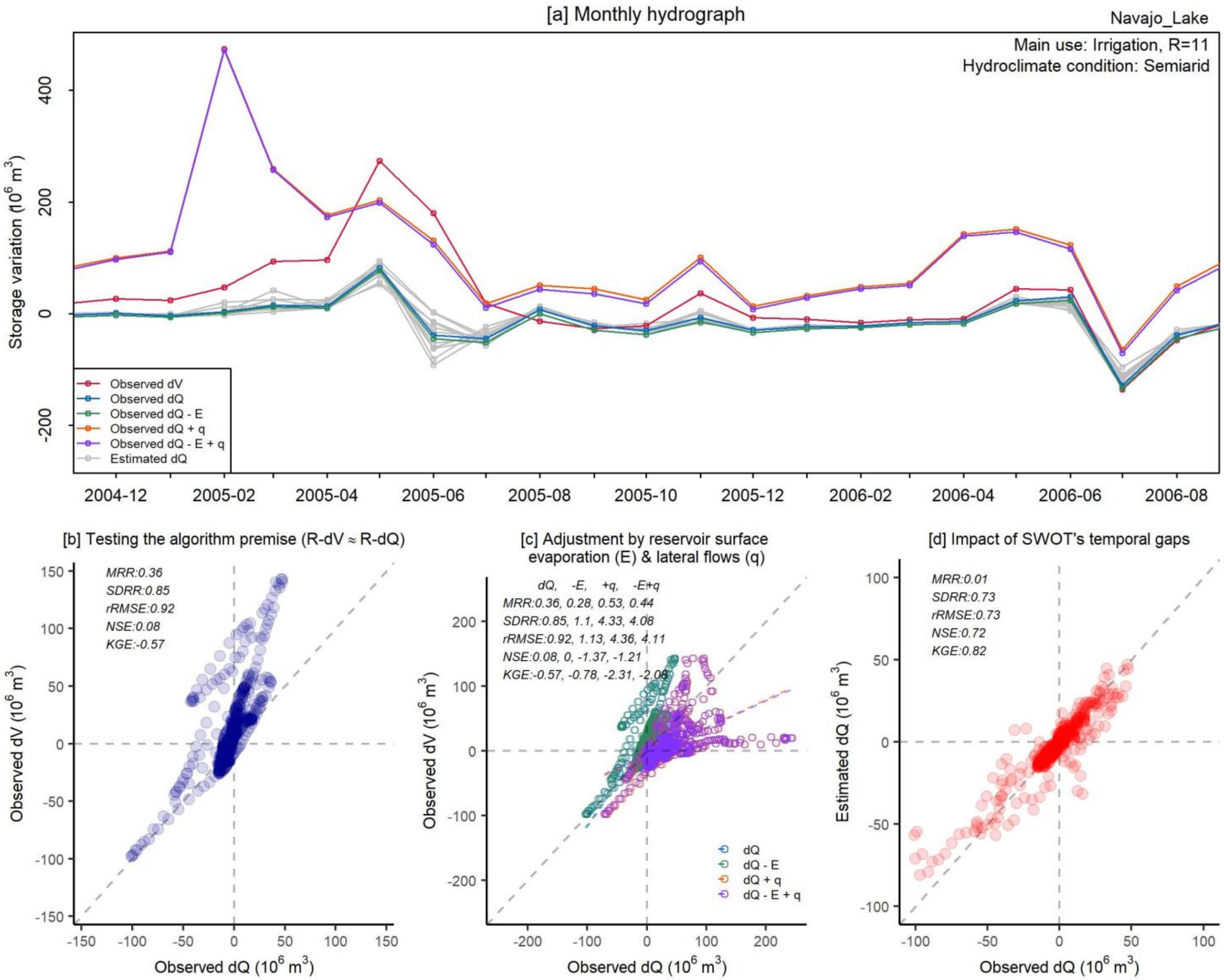
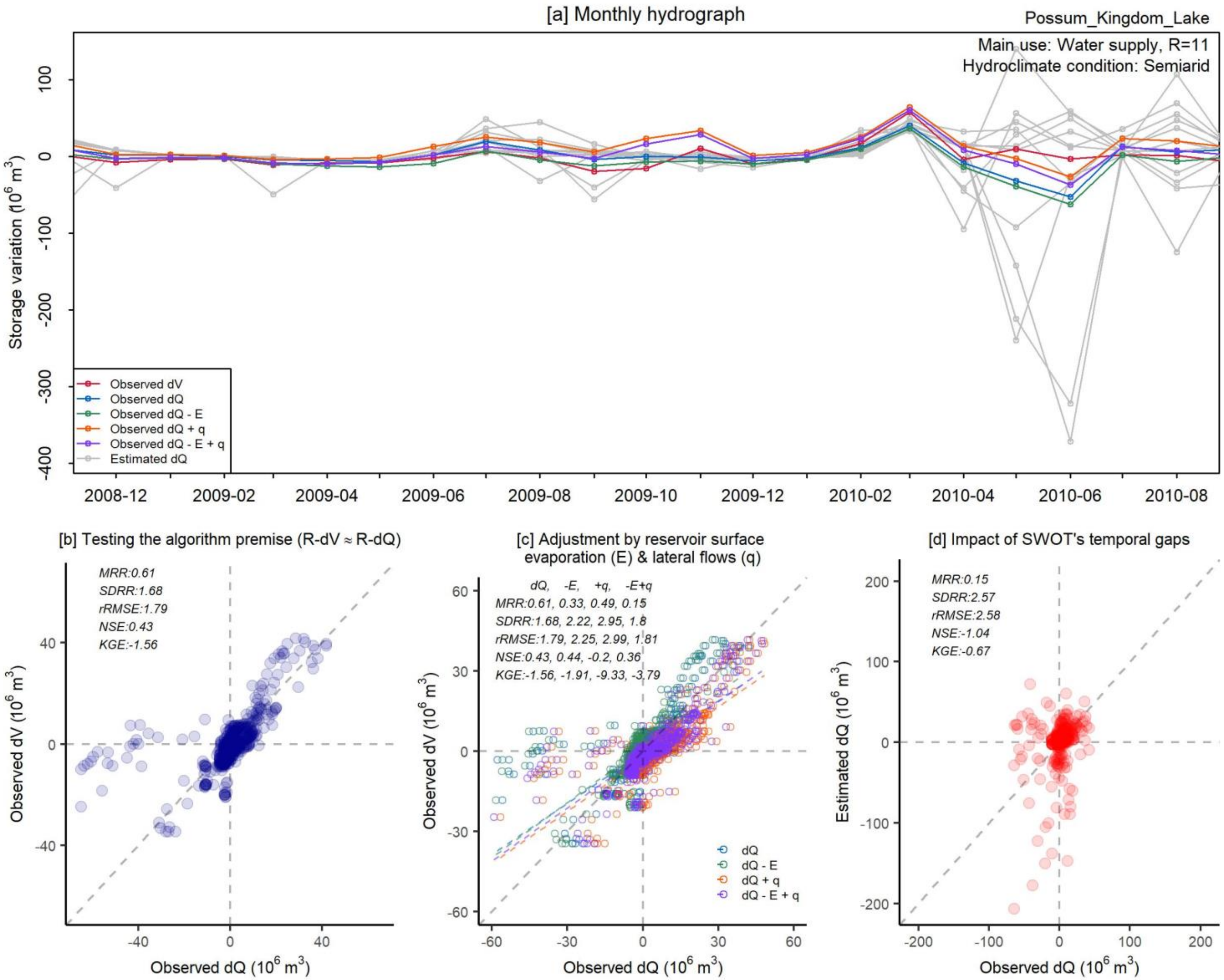


Figure B.1.19 Navajo Lake



* dQ: discharge difference between SWOT-observable inflow and outflow reaches, dV: reservoir storage variation
** E: reservoir surface evaporation, q: lateral flows. Significance level=0.05

Figure B.1.20 Possum Kingdom Lake



* dQ: discharge difference between SWOT-observable inflow and outflow reaches, dV: reservoir storage variation
** E: reservoir surface evaporation, q: lateral flows. Significance level=0.05

Figure B.1.21 Shasta Lake

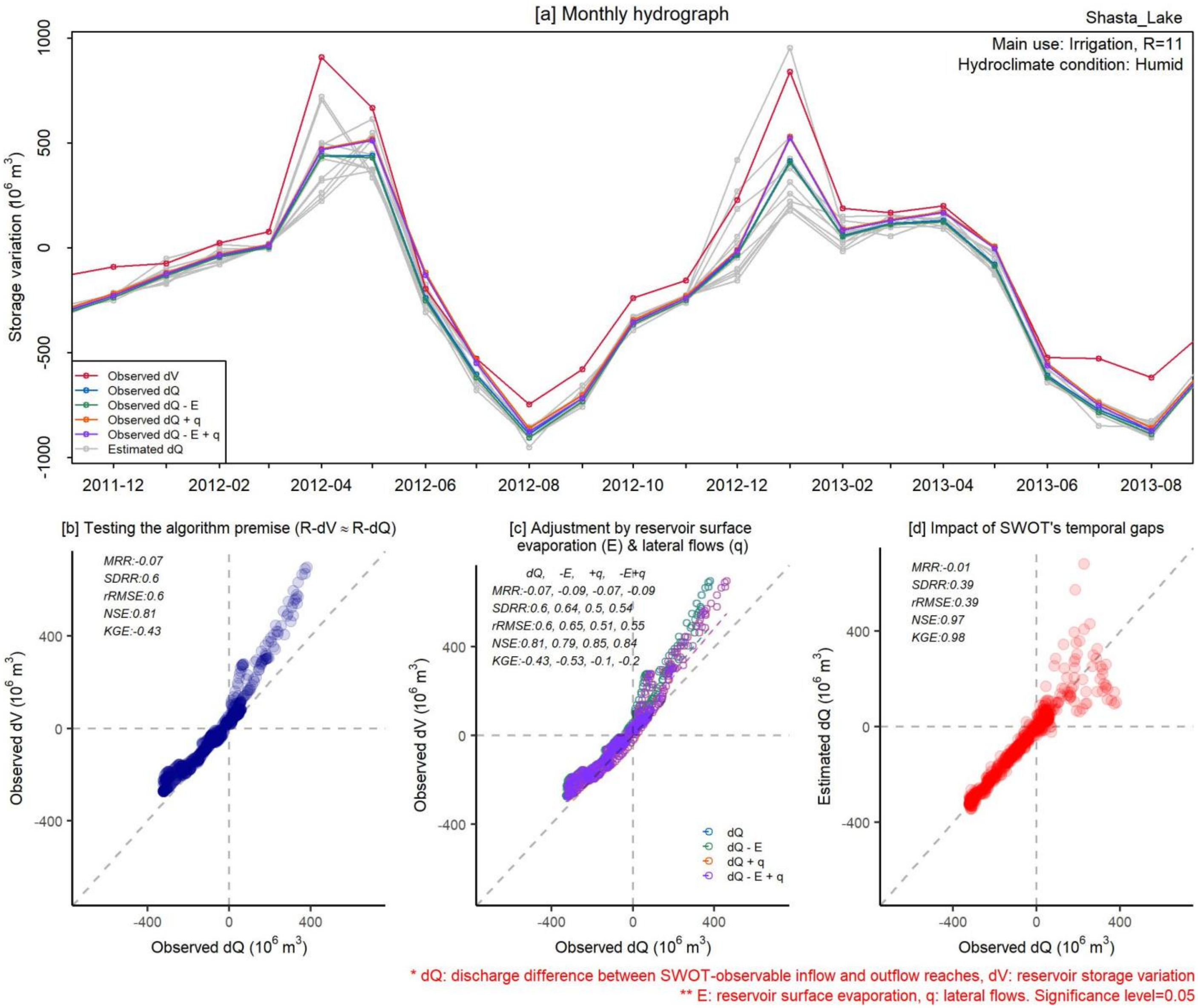


Figure B.1.22 Toledo Bend Lake

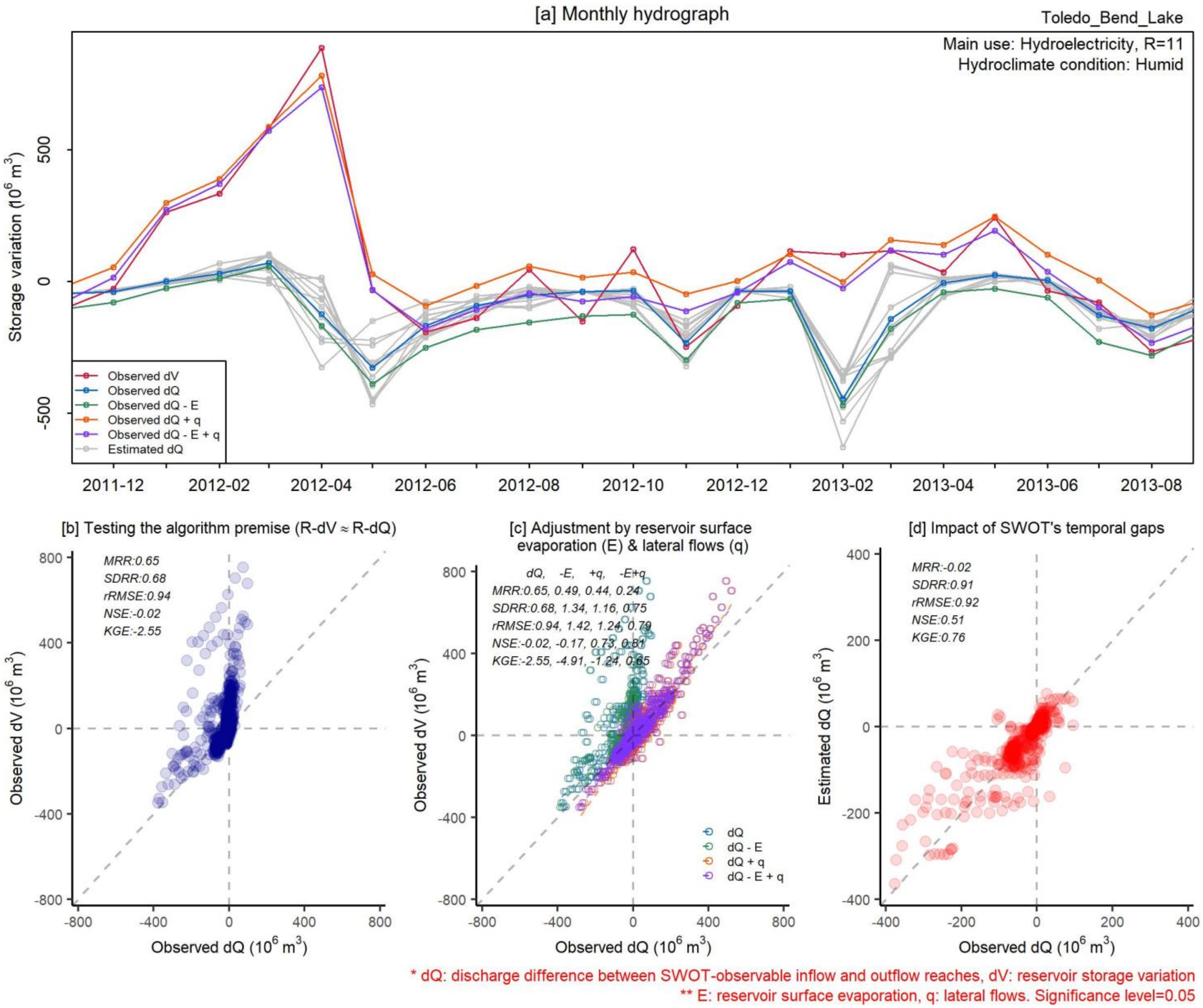


Figure B.1.23 Tuttle Creek Reservoir

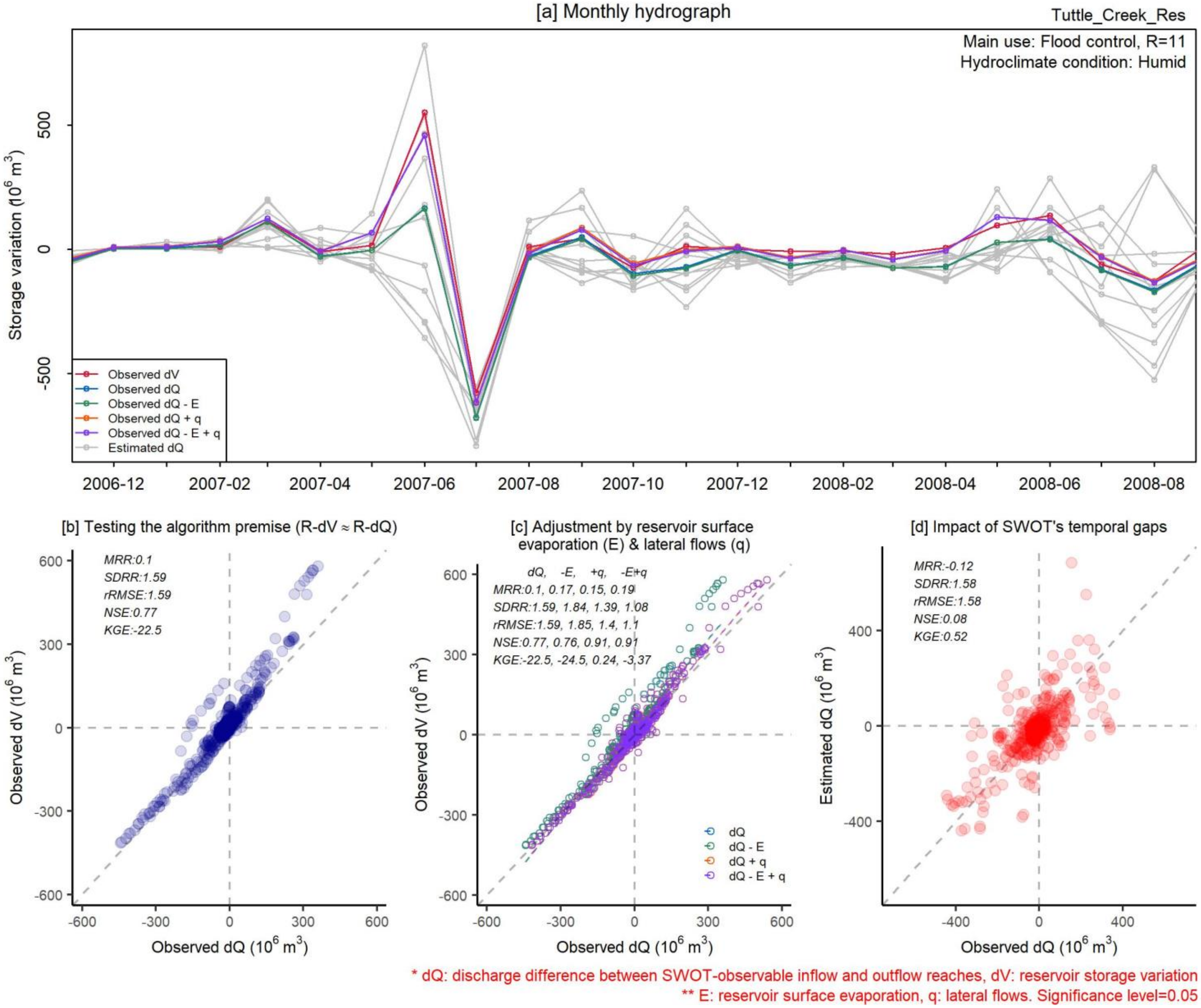
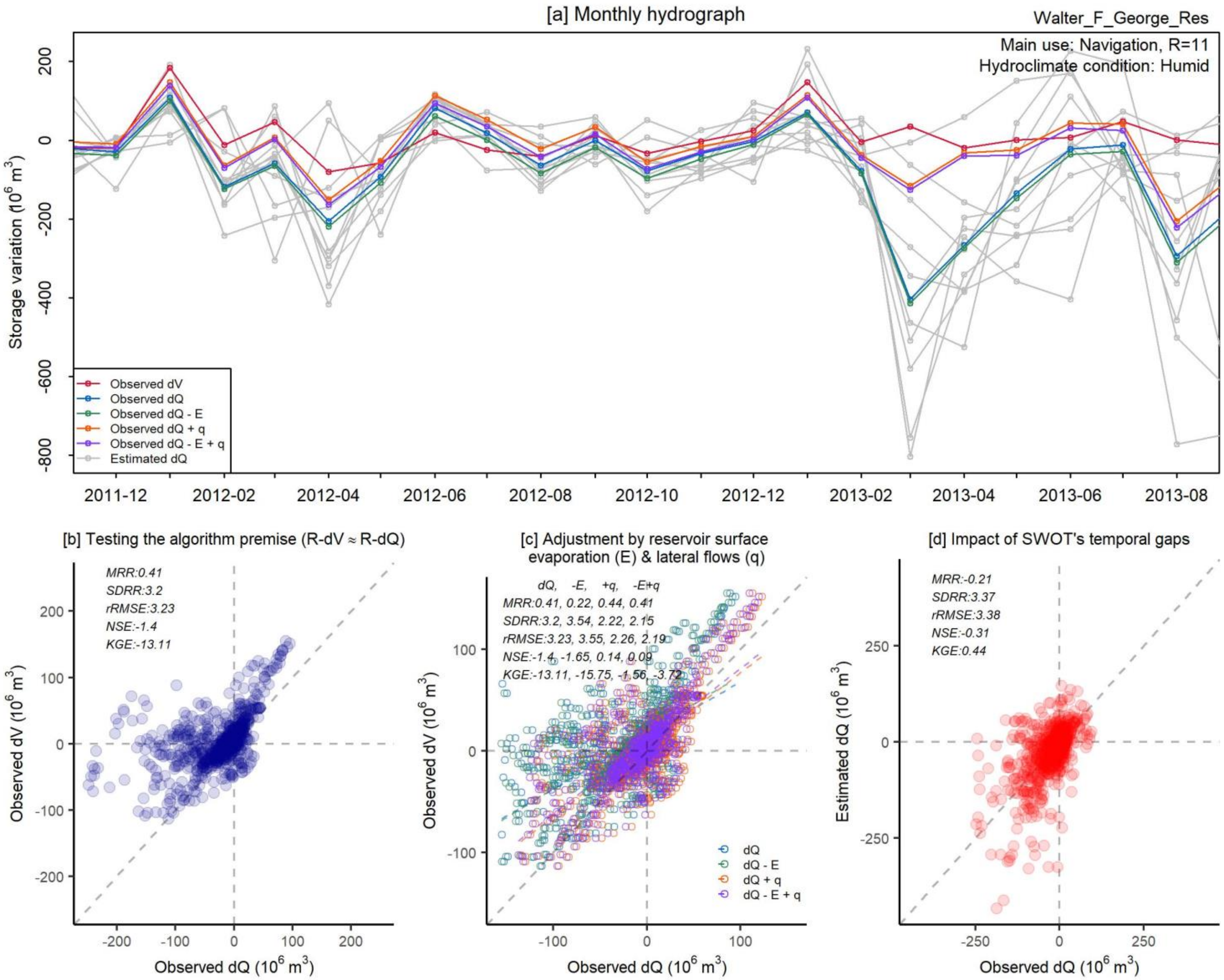


Figure B.1.24 Walter F George Reservoir



* dQ: discharge difference between SWOT-observable inflow and outflow reaches, dV: reservoir storage variation

** E: reservoir surface evaporation, q: lateral flows. Significance level=0.05

Figure B.1.25 West Point Lake

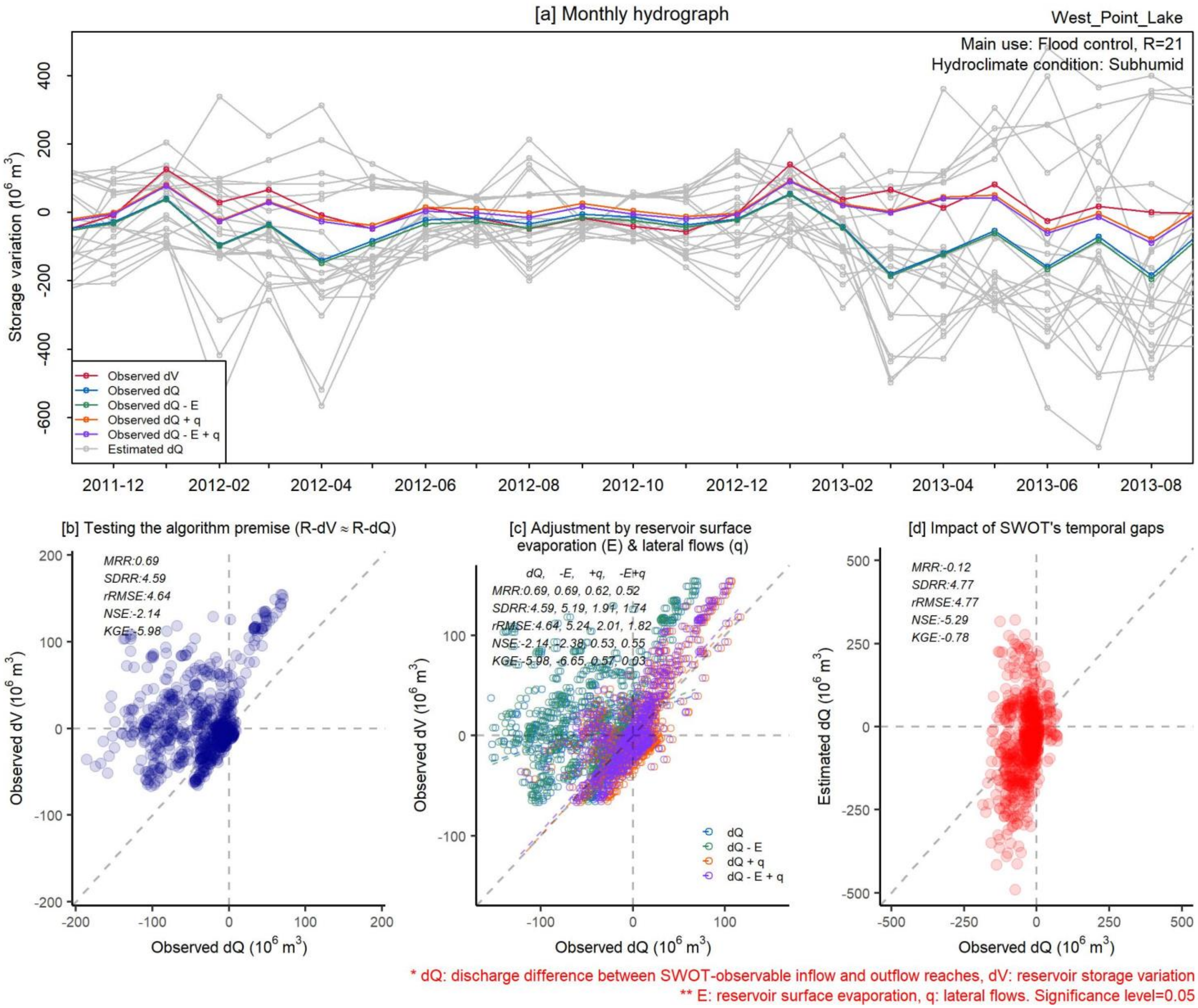
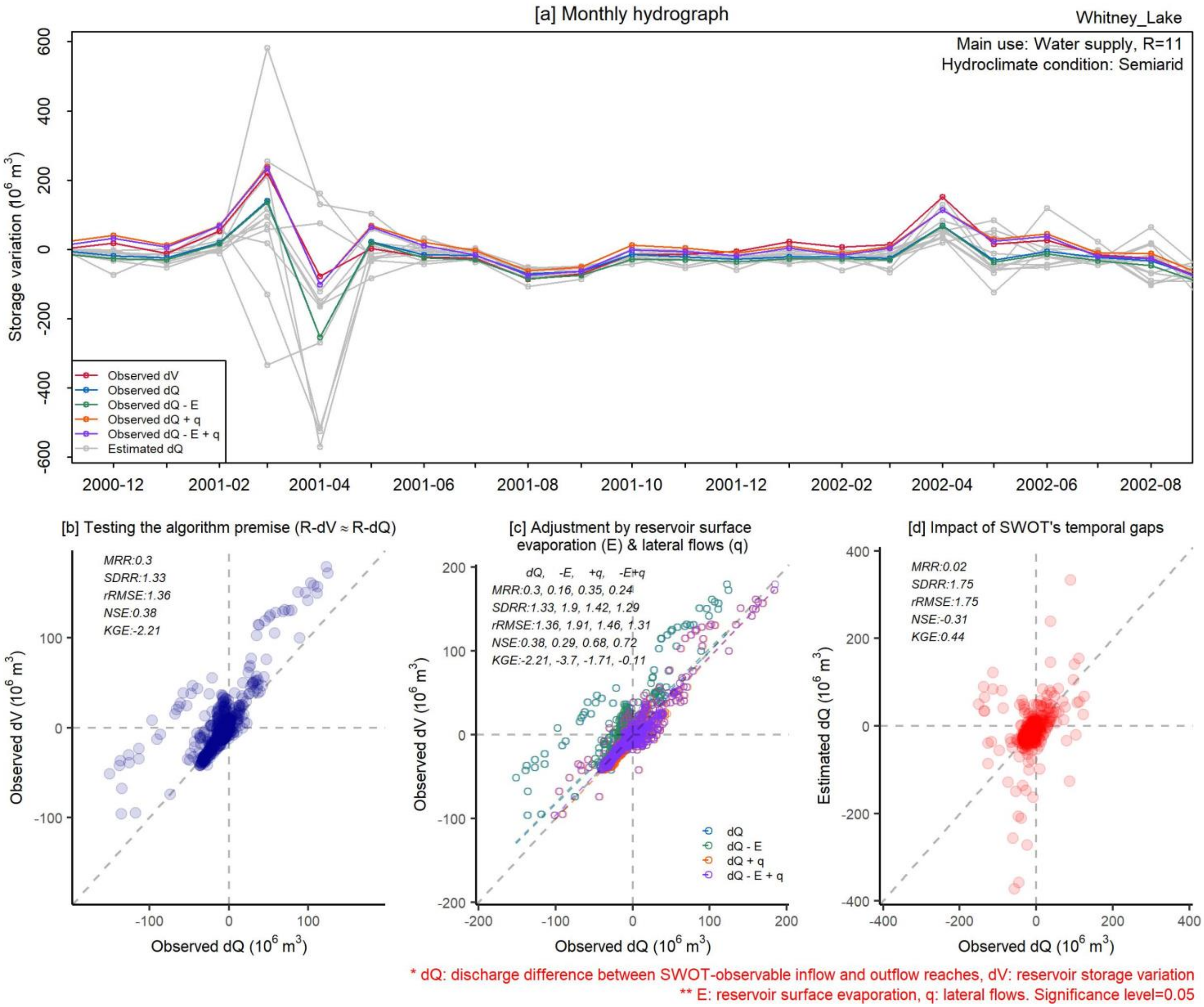


Figure B.1.26 Whitney Lake



Appendix B.2 Overlapping Index and Probability Distribution

Figure B.2.1 Abiquiu Reservoir

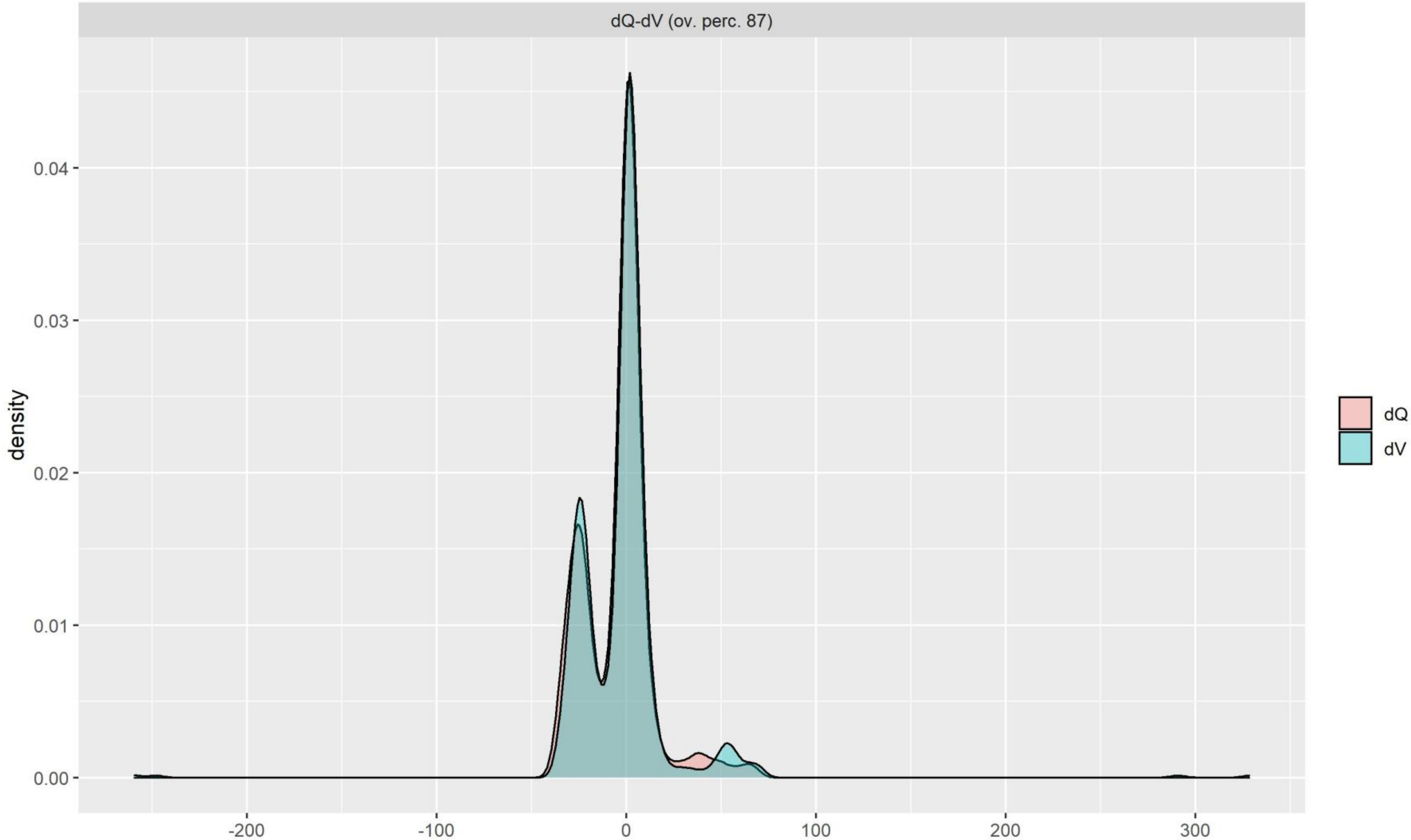


Figure B.2.2 Allatoona Reservoir

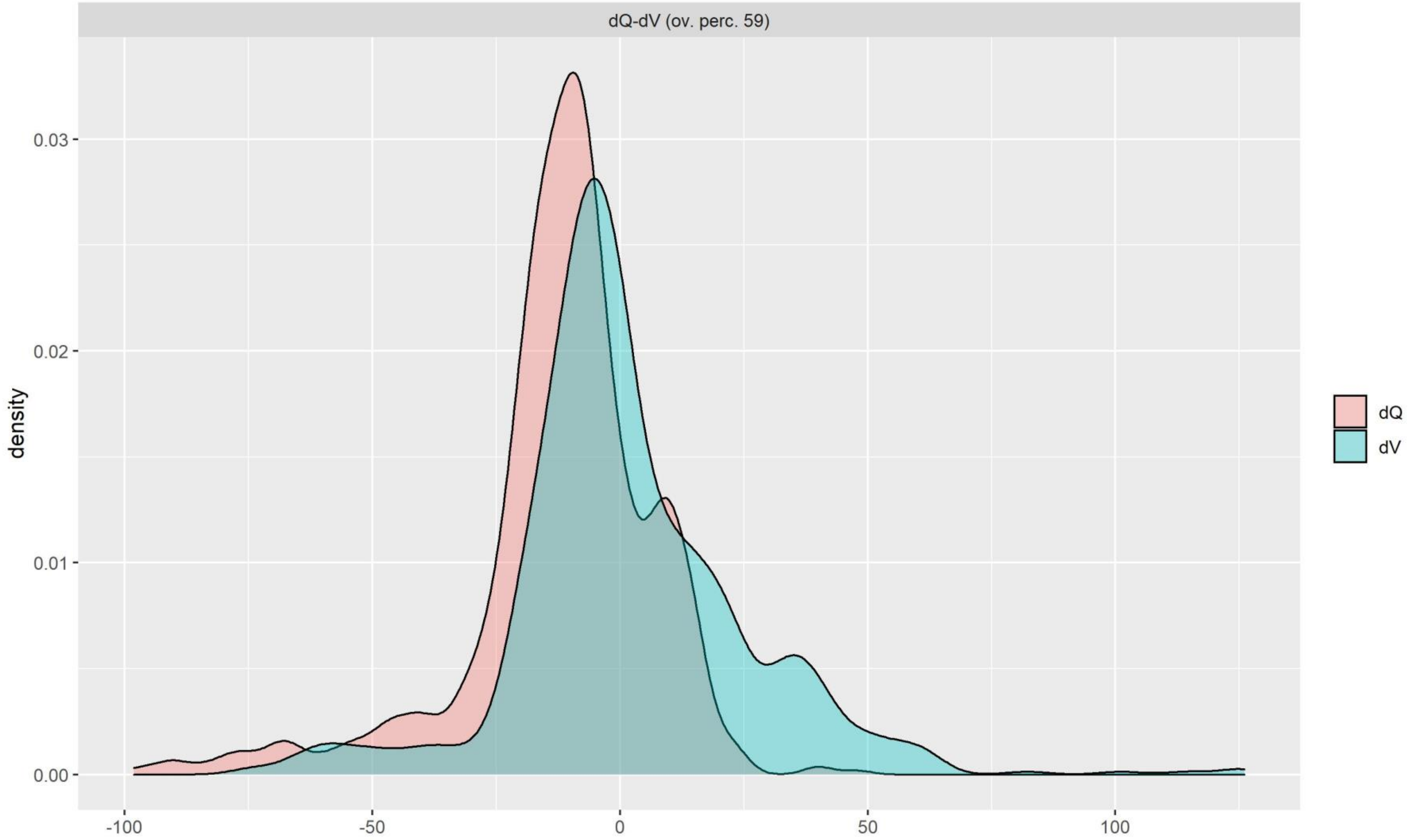


Figure B.2.3 Barlett and Horseshoe Reservoir

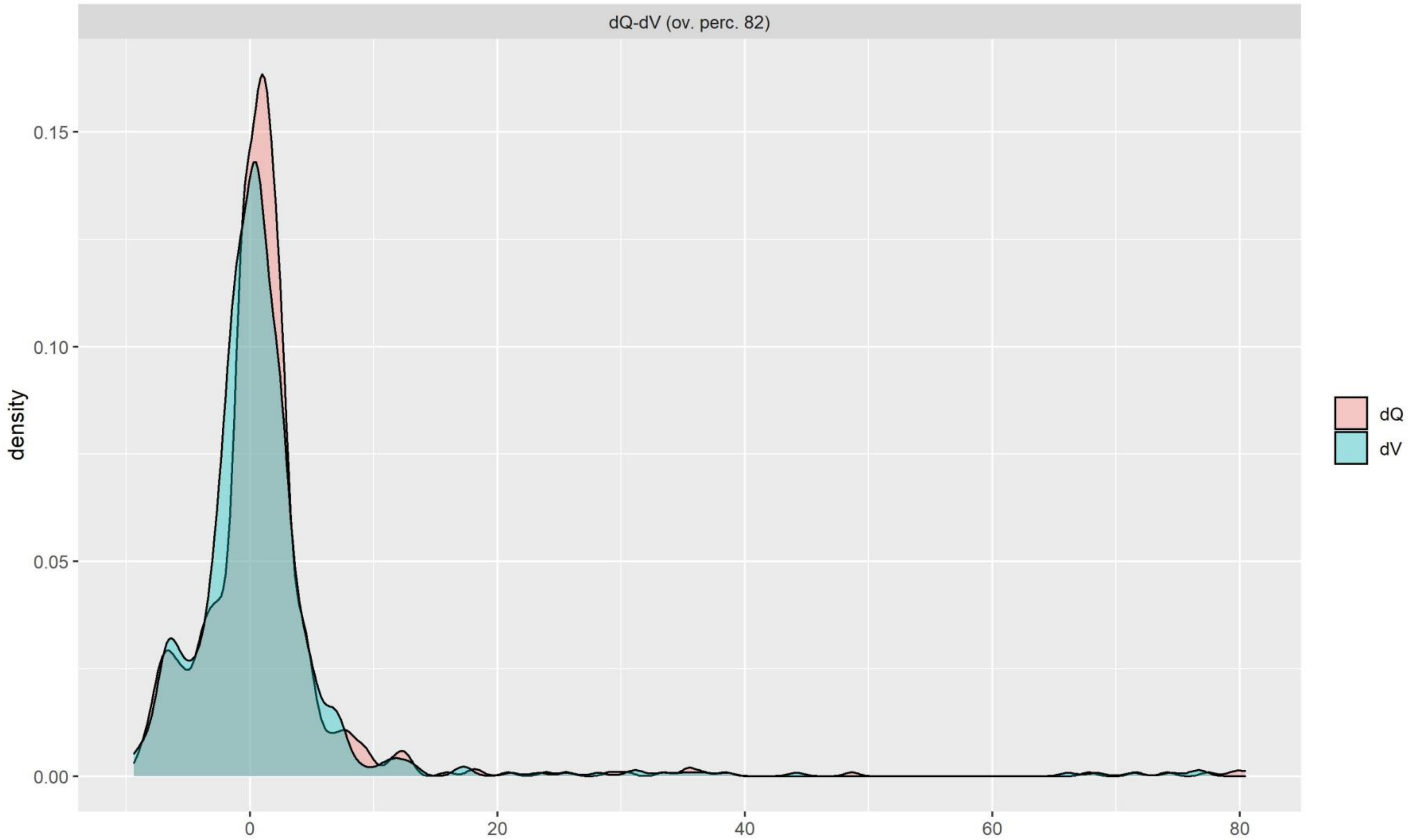


Figure B.2.4 Brantly Lake

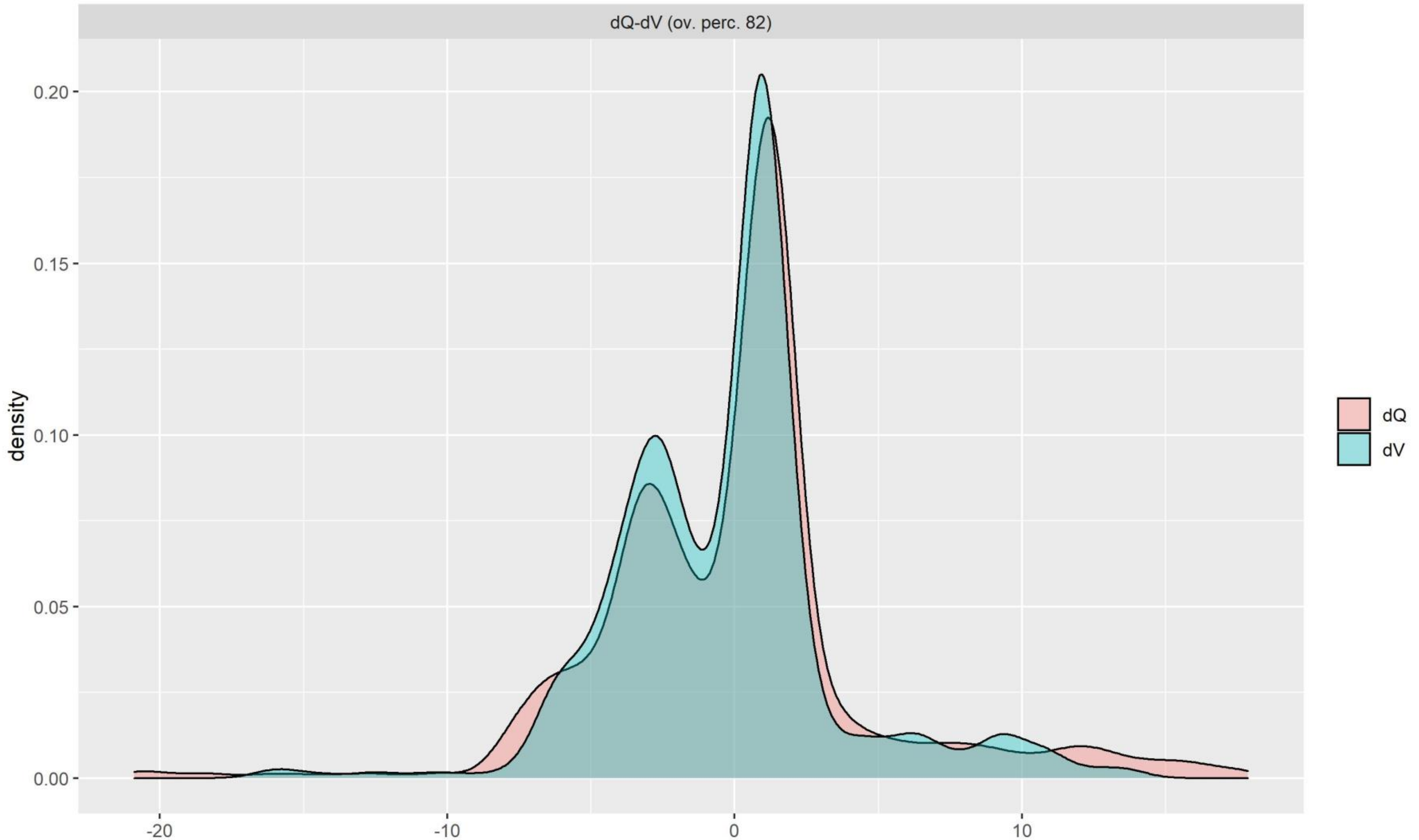


Figure B.2.5 CJ Strike Reservoir

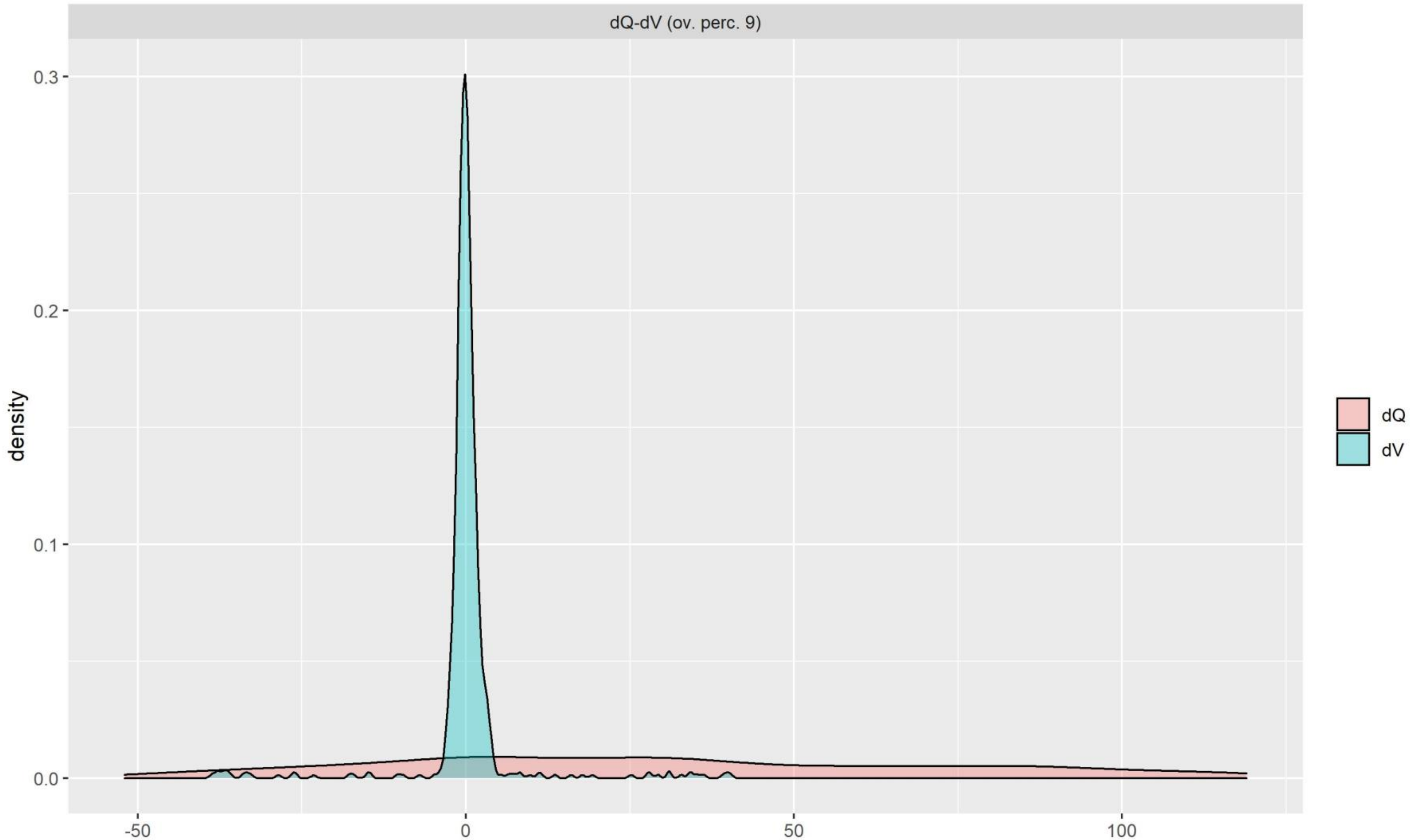


Figure B.2.6 Cochiti Lake

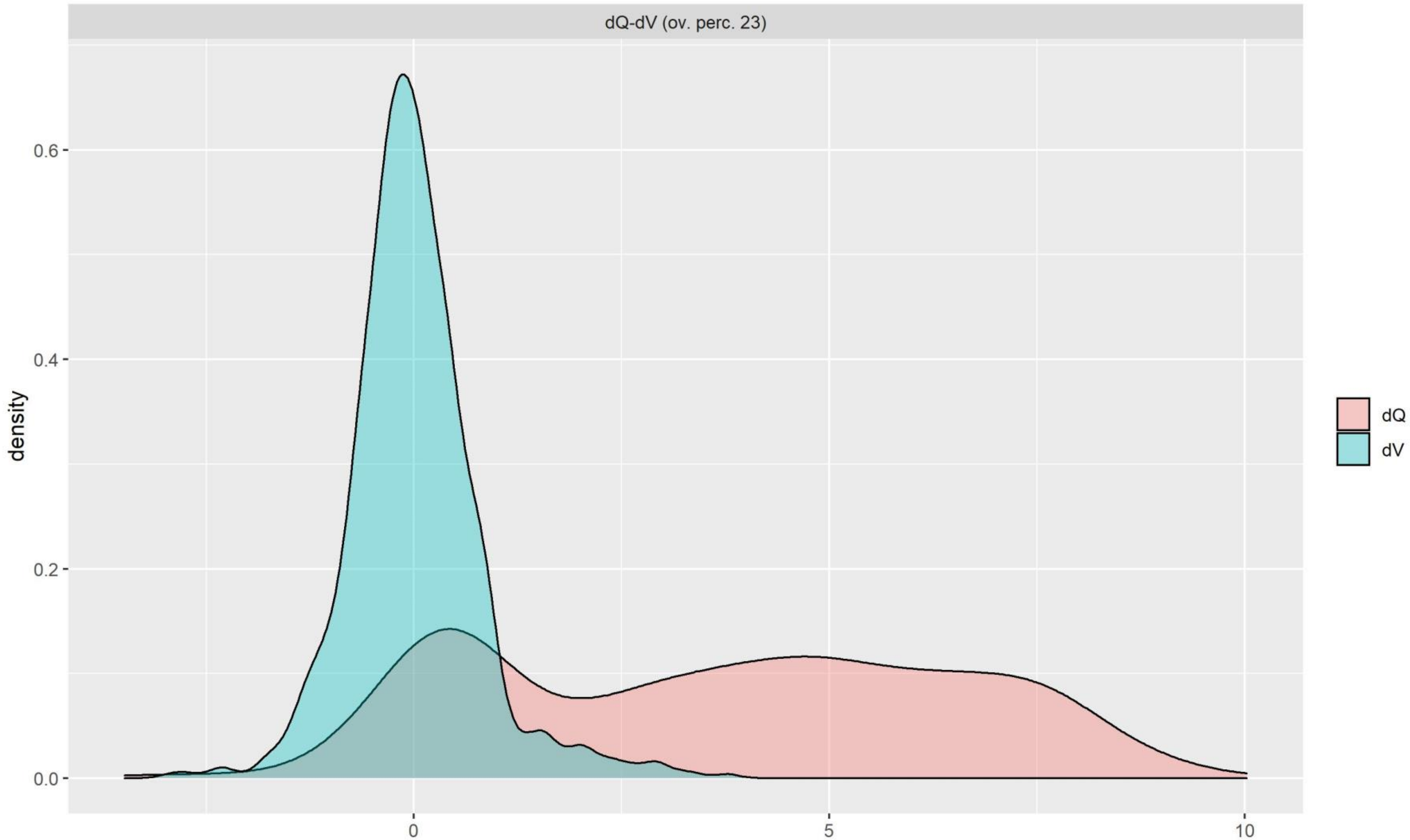


Figure B.2.7 Dworshak Reservoir

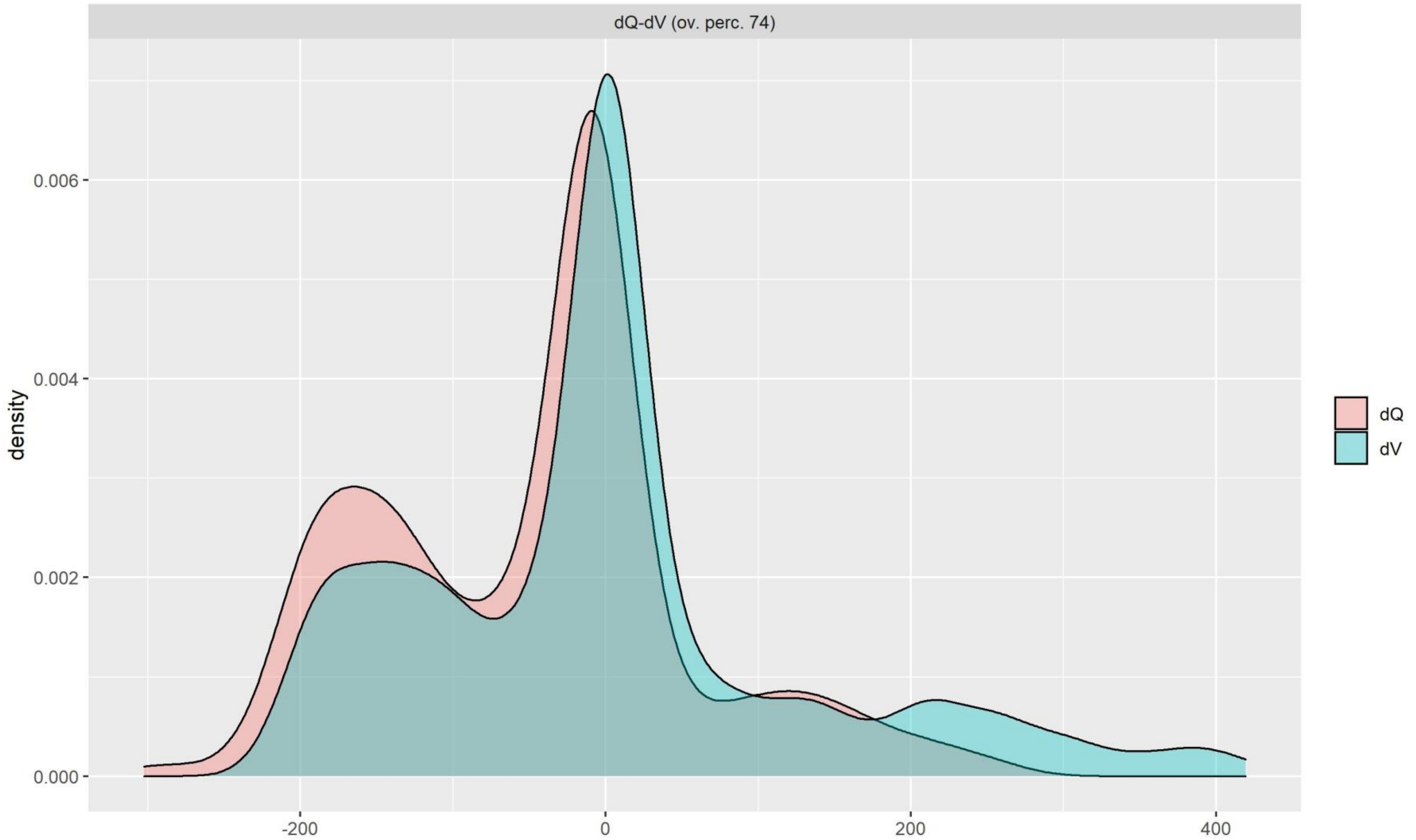


Figure B.2.8 Harlan County Lake

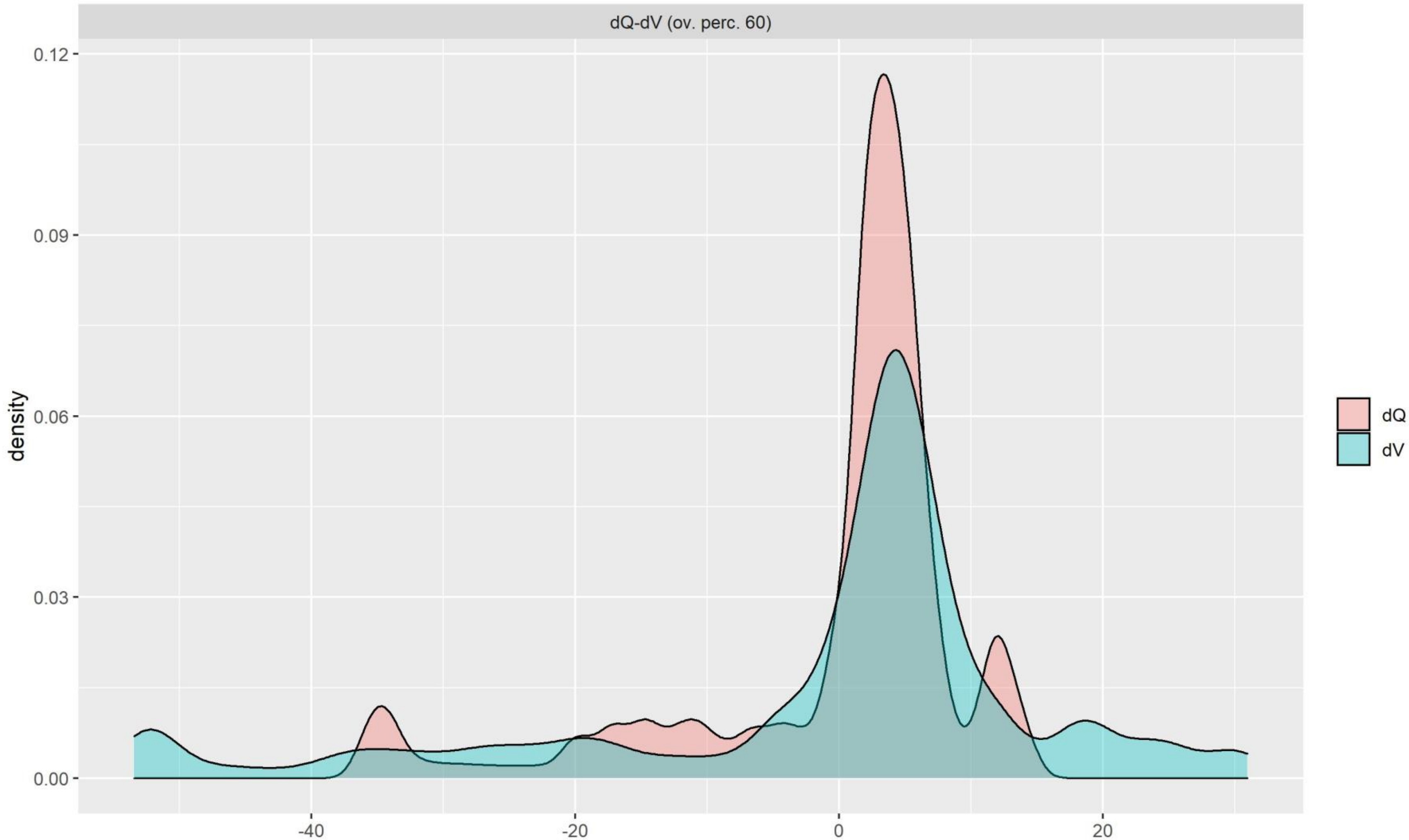


Figure B.2.9 John Martin Reservoir

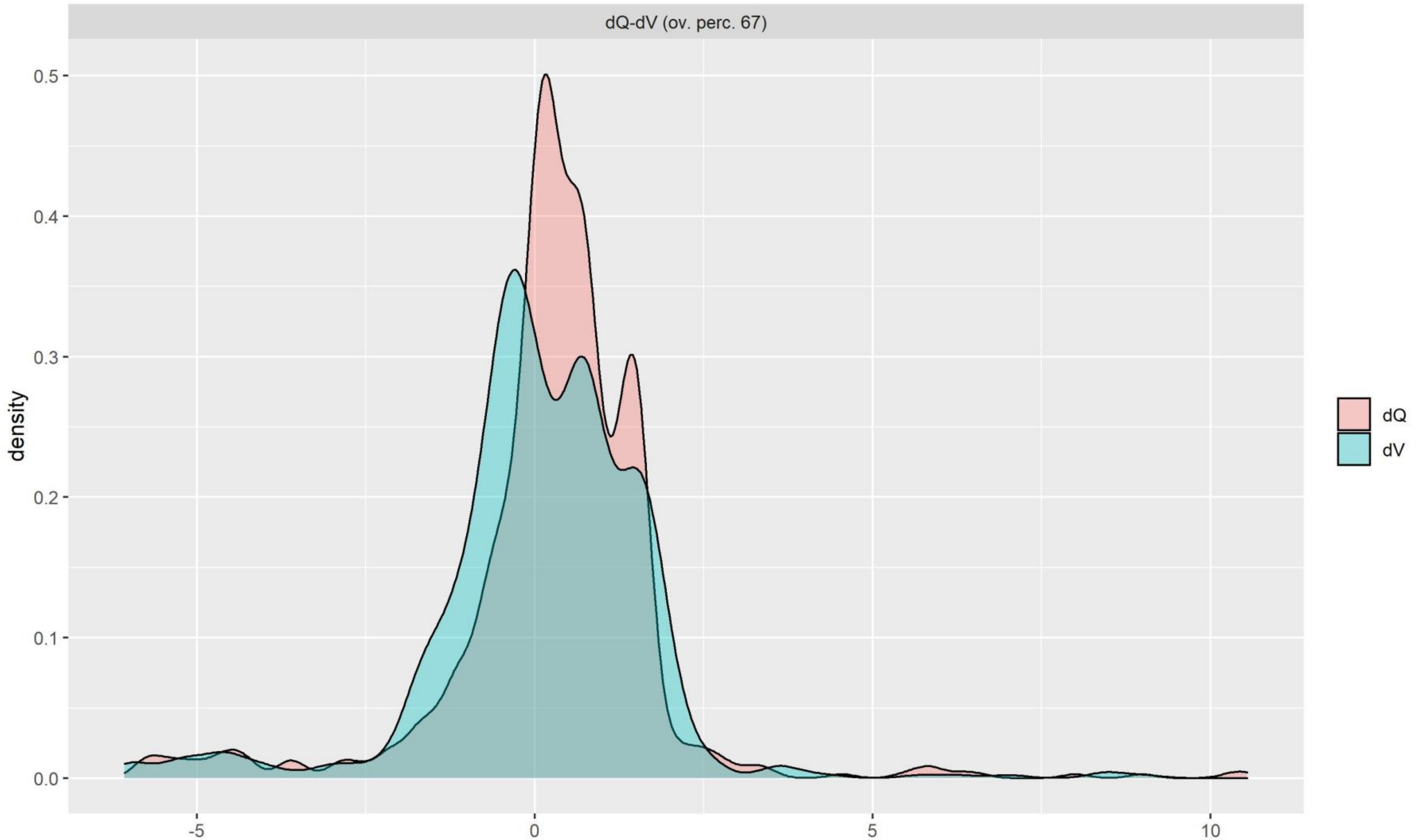


Figure B.2.10 Lake Granbury

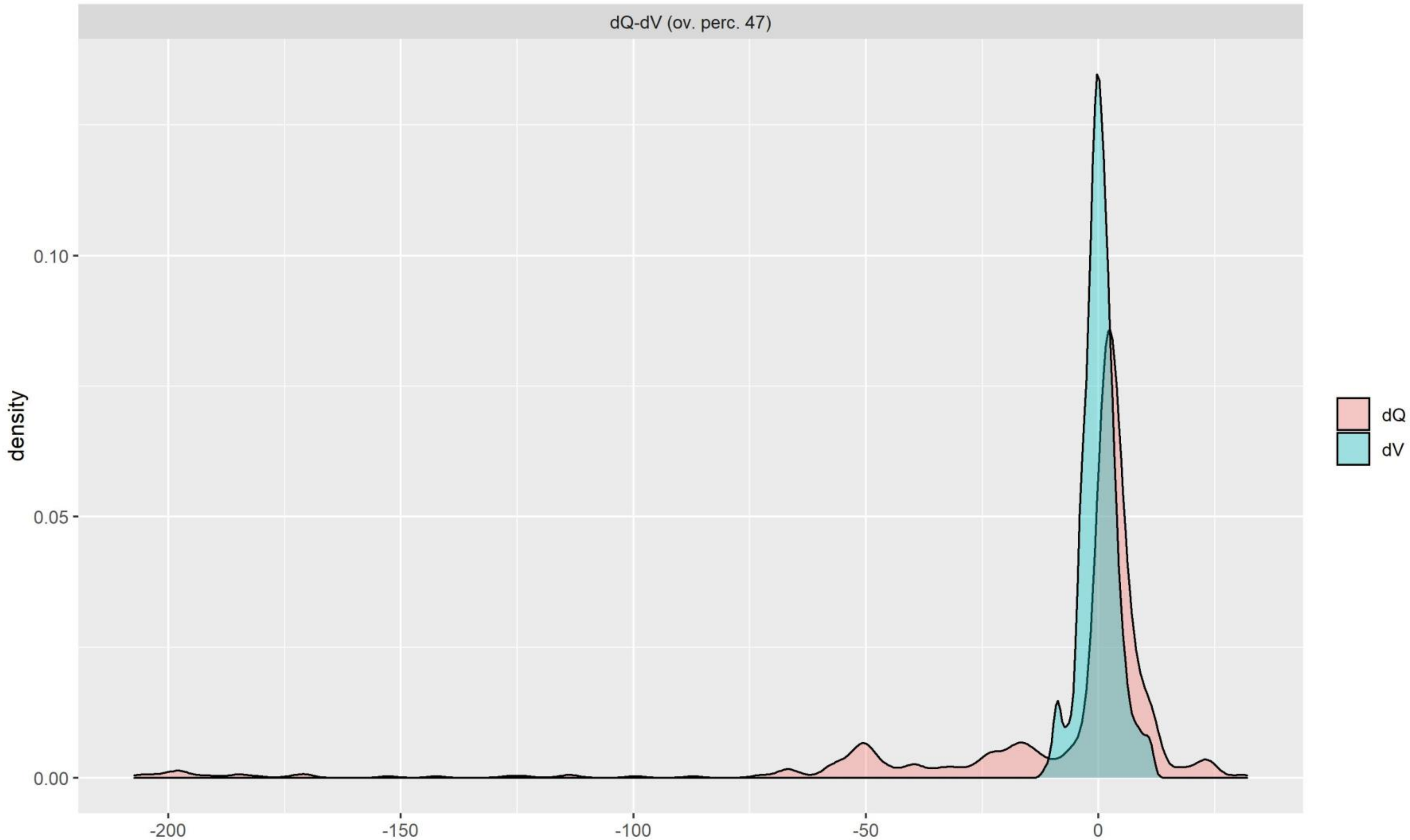


Figure B.2.11 Lake Havasu

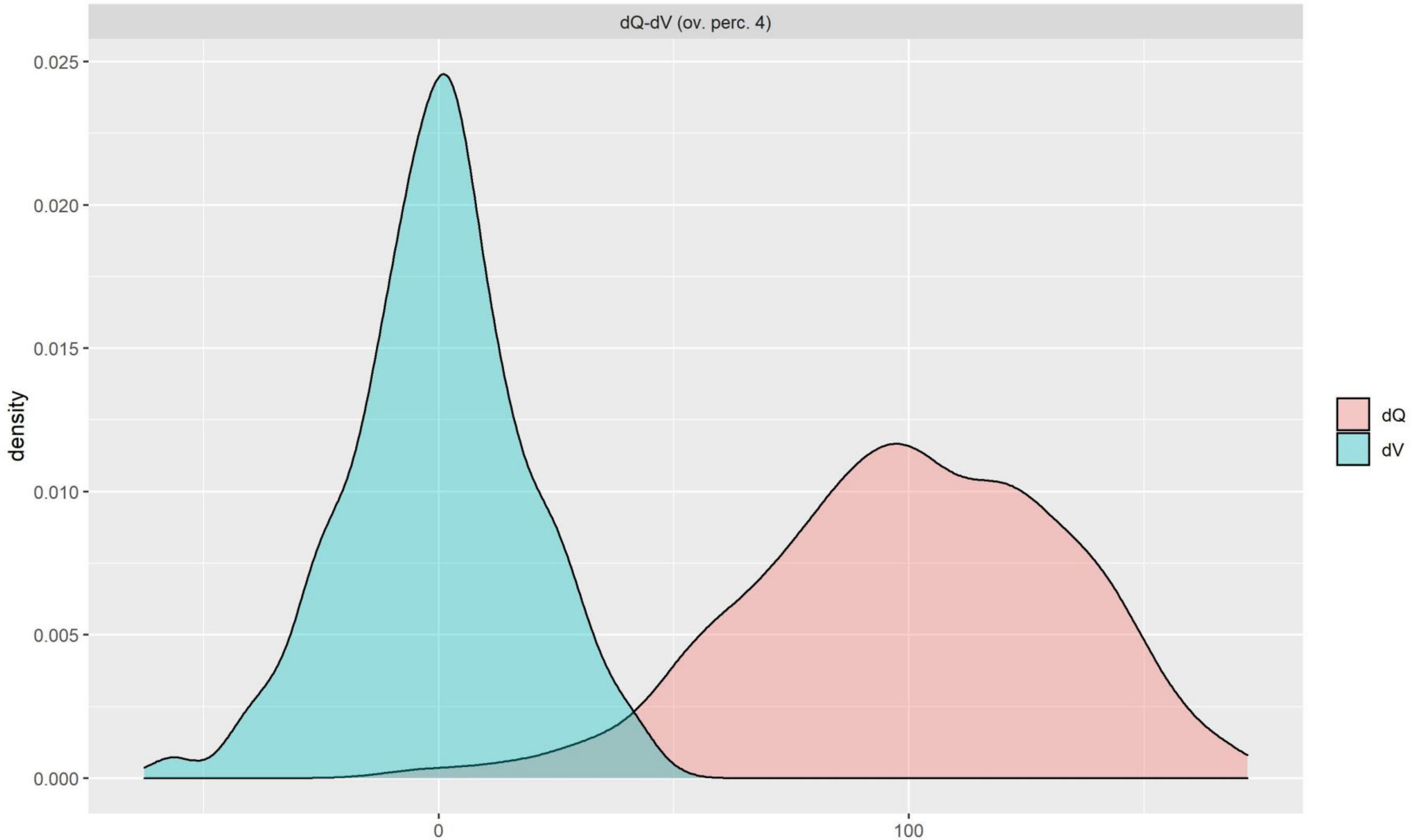


Figure B.2.12 Lake Lanier

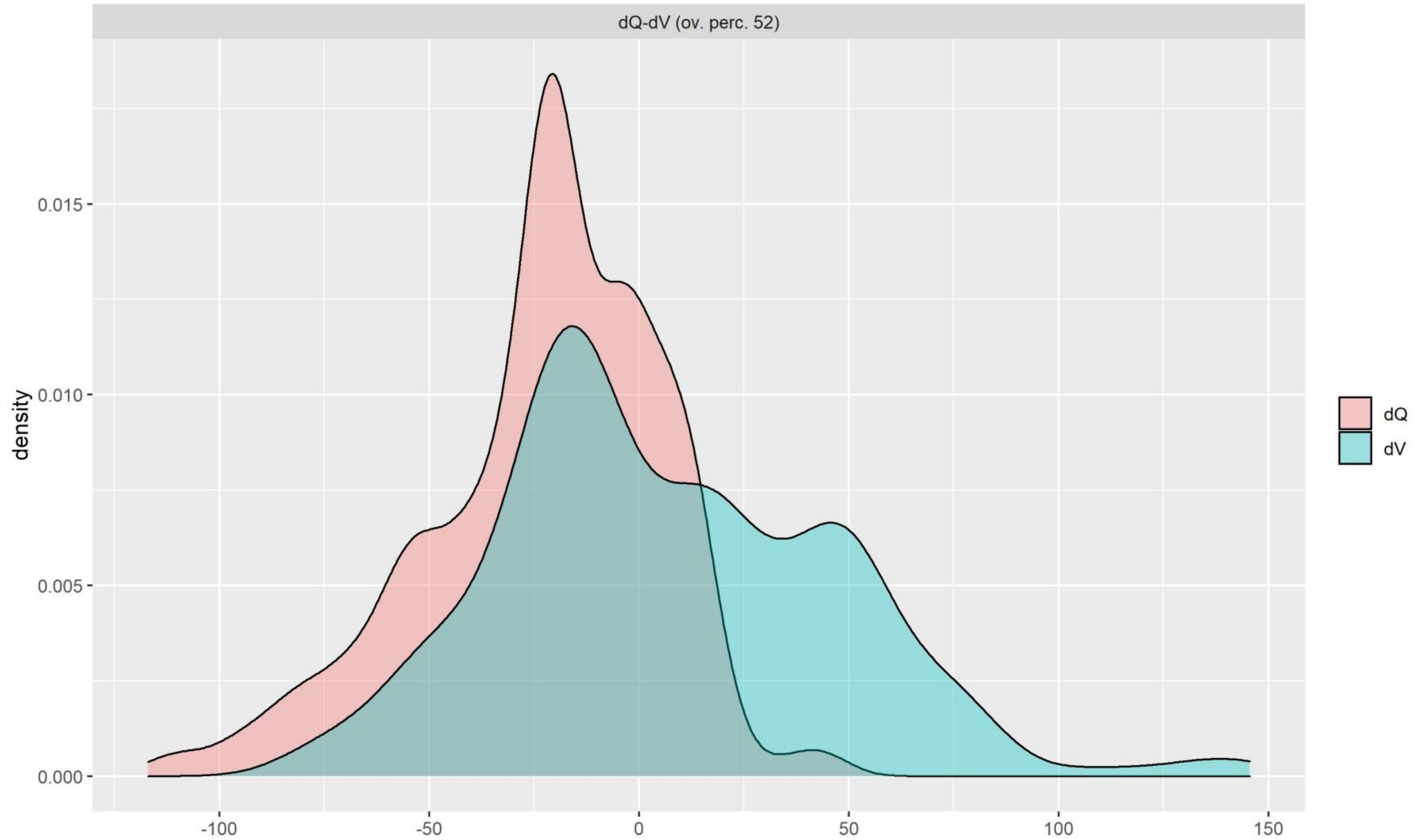


Figure B.2.13 Lake Livingston

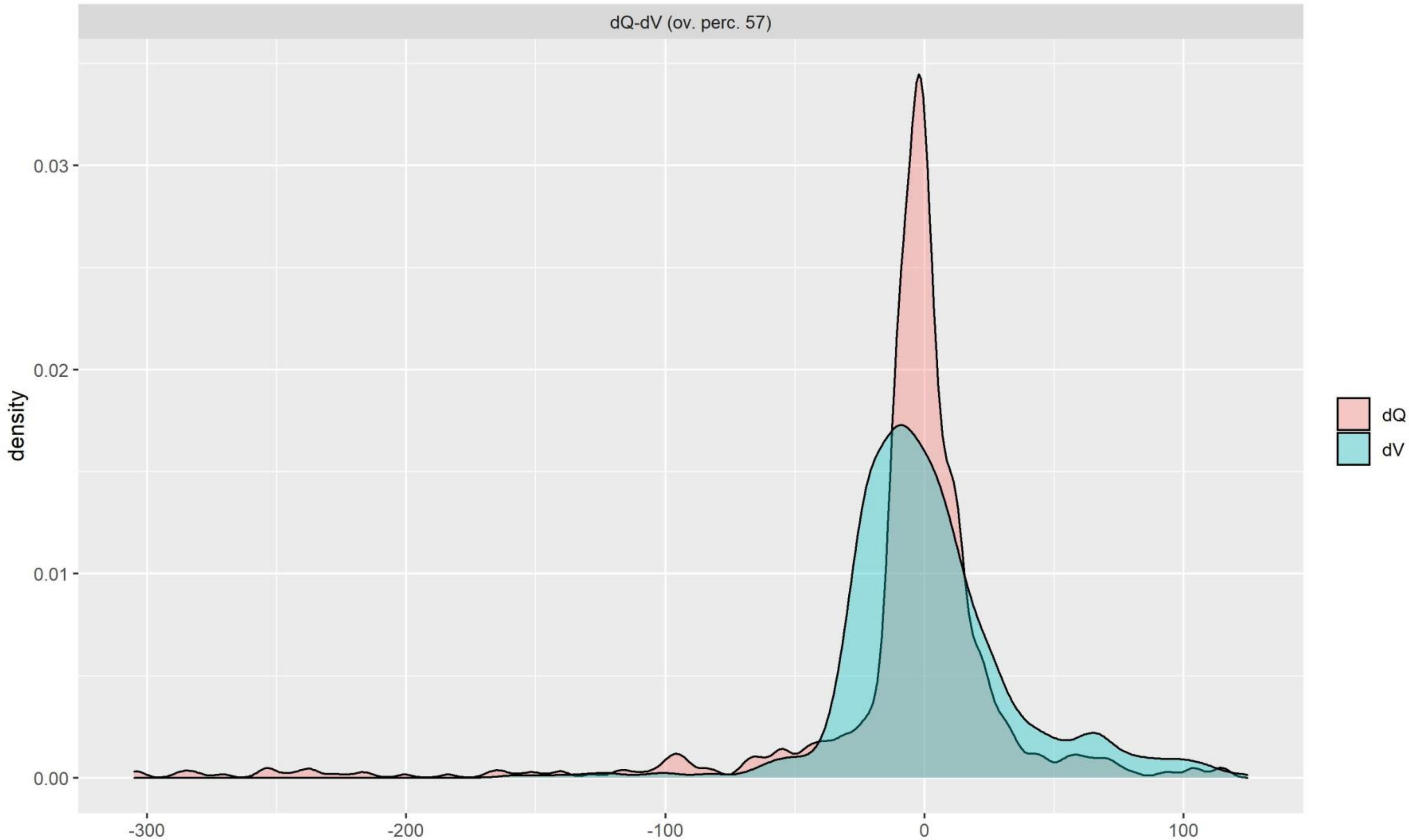


Figure B.2.14 Lake Mohave

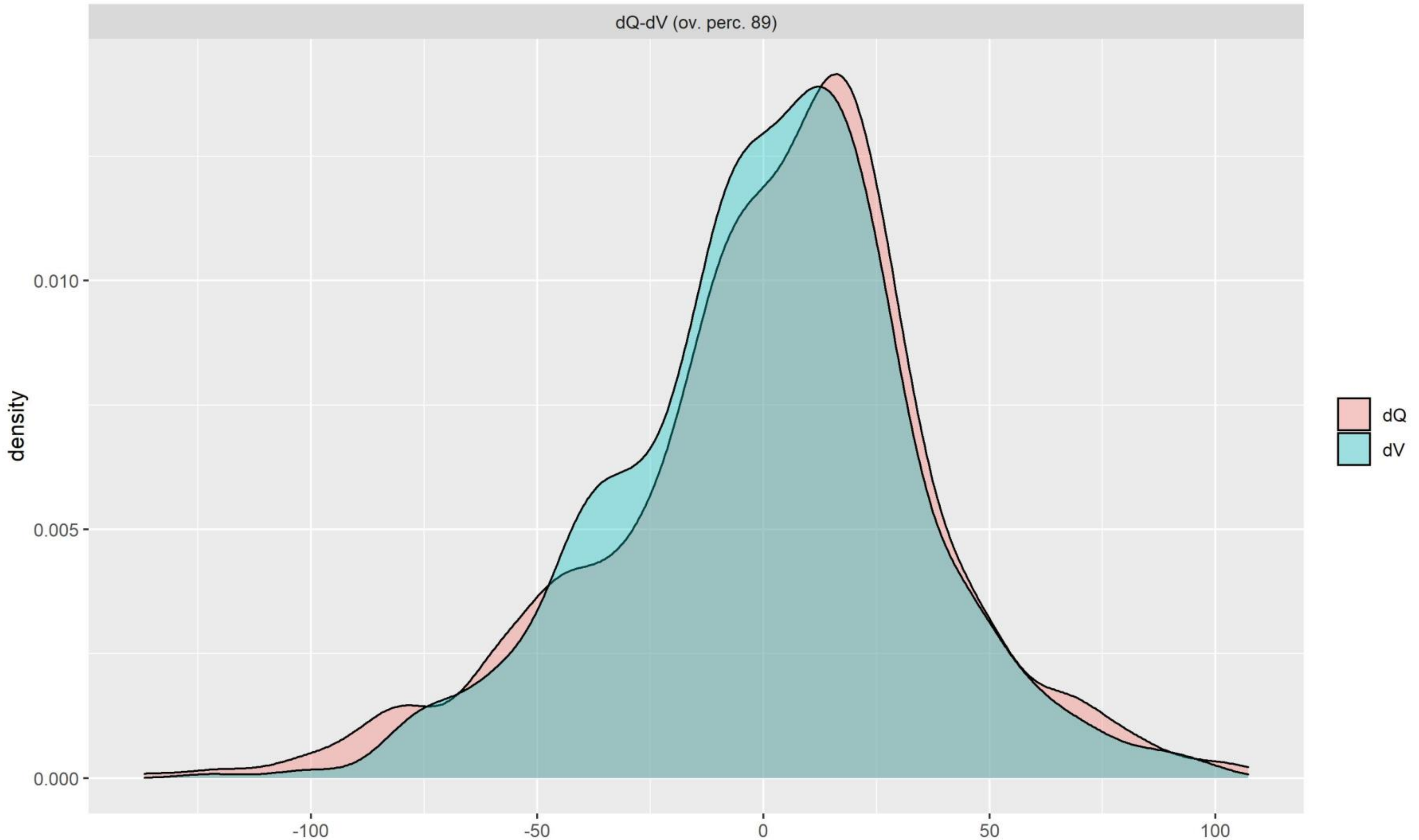


Figure B.2.15 Lake Red Rock

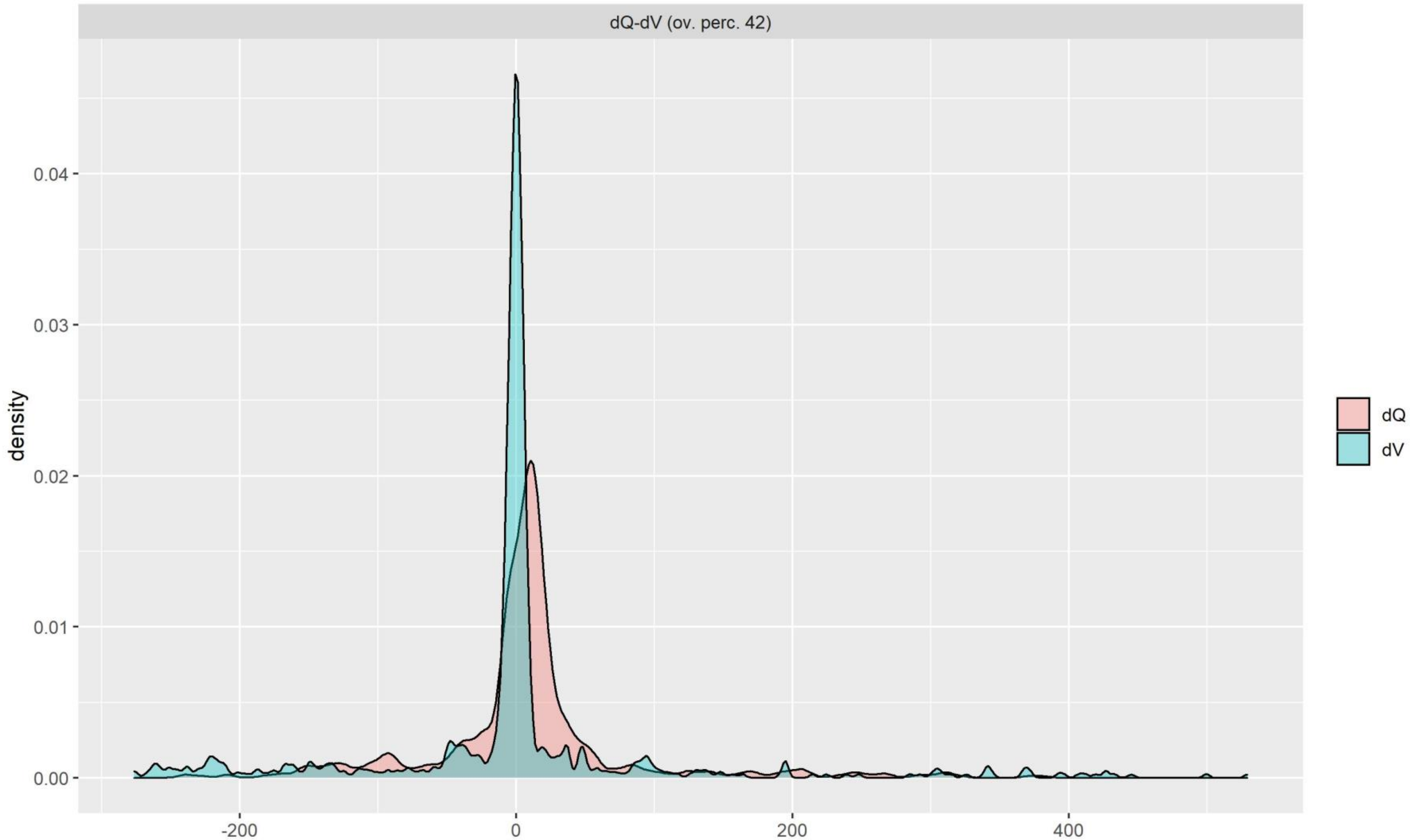


Figure B.2.16 Lake Seminole

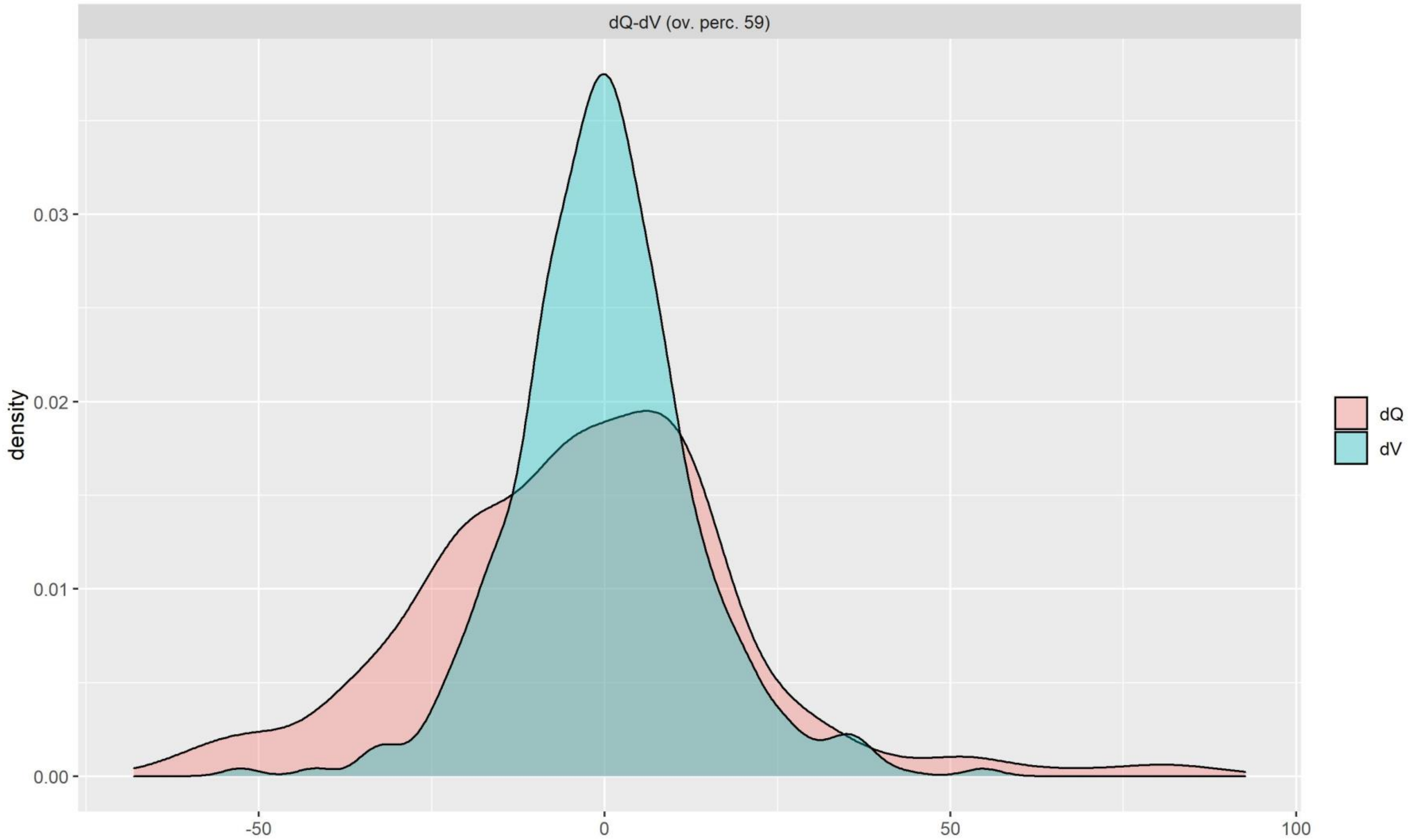


Figure B.2.17 Lake Taxoma

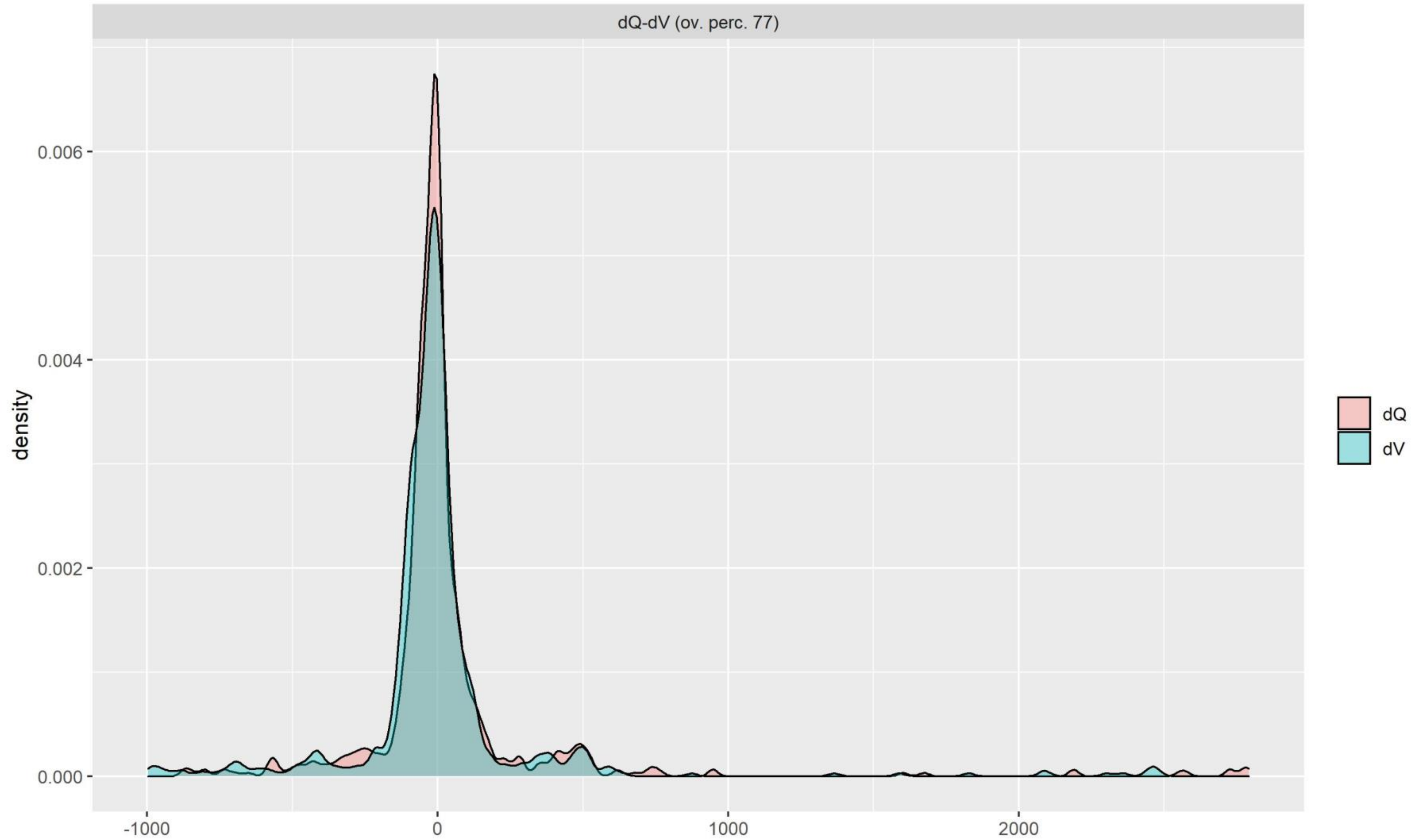


__Figure B.2.18 Milford Lake__

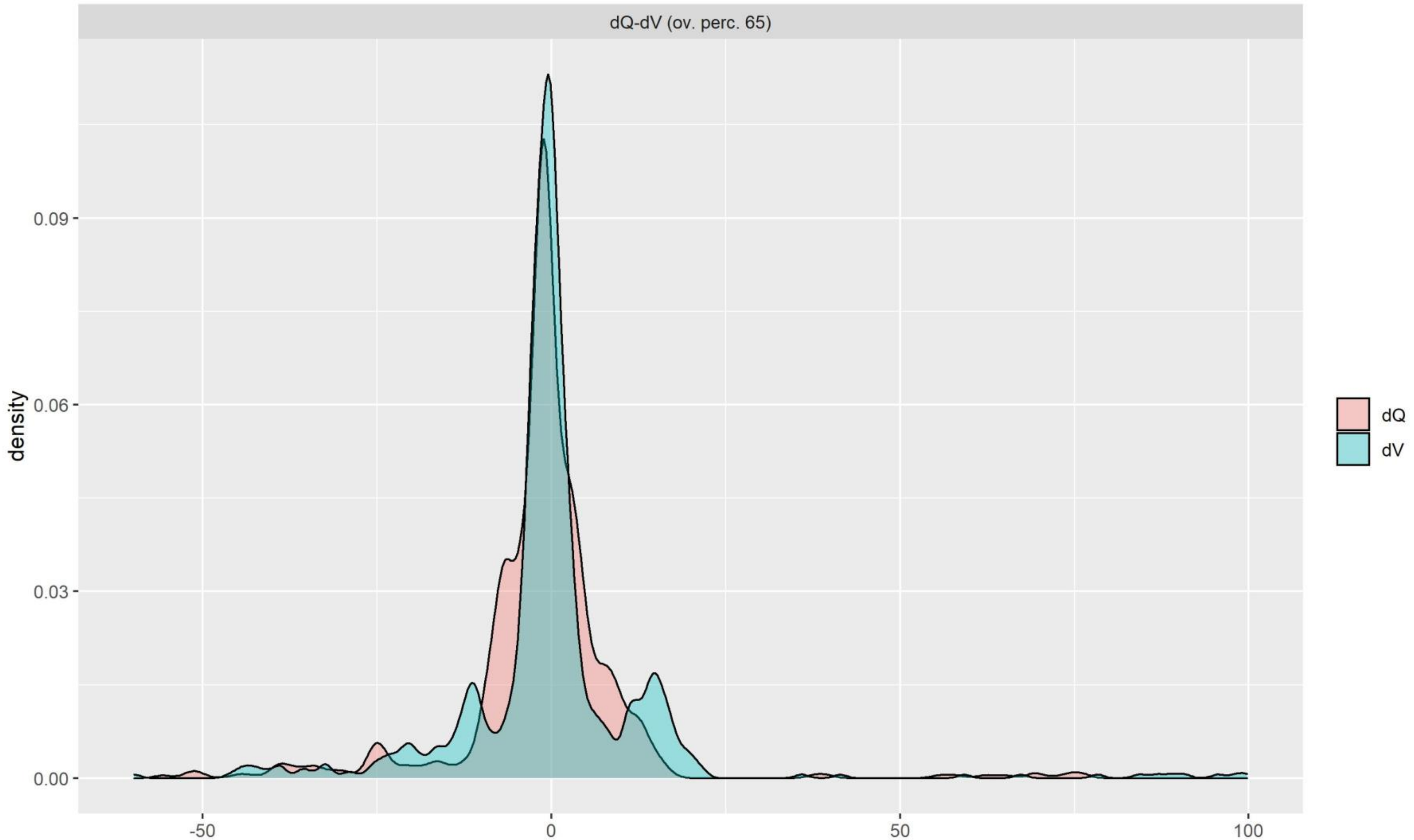


Figure B.2.19 Navajo Lake

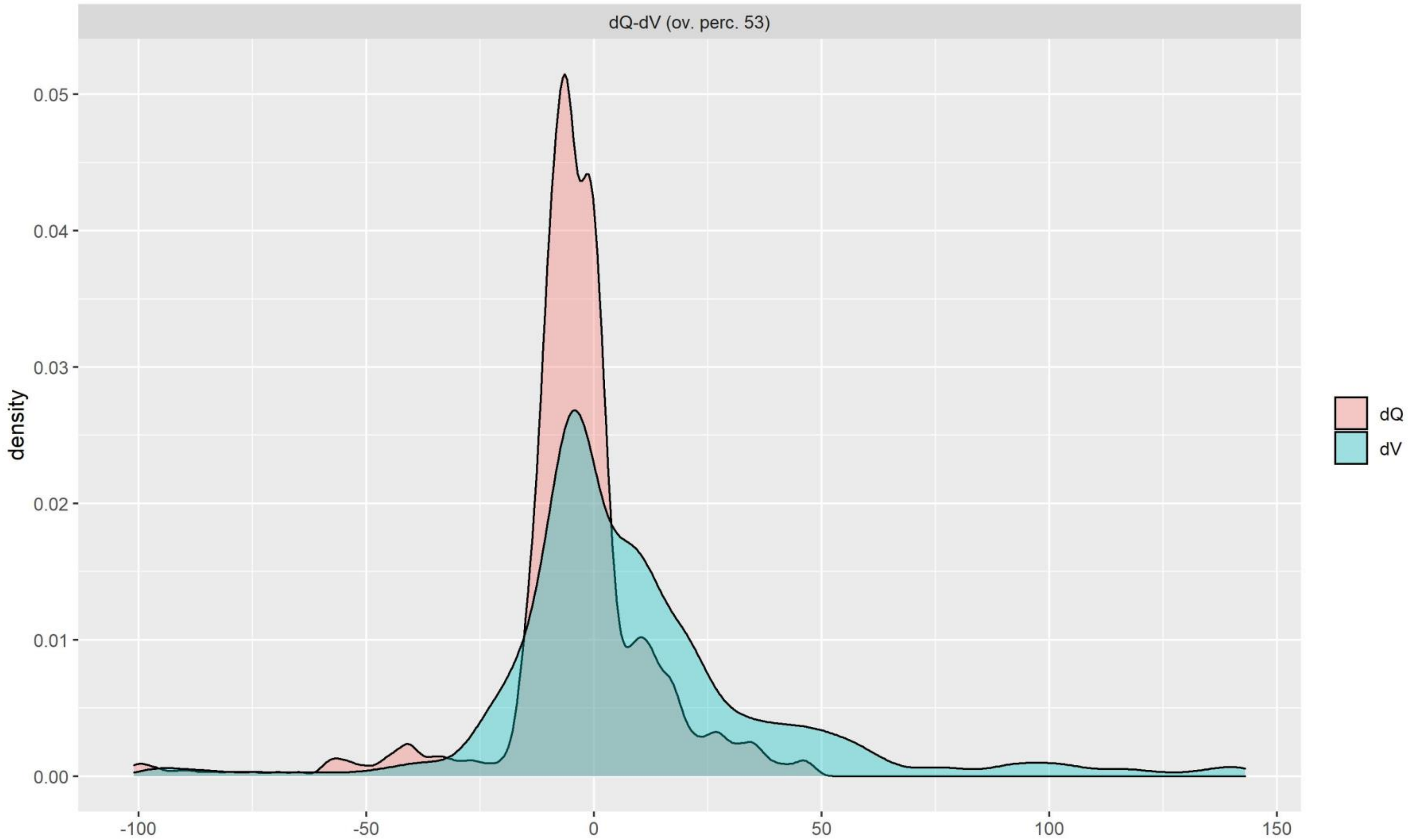


Figure B.2.20 Possum Kingdom Lake

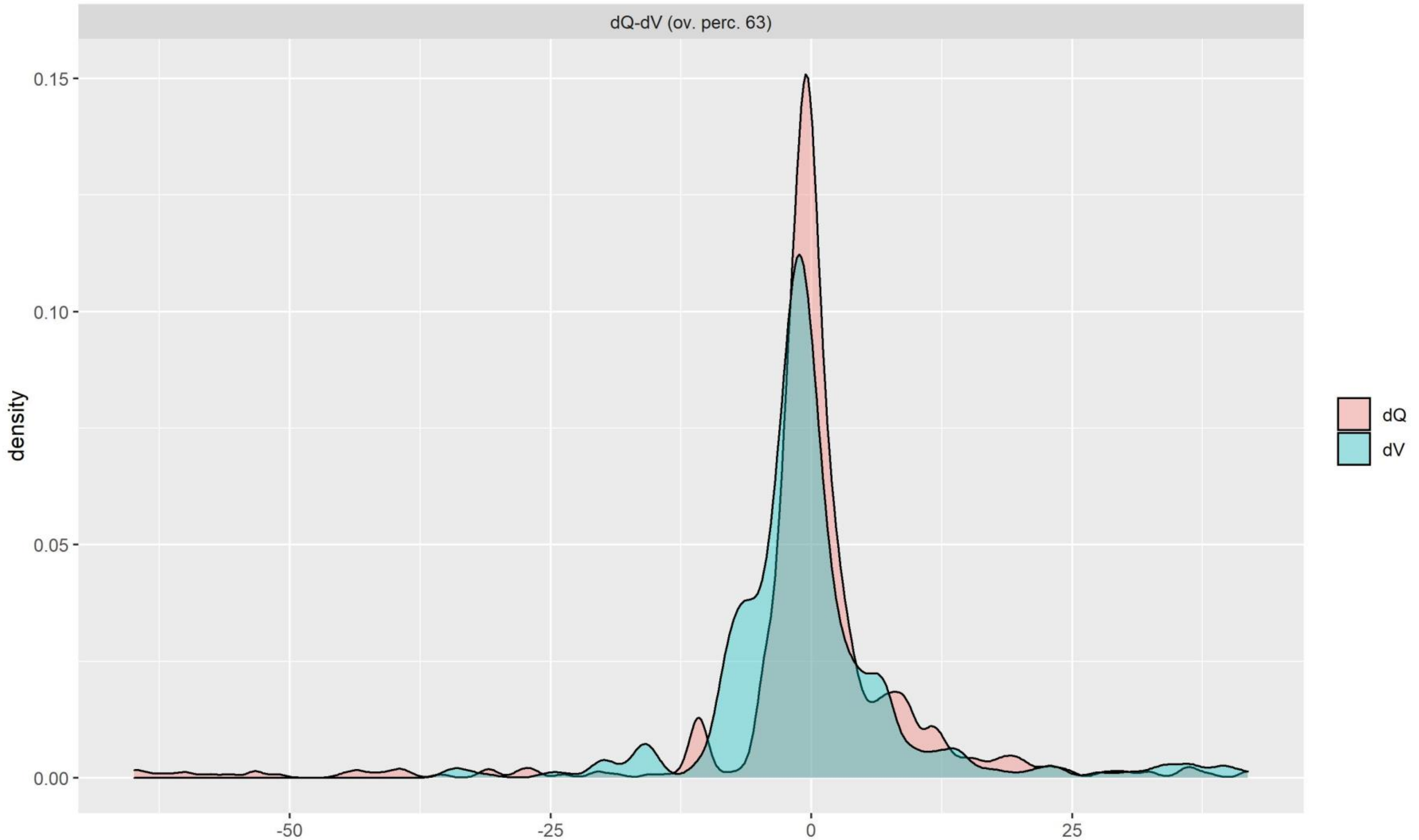


Figure B.2.21 Shasta Lake

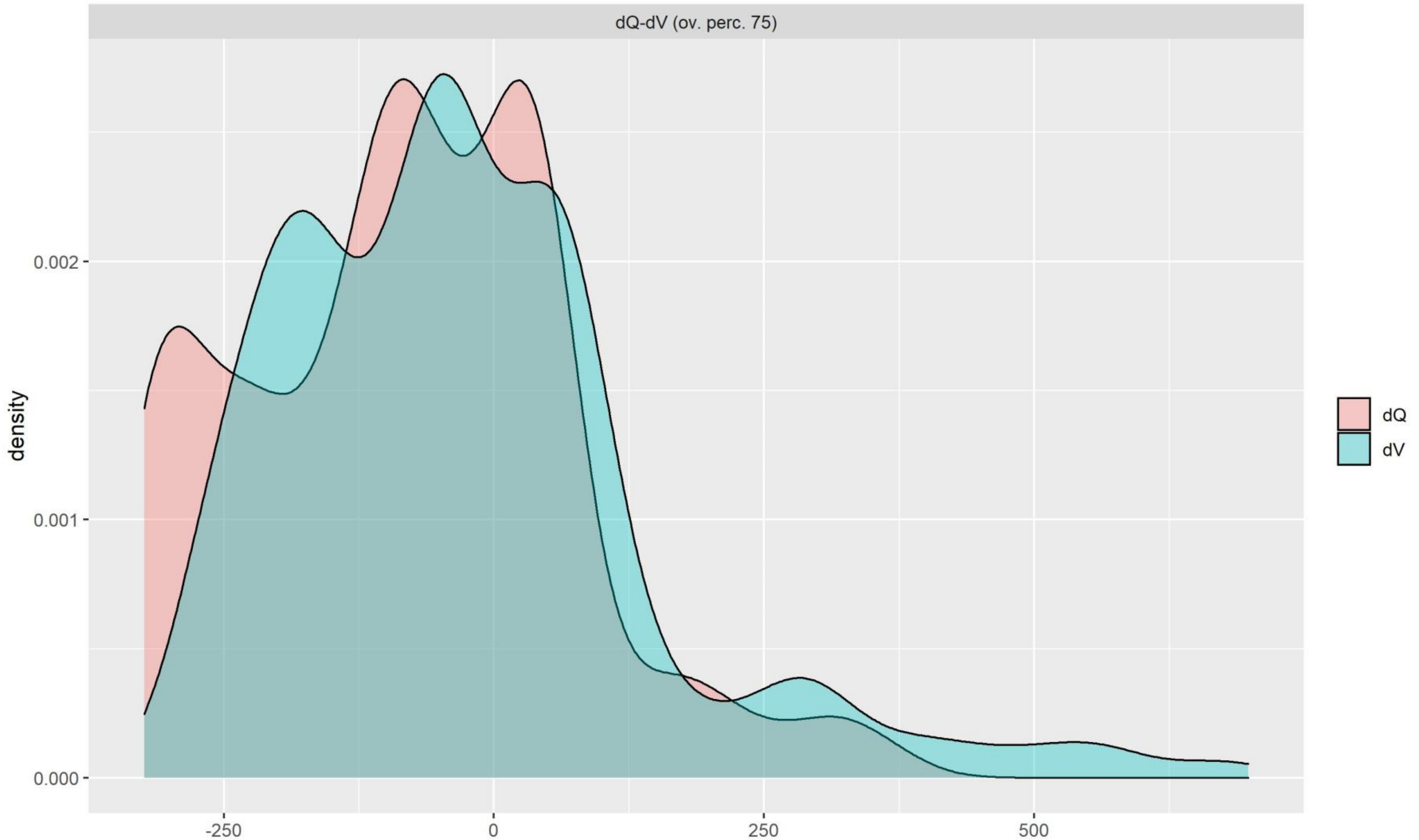


Figure B.2.22 Toledo Bend Lake

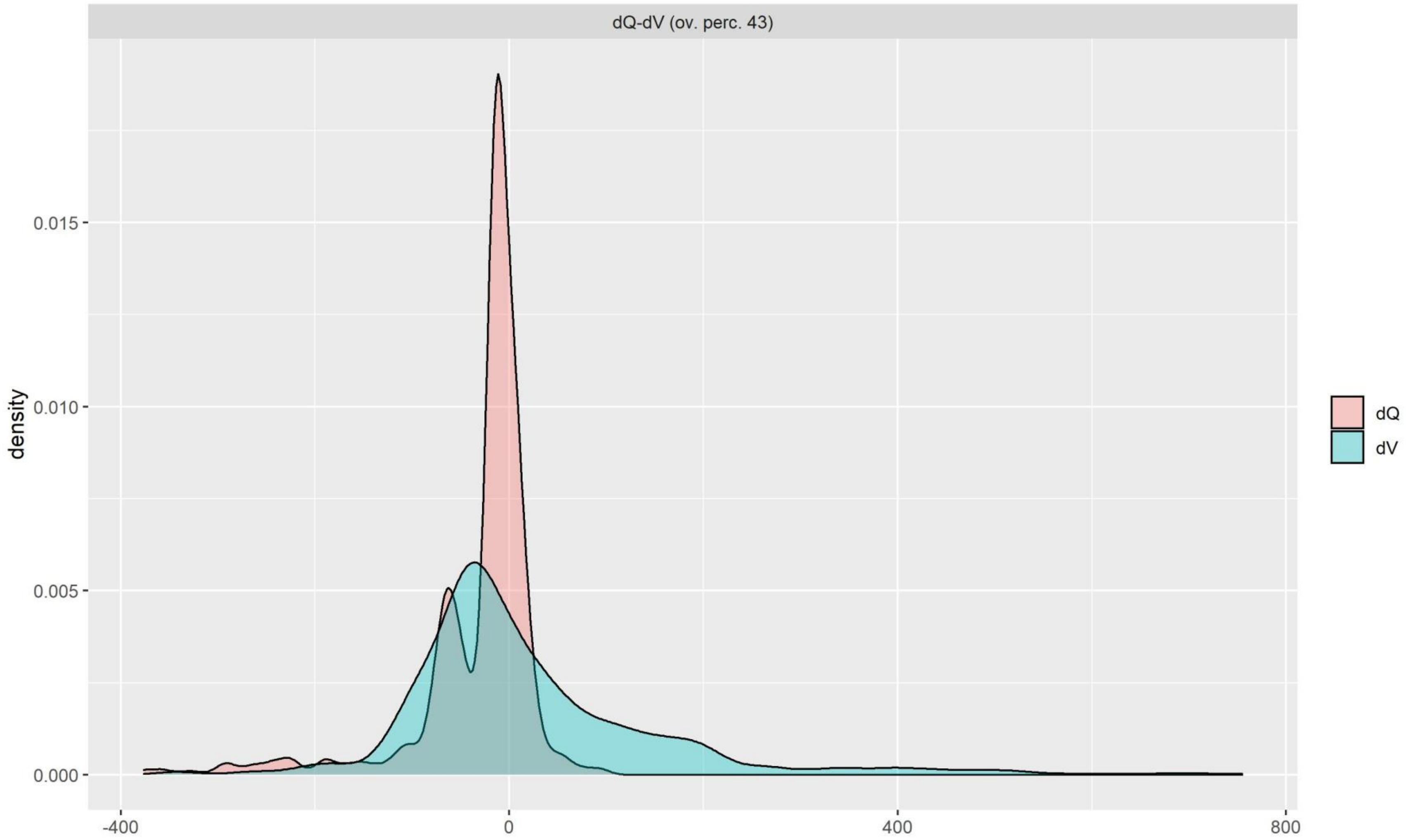


Figure B.2.23 Tuttle Creek Reservoir

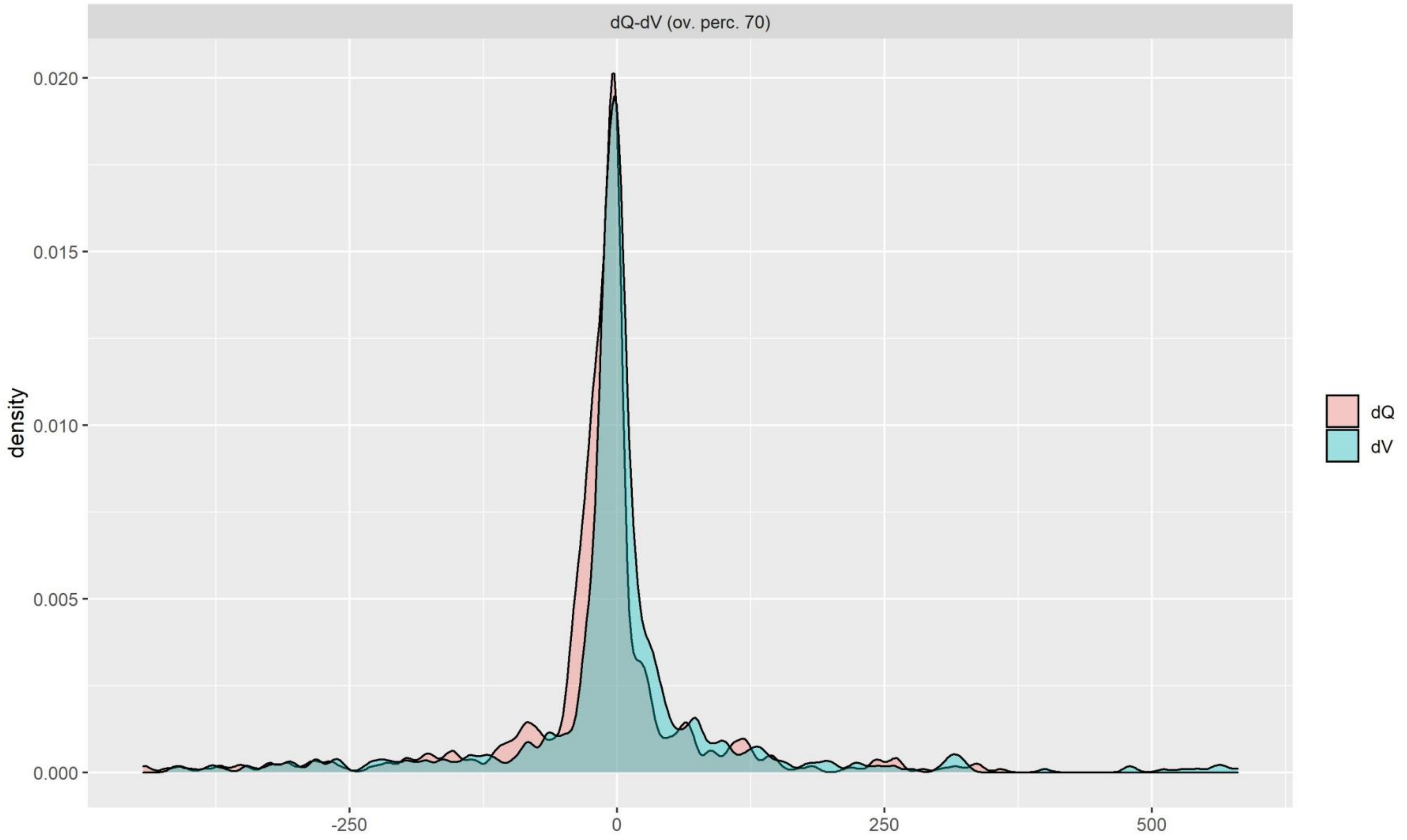


Figure B.2.24 Walter F George Reservoir

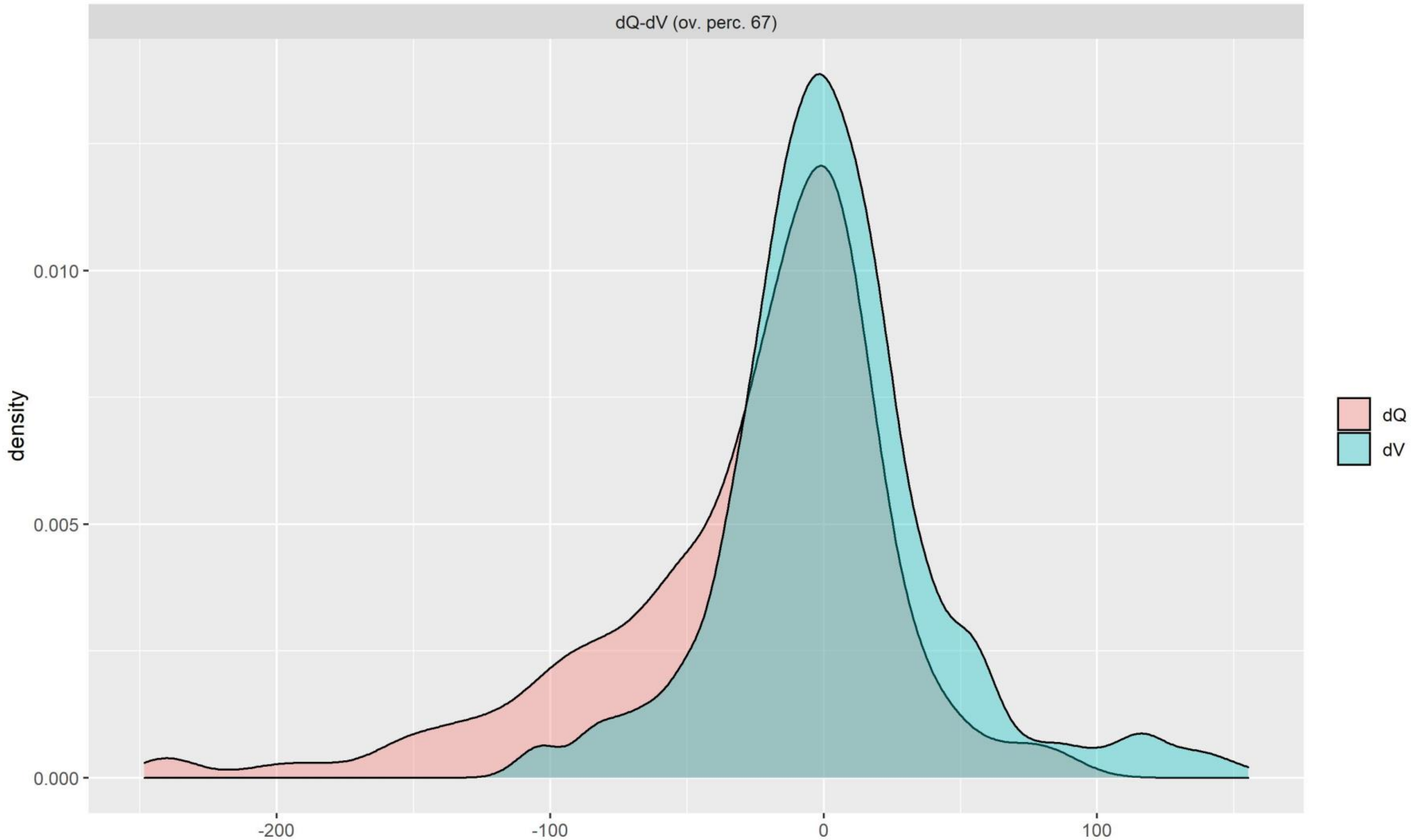


Figure B.2.25 West Point Lake

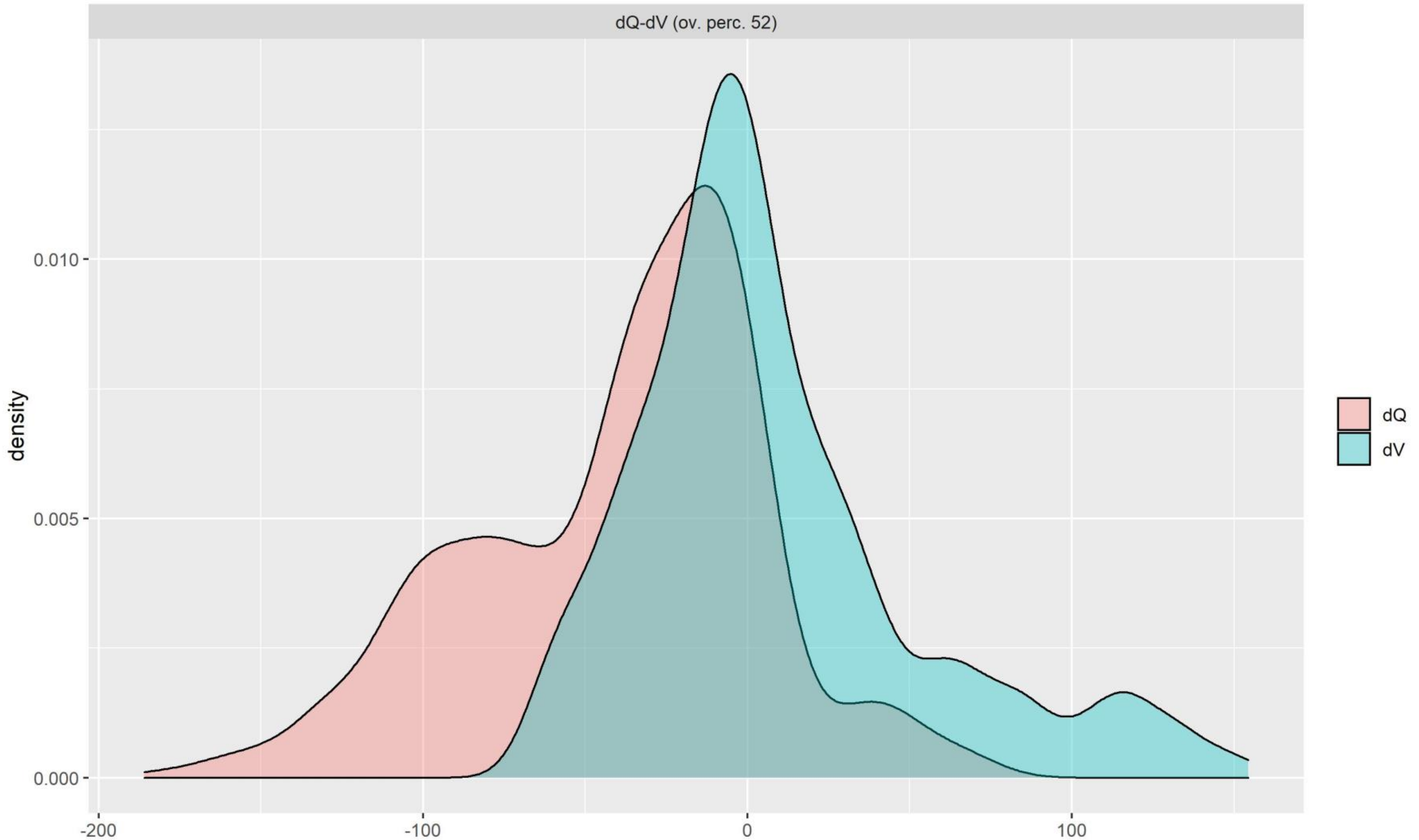
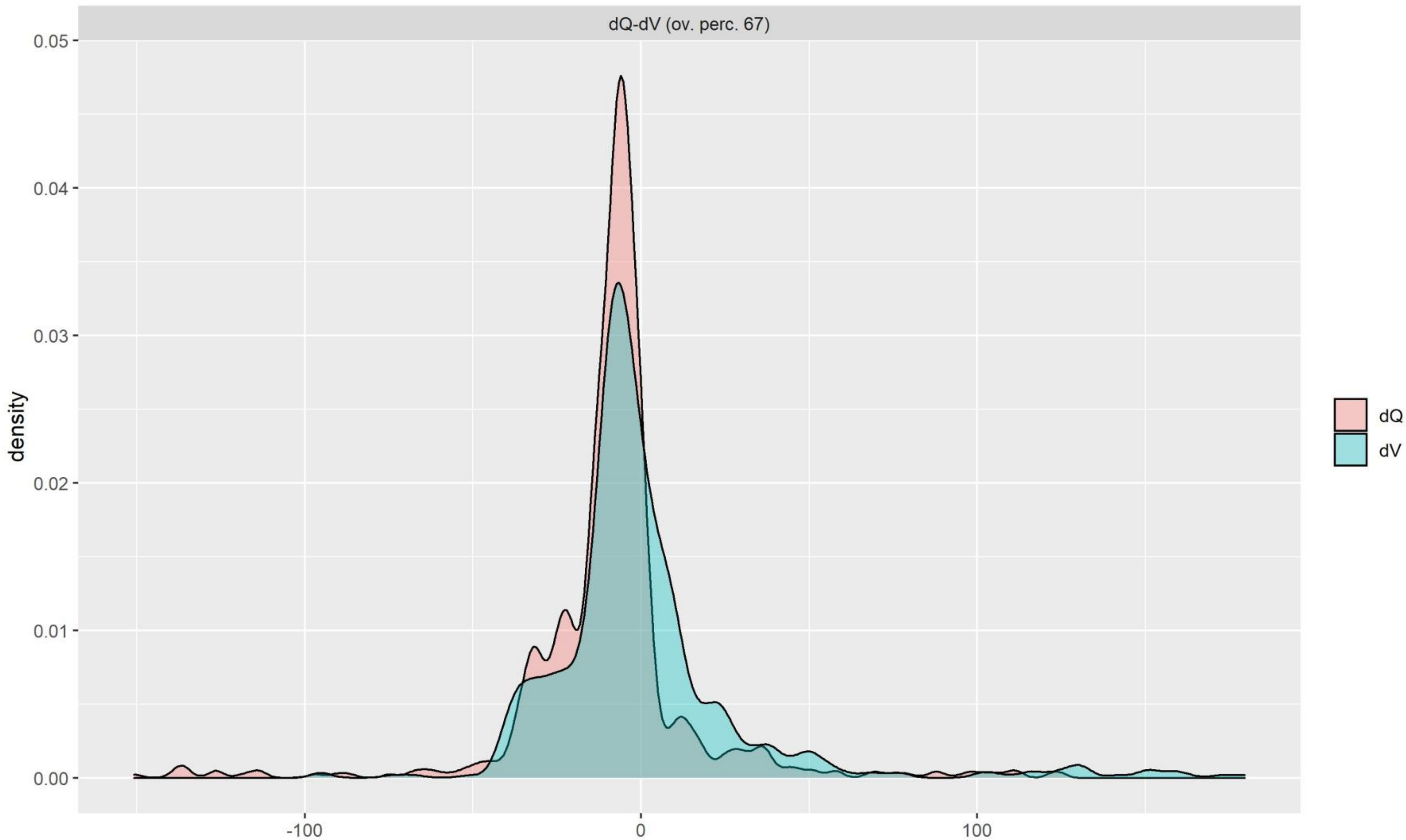


Figure B.2.26 Whitney Lake



Appendix B.3 Daily Hydrographs of dQ and dV

Figure B.3.1 Abiquiu Reservoir

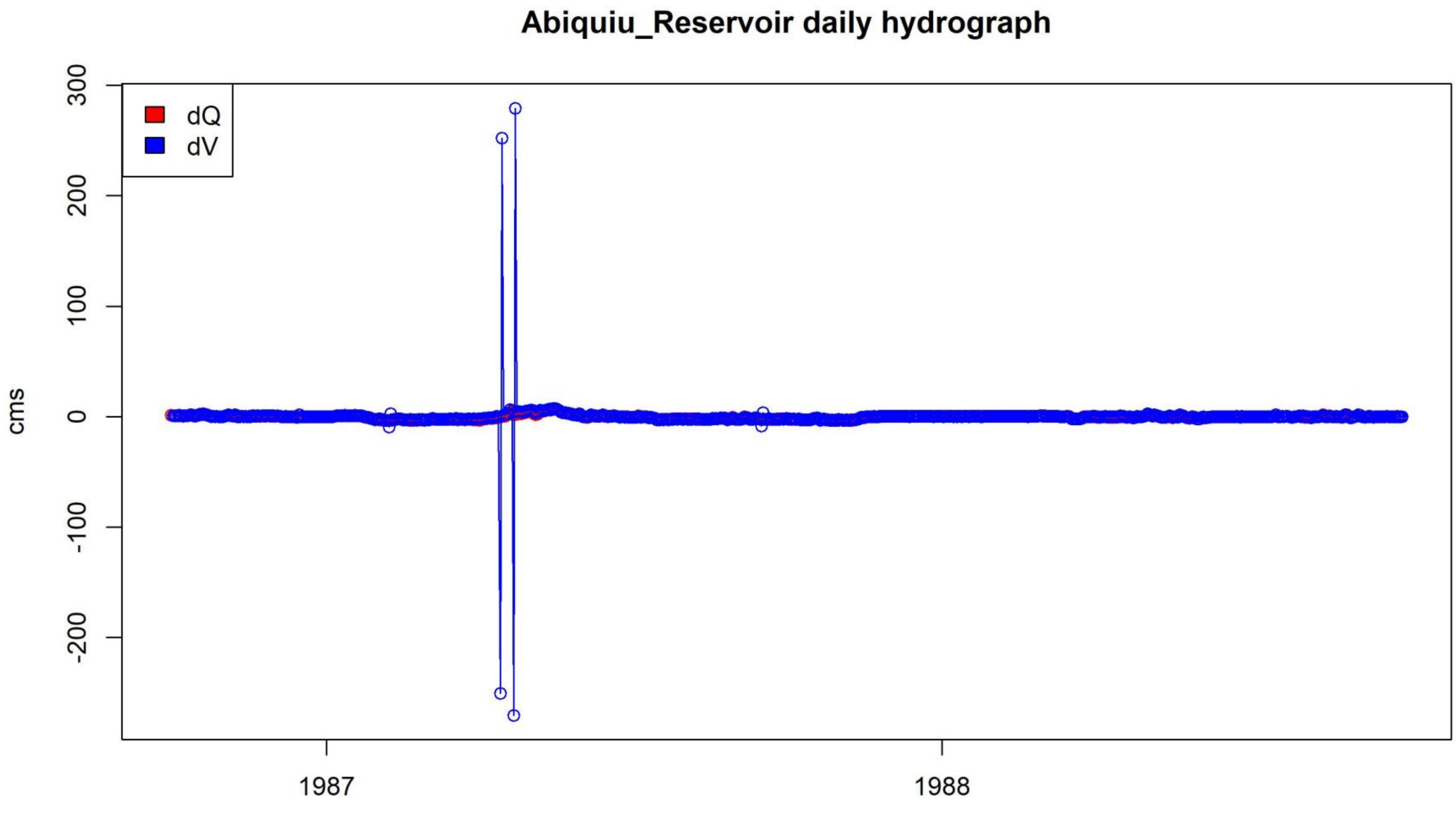


Figure B.3.2 Allatoona Reservoir

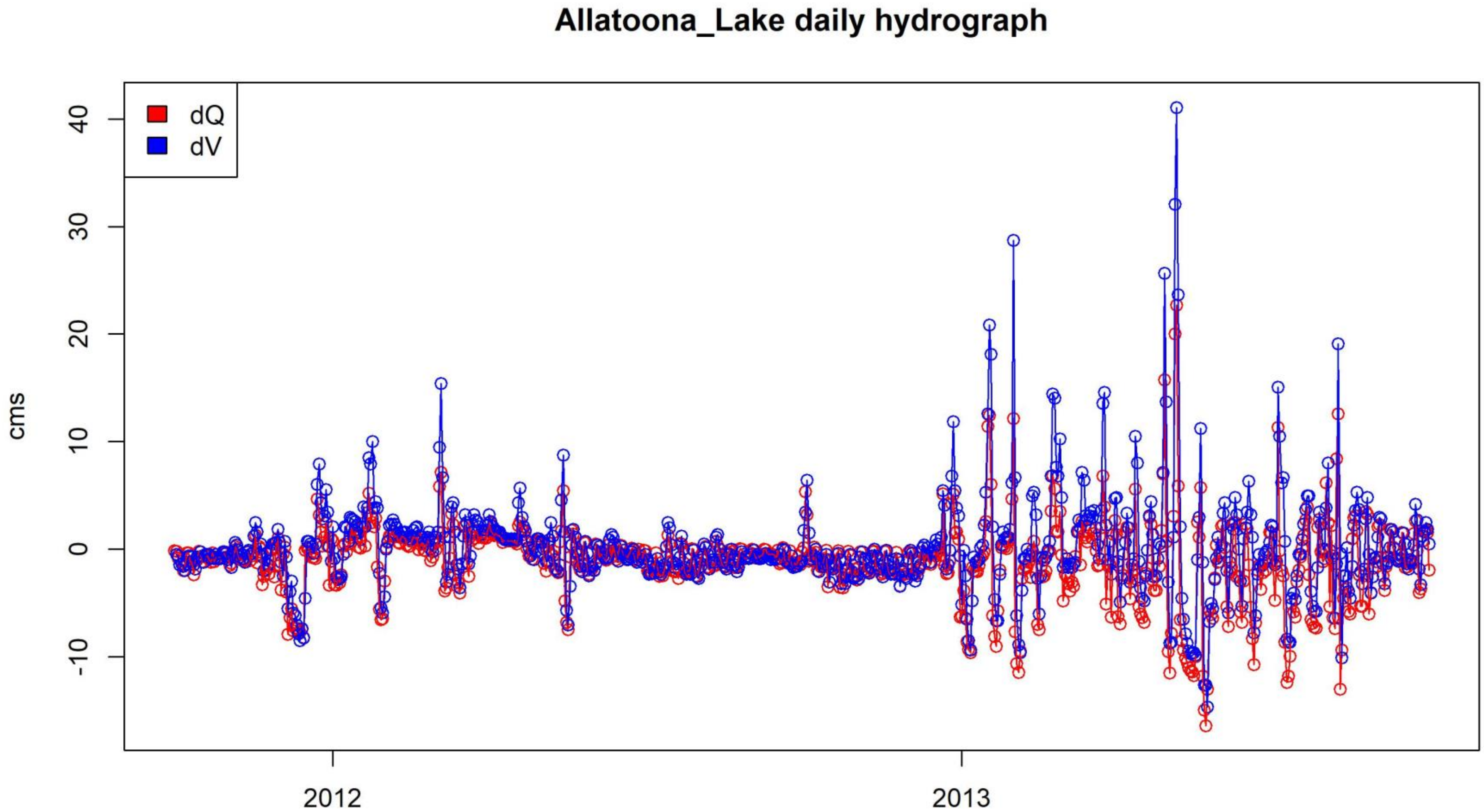


Figure B.3.3 Barlett and Horseshoe Reservoir

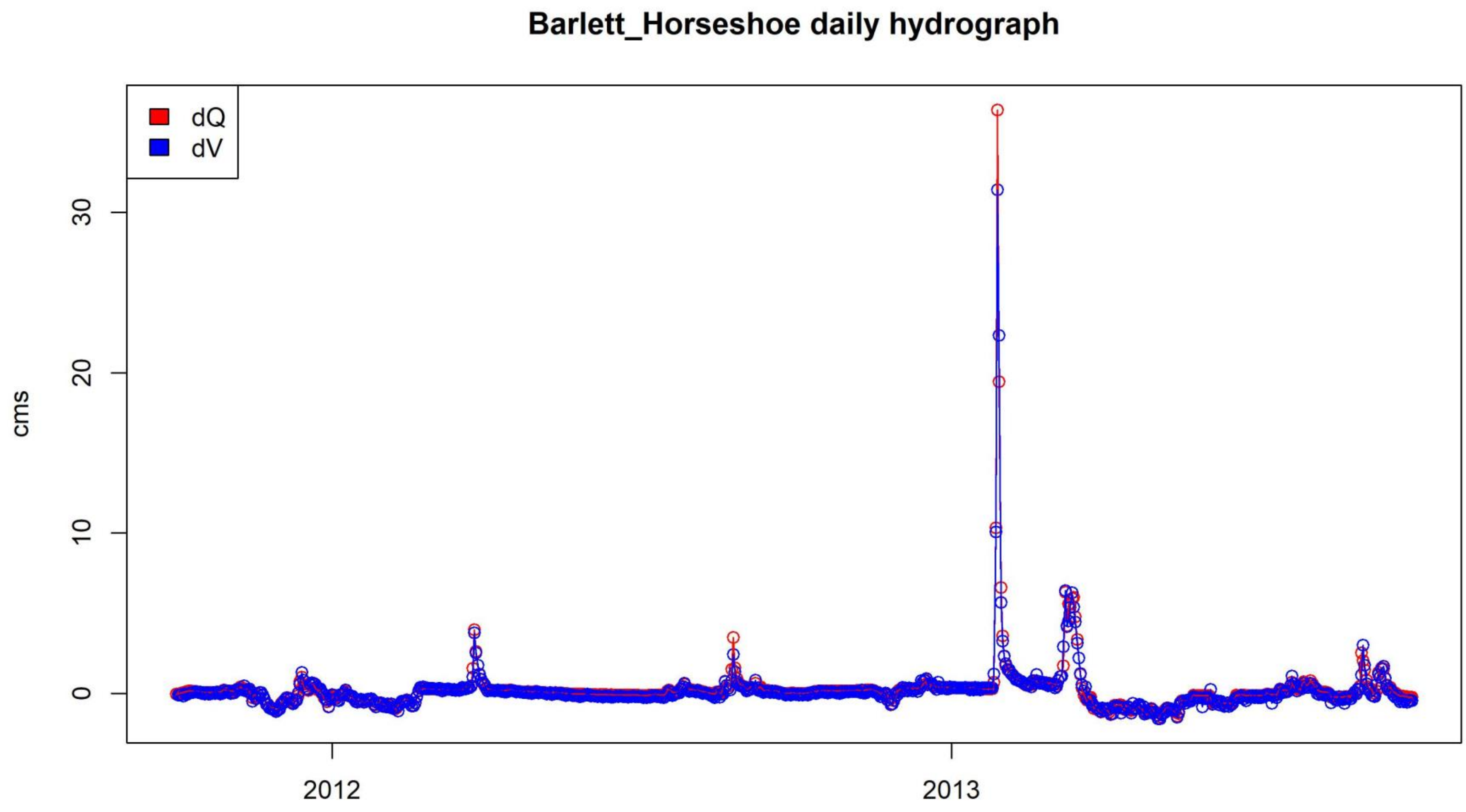


Figure B.3.4 Brantly Lake

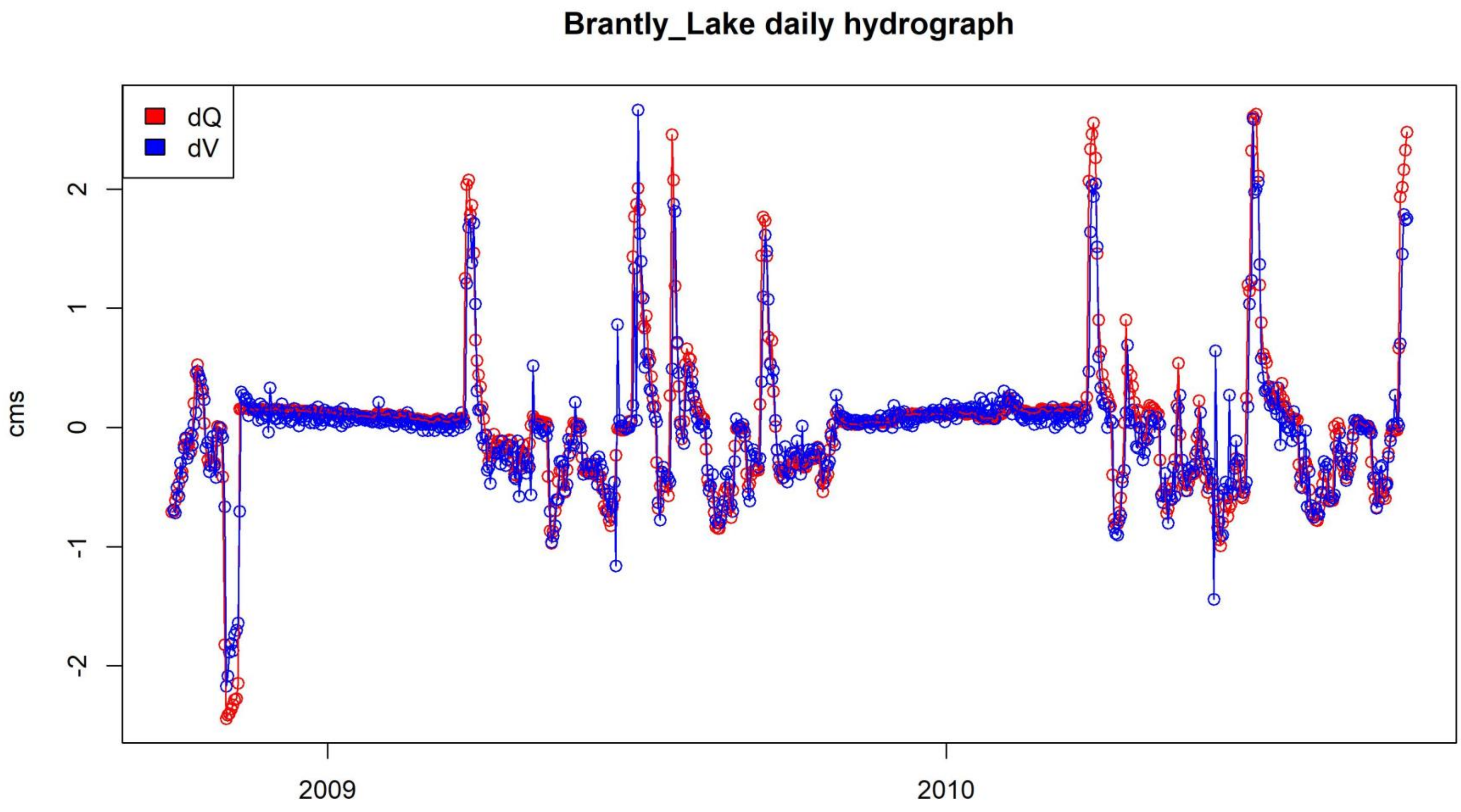


Figure B.3.5 CJ Strike Reservoir

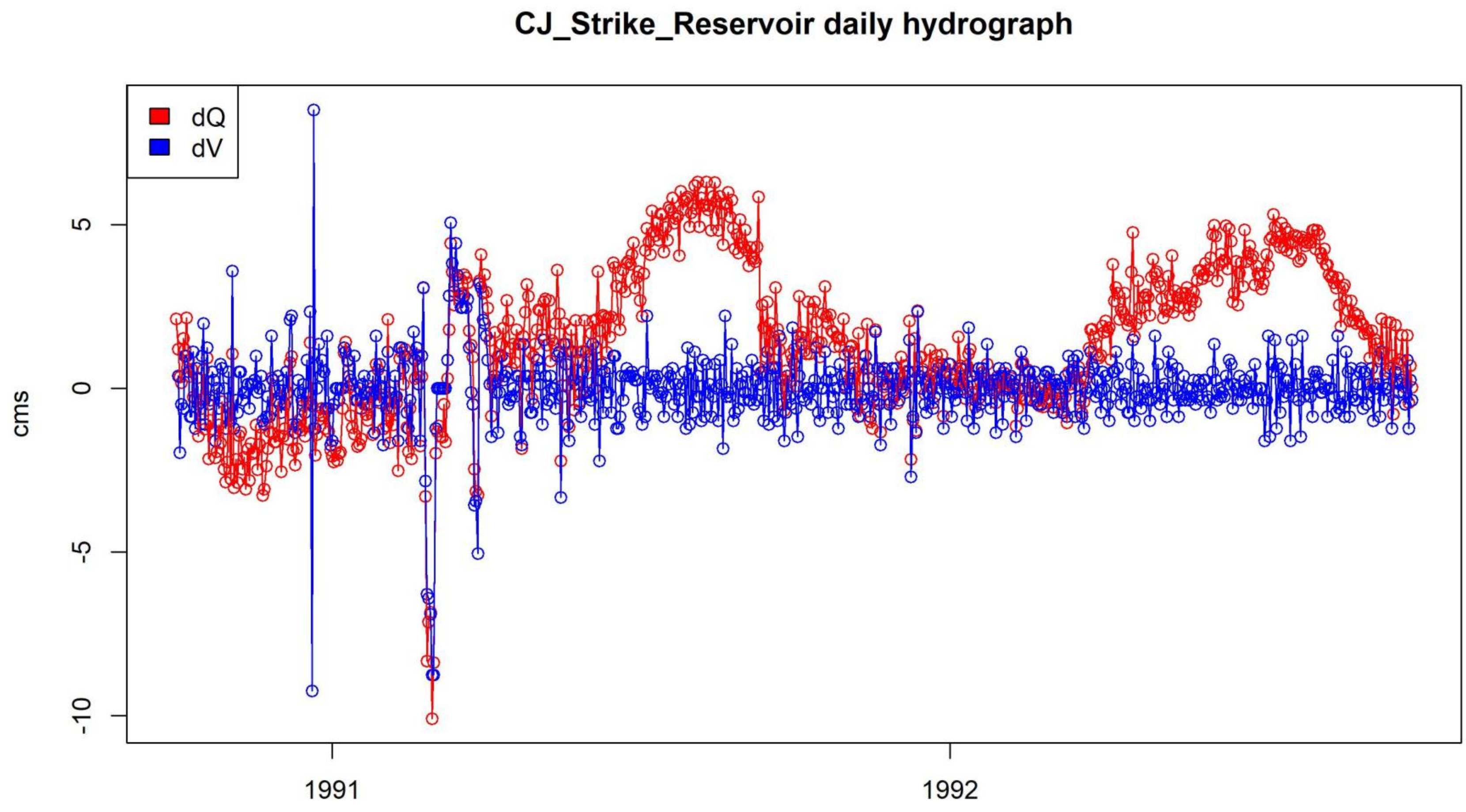


Figure B.3.6 Cochiti Lake

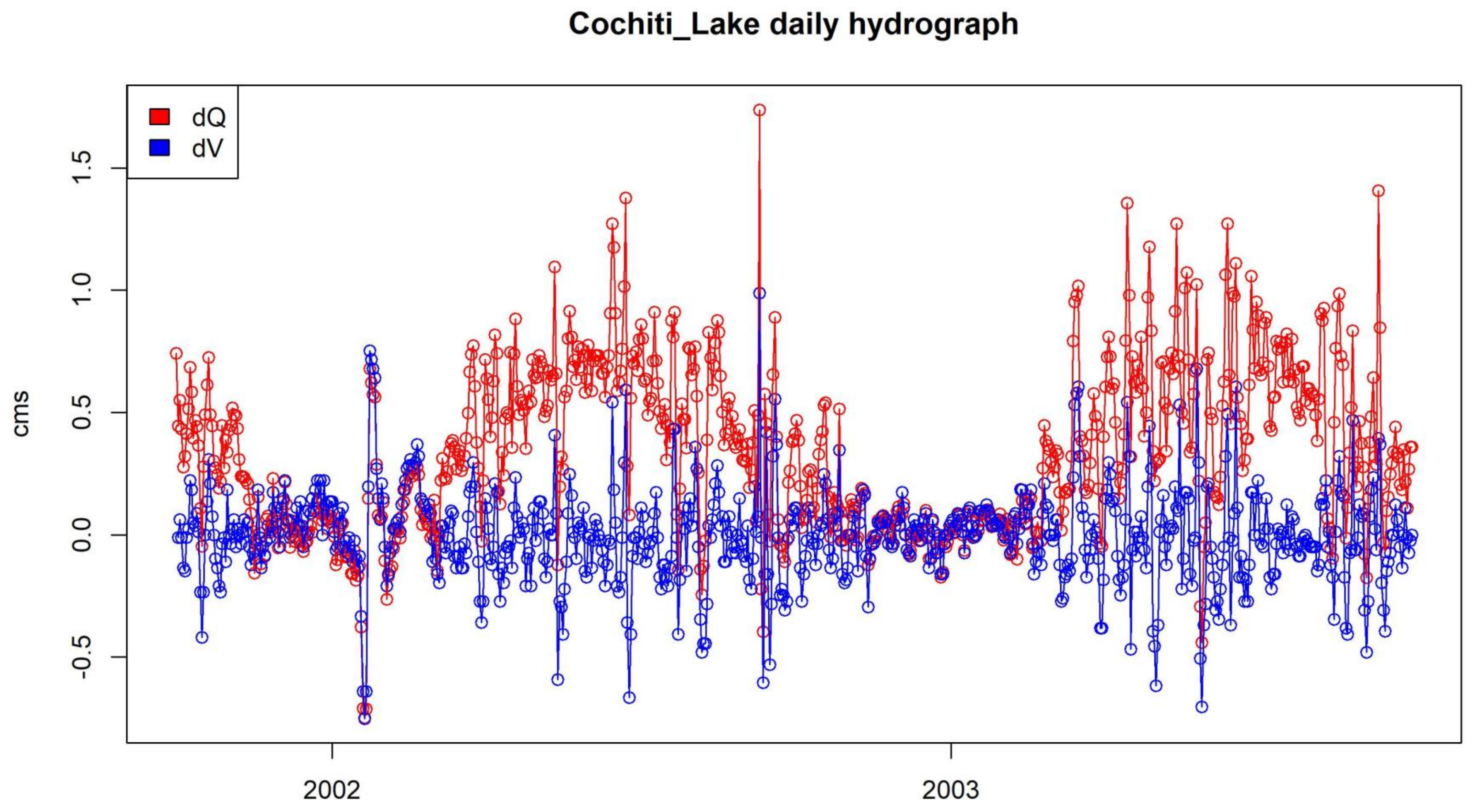


Figure B.3.7 Dworshak Reservoir

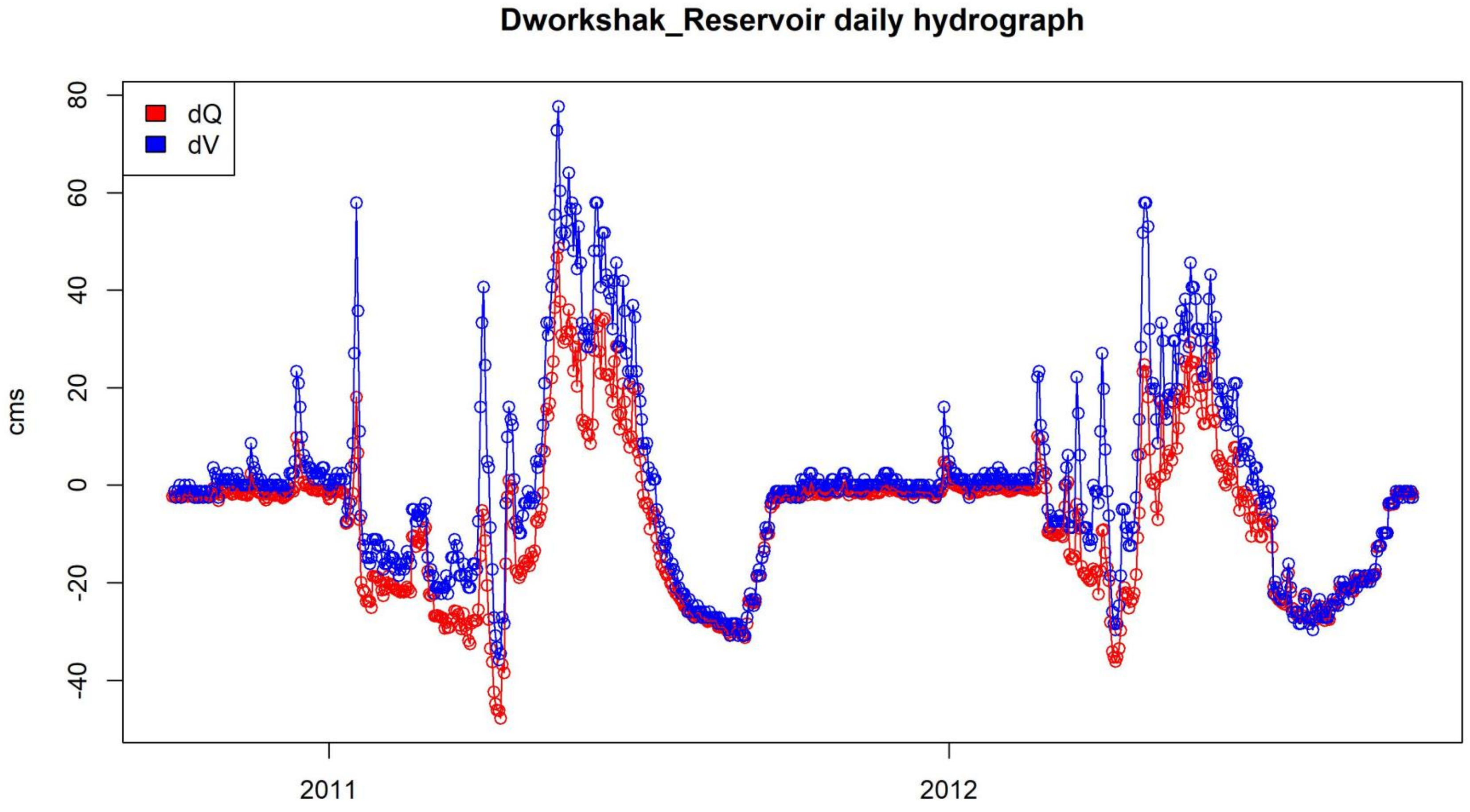


Figure B.3.8 Harlan County Lake

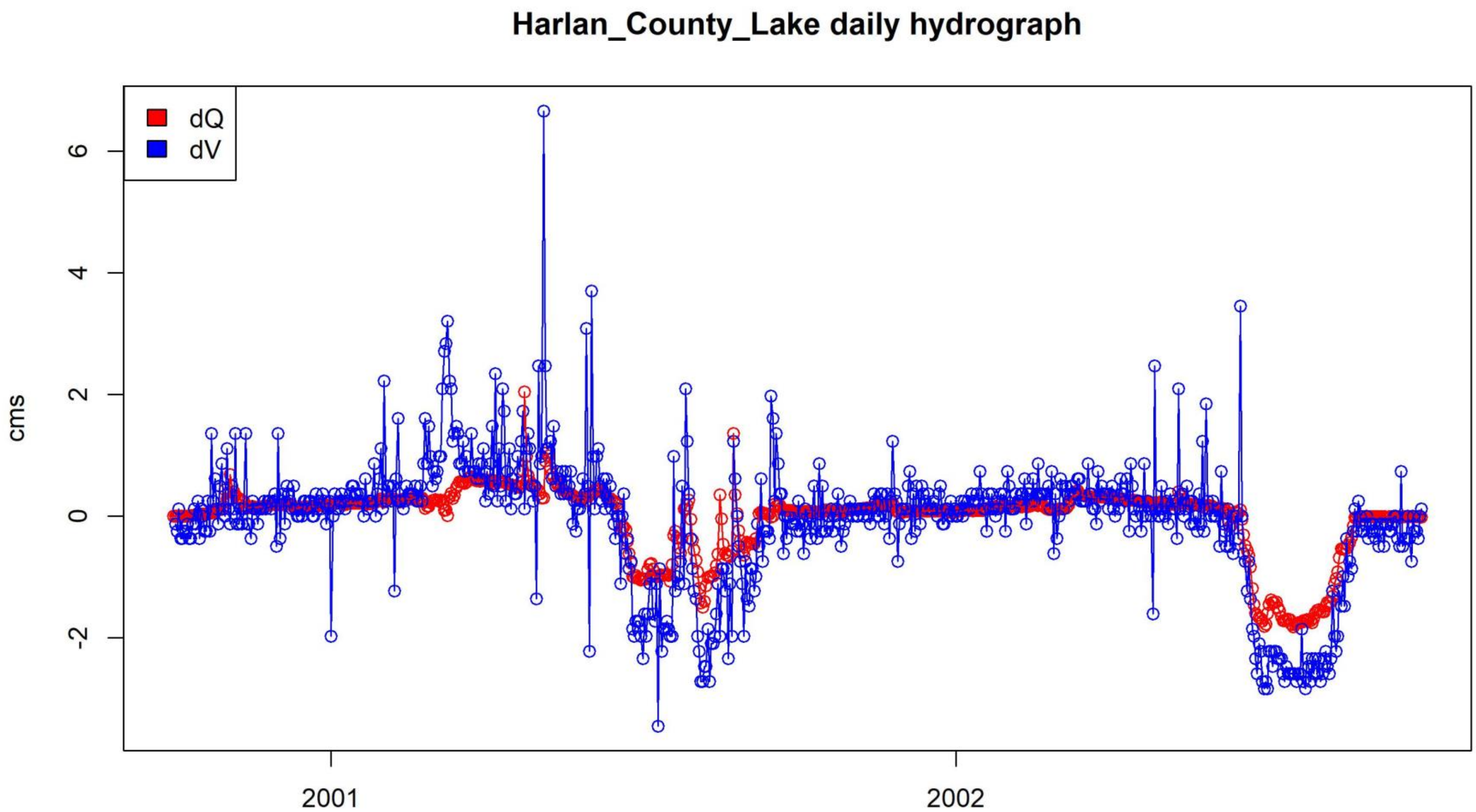


Figure B.3.9 John Martin Reservoir

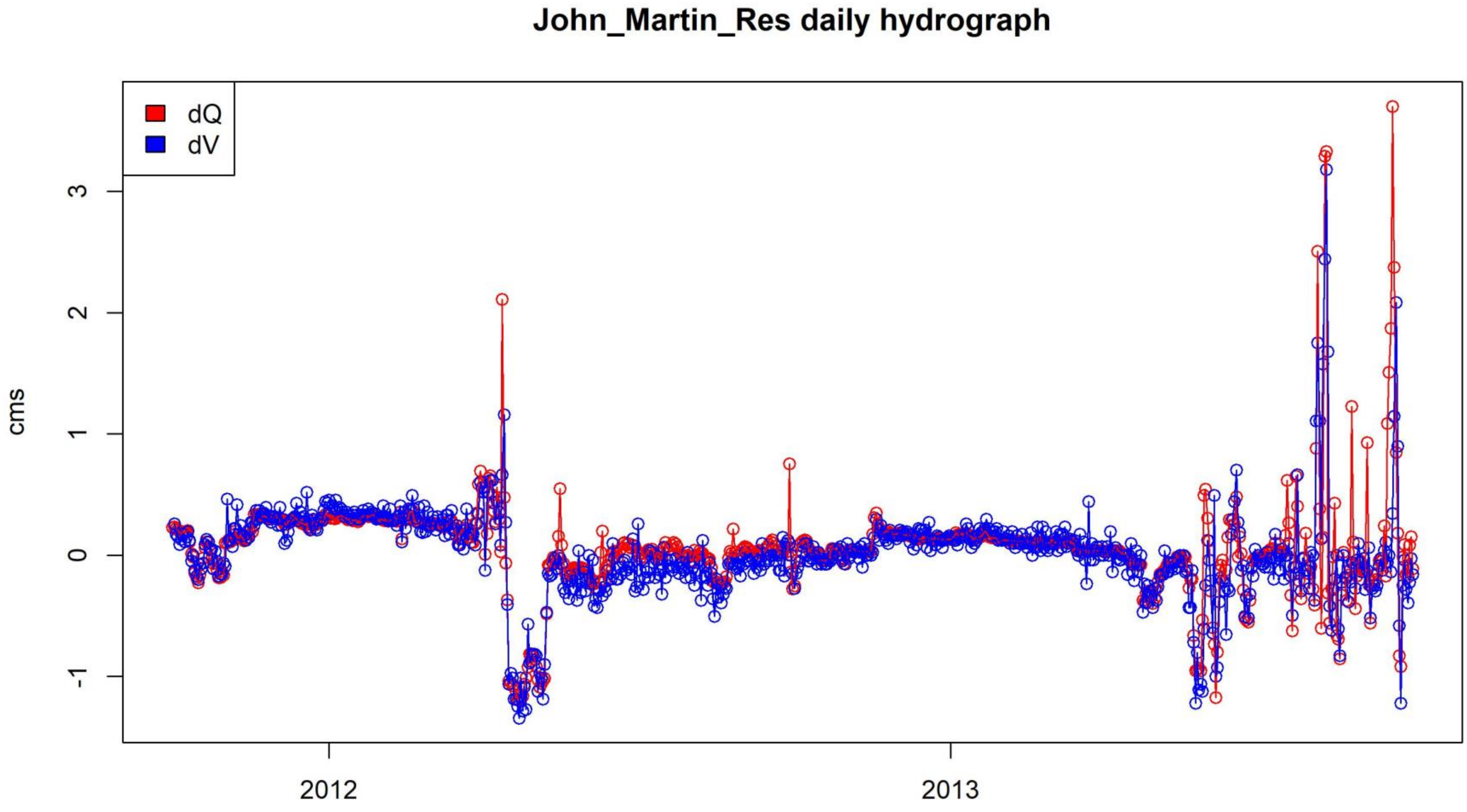


Figure B.3.10 Lake Granbury

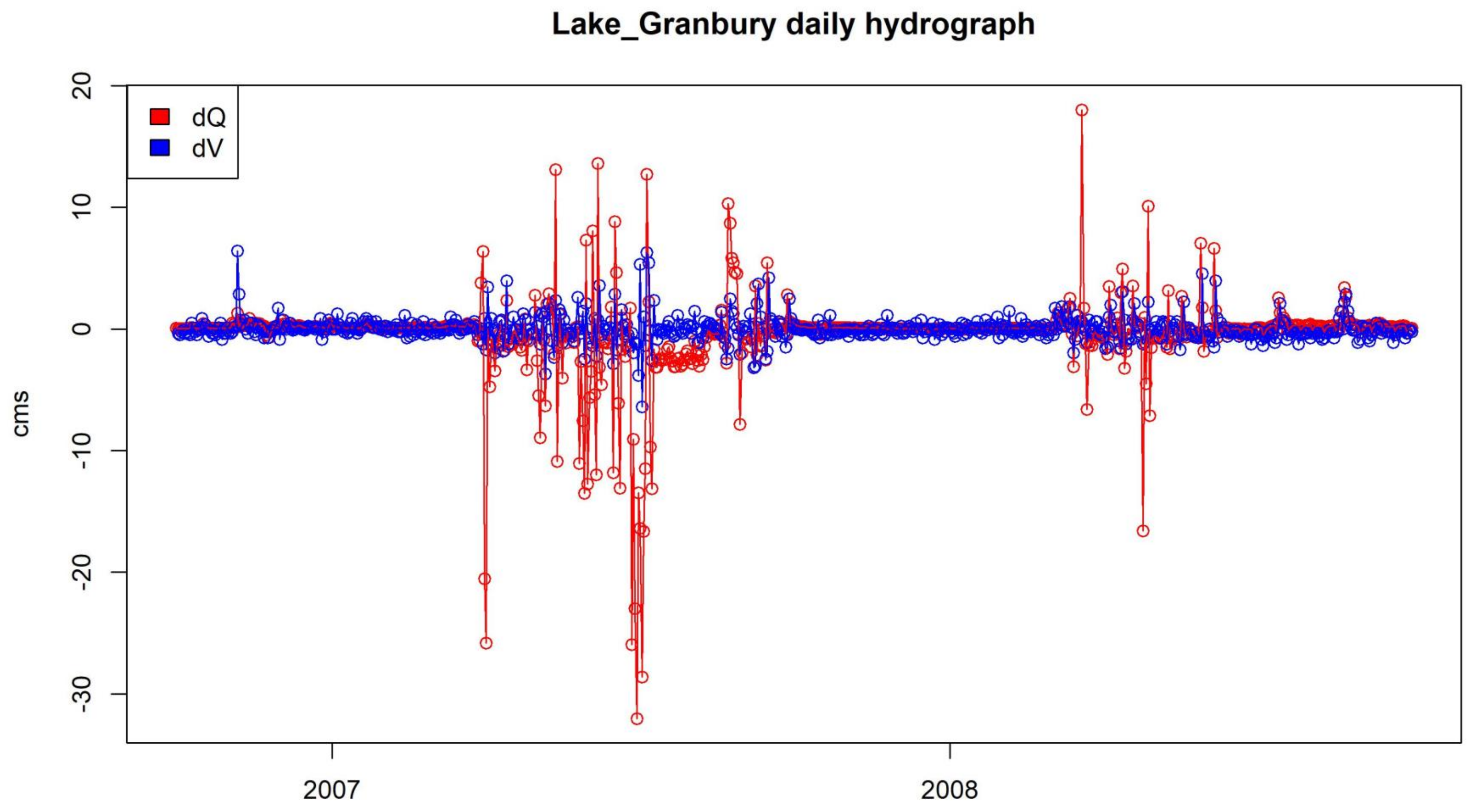


Figure B.3.11 Lake Havasu

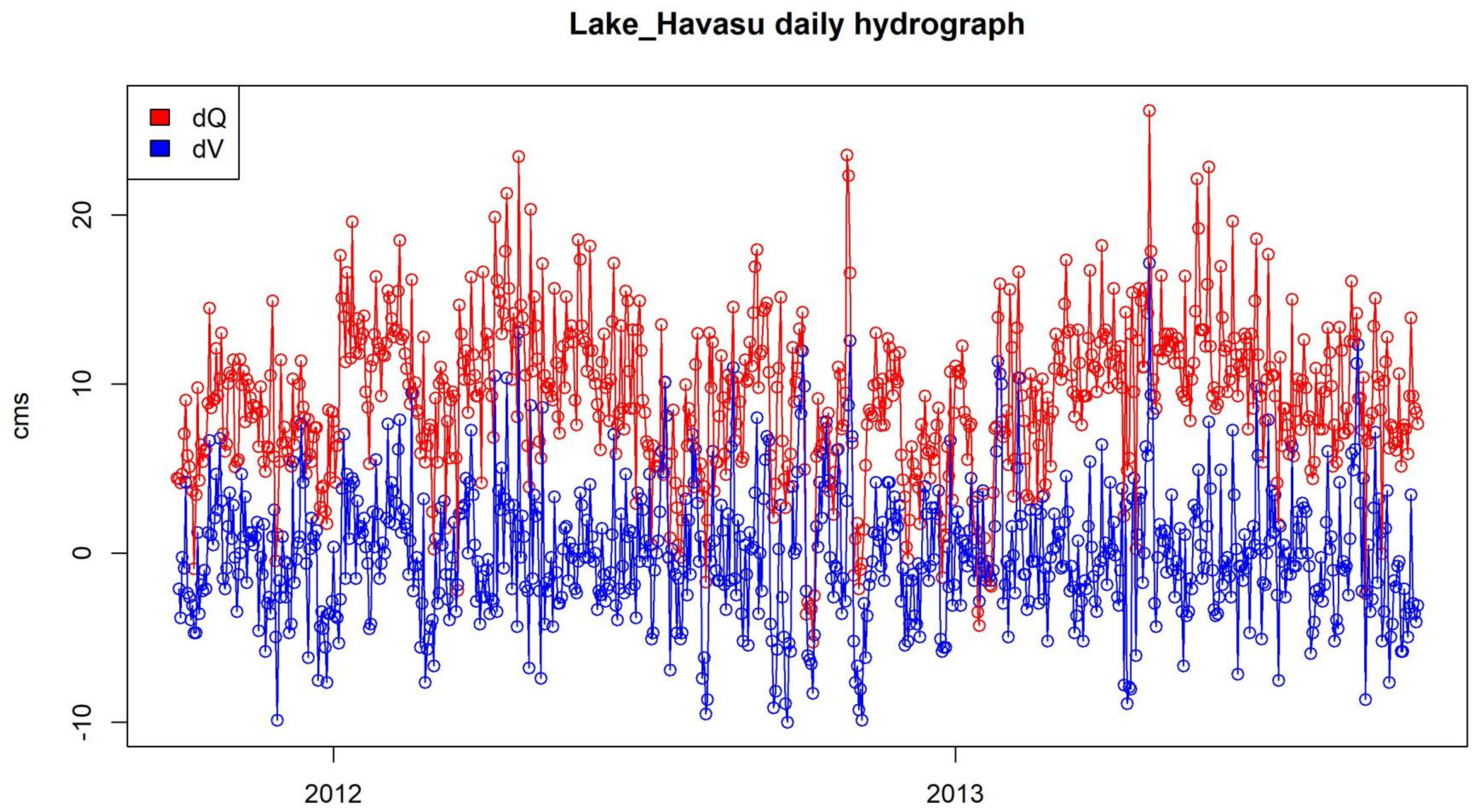


Figure B.3.12 Lake Lanier

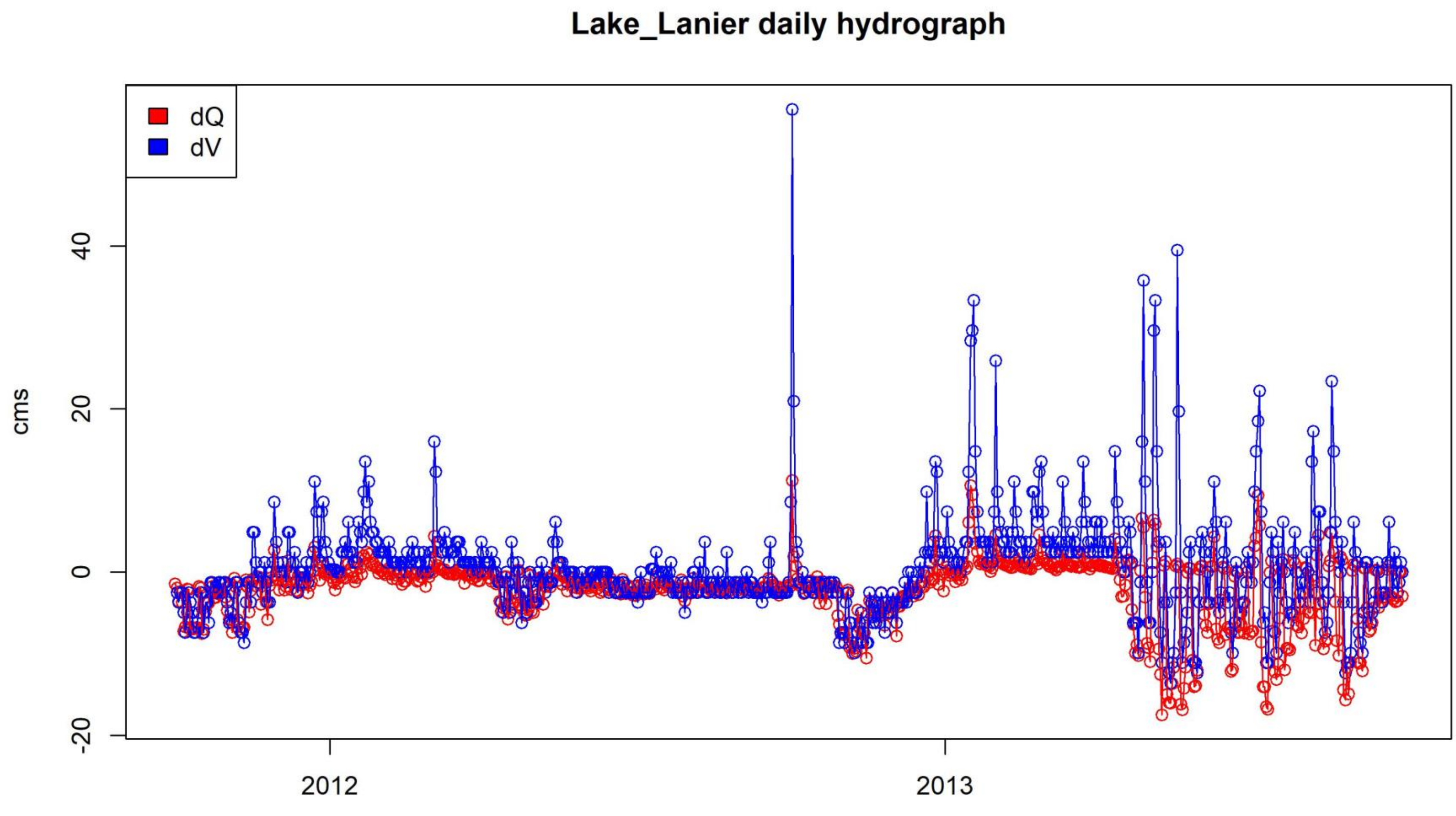


Figure B.3.13 Lake Livingston

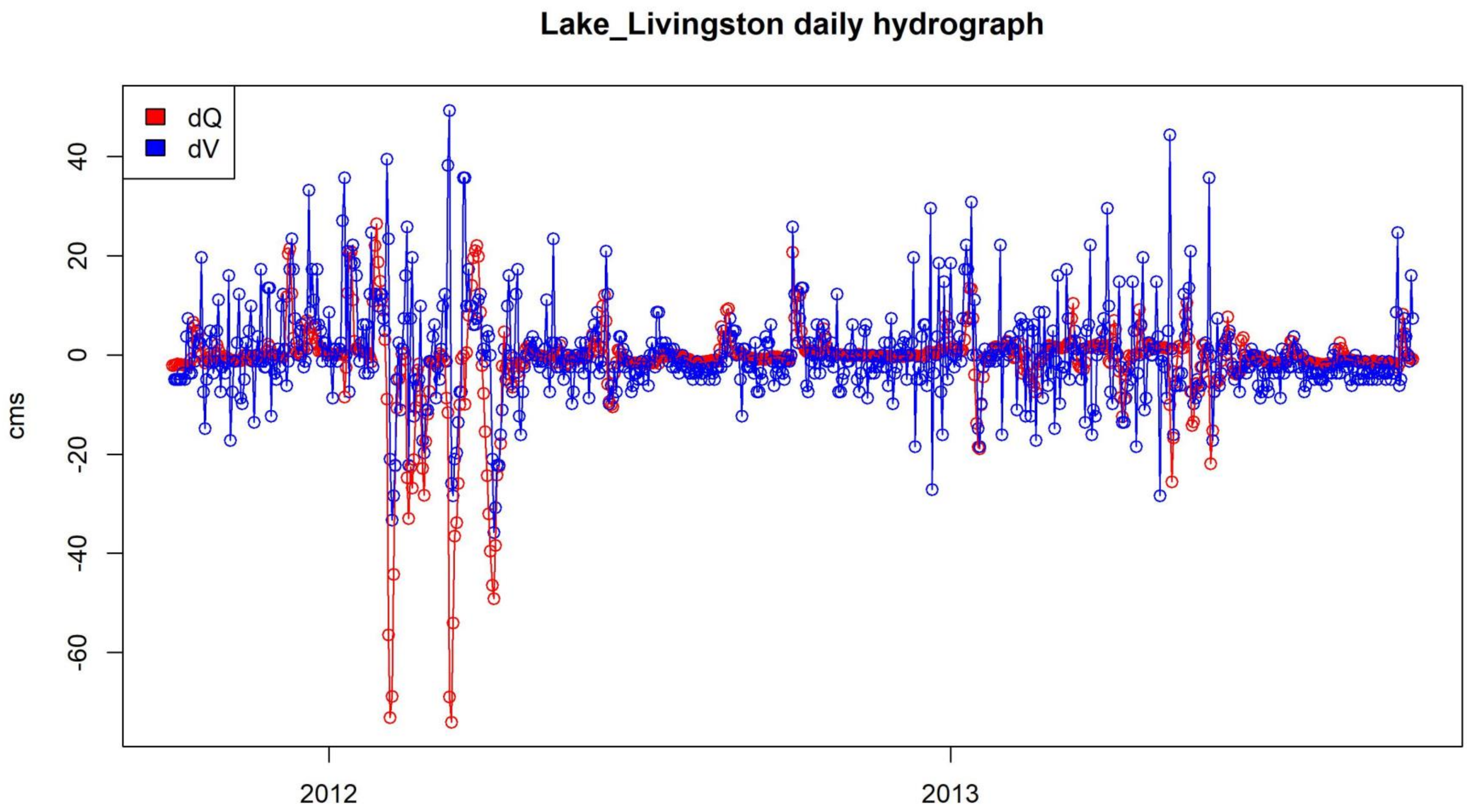


Figure B.3.14 Lake Mohave

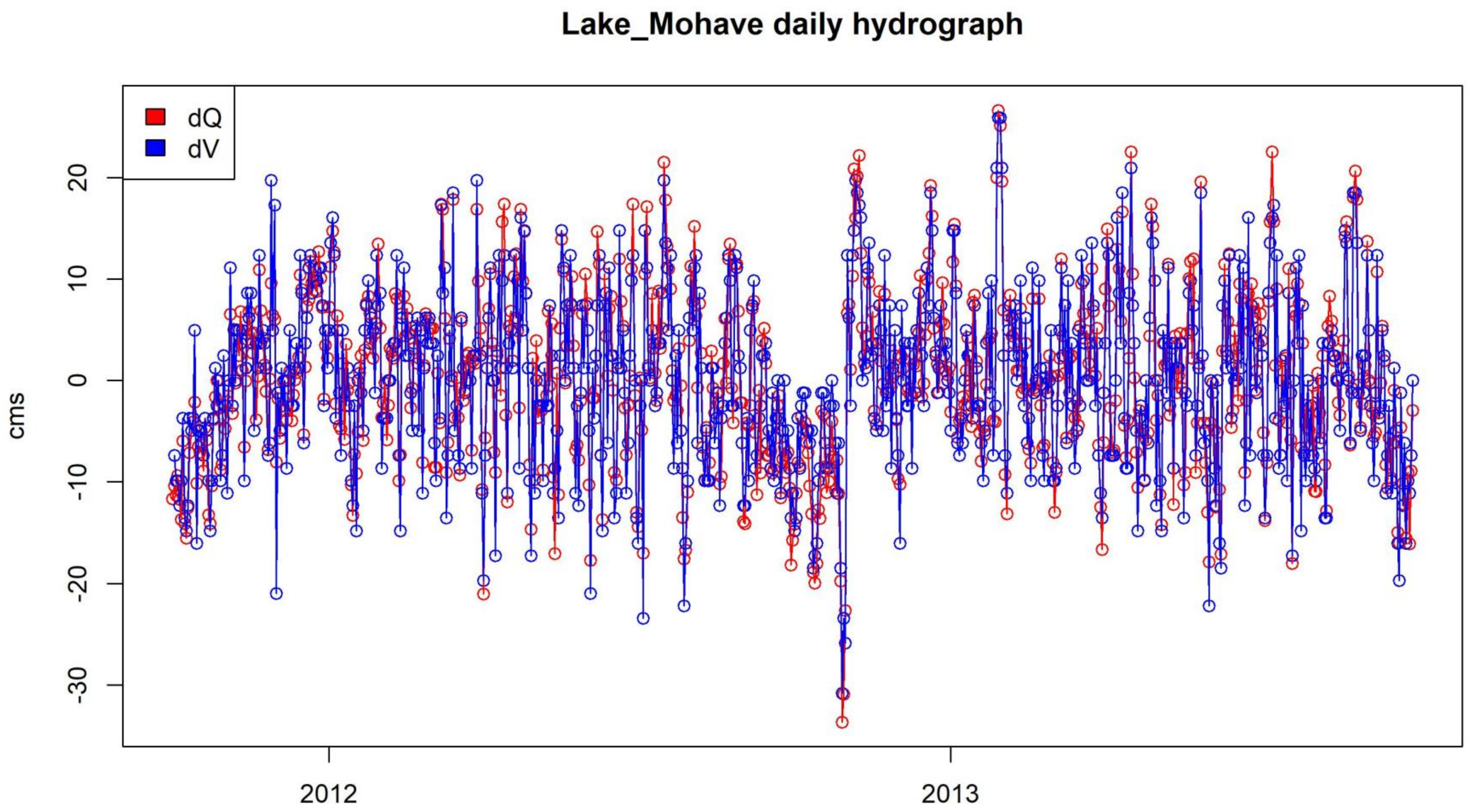


Figure B.3.15 Lake Red Rock

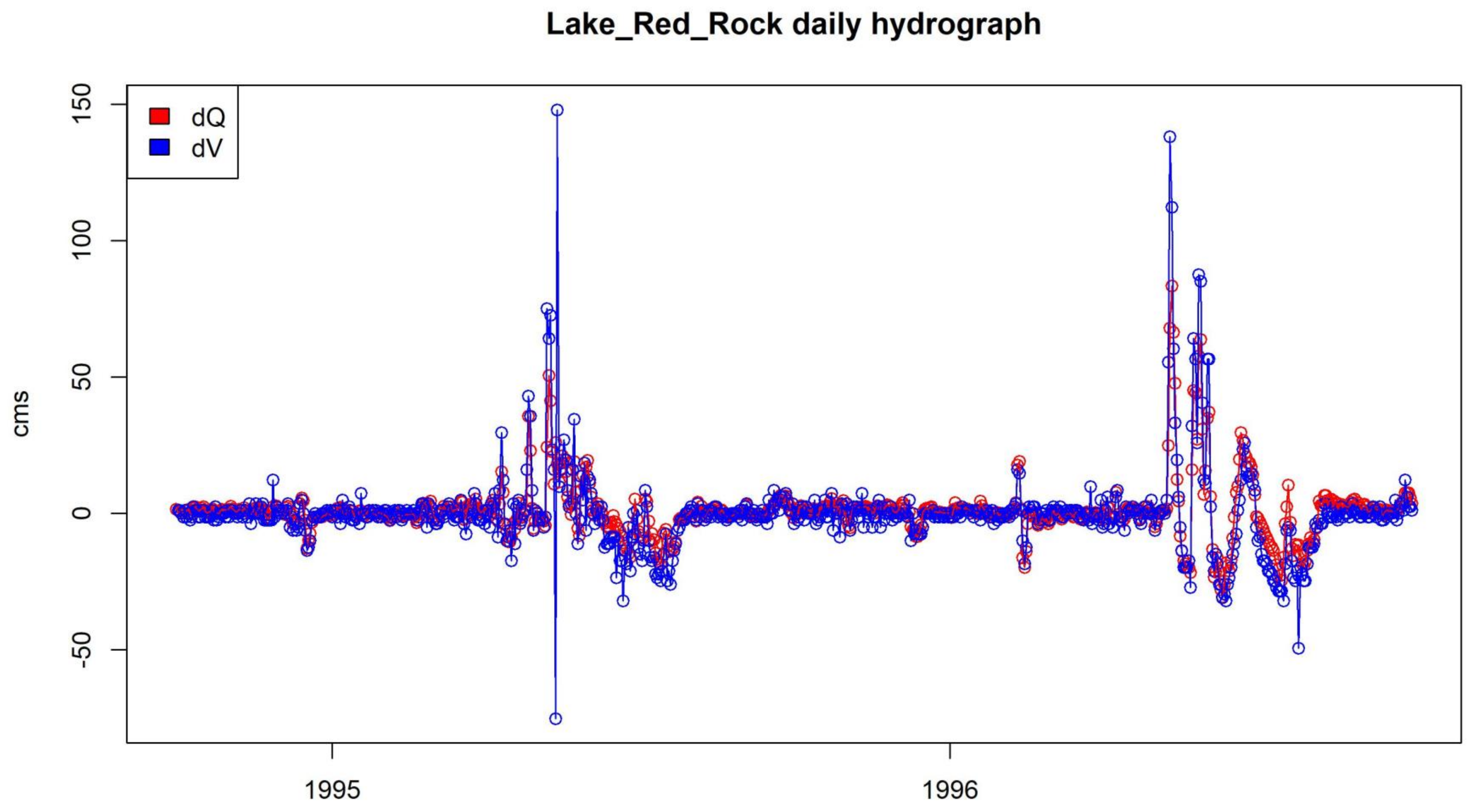


Figure B.3.16 Lake Seminole

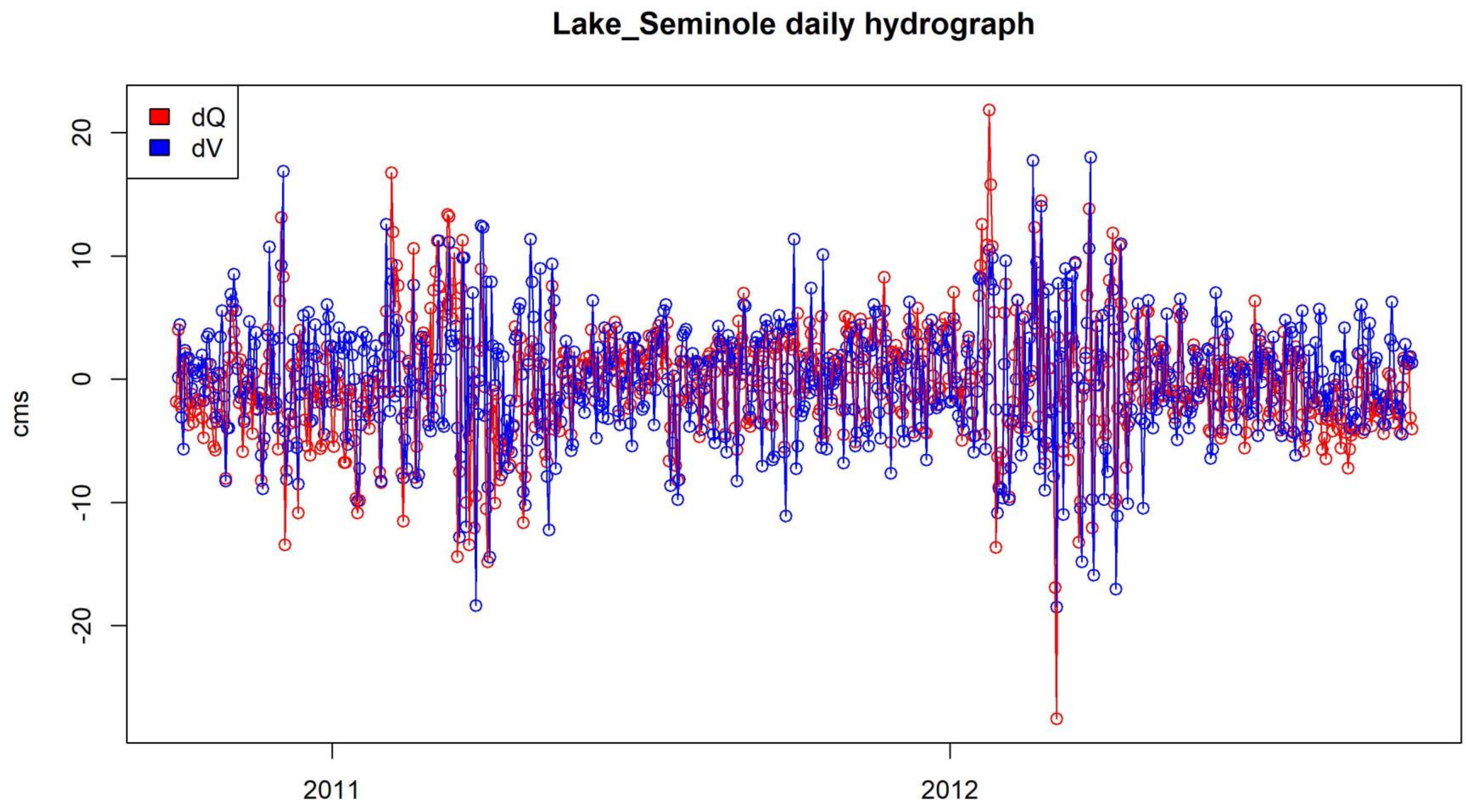


Figure B.3.17 Lake Taxoma

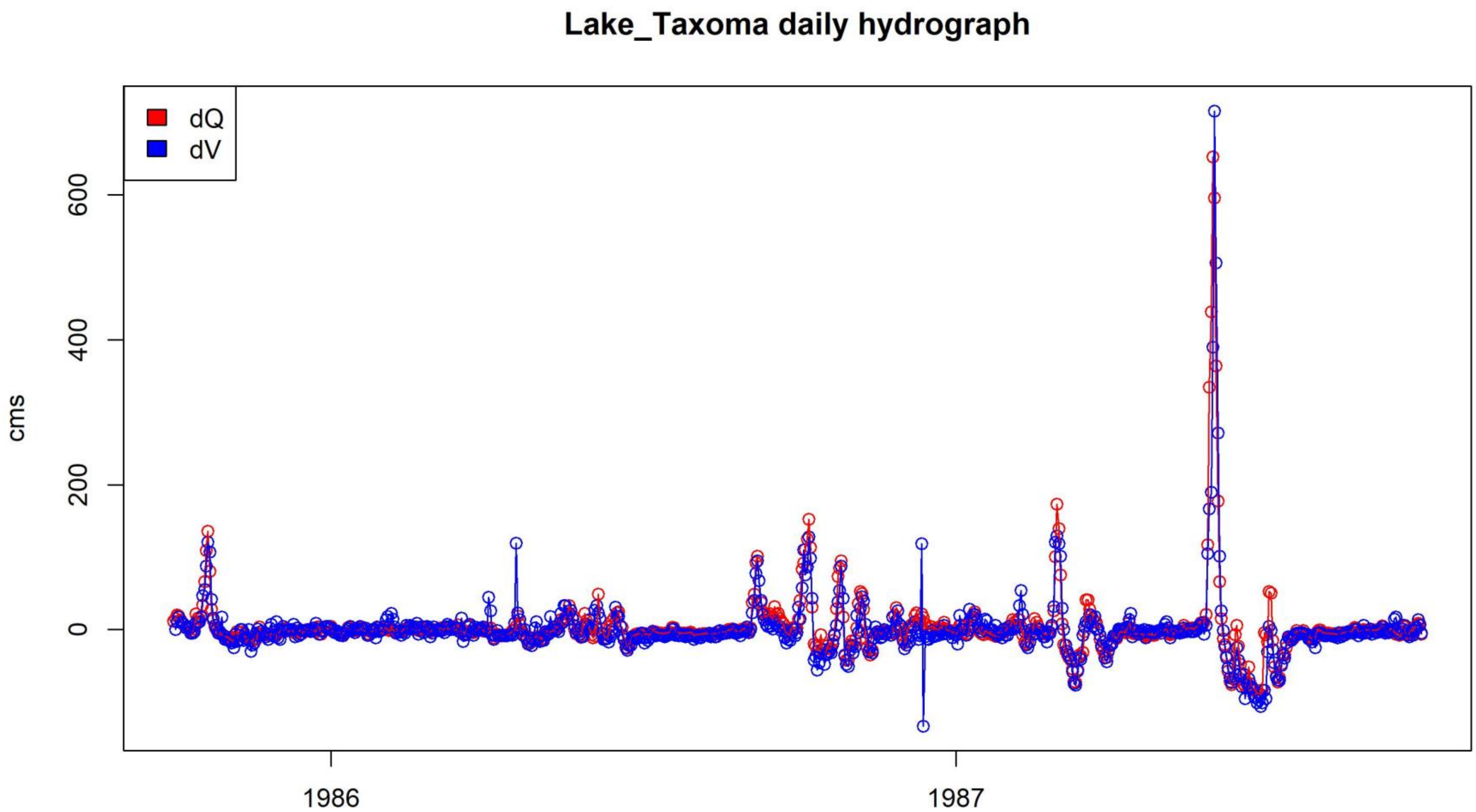


Figure B.3.18 Milford Lake

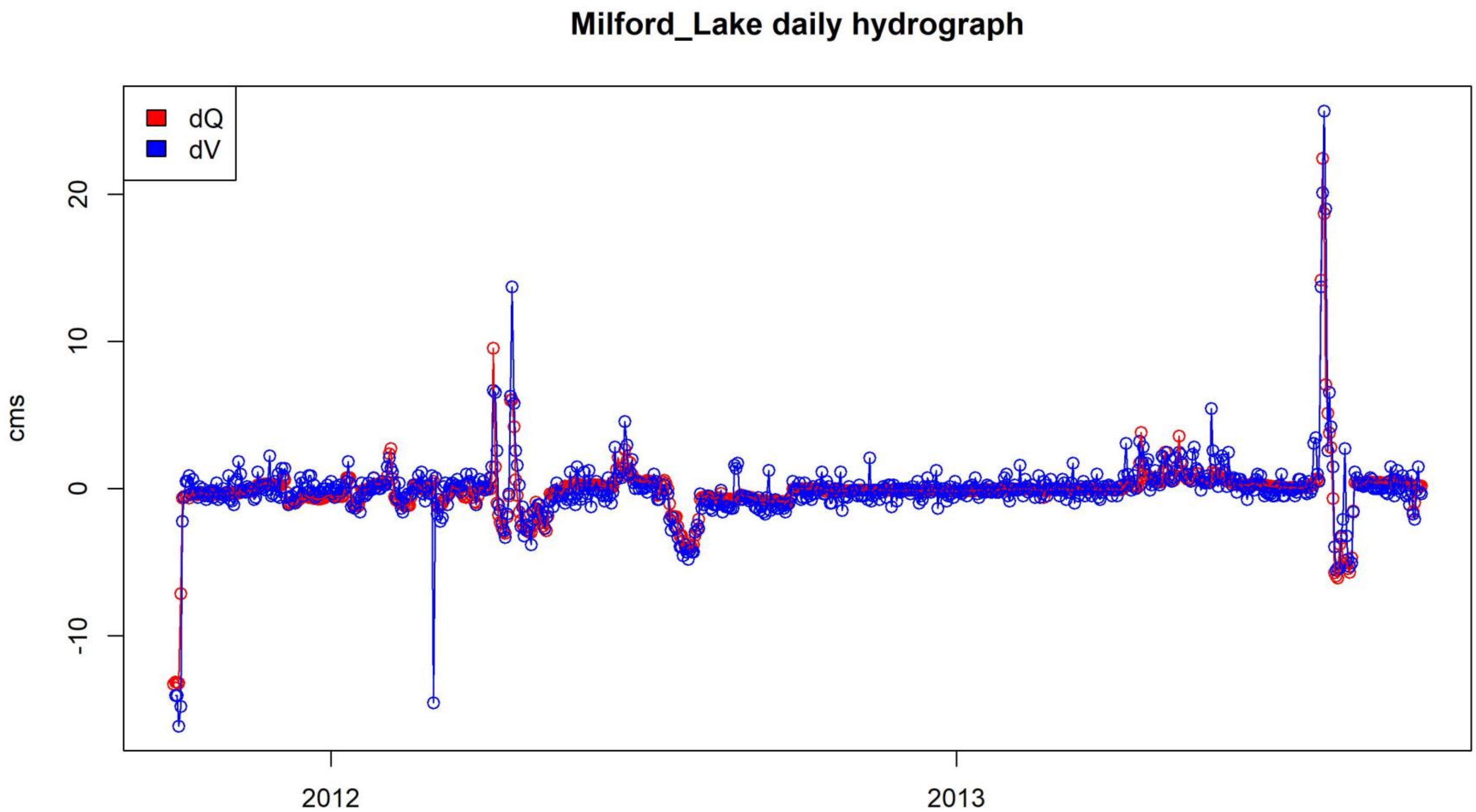


Figure B.3.19 Navajo Lake

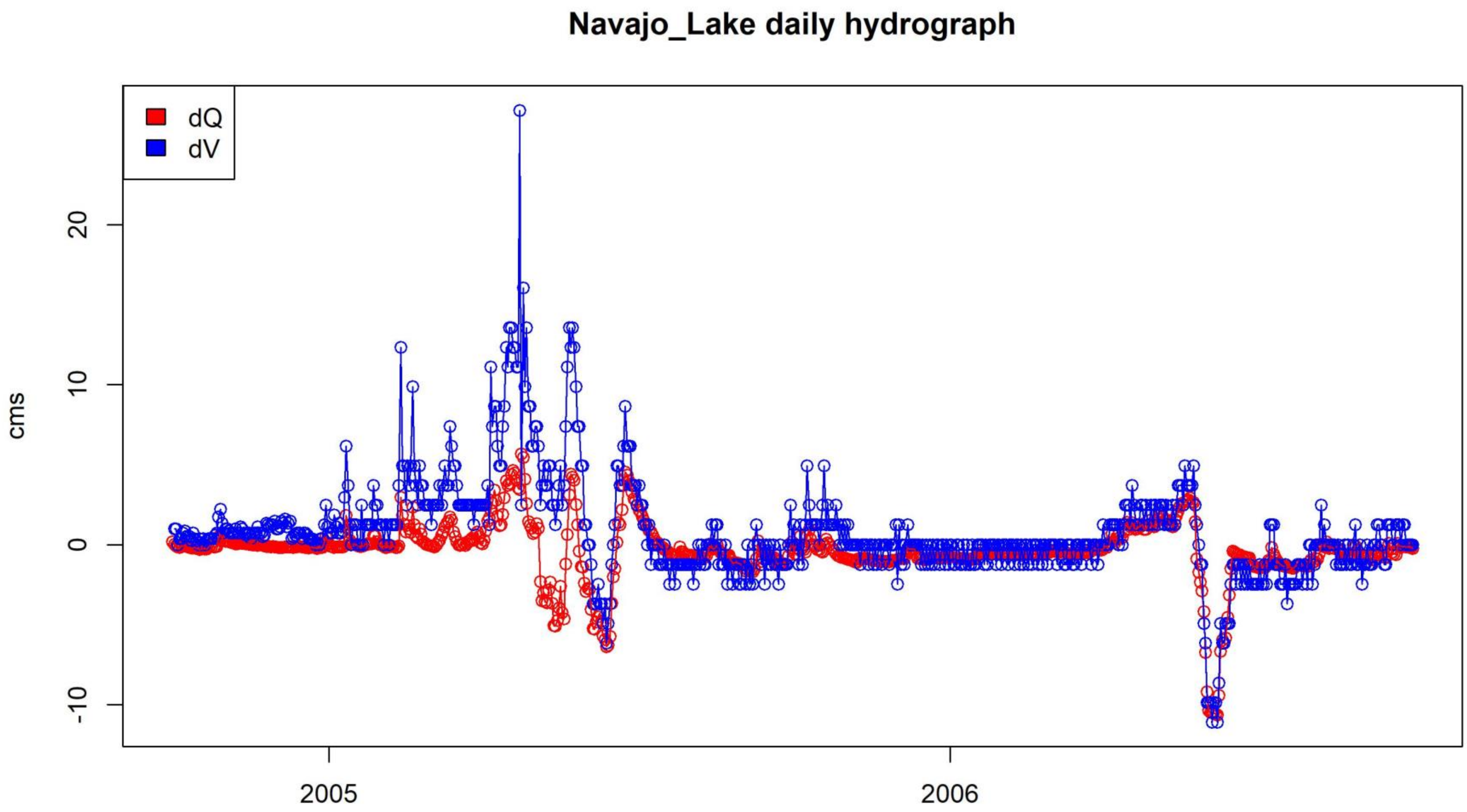


Figure B.3.20 Possum Kingdom Lake

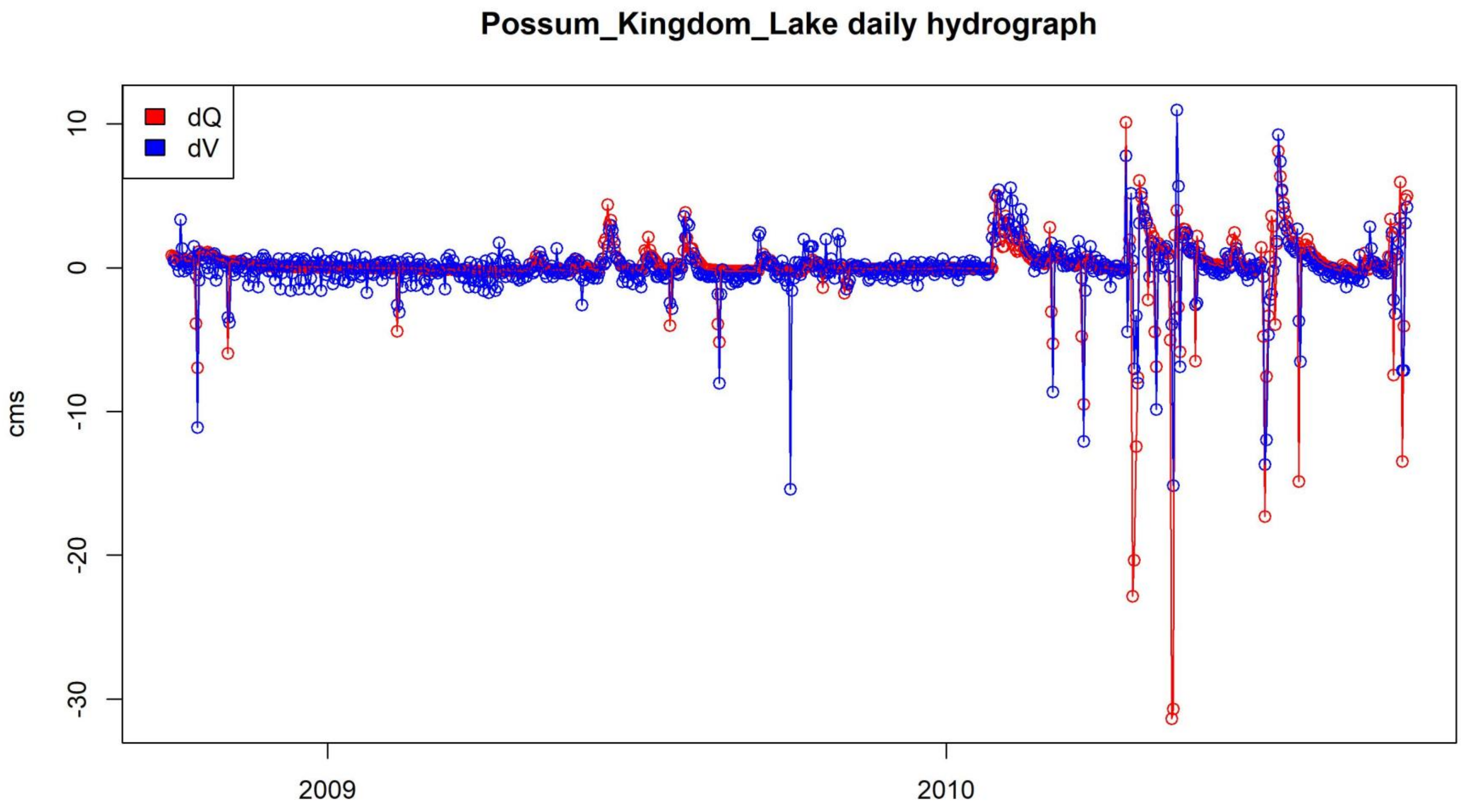


Figure B.3.21 Shasta Lake

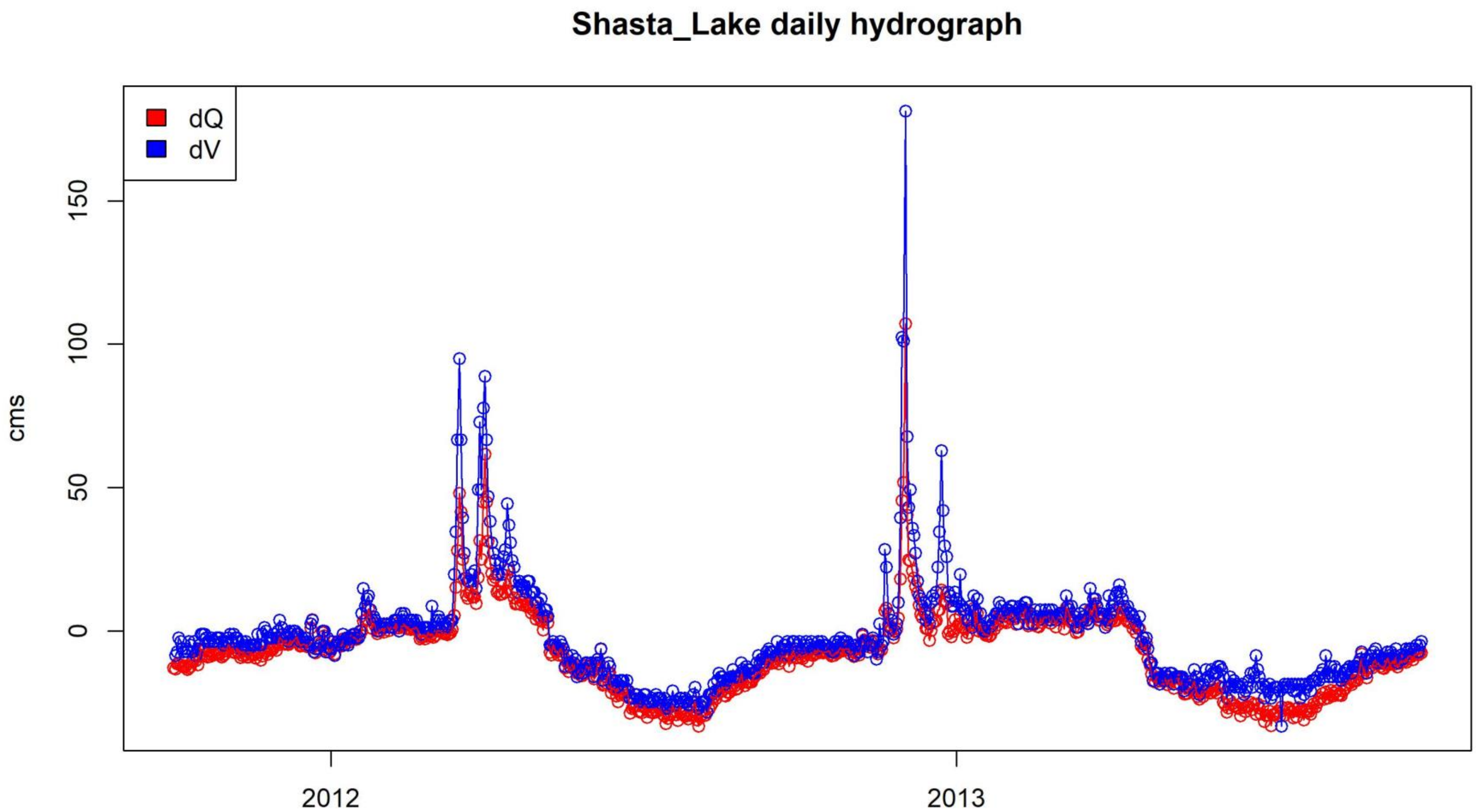


Figure B.3.22 Toledo Bend Lake

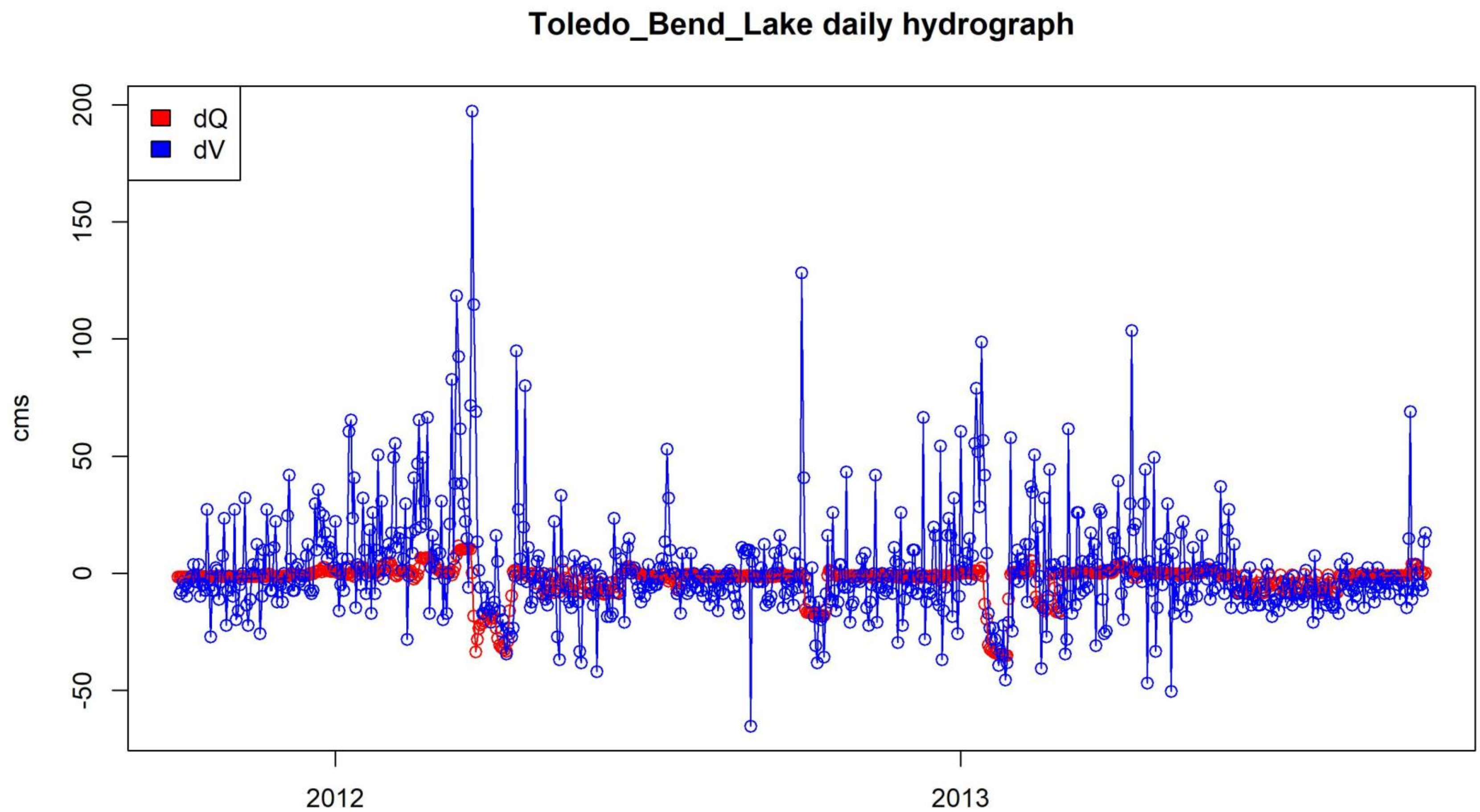


Figure B.3.23 Tuttle Creek Reservoir

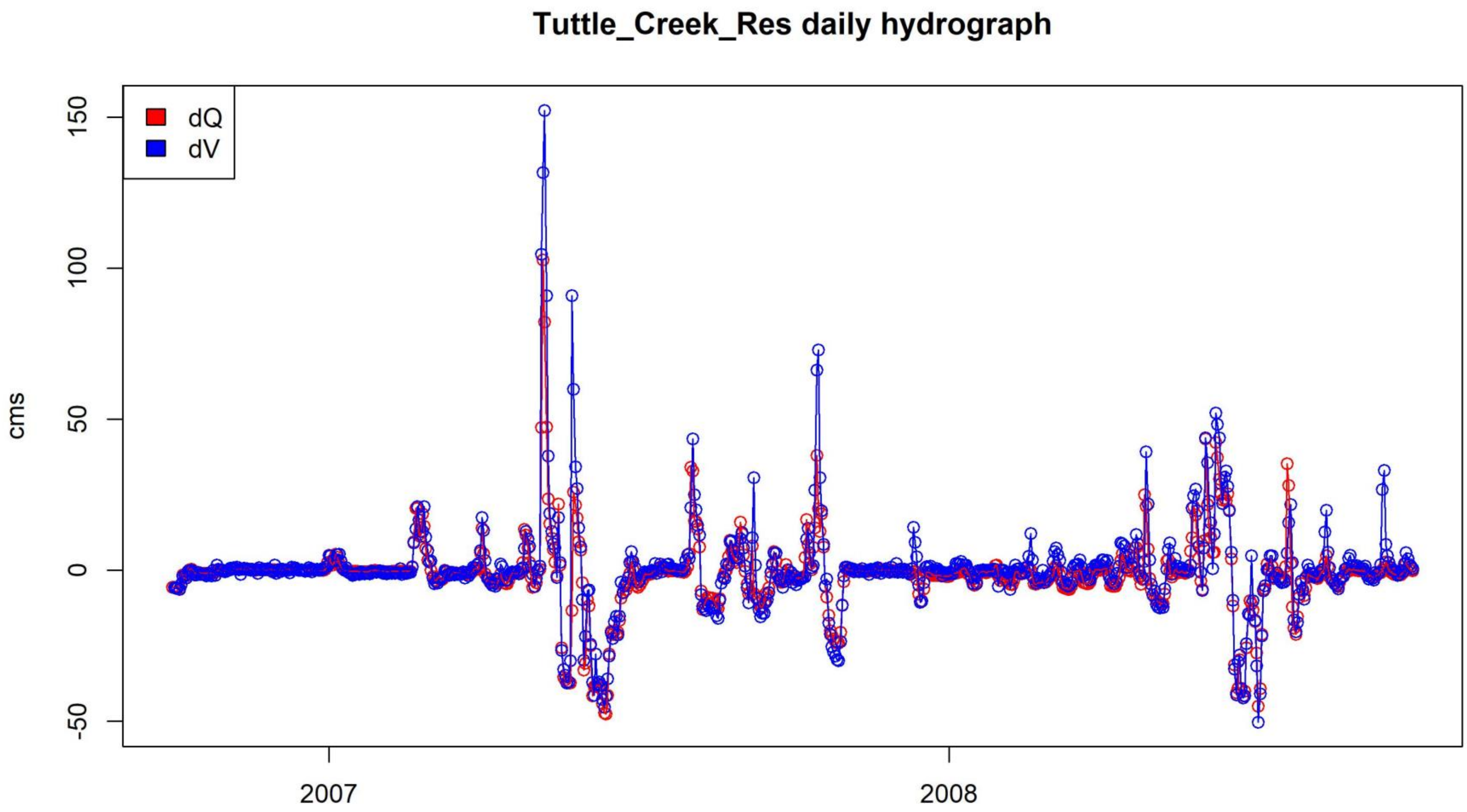


Figure B.3.24 Walter F George Reservoir

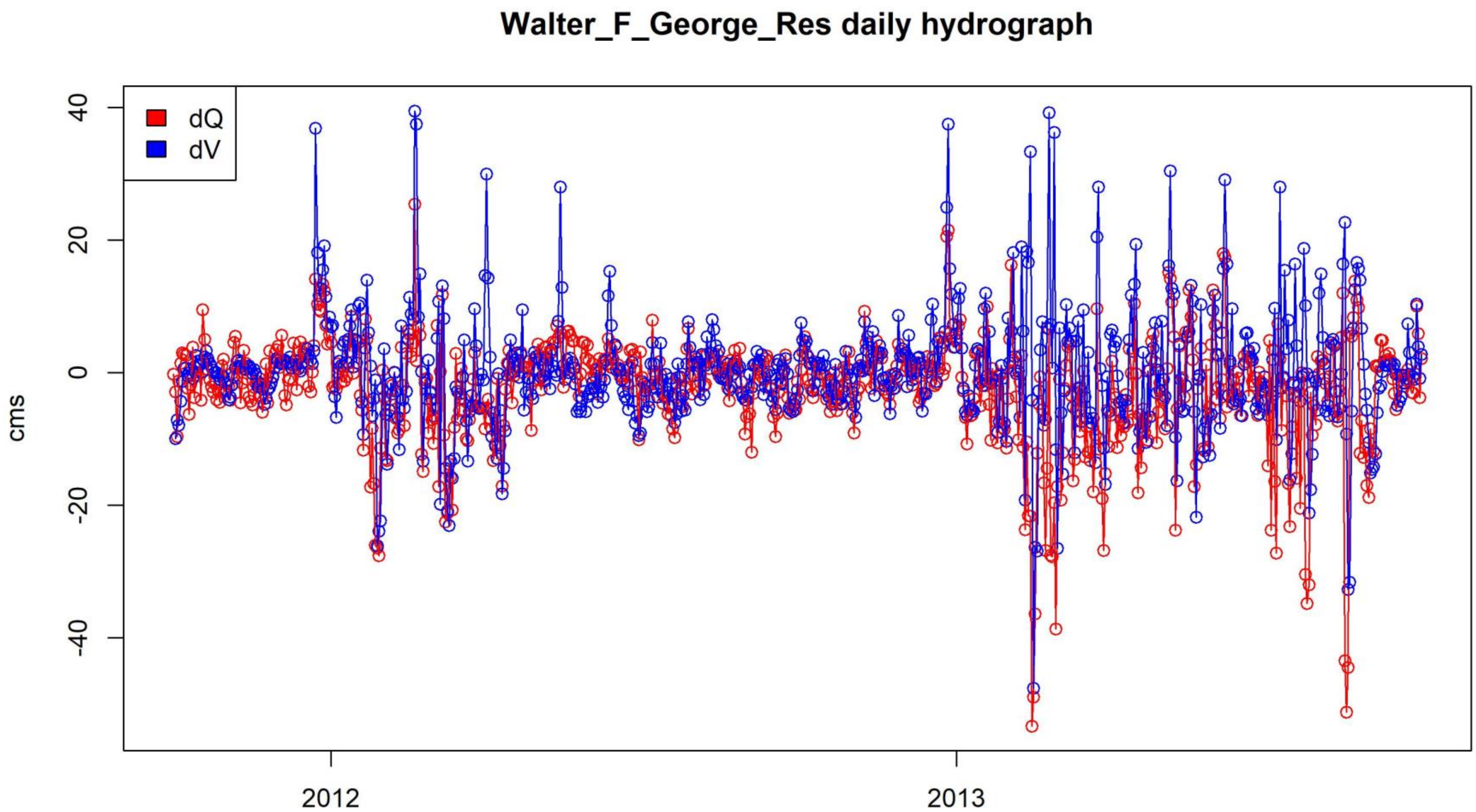


Figure B.3.25 West Point Lake

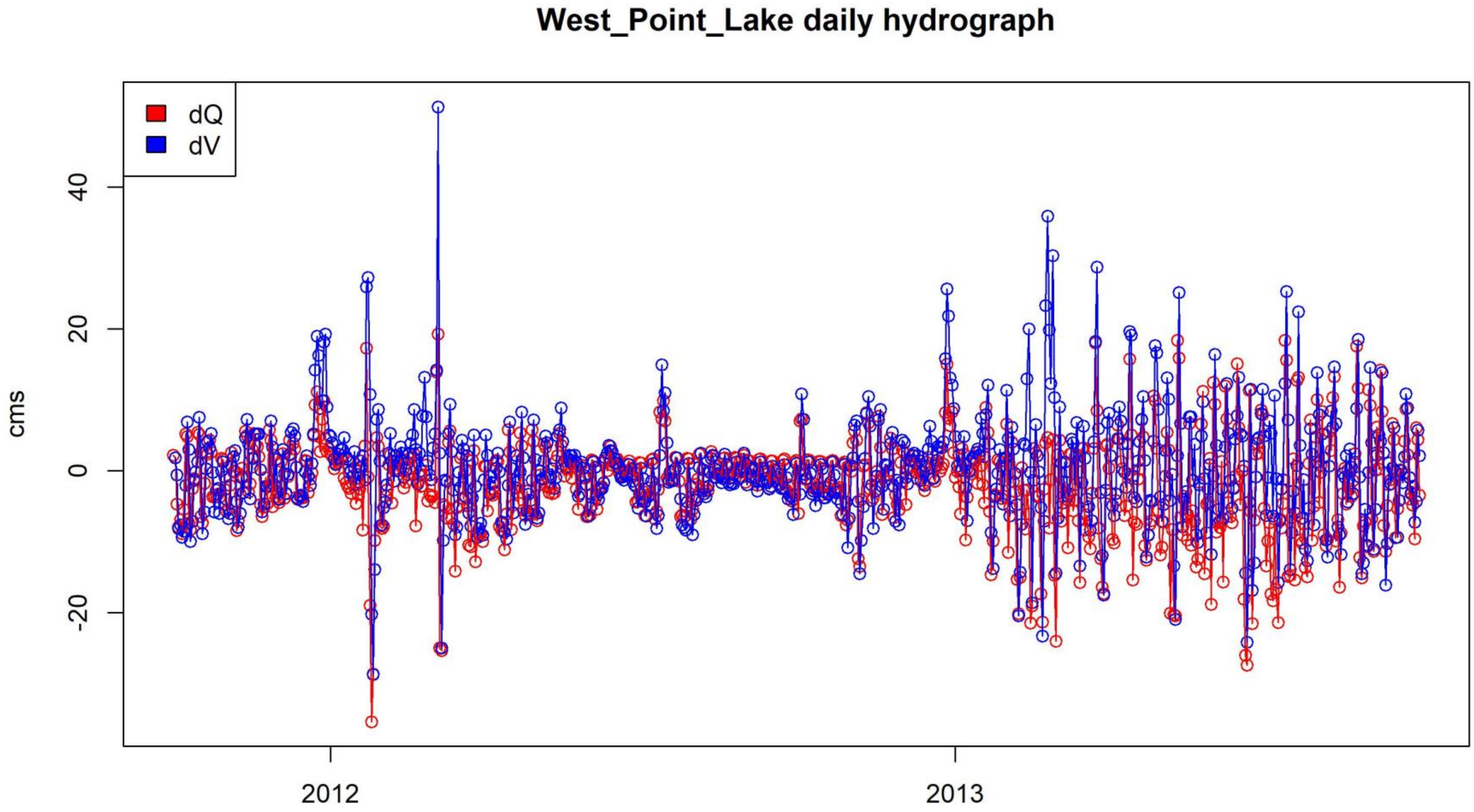


Figure B.3.26 Whitney Lake