On the interaction between bathymetry and climate in the system dynamics and preferred levels of the Great Salt Lake

Ibrahim Nourein Mohammed[1] and David G. Tarboton[1]

Received 21 May 2010; revised 24 October 2010; accepted 19 November 2010; published 17 February 2011.

[1] The Great Salt Lake is a terminal lake whose level is determined by the balance between inflows and outflows. We examine the causes for multimodality in the distributions of lake level and hence volume and area that have previously been examined from a system dynamics perspective. We focus on the role of bathymetry in the dynamics of this system and show that some of the modes that are observed and that represent preferred system states are attributable to features of the bathymetry described using the topographic area-volume relationship. Being a terminal lake, the only "outflow" is evaporation, which depends directly on lake area, which adjusts (with stochastic fluctuations) to balance inflows. Where the topographic side slopes are relatively flat and the lake goes through a large change in area for a given increase in level, the outflux evaporation will go through a corresponding large change, tending to stabilize the lake at that point. Conversely, where topographic side slopes are relatively steep, the stabilizing effect is diminished. This effect was quantified using the derivative of the lake area – lake volume function determined from U.S. Geological Survey topographic and bathymetric surveys. We show how some of the observed modes that represent preferred lake volume states are attributable to peaks in this area-volume derivative, while a complete description of the observed distribution of lake volume requires combining the bathymetry represented by the area-volume derivative with a multimodal lake area distribution that may be connected to multimodality in the aggregate hydrological forcing.

Citation: Mohammed, I. N. and D. G. Tarboton (2011), On the interaction between bathymetry and climate in the system dynamics and preferred levels of the Great Salt Lake, *Water Resour. Res., 47*, W02525, doi:10.1029/2010WR009561.

1. Introduction

[2] The level of closed basin lakes, such as the Great Salt Lake (GSL), is sensitive to the balance between inflows and outflows. Moreover, lake level is sensitive to long-term climatic fluctuations that integrate out high-frequency variability. Previous work [*Lall and Mann*, 1995; *Lall et al.*, 1996] has explored and developed an empirical understanding of the role of climatic variability in the dynamics of the GSL of Utah. *Sangoyomi et al.* [1996] studied the possibility that variations in the volume of the GSL may be described as a low-dimensional nonlinear dynamical system. In addition, *Sangoyomi* [1993] has shown that the GSL volume time series reveals significant interannual and interdecadal fluctuations, which are related to regional climatic variability and are important for understanding and forecasting drought and the long-term availability of water. *Sangoyomi* [1993] also noted that the periods of annual rise and fall of the GSL volume in which the lake rose during the normal decline or fell within the annual rise marked the beginning of a multiyear period of rise or fall, i.e., a regime transition. This prior work did not explicitly consider the interplay between lake volume and surface area, which is a major control on the evaporation outflux.

[3] Fluctuations of the GSL's level are of direct concern to industries and infrastructure along the shore, such as the Salt Lake City Airport, the Union Pacific Railroad, wastewater treatment plants, and Interstate Highway 80 [*Lall et al.*, 1996]. They are also well correlated with regional water supply conditions. During 1983–1986 the Great Salt Lake rose rapidly to its highest level in 100 years and then declined quickly. A pumping project (the West Desert Pumping Project) that cost about $60 million was initiated because of that event and removed more than 3.08 km^3 (2.5 million acre-feet) of water and 695 million tons of salt from the lake from April 1987 to June 1989, and from January 1990 to June 1992, 0.25 km^3 (200,000 acre-feet) of this pumped water and 94 million tons of salt were returned to the lake [*Wold and Waddell*, 1994; *Loving et al.*, 2000]. The following 2 decades (1990–2010) have seen concern that the GSL might be drying up. In examining such behavior we have to consider the dynamics of the GSL because high levels threaten infrastructure, while low levels put large-scale industries and environmental interests in jeopardy.

[4] Previous work [*Lall et al.*, 1996] interpreted multiple modes in the GSL volume probability density function (pdf) as potentially being due to separate attractors in the nonlinear dynamics of the system and suggestive of preferred states. We argue here that the topographic area-volume relationship in the GSL plays an important role in the system

[1]Department of Civil and Environmental Engineering, Utah State University, Logan, Utah, USA.

0043-1397/11/2010WR009561

dynamics because area is a control on the evaporation outflux, the only outflow from the system. The increase in lake area with increasing lake volume increases evaporation and has a stabilizing effect on the volume and level of the GSL. Where, for a given change in volume, the lake area goes through a large change, the outflux evaporation will go through a correspondingly large change, tending to stabilize the volume at that point. On the other hand, if for a given change in volume, the lake area only changes by a small amount, the outflux evaporation will only change by a small amount, resulting in a small stabilizing effect. These considerations suggest that modes of the lake volume distribution should coincide with peaks in the area-volume derivative.

[5] In this paper we test the hypothesis that some modes in the lake volume distribution are attributable to bathymetry. We do this by comparing peaks in the area-volume derivative with modes in the lake volume distribution. We show how some of the observed modes, which represent preferred lake volume states, are attributable to peaks in this area-volume derivative, while a complete description of the observed distribution of lake volume requires combining the bathymetry represented by the area-volume derivative with a multimodal lake area distribution that may be connected to multimodality in the aggregate climate forcing. The overall goal of this work is to better understand how both external forcing and internal feedbacks are involved in the GSL basin system dynamics.

2. Background

[6] The GSL (latitude 40.7°N–41.7°N, longitude 111.9°W–113.1°W) is located in the northeast of the Great Basin, and it is the fourth largest, perennial, closed basin lake in the world (Figure 1). The lake is shallow (average depth of 4–6 m), with a large and variable surface area (3000–6000 km^2), and its salinity ranges from 5% to 28%. Covering

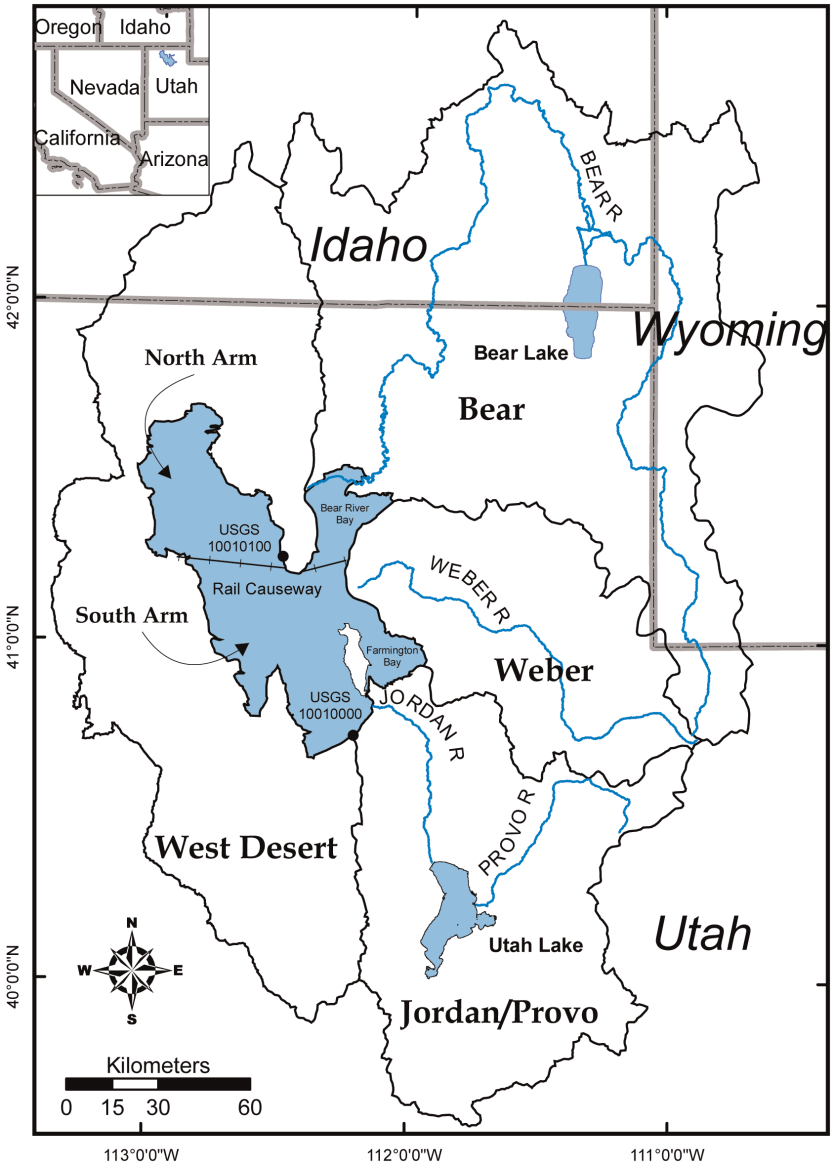

Figure 1. Location of the Great Salt Lake and subbasins that drain to it. The Great Salt Lake (latitude 40.7°N–41.7°N, longitude 111.9°W–113.1°W) is located in northeast Utah (insert).

portions of northern Utah, southern Idaho, and western Wyoming, the GSL drainage has an area of about 55,000 km^2. The GSL inflows are from three major rivers in the Wasatch region, the Bear, the Provo-Jordan, and the Weber, in addition to groundwater seepage. The only outflow is evaporation. The evaporation depends upon meteorological factors, salinity, and the lake surface area, which fluctuates with level.

[7] *Langbein* [1961] was one of the first researchers to study closed lakes, noting that closed lake levels fluctuate more than open lakes because variations in inflow can be compensated only by changes in surface area. *Langbein* [1961] developed an approximate relationship for the coefficient of variation of lake area in terms of rate of evaporation, lake area, lake depth, and drainage area for a number of lakes. He estimated that the response time, defined as the ratio of change in lake volume to the corresponding change in evaporation outflow rate, for the GSL is 9 years. This was based on a linear area-volume relationship, which he acknowledged as approximate but was used for mathematical tractability at that time.

[8] *Lall and Mann* [1995] used spectral analysis to examine the GSL biweekly volume time series and the role of climatic variability in the dynamics of the GSL. Their analysis of the GSL biweekly volume change data from 1847 to 1992 reveals decadal and interdecadal signals. Their work identified that the decadal and the interdecadal signals appeared to be consistent with the El Niño–Southern Oscillation (ENSO) index.

[9] *Moon* [1995], *Moon and Lall* [1996], and *Moon et al.* [2008] studied the relationships between the time variability of the volume of the GSL and selected atmospheric circulation indices. *Moon* [1995] developed and applied nonlinear measures of dependence between selected atmospheric circulation indices and the GSL volume at various lags (presuming that these indices are considered to lead the GSL volume). In addition, he forecasted the volume of the GSL using selected atmospheric circulation indices. The indices considered in his study were the Southern Oscillation index, The Pacific–North America (PNA) climatic pattern, and the central North Pacific climatic pattern. *Moon et al.* [2008] improved *Moon*'s [1995] earlier GSL forecasts by including a local predictive risk measure to guide the forecast process.

[10] *Lall et al.* [2006] developed short-term forecasts (1 year) of the biweekly GSL volume time series using a nonparametric, locally weighted, polynomial regression approach [*Cleveland and Devlin*, 1988]. This approach has the advantage of being easy to understand and implement while addressing an important aspect of predictor sampling design through summarization of the neighborhoods selected and through a reduction in the potential for overparameterization.

[11] Investigations of the GSL time series volume from a dynamical system perspective [*Mann et al.*, 1995; *Abarbanel et al.*, 1996; *Lall et al.*, 1996] suggest that it is one of a few geophysical series whose dynamics can be described by a low-dimensional nonlinear model with limited predictability. In order to demonstrate the relationship between structured low-frequency climate variability, low-order dynamical behavior of the GSL, and the enhanced long-term predictability of the GSL volume, *Abarbanel et al.* [1996] examined a long climatological record of measurements of the GSL volume. They used the global false nearest-neighbors method to choose the embedding dimension d appropriate for describing the GSL volume in multivariate state space. *Abarbanel et al.* [1996] found that an embedding dimension of 4 is sufficient to describe the GSL volume time series, suggesting that there are four degrees of freedom active in the Great Salt Lake volume record.

[12] *Abarbanel et al.* [1996] also suggested that physically based models of climate that are guided by low-frequency spatial and temporal features observed in data and that reproduce the dynamical attributes of the corresponding time series may be more useful for analysis of climate change issues than high-resolution models that lack such guidance. In this regard, long time series such as that of the GSL volume that represent a spatially averaged hydroclimate may provide a useful baseline. *Abarbanel et al.* [1996] analyzed the GSL volume time series, western U.S. precipitation, Northern Hemisphere sea level pressure, and air temperature using multivariate spectral analysis to get the interannual and decadal signals. They identified signals that represent 2, 3–5, 10–12, and 15–20 year intermittent oscillations with slowly varying amplitude and phase characteristics. They argued that the GSL volume responds with a small phase lag to regional precipitation and temperature anomalies, which are, in turn, forced by large-scale atmospheric circulation anomalies.

[13] *Wang et al.* [2009] reported that during the last 4 decades (1970–2010) precipitation in the central Intermountain Region, which supplies water to the GSL, has experienced a pronounced increase in temporal variability. They indicated that the precipitation time series from 1970 onward undergoes an amplified cycle with a period of about 15 years, a time scale that falls between the 2–7 year ENSO cycle and $\leq$25 year Pacific Decadal Oscillation (PDO) variation. *Wang et al.* [2009] also identified a phase shift between PDO and rainfall in which the rainfall cycle lags the PDO cycle by about 3 years. This knowledge related to climate forcing has implications for understanding fluctuations and possible preferred states of the GSL, as well as forecasting GSL levels.

[14] *Shun and Duffy* [1999] studied the precipitation, temperature, and runoff in the Wasatch Front in northern Utah. These quantities are inputs to the GSL basin hydrologic system. *Shun and Duffy* [1999] identified coherent patterns (oscillatory components) of annual, interannual, and decadal oscillations in precipitation, temperature, and runoff from point observations using a space-time form of principle components analysis called multichannel singular spectrum analysis [*Plaut and Vautard*, 1994] and discussed these components in terms of the hydrologic and hydrogeologic processes contributing to streamflow across the Wasatch Front. Multichannel singular spectrum analysis results from *Shun and Duffy* [1999] show that the variance contribution for precipitation, runoff, and temperature at all altitudes is dominated by the annual cycle and harmonics; precipitation shows significant but weak interannual and decadal oscillations, while runoff exhibits strong interannual and decadal oscillations. Their analysis [*Shun and Duffy*, 1999] qualitatively links amplified low-frequency oscillations to large-scale groundwater and base flow.

3. Great Salt Lake Level and Bathymetry Data

[15] Lake level data were retrieved from the U.S. Geological Survey (USGS) (from http://water.usgs.gov/

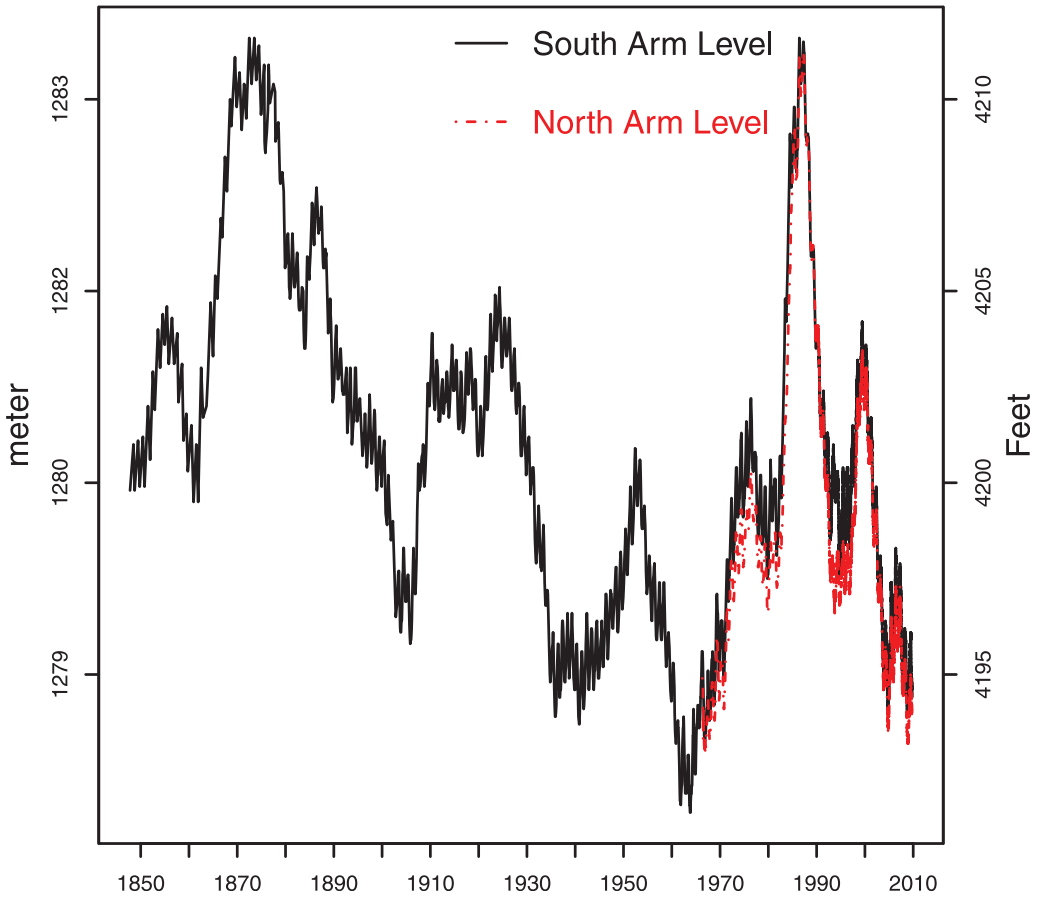

Figure 2. Great Salt Lake levels. The lake was divided into north and south arms by a railroad causeway in 1959. South arm and predivided lake levels (USGS 10010000, Great Salt Lake at Saltair Boat Harbor) are available since 1847. Separate records of level in the north arm (USGS 10010100, Great Salt Lake near Saline) are available since 1966.

data.html) on 13 June 2005 and then were updated (from http://waterdata.usgs.gov/nwis) on 13 May 2010 for the north arm (USGS 10010100) and the south arm (USGS 10010000) (Figure 2). These data incorporate USGS benchmark corrections available at http://ut.water.usgs.gov/greatsaltlake/elevations/gslcorrection.html [*Loving et al.*, 2000]. The earliest level data available are from 18 October 1847. USGS lake elevation measurements were first made in 1875, with lake level values prior to this being estimates based on observer reports. Levels range between the lowest level of 1278.5 m (4194.4 feet), recorded in 1963 (15 October 1963), and the highest level of 1283.7 (4211.6 feet), recorded in 1872, 1873, and 1986 (3 June 1986). The lake was divided by a rail causeway in 1959, and separate north arm level records are available from 15 April 1966.

[16] The bathymetry of the lake has been compiled by the USGS in tables that report the volume and area of the lake for a range of levels. The GSL bathymetry data sources that we are aware of are (1) for the north and south arms, volume-area tables for elevations of 1271.3–1285 m (4171.0–4216 feet) [*Loving et al.*, 2000]; (2) for the south arm, volume-area tables for elevations of 1270–1280 m (4167.0–4200 feet) but excluding Farmington and Bear River Bays [*Baskin*, 2005]; and (3) for the north arm, volume-area tables for elevations of 1270–1280 m (4167.0–4200 feet) [*Baskin*, 2006]. In this work we used the *Loving et al.* [2000] bathymetry tables because they provide estimates of volume and area for levels greater than 1280 m and cover the entire lake. Lake levels together with the bathymetry were used to determine lake area and volume. The relationships between level and area and level and volume for the north arm, the south arm, and the whole lake are depicted in Figure 3.

[17] Given that the lake was divided by a causeway constructed in the lake in 1959, a natural question is what impact the causeway fill had on Lake Bathymetry. We examined this and found the effect of the causeway to be minimal. On the basis of the trapezoidal cross section (11 m top width, 46 m base width, and 16 m height) and 40 km length [*Loving et al.*, 2000, Figure C2] the causeway volume is about 0.018 km^3, which is insignificant in comparison to the volumes depicted in Figure 3.

[18] In 1970, MagCorp constructed a dike to enclose part of the GSL as an evaporation pond for mineral production, which effectively excluded some area from the open lake level-volume-area relationship. This dike was breached in 1986 when the lake was at a high level; then repairs were completed in 1994. In evaluating GSL volume and area from level we used the bathymetry with the MagCorp pond area included prior to 1970 and during the period when the dike

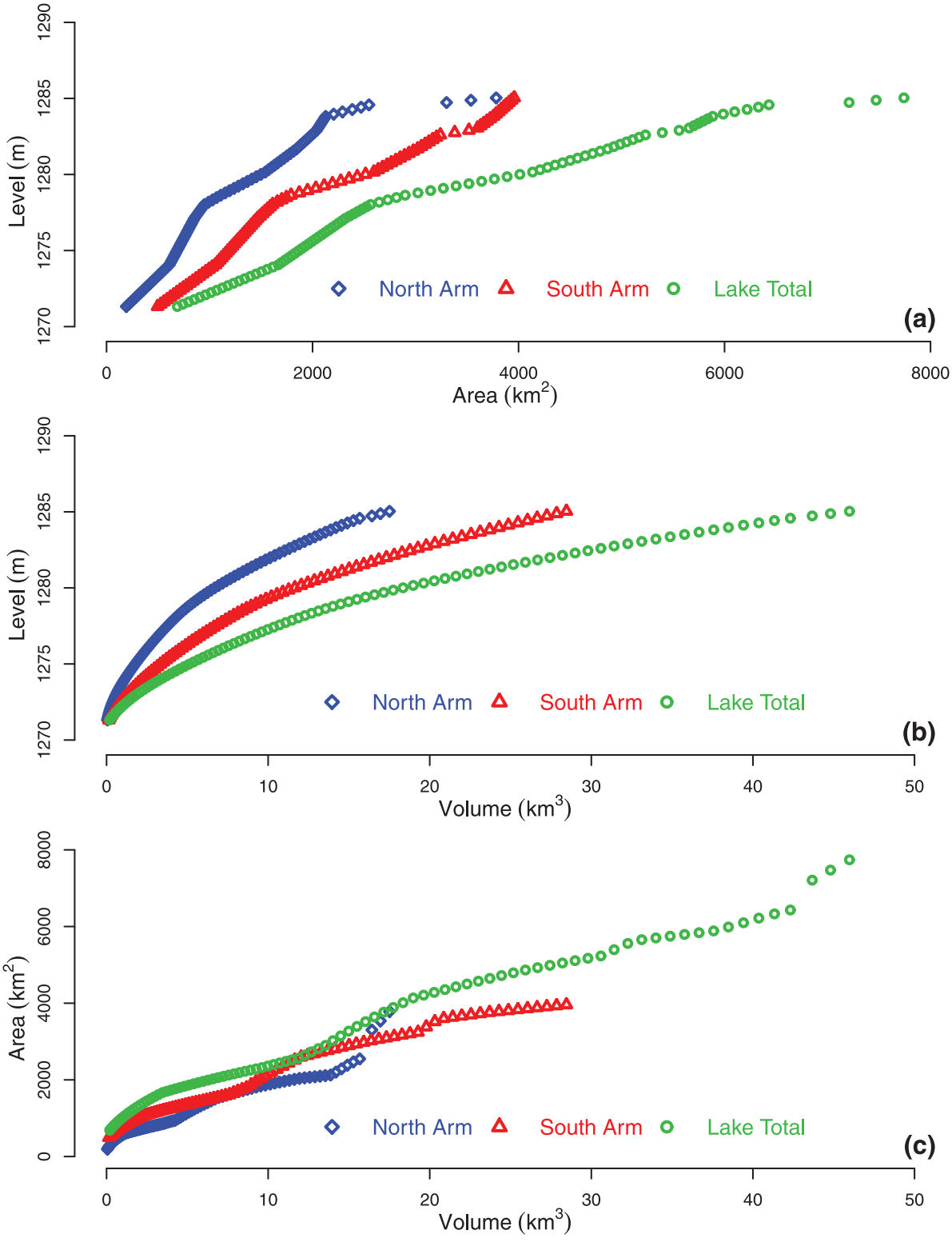

Figure 3. The Great Salt Lake's bathymetry. (a) Area-level (above mean sea level), (b) volume-level (above mean sea level), and (c) volume-area relationships. Data are from *Loving et al.* [2000, Table A-1].

was breached and with the MagCorp pond area excluded outside these periods. *Loving et al.* [2000] includes information on the bathymetry of the MagCorp pond. The MagCorp pond adjustments are not shown in Figure 3 because they are too small to be resolvable at the scale of the figure.

4. Great Salt Lake Inputs

[19] Three rivers, the Bear, Weber, and Jordan, flow into the GSL. The Bear River has been gauged since 1902, the Weber River since 1907, and the Jordan since 1949. The specific stations where streamflow data are available have

changed over the years. A detailed study of streamflow inputs to the GSL was conducted by *Loving et al.* [2000], who estimated the GSL streamflow inputs for 12 years from 1987 to 1998. Loving et al. present regression equations for estimating streamflow at locations where streamflow data are missing. We followed these regressions as closely as possible to compile the monthly streamflow inputs for the period 1 October 1907 to 30 September 2009. However, we did need to extend these methods to be able to estimate streamflow into the GSL for the extended period. Where streamflow data were missing, they was estimated using regression with a nearby station, preferably upstream in the same basin, or where correlation was best. The record for the Jordan River prior to 1949 was estimated using regression with flow in the Bear River.

[20] Direct precipitation on the lake was obtained from the monthly Parameter-Elevation Regressions on Independent Slopes Model (PRISM) climate data (http://www.prism.oregonstate.edu/) [*Daly et al.*, 2008] for the period October 1907 to September 2009. This is on a 2.5 arc min (~4 km) grid. Grid cells falling within the GSL were identified, and data from these grid cells were averaged to produce a time series of monthly precipitation.

5. Theory

[21] A simplified mass balance of the lake suggests that the rate of change in volume can be expressed as

$$\frac{\mathrm{d}V}{\mathrm{d}t} = P + Q - EA - O + R\,, \tag{1}$$

where P is precipitation directly on the lake, Q is flow entering the lake, A is lake area, E is the per unit area evaporation rate from the lake, O is pumping from the lake, and R is return water to the lake from West Desert pumping. Q includes predominantly streamflow but also has a small component of groundwater. The variables O and R are 0 most of the time but have been included here so as to be able to account for the manipulations of the GSL level by the West Desert Pumping Project in 1987–1989 and return flow to the GSL in 1990–1992. Denoting $P + Q - O + R$ as combined input, I, equation (1) can be generalized to

$$\frac{\mathrm{d}V}{\mathrm{d}t} = I - EA\,. \tag{2}$$

[22] If one assumes as a first approximation that the lake is in steady state, equation (2) suggests that area will adjust to balance lake input and output. Specifically,

$$A = \frac{I}{E}\,. \tag{3}$$

[23] So, if for simplicity we think of the lake as being forced primarily by variability in I with E constant and assume that I has probability density function (pdf), $f_I(i)$, then the corresponding pdf of A can be calculated. The pdf of A therefore represents the variability in input (primarily climate) here represented by I. Since A and V are related through the bathymetry, the pdf of V can be calculated as

$$f_V(v) = f_A(a)\frac{\mathrm{d}A}{\mathrm{d}V}, \tag{4}$$

where $f_V(v)$ is the volume probability density function and $f_A(a)$ is the area probability density function. This equation suggests that peaks in the derivative $\mathrm{d}A/\mathrm{d}V$ may be a cause for modes in the pdf of V.

[24] Analysis of the GSL volume time series suggests that the lake generally rises from 1 November to 15 June and falls from 15 June to 1 November [*Mohammed*, 2006]. Recent lake level observations are daily, but the early record is much more sporadic, with a level recorded or estimated only around the date of maximum and minimum each year. For example, the first year of record (1847) has a level recorded on 18 October, and the second year of record (1848) has levels recorded on 4 July and 30 September. To obtain a data set that represented the seasonal cycle and limited bias due to frequent recent observations but much more intermittent observations in the early record, we constructed a lake level time series by selecting the maximum value during the period March–August each year as a peak value and the minimum value during the period September–February as a trough value. The series comprising the sequence of twice-yearly lake levels and volumes and areas determined from these levels was used as the primary data set for examination of multimodality and the role of bathymetry in the relationship between modes in the area and volume distributions in this study. A histogram of lake volumes based on this time series (Figure 4) suggests the possibility of multiple modes in the volume distribution. These modes represent volumes at which clusters or preferred states occur in the lake volume time series, and the question is why these occur. Do they arise from nonlinear dynamics in the system or from the bathymetry or from a combination?

[25] A histogram counts the number of occurrences within predefined bins and is the simplest way of depicting a distribution. However, identifying modes from the histogram requires visual interpretation and is somewhat subjective because of the choice of bin width. It is therefore not very suitable for objective comparison of different distributions. Nonparametric density estimation methods can depict the distribution of data more generally and objectively [*Silverman*, 1986]. Kernel methods are a popular nonparametric approach that we have used here. Kernel methods offer the advantages of being based on local neighborhoods, data driven, and adaptive, requiring weak or limited assumptions about the underlying distribution. The kernel density estimate is defined as

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x - X_i}{h}\right), \tag{5}$$

where X_i, $i = 1, 2, \ldots, n$, represent n samples presumed to be from a probability distribution, h is the window or bandwidth smoothing parameter, and k is the kernel function that satisfies the conditions

$$k(t) > 0$$

$$\int_{-\infty}^{\infty} k(t)\mathrm{d}t = 1. \tag{6}$$

[26] Kernel-based nonparametric density estimation requires selection of the kernel function $k(\,)$ and bandwidth h.

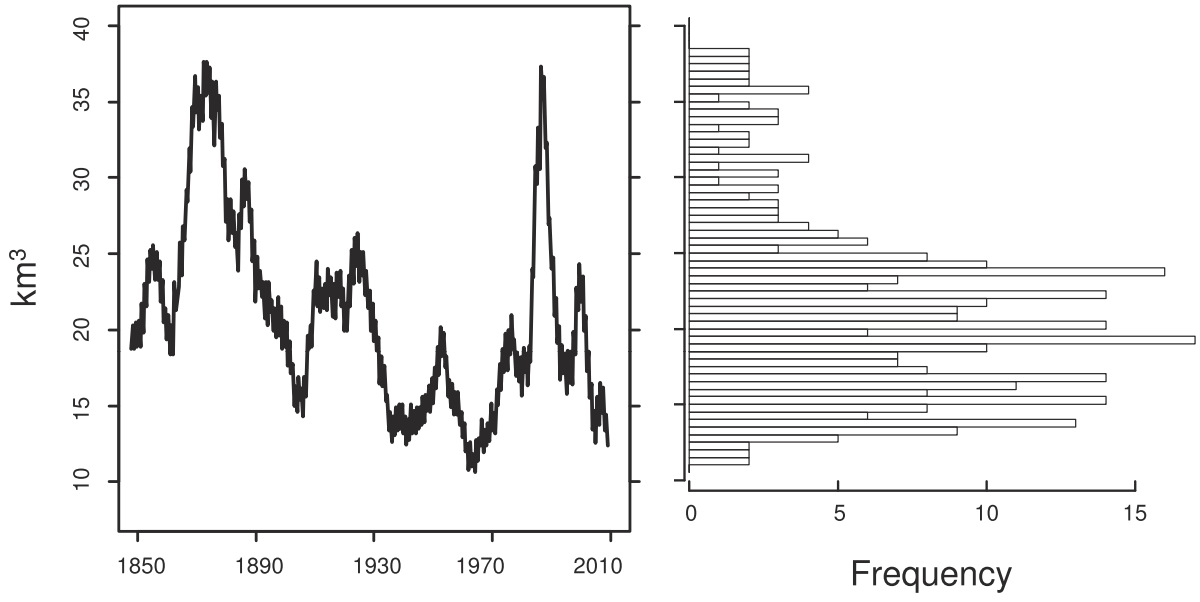

Figure 4. Great Salt Lake volume (1847–2009) and frequency histogram. The histogram counts the number of times volume is within the range associated with each bin.

The method is most sensitive to the selection of h. There are multiple methods that give estimates for the bandwidth h [*Silverman*, 1986; *Scott*, 1992; *Sheather and Jones*, 1991]. Here we used the *Sheather and Jones* [1991] method implemented by the R statistical software package to select bandwidth h [*R Development Core Team*, 2010]. This appears to be a broadly used and well-regarded method for bandwidth selection in the nonparametric statistics community, with a theoretical justification from the examination of kernel density estimate errors and statistical convergence properties (for further details, see *Simonoff* [1996, p. 46], *Bowman and Azzalini* [1997, p. 34], and *Sheather* [2004]). We used a Gaussian function for the kernel $k(\ )$.

6. Results

[27] Kernel density estimates of the GSL's volume for the periods 1847–1992 (289 twice-yearly observations) and 1847–2009 (323 twice-yearly observations) are shown in Figure 5. The period 1847–1992 is shown for comparison with the work of *Lall et al.* [1996], who used this period to examine modes of the GSL volume records from a system dynamics perspective. The pdf estimated for the period from 1847 to 2009 has modes at 15.5, 20.0, and 34.7 km^3. The bandwidth that the *Sheather and Jones* [1991] method estimated for the 1847–1992 data was 1.52 km^3, and for the 1847–2009 data it was 1.47 km^3. These are very close, but for consistency, so that differences are not due to bandwidth differences, Figure 5 was plotted using the average bandwidth of 1.50 km^3. To check sensitivity to bandwidth, we also plotted density estimates with bandwidths of 1.52 and 1.47 km^3 and obtained figures (not shown) that were very similar to Figure 5, with modes at the same locations. We therefore concluded that interpretations were not sensitive to the selection of bandwidth in the range resulting from different data lengths. Examining Figure 5, it is apparent that the years 1993–2009 have resulted in some filling in of the trough (around 18 km^3) between the lowest two modes in the density estimate, making these modes less distinct.

[28] Kernel density estimates of the pdf of the GSL's area for the periods 1847–2009 and 1907–2009 (Figure 6) show modes at $A = 3400$, 4600, and 6020 km^2. The volumes that correspond to these modes are 15.4, 21.3, and 34.1 km^3, respectively. The period 1907–2009 was plotted for later comparison with inflows because this is the period over which inflow data are available. The bandwidth estimate for the 1907–2009 data was 219 km^2, while for the 1847–2009 data it was 188 km^2. Again, in Figure 6 we used an average bandwidth of 203 km^2 for consistency.

[29] The topographic derivative, dA/dV, of the GSL evaluated numerically from area-volume data (Figure 7) has peaks at 14, 18.5, 31, and 43.5 km^3. Our hypothesis is that these peaks result in modes in the lake volume distribution. We see that the first three of these correspond roughly to the modes identified in Figure 5 (14 to ~15.5, 18.5 to ~20, and 31 to ~34.7 km^3), while the fourth bathymetric peak (43.5 km^3) does not have a corresponding mode in Figure 5. This bathymetric peak is above the range over which lake volume fluctuates in the observed record. We should also note that the peak at 18.5 km^3 occurs only in the derivative dA/dV evaluated when the area that is now the MagCorp pond is included as part of the open lake and as such only applies to interpreting lake volumes and areas prior to 1970 and for the period 1986–1994. This spike is the only tangible manifestation of the bathymetry differences due to the MagCorp pond. It is also worth noting that volumes corresponding to modes in the area distribution coincide quite closely with modes in the volume pdf identified above (Figure 5). Both peaks in the area pdf and peaks in dA/dV appear to be contributing to peaks in the volume pdf, and it is thus not possible to attribute volume peaks uniquely to one cause.

[30] Equation (4) suggests that $f_V(v)$ (Figure 5) should be obtained from the product of Figures 6 and 7. We performed

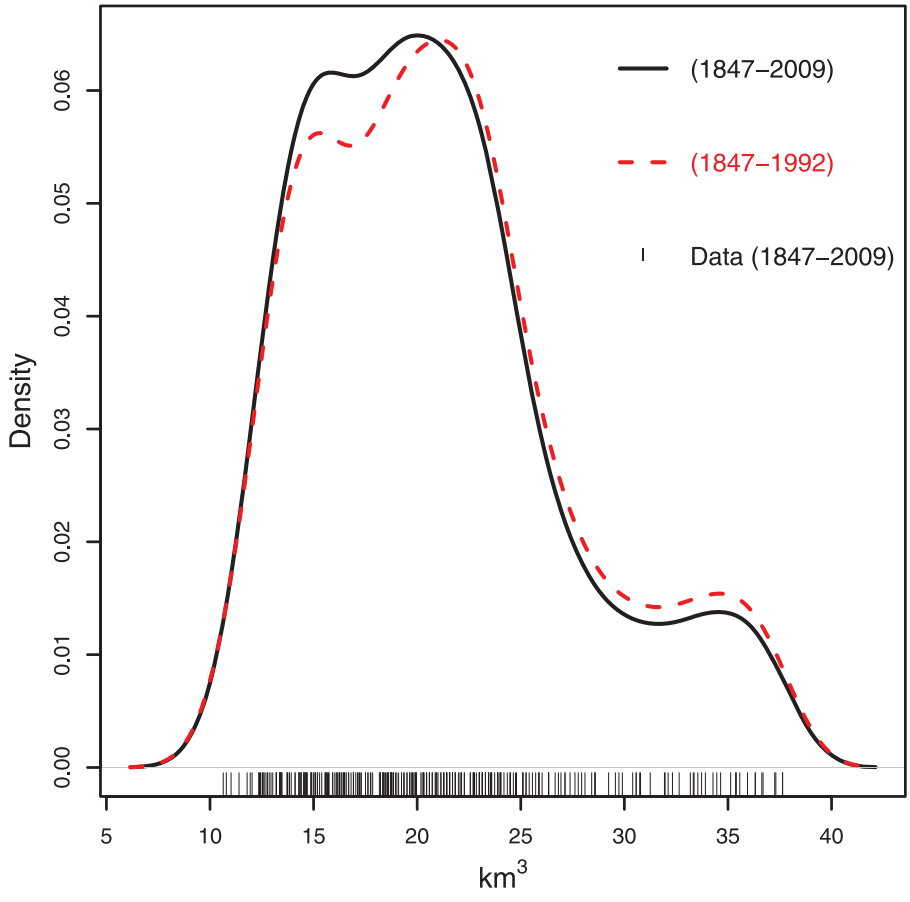

Figure 5. Kernel density estimate of the GSL's volume distribution with bandwidth equal to 1.5 km^3. The bars at the bottom represent the twice-yearly data values in the 1847–2009 volume density estimate.

a sequence of calculations to evaluate the applicability of this idea (Figure 8). The line labeled f_1 is the volume density estimate from the actual data (1847–2009), repeated from Figure 5, but with bandwidth of 1.47 km^3 determined from the full record.

[31] The first test we performed to assess our hypothesis was to evaluate the impact of the multimodality of the $f_A(a)$ distribution only on the volume density distribution multimodality. A constant dA/dV value of 0.12 m^{-1} (close to the average of the data in the center of Figure 7) was assumed in equation (4) and was multiplied by the $f_A(a)$ estimate from the full actual data record (1847–2009) with a bandwidth of 188 km^2 calculated from the data. This resulted in line f_2, which has a shape that is a rescaling of Figure 6 but does not match f_1 (Figure 8a). The density function $f_A(a)$ alone is therefore not sufficient to produce the observed multimodality in $f_V(v)$.

[32] The second test we performed was to evaluate the impact of dA/dV only. A Gaussian distribution was assumed for $f_A(a)$ in equation (4) and was multiplied by dA/dV from Figure 7. This Gaussian distribution had mean and variance determined from the 1847–2009 twice-yearly area data depicted in Figure 6. This multiplication resulted in line f_3 (Figure 8b). We see that multimodality does emerge, so part of our hypothesis that bathymetry is responsible for multimodality seems to be true. However, comparing to f_1, the modes are not in the correct position, and the main mode is not reproduced. Therefore, dA/dV alone is not sufficient to produce the observed multimodality in $f_V(v)$.

[33] The last test used the product of $f_A(a)$ (f_2 with a bandwidth of 188 km^2) and dA/dV from Figure 7. This resulted in line f_4 (Figure 8c). We see that this line is generally close to f_1. The differences between f_4 and f_1, we believe, are due to numerical differences between kernel smoothing over area and over volume and numerical uncertainty (spikiness) in the evaluation of dA/dV from volume and area data. The conclusion is that multimodality in both the input and the bathymetry, manifested by dA/dV, are necessary to explain the multimodality and shape of $f_V(v)$.

[34] In response to a concern that these results may be sensitive to the bandwidth selection method we repeated the procedure using other bandwidth selection methods [*Silverman*, 1986; *Scott*, 1992; *Venables and Ripley*, 2002], which resulted in bandwidth ranging from 1.47 to 1.95 km^3 for lake volume and 178 to 312 km^2 for lake area. We found that the general pattern of Figure 8 was consistent, and thus, the interpretations are robust against selection of bandwidth within these ranges.

[35] In order to further examine the inference from equation (3) that the GSL lake area pdf, $f_A(a)$, is a surrogate of the input pdf we examined the available observed lake inputs (precipitation and streamflow) over 102 water years

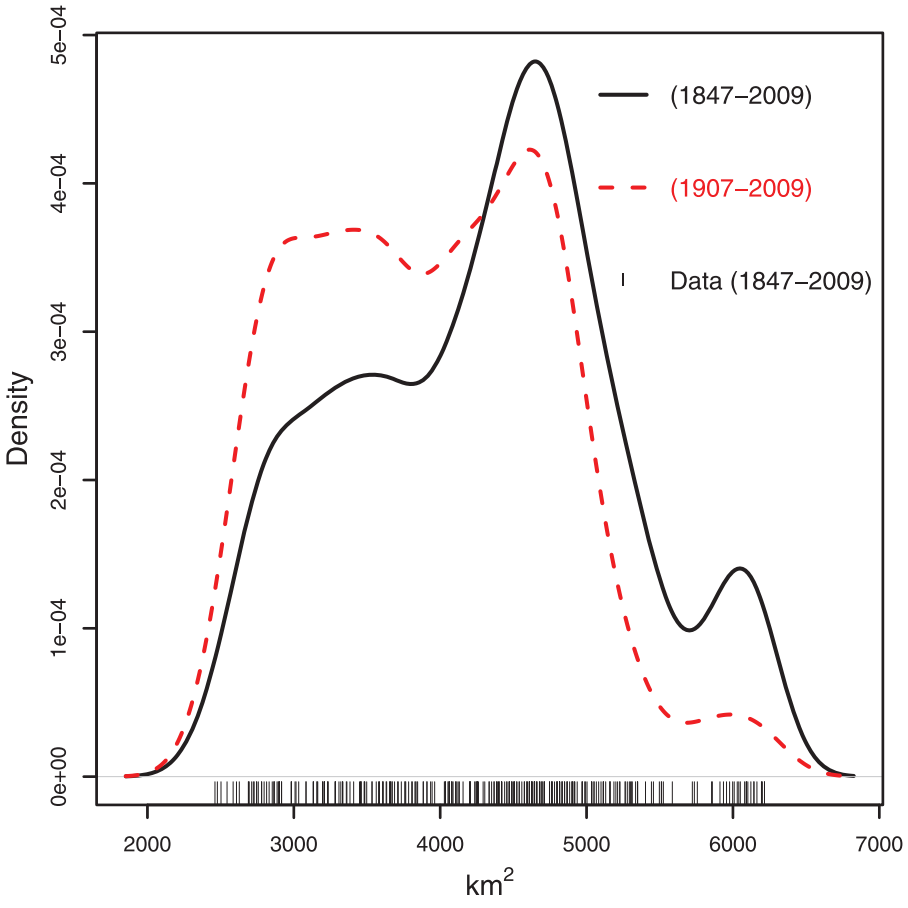

Figure 6. Kernel density estimate of the GSL's area distribution with bandwidth equal to 203 km^2. The bars at the bottom represent the twice-yearly data values used in the 1847–2009 area density estimate.

(1907–2009). The steady state approximation (equation (3)) suggests that lake area should adjust to balance inputs. However, there is a question as to the number of years required for this adjustment, with the likelihood that more recent years should weigh more heavily than earlier years in such a balance. To evaluate this, we derived a series of smoothed inputs using an exponential smoothing (geometric series of weights decreasing with time into the past) [*Holt*, 1957; *Winters*, 1960] to average past annual lake inputs. The exponential smoothing parameter was set to have an integral scale range from 1 to 10 years. We then plotted area times evaporation, the lake outflow, versus these smoothed inputs and evaluated the variability in AE explained by I using R^2 (1:1 line with no regression parameters). We found that the highest R^2 was obtained with an integral scale of 6 years (Figure 9). We did this comparison for evaporation set at the long-term average, $\overline{E}$, estimated from lake volume mass balance (equations (1) and equations (2)), and for a salinity-adjusted evaporation, E', to account for the effect that when the lake level is lower, volume is lower, salinity is higher, and evaporation is suppressed. The calculation of E' assumed that the lake has a constant salt load of 4.08×10^{12} kg and that concentration is derived from volume and salt load [*Loving et al.*, 2000],

$$C = \frac{\text{GSLLoad}}{\text{GSLVolume}}. \tag{7}$$

[36] The salinity correction factor (SCF) gives saline evaporation relative to freshwater evaporation under equivalent conditions [*Waddell and Barton*, 1980]:

$$\text{SCF} = 1 - \left[\frac{0.788C}{1000 + 0.63C}\right]. \tag{8}$$

[37] Equation (7) together with the area-volume relationship in the bathymetry gives concentration as a function of area, $C(A)$. Substituting in (8) results in SCF(A). Given A, we estimated E' as

$$E' = \overline{E}\text{SCF}(A)/\text{SCF}(\overline{A}). \tag{9}$$

[38] The 1:1 fit of AE' versus smoothed I with an integral scale of 6 years is slightly worse than using $\overline{E}$, suggesting that this salinity adjustment does not tangibly benefit these calculations.

[39] We then evaluated the pdf of the GSL inputs and the pdf of the GSL evaporation estimates with and without salinity adjustments (Figure 10). The bandwidth used in Figure 10 is 0.237 km^3, which is the mean bandwidth of the GSL inputs, the average evaporation output, and the salinity-adjusted evaporation output. Inputs smoothed at an integral scale of 6 years have a distribution that has modes at 3.3, 4.9, and 5.9 km^3, in similar locations to evaporation outputs, but the lower mode is higher and the distribution of inputs is

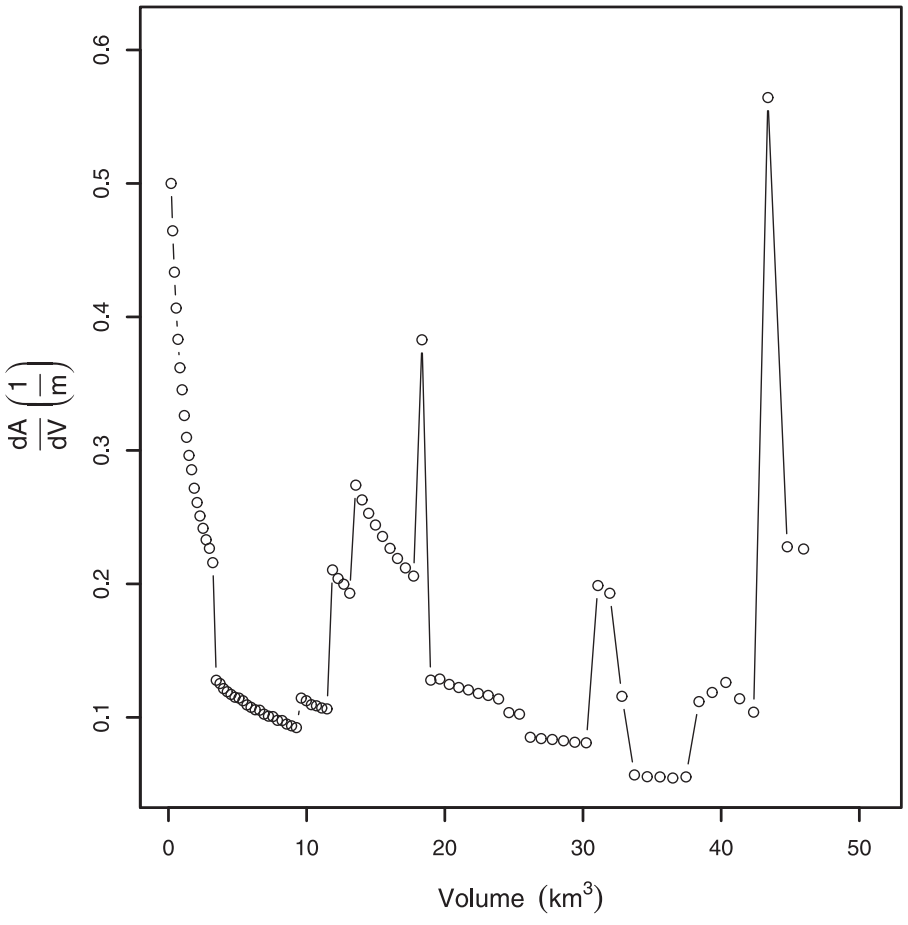

Figure 7. The topographic derivative dA/dV of the Great Salt Lake.

narrower than observed for $A\overline{E}$ and AE'. Dividing these modes by the average annual evaporation (1.036 m) results in corresponding area values of 3200, 4700, and 5700 km^2, which correspond reasonably closely to modes in the area pdf (Figure 6).

7. Discussion and Conclusions

[40] This examination of the modes in the GSL volume and area distribution is part of our effort to understand lake dynamics and the different roles played by bathymetry and hydrologic and climate inputs that drive the system. We have shown the theoretical connection between lake bathymetry and preferred states for a closed basin lake where the only outflow is evaporation and area, and hence, the level adjusts to balance driving inputs. This is important for interpreting lake level and volume records as indicators of system behavior and change. We found that the volume pdf, $f_V(v)$, derived from the area pdf, $f_A(a)$, and dA/dV agrees with the observed volume pdf, validating the derived distribution based on bathymetry (equation (4)) and quantifying the role played by bathymetry in lake modes that are interpreted as preferred system states. However, estimates of $f_V(v)$, derived from combining a constant value for dA/dV with observed $f_A(a)$ or using the observed dA/dV with a unimodal Gaussian $f_A(a)$, resulted in $f_V(v)$ that did not match the observed volume distribution. This indicates that both the bathymetry, quantified through dA/dV, and multimodal $f_A(a)$ are required to obtain $f_V(v)$.

[41] Our hypothesis was that some of the peaks in the volume distribution were attributable to bathymetry. We found that peaks in the area-volume derivative (dA/dV) generally match peaks in the volume distribution within the range of volumes over which the lake fluctuates. This supports the hypothesis. However, we also found that peaks in the area pdf corresponded with peaks in the volume pdf and that peaks in the input pdf corresponded with peaks in the area pdf. Thus, both bathymetry and inputs seem to have a causal role for modes in the volume pdf that represent preferred states. We are thus left concluding that our hypothesis is not fully supported, at least in the sense that modes in the volume pdf are not attributable to bathymetry alone. The coincidence of peaks in the area and input pdf with peaks in the bathymetry area-volume derivative begs the question as to the reason for this coincidence and which came first. Are there interacting dynamics between the sedimentation processes driving the bathymetry and the climate dynamics driving the inputs? The coincidence of these causes for volume pdf preferred states suggests this as a question that would merit further investigation. At the larger scale of historical Lake Bonneville topographic features such as shoreline, terraces and benches have been attributed to levels being persistent for "very considerable periods" [*Gilbert*, 1890] that we would now call preferred

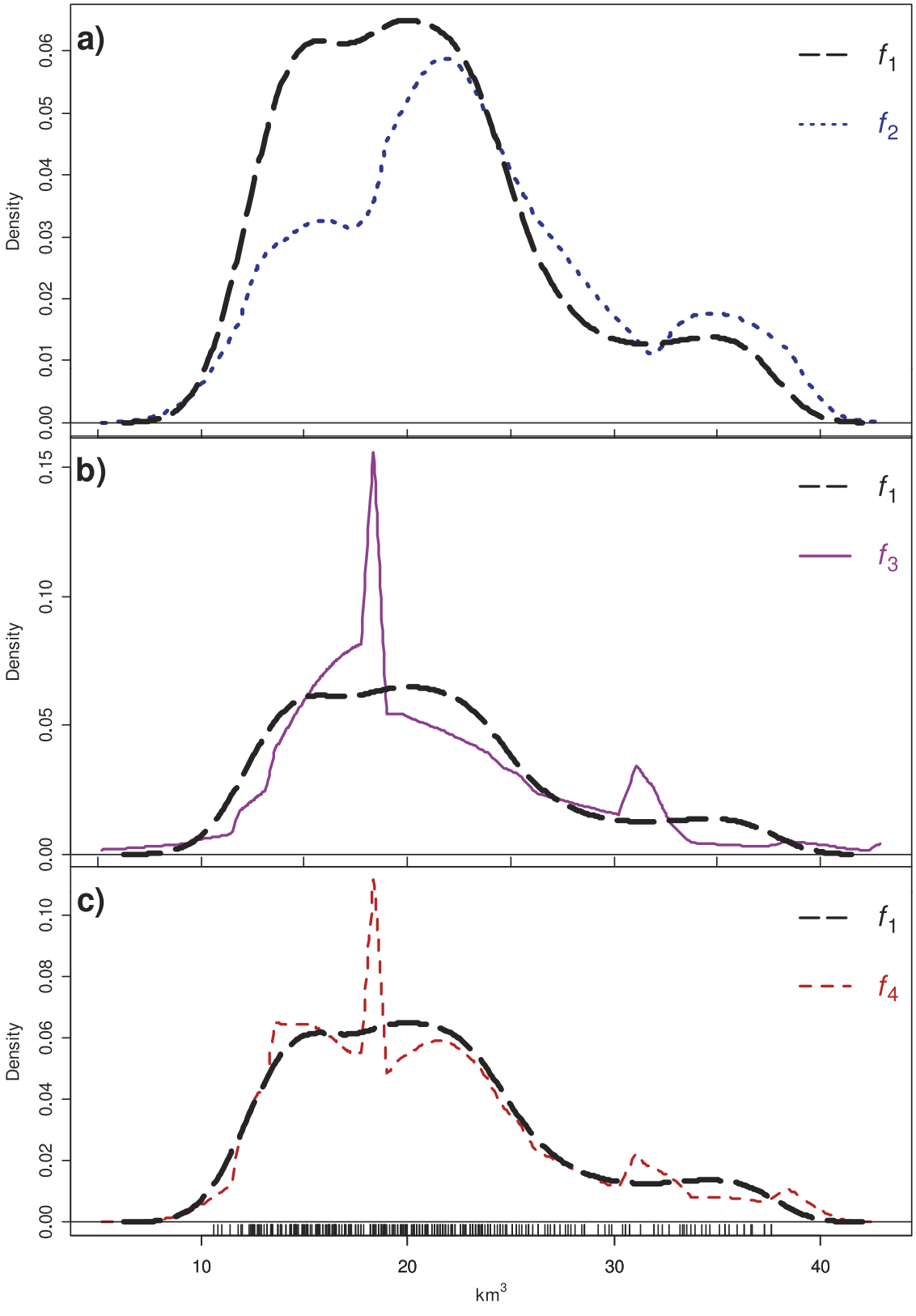

Figure 8. GSL volume density function comparisons: f_1, directly from data; (a) f_2, constant dA/dV used in equation (4); (b) f_3, Gaussian $f_A(a)$ used in equation (4); and (c) f_4, observed $f_A(a)$ and dA/dV used in equation (4). The bars at the bottom represent the total GSL volume twice-yearly data values used in the 1847–2009 volume density estimate.

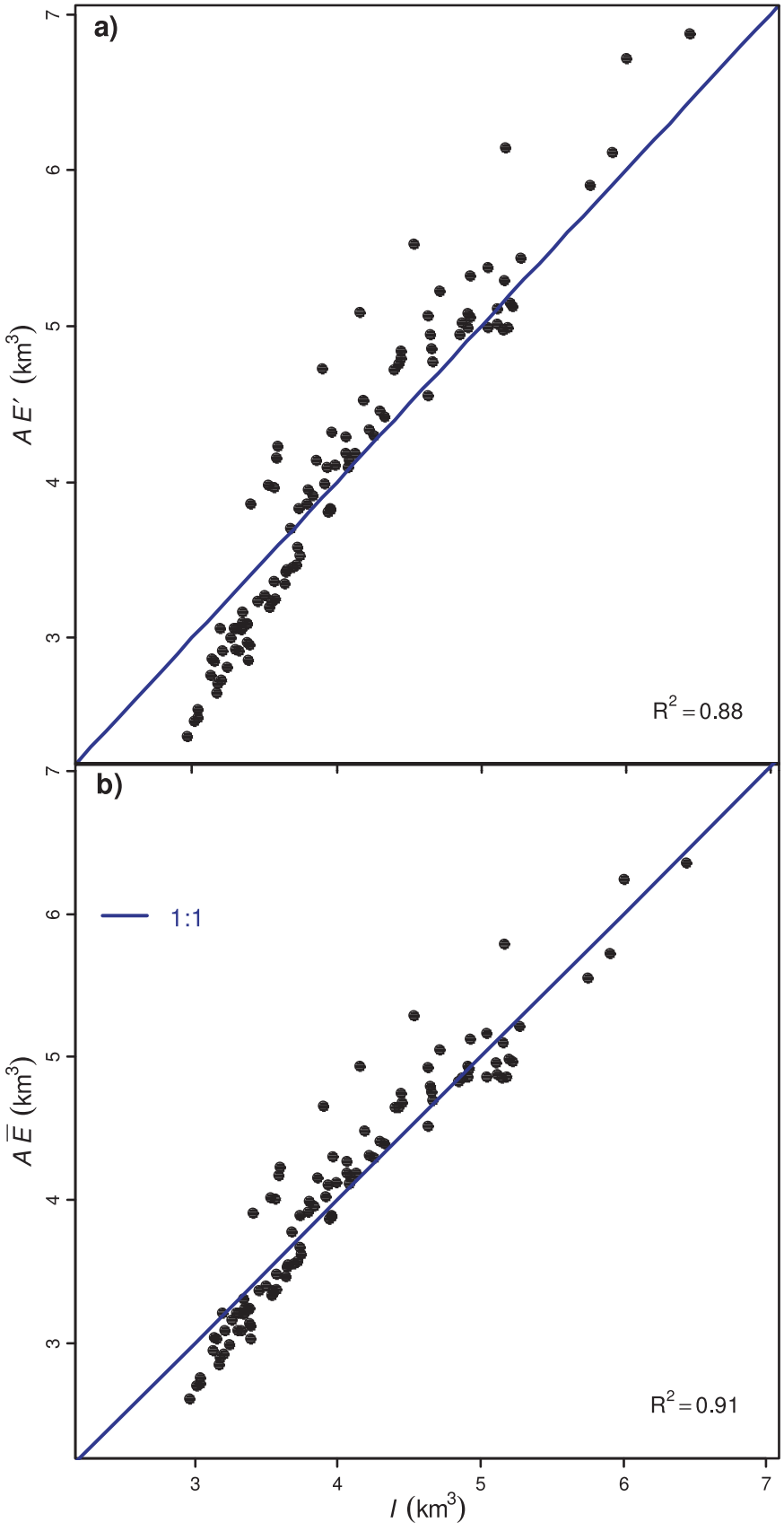

Figure 9. Comparison of Great Salt Lake evaporation output to river and precipitation input exponentially smoothed with 6 year integral scale. Period of record is 1907–2009. (a) Salinity-adjusted evaporation and (b) average evaporation.

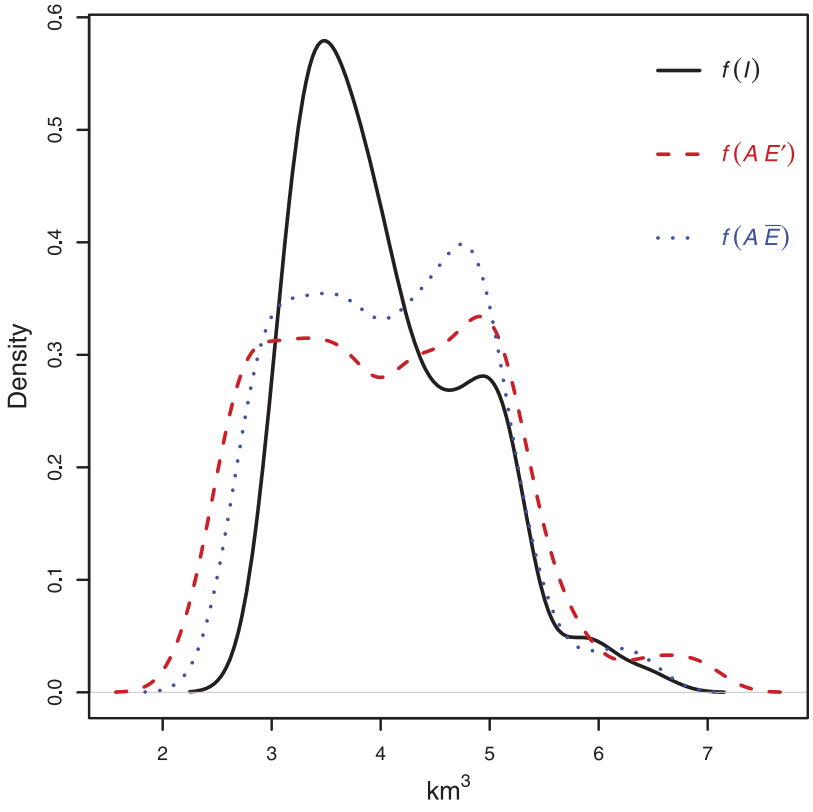

Figure 10. Kernel density estimates of (1) the GSL's inputs exponentially smoothed with a 6 year integral scale, (2) the GSL's area multiplied by average evaporation, and (3) the GSL's area multiplied by salinity-adjusted evaporation for the 1907–2009 time period with bandwidth equal to 0.237 km^3.

states. Could it be that similar effects have formed the present bathymetry of the smaller GSL that is a remnant of Lake Bonneville?

[42] Lake area is (through the assumption of steady state, equation (3)) more directly related to climate inputs that drive lake fluctuations than is lake volume, which first has to be modulated by the bathymetry (dA/dV). We see, comparing Figures 5 and 6, that multimodality in the distribution of lake area is, if anything, more pronounced than multimodality in lake volume. This is suggestive of preferred states in the dynamics of the inputs that drive GSL area and hence level and volume fluctuations.

[43] Directly examining the inputs revealed that inputs and evaporation outflow balance best when an integral scale of 6 years is used to smooth inputs. From this we conclude that response time of the GSL to fluctuations in inputs is close to 6 years. Response time can also be estimated from the ratio of active volume to average flux rate. For the GSL the active volume is estimated from the range of volume values in Figures 4 and 5 as 21 km^3. The average annual input, I, is 3.99 km^3/yr. This is balanced by evaporation. The active volume residence time is thus 21.0/3.99 = 5.3 years. This estimate of response time is consistent with that obtained from selecting the integral smoothing scale that best balances inputs and evaporation outflow. One implication of this is that the effective number of independent data points is smaller than the number of twice-yearly level records used in the analysis. When the period of record is divided by the response time, we get about 30. This is an estimate of the effective number of independent volume data points in this analysis and needs to be borne in mind in interpreting the strength of these results and conclusions as they pertain to modes and preferred states.

[44] A further finding from direct examination of the inputs is that the pdf of inputs smoothed using an integral scale of 6 years is not exactly the same as the pdf of lake area, although modes are in similar places. The differences are likely due to the fact that the GSL dynamics that result in area is not a simple average (with exponential smoothing) of past inputs. The response time is likely to be slower for high lake levels (large lake) than for low lake levels (small lake). Also, the lake is never really in steady state, but rather is continuously adjusting to input fluctuations, and there are uncertainties in the input estimates that may also be responsible for some of the differences in Figures 9 and 10.

[45] In comparing the pdf of lake volume calculated by *Lall et al.* [1996], using data from the period 1847–1992, with the pdf calculated with data from the period 1847–2009 we see that data acquired over the 17 years since *Lall et al.*'s [1996] work was done have changed our estimate of the shape of the volume pdf. This is seen in the reduction of the 18.0 km^3 trough that was evident in the data from the 1847–1992 time period. This begs a question as to why this trough has been filled in and whether this represents a shift in preferred lake volume states. There is also a question as to whether this infilling is statistically significant, especially in light of the effective sample size implied by the response time discussed above.

[46] There are a number of other factors to consider in understanding the fluctuations and preferred states of the GSL. These include salinity, temperature, and human impacts.

[47] Salinity suppresses evaporation when lake levels are low and salinity is high. This stabilizing effect impacts the lake area required to balance inputs as seen in Figure 10, where the lower evaporation, E', associated with higher salinity at lower levels shifts the lower end of the evaporation volume pdf to the left, while higher evaporation, E', associated with lower salinity at higher lake levels shifts the upper end of the evaporation volume pdf to the right. However, including the salinity correction factor in our comparisons of lake balance (Figure 9) made the comparison slightly worse, rather than improving it.

[48] Air temperature also influences evaporation, so one might expect some adjustment of the lake system to global or regional warming. However, there are interesting potential feedbacks. Evaporation from an expansive lake serves to suppress air temperature (or really the effect of advection of heat from around the lake is relatively less), and thus, the total quantity of evaporation is less. Increased evaporation may result in a smaller lake that has less of a cooling influence, thus leading to a positive destabilizing feedback. There may also be a connection between evaporation from the lake and precipitation on the watershed that feeds back into lake levels. Convection triggered over the lake may depend on temperature and lake extent and feedback via precipitation. Understanding all of these feedbacks is beyond the scope of this paper and may require a combination of future physical and statistical modeling investigations.

[49] Further, the GSL is no longer a purely natural system. Its level is managed at the upper end by pumping, and the causeway has divided the lake into two parts. There are two forms of pumping from the GSL. The first is the West Desert Pumping Project, which is used to control lake level when it threatens infrastructure (above 4208 feet, 1282.6 m). Thus, if the lake now approaches these levels, which correspond to a volume of 31 km^3, the pumps (humans intervening and becoming part of the system) will attempt to stop the lake from rising higher, suppressing the occurrence of the highest mode evident in Figures 5 and 6. The second form of pumping is withdrawals into mineral extraction ponds that have been developed on the edges of the lake. We have neglected these in this analysis because historically, these have been quite small and have only impacted the relatively small recent (post 1986 flooding) part of the record.

[50] The presence of the causeway means that the north and south arms can be at different levels. Since 1966, the average level of the south arm has been 0.27 m (0.89 feet) higher than the north arm. This complicates evaluation of the area-volume function, which was implicitly based on a single level in Figures 3 and 7. To assess this, we evaluated the derivative, dA/dV, presuming that the south arm was 0.3 m (1 foot) higher than the north arm and found that while there were small changes in the numerical values as compared to Figure 7, the general pattern was unchanged. We therefore conclude that the separation of the lake due to the causeway does not impact our findings.

[51] Much of the prior work we reviewed examined ways to forecast GSL volumes or levels using a state space dynamical systems perspective. This paper has focused on the physical role played by bathymetry in the determination of preferred volume states. One challenge in forecasting GSL levels related to the use of the state space dynamical systems perspective is that the empirical nature of such models and the variables they work with do not make them conducive to addressing practical questions related to physical and management changes in the system. A dynamical model that forecasts GSL volume on the basis of a state space of prior volumes cannot directly incorporate the effect of pumping or mineral pond withdrawals or alterations in inflow due to land and water resources management in the GSL drainage. In our opinion, a physically based modeling approach that brings together considerations of watershed response, water and salt mass balance, bathymetry, salinity, atmospheric factors affecting evaporation, withdrawals (pumping), and other lake management actions offers a promising path forward for further understanding this system and incorporating this understanding into models that can address some of the management questions being asked.

[52] Overall, this examination of modes in the GSL volume and area distributions has revealed the role played by bathymetry in GSL volume modes and the role played by inputs and response time in lake area modes. Preferred volume states in the GSL align with both peaks in the bathymetry area-volume derivative and peaks in the area pdf related to inputs. Correctly calculating the volume distribution requires knowledge of both. The pdf of area has a somewhat more pronounced multimodality than the pdf of volume, and this appears to be associated with multimodality in the input forcing. Preferred GSL volume states thus arise because of a combination of bathymetry and climate forcing factors.

[53] **Acknowledgments.** This work was supported by the Utah Water Research Laboratory. We are indebted to the valuable review comments from the editor and three anonymous reviewers.

References

Abarbanel, H. D. I., U. Lall, Y.-I. Moon, M. E. Mann, and T. Sangoyomi (1996), Nonlinear dynamics and the Great Salt Lake: A predictable indicator of regional climate, *Energy*, *21*, 655–665, doi:10.1016/0360-5442(96)00018-7.

Baskin, R. L. (2005), Calculation of area and volume for the south part of Great Salt Lake, Utah, *U.S. Geol. Surv. Open File Rep., 2005-1327*.

Baskin, R. L. (2006), Calculation of area and volume for the north part of Great Salt Lake, Utah, *U.S. Geol. Surv. Open File Rep., 2006-1359*.

Bowman, A. W., and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis: The Kernel Approach With S-Plus Illustrations, Oxford Stat. Sci. Ser.*, vol. 18, Clarendon, Oxford, U. K.

Cleveland, W. S., and S. J. Devlin (1988), Locally weighted regression—an Approach to regression-analysis by local fitting, *J. Am. Stat. Assoc.*, *83*, 596–610, doi:10.2307/2289282.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris (2008), Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, *Int. J. Climatol.*, *28*, 2031–2064, doi:10.1002/joc.1688.

Gilbert, G. K. (1890), Lake Bonneville, *Monogr. U.S. Geol. Surv.*, *1*.

Holt, C. C. (1957), Forecasting seasonal and trends by exponentially weighted moving averages, *ONR Res. Memo. 52*, Grad. Sch. of Ind. Admin., Carnegie Inst. of Technol., Pittsburgh, Pa.

Lall, U., and M. Mann (1995), The Great Salt Lake: A barometer of low-frequency climatic variability, *Water Resour. Res.*, *31*, 2503–2515, doi:10.1029/95WR01950.

Lall, U., T. Sangoyomi, and H. D. I. Abarbanel (1996), Nonlinear dynamics of the Great Salt Lake: Nonparametric short-term forecasting, *Water Resour. Res.*, *32*, 975–985, doi:10.1029/95WR03402.

Lall, U., Y. I. Moon, H. H. Kwon, and K. Bosworth (2006), Locally weighted polynomial regression: Parameter choice and application to forecasts of the Great Salt Lake, *Water Resour. Res.*, *42*, W05422, doi:10.1029/2004WR003782.

Langbein, W. B. (1961), Salinity and hydrology of closed lakes, *U.S. Geol. Surv. Prof. Pap.*, *412*.

Loving, B. L., K. M. Waddell, and C. W. Miller (2000), Water and salt balance of Great Salt Lake, Utah, and simulation of water and salt movement through the causeway, 1987–98, *U.S. Geol. Surv. Water Resour. Invest. Rep.*, *00-4221*.

Mann, M. E., U. Lall, and B. Saltzman (1995), Decadal-to-centennial-scale climate variability: Insights into the rise and fall of the Great Salt Lake, *Geophys. Res. Lett.*, *22*, 937–940, doi:10.1029/95GL00704.

Mohammed, I. N. (2006), Modeling the Great Salt Lake, M.S. thesis, 141 pp., Utah State Univ., Logan.

Moon, Y.-I. (1995), Large scale atmospheric variability and the Great Salt Lake, Ph.D. thesis, 140 pp., Utah State Univ., Logan.

Moon, Y.-I., and U. Lall (1996), Atmospheric flow indices and interannual Great Salt Lake variability, *J. Hydrol. Eng.*, *1*, 55–62, doi:10.1061/(ASCE)1084-0699(1996)1:2(55).

Moon, Y.-I., U. Lall, and H.-H. Kwon (2008), Non-parametric short-term forecasts of the Great Salt Lake using atmospheric indices, *Int. J. Climatol.*, *28*, 361–370, doi:10.1002/joc.1533.

Plaut, G., and R. Vautard (1994), Spells of low-frequency oscillations and weather regimes in the Northern Hemisphere, *J. Atmos. Sci.*, *51*, 210–236, doi:10.1175/1520-0469(1994)051<0210:SOLFOA>2.0.CO;2.

R Development Core Team (2010), R: A language and environment for statistical computing, version 2.11.1, R Found. for Stat. Comput., Vienna. (Available at http://www.R-project.org)

Sangoyomi, T. (1993), Climatic variability and dynamics of Great Salt Lake hydrology, Ph.D. thesis, 264 pp., Utah State Univ., Logan.

Sangoyomi, T. B., U. Lall, and H. D. I. Abarbanel (1996), Nonlinear dynamics of the Great Salt Lake: Dimension estimation, *Water Resour. Res.*, *32*, 149–159, doi:10.1029/95WR02872.

Scott, D. W. (1992), *Multivariate Density Estimation: Theory, Practice, and Visualization*, 327 pp., John Wiley, New York.

Sheather, S. J. (2004), Density estimation, *Stat. Sci.*, *19*, 588–597, doi:10.1214/088342304000000297.

Sheather, S. J., and M. C. Jones (1991), A reliable data-based bandwidth selection method for kernel density estimation, *J. R. Stat. Soc., Ser. B*, *53*, 683–690.

Shun, T., and C. J. Duffy (1999), Low-frequency oscillations in precipitation, temperature, and runoff on a west facing mountain front: A hydrogeologic interpretation, *Water Resour. Res.*, *35*, 191–201, doi:10.1029/98WR02818.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, 1st ed., 175 pp., Chapman and Hall, London.

Simonoff, J. S. (1996), *Smoothing Methods in Statistics*, 338 pp., Springer, New York.

Venables, W. N., and B. D. Ripley (2002), *Modern Applied Statistics With S*, Springer, New York.

Waddell, K. M., and J. D. Barton (1980), Estimated inflow and evaporation for Great Salt Lake, Utah, 1931–76, with revised model for evaluating the effects of dikes on the water and salt balance of the lake, *Coop. Invest. Rep. 20*, Utah Div. of Water Resour., Salt Lake City.

Wang, S.-Y., R. R. Gillies, J. Jin, and L. E. Hipps (2009), Recent rainfall cycle in the intermountain region as a quadrature amplitude modulation from the Pacific Decadal Oscillation, *Geophys. Res. Lett.*, *36*, L02705, doi:10.1029/2008GL036329.

Winters, P. R. (1960), Forecasting sales by exponentially weighted moving averages, *Manage. Sci.*, *6*, 324–342, doi:10.1287/mnsc.6.3.324.

Wold, S. R., and K. M. Waddell (1994), Salt budget for West Pond, Utah, April 1987 to June 1989, *U.S. Geol. Surv. Water Resour. Invest. Rep., 93-4028*.

I. N. Mohammed and D. G. Tarboton, Department of Civil and Environmental Engineering, Utah State University, 4110 Old Main Hill, Logan, UT 84322-4110, USA. (ibrahim.mohammed@aggiemail.usu.edu; dtarb@usu.edu)