ABSTRACT

In this paper, we have examined 4 models for Great Salt Lake level forecasting: ARMA (Auto-Regression and Moving Average), ARFIMA (Auto-Regression Fractional Integral and Moving Average), GARCH (Generalized Auto-Regressive Conditional Heteroskedasticity) and FIGARCH (Fractional Integral Generalized Auto-Regressive Conditional Heteroskedasticity). Through our empirical data analysis where we divide the time series in two parts (first 2000 measurement points in Part-1 and the rest is Part-2), we found that for Part-2 data, FIGARCH offers best performance indicating that conditional heteroscedasticity should be included in time series with high volatility.

1 Introduction

The Great Salt Lake (GSL) is one of the largest terminal lake and its position inside the rocky mountains made it an attractive settling area for the people migrating from the East during the 19th century. Located in Utah, its average surface area covers 4,400 km² but is subject to many variations with recent records of 1,529 km² for lowest in 1963 and 5,311 km² for highest in 1987 [1] [2]. This latest flood had a cost of over 200 million dollars in damages. We are far from the prediction of Alfred Lambourne [3] in 1909 who stated that the lake would disappear entirely by 1935.

The impact of the Great Salt Lake over its surrounding area is undisputable and predicting water surface level has become critical to prevent damages to local industries. Therefore, an accurate model could help handle future floods and limit the damages by taking preventive measures.

Many Studies have been done to predict the fluctuations of the Great Salt Lake. Most of them were based on the volume and some on the water-surface level. In [4], a spectral analysis was conducted and demonstrated 14 peaks that pass the 90% significance test. The peaks could be resolved as a combination of four pulsations with a period of 1-, 11-, 14- and 36-years. He then distinguished three different regimes of lake behavior: pre 1924, 1924 – 1986 and 1986 – 1992, based on three modes in kernel probability density estimates of the GSL volume.

In [5], multivariate adaptive regression splines (MARS) were used and helped to forecast the wet and dry seasons of GSL volume time series.

In [6], support vector machines (SVM) are used to forecast the water surface level of the GSL.

The United States Geological Survey (USGS) has been monitoring the water-surface-elevation from the Great Salt Lake since 1875 and on a continuous basis since October 1902. The lake’s northern part started to be supervised in April 1966. There are currently two measuring apparatus (gauges) to observe the water-surface of the lake. The gauges are situated at the boat harbor in Saltair Beach State Park (southern part, since October 1938), and at the Little Valley boat harbor northwest of Saline (north part, since April 1966). Measured elevations are averaged on a daily basis and made available to the public. The Great Salt Lake’s
elevation differs between the southern and northern parts. The Union Pacific Railroad causeway separates the lake into two distinct parts. The water-surface elevation of the southern part of the lake is usually 0.5 to 2 feet higher than that of the northern part because the majority of the inflow to the lake is into the southern part.

First, we will analyze the data of the Great Salt Lake surface water level. Then, we will introduce the concepts behind fractional order modeling as well as the description of model studied. Finally, we will compare the accuracy of the different models to predict the SLC water level.

2 Data Analysis

In this part, a comprehensive analysis of the distributional, statistical, and time series properties of the data will be developed. The purpose is to determine the optimal way to format the data and which characteristics should be emphasized within the models.

The data consist of biweekly observations of the surface water level of the GSL from 1878 to 2000. Figure 1 describes the evolution of water surface of the GSL throughout the considered sample. A rapid observation reveals that the first 1900 or 2000 samples manifest a trend of descending, while the last 1000 samples are in the trend of ascending. Furthermore, figure 5, in the later part, demonstrates that the last 1000 samples possess more volatility than the first 1900 samples. As a consequence, the whole data set is divided into two distinct periods: the first 1900 samples are identified as the first sample period (or sample 1), and the last 1015 samples as the second sample period (or sample 2).

Table 1 outlines the descriptive statistics for sample 1 and sample 2. Both surface water level value and difference in the surface water level are examined (Difference=water level(t)-water level(t-1)). The first interesting result resides in the difference of the standard deviation in the two samples. The second sample presents a larger standard deviation than sample 1, informing more volatility. The two series in sample 1 are left skewed while the surface level in sample 2 is right skewed. Except for the difference of the sample 2, all series exhibits platykurtic, namely their distribution curve is very flat, or plateau-like.

Preliminary to analysis, the stationarity of time series is to be determined. A time series is said to be stationary when "the mean, variance and autocorrelations can usually by well approximated by sufficiently long time averages based on the single set of realizations." [7]. For a stationary process, the effects of the shocks are temporary and the series reverts to its long-run level. Under that condition, any long-term forecast of a stationary series will converge to the unconditional mean of the series. For a non-stationary process, time-dependence exists and matters. A long-run mean does not exist and the variance diverges to infinity.

Standard ARMA analysis requires the assumption that the considered time series is stationary, which can be analyzed using AutoCorrelation Function (ACF) and Partial Autocorrelation function (PACF). ACF determines the correlation between a time series value and the lag of that value. PACF specifies the additional correlation between a time series value and a special lag value, withdrawing the influence of other lag values. When the ACF of a time series declines gradually, the time series is likely to be nonstationary. Figures 2-5 show the ACF and PACF of surface water levels and difference of surface water levels in both samples.

For both samples, surface water levels are not stationary because both ACF and PACF decay slowly, meanwhile, the fast decay in ACF and PACF of the difference of the surface water level shows stationarity. Therefore in the following parts the difference of water level should be used for analysis instead. Furthermore, from Figures 3 and 5 both ACF trail off while PACF in Fig. 3 shows significant spike for the first lag, indicating an AR model with order 1 should be used. Further inspection of ACF can also indicate a tendency in the time series. A slowly decaying ACF depicts a unit root process. For the time series defined in (1), $y_t$ is stationary and the estimate of $a_1$ is efficient when $|a_1| < 1$. If $a_1 = 1$, the variance becomes infinitely large as time develops and there is a unit root. However, discriminating a unit root process from a stationary process using ACF analysis only can be very challenging. Indeed, a near unit root process will have a similar shape to a real unit root process. For example, when $a_1 = 0.99$ the ACF analysis exhibits a similar gradual decay as a nonstationary process. Therefore, ACF or PACF analysis does not suffice to assert the existence of a unit root, and Dickey-Fuller test and Phillips-Perron Test should be performed. Phillips Perron test incorporates an automatic correction to the Dickey-Fuller test procedure to allow autocorrelated residuals. Both Dickey-Fuller and Phillips-Perron Tests results point a unit
root in the surface water level, while the null of unit root is rejected for the difference of the surface water level, endorsing the difference of the surface water level as the data employed for modeling.

Suppose a time series $y_t$ defined as:

$$y_t = a_1 y_{t-1} + \varepsilon_t$$  \hspace{1cm} (1)

In the time series as in (1), $\varepsilon$ is an evenly distributed white noise process with mean 0 and variance $\sigma^2$. While empirical proved that variance can be changing. One way is to model the conditional variance using an AR process:

$$\sigma_t^2 = a_0 + \sigma_{t-1}^2 + \sigma_{t-2}^2 + \ldots + \sigma_{t-q}^2 + \nu_t$$ \hspace{1cm} (2)

where $\nu_t$ is a white noise process. An equation like (2) is called an AutoRegressive Conditional Heteroskedastic (ARCH) model. To test whether a time-series' variance is correlated, ARCH LM test and Ljung-box Q tests should be performed.

Figure 6 describes the considered series (difference of surface water level) from 1878 to 2000. Variance clusters change over time can be observed, i.e. large difference in water level tend to be followed by oppositely large differences. Furthermore, ARCH LM test and Ljung-Box Q statistics results reveal a significant serial correlation in residuals, which shows conditional heteroscedasticity in the time series and therefore recommends the use of ARCH modeling.

3 Long memory & Hurst parameter

Some empirical researches demonstrate the existence of long memory in time series data. This long memory, also named long run dependence, reflects when autocorrelations’ decay is slow while the process remains stationary. For a weakly stationary time series where sample mean $\mu_t$ is independent of $t$ and correlation $\rho(t+h,t)$ is independent of $t$ for each $h$, the autocorrelation function (ACF) is defined as:
\( \rho(k) = \frac{E[(x_t - \mu)(x_{t+k} - \mu)]}{\sigma^2} \)  \( (3) \)

where \( \mu \) is the mean and \( \sigma^2 \) is the variance.

If \( \sum_{k=-\infty}^{\infty} |\rho(k)| < \infty \), the time series exhibits short memory or short range dependence, or weak dependence; while if \( \sum_{k=-\infty}^{\infty} |\rho(k)| = \infty \) as \( k \to \infty \), the time series reveals long memory.

If the correlation function \( \rho(k) \) satisfies \( \rho(k) \sim C_p |k|^{-2(1-H)} \) as \( k \to \infty \) where \( C_p \) is a constant, \( C_p > 0 \) and the symbol \( \sim \) means "asymptotically equal to". \( H \) is the Hurst parameter (whose name was taken from the hydrologist who pioneered the topic [Hurst, 1951]). The Hurst parameter indicates the degree of Long Range Dependence (LRD). A Hurst parameter \( 0 < H < 0.5 \) stands for a negatively correlated process, or anti-persistent process. Different techniques exist to estimate the Hurst parameter such as The R/S parameter, aggregated variance, periodogram, Variance Residuals and so on. Table 2 summarizes the Hurst parameter estimation using different techniques for the surface water level and the difference from year 1878 to 2000. For the surface water level, most methods show that the Hurst parameter is between 0.5 and 1, and for the difference of the surface water level which is the variable of interest, all of the methods shows the existence of long run dependence. Therefore, long memory is recommended in the model.
4 Description of Models

In this section, four different models: ARMA, ARFIMA, GARCH, FIGARCH will be described.

4.1 ARMA(p, q)

Let \( \{x_t\}_{t=1}^{\infty} \) be a stationary process defined by

\[
\Phi(B)x_t = \theta(B)\varepsilon_t
\]  

(4)

where \( \Phi(B) \) is the autoregressive polynomial operator, \( \Phi(B) = 1 - a_1L - a_2L \cdots - a_pL \); \( \theta(B) \) is the moving average polynomial operator and \( \theta(B) = 1 + \beta_1L + \beta_2L \cdots + \beta_qL \). \( \varepsilon_t \) is a white noise process normally distributed with zero mean and finite variance \( \sigma_{\varepsilon}^2 \).

4.2 ARFIMA(p, d, q)

Empirical researches demonstrate the existence of long memory in time series data. The study of this long memory phenomenon has received much attention in recent years. ARFIMA, fractionally differentiated ARMA model, has been widely used in time series data analysis.

In [8] and [9], the ARFIMA(p, d, q) model is defined as:
\[ \Phi(B)(1-B)^d x_t = \theta(B) \varepsilon_t \]
\[ (1-B)^d = \sum_{k=0}^{\infty} (dk)(-B)^k = \sum_{k=0}^{\infty} \pi_k B^k \]
\[ \pi_k = \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} = \prod_{1 \leq j \leq k} \frac{j-1-d}{j} \]
\[ \Gamma(x) = \begin{cases} \int_0^\infty t^{x-1} e^{-t} dt \quad , x > 0 \\ \infty \quad , x = 0 \\ x^{-1} \Gamma(1+x) \quad , x < 0 \end{cases} \]

where \((1-B)^d\) is the fractional differentiating operator and \(\Gamma(.)\) is the gamma function.

For \(d \in (-0.5;0.5)\) the process defined is stationary and invertible.

### 4.3 GARCH(1,1)

In ARMA(p,q) processes, the variance of the disturbance term is assumed to be constant, which is also called Homoskedastic. While many time series data exhibits volatility clustering. Engle, in [7], defines a stochastic process whose conditional variance is a linear function of the square of the estimated residuals, called an autoregressive conditional heteroskedastic (ARCH) model. In [10], Bollerslev extends Engle’s work by allowing the conditional variance to be an ARMA process. Such model is named generalized ARCH model, or GARCH model. The key features of GARCH models are both autoregressive and moving average components included in the heteroskedastic variance.

The GARCH (1,1) model is the most popular model in empirical research and is defined as:

\[ x_t = \alpha_0 + \alpha_1 x_{t-1} + \varepsilon_t \]
\[ \varepsilon_t = h_t^{1/2} z_t \]
\[ h_t = \omega + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} \]

with \(\omega \geq 0, \beta_1 \beta_2 \geq 0\). \(h_t\) is the conditional variance. \(\varepsilon_t\) is the residual of the process. \(z_t\) is defined as the standardized residuals:

\[ z_t = \frac{\varepsilon_t}{\sqrt{h_t}} \quad (5) \]

Moreover, the distribution of \(z_t\) conditional on previous \(x\) is then:

\[ z_t | x_{t-1}, x_{t-2} \ldots \sim N(0,1) \quad (6) \]

Such process is stationary if and only if \(\beta_1 + \beta_2 < 1\). (10). In [11], this condition satisfied, the unconditional variance is finite, the unconditional kurtosis is always greater than 3 and can be infinite, the correlation is zero between any two variables in the time series data, and the correlation is positive between any squared residuals.

### 4.4 FIGARCH(1,d,1)

The processes in ARCH model developed by Engle and GARCH model by Bollerslev are short memory processes since the response of a shock on the conditional variance decreases at an exponential rate. However, researches have proven the existence of long run dependence in the conditional variance process. Volatility tends to change slowly over time, and the autocorrelations is dominated by a hyperbolic rate of decay. Similar as the development of research in conditional mean in time series data, vast research boomed in recent years. Using daily returns of the S&P500 index, Bollerslev and Mikelsen [12] constructed and evaluated Fractionally Integrated Generalized Autoregressive Conditional Heteroskedastic model (FIGARCH \((1,d,1)\)). Their results provided evidence against short memory specifications where \(d = 0\) and reject integrated process where \(d = 1\). In their research, it shows that the effects of a shock on the conditional variance decrease at a hyperbolic rate when \(d\) is between 0 and 1, which is different from ARFIMA model where \(0 < d < 0.5\).

Comparing with ARCH models who focus on the short term volatility specification and forecasting, FIGARCH model considers a finite persistence of volatility shocks. FIGARCH model impose an ARFIMA structure on the variance. When \(d = 1\), the FIGARCH model will reduce to an integrated GARCH model. When \(d = 0\), the FIGARCH model will reduce to GARCH model.

An AR(1)-FIGARCH(1,d,1) is defined as:

\[ x_t = \alpha_0 + \alpha_1 x_{t-1} + \varepsilon_t \]
\[ \varepsilon_t = h_t^{1/2} z_t \]
\[ h_t = w + \beta_1 h_{t-1} + (1 - \beta_1)(1 - \phi_1 L)(1 - L)^d \varepsilon_t^2 \]

where \(z_t\) is a white noise with mean 0 and variance 1, \(x_t\) follows a AR(1) process and the conditional variance \(\varepsilon_t^2\) is represented by a FIGARCH process:

\[ [1 - \beta_1(L)] h_t = w + \beta_1 h_{t-1} + (1 - \beta_1)(1 - \phi_1 L)(1 - L)^d \varepsilon_t^2 \quad (7) \]

where all the roots of \(\phi_1 L\) and \([1 - \beta_1(L)]\) lie outside of the unit circle and \(0 < d < 1\). The conditional variance of the FIGARCH process can be written as:
\[ h_t = \frac{w}{1 - \beta_1(L)} + \lambda(L)\varepsilon_t^2 \] (8)

where \( \lambda(L) = (1 - (1 - \phi_1 L)(1 - L)^d)[1 - \beta_1(L)]^{-1} \).

5 Results

Each sample is split in two parts: the estimation part and forecasting part. Model parameters will be estimated from the first \( T - 24 \) observations (\( T \) is the sample size), and 24 forecasts will be generated, giving two observations per month over one year. The efficiency of each model will be estimated by the performance of the forecasting accuracy, which will be measured using Mean Squared Error, Median Squared Error and Mean Absolute Error.

ARMA(1,0) and GARCH(1,1) are chosen as models since: first, it is the most simple specification and most widely used by the literature, and second, results from ACF and PACF analysis graphs clearly demonstrate that ACF trails off while PACF shows significant spike for the first lag.

The parameters estimated by ARMA, GARCH, FIGARCH, ARFIMA models are summarized in table 3 and table ??2. Except for the GARCH model from sample 2, all parameters are statistically significant at 1%. \( d \) parameters in FIGARCH in both samples are between 0 and 1, indicating the stability of the process. However, \( d \) parameters in both FIGARCH models are greater than 0.5, signifying that the process is not stationary.

24 Forecasts (1 year) by ARMA, ARFIMA, GARCH, FIGARCH will be presented and compared by Mean squared error (MSE), Median Squared Error(MedSE), and Mean Absolute Error(MAE) as shown in Tables ?? and ?? for two samples. Based on these statistics, FIGARCH forecasts are generally more accurate (smaller MSE), and less biased, (smaller variance than most results shown by ARMA, GARCH and FIGARCH). The performance of FIGARCH is excellent in sample 2, (MSE is significantly lower than the others, indicating that conditional heteroscedasticity should be used in periods of high volatility). FIGARCH forecast is better than the other models for overall performance as also illustrated in Figs. 7 and 8.

Long term forecasting is also considered to check the proficiency of FIGARCH modeling. In Figure 9, the blue line represents the real Great Salt Lake surface water level while the red line represents the ten years out of sample forecast. The FIGARCH forecast is capable to catch the yearly rise and fall of the water levels properly as well as the overall evolution of the surface water level.

6 Conclusions and Discussions

The empirical research done in this paper shows that the difference of the surface water level in The Great Salt Lake demonstrates long run dependence as proved by Hurst parameter value and variance clusters. The whole time series data set is divided into two samples: the first sample in a descending trend and less volatile, and the second sample in ascending trend and more volatile. Four models: ARMA, ARFIMA, GARCH and FIGARCH are estimated and compared by two samples. The results show that the overall performance of FIGARCH is better especially for the second sample, indicating that conditional heteroscedasticity should be considered in times series with high volatility. The FIGARCH model successfully forecasts the changes of the water level in short term-one year as well as the long term ten years.

ACKNOWLEDGMENT

This research was sponsored in part by Utah Water Research Laboratory’s Research Initiative Seed Grant (2006-2007). C. Tricaud was supported by Utah State University (USU) Presidential
Figure 7. DRAFT OF FORECASTS VS. REAL VALUE FOR SAMPLE 1


REFERENCES
Figure 8. DRAFT OF FORECASTS VS. REAL VALUE FOR SAMPLE 2


Figure 9. DRAFT OF FORECASTS VS. REAL VALUE OVER 10 YEARS