Effective Water Management in the Mahaweli Reservoir System; Analyzing the Inflow of the Upmost Reservoir

W. M. N. Dilini\textsuperscript{a}  
Liwan Lyanage-Hansen\textsuperscript{b}  
M.T.D. Attygalle\textsuperscript{a}  
K. D. W. Nandalal\textsuperscript{c}

Abstract: Mahaweli cascaded reservoir system is built contiguous to the Mahaweli river, enhancing the water storage and transferring ability to reinforce the needs of water in accordance with the climatic changes. The foremost requirement of the system is to provide water for irrigation and to produce hydroelectricity to the country that has given rise to conflicting demands of water requirements from the two sectors. Forecasting future water availability is crucial to predict the hydroelectricity generation capacity while maintaining a balance between the provisions of water to both sectors. Therefore, modeling the changes of in-flow of the upmost reservoir is substantial to effective water management. Literature has not revealed much evidence for an existence of an accepted statistical model in predicting the inflow, of the system considered. This research is based on the methodologies to test for the seasonality, stochasticity and non-linearity of the in-flow, in advance to fitting a suitable model.

Key words: In-flow; Stochasticity; Non linearity; Seasonality.

I. Introduction

Hydroelectricity is the production of electrical power through the use of the gravitational force of falling or flowing water. As with most other countries, in Sri Lanka too, hydroelectricity is one of the main sources to cater to the increasing demand of electricity. Mahaweli hydropower system which originated under the Mahaweli Ganga Development Programme, is the largest integrated rural development multi-purpose programme based on water resources of Mahaweli and allied six river basins. The foremost requirement of the system is to provide water for irrigation and to produce hydroelectricity to the country that has given rise to conflicting demands of water needs mainly from the two sectors power generation and agriculture.

Mahaweli hydro power complex consists of six major power stations which had an installed capacity of 660 MW. Currently, it contributes to around 13% electrical energy to the country annually. The major power stations coming under Mahaweli complex are, Kotmale, Victoria, Randenigala, Rantambe, Ukuwela, Bowatenna and Nilambe covering 1268 sqkm of

\textsuperscript{a}Department of Statistics, University of Colombo, Colombo 03, Sri Lanka.  
\textsuperscript{b}School of Computing and Mathematics, University of Western Sydney, Australia.  
\textsuperscript{c}Department of Civil Engineering, University of Peradeniya, Peradeniya, Sri Lanka.  
*Corresponding author: Liwan Lyanage-Hansen l.liyanage@uws.edu.au  
http://dx.doi.org/10.14453/isngi2013.proc.15
Mahaweli basin. This research is carried out based on these five reservoirs starting from Kothmale flowing down the line.

Hydroelectricity is the production of electrical power through the use of the gravitational force of falling or flowing water. As with most other countries, in Sri Lanka too, hydroelectricity is one of the main sources to cater to the increasing demand of electricity. Mahaweli hydropower system which originated under the Mahaweli Ganga Development Programme, is the largest integrated rural development multi-purpose programme based on water resources of Mahaweli and allied six river basins. The foremost requirement of the system is to provide water for irrigation and to produce hydroelectricity to the country that has given rise to conflicting demands of water needs mainly from the two sectors power generation and agriculture.

![Figure 1. Schematic diagram for the cascaded reservoir system in Mahaweli river basins.](Image)

Mahaweli hydro power complex consists of six major power stations which had an installed capacity of 660 MW. Currently, it contributes to around 13% electrical energy to the country annually. The major power stations coming under Mahaweli complex are, Kothmale, Victoria, Randenigala, Rantambe, Ukuwela, Bowatenna and Nilambe covering 1268 sqkm of Mahaweli basin. This research is carried out based on these five reservoirs starting from Kothmale flowing down the line.
In Sri Lankan context, accurate prediction of the generation of the hydroelectricity from the Mahaweli system is of vital importance to determine the ability of satisfying the future electricity needs and to minimize the wastage of resources due to inaccurate predictions. Forecasting the hydroelectric generation, however, requires a comprehensive knowledge of future availability of water resources.

Modeling the in-flow to the upmost reservoir Kothmale is substantial in that regard, as the whole system depends upon the changes in the natural behavior of this in-flow. Thus a better understanding of the natural inflow is of utmost importance. Original data for the in-flow are recorded on a daily basis and these were initially aggregated to weekly, monthly and annually by calculating the averages. The data set consists of daily data values from 1997 to 2011 which includes 14 years of data altogether. The number of observations for the Kothmale reservoir is 5478 and those were recorded as 365 per each year with the exception of leap years.

This paper presents the analysis of the basic characteristics of the hydrological process for the upmost natural in-flow in order to understand the stream-flow process using univariate time series.

II. Related Literature

The focus of the analysis is based on the stochasticity, stationarity and linearity of the in-flow data. The Literature related to analyzing these characteristics of the in-flow was explored and the techniques found were implemented.

A. Stochasticity

A stochastic process is a statistical process involving a number of random variables depending on a variable parameter. If the variable parameter is the time, then the process becomes a stochastic time series in this analysis the in-flow series. The most important aspect of the stochasticity is stationarity. A series \{Xt\} is called stationary if its statistical properties do not change with time. More precisely, \{Xt\} is said to be completely stationary if, for any integer k, the joint probability distribution of xt, xt+1, ..., xt+k is independent on the time index t. There can be two types of stationarity tests, namely the trend test and the unit root test. The first test is done to determine if the values of a series have a general increase or decrease with the time increase while the latter test is done to determine if the distribution of a series is dependent on the parameter ‘time’.

The Mann–Kendall trend test is one of the widely used non-parametric tests to detect significant trends in time series. This test, being a function of the ranks of the observations rather than their actual values, is not affected by the actual distribution of the data and is less sensitive to outliers. On the other hand, parametric trend tests, although more powerful, require the data to be normally distributed and are more sensitive to outliers. The Mann–Kendall test is therefore more suitable for detecting trends in hydrological time series, which are usually skewed and may be contaminated with outliers. This test has been used with environmental time series.

B. Seasonal Unit Roots
Most of the time series data are evident of substantial seasonal fluctuations rather than the usual trend. The trend of time series data can be addressed by the so called Augmented Dickey-Fuller (ADF) test in order to identify the fulfillment of stationary assumptions. If seasonality is present in time series data, there are no such tests in common practice and hence most of the researchers have overcome this issue simply by using seasonally adjusted data.

The seasonal adjustment can lead to destroy important relationships between economic time series data\(^2\). Since the seasonal component is not only present in the economic time series data but also severely occurring in environmental data, a testing procedure should be there to identify the seasonal fluctuations in time series data without removing the seasonality. There are many tests that are proposed for testing seasonal unit root tests such as Dickey-Hasza-Fuller (DHF) test\(^3\), OCSB test\(^4\). Among these seasonal unit root tests the HEGY test\(^5\) has the advantage of testing seasonal unit root at each frequency separately\(^6\). According to the work carried out by Empananza\(^7\) the methods introduced by Hylleberg et al.,\(^6\) were generalized in a way that p-values for any periodicity can be calculated.

C. Non-Linearity
Stream flow processes are commonly perceived as nonlinear. They could be governed by various non linear mechanisms acting on different temporal and spatial scales\(^8\). He further elaborates that the BDS test is a non parametric method for testing for serial independence and non linear structure in a time series based on the correlation integral of the series.

D. Seasonality and Spectral Analysis
The dynamics of stream flow are often dominated by annual variations. How well the seasonality is captures is very important criterion for assessing a stochastic model for streamflow\(^8\). Out of the cycle plots\(^9\) to study the behavior of seasonal time series, seasonal subseries plots is a cycle plot that can be effectively used in detecting seasonality in a time series. However, this plot is only useful if the period of the seasonality is known in advance. A spectral plot is a graphical technique for examining cyclic structure in the frequency domain. Uneven long term variations in the series, other than trend, are called cycles. Spectral analysis is a statistical approach to detect regular cyclical pattern, or periodicities. In Spectral Analysis the data are transformed with a finite Fourier transformation and decomposed into waves of different frequencies. Thus the time series data are expressed in terms of sine and cosine components.

III. Analysis

According to Figure 2 (left), annual average inflow shows fluctuations between the range from 1.6 mcm to 2.5 mcm with the exception of 3.36 mcm for the year 2010. The yearly data series according to Figure 2 (right) for monthly data, seasonality can be seen but the length of periodicity of the seasonality is not very clear. The data seem to be not having significant trends and they seem to be somewhat stationary.

A. Testing for Stochasticity
In order to detect the nature of stochasticity, the stationarity of the in-flow series is observed. Testing for the trend in the in-flow time series, in other words, determining whether the values of a series increase or decrease with time, is presented in Table 1.
Table 1: Mann Kendall Trend Test results for monthly inflow series of year 1997 – 2011.

<table>
<thead>
<tr>
<th>Month</th>
<th>Score</th>
<th>Tau</th>
<th>2-sided value</th>
<th>p</th>
<th>Month</th>
<th>Score</th>
<th>Tau</th>
<th>2-sided value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>17</td>
<td>0.162</td>
<td>0.42848</td>
<td>0.42848</td>
<td>July</td>
<td>-4</td>
<td>-0.0383</td>
<td>0.42848</td>
<td>0.88183</td>
</tr>
<tr>
<td>February</td>
<td>22</td>
<td>0.211</td>
<td>0.29811</td>
<td>0.29811</td>
<td>August</td>
<td>-48</td>
<td>-0.459</td>
<td>0.29811</td>
<td>0.019872</td>
</tr>
<tr>
<td>March</td>
<td>40</td>
<td>0.383</td>
<td>0.053314</td>
<td>0.053314</td>
<td>September</td>
<td>-8</td>
<td>-0.0766</td>
<td>0.053314</td>
<td>0.72872</td>
</tr>
<tr>
<td>April</td>
<td>13</td>
<td>0.124</td>
<td>0.55262</td>
<td>0.55262</td>
<td>October</td>
<td>-21</td>
<td>-0.2</td>
<td>0.55262</td>
<td>0.3223</td>
</tr>
<tr>
<td>May</td>
<td>10</td>
<td>0.0957</td>
<td>0.65565</td>
<td>0.65565</td>
<td>November</td>
<td>-1</td>
<td>-</td>
<td>0.65565</td>
<td>1</td>
</tr>
<tr>
<td>June</td>
<td>-2</td>
<td>-</td>
<td>0.0191</td>
<td>0.96048</td>
<td>December</td>
<td>15</td>
<td>0.143</td>
<td>0.96048</td>
<td>0.48842</td>
</tr>
</tbody>
</table>

Observing the results, it can be concluded that, only the month August can be detected to have a trend in the inflow series at 5% level of significance. The lack of having monthly trends in the series can be accepted to a certain extent as environmental data must be having more seasonality than trends.

B. Testing for Seasonality

The seasonality in the inflow series was observed on a monthly basis since monthly flows usually exhibit strong seasonality according to Literature. The graphical technique “seasonal sub series plot” was used for this purpose.

Seasonal sub series plot for monthly data was initially plotted (left) and the same plot for log series was also obtained (right) due to the higher variability in the raw monthly series. The second plot reveals a seasonality pattern. The inflow series shows a peak in November whereas July and October also shows a value very close to that. The inflow series exhibits very low occurrence for the months of February and March. Starting from March, inflow series steadily increase until July and then show up a slow decrease until October. Then, the series again rises in the November and decreases until March.
The inflow series heavily depend on the rainfall patterns and river flow patterns. The rainfall also acts as one of the factors that affect the river flow pattern. Paddy cultivation in Sri Lanka is being done for two main seasons “yala” (October - March) and “Maha” (April – September) considering the rainfall variation. The inflow series exhibit a pattern somewhat close to this grouping which suggests incorporating the seasonality in the rainfall to predict the inflow for further studies.

C. Testing for Stationarity

As seasonal decomposition may lead to loss of important information, stationarity (existence of seasonal unit roots) was tested in the presence of seasonality using the HEGY procedure (Table 2).

<table>
<thead>
<tr>
<th>Statistic</th>
<th>t1=</th>
<th>t2=</th>
<th>F1=</th>
<th>F2=</th>
<th>F3=</th>
<th>F4=</th>
<th>F5=</th>
</tr>
</thead>
<tbody>
<tr>
<td>t1=</td>
<td>2.82</td>
<td>2.97</td>
<td>12.22</td>
<td>17.63</td>
<td>11.21</td>
<td>12.12</td>
<td>8.28</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0423</td>
<td>0.0239</td>
<td>0.0003</td>
<td>0.0000</td>
<td>0.0007</td>
<td>0.0003</td>
<td>0.008</td>
</tr>
</tbody>
</table>

Calculated p-values for all seasonal unit roots (monthly) are less than 0.05. Thus, there is enough evidence to reject the null hypothesis of the existence of a unit root. Therefore it can be concluded that the monthly series of inflow does not have seasonal unit roots that suggests stationarity. The same procedure for weekly data also provided evidence at 10% level of significance that the series does not have seasonal unit roots.

D. Testing for Non-Linearity

BDS test for linearity/non-linearity assumes that the series of interest is stationary. Since the monthly series and weekly series gave up satisfactory results BDS test can be applied here to check for non-linearity (Table 3).
Table 3: BDS test results of the inflow series from 1997 – 2011 for four time scales

<table>
<thead>
<tr>
<th>Time Scale</th>
<th>Dimension = 2</th>
<th>Dimension = 3</th>
<th>Dimension = 4</th>
<th>Dimension = 5</th>
<th>Dimension = 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistics</td>
<td>Statistic</td>
<td>Statistics</td>
<td>Statistic</td>
<td>Statistic</td>
</tr>
<tr>
<td></td>
<td>p - value</td>
<td>p - value</td>
<td>p - value</td>
<td>p - value</td>
<td>p - value</td>
</tr>
<tr>
<td>Daily</td>
<td>0.08484</td>
<td>0</td>
<td>0.14454</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.181596</td>
<td>0</td>
<td>0.200795</td>
<td>0</td>
<td>0.207178</td>
</tr>
<tr>
<td>Weekly</td>
<td>0.071247</td>
<td>0</td>
<td>0.109348</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.132478</td>
<td>0</td>
<td>0.144014</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Monthly</td>
<td>0.013696</td>
<td>0.0191</td>
<td>0.013696</td>
<td>0.0064</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.032379</td>
<td>0.0035</td>
<td>0.036281</td>
<td>0.0017</td>
<td>0</td>
</tr>
<tr>
<td>Annual</td>
<td>0.001208</td>
<td>0.9604</td>
<td>-0.106511</td>
<td>0.0081</td>
<td>-0.053507</td>
</tr>
<tr>
<td></td>
<td>-0.053507</td>
<td>0.2851</td>
<td>-0.073135</td>
<td>0.1811</td>
<td>-0.036281</td>
</tr>
<tr>
<td></td>
<td>0.013963</td>
<td>0.8014</td>
<td>0.013963</td>
<td>0.8014</td>
<td>0</td>
</tr>
</tbody>
</table>

The test statistic for BDS test as indicated above, decreases with the increase of time scale, from daily to annual. Therefore, it can be stated that non linearity decreases with the increase in time scale. Further considering the p-values, it can be concluded that the strongest non-linearity is viewed for the daily series. Weekly and monthly series are also non linear at 5% level of significance. However, it can also be concluded that the annual series is linear as it’s p-value is 0.96. This result is in agreement with the statement ‘All annual series are linear’.

However, the seasonality in the inflow series is assumed to be annual in the previous analyses and it might not be the actual case for in-flow data.

E. Spectral Analysis

Initially, the monthly series was tested to verify the existence of an annual periodic component as tested above. The above seasonality test is carried out assuming that data contain annual periodic components. The validity of this assumption was checked using the periodogram drawn for monthly data. An annual periodicity corresponds to a period of 12 as the inflow series with this data are monthly values. Since the period and frequency are reciprocals of each other, period of 12 corresponds to a frequency of 1/12 (=0.083). When analyzing the periodogram, it can be seen that the highest peak is at frequency 0.083.

Spectral density function, which is a smooth version of the periodogram is analyzed for the remaining peaks. Although the density plot consists with distinct peaks, starting from the lowest frequency peak for the period 12, those peaks are not equally spaced. Thus, the existence of annual periodic components cannot be confirmed.
IV. Future Work

The major advantage of Artificial Intelligence modeling is the considerable ability to map input-output pattern without requiring prior knowledge about the factors that affect the forecasting parameters, and have forecasted daily levels using adaptive network-based fuzzy inference system\textsuperscript{11}. In order to handle the random nature of the problem, the literature suggests that neural networking could be a supporting methodology to deal with such complex problems.

Another set of researchers on the other hand state that natural processes are complex, and it is sometimes not possible to build a single global model that adequately captures the system behaviour\textsuperscript{12}. They claim that the training data can be split into a number of subsets, and separate specialized models can be built on each subset. These models are called local or expert models, and sometimes called as a committee machine. Thus, a combination of several methodologies, such as time series analysis and neural networking may resolve and find a reasonable solution to the problem of this study.

V. Conclusion

The inflow series does not consist with monthly trends and seasonal unit roots at frequencies monthly and weekly. The reason to these results may due to the absence of annual periodicity which was detected through spectral analysis. High dependency of the series with the rainfall patterns and river flow patterns can be visualized through the monthly seasonal patterns of the data. Non linearity of the series decreases with the increase in time scale. The analysis concludes confirming the chaotic and volatile nature of inflow series, which need to be catered through data driven methodologies described in section 4.

References

6Meng, X., and He, C., Testing Seasonal Unit Roots in Data at Any Frequency an HEGY approach, Dalarna University, 2012.