Reconstruction of gappy mean sea level data

Quang-Hung Luu* & Pavel Tkalich
Tropical Marine Science Institute, National University of Singapore, 18 Kent Ridge Road, S119227, Singapore
*[E-mail: tmslqh@nus.edu.sg]

Received 12 August 2013; revised 16 October 2013

Missing data is a common problem in sea level records, and yet gappy records may hinder estimates of sea level trend and variability. Sea level may be affected by various scale phenomena, ranging from global climate change to interannual sea level pattern caused by coupled ocean-atmosphere oscillations, as well as by many other regional phenomena. It could be expected that such complexity make sea level analysis more difficult; however, if relationships between the above phenomena and sea level values are established for continuous parts of the record, they can be used for reconstruction of the gappy parts. In this study, we derive simple and robust method of reconstruction of the gappy data in the Singapore Strait using established relationships of sea level on the El Niño-Southern Oscillation (ENSO) and the Asian monsoon. The reconstructed mean sea level is verified against observations, and allows for more accurate estimation of sea level trend and variability in the Singapore Strait.

[Keywords: mean sea level, reconstruction of gappy data, sea level rise, monsoon, ENSO.]

Introduction

Missing data is a common problem in sea level analysis. For tide gauge records gappy data may hinder estimates of mean sea level trend and variability, which thereafter may have a far-reaching effect on the conclusions.

Sea level is governed by various scale phenomena, from global to regional. Global signals due to climate change and variability are having a typical temporal scale in the order of decades to centuries. Interannual variability may be modulated by coupled ocean-atmosphere oscillations such as El Niño-Southern Oscillation (ENSO) and North Atlantic Oscillation (NAO). At annual and seasonal time-scale sea level is driven by regional forces such as monsoon system in the Pacific and Indian Ocean or by the annual cycle of the Tropical Atlantic Variability in the Atlantic Ocean.

Global mean sea level has been rising at a rate of 1.7 mm yr⁻¹ since the 19th century due to the climate change. The sea level rise is faster in recent decades, with the rate of 3.3 mm yr⁻¹ established using satellite data and 2.8 mm yr⁻¹ from tide gauge records. In the Singapore Strait, mean sea level has been rising at the rate 1.2–1.7 mm yr⁻¹ during past 35 years.

At smaller temporal and spatial scales rising and dropping of mean sea level is regionally driven; and higher rates are observed at western tropical Pacific encompassing East Sea of Vietnam/South China Sea (ESV/SCS). Across the Pacific and Indian Oceans, sea level is affected by the ENSO teleconnection, observed through the increase in surface pressure over the Australian and Indonesian regions, the decrease in air pressure over the eastern and central Pacific, the weakening in trade winds in the south Pacific, or the redistribution of warm waters toward the east Pacific. In the ESV/SCS, interannual variability of current and sea level is highly correlated with ENSO. ENSO signals are carried into the ESV/SCS probably by the intrusion of waters from the Pacific through the Luzon Strait. In the Singapore Strait, recent work displayed a highcorrelation (0.7) between an ENSO index, Multivariate ENSO Index (MEI), and annual sea level variability having range ±5 cm, where drops are associated with El Niño events, and rises are correlated with La Niña episodes.

At annual scale, ESV/SCS experiences seasonal sea level variability dominated by the northeast (NE) monsoon during November–February and southwest (SW) monsoon during June–August. At Sunda Shelf, sea level surge due to monsoon wind is amplified by the shallow depth. During NE monsoon level anomaly is positive and in the order of 30 cm; while it is negative in the order of 20 cm during SW monsoon.
Present approaches to resolve the missing data problem rely on either EOF-based technique, or sea level records themselves. The first approach\textsuperscript{15,16} reconstructs the sea level using dominant modes from external model runs or satellite altimetry observations. Though it is successful in obtaining global mean sea level, resolution of reconstructed sea level from this technique is still coarse, which is inapplicable to regional scale where distribution of tide gauge station is dense and the accuracy requires higher such as the Singapore Strait. In addition, the approach is complicated for those who prefer a quick and robust solution, and also limited by availability of data such as satellite observation (only since 1993).

The second approach seeks for a simpler reconstruction directly using tide gauge records. For example, Wahl et al.\textsuperscript{4,6} proposed the k-factor method to compute mean sea level from mean tide level, before applying the Monte-Carlo autoregressive padding to reconstruct the missing time series. This approach uses only tide gauge records themselves to fill the missing data, while neglecting the fact that sea level might be governed by a phenomenon having continuous records.

In this paper, we propose a novel approach which assumes that in addition to tide gauge records, external correlated forcings (such as regional coupled atmosphere-ocean phenomena and indices) could be successfully incorporated into the reconstruction process. For demonstration, wind field from the ESV/SCS and ENSO index are good candidates for the purpose. The structure of the paper is as follows: at first, datasets and methods to represent mean sea level are defined; then, results and associated discussions are described; finally, main findings are summarized in conclusions.

Materials and Methods

We illustrate the data reconstruction methodology using monthly mean sea levels at tide gauge stations in the Singapore Strait, which consist of major regional phenomena and signals including annual Asian monsoon, interannual variability of the ENSO, and global warming. Research-quality tide gauge data \( h(t) \) are obtained from the Permanent Service for Mean Sea Level\textsuperscript{21} (http://www.psmsl.org /data /obtaining/). At this study, we assume that functions describing these phenomena are orthogonal (independent) in the sense that they govern different time-scales by different physical mechanisms. Thus, total mean sea level in Singapore Strait (\( \eta \)) with respect to time (\( t \)) being represented as the sum of mean sea level changes induced by Asian monsoon (\( \eta_m \)), variability of the ENSO (\( \eta_e \)), and global warming (\( \eta_g \)) such that

\begin{equation}
\eta(t) = \sum \eta_i(t) = \eta_m(t) + \eta_e(t) + \eta_g(t) \quad \ldots (1)
\end{equation}

Analyzing tide gauge records in conjunction with satellite altimetry and wind reanalysis in the ESV/SCS, Tkalich et al.\textsuperscript{14} suggested that the wind over its central part is arguably the most important factor determining the observed daily-to-annual variability of sea level in the Singapore Strait, lagging by 1-2 days. We compute monthly averaged wind at location 110-115\degree E, 15-20\degree N from the ERA-Interim dataset, the latest global atmospheric reanalysis produced by the European Centre for Medium-Range Weather Forecasts dating back to 1979\textsuperscript{17}. Following suggestion of Tkalich et al.\textsuperscript{14} we assume the monsoon-induced mean sea level be linearly proportional to wind shear stress (\( \tau \)) with the lag \( \phi_m \)

\begin{equation}
\eta_m(t) = a_m + b_m \frac{\tau(t+\phi_m)}{\|\tau(t+\phi_m)\|} \quad \ldots (2a)
\end{equation}

The wind stress is calculated from the square of wind speed (at 10m above the sea surface) (\( W \)) and drag coefficient (\( C_z \)) following Wu\textsuperscript{18}; the scaling operator \( \|.\| \) is defined as \( \|z\|=\max|z(t)| \). At the monthly scale, we can neglect the lag to obtain a form of the monsoon-induced mean level as

\begin{equation}
\eta_m(t) = a_m + b_m \frac{C_z(W(t))^2}{\|C_z(W(t))^2\|} \quad \ldots (2b)
\end{equation}

In relating sea levels to ENSO events, the MEI is used\textsuperscript{19, 20}. Starting from 1950, the index consists of sea-level pressure, zonal and meridional components of the surface wind, surface temperature and total cloudiness fraction of the sky, which have been observed over the tropical Pacific for many years. Our preliminary analysis reveals that ENSO events lead the sea level anomaly in Singapore Strait by about 2 months (\( \phi_e=2 \) month). It is suggested that there is a linear dependency between this index and ENSO-induced sea levels in the Singapore Strait\textsuperscript{9}.

\begin{equation}
\eta_e(t) = a_e + b_e \frac{\text{MEI}(t+\phi_e)}{\|\text{MEI}(t+\phi_e)\|} \quad \ldots (3)
\end{equation}
Global mean sea level has been rising due to the global climate change. In the Singapore Strait, the measured regional mean sea level is also rising. Thus, we describe this dependency as

\[ \eta_g(t) = a_g + b_g t \quad \cdots (4) \]

In summary, equations (1), (2), (3), (4) with definition \( a=a_m+a_e+a_g \) give us the representation of mean sea level through its relationship with the governing phenomena

\[ \eta(t) = a + b_m \frac{C_z(t,W)W^2(t)}{\|C_z(t,W)W^2(t)\|} + b_e \frac{\text{MEI}(t+\phi_e)}{\|\text{MEI}(t+\phi_e)\|} + b_g t \quad \cdots (5) \]

The set of unknown parameters \((a, b_m, b_e, b_g)\) in equation (5) are determined using data from tide gauge records. For each station, records during a given period are characterized by a distinct set of parameters. When these parameters are derived, the equation (5) defines the reconstructed time series. The cost function is the sum of squared residuals between reconstructed and observed time series.

\[ \varepsilon^2(a, b_m, b_e, b_g) = \sum_t [\eta(t) - h(t)]^2 \overset{\text{min}}{\rightarrow} \varepsilon_{\text{min}}^2 \quad \cdots (6) \]

We use least square fitting to minimize this cost function, which leads to the conditions

\[ \frac{\partial \varepsilon^2}{\partial a} = 0; \quad \frac{\partial \varepsilon^2}{\partial b_m} = 0; \quad \frac{\partial \varepsilon^2}{\partial b_e} = 0; \quad \frac{\partial \varepsilon^2}{\partial b_g} = 0; \quad \cdots (7) \]

Equations (7) gave us a system of linear equations to determine the unknowns \( x=(a, b_m, b_e, b_g)^t \).

\[ \mathbf{A} x = \mathbf{y} \quad \cdots (8) \]

where \( \mathbf{A} = \{A_{ij}\}_{i,j=1,4} \) and \( \mathbf{y} = \{y_j\}_{j=1,4} \). In fact, elements of matrices \( \mathbf{A} \) and \( \mathbf{y} \) can be defined in a more general case that adds more forcings into consideration.

Here, we explicitly got elements of the matrix \( \mathbf{A} \) as

\[ A_{11} = N \]
\[ A_{22} = \sum_t \left( \frac{C_z(t,W)W^2(t)}{\|C_z(t,W)W^2(t)\|} \right)^2 \]
\[ A_{33} = \sum_t \left( \frac{\text{MEI}(t+\phi_e)}{\|\text{MEI}(t+\phi_e)\|} \right)^2 \]
\[ A_{44} = \sum_t t^2 \]
\[ A_{12} = \sum_t \frac{C_z(t,W)W^2(t)}{\|C_z(t,W)W^2(t)\|} \]
\[ A_{13} = \sum_t \frac{\text{MEI}(t+\phi_e)}{\|\text{MEI}(t+\phi_e)\|} \]
\[ A_{14} = \sum_t t \]
\[ A_{23} = \sum_t \frac{C_z(t,W)W^2(t)}{\|C_z(t,W)W^2(t)\|} \frac{\text{MEI}(t+\phi_e)}{\|\text{MEI}(t+\phi_e)\|} \]
\[ A_{24} = \sum_t \frac{C_z(t,W)W^2(t)}{\|C_z(t,W)W^2(t)\|} t \]
\[ A_{34} = \sum_t \frac{\text{MEI}(t+\phi_e)}{\|\text{MEI}(t+\phi_e)\|} t \]

and the matrix \( \mathbf{y} \) as

\[ y_1 = \sum_t h(t) \]
\[ y_2 = \sum_t \frac{C_z(t,W)W^2(t)}{\|C_z(t,W)W^2(t)\|} h(t) \]
\[ y_3 = \sum_t \frac{\text{MEI}(t+\phi_e)}{\|\text{MEI}(t+\phi_e)\|} h(t) \]
\[ y_4 = \sum_t t \cdot h(t) \quad \cdots (10) \]

We conducted 5 experiments to reconstruct monthly mean sea level at Raffles Lighthouse during 1980-2010 using different configurations given in Table 1. The method is then applied to reveal the characteristics of tide gauge records at Sembawang, Sultan Shoal and Tanjong Pagar.
Results and Discussions

Time series of reconstructed mean sea level in EXP A captures the seasonal variability in mean sea level as displayed in Fig. 1a. The reconstructed data are correlated with observations with coefficient 0.69 and have a root mean square difference (RMSD) of 5.3 cm against the tide gauge records (Fig. 3). Since this experiment considers only wind, it is in line with previous study of Tkalich et al.\textsuperscript{14} who showed that wind stress over the ESV/SCS is the major driver for monsoonal variability of sea level in the Singapore Strait.

The experiment EXP B explores the relationship between the ENSO teleconnection and regional mean sea level in the Singapore Strait. The correlation between MEI and monthly mean sea level is 0.43 in which the deviation is smaller than that of observation (Fig. 3). The reconstructed time series seem to be the moving average of monthly tide gauge records (Fig. 1a). In fact, at annual scale, mean sea level at Raffles Lighthouse is highly modulated by ENSO with the correlation coefficient of 0.7\textsuperscript{9}.

In the experiment EXP C, we take into account only the effect of global warming (Equation 4). The reconstructed time series are very close to the regression line, with low correlation coefficient as of 0.26 (Table 1). Obtained parameter exhibits an increasing rate of 2.3 mm yr\textsuperscript{-1} (Fig. 1a and Table 2).

When all three components involving monsoonal wind, ENSO and global warming are simultaneously used in the experiment EXP E, the reconstructed time series become highly correlated with tidal records (Fig. 1a), with the correlation coefficient 0.77 and the RMSD of 4.7 cm (Table 1). The improvements are clearly seen during both NE and SW monsoon seasons (Fig. 1). As the monthly mean sea levels peak at NE monsoon and trough at SW monsoon, the enhancement is also observed in the statistical comparison (Fig. 2). The experiment suggested that the combination of these three forces is essential to describe mean sea level variability in Singapore waters.

In the experiment EXP D where effect of global warming is taken into account, the correlation coefficient is 0.76 (Table 1). The RMSD is 4.8 cm, which is slightly lower than in the previous experiments. The improvements are clearly seen during both NE and SW monsoon seasons (Fig. 1).

<table>
<thead>
<tr>
<th>Experiment</th>
<th>(b_m)</th>
<th>(b_e)</th>
<th>(b_g)</th>
<th>CORR</th>
<th>RMSD (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXP A</td>
<td>✓</td>
<td></td>
<td>0.69</td>
<td></td>
<td>5.3</td>
</tr>
<tr>
<td>EXP B</td>
<td>✓</td>
<td></td>
<td>0.43</td>
<td></td>
<td>6.6</td>
</tr>
<tr>
<td>EXP C</td>
<td>✓</td>
<td>✓</td>
<td>0.26</td>
<td></td>
<td>7.1</td>
</tr>
<tr>
<td>EXP D</td>
<td>✓</td>
<td>✓</td>
<td>0.76</td>
<td></td>
<td>4.8</td>
</tr>
<tr>
<td>EXP E</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>0.77</td>
<td>4.7</td>
</tr>
</tbody>
</table>

1—Mean sea level time series for Raffles Lighthouse (1980-2010) for (a) individual components and (b) their different combination.
warming is eliminated, the correlation coefficient reduces to 0.76, and RMSD error increases by 0.1 cm in comparison to the experiment EXP E (Fig. 3). The peaks during recent NE monsoon seasons are not depicted accurately (Fig. 1). Thus, the global warming should be incorporated into the reconstruction of missing data.

Results are similar for other tide gauges in Singapore waters including Sembawang, Sultan Shoal, and Tanjong Pagar. Wind stress over the ESV/SCS is the dominant factor controlling the variability of mean sea level in the Singapore Strait, followed by ENSO and the global warming. Added nonlinearities also improve the representation of mean sea level in the Strait.

Fitted coefficients of reconstructed mean sea level for these tide gauge station data are given in Table 2. It shows that under the same forcing, measured mean sea level responses differently. For instance, monsoonal dimensionless parameters ($b_m$) for Sembawang and Tanjong Pagar tide gauge data are 0.62 and 0.54, respectively, which is as double as (or 50% higher than) these for Sultan Shoal (0.35) and Raffles Lighthouse (0.45). It implies that influences of wind from the ESV/SCS on the mean sea levels at Sembawang and Tanjong Pagar are much stronger than these at Raffles Lighthouse and Sultan Shoal. This result is consistent with climatological results of Tkalich et al. (2013), who related this difference to the larger exposure of Sembawang and Tanjong Pagar locations to the phenomena in the ESV/SCS.

An interesting implication from these results is that once we archive a characteristic set of parameters for each tide gauge, we could execute statistical downscaling of global and regional signals to Singapore Strait.

Conclusions

In this paper, we have presented a simple and
robust method to reconstruct the missing data at tide gauge records. The method utilizes quantification of the underlying phenomena to build a set of parameters, which is then used to reconstruct the variability and trend of monthly mean sea levels. We demonstrated the method at Raffles Lighthouse data, and further verified at three other tide gauge stations including Sembawang, Sultan Shoal and Tanjong Pagar. Our reconstruction is qualitatively consistent with previous findings on the dominant role of seasonal wind from ESV/SCS, the interannual impact of ENSO, and the long-term global sea level rise.

We derived a characteristic set of parameters for selected four tide gauges in the Singapore Strait. These sets quantify distinct effects of monsoonal wind, ENSO and global warming at each tide gauge. An important application of this set is to project the mean sea level in Singapore Strait to the future using global model projections of regional wind speed, ENSO and global sea level.

The reconstruction successfully captures the basic variability in the mean sea levels; however, representations of peaks during many NE monsoons are not very well resolved. It implies that additional governing physics or alternative forms of nonlinearity and their combinations are yet to be identified. This would be an interesting subject for a follow up study.

Acknowledgements

Authors thank Guillaume Maze for the script plotting Taylor diagram.

REFERENCES