Journal of Geodetic Science
• 1(3) • 2011 • 221-232 DOI: 10.2478/v10156-010-0025-0 •

Empirical Modeling and Impact of Transient Effects on the Mean Sea Level Trend Estimates from the Global Tide Gauge Data

Research article

H. Bâki Iz and H. M. Ng*

Department of Land Surveying and Geo-Informatics The Hong Kong Polytechnic University Hong Kong S.A.R., China.

Abstract:
The mean sea level trend estimates from shorter records (less than 50 years) are easily influenced by a number of additional transient effects including atmospheric pressure variations, interannual and decadal changes in the mean sea level or various station dependent disturbances, which are not always accounted for in previous studies because their influences may be negligible for stations longer than 100 years or simply such information may not have been available. In this study, we detected and modeled 8,072 transient changes (a nearly periodic, or non-periodic variation, a shift or an episodic change in the mean sea level that may last several months or longer) from all of the globally distributed 1,862 tide gauge stations with approximately 47,000 years of tide gauge data in the Permanent Service Mean Sea Level repository. It was shown that 1,264 out of 1,862 globally distributed tide gauge station solutions were affected significantly by modeling transient changes in the mean sea level. The solutions R^2 values improved at a 95 % confidence level with the inclusion of new empirical model parameters representing transient changes as mean shifts as well as the trend estimates that fall within [-1 to +3] mm/year interval.

Keywords:
Mean Sea Level • Tide Gauge • Global MSL Trend • Climate Disturbances

Received 13 April 2011; accepted 16 May 2011

1. Introduction

Understanding the global and local mean sea level (MSL) changes is important for global climate change studies. The Intergovernmental Panel on Climate Change (IPCC) reported that global average sea level rose at an average rate of 1.8 [1.3 to 2.3]mm per year over 1961 to 2003 and at an average rate of about 3.1 [2.4 to 3.8]mm per year from 1993 to 2003 (IPCC, 2007), which largely relied on the long-term (>50 years) tide gauge records distributed around the world (Bindoff et al., 2007).

The recent MSL trend analyses use the data sets extracted from the Permanent Service for Mean Sea Level (PSMSL) repository (Woodworth & Player 2003). In these studies, the number of available tide gauge (TG) records is restricted by the different selection criteria such as completeness of the records, referencing to different geographic regions, exclusion of records in tectonically active regions (Gutenberg 1941; Fairbridge & Krebs 1962; Gornitz et al. 1982; Barnett 1983a; Peltier & Tushingham 1989; Douglas 1991), by the record lengths (Gornitz et al. 1982; Gornitz & Lebedeff 1987; Barnett 1983a; Douglas 1991), and as a result of careful corrections and verification of tide gauge vertical datum and benchmark changes. Currently, only less than 10 % of the existing tide gauge records span over 100 years and the restricted number of globally distributed tide gauge stations biases the global MSL trend estimates, thereby there is an urgent need for the use of additional tide gauge stations with shorter records. Nevertheless, the estimates from shorter records are easily influenced by unmodeled effects

*E-mail: lshbiz@polyu.edu.hk

VERSITA

such as local and regional interannual or decadal fluctuations in the MSL which can not be eliminated or reduced effectively using tide gauge data averaged over long time intervals (Iz 2006).

In this study, we examine the existing PSMSL tide gauge records and show that 1,264 out of 1,862 globally distributed tide gauge stations in the PSMSL repository are significantly influenced by transient MSL changes, sometimes abrupt, but mostly subtle MSL transient variations that are difficult to detect. These changes occur, in addition to the well-known periodic MSL variations due to ocean tides, not well known variations caused by lunar nodal spectrum and changes, long-term mass, and volume changes due to melting of ice and thermal expansion, salinity, etc. Some of these changes are induced by almost periodic components of air pressure variations, global and regional temperature changes, interannual and decadal variations in ocean circulation including wind, episodic sea level changes caused by eddies, thermal and salinity (steric) variations, nearby ice-sheet/glacier melt or river discharges, changes in the ocean basin due to the processes including glacial isostatic adjustment or tectonics. Local vertical crustal movements due to tectonic motion, anthropogenic subsidence and glacial isostatic adjustment, elastic loading from atmospheric, tide, or hydrologic loading, also affect the tide gauge records (Vanicek 1978). Furthermore, some of these records experiences datum changes caused by instrument or benchmark replacements, and all of them are subject to random measurement errors at each station.

The response of the MSL to these effects can be modeled if additional information exists. The impact of pressure variations, for instance, can be included in the model if local atmospheric pressure data is available (Iz & Shum 1998, Ponte 2006). In other cases, such data are hard to come by for all of the tide gauge stations. In the absence of auxiliary information about their sources, changes in the MSL can be modeled empirically to alleviate their impact on the estimation of local MSL trends. However these changes are sometimes not visually discernable and buried in the noise of the tide gauge data and they may still have an impact on the MSL trend estimates (Iz 2006).

In this long list of factors that affect the MSL time series, periodic variations are represented by trigonometric series in modeling the tide gauge data whereas aperiodic (transient) or almost periodic (episodic) MSL changes require additional information that is not always available for all of the tide gauge series.

Iz (2006) quantified that the impact of transient or episodic changes on trend estimates may be negligible for very long data series (longer than 100 years) but influential for shorter series up to 0.5 mm/year as a function of series length, the epoch of the transient effect and its magnitude. Because of the availability of spatially well distributed stations is equally important for monitoring global MSL changes as well as availability of long records, shorter series can also be used in the pool of tide gauge stations for the analysis of global MSL studies by improving the current models, which are appropriate for the long series but not the shorter ones.

In the following sections, we concentrate on modeling the transient or episodic variations in the MSL. We first give an example in which the tide gauge data were contaminated with transient or episodic changes and discuss an algorithm to detect the presence of these variations and their starting epochs. Subsequently, we establish a model which is blind to the source of the transient effects but account for their lump-sum impact on the MSL trend. These changes are represented by binary (indicator) variables together with trend parameters and trigonometric functions representing the periodic changes in the data. A comparison of a sample solution with and without modeling transient changes demonstrates that systematic effects may unduly bias the trend estimates and influence the solution statistics. Another nearby tide gauge station series is also analyzed to validate the new model. Finally, we use the newly parameterized model for solving the local trend estimates for all the 1,862 stations with the metric tide gauge data in the PSMSL repository including almost all of the global tide gauge station records and analyze the results.

2. Modeling of MSL Tide gauge Data for Local Trend Estimation

The following trigonometric model can be used to explain a number of periodic variations in the PSMSL data including a long trend.

$$y_t = a + b(t - \bar{t}) + \sum_{h=1}^{3}\left[\alpha_h \cos\left(\frac{2\pi}{P_h}\right)(t-\bar{t}) + \gamma_h \sin\left(\frac{2\pi}{P_h}\right)(t-\bar{t})\right] + e_t \quad (1)$$

In this expression y_t represents the monthly averaged PSMSL time series data, a is the y-intercept, b represents the local MSL trend, t is the time tag for the averaged values of the tide gauge data, and $\bar{t}$ refers to the middle epoch of the series. The Greek letters α and γ represent the magnitudes of periodic functions due to the tidal constituents with periods, P_h, including semi-annual, annual, and nodal (18.6 years) periods.

The random variable e_t represents the lump-sum effect of random instrument errors and unmodeled stochastic effects in the MSL with the following properties:

$$E[e_t] = 0, \quad Var[e_t] = \sigma^2. \quad (2)$$

The current solutions assume that the errors are not serially correlated and stationary in time.

The unknown parameters in the above formulation can be estimated from the tide gauge data using the least squares method. The influence of the unmodeled random or systematic effects attenuates with averaging, thereby reducing their impact on the solution. In this solution, unmodeled periodic and almost-periodic variations, if present, are either absorbed by the coefficients of the existing trigonometric model parameters or present in the residuals without significantly biasing the trend estimates. However,

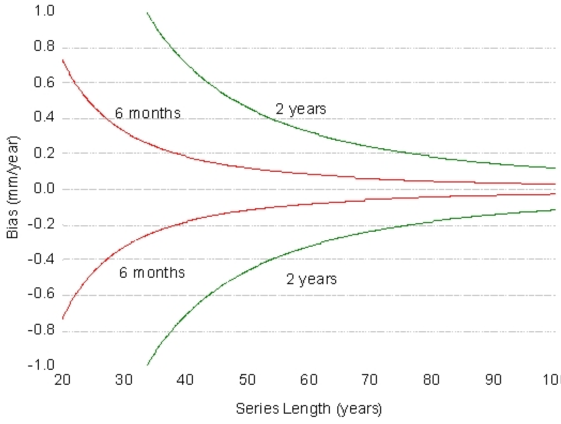

Figure 1. A trend estimate is biased by a transient effect if it is not included in the model. In the above cases, a 100 mm MSL change occurs at the beginning and at the end of each series of different length, and lasts for 6 months and 2 years, respectively (Iz 2006).

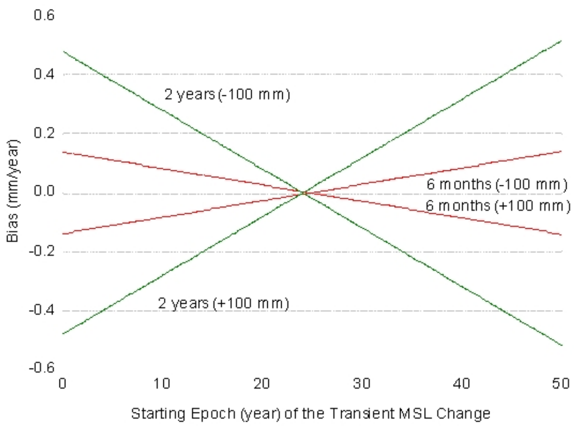

Figure 2. In this graph, the magnitude of the trend estimate bias is a function of the location of a transient MSL change in the series for two scenarios. In the first case, the change occurs at the end of the series, and lasts 6 months; in the latter, it lasts for 2 years (Iz 2006).

as illustrated in Fig. 1 and Fig. 2, the impact of these effects can be significant on the MSL trend estimates if they are not included in modeling tide gauge series shorter than 100 years. Fig. 1, for instance, shows that a 100 mm MSL change that occurs at the beginning and at the end of each series of different lengths and lasts for 2 years can bias the trend estimate of a 50-year long tide gauge series by as much as 0.5 mm/year (Iz, 2006). The maximum bias occurs if the change takes place at the beginning or at the end of the series and the trend bias vanishes if the change happens in the middle of the series (Fig. 2).

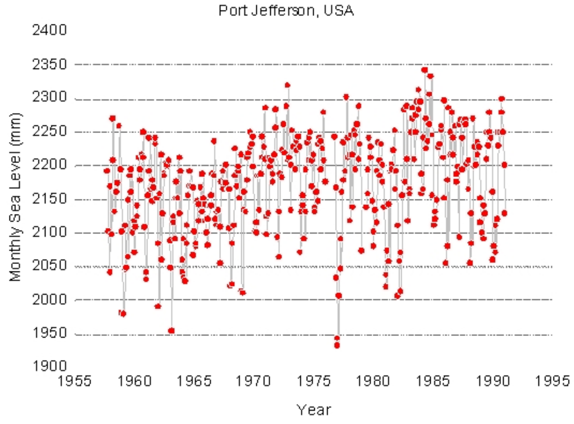

Figure 3. Monthly averaged tide gauge data from Port Jefferson, USA.

3. Detection and Modeling of Transient Changes in Tide Gauge Series: A Sample Solution

In this section, the tide gauge data (Fig. 3), Port Jefferson in the USA, from the PSMSL repository is analyzed as an example to demonstrate the detection and estimation of transient change in the MSL. Another nearby station series, New York City USA, is also considered to validate the new model results. The nearby station tide gauge series coincidentally were not influenced by any transient changes during the same overlapping period with the Port Jefferson station tide gauge data series.

PSMSL considers the Port Jefferson Metric and RLR records to be the same, apart from an overall offset, and classifies it as a good record (Woodworth 2005). The data set spans over 33 years and may contribute to the pool of tide gauge stations in the study of global MSL.

The local trend of the model given by (1) from the Port Jefferson, USA tide gauge station data was estimated using the ordinary least squares solution. The model includes semi-annual, annual, and nodal MSL variations in addition to the trend and the intercept. The estimated trend is 1.9 mm/year with a standard deviation of 0.3 mm/year (Table 1). The R^2 value (the coefficient of determination) of the solution indicates that the model can only explain up to 50 % of the variation in the tide gauge data which is not a very impressive model performance. Fig. 4 shows the residuals of the solution (RMS=53.2 mm) which are now free from any periodic changes in the MSL, but may be contaminated by a number of transient variations, which are also visually evident in the residuals, whose starting epochs can be detected by statistical methods.

In this study, we use the Cumulative Sum (CUSUM) method for the detection of starting epochs, change point (CP) of transient or episodic change in the MSL (mean shifts), from the residuals of solution models that do not include the parameters for such effects. Although there are a number of competing methods for

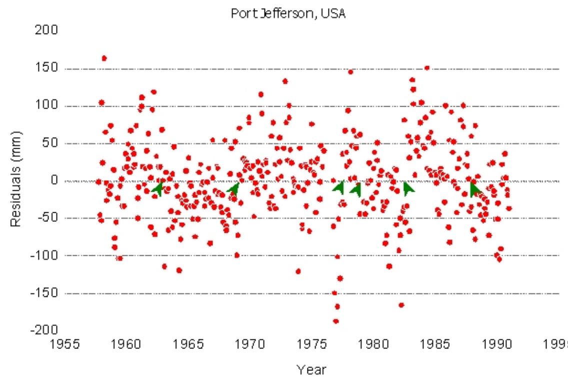

Figure 4. Port Jefferson station residuals (RMS = 53.2 mm) exhibit no visually evident periodic changes in the MSL, but still contaminated by a number of transient variations, whose starting epochs, indicated by green arrows, were detected by the CUSUM algorithm.

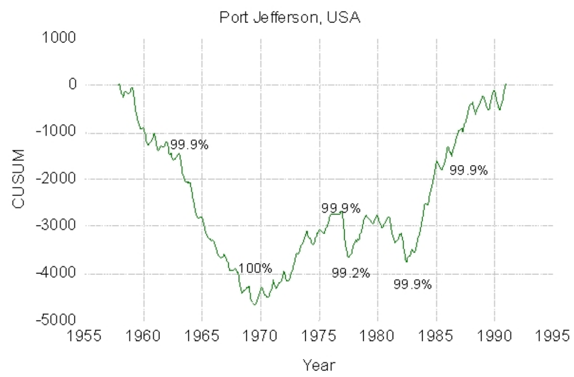

Figure 5. CUSUM Chart for Port Jefferson station residual ranks. The percentages indicate the probability of a change in the MSL at the corresponding epoch.

CP detection (mean shift analysis), our choice is motivated by the method's simplicity and its ease of automation for programming for a large number of solutions. The algorithm is discussed in the appendix.

The CUSUM chart displayed in Fig. 5 created by using ranks of the residuals of the tide gauge series shows the presence of 6 CPs with probabilities larger than 99 %. The probabilities were calculated by bootstrapping the relevant portion of the series with 1,000 replications. As discussed earlier, these transient changes can be eliminated if their origins are known, by preprocessing the tide gauge data using auxiliary information in the analysis. For instance, MSL changes induced by the pressure variations can be eliminated by incorporating barometric information into the model (Iz and Shum 2000). Alternatively, these changes can be solved using other model parameters regardless of their origins and without requiring any additional information about their sources.

We introduce binary variables into the model (1) for this purpose. The new independent variables are linked to the unknown coefficients which represent the magnitude of the linear changes in the MSL (mean shifts). Binary variables, also known as dummy or indicator variables in statistical applications (Neter et al. 1996), take the value 0 if there are no changes in the mean of tide gauge data or 1 when there is a change in the MSL for a period of time until the next change occurs.

Taking into consideration the statistically significant MSL changes and their epochs that are determined from the CUSUM analysis of residuals of the previous solution, we formulate a new model for the tide gauge observations at the Port Jefferson station as,

$$\begin{aligned} rcly_t &= a_0 + \sum_{i=1}^{n} \Delta a_t^i \delta + b(t-\bar{t}) \\ &+ \sum_{h=1}^{3}\left[\alpha_h \cos\left(\frac{2\pi}{P_h}\right)(t-\bar{t}) + \gamma_h \sin\left(\frac{2\pi}{P_h}\right)(t-\bar{t})\right] \\ &+ \varepsilon_t \end{aligned} \tag{3}$$

In the above expression, δ is the binary variable which is 0 for a tide gauge record if there is no change in the residuals of observations, and 1 if there is a change in the mean value of the residuals of observations as determined by an earlier solution using equation (1), and CUSUM analysis. If the total number of CPs (mean shifts) is denoted by n, then $i = 1, \ldots, n$, and Δa_t^i is the amount of MSL change to be estimated with respect to the previous change in the MSL for the i^{th} station data at an epoch t. Note that this estimate can also be made to refer to a change with respect to a datum defined by the intercept with an appropriate arrangement of the elements in the design matrix of the observation equations. This model is used again to estimate the unknown parameters, including the mean shift parameters from the Port Jefferson tide gauge station data. The new model for the Port Jefferson data has now 6 additional parameters. In this solution, modeling the mean shifts resulted in a substantial 58 % decrease in the trend estimate from 1.9 mm/year down to 0.8 mm/year, despite the modest 5% improvement in the R^2 and the RMS values of the residuals (Table 1). This is partly because the new models, a more realistic representative of the data, enabled a better separation of the trend parameter from the other model parameters, and mostly because two of the 6 CPs are located at both ends of the series which are influential in estimating the MSL trend (Iz and Shum 2000, Iz 2006).

The residuals displayed in Fig. 6 are now free from marked systematic variations that were present in the earlier solution's residuals. This sample solution demonstrates that transient MLS variations, if not included in the solutions, can significantly influence the trend estimates as also shown theoretically by Iz (2006) for series as long as 40 to 50 years.

Notwithstanding the marked changes in the estimates and the improvement in the coefficient of variations, there is a slight increase

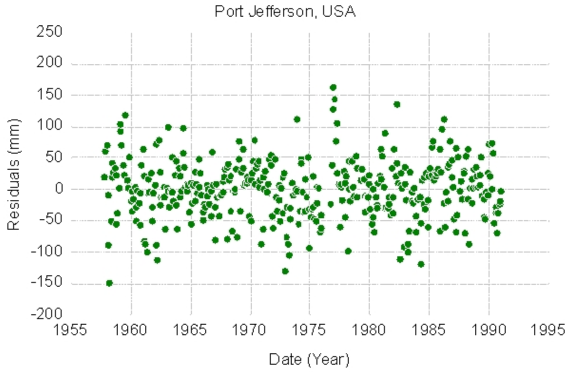

Figure 6. Residuals from the least square solution using the model that includes the transient changes.

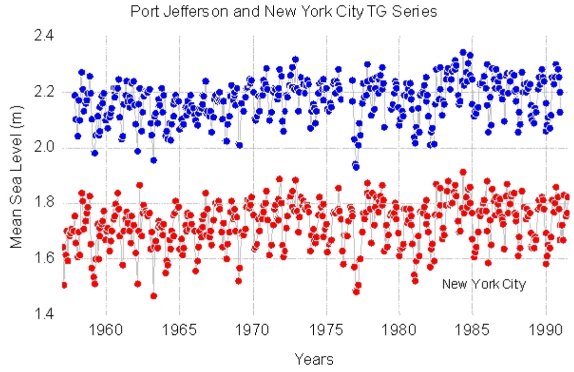

Figure 7. Port Jefferson (top) and New York City tide gauge time series are displayed. The tide gauge stations are approximately 100 km apart from each other. There are no statistically significant change points in the New York tide gauge data during the same period.

in the standard error of the newly estimated trend parameter (0.3 vs 0.4 mm/year in Table 1), a phenomenon which also exists in other solutions to be reported in subsequent sections. The introduction of the change points as new parameters in the new models increased the standard errors of the trend estimates and seems counterintuitive at a first glance since one will expect improvements in the standards error to decrease rather than increase in magnitude. Yet, as it was shown by Potluri (1971) that, we quote his Theorem 2: In the classical regression model, omission of a variable specified by the truth decreases the variance of all the least square estimates. Hence, if our equation 3, which is a trigonometric regression model, is the true model, then the omission of the CPs will decrease the standard errors of the estimates. Moreover, it is also well known that even the introduction of spurious parameters do not bias the estimates in a least squares solution as long as the new parameters do not create harmful collinearity. In any case, we analyzed another nearby tide gauge station data to validate the new model solution which is the topic of the next section.

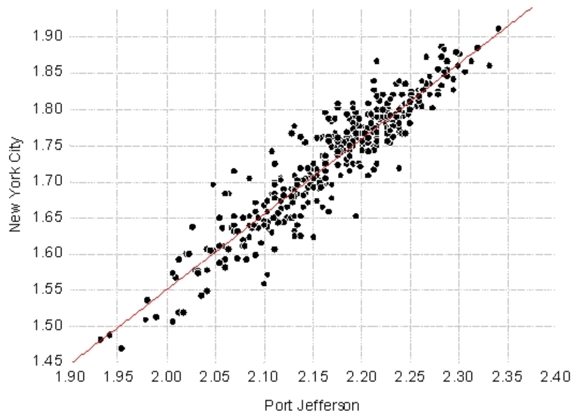

Figure 8. Tide gauge data from the New York City and Port Jefferson stations are highly correlated (the PearsonŠs correlation coefficient is 0.93). The plot is generated using the monthly averaged data from both stations during 1958-1991.

4. Model Validation

The New York City USA tide gauge station is located approximately 100 km away from the Port Jefferson station. Its time series data shown in Fig. 7 were analyzed to validate the new model solution. Both series are in close agreement over the common 33 year period and exhibit a strong correlation, 0.93, (Fig. 8) because of their proximity.

Table 1 shows the trend estimates for the New York City tide gauge series using models with and without CPs parameterization data for the overlapping 1958-1991 period. As it turned out, there are no statistically significant CPs in the New York City time series during the 33 years period, which also suggest that the mean shifts in the Port Jefferson data are station specific changes rather that caused by major unmodeled climatic disturbances.

Table 1. Solution statistics with and without transient changes for modeled (CP) for the Port Jefferson and New York City TG time series data. The bootstrapped probabilities for all the CPs are larger than 99 %. Note that there are no statistically significant CPs detected in the New York City time series during the 33 years period.

Station Name Series Length	Port Jefferson, USA 33 Years		New York City, USA 33 Years
	No CP	6 CPs	No CPs
RMS residuals (mm)	53.2	48.5	58.3
R^2	0.50	0.55	0.33
Trend (mm/year)	1.9±0.3	0.8±0.4	1.0±0.3

The 0.9 mm/year difference in the estimated Port Jefferson and New York City tide gauge solutions (1.9±0.3 and 1.0±0.3 mm/year respectively) without modeling CPs is statistically different as revealed by a two-tailed t-test at 95 % confidence level. For the new model solution however, the trend estimate for the Port Jefferson station is 0.8±0.4 mm/year is in better agreement with the 1.0±0.3 mm/year New York City trend and the null-hypothesis for the 0.2 mm/year difference in the trends cannot be rejected at 95 % confidence level using a two-tailed t-test.

These results support unequivocally the appropriateness of the new model parameterization using transient effects. In the next section, we extend the new modeling and analysis approach in detecting the transient changes in the MSL in the series, and in estimating the local MSL trends for all of the globally distributed tide gauge stations in the PSMSL repository.

5. Global Tide Gauge Data used in the Analysis of MSL Trend Estimates with Empirically Modeled Transient Effects

Permanent Mean Sea Level, PSMSL, repository maintains a tide gauge database from over 1800 stations since 1933. About 200 national authorities around the world provide data to PSMSL in the monthly and yearly formats. Until February 2001, PSMSL contained more than 47,000 station-years of sea level data and received approximately 2,000 station-years of data each year (Woodworth & Player 2003).

PSMSL offers Metric and Revised Local Reference (RLR) data (Permanent Service for MSL 2001). Metric data is the raw data directly received from the authorities. The RLR data contains monthly and annual MSL data referenced to a common datum. The datum is defined 7 m below the global MSL to avoid negative monthly and annual MSL values. Only two thirds of the stations in the PSMSL database, however, have been adjusted to a common datum. The recent MSL trend analyses use the data sets extracted from the PSMSL repository. The earlier studies utilize the Revised Local Reference (RLR) data set, which contains those stations that passed a consistency check. At each station, the RLR data are reduced to a common datum according to their benchmark histories by the PSMSL. Currently, only about two-thirds of the stations are converted into the RLR data set.

In the following solutions, all of the monthly PSMSL Metric data are used because the new models can accommodate most of the problem areas in this complete data set. This is the entire data in the repository in 2001, which span records up to 181 years for various stations for a total of 1,862 tide gauge stations.

6. Estimated Mean Shift Parameters

Two sets of solutions were calculated for comparison using models based on (1) and (3) for all the globally distributed tide gauge stations as we did in the previous sections for the Port Jefferson tide gauge station data. The first solution is a special case of the second one in which all of the binary variables are set to zero, i.e., it is assumed that there are no transient changes in the series and the second solution includes all of the statistically significant CPs (at 95% confidence level) calculated from the residuals of the first solution. As before, the random tide gauge data errors were assumed to be homogeneous with zero expected values as described by (2).

The solutions show that 1,264 stations were affected by transient changes out of 1,862 tide gauge stations in the repository (68 %). Overall, 8,072 statistically significant CPs are detected at 95 % confidence level for the 1,264 stations with an average of 2 CPs per station per decade. Fig. 9, the histogram of the estimated magnitudes of the transient changes in the MSL, exhibits a nearly zero mean bell shaped distribution and contaminated with a number of very large variations on both tails, which can be attributed to datum changes due to instrument relocations, subsidence, etc. because of their unusually larger magnitudes (remember that we are using the large pool of metric data rather than the reduced level data from the PSMSL repository). The 67% of the estimated mean shifts at statistically significant CPs falls within a [-15, +15] cm interval. This result is in agreement with the transient changes induced by eddies, temperature, and pressure variations to nearby tide gauge stations and shows an overall equilibrium for the globally distributed tide gauge stations (Pugh 2004).

Fig. 10 is a snapshot of the distribution of the cumulative sum of estimated mean shift parameters at each station calculated for all 1,264 stations. Over 86 % of the stations' cumulative mean shift estimates are within a [-15, +15] cm interval. The large number of mean shifts at the tail values is again likely to be due to the same station location-specific events, such as instrument relocation, that are more persistent, rather than transient in nature. The observed symmetry in the histogram of the cumulative mean shifts at each station (Fig. 10) is indicative of a global equilibrium in the transient MSL changes that are experienced by the tide gauge stations all over the world.

As far as the standard errors of the estimated mean shift parameters are concerned, over 75 % them are below ±3 cm as shown in Fig. 11. Only 14% of the standardized mean shift estimates, which are calculated by dividing the estimates by their standard deviations, fall within the interval of a ±1 signal-to-noise ratio (Fig. 12) evidencing the widespread statistically significant presence of transient changes in the MSL globally.

6.1. Estimated MSL Trend Parameters

The histogram in Fig. 13 displays the magnitudes of the local MSL trend estimates calculated from two different solutions for all of the global 1862 stations using equation (1) and (3), with and without CPs included. Although only 1,264 stations are affected by the CPs, the histogram is created including all stations to give a better overview for the effect of the modeling transient changes in the MSL.

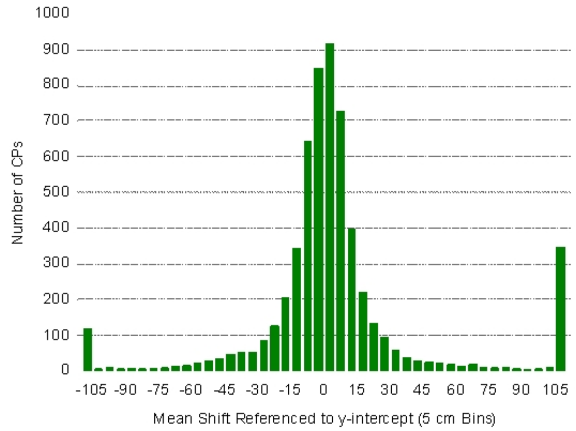

Figure 9. The frequency and magnitude of the estimated mean shifts with respect to the y-intercept of each tide gauge station solution.

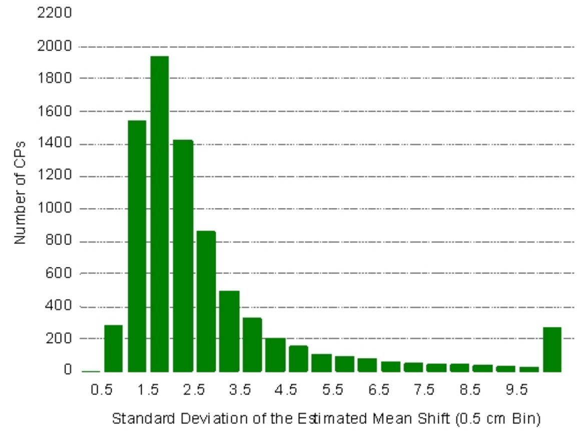

Figure 11. The distribution of the standard errors (deviations) of the estimated mean shifts (based on the a posteriori variance of unit weight).

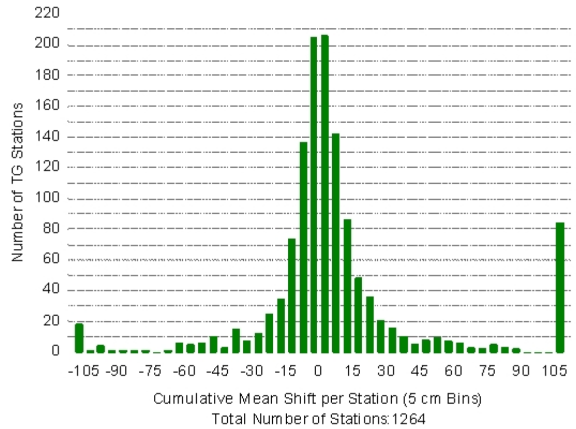

Figure 10. The distribution of the cumulative sum of the estimated mean shift parameters at each station for a total 1264 stations.

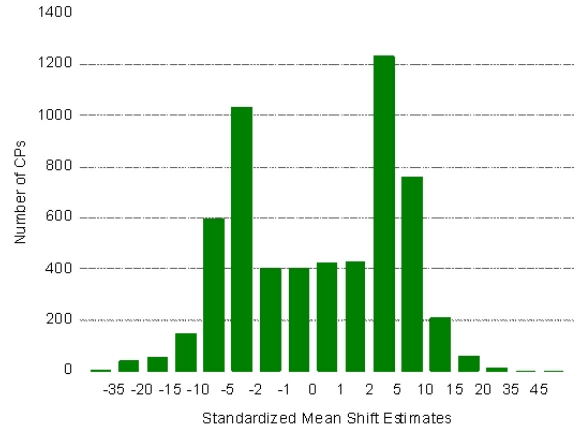

Figure 12. Standardized mean shift estimates (calculated by dividing the mean shift estimates by their standard deviations). Note that the bin intervals are not equal.

The impact of modeling transient changes in the MSL is an overall change in the magnitudes of MSL trend estimates which fall within the interval [-1, +3] mm/year as illustrated in Fig. 13, which refers to the estimates as obtained directly from the least squares solutions. The number of trend estimates within this interval was decreased with the inclusion of the transient changes in the models. The effect of post glacial rebound, PGR calculated from the ICE-4G (VM2) model (Peltier 1994, 1996, 1998, 2001), is included in Fig. 14. Almost all of the changes in the MSL trends occur for shorter tide gauge series as shown Fig. 15, which is produced using the difference of the trend estimates calculated with model solutions with and without parameterization for transient effects. This is not a surprising result considering the theoretically demonstrated dependency of the unmodeled MSL level variations on the series length (Iz 2006). If we assume that the model solutions with transient effects are better, then the difference between them and the model solutions that use the transient change as models parameters is the model bias due to the omission of the effect of the transient changes in the MSL. Given the fact that almost all of the differences in Fig. 15 follow the same trend defined by the theoretically expected bias in the estimates (scaled solid line) as a result of unmodeled systematic effects displayed in Fig. 1, support this assumption.

Table 2 shows 21 tide gauge stations' estimated trends sampled from nine oceanic regions together with their estimates from previous studies. The trend columns in Table 2 are the estimated MSL trends in mm/year with and without CPs. All of the trend values are corrected for the effect of PGR calculated from the ICE-4G (VM2) model (Peltier 1994, 1996, 1998, 2001). Some of the

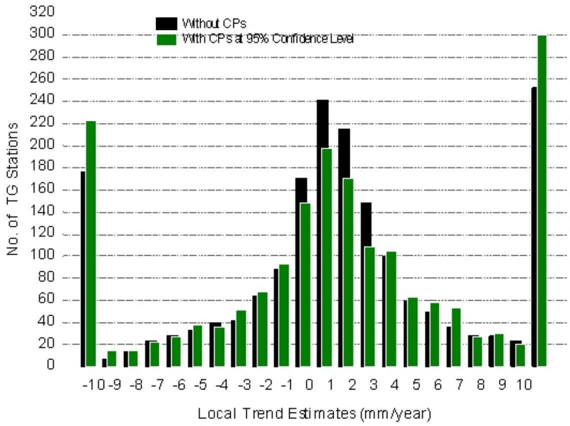

Figure 13. Local MSL trend estimates calculated using models with and without CPs for all of the stations in the PSMSL repository. No PGR corrections were applied.

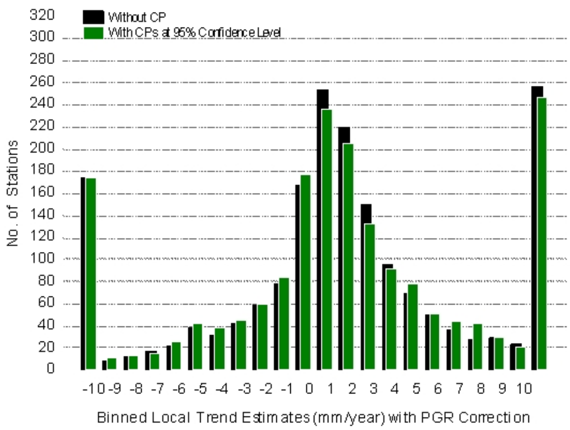

Figure 14. Local MSL trend estimates calculated using models with and without CPs for all of the stations in the PSMSL repository. The MSL trend estimates are corrected for the PGR effect.

station trends change significantly with the application of the ICE-4G model for PGR. There are also significant differences between with CP and without CP solutions. Nevertheless, those solutions with CPs that are in better agreement with the previous estimates may have included data such as atmospheric pressure to account for the effect of transient changes. The differences between earlier solutions and the solutions generated in this study, where transient MSL variations are not modeled in both cases, are mostly due to the additional data (over a decade) in the current solutions which are influential on the estimates (see Iz & Shum 1998, and Iz 2006 for further discussions on the end effects of data in tide gauge series).

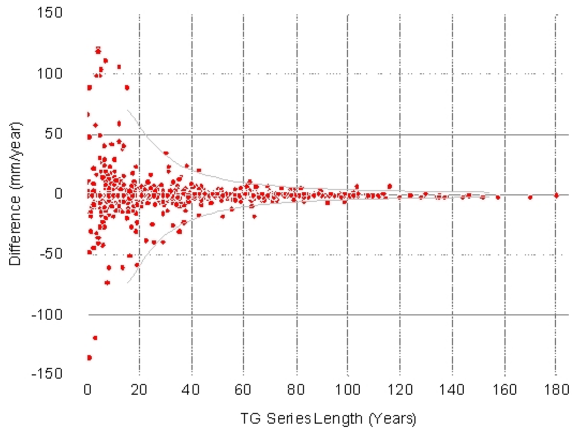

Figure 15. Changes in the local trend estimates as a result of modeling transient effects. The differences follow the theoretical bias pattern (Iz 2006), shown as a solid line (scaled), due to the amalgamated unmodeled systematic effects (theoretical) displayed in Fig. 1.

7. Coefficient of Determination and Model Performance

The R^2 values using both models were also calculated for all of the tide gauge stations data in the PSMSL repository to assess how well the models explain the MSL variability in the tide gauge series -- the closer is the R^2 to unity, the better the fit (Neter et al. 1996). In this application however, larger R^2 values also ensure that the estimates are biased less and their uncertainties are more reliable because unmodeled systematic effects have considerable impact on the solution (Iz 2006).

Fig. 16 shows the R^2 values of 1862 station solutions using both models. Among all 1,862 stations, 1,264 stations solutions are improved as a result of modeling transient variations that are significant at a 95% confidence level. The improvements over half of the 1,264 stations are below 20% change in the R^2 values, while the remaining improvements reach over 90%.

8. Conclusion

We have used CUSUM charts to detect transient changes in the MSL as indicated by CPs. We have developed statistical models that incorporate these effects and validated the trend estimates by comparing them to the estimates from a nearby tide gauge series free from transient effects. We subsequently showed that MSL for 1,264 out of 1,862 tide gauge stations in the PSML repository are contaminated with transient effects.

We have modeled, estimated, and demonstrated, as evidenced in Fig. 1 that unmodeled transient changes affect local MSL trends depending on the magnitude, location of the effects, and length of the tide gauge series as also explained theoretically in Iz (2006). Modeling of transient changes in the MSL also improved the R^2

Table 2. Sampled Sea Level Trends from PSMSL Metric Data with and without CP. PGR values are calculated using ICE-4G (VM2) (Peltier, 2001). This table presents 21 tide gauge sites in nine oceanic regions that show the sea level trends estimated with and without CP, and also the previous estimation from [Douglas, 1991].

Region	Series Length (Year)	Latitude	Longitude	PGR Correction mm/year	Trend* without CP mm/year	Trend with CP at 95% mm/year (No. of CPs)	PE**
North Sea							
Aberdeen II	104	57°09'N	2°05'W	-0.7	0.9±0.0	1.1±0.3 (5)	1.7
North Shields	99	55°00'N	1°26'W	-0.5	1.3±0.1	1.9±0.2 (6)	2.4
English Channel							
Newlyn	86	50°06'N	5°33'W	0.5	0.2±0.1	0.2±0.3 (5)	1.2
Brest	171	48°23'N	4°30'W	0.5	0.4±0.0	- 0.1±0.1 (6)	0.8
Atlantic							
Cascais	104	38°41'N	9°25'W	0.0	1.3±0.0	1.4±0.2 (9)	1.2
Tenerife	57	28°29'N	16°14'W	0.0	1.1±0.1	0.4±0.4 (3)	1.8
Mediterranean							
Marseille	112	43°18'N	5°21'E	0.0	1.4±0.1	1.1±0.1 (5)	1.7
Genova	89	44°24'N	8°54'E	-0.0	0.3±0.0	0.2±0.2 (3)	1.3
Trieste	90	45°39'N	13°45'W	0.0	0.3±0.1	0.2±0.3 (4)	1.3
Pacific							
Honolulu	96	21°19'N	157°52'W	-0.1	1.5±0.0	2.1±0.8 (59)	1.7
North American West Coast							
San Francisco	147	37°48'N	122°28'W	0.3	1.3±0.0	3.7±0.8 (76)***	1.6
Central America							
Balboa	88	08°58'N	79°34'W	0.0	1.4±0.1	1.3±0.9 (36)	1.6
Cristobal	72	09°21'N	79°55'W	-0.0	1.4±0.1	1.7±0.5 (20)	1.4
SE North America							
Key West	87	24°33'N	81°48'W	0.5	1.8±0.1	2.1±0.3 (10)	2.8
Charleston I	80	32°47'N	79°56'W	0.5	2.8±0.1	4.3±0.7 (13)	2.9
Hamption Rds	74	36°57'N	76°20'W	0.8	3.5±0.1	4.0±0.6 (11)	3.5
Baltimore	99	39°16'N	76°35'W	1.0	2.1±0.1	4.5±0.5 (20)	2.4
Atlantic City	87	39°21'N	74°25'W	1.4	2.6±0.1	2.9±0.5 (15)	2.6
New York	145	40°42'N	74°01'W	1.2	1.6±0.0	1.3±0.3 (7)	2.1
NE North America							
Portland	89	43°40'N	70°15'W	0.2	1.6±0.1	2.2±0.4 (19)	2.0
East port	66	44°54'N	66°59'W	0.5	1.7±0.1	2.0±0.4 (15)	2.7

* Trend estimates are corrected fro PGR

** PE: Previous estimates reported in [Douglas, 1991] and were corrected by ICE-4G (VM2).

*** Reported estimates vary for this station: 1.3 mm/year (solution with no CP), 1.6 mm/year (Douglas 1991), 2.1 mm/year (NOA, 2005), 3.7 mm/year with CP (solution with CP) which is likely to be an aberration since it includes a large number of CP compared to the other solutions. Note also that this station is subject to potential vertical movements by frequent earthquakes in this area.

values, which ensure that the estimates are less biased and their uncertainties are more reliable. Modeling transient changes in the MSL had also an overall impact on the magnitudes of MSL trend estimates that fall within the interval [-1, +3] mm/year. The number of trend estimates within this interval, which also overlaps with the interval reported by Bindoff et al (2007) for the global MSL, decreased with the inclusion of the transient changes in the models.

The presence of large number of statistically significant transient MSL changes in the tide gauge data (8,072 CPs) and their impact

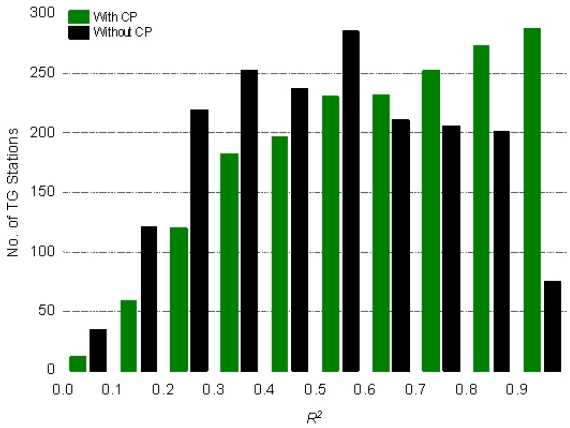

Figure 16. R^2 (coefficient of determination) of model solutions with and without transient effects.

on the local solutions give a more precise meaning to the local trend estimates calculated using models that account for them. In other words, the fact that significant numbers of transient MSL changes are attributable to variations in the local and regional air pressure, wind, currents, river discharges etc., modeling their influences in analyzing the tide gauge data, enables the resulting trend estimates to be more representative of the influence of the time-varying MSL gradients due to the total steric and long-term eustatic effects. Better yet, if the transient changes include also the impact of the total steric effects (thermo and halosteric) on the MSL because the ocean surfaces respond faster to temperature changes and density variations, then the estimated long term trends are more representative of the eustatic changes in the MSL rather than the lump-sum effects of MSL trends of total steric and eustatic origins. This separation explains also for the reduced number of stations having MSL trend estimates within the [0 to 3] mm/year interval as compared to the earlier solutions which do model transient changes in the MSL.

Under this scenario, the magnitudes and spatial and temporal distribution of the transient MSL level changes can also reveal clues about the peculiarities of the recent changes in the global MSL, which deserve further investigation.

9. Appendix: Change Point Detection

Control charts and change point (CP) analysis are used to determine whether the mean of a stochastic process has shifted. The Cumulative Sum Chart (CUSUM) chart, a graphical method of CP detection, was first introduced by Page (1954). The underlying mathematical principles involved in its construction were elaborated in Ewan (1963), Johnson (1961), and Johnson and Leone (1962). Mertikas and Rizos discussed their geodetic applications in 1997.

CP analysis complements the CUSUM charts by trying to answer the following questions: Did a change occur? Did more than one change occur? When did any changes occur? With what confidence can we conclude that a change occurred?

Consider the cumulative sums, $s_0, s_1 \ldots, s_n$, that are calculated from data $x_1, \ldots, x_n$, which are assumed to be random in nature, as follows:

$$s_i = s_{i-1} + (x_i - \bar{x}), \quad i = 1, 2, \ldots n \tag{4}$$

where,

$$s_0 = 0 \quad \text{and} \quad \bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} \quad . \tag{5}$$

Plots of CUSUMs against i generate CUSUM charts. If there are no changes in the mean, CUSUM charts display a steady straight path, since the data is random in nature. Otherwise, a segment of the CUSUM chart with an upward slope indicates a period where the values tend to be above average or a segment with a downward slope indicates a period of time where the values tend to be below the overall average.

A sudden change s_{diff} in the slope of the CUSUM occurs with a sudden shift in the average. The magnitude of the sudden change, $|s_{diff}|$,can be estimated from

$$s_{diff} = \max_{i=0,\ldots,n} s_i - \min_{i=0,\ldots,n} s_i \tag{6}$$

A bootstrap analysis by random reordering of the elements of the series with replacement (Efron & Tibshirani 1993) gives the confidence level whether a mean shift has occurred. s_{diff} calculated from a large number of bootstrapped series tends to stay at about zero or the mean value of the series when compared with the $s_{diff}^{original}$ of the original series. Hence, a total of N bootstraps, for which p is the number of bootstraps with $s_{diff} < s_{diff}^{original}$, gives the confidence level as

$$Confidence\ level = 100 \times \frac{p}{N} \quad . \tag{7}$$

A high confidence level is strong evidence that a change has indeed occurred. Once a change has been detected, the location of the change needs to be detected. Among others, a natural and simple approach is to use the following expression:

$$s_m = \max_{i=1,\ldots,n} |S_i| \tag{8}$$

where s_m is the point furthest from zero in the CUSUM chart. The m^{th} point gives the last data before the change occurred and the $m+1$ point refers to the first data after the change.

The whole procedure can be applied to the series that are generated by replacing the original data with their ranks. This approach is more resistant to the outliers and large variations in the data, which was the preferred approach in this study.

Acknowledgment

The ICE-4G (VM2) model corrections were provided by C.K. Shum. The critical comments by C.K. Shum and an anonymous reviewer are gratefully acknowledged.

References

Barnett, T.P., 1983, Recent changes in sea level and their possible causes. Clim. Change (5):15-38.

Bindoff, N.L., J. Willebrand, V. Artale, A. Cazenave, J. Gregory, S. Gulev, K. Hanawa, C. Le Quéré, S. Levitus, Y. Nojiri, C.K. Shum, L.D. Talley, and A. Unnikrishnan, 2007, Observations: Oceanic climate change and sea level, in *Climate Change 2007: The Physical Science Basis*, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller, Cambridge University Press, Cambridge, pp. 385-432.

Douglas B.C., 1991, Global sea level rise. Journal of Geophysical Research 96(C4):6981-6992.

Efron B., R.J.Tibshirani, 1993, An introduction to the bootstrap. Chapman and Hall/CRC Publs.

Ewan W.D., 1963, When and how to use CUSUM charts. Technometrics (5):1-32.

Fairbridge R.W. and Krebs OA Jr, 1962, Sea level and the southern oscillation. Geophys. J. (6):532-545.

Gornitz V., S. Lebedeff, 1987, Global sea-level changes during the past century, In D Nummedal et al. (Editors) Sea-level fluctuation and coastal evolution. Soc. Econ. Paleontol. Miner., Spec. Publi. (41):3-16.

Gornitz V., S. Lebedeff, J. Hansen, 1982, Global sea level trend in the past century. Science (215):1611-1614.

Gutenberg B., 1941, Changes in sea level, postglacial uplift, and mobility of the earthŠs interior. Bull. Geol. Soc. Am. (52):721-772.

IPCC Climate Change, 2007, Synthesis Report. On line access http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf.

Iz H.B. and, C.K. Shum, 2000, Mean Sea Level Variations in the South China Sea from Four Decades of Tidal Records in Hong Kong, Marine Geodesy Vol.23, (4), pp-221-233.

Iz H.B. and H.M. Ng, 2005, Are the Global Tide Gauge Data Stationary in Variance? Marine Geodesy, (3):209-217.

Iz H.B., 2006, How do Unmodeled Systematic MSL Variations Affect Long Term Sea Level Trend Estimates from Tide Gauge Data? Journal of Geodesy, DOI 10.1007/s00190-006-0028-x.

Johnson N.L., 1961, A simple theoretical approach to cumulative sum control charts. Journal of the American Statistical Association (56):83-92.

Johnson N.L., F.C. Leone, 1962, Cumulative sum control charts - mathematical principles applied to their construction and use. Industrial Quality Control (18):15-21.

Mertikas S.P., C. Rizos, 1997, Online detection of abrupt changes in the carrier phase measurements of GPS Journal of Geodesy (71):469-482.

Neter J.M., H. Kutner C.J. Nachtsheim, W. Wasserman, 1996, Applied linear statistical models. Richard D. Irwin, Illinois, USA, 1408.

Page E.S., 1954, Cumulative sum charts. Technometrics (3):1-9.

Peltier W.R., 1994, Ice-age paleotopography. Science (265):195201.

Peltier W.R., 1996, Mantle viscosity and ice-age ice-sheet topography. Science (273):13591364.

Peltier W.R., 1998b, Implicit ice in the global theory of glacial isostatic adjustment. Geophys. Res. Lett. (25):39573960.

Peltier W.R., A.M .Tushingham, 1989, Global sea level rise and greenhouse effect: Might they be connected? Science (244):806-810 International Geophysics Series 75.2001.

Peltier W.R. ICE4G (VM2), 2001, Glacial isostatic adjustment corrections. In Sea level rise: History and consequences. Douglas BC, Kearney MS, and SP Leatherman (Eds.). Academic Press. Permanent Service for MSL, 2001, Home page, http://www.pol.ac.uk/psmsl/. (accessed on 24th September 2002).

Ponte, R.M., 2006, Low-Frequency Sea Level Variability and the Inverted Barometer Effect. J. Atmos. Oceanic Technol., 23, 619-629.

Pugh D., 2004, Changing Sea levels Effects of Tides, Weather and Climate.Cambridge University Press, Cambridge. pp.165 -167.

Potluri R., 1971, Some Notes on Misspecification in Multiple Regressions. The American Statistician, Vol. 25, No. 5. pp. 37-39.

Vanicek P., 1978, To the problem of noise reduction in sea level Used in vertical crustal motion detection, Physics of the Earth and Planetary Interiors, 17, 265-280.

Woodworth P.L, 1990, A search for acceleration in records of European MSL. Int. J. Climatol. 10:129-143.

Woodworth P.L., R. Player, 2003, The permanent service for MSL: An update to the 21st century. Journal of Coastal Research 19(2):287-295.

Woodworth P.L., 2005, Personal communication.