TIDE GAUGE LOCATION AND THE MEASUREMENT OF GLOBAL SEA LEVEL RISE

Michael Beenstock[1], Daniel Felsenstein*[2], Eyal Frank[1], Yaniv Reingewertz[1]

1.Department of Economics, Hebrew University of Jerusalem

2.Department of Geography, Hebrew University of Jerusalem

*Corresponding author

msdfels@mscc.huji.ac.il

Abstract

The location of tide gauges is not random. If their locations are positively (negatively) correlated with sea level rise (SLR), estimates of global SLR will be biased upwards (downwards). Using individual tide gauges obtained from the Permanent Service for Mean Sea Level (PSMSL) during 1807 – 2010, we show that tide gauge locations in 2000 were independent of SLR as measured by satellite altimetry. Therefore these tide gauges constitute a quasi-random sample, and inferences about global SLR obtained from them are unbiased. Using recently developed methods for nonstationary time series, we find that sea levels rose in 7 percent of tide gauge locations and fell in 4 percent. The global mean increase is 0.39 – 1.03 mm/year. However, the mean increase for locations where sea levels are rising is 3.55 - 4.42 mm/year.

These findings are much lower than estimates of global sea level (2.2 mm/year) reported in the literature and adopted by IPCC (2014). These estimates make widespread use of imputed data for locations which do not have tide gauges. We show that although tide gauge locations in 2000 are uncorrelated with SLR, the global diffusion of tide gauges during the 20th century was negatively correlated with SLR. This phenomenon induces positive imputation bias in estimates of global mean sea levels because tide gauges installed in the 19th century happened to be in locations where sea levels happened to be rising.

Acknowlegement: we acknowledge the useful and insightful comments of two anonymous referees.

Introduction

The debate over sea level rise (SLR) has been heavily informed by the availability of tide gauge data feeding the Permanent Service on Mean Sea Levels (PSMSL). However, the global coverage of tide gauges is patchy and incremental and is often influenced by local contingencies. Historically, tide gauges were installed in sea ports to service commercial shipping and to mitigate the vagaries of flooding and ebbing currents. They were not intended for measuring sea level and only came to be used for this purpose in the early 20th century (Emery and Aubrey 1991). The distribution of tide gauges is dominated by coastal sites (ports, rivers, estuaries etc) with inadequate coverage of islands and the open sea, and by a location bias that favors the northern hemisphere (Pirazolli 1993, Douglas 2001, Mitchum 2010).

Although coverage has greatly improved, especially in the second half of the 20th century, there are still many locations where there are no tide gauges at all, such as the southern Mediterranean, while other locations are under-represented, especially Africa and South America. The geographic deficiency of tide gauge location is compounded by a temporal deficiency due to the limited number of tide gauge with long histories. Indeed, the number of tide gauges recording simultaneously at any given period is limited (Groger and Plag 1993), although this problem is becoming less acute with the passage of time.

To mitigate these spatial and temporal deficiencies, data imputation has been widely used. For example, Chambers et al (2002) and Church at al (2004) supplement tide gauge data with data on satellite altimetry, and Jevrejeva et al (2006) impute virtual tide gauge data. Imputation naturally increases the more we delve into the history of sea levels because tide gauge coverage becomes thinner the more we step back in time. Indeed, estimates of global sea levels in the 19th century are almost entirely imputed. See Christiansen et al (2010) for a critical discussion of these data imputations.

Had the location of tide gauges been determined as a random sample of the population of sea levels, it should not have mattered that some locations do not have tide gauges. However, the global diffusion of tide gauges has not followed a planned statistical design. The richer countries have installed more tide gauges and the Southern

Hemisphere is under-represented in PSMSL. Indeed, this kind of under-representation has varied over time.

Our central purpose is to note that, just because a sample does not happen to be random, this does not necessarily mean that parameters estimated from it must be biased. Therefore, estimates of global sea level obtained from PSMSL data are not automatically biased simply because PSMSL happens to be a non-random sample. Non-random samples will be quasi-random and informative if the observations are selected independently of the outcome of interest (Heckman 1976). Therefore, if the location of tide gauges happens to be independent of SLR, a non-random sample such as PSMSL may still be used to make unbiased inferences about global sea level. If, however, tide gauges happen to be located where SLR is smaller, the sample mean will be biased downward, in which case inferences based on PSMSL will underestimate global SLR. If, on the other hand, tide gauges happen to be located where SLR is larger, inferences based on PSMSL will overestimate global SLR.

Reasons may be posited why tide gauge location should be both positively and negatively correlated with SLR. Given everything else, if SLR is larger, governments might invest more in the installation of tide gauges for precautionary reasons and monitoring (Tai 2001). Where SLR is zero, there is no need for monitoring, and it is less likely therefore that governments will incur the investment and current costs of tide gauges. Tide gauge locations can be negatively correlated with SLR in cases where their installation was related to harbor operation (falling sea levels and difficulties in maneuvering ships under these conditions). If the prevalence and diffusion of tide gauges happen to be independent of SLR, we maintain that it is reasonable to assume that estimates of global SLR based on PSMSL are unbiased, in which case recourse to data imputation would be unnecessary.

Surprisingly, tide gauge location has not been an issue in the study of SLR. In this paper we investigate for the first time whether PSMSL is a quasi random sample for purposes of estimating global sea level. Specifically, we investigate whether the location of tide gauges is independent of SLR. The obvious methodological difficulty is that measures of SLR are not available where tide gauges have not been installed. We overcome this methodological difficulty by using satellite altimetry data, which since the early 1990s includes measurements of sea levels for the entire globe within 75 degrees latitude north and south. We have matched these data to the locations of PSMSL tide gauges in 2000.

Our main finding is that the tide gauges in PSMSL constituted a quasi random sample in 2000 because their locations were independent of SLR. We also show, however, that the global diffusion of tide gauges in the 19th and early 20th centuries was correlated with SLR. Specifically, the tide gauges installed before the second half of the 20th century were more likely to have been in locations where sea levels were rising. Three important conclusions follow from these results. First, the tide gauge data in PSMSL may be used to obtain unbiased estimates of global sea level rise without the need for data reconstruction or imputation. Second, there is no evidence from PSMSL of global sea level rise. In most locations sea levels are stable. In a minority of locations sea levels are rising, and in a smaller minority sea levels are falling. Third, the claim that sea levels are rising globally (IPCC 2013) is an artifact induced by the use of imputed data.

2. Methodology

2.1 Testing Nonstationarity

To estimate SLR in specific locations over time we use statistical tests for non-stationarity. A time series is non-stationary when its sample moments depend on when the data are measured. Trending variables must be non-stationary because their means vary over time. Therefore, if a time series, such as sea level, happens to be stationary it cannot have a trend by definition. This insight serves as our methodological basis for testing hypotheses about sea level rise.

Let Y_{jt} denote the sea level record of tide gauge j in time period t when the number of observations is denoted by T_j. Dickey and Fuller (1979) suggested the "augmented Dickey-Fuller" test (ADF) for stationarity[1] (dropping j for convenience):

[1] Standard statistical tests based on the normal distribution (t, chi-square and F tests) assume that the data are stationary. Obviously, these tests cannot be used to test for stationarity.

$$\Delta Y_t = \alpha + \gamma t + \beta Y_{t-1} + \sum_{i=1}^{p} \lambda_i \Delta Y_{t-i} + e_t \qquad (1)$$

where e is assumed to be identically and independently distributed (iid) with variance σ^2, and p is chosen so that e is iid. There are three main cases.

Case 1: If $\beta < 0$ Y is trend stationary; it has a deterministic time trend equal to γ/β.

Case 2: If $\beta = \gamma = 0$ Y is difference stationary; it has a stochastic time trend equal to α.

Case 3: If $\alpha = \gamma = \beta = 0$ Y is nonstationary but has no trend because it is a random walk.

Case 4: If $\beta < 0$ and $\gamma = 0$ Y is stationary; it is stable and mean-reverts to $-\alpha/\beta$.

Dickey and Fuller have calculated critical values for testing these cases. The ADF statistic is the OLS estimate of β divided by its standard deviation. This looks like a t-statistic but it does not have a t-distribution. In fact the critical value (p = 0.05) of ADF when there are 100 observations is -2.89 rather than -1.96 as suggested by the conventional t-statistic for case 4 and it is -3.43 in case 1. If the ADF statistic is larger (less negative) than these critical values the null hypothesis is rejected. For example, in case 4 the null hypothesis of nonstationarity cannot be rejected if ADF > -2.89. In summary, the null hypothesis in ADF tests is that the data are nonstationary.

If time series for sea levels at tide gauge j can be classified as cases 3 and 4 they have neither deterministic nor stochastic time trends. Sea level is neither rising nor falling; it is stable. We use the ADF statistic to classify tide gauges in terms of these 4 basic cases. Further variants allow for quadratic or higher order deterministic or stochastic trends. Allowing a quadratic deterministic trend would involve specifying t^2 in equation (1). In the stochastic case a quadratic trend would involve specifying $\Delta^2 Y_t$ as the dependent variables in equation (1). Alternatively, both types of trend may be present.

Another issue involves discrete breaks in sea levels and/or their trends. For example, in case 4, a change in α_j may create the illusion that the data are non-stationary when the opposite is true (Perron 1989). In this case the critical value of ADF naturally becomes more stringent; it is -4.5 instead of -2.89 for case 4 because it penalizes data-mining that is involved if the date of the structural break is unknown.

Phillips and Perron (1988) have suggested a nonparametric alternative to the ADF statistic. Instead of augmenting equation (1) by specifying lagged dependent variables to capture serial correlation, they set p = 0, and calculate robust (HAC) estimates for the standard deviation of β. However, critical values for the Phillips-Perron statistic (PP) are the same as for ADF. Maddala and Kim (1998, p 81) note that ADF might be more reliable than PP.

Whereas ADF and PP test the null hypothesis that the data are nonstationary the KPSS statistic (Kwiatkowski, Phillips, Schmidt and Shin 1992) tests the null hypothesis that they are stationary. Let d denote the order of differencing required to make Y_{jt} stationary. ADF treats d = 1 as the null whereas KPSS treats d = 0 as the null, i.e. it takes the form of a Lagrange multiplier test in which d is assumed to be zero. KPSS is based on the following model:

$$Y_t = \eta + \delta Z_t + \theta t + \varepsilon_t \quad (2a)$$

$$\Delta Z_t = \gamma + v_t \quad (2b)$$

where Z is an unobserved nonstationary component of Y, and ε and v are independent iid random variables, where ε is the stationary component of Y and v is the stationary component of Z. Since Z is generated by a random walk, it must be non-stationary. Here too there are four cases.

Case 1: If δ = γ = 0 the data are trend stationary and have a deterministic trend θ.

Case 2: If θ = 0 the data are difference stationary and have a stochastic trend γ.

Case 3: If θ = γ = 0 the data are nonstationary because their DGP is driftless random walk.

Case 4: If θ = δ = γ = 0 the data are stationary; they have no trend and they mean-revert to η.

If cases 3 and 4 apply sea levels are stable. Just as in ADF tests structural breaks in levels and trends might induce error in these tests, so structural breaks in η, θ and γ might induce error in KPSS tests. Lee and Strazicich (2001) and Korozumi (2002) have extended the KPSS test to the case in which there is a structural break in either levels or trends.

Rejecting the null hypothesis of stationarity is not conceptually equivalent to accepting the null hypothesis of non-stationarity, and vice-versa[2]. Therefore ADF and KPSS tests are not mirror images. In any case, if the KPSS statistic rejects the null

[2] Failing to prove guilt is not equivalent to proving innocence, and vice-versa.

hypothesis d = 0, this does not necessarily mean that Y is nonstationary because the true value of d might be 0.4, or some other value less than 0.5, at which Y is nonstationary (see below). If ADF rejects its null (d = 1) and KPSS fails to reject its null (d = 0), we may be reasonably confident that the data are stationary. Ambiguity arises, however, if ADF and KPSS conflict because e.g. DF and KPSS do not reject their respective null hypotheses. If the null hypothesis is that sea levels are stable, the KPSS test is appropriate. If the null hypothesis is that sea levels are rising, the ADF test is appropriate.

In principle, we think that the burden of proof falls upon those claiming that sea levels are rising, in which case the appropriate test is KPSS. On the other hand, the belief that sea levels are rising has become so widespread (IPCC 2013), that the burden of proof may have shifted onto those claiming that sea levels are not rising, in which case the appropriate test would be ADF. We use both tests. In any case, if ADF tests reject the hypothesis that sea levels are rising, and KPSS tests cannot reject the hypothesis that sea levels are not rising, this naturally increases confidence in the claim that sea levels are not rising.

2.2 Trends v Long Memory

We have already mentioned that d may not be an integer such as 0 or 1. In fact the threshold for nonstationarity is ½ rather than 1 (Granger and Joyeux 1980). If d lies between 0 and ½ the time series has "long memory" because its mean-reverting properties are protracted. Such models specify the dependent variable in equation (1) as $\Delta^d Y_t$ where d is a parameter to be estimated that falls in the range 0 - ½.

Long memory may be mistaken for trend or difference stationarity especially if d is large (e.g. 0.4). This happens because mean-reversion, which varies inversely with d, is slow. Therefore, tide gauges which are classified using the methodology in section 2.1 as nonstationary should be checked to see whether they have in fact long memory. If they do, sea levels do not have a trend, but are particularly slow to revert to their long-run level. To estimate d we use the two-stage procedure suggested by Geweke and Porter–Hudak (1983) and the semi-parametric procedure suggested by Robinson (1995). The former is not consistent for d > 1 whereas the latter is[3]. This check need not be applied here to those tide gauges which are classified as stationary with d = 0, because long memory implies stationarity too.

These estimators involve calculating the periodogram of the data at $n = T^a$ frequencies $\omega_j = 2\pi j/T$ (j = 1,2,…,n). The standard default for a is ½. For tide gauges with the shortest histories (120 monthly data points) this means that n = 11, $\omega_1 = 0.052$ and $\omega_n = 0.576$. For the oldest tide gauge (1807) it means that n = 50, $\omega_1 = 0.026$ and $\omega_n = 0.1257$. Since d is estimated from only 11 ordinates for tide gauges with short histories, we experiment by increasing a. For example when a = 0.7 n increases to 29 for these tide gauges, ω_1 remains unchanged and ω_n increases to 1.519. Increasing a not only increases the number of ordinates, it also increases the range of frequencies that are used to estimate d. We set a so that the number of tide gauges with long memory ceases to sensitive to a.

2.3 Selection Model for Tide Gauges

In this sub-section we propose a test for potential sample selection bias induced by the non-random location of tide gauges in PSMSL. International organizations do not commission tide gauges. The decision to commission tide gauges is made by the relevant governmental authorities in each country. Let T_k* be the benefit to the relevant authority of having a tide gauge in location k. Suppose that T_k* is hypothesized to depend on covariates X_k in which tide gauge j is located[4], These covariates include GDP per head since richer countries can afford more tide gauges, population as a measure of risk exposure, and length of coastline. In addition, T_k* is hypothesized to depend on SLR since if sea level is rising there may be a greater need for monitoring. Finally, e is assumed to be a normally distributed random variable with zero mean capturing unobservable phenomena which determine benefit from tide gauges in the location:

$$T_k^* = \delta + \theta X_k + \phi SLR_k + e_k \qquad (3)$$

[3] Sowell (2002) has suggested a one-step estimator by maximum likelihood, which, however, is difficult to apply. Two-step estimators estimate the power spectrum or periodogram in the first stage, and then estimate d by on OLS regression involving the estimated frequencies obtained from the first stage.

[4] We avoid here the decision by the government regarding the precise location of the tide gauges it constructs.

Let T_k denote a dummy variable that equals 1 if there is at least one tide gauge in location k and zero otherwise. Since T_k* is not observed, we assume that there will be a tide gauge in location k if the benefit outweighs the costs of construction and maintenance denoted by C, i.e. $T_k = 1$ if $T_k^* > 0$. Since T is a dummy variable the estimates of δ, θ and ϕ may be obtained from a probit regression of T on X and SLR.

If ϕ is zero the location of tide gauges in PSMSL is independent of SLR, in which case the data are quasi-random (Heckman 1976). In this case despite the fact that PSMSL is not a random sample, estimates of SLR are unbiased. If ϕ is positive (negative) estimates of SLR will be biased upwards (downwards).

An alternative statistical procedure is to estimate equation (3) by censored regression (tobit model) in which the dependent variable is the number of tide gauges in a specific country. The dependent variable is censored because many locations have no tide gauges. If the number of tide gauges in the location is conditionally independent of SLR, PSMSL may be regarded as a quasi random sample for purposes of estimating global SLR.

2.4 Global Mean Sea Level

Investigators of SLR have sought to construct measures of global mean sea level (GMSL), which are intended to convey information about global trends over time. Our objective is not to estimate GMSL. Instead, we estimate SLR for the individual tide gauges in PSMSL. Since these tide gauges were installed at different dates estimates of SLR refer to different time periods. This raises a question about the comparability of these estimates over time. In this subsection we clarify the relationship between our work and GMSL.

Suppose that the world consists of only two locations, A and B, and that tide gauge A was installed in 1900 and tide gauge B was installed in 1950. Suppose that since 1900 SLR in A is estimated at 3mm/year and since 1950 SLR in B is estimated at 1mm/year. To compare the former to the latter we need to estimate SLR in A since 1950. Suppose that the trend in A is stable so that since 1950 SLR in A was 3mm/year. Therefore, since 1950 GMSL grew at 2mm/year. We cannot estimate GMSL during 1900 – 1950 because there are no data for B. Such data can be imputed, but this might induce imputation bias.

In what follows we use all the data that are available for each tide gauge. Therefore, in the case of tide gauge A we would use the data since 1900. However, we carry out statistical tests described in section 2.1 to determine whether SLR in A is stable or not. If SLR is stable we could, in principle, have used our results to construct GMSL since 1950. The reason why we don't is because there are more than a thousand tide gauges, which were installed at various dates (see below). Therefore it is impossible to construct GMSL prior to 1950 unless one is prepared to impute data for tide gauges installed between 1900 and 1950, or some other base year[5]. As GMSL approaches the base year the rate of data imputation naturally becomes increasingly frequent.

By contrast our approach allows us to estimate recent GMSL. For example, if the trends in A and B are stable, we may conclude that since 1950 GMSL grew by 2mm/year. Prior to 1950 GMSL cannot be calculated without imputing data for location B. In practice, therefore, we may use our results to inform about GMSL during the last few decades only. These estimates will be unbiased if the PSMSL tide gauges constitute a quasi-random sample. We cannot say anything about GMSL during the 19th and 20th centuries, and refrain from doing so.

2.5 The Wood and the Trees

Whereas Douglas (2001), Church et al (2004), Jevrejeva et al (2006) and others test hypotheses about global sea levels “top down” using globally averaged data, and by imposing a common statistical model on all tide gauges, our methodology is “bottom up”. It allows for heterogeneity by allowing each tide gauge to have its own statistical idiosyncrasies, and inferences about global sea level are made from these local results. The “top down” approach imposes homogeneity and therefore runs the risk of aggregation bias induced by imposing invalid statistical models locally.

The testing procedure proposed in sections 2.1 uses individual tide gauge data independently. Im et al (2003) suggest a test for nonstationarity in heterogeneous panel data such as PSMSL in which the null hypothesis is d = 1 as in ADF. Hadri (2000) has suggested a test for stationarity in heterogeneous panels in which the null hypothesis is d = 0 as in KPSS. Robinson (1995) has suggested a test for nonstationary panel data when d is fractional. We do not use these tests for two reasons. First, they assume that the panel

[5] For example, in Jevrejeva et al (2006) the base year is 1810 and in Church and White (2006) it is 1870.

units are independent, which is unlikely to apply to tide gauges since sea levels are spatially related for physical reasons. Second, they are designed for balanced panels (the units are observed over the same period) which does not apply to PSMSL. However, our bottom-up approach risks missing the wood for the trees; it might falsely reject the hypothesis sea levels are rising globally.

Suppose again that the world has two locations A and B. If tide gauge A does not happen to have a trend that is statistically significant, it is classified as trend-free, and its SLR is assumed to be zero. If, on the other hand, tide gauge B has a trend which is statistically significant and positive, its SLR is classified as positive. It is tempting to conclude that sea level rise is not global because it only occurs in B. This conclusion would be correct if A and B are independent. However, because the oceans constitute an integrated system, A and B are unlikely to be independent. This dependence gives rise to two potential problems. First, the trend in A might be jointly significant with the trend in B. Second, the trend in B might be jointly insignificant with the trend in A. As noted above, this bottom-up approach, which involves individual modeling of each tide gauge data, is exposed to the danger of missing the wood for the trees.

This risk may be elucidated as follows. For illustrative reasons we assume that sea levels have linear deterministic time trends, but the point generalizes to stochastic and nonlinear trends. The global average sea level is $S = ½(S_A + S_B)$ where S_A and S_B denote sea levels in A and B respectively. Suppose S_t is regressed on t as in the top-down approach:

$$S_t = \alpha + \beta t + u_t \qquad (4)$$

where β is the global trend. The estimate of β is:

$$\hat{\beta} = \frac{\sum S_t (t - \bar{t})}{\sum (t - \bar{t})^2} = \frac{\frac{1}{2}[\sum S_{At}(t - \bar{t}) + \sum S_{Bt}(t - \bar{t})}{\sum (t - \bar{t})^2} = \tfrac{1}{2}(\hat{\beta}_A + \hat{\beta}_B) \qquad (5)$$

The estimated global trend is an average of the local trends, $\hat{\beta}_A \; and \; \hat{\beta}_B$. The variance of the estimated global trend is:

$$\mathrm{var}(\hat{\beta}) = \tfrac{1}{4}[\mathrm{var}(\hat{\beta}_A) + \mathrm{var}(\hat{\beta}_B) + 2\,\mathrm{cov}(\hat{\beta}_A \hat{\beta}_B)] \qquad (6)$$

and the t - statistic for the global trend is:

$$t_{\hat{\beta}} = \frac{\hat{\beta}}{sd(\hat{\beta})} = \frac{\hat{\beta}_A + \hat{\beta}_B}{\sqrt{\mathrm{var}(\hat{\beta}_A) + \mathrm{var}(\hat{\beta}_B) + 2\,\mathrm{cov}(\hat{\beta}_A \hat{\beta}_B)}} \tag{7}$$

The t – statistics for the local trends are:

$$t_A = \frac{\hat{\beta}_A}{sd(\hat{\beta}_A)} \quad and \quad t_B = \frac{\hat{\beta}_B}{sd(\hat{\beta}_B)}$$

Substituting these local t –statistics into equation (7) we obtain the relationship between the global (top-down) t-statistic and the local (bottom-up) t – statistics:

$$t_{\hat{\beta}} = \frac{t_A sd(\hat{\beta}_A) + t_B sd(\hat{\beta}_B)}{\sqrt{\mathrm{var}(\hat{\beta}_A) + \mathrm{var}(\hat{\beta}_B) + 2\,\mathrm{cov}(\hat{\beta}_A \hat{\beta}_B)}} \tag{8}$$

Equation (8) establishes that the global t – statistic is a weighted average of the local t – statistics. We may write $\mathrm{cov}(\hat{\beta}_A \hat{\beta}_B) = rsd(\hat{\beta}_A) sd(\hat{\beta}_B)$, where r denotes the correlation coefficient between $\hat{\beta}_A$ and $\hat{\beta}_B$. The weights in equation (8) only sum to one when r = 1 as may be seen in case 1 in the table. In this case if the local trends are individually not significant, nor is the global trend. Case 2 shows that if the tide gauges are independent the global trend remains insignificant, but the global t-statistic increases. If the tide gauges are sufficiently negatively correlated (case 3), the global trend is statistically significant. Case 4 illustrates the possibility that the global trend is not statistically significant despite the fact the SLR is significantly positive in A. Case 5 illustrates the converse; the global trend is statistically significant and positive despite the fact that SLR is local.

The Relation between Global and Local t-statistics for Sea Level Rise

Case	t_A	t_B	R	t
1	1.5	1	1	1.2
2	1.5	1	0	1.664
3	1.5	1	-0.5	2.268
4	2.5	0	0	1.387
5	0	2.5	-0.5	2.835

Notes: $sd_A = 2$, $sd_B = 3$

If sea levels are positively correlated, and according to the bottom-up approach sea levels are not rising, there is no danger of the wood being missed for the trees. This danger arises, however, when sea levels are negatively correlated, as in case 3. On the other hand, in the top-down approach there is always a danger of mistaking the trees for the wood, as in case 5, regardless of the correlation between tide gauges. If the correlation is sufficiently negative the global trend can be positive and statistically significant despite the fact the in A sea level is falling significantly, while in B it is increasing insignificantly.

In case 2 the global t–statistic exceeds its local counterparts because the local trends are uncorrelated. Case 3 shows that if sea levels are negatively correlated the global t–statistic exceeds its critical value. This is a case of missing the wood for the trees since the local trends are not individually significant. In case 4 sea level is flat in B, but rising in A. The global trend is not significant. In case 5 the global trend is significant despite the fact that there is no global trend.

The important conclusion is that the more positively correlated are sea levels in different locations the less likely it is that the wood will be missed for the trees in a bottom-up approach such as ours. On the other hand, the top-down approach is more likely to see woods when there are no trees since the global t-statistic varies inversely with r.

3. Data

When the data were retrieved in 2011 PSMSL (www.pol.ac.uk/psml) comprised data ending in December 2009, comprising 564,552 monthly observations of 1,197 tide gauge stations observed between 1807 and 2009. Figure 1 plots the number of reporting tide gauges over time. The reporting periods for these tide gauges vary and range from 6 months to 203 years, with a mean of 39 years. It is clear from Figure 1 that not only were new tide gauges commissioned, but some tide gauges were also decommissioned. The number of tide gauges decreased sharply after 1995 because of reporting delays which are substantial. We exclude tide gauges with continuous records of less than 10 years since a decade is insufficiently long to check for non-stationarity. Out of the 1197 tide gauges in PSMSL we excluded 197 with fewer than 10 years of consecutive data. This leaves 1000 tide gauges which we use in our analysis.

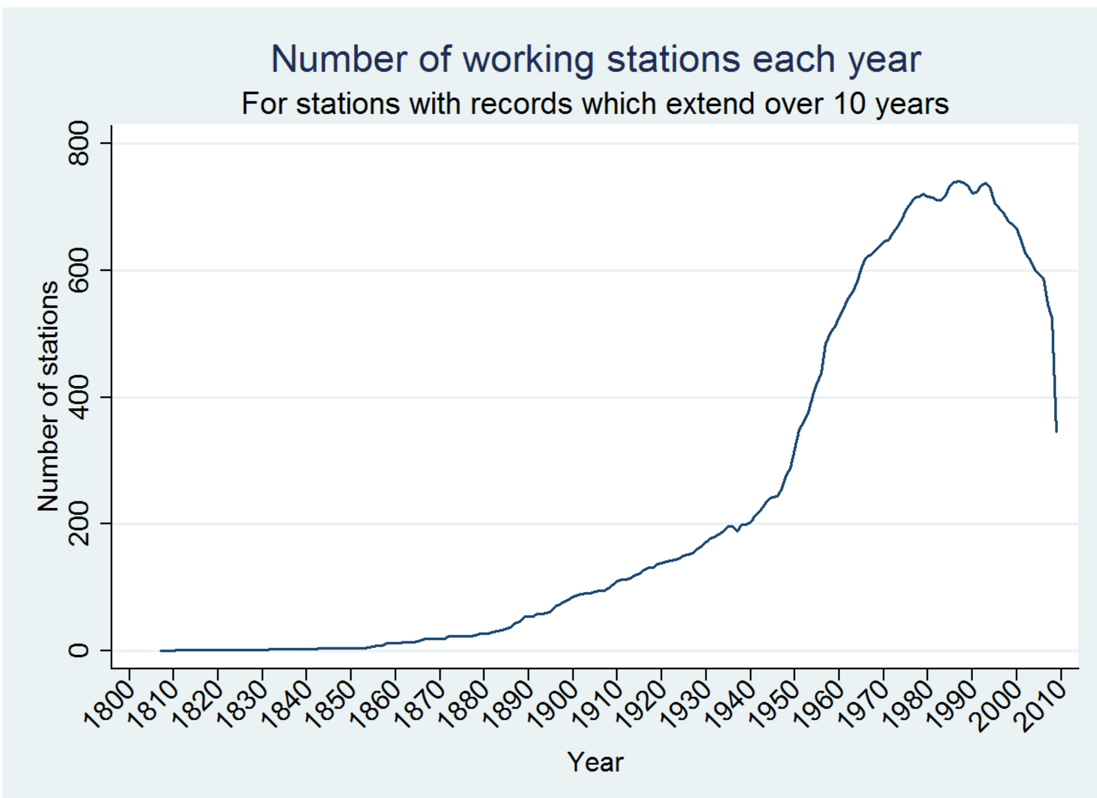


Figure 1: Number of Reporting Tide Gauges

Only 140 of these 1000 tide gauges have unbroken data records. Figure 2 presents the distribution of the number of monthly missing values in the dataset. Forty tide gauges are missing one data point. In some cases missing data are consecutive, while in other cases they are sporadic. We divided missing values into three categories: i) sporadic missing values, ii) 2 - 12 consecutive missing values (up to 12 months), iii) over 12 consecutive missing values. In the first case we imputed the missing value by the average of the data before it and after it. In the second case imputation is based on interpolation. In the third case the data are split and treated as separate segments. Out of the 1000 tide gauges in the analysis, 823 are un-split, 148 are split once, 27 are split twice and 2 are split three times. This makes a total of 1208 continuous data segments in all. In what follows we distinguish between tide gauges and segments[6].

Map 1 shows that the global diffusion of tide gauges has been far from random. In 1900 there were almost no tide gauges in the southern hemisphere and they were concentrated in the Baltic Sea. A century later the coverage is more comprehensive, but Africa and South America are under- represented. By contrast Japan which had only one tide gauge in 1900 had more than a hundred in 2000.

[6] Examples of split tide gauges may be seen in the Graphical Appendix: Neuville (Canada) in Figure A1, Vancouver in Figure A2, Brest and Swinousjcie (Poland) in Figure A4.

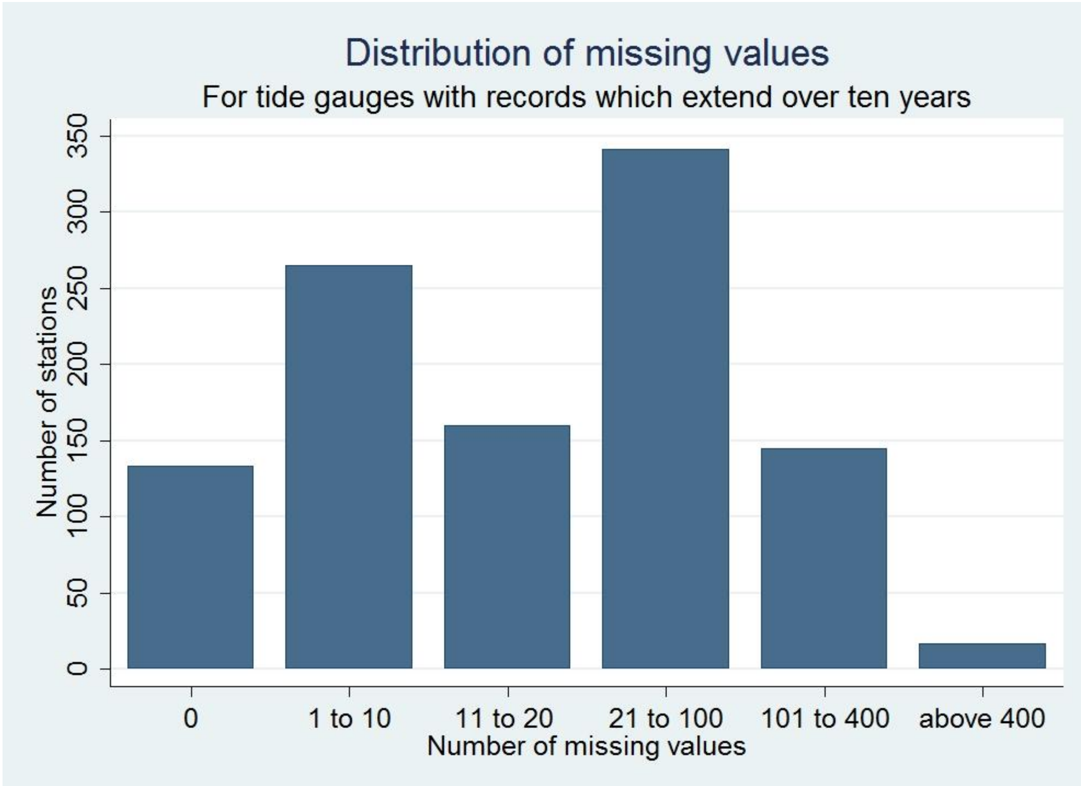


Figure 2: Distribution of Missing Values

4. Results

We begin by carrying out ADF and PP tests on 1208 segments of the 1000 tide gauges mentioned in section 3, which test the null hypothesis of sea level rise. Thereafter, we carry out KPSS tests for these tide gauge segments, which tests the opposite null hypothesis of sea level stability. Finally, we test the hypothesis that the apparent rise in sea levels is mistaken for the statistical phenomenon of "long memory" in tide gauge records. The KPSS tests indicate more widespread evidence of sea level rise whereas according to the ADF tests there is the least evidence of sea level rise.

Since the number of tide gauge segments is large and for each tide gauge segment several models are estimated the number of individual results approaches 10,000. We have therefore decided to summarize the results rather than provide details of all the models, which are available from the authors.

4.1 Classification of Tide Gauges

The Augmented Dickey-Fuller statistics (ADF) show that in the vast majority of segments and tide gauges sea levels are not rising. In only 22 segments and 20 tide gauges is there evidence of statistically significant sea level trends. These trends are stochastic rather than deterministic. The ADF tests were carried out with 6 monthly augmentations designed to correct the test statistic for serial correlation. Robustness checks show that these results do not depend on plausible changes in the number of augmentations (p in equation 1). These results show that the null hypothesis of rising sea levels is overwhelmingly rejected, sea level rise occurs in a small minority of tide gauge locations, and there is no evidence of global sea level rise.

The number of tide gauges with statistically significant trends is even smaller using Phillips-Perron tests, according to which as few as 5 segments and 5 tide gauges have trends. This difference between ADF and PP tests in which the null hypothesis is the same is well documented (Maddala and Kim 1998). It is also well documented that ADF tests may over-reject the null hypothesis.

On the other hand, the null hypothesis of sea level stability is supported the KPSS tests but less overwhelmingly. In Table 1 we report the number of tide gauges and segments that have time trends according to the KPSS tests. Recall that in these tests the null hypothesis is that SLR is zero.

Table 1: KPSS Classification of Tide Gauges and Segments by SLR

SLR = 0		SLR > 0		SLR < 0	
Segments	Tide Gauges	Segments	Tide Gauges	Segments	Tide Gauges
769	610	389	349	50	41

The classification process, reported in Table 1, comprises the following steps:

1. First we test whether sea levels are stationary in each of the 1208 segments. These tests show that in 556 cases we cannot reject the null hypothesis of stationarity. Therefore in these cases there cannot be a significant trend by definition.
2. Next, we test the null hypothesis of trend stationarity in the remaining 652 (1208 – 556) segments. We cannot reject this hypothesis in 439 cases, out of which 389 have positive time trends (SLR positive) and 50 have negative time trends (SLR negative).
3. Out of the 213 (1208 – 556 – 439) thus far unclassified segments we investigate why they are neither stationary nor trend stationary ($\gamma = 0$ in equation 2a). There

are three possible reasons. First, they might be difference stationary. Second, they might be driftless random walks, in which case there is no trend and the non-stationarity is induced by the variance which increases over time. Third, they might not be difference stationary, e.g. they might be stationary in second differences. In this case the trend is either accelerating or decelerating. We find that all 213 cases belong to the second category, and are listed as driftless random walks in Table 1. Indeed, the smallest p-value of the estimated stochastic trend is as large as 0.58.

The 769 segments reported in Table 1, which according to KPSS have no trend, consist of the 556 segments in step 1 and the 213 in step 3. Notice that SLR is positive in 389 segments and negative in 50 as classified in step 2. Table 1 presents the results for tide gauges as well as segments. Notice that some tide gauges changed their classification, e.g. they did not trend in one segment, but trended in another. There are 70 such segments involving 24 tide gauges. These tide gauges are identified separately below (Map 3).

The distribution of the KPSS time trends for tide gauges[7] is plotted in Figure 3 where the horizontal axis measures SLR in mm/year. Note that Figure 3 refers only to those tide gauges for which SLR is non-zero; it excludes the 610 tide gauges for which SLR is classified as zero. The conditional simple mean is 2.82 mm/year. This is made up of 3.55 mm/year for the tide gauges with positive trends and -5.97 mm/year for the tide gauges with negative trends. The unconditional mean is only 1.03 mm/year since SLR is zero in the majority of tide gauges, which are not featured in Figure 3. We have stated in section 2.4 that although we are unable to estimate GMSL over time, we may infer the recent rate of increase in GMSL. This is in fact equal to the unconditional, or grand, mean increase of 1.03mm/year. In section 4.4 we will present additional results, which lead us to conclude that the unconditional mean is lower than this.

As mentioned above, rejecting the null hypothesis that SLR is zero is not tantamount to accepting the alternative that SLR is positive. In fact, we find that it is much easier to reject this alternative than to reject the null hypothesis that SLR is zero. If SLR is accelerating, sea levels should be nonstationary in first differences, but stationary

[7] We omit 24 tide gauges whose classification varied across segments. For the remaining 325 tide gauges with multiple segments we use a weighted average in Figure 3.

in second differences. In none of the tide gauges and segments do the Dickey-Fuller and KPSS statistics support the accelerationist hypothesis.

We apply the tests suggested by Clemente at al (1998) and Lee and Strazicich (2001) to check for structural breaks in the estimated sea level trends. These tests are particularly relevant to the tide gauges dating back to the 19th and early 20th centuries. We distinguish between the following taxonomy of structural breaks:

i) The sea level trend is not fixed because there is continuous acceleration in SLR.

ii) SLR accelerates discontinuously, i.e. SLR is positive and stable, but it eventually increases.

iii) SLR is zero but it eventually turns positive, or vice-versa.

iv) SLR is zero but there is a discrete increase (decrease) in sea levels

Since, as mentioned, none of the tide gauge data had to be differenced twice to make them stationary, we may rule out continuous acceleration in SLR (case i). Nor do the structural break tests suggest that SLR accelerates discontinuously (case ii). We have already taken note of case iii since there are 26 tide gauges in which SLR is zero in one segment but positive in another. In case iv a discrete increase in sea levels might be mistaken by the KPSS test for a trend, which would over-estimate the incidence of SLR in Table 1. However, we have been unable to find such cases. In summary, structural break analysis suggests a few case iii tide gauges, but no tide gauges in the other categories.

4.2 Eye-ball Confirmation

In a Graphical Appendix (Appendix 1) we show that the KPSS classifications in Table 1 are consistent with "eye-ball tests" of the data. For these purposes we use tide gauges with at least 50 years of data. Figure A1 plots data for 10 tide gauges that are sampled randomly from the tide gauges which have no trend according to the KPSS test in Table 1. The tide gauge at Honolulu and perhaps the tide gauge at Newport (US) appear misclassified; SLR is positive rather than zero. Figures A2 and A3 refer to the KPSS classifications for positive and negative trends and record the estimated time trends. Since some of these trends are small, they are occasionally difficult to discern by the naked eye. All these tide gauges seem to be correctly classified but for the possible exception of Hamina, where there appears to be no trend. Therefore, out of the 30 tide gauges in Figures A1 – A3 the overwhelming majority appear to be correctly classified. Figure A4 plots data for tide gauges with long histories. It shows that these tide gauges are unusual in that many of them exhibit positive trends.

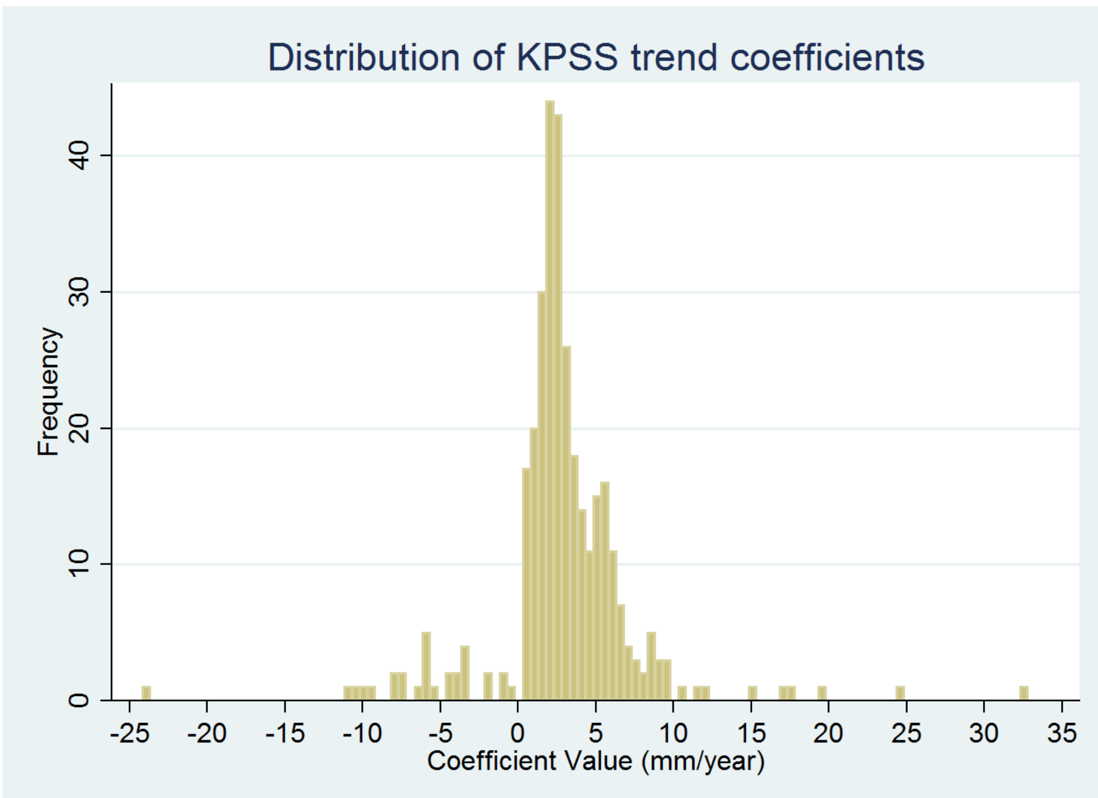


Figure 3: Distribution of KPSS Trend Coefficients

4.3 The Wood and the Trees

As mentioned in section 2.5, the focus on individual tide gauges might miss the wood for the trees; sea levels can rise globally even though locally they do not appear to be rising. To check for this we note that out of the 769 segments which are classified as trend-free, 42 percent of these non-significant trends are negative. In any case these insignificant trends are most probably positively correlated since the oceans form an interrelated system. Indeed, the KPSS residuals (ε in equation 2a) are positively spatially autocorrelated for common observations. Had they been negatively correlated and had the insignificant trends been predominantly positive there would have been a prima facie case that the wood has been missed for the trees. Since neither apply, we are confident that our bottom-up approach has not overlooked a positive global trend in sea level.

4.4 Long Memory

As mentioned in section 2.2, just because the KPSS statistic rejects the null hypothesis that d = 0 does not necessarily mean that the data are nonstationary since they might simply have long memories. According to Table 1 there are 389 tide gauge segments for which SLR appears positive. For these segments we estimate d and test the hypothesis that d < ½ using the semi-parametric estimator due to Robinson (1995)[8]. The number of tide gauges with long memory was found to vary directly with the power (a) of the number of observations (T^a). However, the number of tide gauges stabilized at a = 0.75.

Setting the test size at 5% it turns out that 297 of these 389 segments have estimates of d that are less than ½ in absolute value when a = 0.75. Sea levels at these segments have long memory which creates the misleading impression that they are rising. Therefore the number of tide gauge segments with positive sea level trends is 92 (389 - 297). Similar results are obtained by the Geweke – Porter-Hudak estimator. This means that long memory goes some way in explaining the contradiction between the KPSS and ADF results.

Since the estimates of d might be sensitive to seasonality, we seasonally adjusted the data using deterministic monthly components, and then estimated d using the seasonally adjusted data. For these purposes we used a sample of 233 tide gauge segments (out of 389) with sufficiently long records (at least 30 years) to estimate the seasonal components. The proportion of tide gauge segments with long memory increased from 76 percent to 79 percent. Therefore, our findings do not seem to be sensitive to seasonality in sea levels.

When the 297 tide gauge segments with long memory are excluded from the calculation of global mean sea level rise, the unconditional mean sea level rise is 0.39

[8] Unlike the more popular estimator or Geweke and Porter-Hudak (1981), Robinson's estimator is consistent for d > 1.

mm/year instead of 1.03mm/year according to the KPSS classification. This reduction is natural because tide gauges which previously were classified (by KPSS) to have sea level rise are now classified as having stable sea levels with long memory. However, the conditional mean when SLR > 0 increases from 3.55 mm/year to 4.24 mm/year. This stems from the fact that the tide gauges classified with long memory had estimated trends that were less than the average. Indeed, this is why the unconditional mean does not decrease by more than it did.

4.5 Tests for Sample Selection Bias in PSMSL Tide Gauge Locations

We test for sample selectivity in tide gauge locations by using the probit model described in section 2.3. The test is applied at the country level rather than the local level, and therefore the dependent variable is the number of tide gauges by country, which is zero if there are no tide gauges and positive otherwise. We carry out two kinds of tests. The first is based on the probit model in which the dependent variable is zero if the country has no tide gauges and one otherwise. The second test is based on the censored regression[9] model in which the dependent variable is the number of tide gauges operating by country, which in many cases is zero. The independent variables include GDP per capita, population and coastline lengths of the country[10]. We use satellite altimetry data for 1993 – 2010 (Map 2) to measure SLR for all countries[11] including the countries that have no tide gauges at all. Specifically we calculate mean SLR for each country. SLR may vary within countries because the coastline is long (as in Chile) or because there is more than one coast (as in France). The main question is whether this latter variable has a statistically significant effect on both the existence of tide gauges and their number. If not, the PSMSL sample of tide gauges is quasi random because the location of tide gauges is independent of SLR as measured by satellite altimetry.

Results are reported in Table 2. The main conclusion from the censored regression model (tobit) is that the number of tide gauges varies directly, as expected, with GDP per capita, population and length of coastline. It also varies directly with SLR (as measured by satellite altimetry) but this effect is not even remotely statistically

---

[9] Also known as Tobit regression. Censored regression assumes that the dependent variable is continuous if it is not zero.

[10] Data sources are provided in an Appendix 1.

[11] Satellite data do not cover Arctic and Antarctic coasts.

significant. Indeed, the p-value is almost 1. Therefore, although the location and number of tide gauges in PSMSL is obviously not random, it is quasi random because it is independent of SLR. The probit results differ from their tobit counterparts insofar as length of coastline does not affect the probability of a country having at least one tide gauge, but SLR has a positive effect, which is (marginally) statistically significant at conventional values (p = 0.05).

Taken together the two models suggest that SLR is at most weakly related to whether or not a country has at least one tide gauge, but it is unrelated to the number of tide gauges if there is at least one. We attach more importance to the tobit results because the classifications in Table 1 refer to tide gauges rather than countries. Since tide gauge locations are independent of SLR, they are quasi-random, and our estimates of global SLR are consequently unbiased. On the other hand, the probit results suggest that, if anything, we are likely to over-estimate SLR because countries with tide gauges happen to be located in parts of the world where SLR happens to be larger.

Table 2: Tobit and Probit Regressions for the Location of Tide Gauges

	Tobit	**Probit**
Intercept	-12.12 (3.129)	-0.068 (-0.35)
GDP/capita	1.98 (9.041)	3.76E-5 (2.46)
Population	7.37 E-8 (4.75)	8.63E-9 (2.58)
SLR	0.00931 (0.011)	0.089 (2.17)
Coastline	0.334 (4.28)	4.68E-7 (0.08)
Pseudo R^2	0.086	0.1404
Observations	164	164

Dependent variable: Tobit, number of tide gauges in each country in 2000. Probit, dummy variable = 1 if there is at least one tide gauge in each country in 2000. t – statistics in parentheses. Method of estimation: maximum likelihood.

The results in Table 2 indicate that in 2000 there is no evidence that tide gauge locations are correlated with SLR. Ideally, we would like to carry out this test for various years during the 20th century. Unfortunately this is impossible because satellite altimetry data are not available prior to 1992. We therefore cannot rule out the possibility that although tide gauge locations are independent of SLR in 2000, the same was not true, for example, in 1950.

4.6 Tests for Sample Selection Bias in the Location of Veteran Tide Gauges

Although we cannot test for sample selectivity prior to 2000, we can test whether there is sample selectivity in the location of veteran tide gauges, i.e. tide gauges that date back to the 19th and the early 20th centuries. These tide gauges are particularly important because of their role in data imputation at a time when there were relatively few tide gauges (Figure 1) and their geographical coverage was substantially less than it is today. In particular we investigate whether these veteran tide gauges are more likely to record sea level rise relative to their "younger" and more recent counterparts, and especially those that were installed since 1950.

Specifically, we test whether SLR was more likely to be positive among veteran tide gauges than among more recent tide gauges. Obviously, the latter tide gauges cover shorter observation periods than the former. For example, tide gauge (A) installed in 1950 records sea levels since 1950 in its location, whereas tide gauge (B) installed in 1850 covers 1850 – 1950 as well as 1950 onwards in its location. Our test essentially compares the two tide gauges over the same (shorter) time period for which the data overlap, i.e. A and B are compared since 1950. Recall that we have already checked that SLR since 1850 is stable in tide gauge B. If the test shows that SLR was more likely to be positive among the veteran tide gauges, it means that in 1900 our method would have led to the conclusion that SLR was greater globally than in 2000. It also means that imputed data based on these veteran tide gauges would over-estimate SLR.

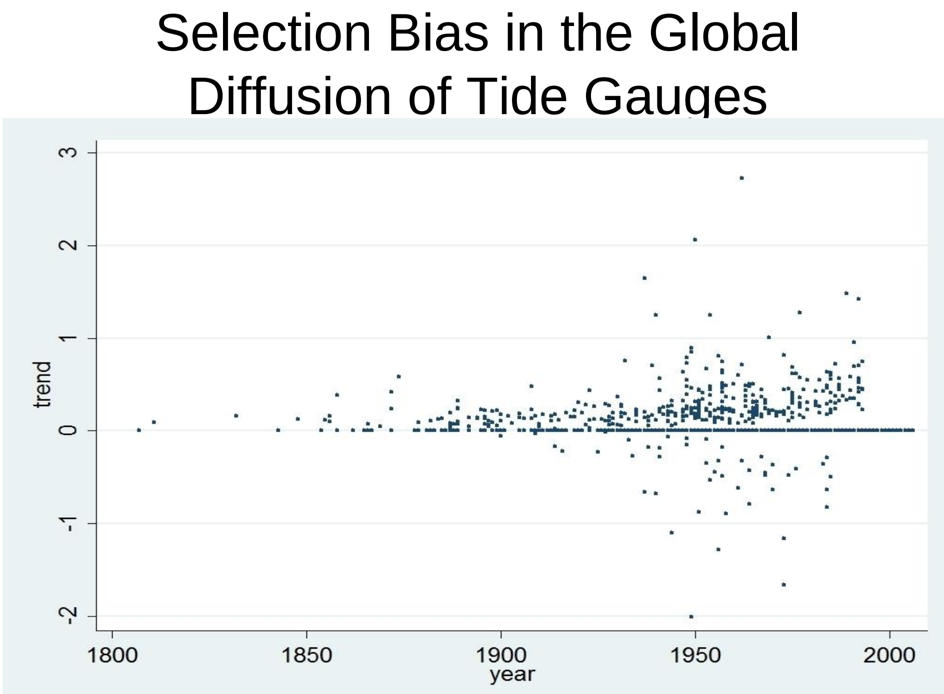


Figure 4: Selection Bias in the Global Diffusion of Tide Gauges

Figure 4 plots our (KPSS) estimates of sea level rise by year of tide gauge installation. It shows that for tide gauges installed in the 19th century SLR was either zero or positive. For tide gauges installed in the 20th century there is much more heterogeneity; SLR may be negative as well as positive and zero. In Table 3 we report a multinomial logit model which estimates the relationship between SLR status (positive, zero, negative) for individual tide gauges and their year of installation. Since in the majority of cases sea levels are classified as stable, the baseline is SLR = 0. The first column refers to the probability that SLR is negative and the second column refers to the probability that SLR is negative as a function of the year in which the tide gauge was commissioned. As indicated by the chi-square statistic the parameters are jointly significantly different from zero; the p-value of the multinomial logit model is close to zero.

Table 3 Multinomial Logit Model

Base: SLR = 0

	SLR < 0	SLR > 0
Intercept	-11.9709	0.02723
Year - 1807	0.15205	0.006851
$(\text{Year- }1807)^2$	-0.0005812	-0.0000633
Chi-square	31.33	
N	1000	

Since the oldest tide gauge dates back to 1807, the explanatory variables in Table 3 are expressed as the year in which the tide gauge was initiated minus 1807. A second order polynomial fits the data best. We interpret the parameter estimates in Table 3 by computing the predicted probabilities of SLR (positive, negative or zero) in Table 4.

Table 4 Probabilities for SLR

	SLR < 0	SLR = 0	SLR > 0
1850	0	0.45	0.55
1900	0.03	0.47	0.51
1950	0.07	0.55	0.38

The parameter estimates imply that the probability of tide gauges being installed in locations of rising sea levels has decreased over time, whereas the probability of their installation in locations where sea levels are stable has increased over time, as has the probability of their installation in locations where sea levels are decreasing. Therefore, although the current location of tide gauges in the PSMSL data is independent of SLR, the same cannot be said of the global diffusion of tide gauges. Tide gauges installed more recently have been in locations where sea levels were less likely to have increased. Alternatively, the tide gauges installed in the 19th century were in locations where sea levels were more likely to have increased.

4.7 Mapping SLR

In Map 3 we display the geographical distribution of the tide gauge classification based on the exclusion of tide gauges estimated to have long memory. Since this classification refers to tide gauge segments, and since segments can have conflicting classifications SLR is flagged as 'positive' (negative) if all the segments for that tide gauge are positive (negative) and as 'conflicting' otherwise. We also present the tide gauges as circles whose size is proportionate to the number of observations per tide gauge. Consequently, the largest circles mainly refer to the tide gauges installed in the 19th century, and the smallest circles mainly refer to tide gauges installed in the second half of the 20th century. Sea level rise is shown in red, sea level fall in green, stable sea levels in yellow, and conflicting classifications in black.

The overwhelming picture is yellow (no trend) because according to the results in section 4.4 sea levels are trending in only a minority of locations. We therefore concentrate on the regions in which sea levels are classified to be rising. There are only a few clusters of SLR in a predominantly trend-free world. The main areas of SLR include the Baltic Sea and the Russian Arctic. Map 3 would naturally be "redder" had we used the KPSS classification. It would also have included the Ring of Fire and the Atlantic coast of the United States as concentrations of sea level rise. No significant evidence of sea level rise can be discerned along the coast lines of the north-west and east sides of the North Atlantic, the north-east Pacific, the northern parts of the Indian Ocean, and the whole of Africa and South America. Finally, there is evidence of sea level fall in Alaska and the western coast of India.

It is notable that tide gauges with sea level rise are co-located with tide gauges that are trend-free. The same applies to the minority of tide gauges where sea levels are falling. This surprising pattern may also be found in Map 2 which plots the satellite altimetry data where dark blue (large negative SLR) grid points are located in the vicinity of dark red (large positive SLR) grid points.

Map 3 underscores visually the observation in section 4.6 that the tide gauges installed in the 19th century (represented by large circles) are likely to be red (sea level rise) whereas the more recent tide gauges (represented by small circles) are more likely to be yellow (stable sea levels). The map would look of course more yellow and less red had each tide gauge not been weighted by its vintage, i.e. had all the circles been the same size.

5. Discussion

This paper breaks new ground by studying tide gauge location to make two contributions to the study of global sea level rise. The first contribution is methodological and the second substantive. We claim that although extant tide gauges in PSMSL do not constitute a random sample of the world's coastlines, and are therefore selective, they constitute a quasi-random sample. A sample is quasi-random if sample selectivity is independent of the parameter of interest. In the present context, PSMSL constitutes a quasi-random sample for SLR if the location of tide gauges is independent of SLR.

We show that the location of tide gauges in PSMSL depends on a variety of factors such as GDP per head, population and coastline length. However, their location in 2000 is independent of SLR. This means that PSMSL may be used without recourse to the widespread practice of using imputed data for locations that do not happen to have tide gauges. A probit selection model suggests, if anything, that the probability of tide gauge presence varies directly with SLR in which case global sea level would be over-estimated.

Another methodological contribution is to apply recent developments in the statistical analysis of nonstationary time series data to test hypotheses about sea level rise. Specifically, we exploit the principle that if sea levels are rising the data cannot be stationary. We find that for the majority of tide gauges in PSMSL we cannot reject the hypothesis that the data are stationary.

Although the location of tide gauges in 2000 is a quasi-random sample, the same does not apply to tide gauges in 1900. This is because tide gauges with long histories were more likely to be installed in locations where sea levels happened to be rising. By contrast, the global proliferation of tide gauges that occurred especially during the second half of the 20th century was concentrated in locations where sea levels were more likely to be stable and even falling. Since data imputation gives greater weight to veteran tide gauges with long records, we suspect that this overweighting may induce positive bias in estimates of global sea level rise[12].

Scientists are impaled on the horns of a methodological dilemma. If they do not use imputed data, sample selection bias may be induced in estimates of global SLR. And if they resort to imputed data, measurement error might induce bias in estimates of global SLR. We resolve this dilemma by showing that PSMSL tide gauges are quasi-random. We hope that this concept of quasi-randomness is helpful and insightful.

This study shows that there is no sample selection bias in the conservative methodology because tide gauge locations are independent of SLR. On the other hand, imputation bias may be present because, as shown in this study, the global diffusion of tide gauges is negatively correlated with SLR. Since imputation attaches increasing importance to tide gauges with long histories, and since their locations are positively correlated with SLR, reconstructed estimates of global SLR have positive bias.

The substantive contribution of the paper is concerned with recent sea level rise in different parts of the world. Our estimates of global SLR are markedly smaller than estimates obtained using imputed data. While we find that sea levels are rising in 8 – 30 percent of tide gauge locations, SLR is not a global phenomenon. Consensus estimates of recent global mean sea level rise are about 2.2 mm/year. Our estimate is much smaller and ranges between 0.39 – 1.03 mm/year. We suggest that the difference between the two estimates is induced by the widespread use of imputed data which inform the consensus estimates adopted by IPCC in its Fifth Review[13].

Finally, sea level rise (or fall) is local rather than global and is concentrated in the southern Baltic, the Ring of Fire, and the Atlantic coast of the US. By contrast the north-west Pacific coast and north-east coast of India are characterized by sea level fall. In the minority of locations where sea levels are rising the mean increase is 3.55 – 4.24 mm/year, and in some locations it is as large as 9 mm/year. The fact that sea level rise is not global should not detract from its importance in those parts of the world where it is a serious problem.

[12] Douglas (2001) writes (p 56), "It is that the longer the period of sea level variation, the greater the spatial extent of that signal. Thus very long records do not require the coverage needed by the shorter ones to establish a value of global change." This statement assumes incorrectly that the tide gauge locations providing these long records are uncorrelated with SLR.

[13] The Fifth Review has adopted the thesis of Vermeer and Rahmstorf (2009) who attribute global sea level rise to global warming. However, our results question whether global sea levels are in fact rising.

References

Bâki Iz H, and Ng H.M. 2005. Are Global Tide Gauge Data Stationary in Variance? *Marine Geodesy*, 28: 209–217,

Chambers, D. P., Mehlhaff C.A., Urban T.J, and Fuji D. (2002) Low-Frequency Variations in Global Mean Sea Level: 1950-2000, *Journal of Geophysical Research*, 107, 3026-3034.

Christiansen, B., T. Schmith, and P. Thejll 2010. A surrogate ensemble study of sea level reconstructions. *Journal of Climate*, 23: 4306-4326

Church, J. A., and N. J. White 2006. A 20th century acceleration in global sea-level rise. *Geophysical Research Letters*, 33: L01602.

Church, J. A., N. J. White, R. Coleman, K. Lambeck, and J. X. Mitrovica 2004 Estimates of the regional distribution of sea level rise over the 1950–2000 Period. *Journal of Climate*, 17: 2609-2625.

Clemente, J., A. Montanes and M. Reyes 1998. Testing for a unit root in variables with a double change in the mean, *Economics Letters*, 59: 175-182.

Douglas B.C.(2001) Sea Level Change in the Era of the Recording Tide Gauge, pp 37-63 in Douglas B.C, Kearney M.S and Leatherman S.P., *Sea Level Rise: History and Consequences*, Academic Press, San Diego, CA.

Dickey, D.A. and W.A Fuller 1981 Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49: 1057-1079.

Emery K.O and Aubrey DF.G (1991) *Sea Levels, Land Levels and Tide Gauges*, Springer New York.

Geweke J. and S. Porter - Hudak 1983 The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 2: 221-238.

Granger C.W.J and R. Joyeux 1980 An introduction to long memory models and fractional differencing. *Journal of Time Series Analysis*, 1: 15-39.

Groger M and Plag H.P. (1993) Estimation of Global Sea Level Trend: Limitations from the Structure of the PSMSL Global Sea Level Data, *Global and Planetary Change*, 8, 161-179.

Heckman J.J. 1976 The common structure of statistical models of truncation, sample selection and limited dependent variables. *Ann Econ Soc Measure* 5:475–492

Jevrejeva, S., A. Grinsted, J. C. Moore, and S. Holgate 2006 Nonlinear trends and multiyear cycles in sea level records. *J. Geophys. Res*, 111: C09012.

Klein, M. and M. Lichter 2009 Statistical analysis of recent Mediterranean Sea-level data. *Geomorphology*, 107: 3-9.

Korozumi, E. 2002 Testing for stationarity with a break. *Journal of Econometrics*, 108: 63-99.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin 1992 Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54: 159-178.

Lee, S. and M. Strazicich 2001 Testing the null of stationarity in the presence of a structural break. *Applied Economic Letters,* 8: 377-382.

Mitchum G.T., Nerem R.S. Merrifield M.A. and Gehrel W.R (2010) Modern Sea Level Change Estimates pp 122-142 in Church J.A., Woodworth P.C., Aarup T and Wilson W.S (eds) *Understanding Sea Level Rise and Variability*, Blackwell, UK

Perron P (1989) The Great Crash, the Oil Shock and the Unit Root Hypothesis, *Econometrica,*57, (1989), 1361-1402.

Peltier, W. R. 2001 Global glacial isostatic adjustment and modern instrumental records of relative sea level history. *Sea Level Rise: History and Consequences*, edited by B. C. Douglas, M. S. Kearney, and S. P. Leatherman. San Diego: Academic Press

Pirazolli P.A. (1993) Global Sea Level Changes and their Measurement, *Global and Planetary Change*, 8, 135-148.

Phillips, P.C.B and Perron P. 1988 Testing for a Unit Root in Time Series Regression, *Biometrika*, 75, 335–346

Robinson, P.M. 1995 Gaussian semiparametric estimation of long run dependence. *Annals of Statistics*, 22: 1047-1072.

Said, S. and D. Dickey 1984 Testing for unit roots in autoregressive-moving-average models with unknown order. *Biometrica*, 71: 599-607.

Sowell, F.B. 2002 Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, 53: 185-188.

Tai C.K 2001 Infering the global mean sea level from a global tide gauge network, *Acta Oceanol. Sin*., 30 (4), 102-106.

Vermeer, M. and S. Rahmsdorf 2009 Global sea level linked to global temperature. *PNAS*, 106: 21527-21532.

Appendix 1: Data

Tide Gauges

The data are obtained from PSMSL's website. This dataset consists of a monthly calendar record per tide-gauge of mean sea level. While each station supplies its own record on a metric scale (also known as the raw data) PSMSL coverts these data using a joint datum for all stations. These datum reductions create a Revised Local Reference measure to which we apply Peltier's VM2 GIA correction (Peltier 2001). RLR measures deviations in millimeters from the 7000mm below mean sea level datum.

Data for Table 2

Data on GDP per capita were taken from the World Bank (Indicators: GDP Current US$ and Population, total) and the United Nations (Per capita GDP at current prices – US$ and Total population). We used the World Bank as the primary source of data and supplemented it with the UN's records to fill missing values where applicable.

Data for Map2

The data on satellite altimetry reported in Map 2 is obtained using the gridded, multi-mission Ssalto/Duacs data since 1993 available on the AVISO website. The data on coastline lengths is from http://chartsbin.com/view/ofv.

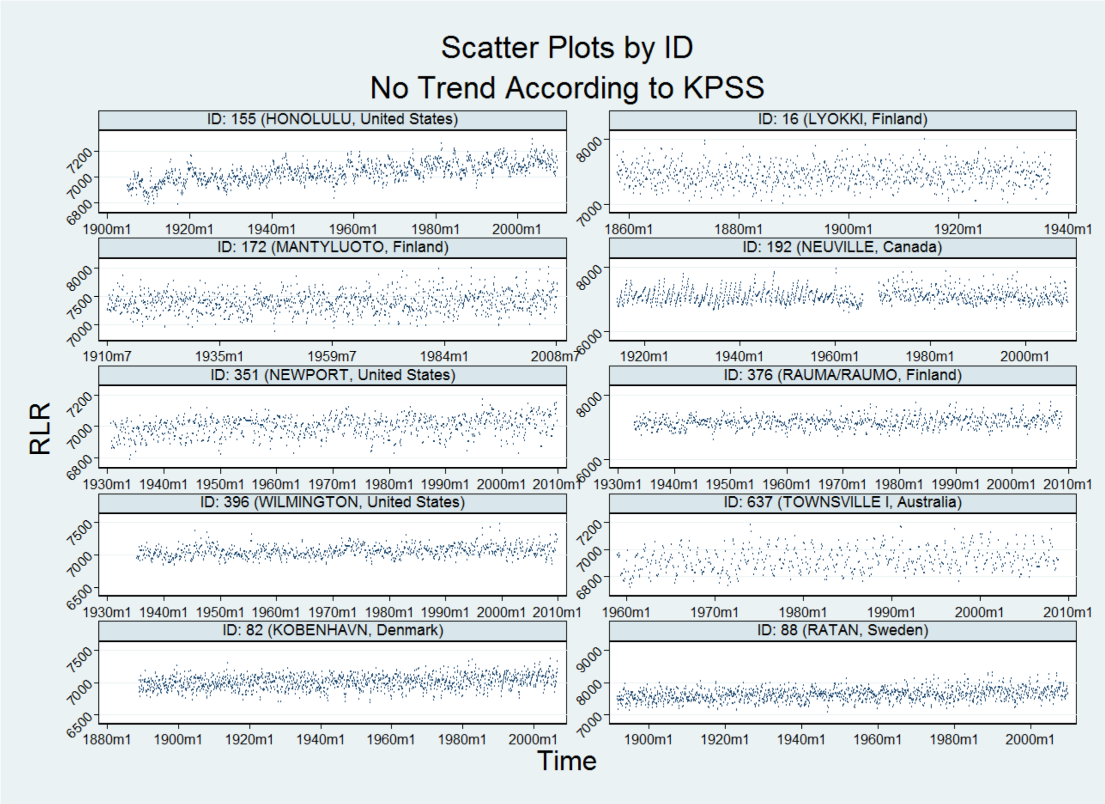


Figure A1: Scatter Plots by ID: No Trend (KPSS)

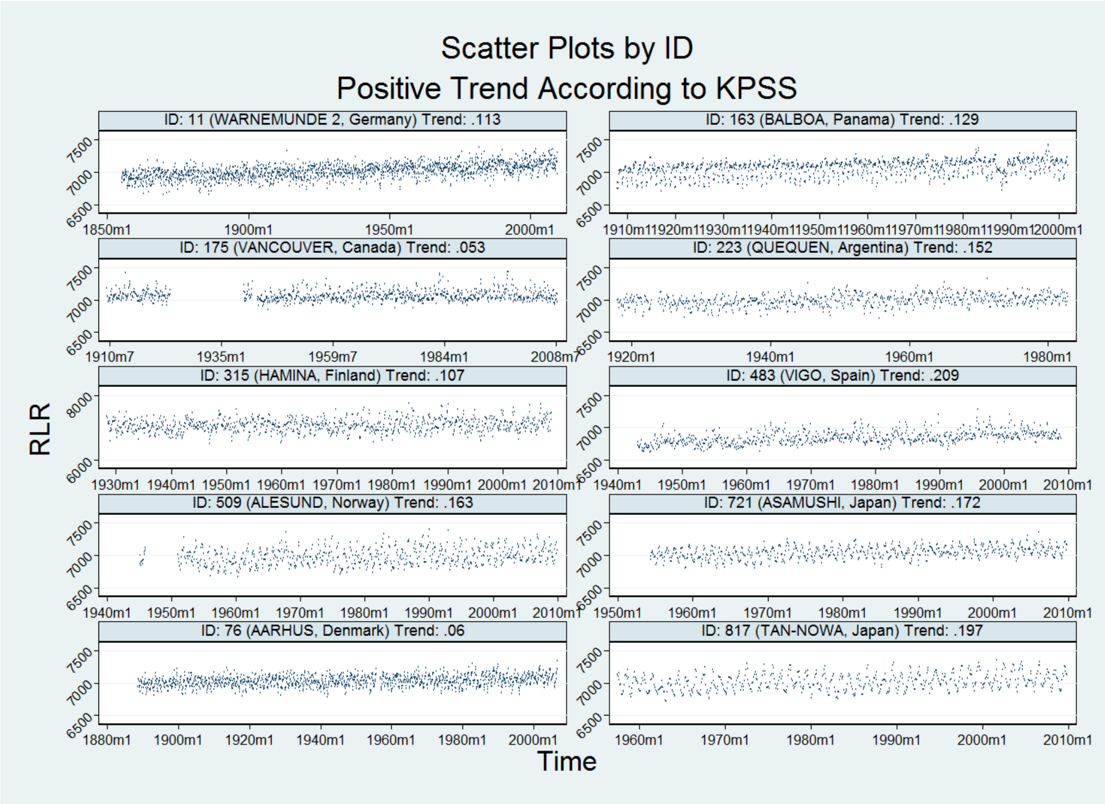


Figure A2: Scatter Plots by ID: Positive Trend (KPSS)

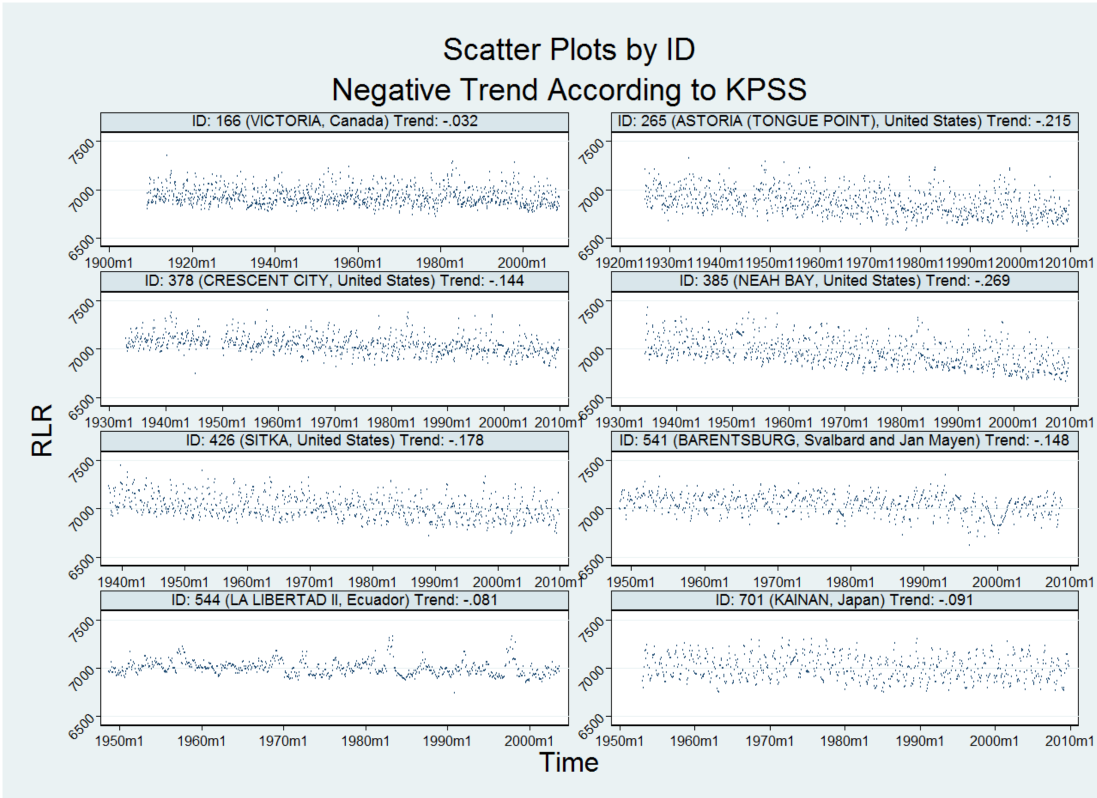


Figure A3: Scatter Plots by ID: Negative Trend (KPSS)

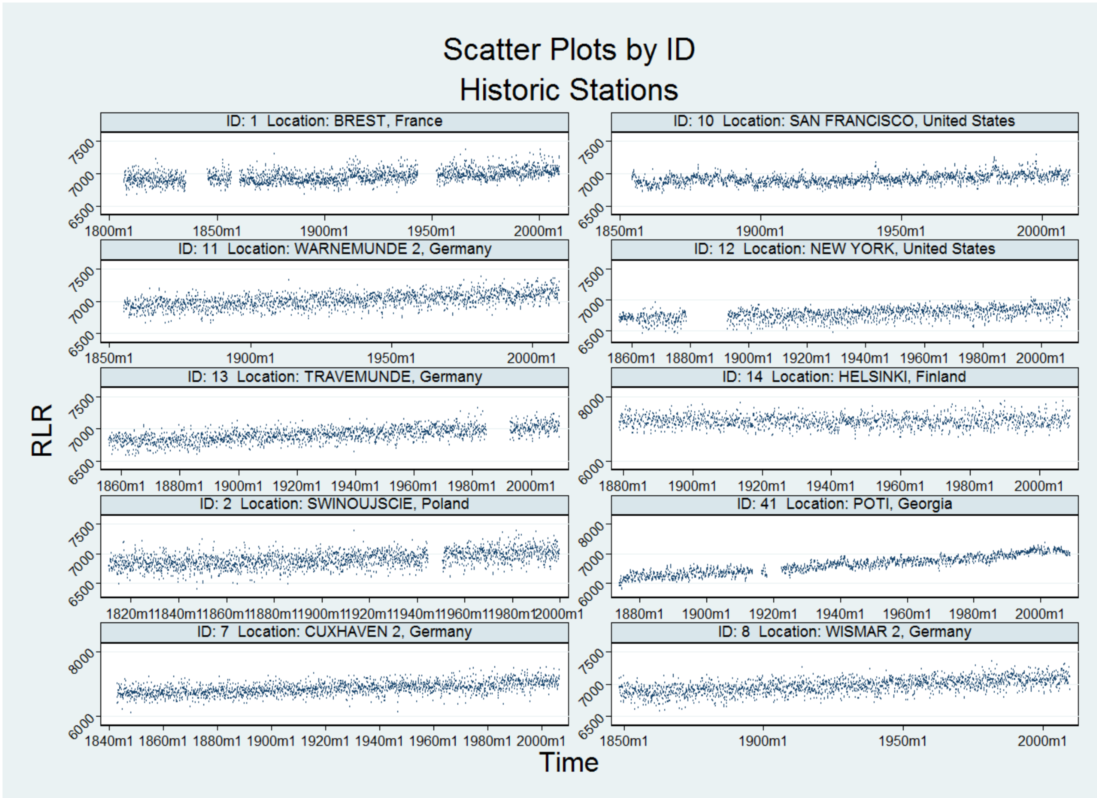
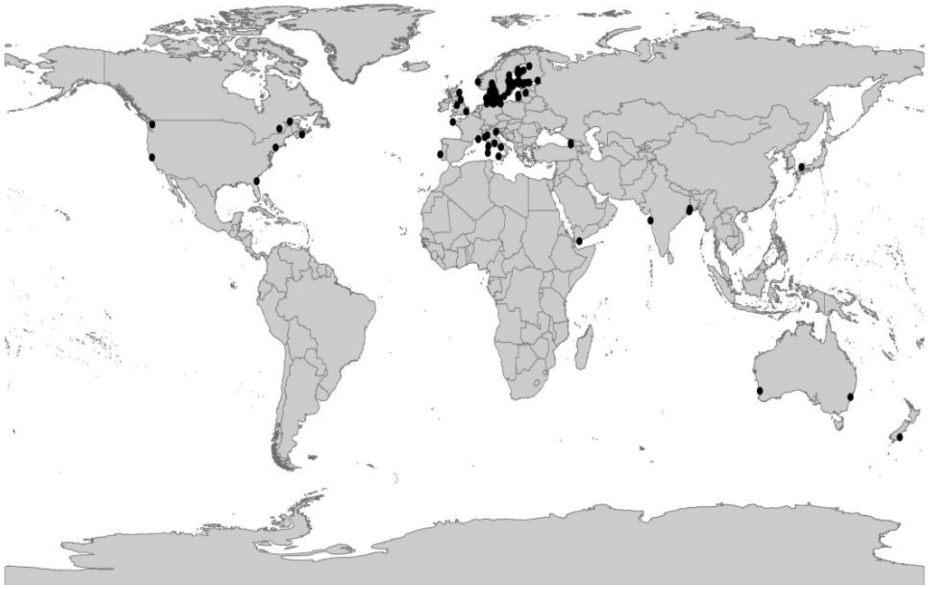


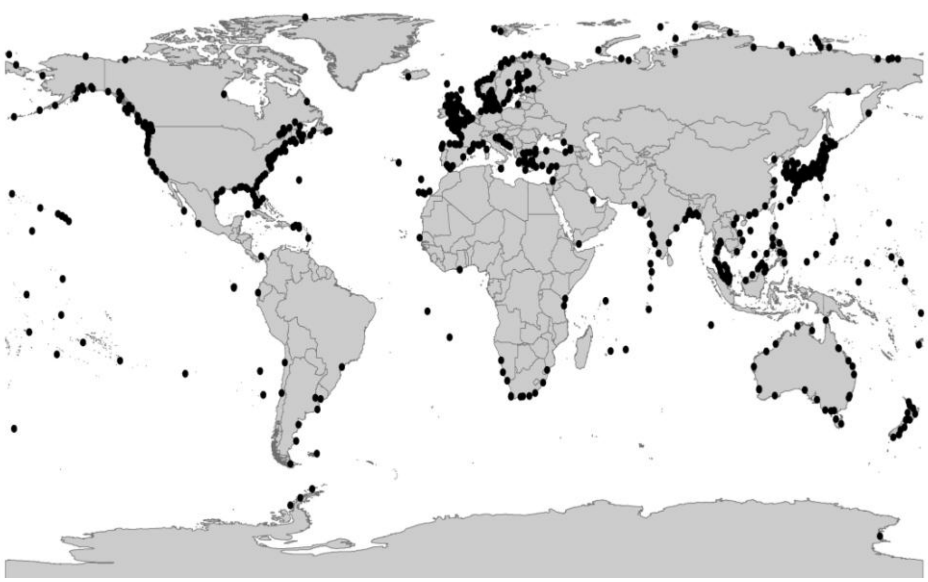
Figure A4: Scatter Plots by ID: Historic Stations

Map 1: The Global Coverage of Tide Gauges

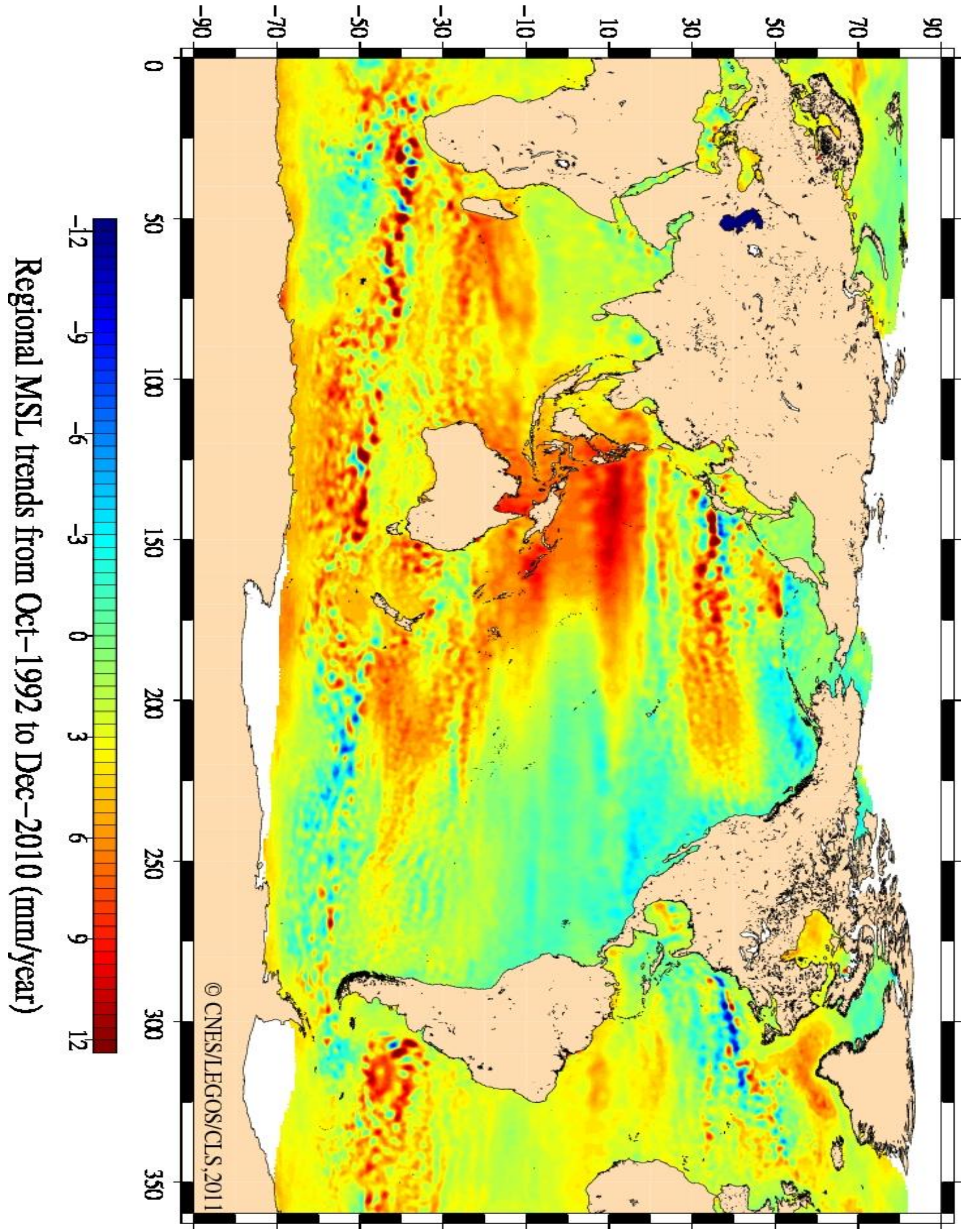
Tide Gauges in 1900



Tide Gauges 2000



Map 2: Global SLR using satellite altimetry data for 1993 – 2010



Map 3: Tide Gauges classified by seal level trends and number of observations