

TIDE GAUGE LOCATION AND THE MEASUREMENT OF GLOBAL SEA LEVEL RISE

Michael Beenstock¹

Daniel Felsenstein^{2*}

Eyal Frank¹

Yaniv Reingewertz¹

¹Department of Economics, Hebrew University of Jerusalem, Jerusalem 91905,
Israel

²Department of Geography, Hebrew University of Jerusalem, Jerusalem 91905,
Israel, Tel 972-2-5883343; msdfels@mscc.huji.ac.il

*Corresponding author

Abstract

The location of tide gauges is not random. If their locations are positively (negatively) correlated with SLR, estimates of global SLR will be biased upwards (downwards). We show that the location of tide gauges in 2000 is independent of SLR as measured by satellite altimetry. Therefore PSMSL tide gauges constitute a quasi-random sample and inferences of SLR based on them are unbiased, and there is no need for data reconstructions. By contrast, tide gauges dating back to the 19th century were located where sea levels happened to be rising. Data reconstructions based on these tide gauges are therefore likely to over-estimate sea level rise.

We therefore study individual tide gauge data on sea levels from the Permanent Service for Mean Sea Level (PSMSL) during 1807 – 2010 without recourse to data reconstruction. Although mean sea levels are rising by 1mm/year, sea level rise is local rather than global, and is concentrated in the Baltic and Adriatic seas, South East Asia and the Atlantic coast of the United States. In these locations, covering 35 percent of tide gauges, sea levels rose on average by 3.8mm/year. Sea levels were stable in locations covered by 61 percent of tide gauges, and sea levels fell in locations covered by 4 percent of tide gauges. In these locations sea levels fell on average by almost 6mm/year.

Acknowledgment; Thanks to Michal Lichter for assistance with Map 3.

Introduction

The debate over sea level rise (SLR) has been heavily informed by the availability of tide gauge data. However, as is well known, the global coverage of tide gauge stations is patchy and incremental and is often influenced by local contingencies. Although the coverage has greatly improved, especially in the second half of the 20th century, there are still many locations where there are no tide gauges, such as the southern Mediterranean, while other locations are under-represented, especially in Africa and South America. Surprisingly therefore, the study of tide gauge locations has not been an issue in the study of SLR.

The imperfect coverage of tide gauges has induced many investigators to engage in data reconstruction (Church et 2004, Church and White 2006, Jevrejeva et al 2008, Merrifield et al 2009 and Woodworth 2009). The data used in these studies are invariably obtained from the tide gauges feeding the Permanent Service for Mean Sea Level (PSMSL). However, the actual measurements in the PSMSL data are usually supplemented because the global coverage by tide gauges is incomplete. Satellite altimetry data are used since the 1990s, and prior to that, data are reconstructed for locations where there were no tide gauges (Christiansen et al 2010). Differences in estimates of global SLR may be attributed to the different reconstruction methods that have been used.

The motivation for using data reconstructions presumably stems from the incomplete global coverage of tide gauges, and the concern that the sample of tide gauge locations in PSMSL may not be representative. Had the location of tide gauges been determined as a random sample of the population, it should not have mattered that some locations do not have tide gauges. However, the global diffusion of tide gauges has not followed such a statistical design. The richer countries have been able to afford more tide gauges and the Southern Hemisphere is under-represented in PSMSL. Indeed, this kind of under-representation has varied over time.

Just because a sample does not happen to be random does not necessarily mean that parameters estimated from it must be biased. Therefore estimates obtained from PSMSL data are not automatically biased estimates of global SLR simply because PSMSL happens to be a non-random sample. Non-random samples will be quasi-random if the units participating in the sample 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 infer unbiased estimates of global SLR. 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.

Therefore the crucial question is: are tide gauges more or less likely to be located where SLR is greater? There is an a priori reason why tide gauge location should be positively 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 2011). Where SLR is zero, there is no need for monitoring, and it is less likely therefore that the government will incur the investment and current costs of tide gauges. It is harder to propose an a priori reason why tide gauge locations should be negatively correlated with SLR. 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 reconstruction would be unnecessary.

We use satellite altimetry data to show that as of 2000, tide gauge locations were independent of SLR. However, in the 19th and early 20th centuries tide gauge locations were correlated with SLR. This means that the global diffusion of tide gauges is negatively correlated with SLR. It also means that the tide gauge data currently in PSMSL constitute a quasi-random sample for the purposes of estimating current global SLR. Our main result is that sea level is rising in about a third of tide gauge locations. However, the same was not true in the past. The tide gauges installed a century ago and before, were more likely to be located where SLR was larger.

In this paper we do not use data reconstructions. This conservative approach makes inferences about global SLR using actual data not supplemented by reconstructed data, and is valid provided the data are a quasi-random sample. Reconstructed data are inevitably noisy estimates of their true, unknown counterparts. If their measurement errors are not independent of SLR, estimates of SLR obtained from combining actual and reconstructed tide gauge data will generate biased estimates of global SLR.

Researchers are impaled on the horns of a methodological dilemma: to reconstruct or not to reconstruct. There are dangers in both. On the one hand, the

conservative methodology might induce sample selection bias. On the other hand, the use of reconstructed data might induce reconstruction bias. This study shows that there is no sample selection bias in the conservative methodology because tide gauge locations are independent of SLR. However, reconstruction bias may be present because, as shown in this study, the global diffusion of tide gauges is correlated with SLR.

In addition to the methodological implications of selectivity in tide gauge locations, we also draw attention to recent methodological developments concerning the estimation of time trends. Time series data are non-stationary when their sample means and/or variances depend upon time. If SLR is positive, mean sea levels rise, in which case sea level data must be non-stationary (Bâki Iz and Ng 2005). We use stationarity tests to determine whether sea levels are trend-free, and estimate the trend for each tide gauge otherwise. Some investigators have used time series data on changes in sea levels rather than absolute levels because it is difficult to compare absolute sea levels in different locations. We show that this common practice might not be innocent; it may generate spurious trends in estimates of SLR. If sea levels are non-stationary, changes in sea levels will tend to be stationary. However, if sea levels are stationary, changes in sea levels might induce spurious time trends.

2. Methodology

2.1 Non-stationary Data

A time series is non-stationary when its sample moments depend on when they 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.

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 (1981) suggested the following test for stationarity¹:

$$\Delta Y_{jt} = \alpha_j + \beta_j Y_{jt-1} + e_{jt} \quad (1)$$

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

where e_j is assumed to be identically and independently distributed (iid) with variance σ_j^2 . If $\beta_j = 0$, equation (1) is a random walk with drift α_j , which represents the time trend in Y_j . If, however, β_j is negative Y_j ceases to be a random walk with drift in which case Y_j is trend-free². The distribution of estimates of β_j is not normal under the null hypothesis that $\beta_j = 0$. The Dickey-Fuller statistic (DF) looks like a t-statistic because it is defined as the estimate of β_j divided by its standard deviation, but it does not have a t-distribution. In fact the critical value ($p = 0.05$) of DF when there are 100 observations is -2.89 rather than -1.96 as suggested by the conventional t-statistic. If the DF statistic is smaller than -2.89, Y_j is stationary, in which case it does not have a trend that is statistically significant. If $DF > -2.89$, Y_j is not stationary. In this case we cannot reject the null hypothesis that $\beta = 0$, and we cannot rule out the possibility that there is a trend in Y_j .

A variety of tests have been suggested to distinguish between breaks in time trends and levels. For example, a permanent change in the level may create the illusion that the data are non-stationary when the opposite is true. In terms of equation (1) β_j is negative when α_j changes from some date. These tests (Clemente et al. 1998) penalize the 'data mining' that is typically involved when investigators search for a break-point at some unknown date. Indeed, because of this, the critical value of DF is -4.5 instead of -2.89.

The DF statistic tests the null hypothesis that the data are non-stationary. The KPSS statistic (Kwiatkowski, Phillips, Schmidt and Shin 1992) tests the null hypothesis that the data are stationary. KPSS is based on the following model:

$$Y_{jt} = \eta_j + \delta_j Z_{jt} + \varepsilon_{jt} \quad (2a)$$

$$\Delta Z_{jt} = \gamma_j + v_{jt} \quad (2b)$$

where ε and v are independent iid random variables. Since Z is generated by a random walk, it must be non-stationary. Therefore, Y_j is non-stationary if δ_j is non-zero, in which case the trend in Y_j is γ_j . Lee and Strazicich (2001) and Korozumi (2002) have extended the KPSS test to the case in which there is a structural break in either the level or the trend in the time series.

Rejecting the null hypothesis of stationarity is not conceptually equivalent to accepting the null hypothesis of non-stationarity, and vice-versa³. Therefore DF and

² In this case Y_j mean-reverts to $-\alpha_j/\beta_j$.

³ Failing to prove guilt is not equivalent to proving innocence, and vice-versa.

KPSS tests are not mirror images. If DF rejects non-stationarity and KPSS fails to reject stationarity, we may be reasonably confident that the data are stationary. Ambiguity arises, however, if DF and KPSS conflict because e.g. DF does not reject non-stationarity while KPSS does not reject stationarity or vice-versa. 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 DF test is appropriate. We use both of these tests. However, we prefer the KPSS test for two reasons. First, if the claim is that sea levels are rising, the null hypothesis should be that they are stable. Second, as we explain in section 4, KPSS test statistics are more likely to reject the null hypothesis that sea levels are stable, than the DF test statistics are likely not to reject the null hypothesis that sea levels are rising. Therefore, the KPSS tests are more conservative and are more likely to show that sea levels are rising.

The DF test assumes that e in equation (1) is serially independent. If it is not, various corrections have been suggested including the augmented DF statistic (ADF), in which lags of the dependent variable in these equations are specified as covariates (Said and Dickey, 1984), and the Phillips-Perron statistic (Phillips and Perron, 1988), in which robust standard errors are calculated for the estimate of β_j using a nonparametric correction for serial correlation. Similarly, the KPSS test assumes that ε in equation (2a) is serially independent. If it is not, the KPSS test uses a nonparametric correction for serial correlation.

2.2 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. Let T_j^* be the benefit to the relevant authority of having a tide gauge in location j . Suppose that T_j^* is hypothesized to depend on variables such as GDP per capita of the country in which j is located, population, and length of coastlines denoted by vector X_j . In addition, T_j^* is hypothesized to depend on SLR since if sea level is rising there may be a greater need for monitoring. Finally, e is a normally distributed random variable with zero mean capturing unobservable phenomena which determine benefit from tide gauges in the location:

$$T_j^* = \delta + \theta X_j + \phi SLR_j + e_j \quad (5)$$

Let T_j denote a dummy variable that equals 1 if there is a tide gauge in location j and zero otherwise. Since T_j^* is not observed, we assume that there will be a tide gauge in location j if the benefit is positive, i.e. $T_j = 1$ if $T_j^* > 0$. 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. 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).

2.3 Spurious Time Trends

Suppose SLR is zero, and the trend is estimated using data on *changes* in sea levels as in Church and White (2008). The estimate of the average change in sea level during T time periods is defined as the change in sea level over the entire period divided by the number of observations, i.e. $\tau = (Y_T - Y_1)/T$. Therefore, if the last observation in the data (Y_T) just happens to exceed the first observation (Y_1) τ is positive. Since SLR is zero a spurious time trend is estimated. The reason is that all the interceding observations, which are stable, have been ignored by looking at changes rather than levels.

The t-statistic for τ is defined as τ divided by its standard deviation. If SLR is zero it may be shown that this t-statistic equals:

$$t_\tau = \frac{Y_T - Y_1}{\sqrt{2sd(Y)}} \quad (6)$$

If t_τ exceeds 1.96 the time trend will appear to be statistically significant, and is therefore spurious. In section 4.3 we provide illustrations of spurious time trends, which are "statistically significant".

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 consisted 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 reconstructed, but this might induce reconstruction 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 use our results to construct measure of 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 as indicated in Figure 1. Therefore it is impossible to construct GMSL prior to 1950 unless one is prepared to reconstruct data for tide gauges installed between 1900 and 1950, or some other base year⁴. As GMSL approaches the base year the rate of data reconstruction naturally becomes increasingly large.

By contrast our approach allows us estimate recent GMSL rise. 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 reconstructing 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.

3. Data

When the data were retrieved (2011) PSMSL (www.pol.ac.uk/psml) comprised (in December 2009) 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.

⁴ For example, in Jevrejeva et al (2006) the base year is 1810 and in Church and White (2006) it is 1870.

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.

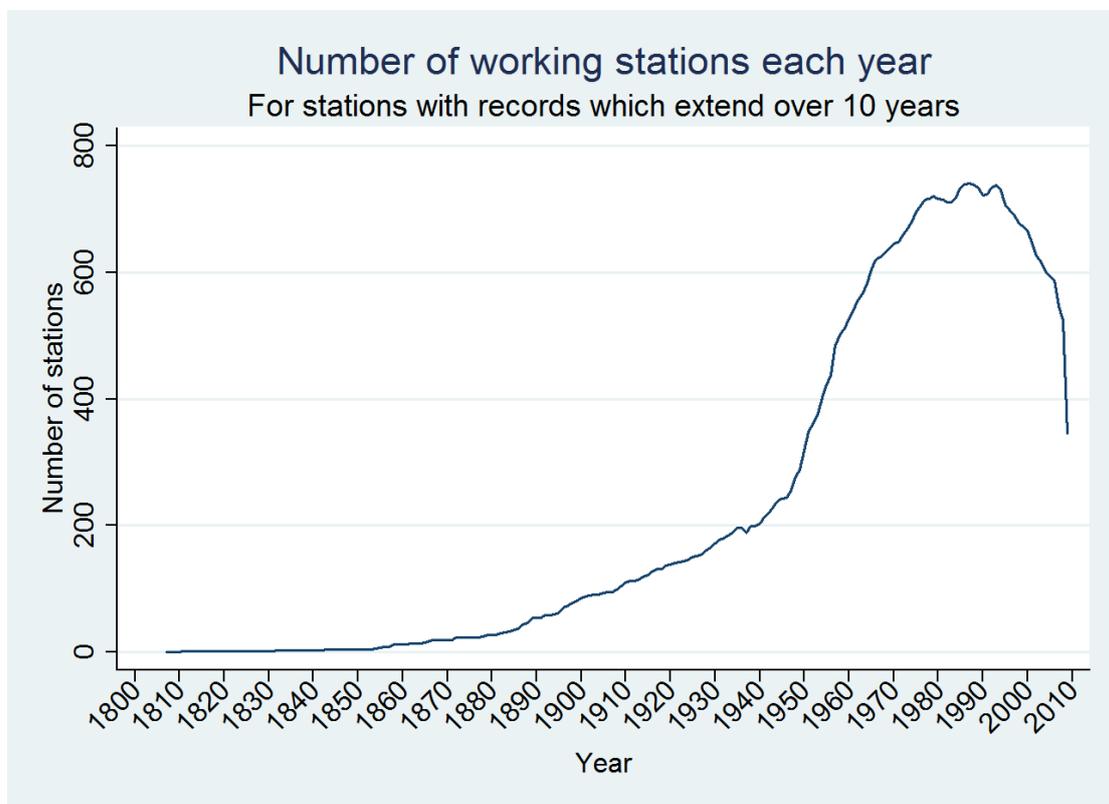


Figure 1

Map 1 shows that the global diffusion of tide gauges has been far from random. In 1990 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-sampled. By contrast Japan which had only one tide gauge in 1900 had more than a hundred in 2000.

Only 140 (Figure 2) of these 1000 tide gauges had complete data records. Many records are incomplete. Figure 3 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 clustered, i.e. the data gaps 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 (see Data Appendix). In the third case the data are split and treated as separate segments. These imputations and interpolations should be distinguished from reconstructions because they refer to existing tide gauges rather than to tide gauges that do not exist. Also, they constitute a tiny fraction of the data. Out of the 1000 tide gauges in the analysis, 823 are unsplit, 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 segments⁵.

⁵ Examples of split tide gauges may be seen in the Graphical Appendix: Neuville (Canada) in Figure A1, Vancouver in Figure A2, Brest and Swinoujcie (Poland) in Figure A4. Klein and Lichter (2009), for example, estimate time trends from such data assuming that none of the data are in fact missing.

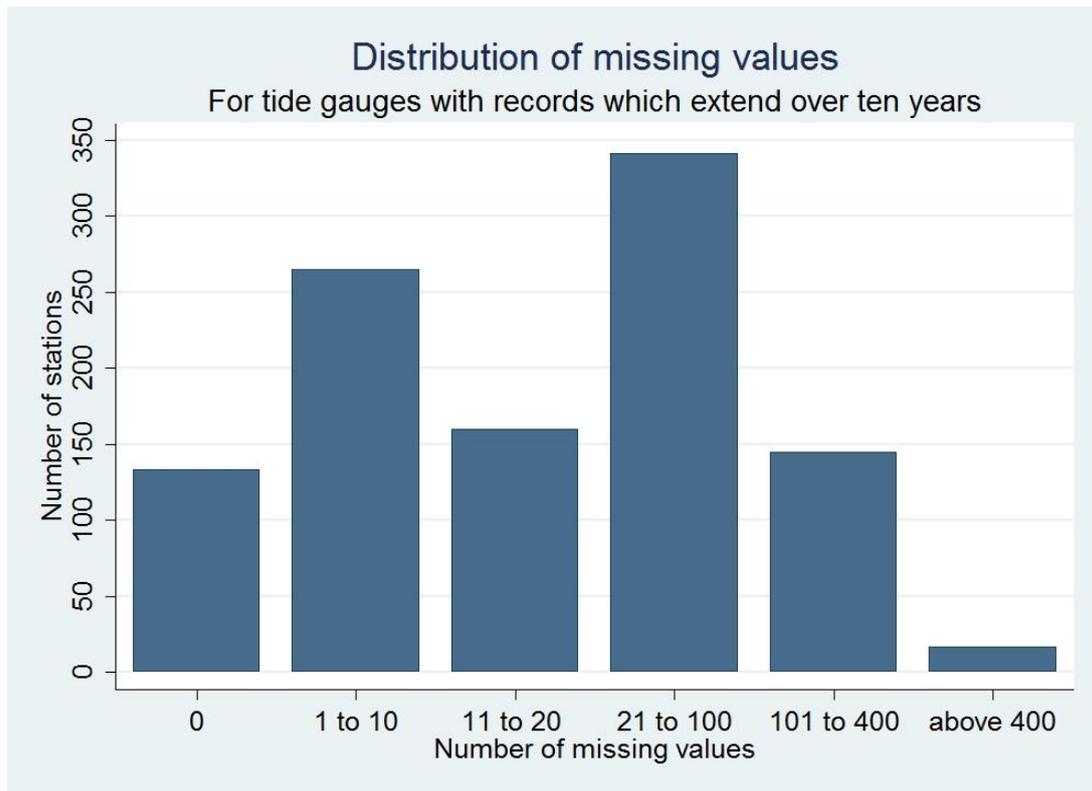


Figure 2

4. Results

4.1 Classification of Tide Gauges

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.

The distribution of the KPSS time trends for tide gauges⁶ is plotted in Figure 3 where the horizontal axis measures SLR in mm/year. The conditional mean is 2.93 mm/year. This is made up of 3.79 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 610 tide gauges, which are not featured in Figure 3.

As mentioned, 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. The Augmented Dickey-Fuller statistics (ADF), which test the alternative hypothesis,

⁶ 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.

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 a significant sea level trend. The ADF tests were carried out with 6 monthly augmentations designed to correct the test statistic for serial correlation. We also carried out a non-parametric alternative suggested by Phillips and Perron (1988) according to which only 5 segments and 5 tide gauges have trends. These test statistics indicate that statistically significant sea level trends occur in only a small minority of cases.

If SLR is accelerating, sea levels should be nonstationary in first differences, but stationary 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 et 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.

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.

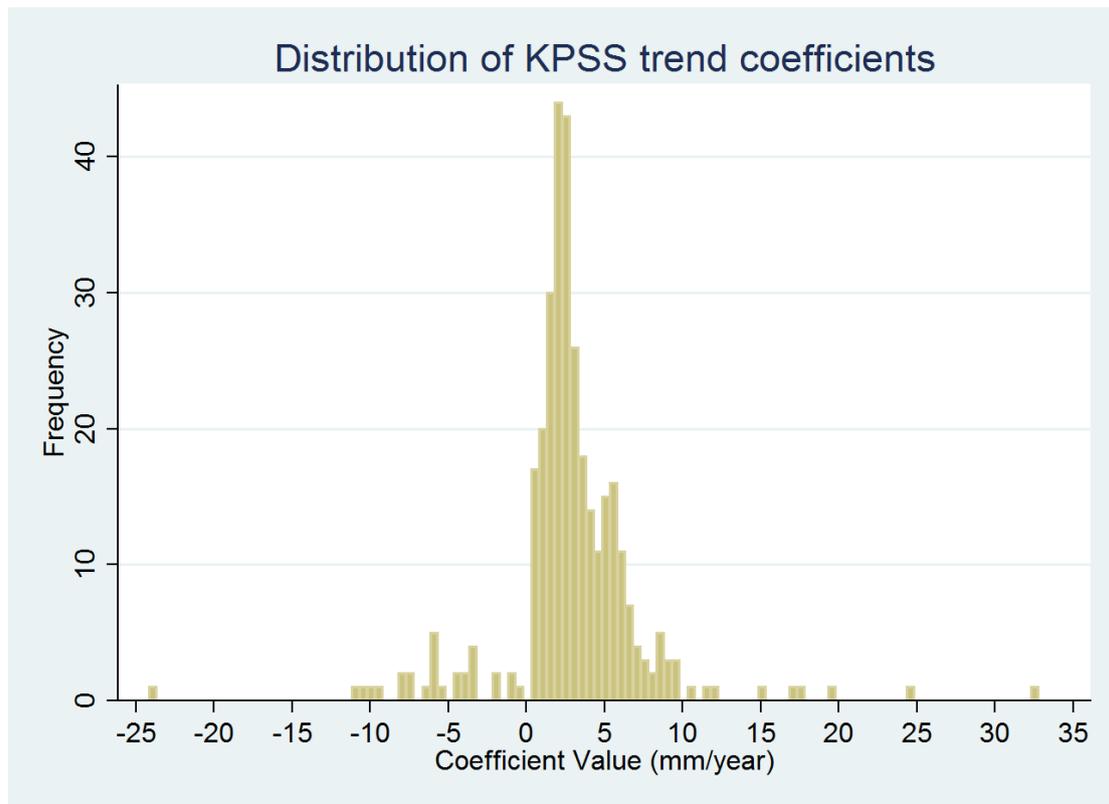


Figure 3

In a Graphical Appendix 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 appears misclassified and perhaps the tide gauge at Newport (US). 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.

4.2 Tests for Spurious Time Trends

We apply the spurious time trend test, described in section 2.3, to 545 of the 610 tide gauges, which according to the KPSS test are trend-free, and for which the final observation exceeds the first. In 92 of these cases the t – statistic of the spurious time trend exceeds 1.96. The largest t – statistic is 4.5. The histogram of these t - statistics is reported in Figure 4. Therefore, the incidence of spurious time trends is widespread. This means that estimates of SLR based on changes in sea levels rather than sea levels themselves are inclined to be spurious.

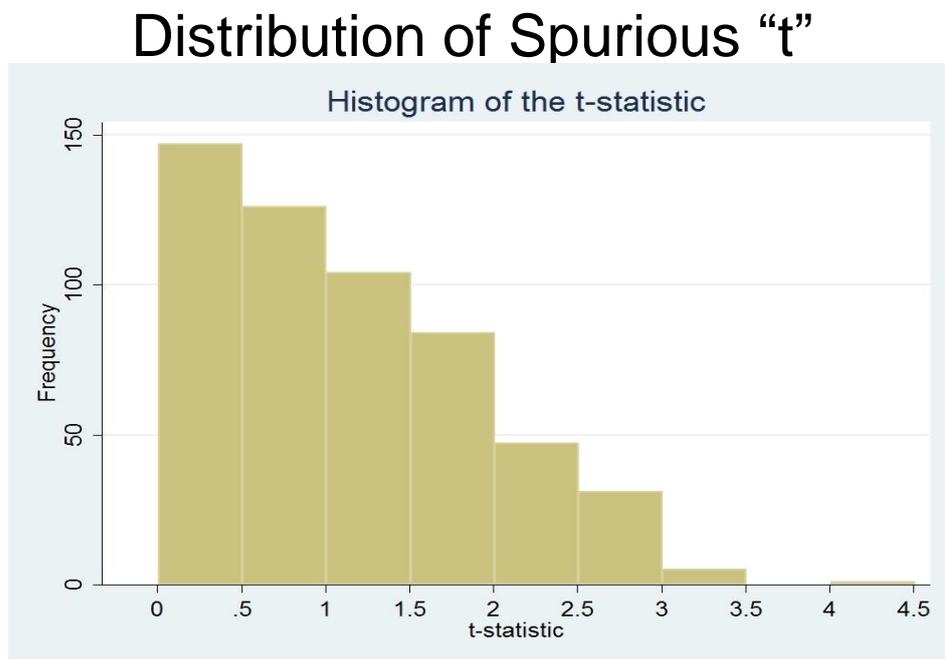


Figure 4

4.3 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 2.2. 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 regression⁷ model in which the dependent variable is the number of

⁷ Also known as Tobit regression. Censored regression assumes that the dependent variable is continuous if it is not zero.

tide gauges operating by country, which in many cases is zero. The independent variables include GDP per capita, population and coastline lengths of the country⁸. We use satellite altimetry data for 1993 – 2010 (Map 2) to measure SLR for all countries⁹ 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 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.

⁸ Data sources are provided in an appendix.

⁹ Satellite data do not cover Arctic and Antarctic coasts.

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 R²	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. What matters for our purposes is that the tide gauge locations used in Table 1 constitute a quasi-random sample.

4.3 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 reconstructions 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 the conservative method would have led to the conclusion that SLR was greater globally than in 2000. It also means that data reconstructions based on these veteran tide gauges would over-estimate SLR.

Selection Bias in the Global Diffusion of Tide Gauges

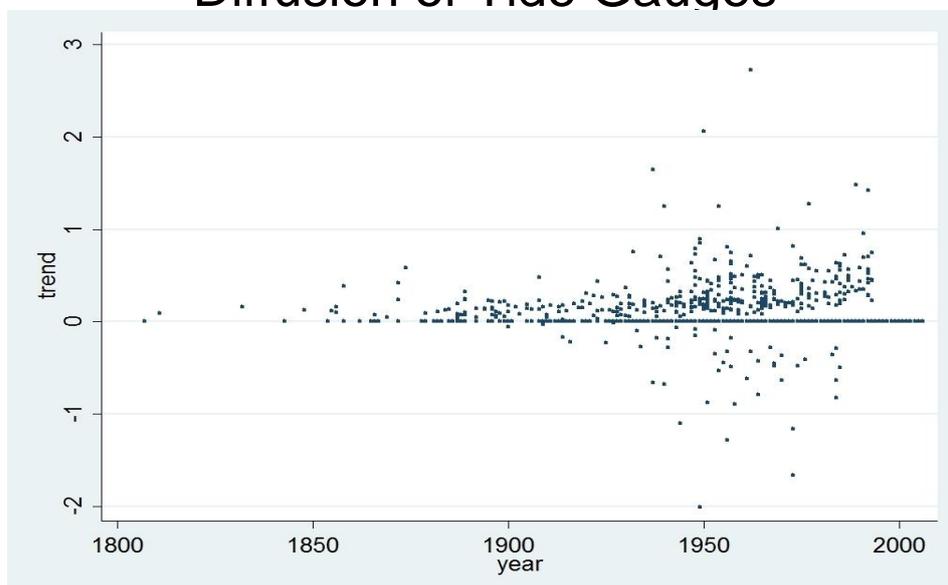


Figure 5

Figure 5 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. 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
(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.

4.4 Mapping SLR

In Map 3 we display the geographical distribution of the KPSS tide gauge classification presented in Table 1. Given that the data for each point may not always represent a continuous series, SLR is flagged as ‘positive’ (negative) if all the segments for that tide gauge are positive (negative) and as ‘conflicting’ if the KPSS results across the segments are not consistent. We also present the tide gauges by circles proportionate to the number of observations per gauge. This indicates the effect of time on the accumulation of observations per station.

The overwhelming picture is yellow (no trend) because according to KPSS sea levels are trending in only a minority of locations. We therefore concentrate on the regions in which sea levels, according to KPSS, are rising, indicated in red. There are only a few clusters of SLR in a predominantly trendless world. The main areas of SLR include the Baltic Sea, the Atlantic coast of the US, the Ring of Fire (Japan, SE China and Malaysia) and the Russian arctic. In contrast, no significant positive trends 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, negative trends (indicated in green) may be found in Alaska and the western coast of India.

It is notable that tide gauges with positive trends are co-located with tide gauges that are trend-free. The same applies to the minority of tide gauges that happen to have negative trends. 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.

5. Discussion

This paper uniquely uses the study of tide gauge location to make two contributions to the study of global SLR. 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 data reconstructions for locations that do not happen to have tide gauges. We refer to this as the "conservative methodology" because it avoids the use of data reconstructions.

Our estimates of global SLR obtained using the conservative methodology are considerably smaller than estimates obtained using data reconstructions. While we find that sea levels are rising in about a third of tide gauges, SLR is not a global phenomenon.

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 veteran tide gauges 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 reconstructionists give greater weight to veteran tide gauges, we suspect that this overweighting may induce positive bias in SLR.

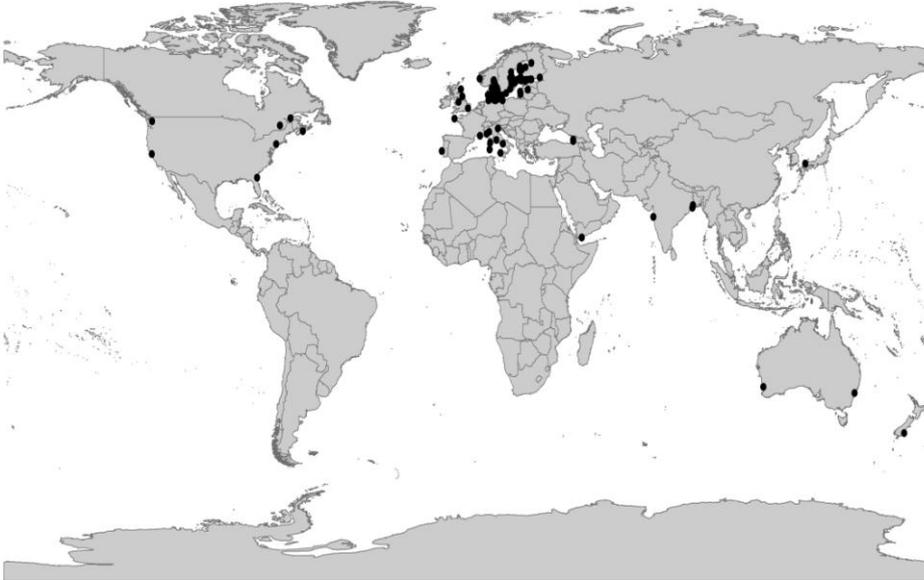
Scientists are impaled on the horns of a methodological dilemma. If they do not use data reconstructions, sample selection bias may be induced in estimates of global SLR. And if they resort to data reconstructions, measurement error might induce bias in estimates of global SLR. We resolve this dilemma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

The substantive contribution of the paper is concerned with recent SLR in different parts of the world. Consensus estimates of recent GMSL rise are about 2mm/year. Our estimate is 1mm/year. We suggest that the difference between the two estimates is induced by the widespread use of data reconstructions which inform the consensus estimates. There are two types of reconstruction. The first refers to reconstructed data for tide gauges in PSMSL prior to their year of installation. The second refers to locations where there are no tide gauges at all. Since the tide gauges currently in PSMSL are a quasi-random sample, our estimate of current GMSL rise is unbiased. If this is true, reconstruction bias is approximately 1mm/year.

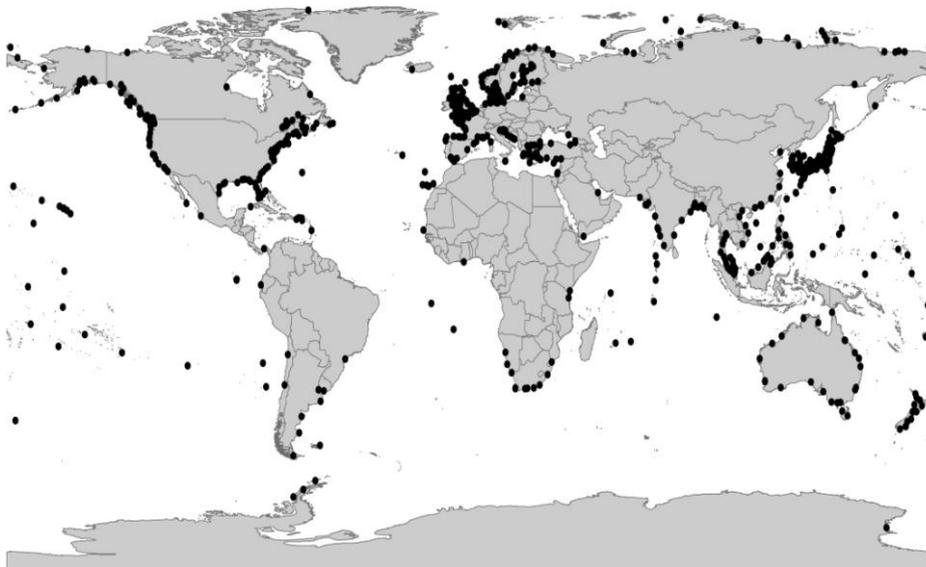
Sea level rise is regional rather than global and is concentrated in the southern Baltic, the Ring of Fire, and the Atlantic coast of the US. By contrast the north-west Pacific coast and north-east coast of India are characterized by sea level fall. In the minority of locations where sea levels are rising the mean increase is about 4 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.

Map 1: The Global Coverage of Tide Gauges

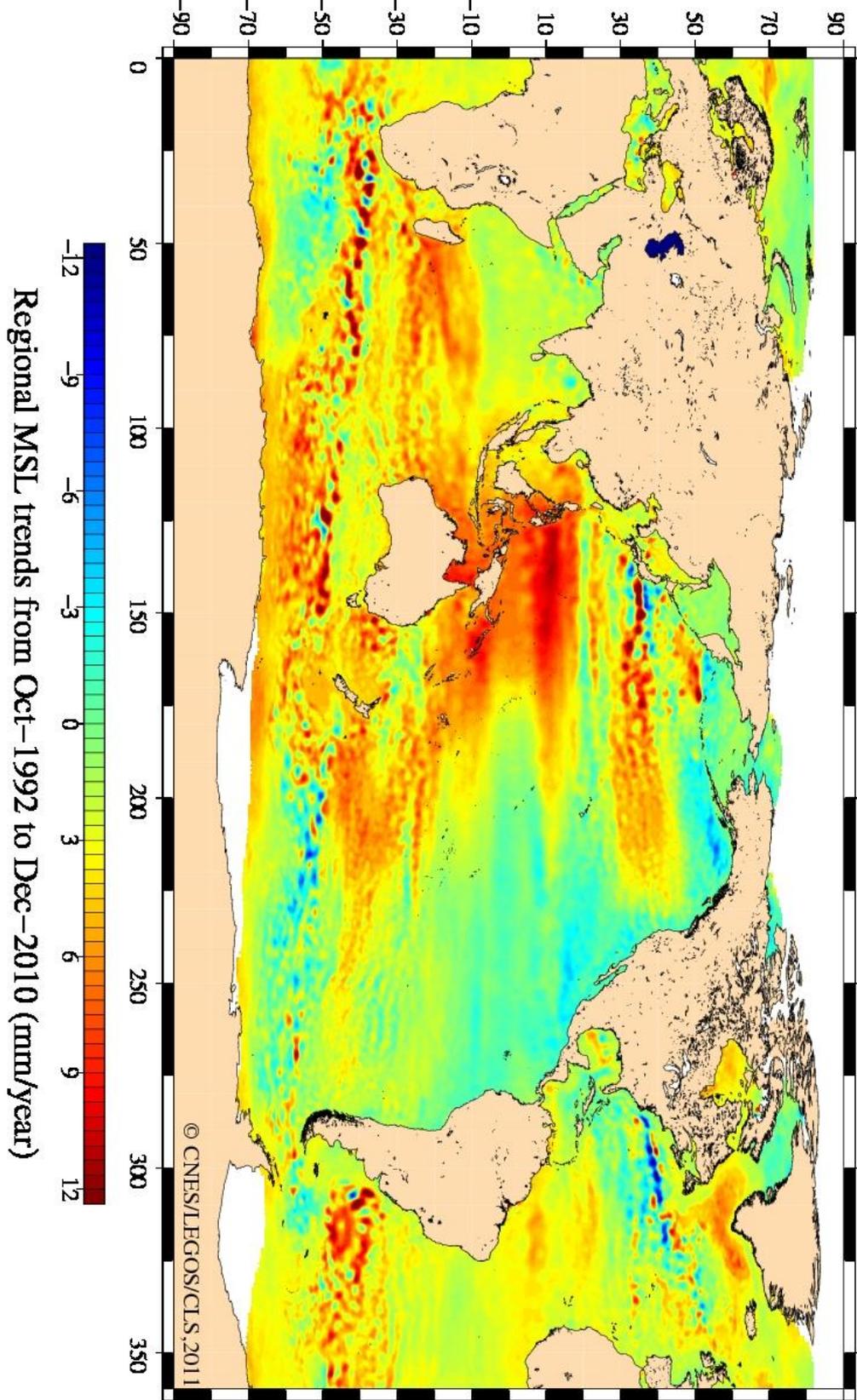
Tide Gauges in 1900



Tide Gauges 2000



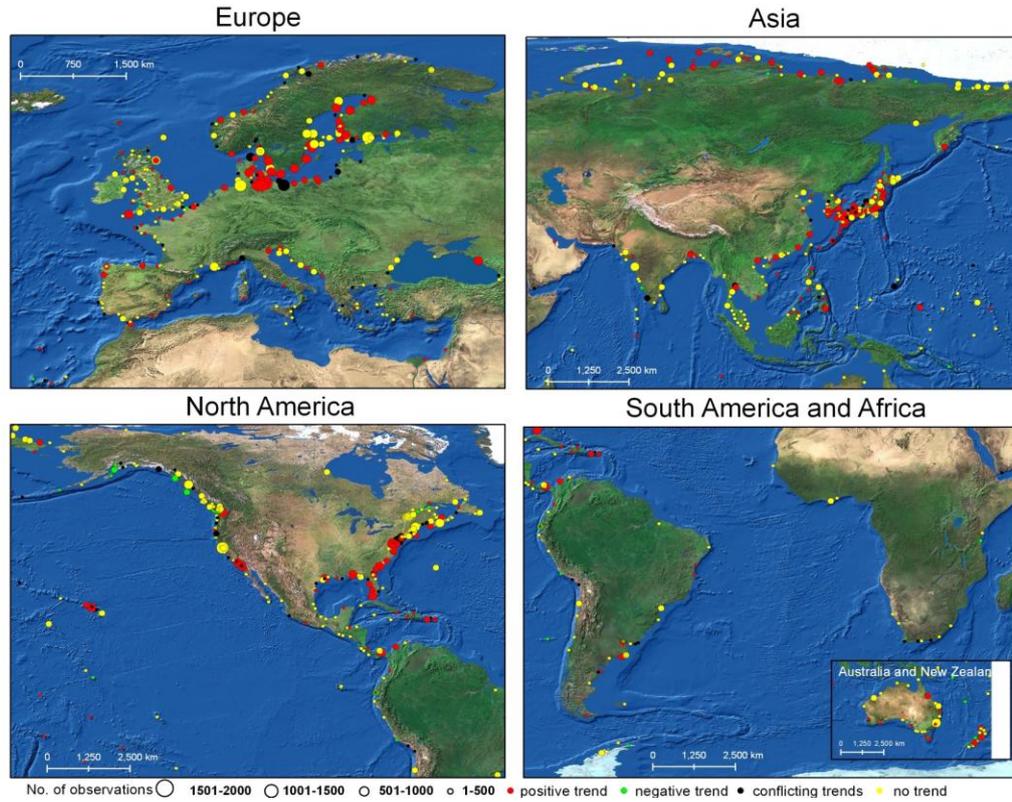
Map 2: Global SLR using satellite altimetry data for 1993 – 2010



Map 3: SLR trends by Tide Gauges; Results of KPSS tests for Stationarity by Location of Tide Gauge and Number of Observations

Tide Gauges – KPSS Classification

Weighted by sample size



References

- Bâki Iz H, and Ng H.M. 2005. Are Global Tide Gauge Data Stationary in Variance? *Marine Geodesy*, 28: 209–217,
- 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., White N.J, Aarup T., Wilson S., Woodworth P.L., Domingues C.M, Hunter J.R., and Lambeck K., 2008 Understanding global sea levels: past, present and future. *Sustainability Science* 3: 9-22.
- 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.
- Dickey, D.A. and W.A Fuller 1981 Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49: 1057-1079.
- 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
- Jeverjeva, S., J. C. Moore, A. Grinsted, and P. L. Woodworth 2008 Recent global sea level acceleration started over 200 years ago? *Geophysical Research Letters*, 35: L08715.
- 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.

- Merrifield, M. A., S. T. Merrifield, and G. T. Mitchum 2009 An anomalous recent acceleration of global sea level rise. *Journal of Climate*, 22: 5772-5781.
- 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
- Phillips, P.C.B and Perron P. 1988 Testing for a Unit Root in Time Series Regression, *Biometrika*, 75, 335–346
- Said, S. and D. Dickey 1984 Testing for unit roots in autoregressive-moving-average models with unknown order. *Biometrika*, 71: 599-607.
- Tai C.K 2001 Inferring the global mean sea level from a global tide gauge network, *Acta Oceanol. Sin.*, 30 (4), 102-106
- Woodworth, P. L., White N.J., Jevrejeva S., Holgate S.J., Church J.A. and Gehrels W.R. 2009 Evidence for the accelerations of sea level on multi decade and century time scales. *International Journal of Climatology* 29: 777-789.

Data Appendix

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 converts 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.

We deal with missing values in three ways depending on their number and timing: i) sporadic missing values, ii) 2 - 12 consecutive missing values (up to 12 months), iii) over 12 consecutive missing values. Single missing values were replaced by the mean of the periods before and after the gap. Gaps of between 1-12 consecutive missing values were replaced by interpolated values using the *ipolate* command in STATA which linearly interpolated missing values based on the entire given segment. Tide gauges with gaps of over 12 consecutive missing values were split into separate segments of continuous observations.

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.

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>.

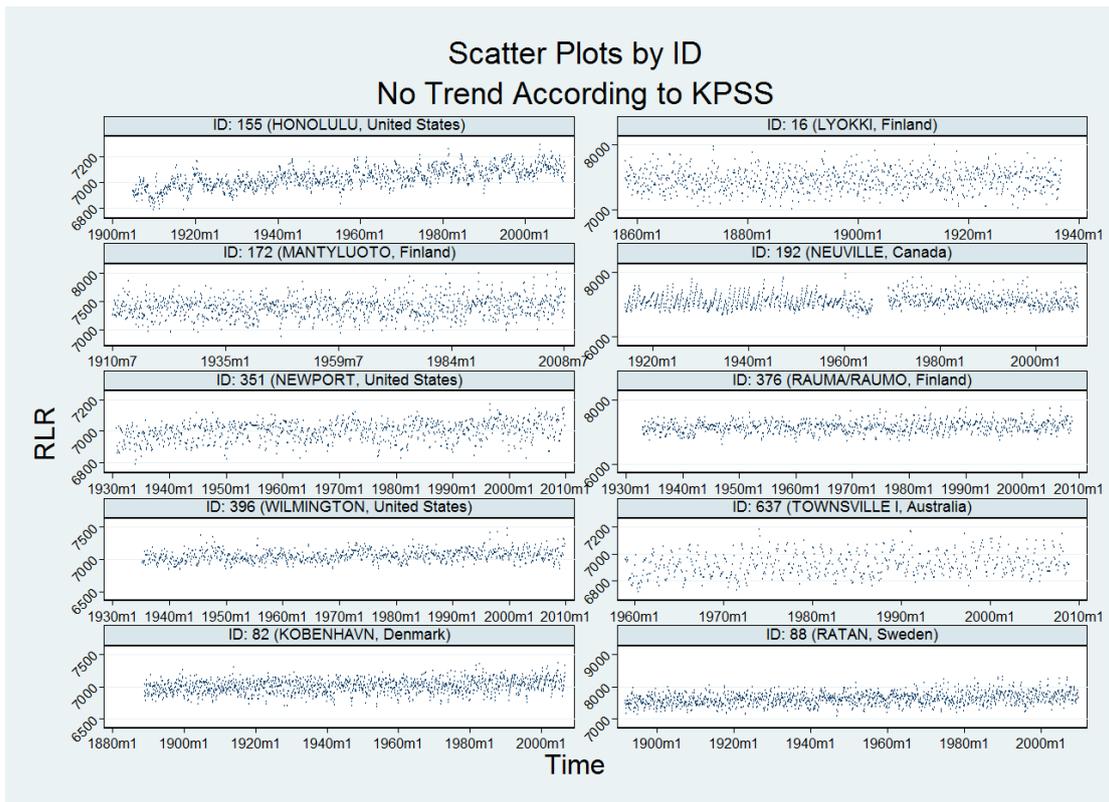


Figure A1

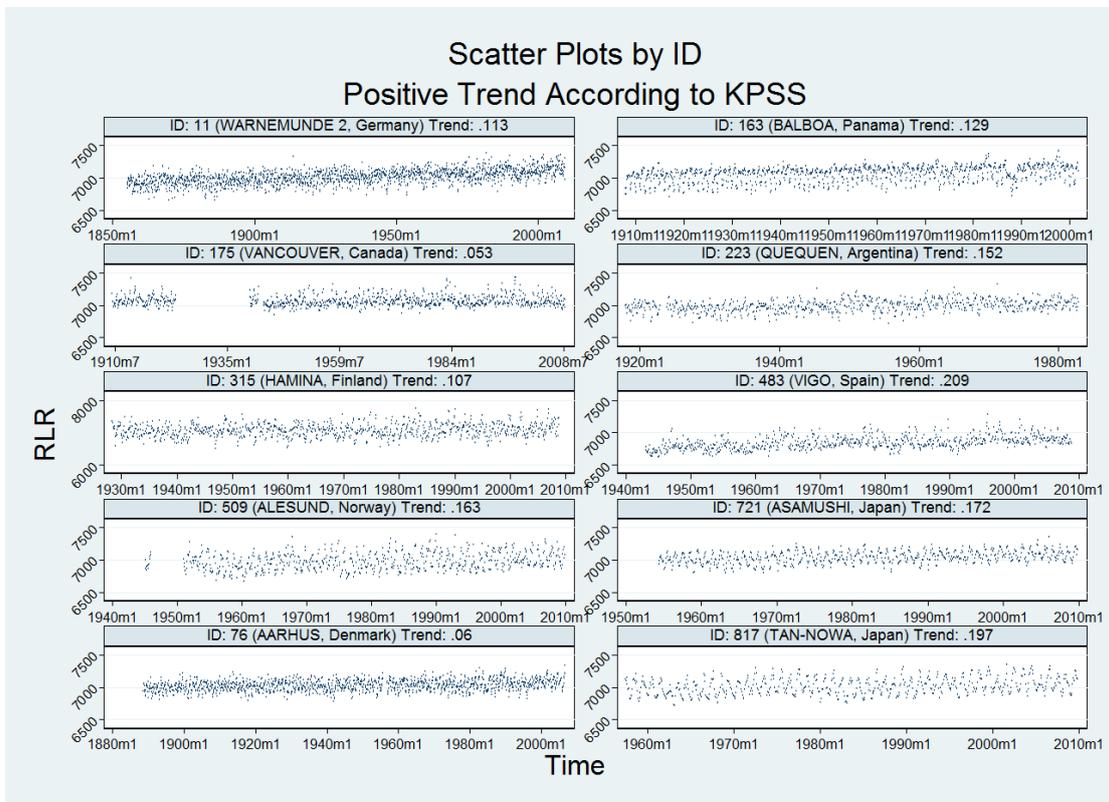


Figure A2

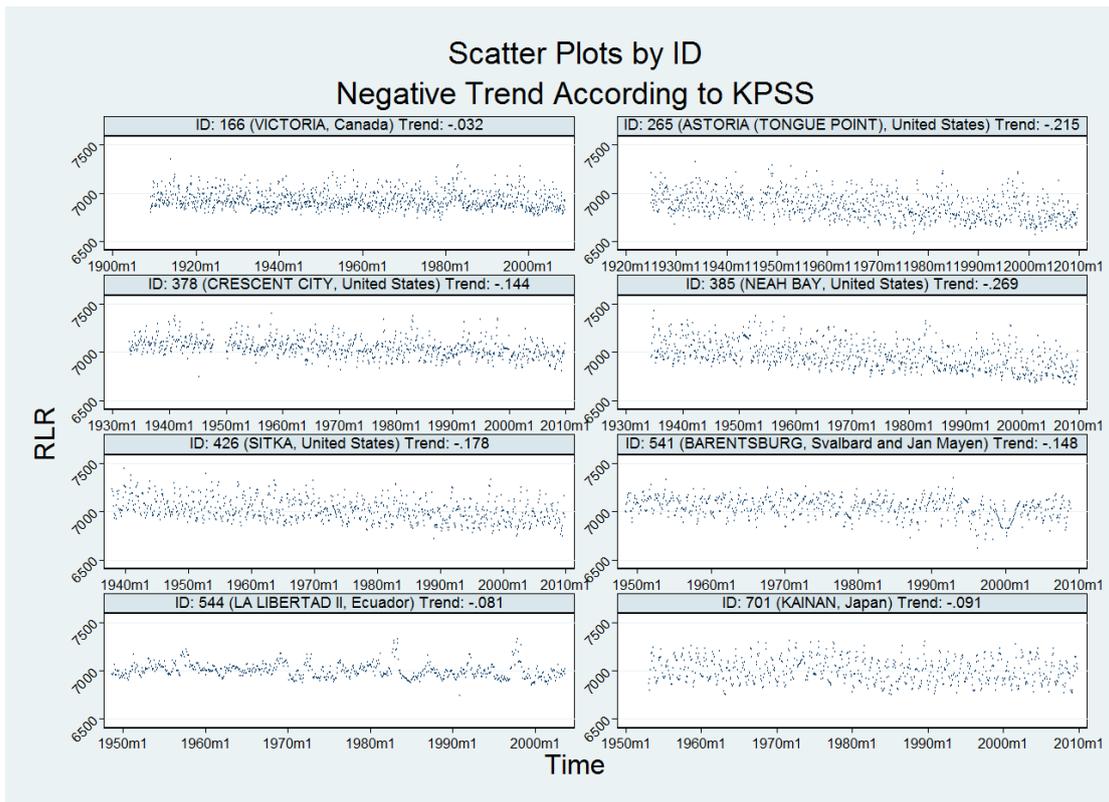


Figure A3

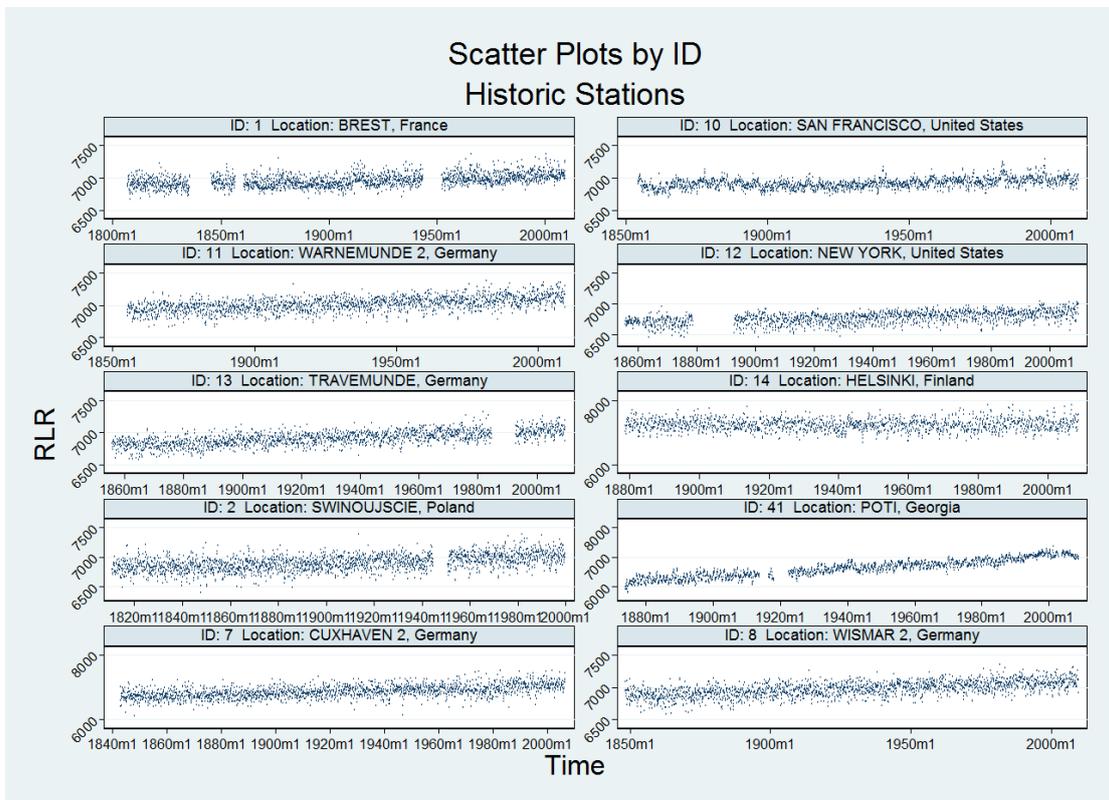


Figure A4