

REPORTS

Sea Level Rise During Past 40 Years Determined from Satellite and in Situ Observations

Cecile Cabanes, Anny Cazenave, Christian Le Provost

The 3.2 ± 0.2 millimeter per year global mean sea level rise observed by the Topex/Poseidon satellite over 1993–98 is fully explained by thermal expansion of the oceans. For the period 1955–96, sea level rise derived from tide gauge data agrees well with thermal expansion computed at the same locations. However, we find that subsampling the thermosteric sea level at usual tide gauge positions leads to a thermosteric sea level rise twice as large as the “true” global mean. As a possible consequence, the 20th century sea level rise estimated from tide gauge records may have been overestimated.

Coastal tide gauges have provided the main technique by which sea level change has been measured during the past century. For about a decade, sea level has been monitored worldwide by altimeter satellites, in particular by Topex/Poseidon, with global coverage, high spatio-temporal resolution, and direct tie to Earth’s center of mass (1). The global mean sea level rise observed by Topex/Poseidon amounts to 2.5 ± 0.2 mm/year between January 1993 and December 2000 (2). On time scales longer than 1 year, global mean sea level change results from two main causes: (i) volume change due to seawater density change in response to temperature and salinity variations (the two parameters having opposite effects on sea level) and (ii) mass change due to exchange of water with atmosphere and continents, including glaciers and ice sheets, through precipitation, evaporation, river runoff, and ice melting. The recent availability of long time series of global ocean temperatures down to 3000 m, for the period 1945 through 1998 (3), has made it possible to quantitatively estimate the thermal contribution to the sea level change observed during the 1990s.

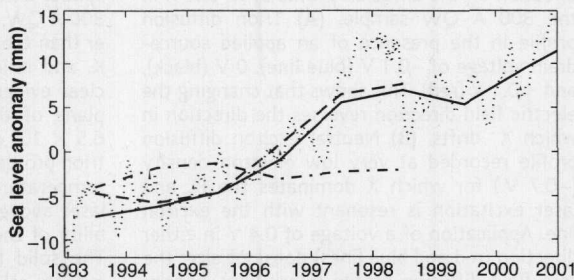
We used the yearly mean temperature data (available for the upper 500 m only) to compute the thermosteric sea level (4) and compared it with the yearly averaged sea level derived from Topex/Poseidon (Fig. 1). The thermosteric sea level rise for 1993–98 amounts to 3.1 ± 0.4 mm/year, in agreement with the 3.2 ± 0.2 mm/year rate measured by Topex/Poseidon over the same time span (5). The residual sea level (observed minus thermosteric sea level; also shown in Fig. 1) presents a small, not significant, trend of 0.2 ± 0.2 mm/year. Apart from uncertainties in observed and computed sea level rise, the

residual trend would reflect unknown contributions from three sources: deep (500 to 3000 m) thermosteric change, water mass addition to the oceans due to exchange with atmosphere and continents, and halosteric (i.e., due to salinity change) variations. The above results suggest that these components contribute negligibly (less than 5%) to the observed sea level rise. The agreement between the Topex/Poseidon-observed and thermosteric sea level trends (2, 4) for 1993–98 is striking, particularly in the tropics and Northern Hemisphere (Fig. 2). The thermosteric trend map reproduces well the Eastern Pacific sea level rise associated with the 1997–98 El Niño–Southern Oscillation (ENSO) event, as well as the Western Indian Ocean rise. Sea level patterns in the equatorial and Northern Atlantic also are well reproduced in shape and magnitude by the thermosteric map. Some discrepancy is observed in the southern oceans, where the positive trends observed by Topex/Poseidon (6) are larger than the thermosteric contribution, a likely consequence of sparse temperature coverage in remote southern regions. The quantitative comparison presented here shows that, for recent years, warming of the upper oceans almost fully accounts for the global mean sea level rise observed by Topex/Poseidon. Thus, other climatic contri-

butions to the 1993–98 observed sea level rise, due to water mass exchange with the atmosphere and with continental reservoirs as well as deep ocean (below 500 m) thermal effects and salinity contribution, may globally counterbalance each other.

The third assessment report of the Intergovernmental Panel on Climate Change (IPCC) (7) estimates the various factors that have contributed to the 20th century sea level rise. The largest contribution (0.7 mm/year sea level rise) arises from thermal expansion due to warming of the oceans that mainly occurred since the 1950s (8). Melting of continental glaciers produces 0.2 to 0.4 mm/year sea level rise (7). Estimated Greenland and Antarctica mass imbalance (accounting for a long-term readjustment since Last Glacial Maximum plus a climate-related response) contributes -0.2 to 0.6 mm/year (7). The least certain contribution is the change in terrestrial water storage that results partly from human activities, which is in the range of -1.1 to $+0.4$ mm/year with a median value of -0.35 mm/year (i.e., corresponding to sea level drop) (7). The sum of these contributions ranges from -0.8 to 2.2 mm/year, with a median value of 0.7 mm/year (7). Values for the 20th century sea level rise based on tide gauges records, published during the 1990s, are in the range 1 to 2 mm/year (7). The most recent global analyses (9, 10), which use the longest tide gauge records available (≥ 70 years), report a rate of rise closer to 2 mm/year: 1.71 ± 0.55 mm/year (9) and 1.84 ± 0.35 mm/year after correcting for postglacial rebound (10). The third IPCC report (7) adopts a best estimate of 1.5 ± 0.5 mm/year for the observed 20th century sea level rise and notes that the sum of climate-related components is low compared with the observational estimates. In effect, these observed values (1.5 mm/year or 1.8 mm/year) are more than twice as large as the revised estimate of total climate contributions, although there is complete overlap between the range of the sum of contributions (7) and the observed range. It would appear that either the climate-related processes causing sea level rise have been underestimated or the sea level rise observed with tide gauges is biased toward values too high. The latter possibility may arise from the fact that tide gauges are located at continents or island coastlines and

Fig. 1. Global mean sea level curves. Dotted curve, observed at 10-day interval by Topex/Poseidon for 1993–2000. Solid curve, yearly averaged sea level from Topex/Poseidon. Dashed curve, thermosteric component computed from global temperature data (3) down to 500-m depths for 1993–98. Dashed-dotted curve, residual (Topex/Poseidon minus thermosteric) sea level.



Laboratoire d'Etudes en Géophysique et Océanographique Spatiales, Centre Nationale d'Etudes Spatiales, 18 Avenue Edouard Belin, Toulouse 31400, France.

hence do not globally sample the spatial variations of the sea level change. The recent availability of global sea temperature data (3) can help address the second possibility.

We computed the thermosteric sea level,

We computed the thermosteric sea level,

for the period 1955–96, using the 5-year mean temperature data over the depth range 0 to 3000 m (3). The 60°S to 60°N average shows a mean thermosteric trend of 0.50 ± 0.05 mm/year (Fig. 3). To test the agreement

0.05 mm/year (Fig. 3). To test the agreement

between this average and the value that would be calculated by using only data from tide gauge positions, we computed a pseudo “global mean” steric sea level time series by subsampling the global thermosteric sea level grid at locations close to tide gauge sites (11), using 25 of the 27 stations selected by Douglas (9). The pseudo global mean thermosteric sea level time series, superimposed to the “true” global mean in Fig. 3, is 1.4 ± 0.10 mm/year, a value more than two times as large as the true global mean trend (0.5 ± 0.05 mm/year). Twenty-three out of the 25 sites are located in positive trend regions (Fig. 4). Considering the substantial regional variations in thermosteric trends, it appears clear that these sites fail to correctly sample the global variation and that averaging thermosteric sea level at these sites is not representative of the global mean.

We further checked whether the computed thermosteric sea level rise correctly reproduces the tide gauge–derived sea level rise. For that purpose, we considered tide gauge records from the Permanent Service for Mean Sea Level (PSMSL) (12) at the 25 sites. The “average” tide gauge–derived sea level curve over 1955–96, after performing regional grouping (13), is shown in Fig. 3. The observed (i.e., tide gauge–derived) mean sea level rise for 1955–96 is 1.6 ± 0.15 mm/year, a value that agrees well with the pseudo global mean thermosteric rise (1.4 ± 0.10 mm/year). The tide gauge–derived sea level curve displays decadal oscillations that are smaller in the thermosteric sea level curve. Inspection of individual tide gauge records indicates that stations of the northeast U.S. coast are largely responsible for the observed decadal variability, which originates from North Atlantic wind forcing (14). In our computation, we did not account for the halosteric component because global gridded salinity data are not yet available. Antonov *et al.* (15) showed that the halosteric contribution is quite substantial in the subpolar part of the North Atlantic, especially in the Labrador Sea, where it nearly counteracts the thermosteric contribution. However, in terms of global mean, the halosteric sea level rise has not exceeded 0.05 mm/year over the past 40 years (15). We checked whether neglect of the haline component when computing the steric sea level at the tide gauge sites would change our results. At each of the 25 stations, we computed difference time series of the tide gauge–derived minus thermosteric sea level and fitted a linear trend to these difference time series. The mean difference trend is 0.25 ± 0.14 mm/year, a value that represents an upper bound of the neglected halosteric and other climatic contributions.

Our study has demonstrated that the global estimate of the thermal expansion component is substantially smaller than the value obtained if the same field is subsampled at the tide gauge positions used to compute the

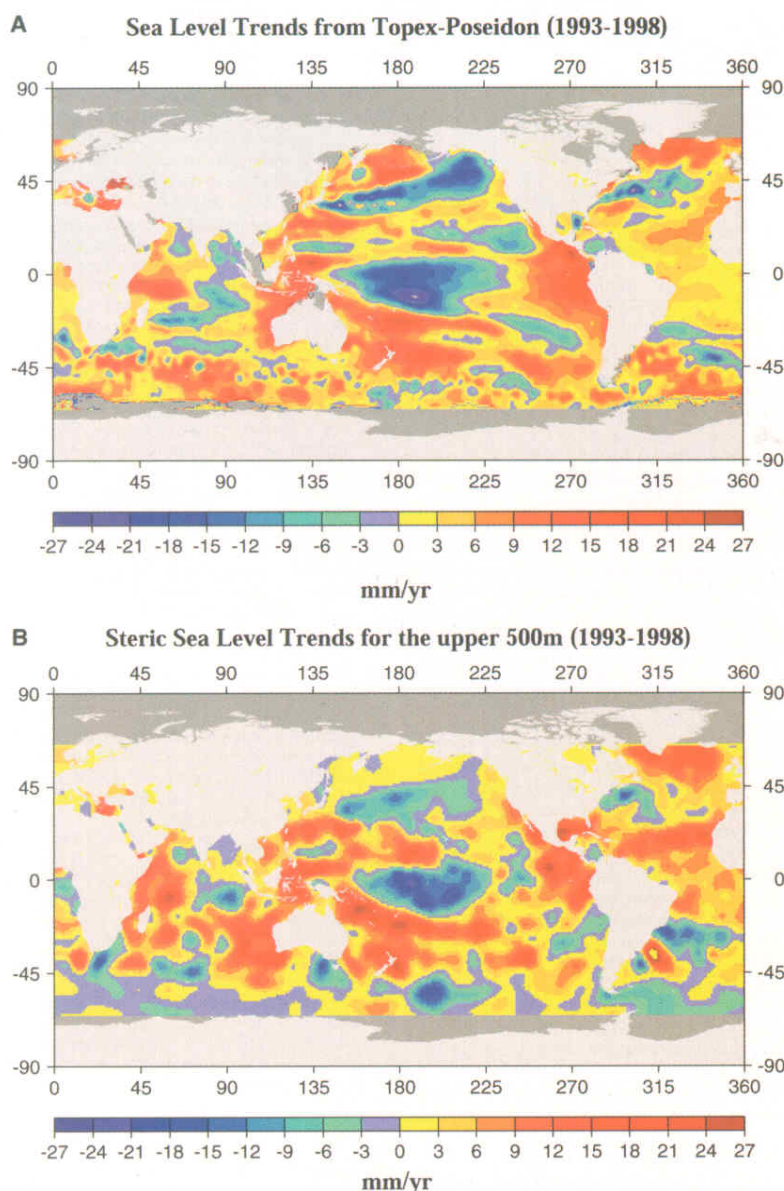
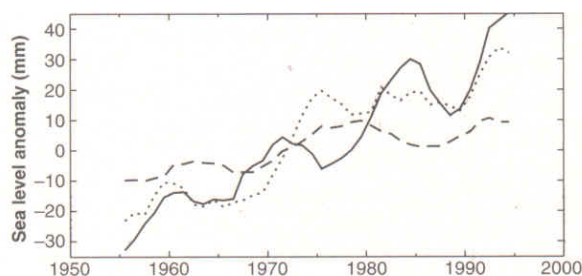


Fig. 2. Map of the geographical distribution of sea level trends over 1993–98 computed from Topex/Poseidon altimetry (A) and from the thermosteric sea level data (B).

Fig. 3. Mean sea level curves for 1955–96. Dashed curve, global mean thermosteric component computed with data from (3) down to 3000-m depths. Dotted curve, pseudo global mean thermosteric sea level computed by subsampling the global data set at the 25 tide gauge sites. Solid curve, observed sea level curve based on the PSMSL records at the 25 tide gauge sites.



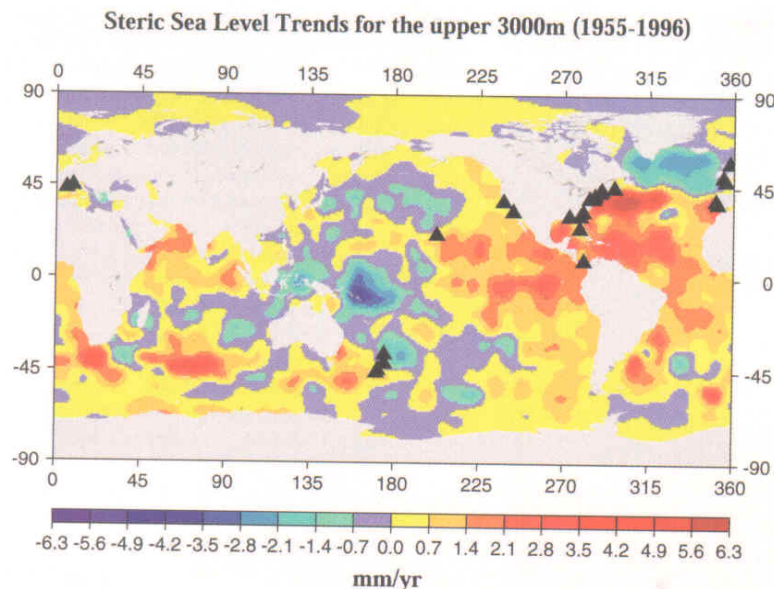


Fig. 4. Map of the geographical distribution of thermosteric sea level trends for 1955–96 computed with temperature data from (3) down to 3000-m depths. Black triangles show the locations of the 25 tide gauges.

20th century global mean sea level rise. It is generally assumed that spatial variation of sea level rise is caused by nonuniformity in thermal expansion, other contributions leading rapidly to uniform sea level change. Thus, the reported difference may reflect an overestimate of the sea level rise for the past decades, caused by the uneven distribution of the tide gauges and limited geographical sampling available from historical records. Even though the global tide gauge network has been considerably extended during the 1990s (16), recent sea level rise estimates based on the tide gauges still substantially depart from the global mean measured by Topex/Poseidon (17). Because of temperature data availability, we limited our analysis to the second half of the 20th century, but it should be noted that the mean sea level rise computed with this 40-year-long tide gauge record agrees well with values based on longer records (9, 10). Thus, our conclusion that the tide gauge-derived sea level rise for the past few decades has been overestimated possibly holds for the whole 20th century. This would reconcile observed sea level rise and estimate of climate-related contributions (on the order of 0.7 mm/year) as reported by the third IPCC assessment report (7).

References and Notes

1. L. L. Fu, A. Cazenave Eds., *Altimetry and Earth Science, A Handbook of Techniques and Applications*, vol. 69 of *International Geophysics Series* (Academic Press, London, 2001).
2. We analyzed Topex-Poseidon altimetry data (merged Geophysical Data Records available from the AVISO/Altimetry data center, Centre National d'Etudes Spatiales, Toulouse, France) from January 1993 through December 2000. Geophysical and instrumental corrections

applied to the data are the most recent version provided by AVISO. In addition to the onboard oscillator drift, the 1 mm/year radiometer drift is taken into account up to December 1996. A 7-mm bias is applied to the data beyond February 1999 to account for the vertical offset after the switch to the redundant Topex altimeter (18). Otherwise, the data processing is identical to that used in (6) and (19). For sea level trends computations, two different processing were applied depending on the required representation (global mean sea level trend or geographical distribution of the regional trends). For the former, we averaged the along-track sea level data over each 10-day orbital cycle, applying an equi-area weighting. Only data between 60°S and 60°N are considered because of noisy data at higher latitudes. The mean trend is based on a simple least squares fit to the sea level time series. The associated uncertainty is the formal error based on the scatter of the fit. To estimate regional sea level trends, we interpolated along-track sea level data onto regular 1° by 1° grids. Finally, we low-pass filtered the raw time series at each grid point to remove short-period (≤ 1 year) signal and fitted a linear trend to the filtered time series through a least squares inversion.

3. S. Levitus, C. Stephens, J. I. Antonov, T. P. Boyer, *NOAA Atlas NESDIS 40* (U.S. Government Printing Office, Washington, DC, 2000).
4. To compute gridded time series of the thermosteric (i.e., due to temperature change) sea level, we first converted the gridded temperature anomalies in terms of density anomalies at each standard level using the classical expression for the equation of state of the ocean (20). Then we vertically integrated the density anomalies to compute the thermosteric sea level anomaly. Regional trends were computed through a least squares fit of the thermosteric sea level time series at each grid mesh.
5. Using Topex/Poseidon data over the years 1993 through 1998, Nerem and Mitchum (21) found a rate of sea level rise of 2.5 ± 1.3 mm/year. Their rate uncertainty is based not only on the formal error but also on instrumental drift errors.
6. C. Cabanes, A. Cazenave, C. Le Provost, *Geophys. Res. Lett.* **28**, 9 (2001).
7. J. Church et al., in *Climate Change 2001, The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental*

Panel on Climate Change, J. T. Houghton et al., Eds. (Cambridge Univ. Press, Cambridge, 2001), pp. 639–694.

8. S. Levitus, J. I. Antonov, T. P. Boyer, C. Stephens, *Science* **287**, 2225 (2000).
9. B. C. Douglas, in *Sea Level Rise, History and Consequences*, B. C. Douglas, M. S. Kearney, S. P. Leatherman, Eds., vol. 75 of *International Geophysics Series* (Academic Press, London, 2001), pp. 37–64.
10. W. R. Peltier, in *Sea Level Rise, History and Consequences*, B. C. Douglas, M. S. Kearney, S. P. Leatherman, Eds., vol. 75 of *International Geophysics Series* (Academic Press, London, 2001), pp. 65–95.
11. Among the 27 sites selected by Douglas (9), we did not consider Trieste (North Adriatic Sea) and Buenos Aires (South Atlantic) because of the poor coverage of temperature data at these sites. We performed a series of tests to derive the thermosteric sea level time series at the tide gauge location. We successively considered the three, five, and ten closest grid points to produce spatial averages of the thermosteric sea level, after deleting values larger than three standard deviations. We first averaged the thermosteric sea level data by regions (nine in total) to avoid excessive regional weight when computing the pseudo global mean. When averaging, we applied a $\cos(\text{latitude})$ weighting. The corresponding thermosteric sea level time series were found to be very similar, indicating that the averaging procedure is not crucial. To compute the tide gauge-based sea level time series, we also made a series of tests consisting of selecting different locations. We considered the 17 tide gauge sites from (22). We also considered a sample of 32 sites that included the 25 sites of the nominal case plus seven island sites in the Western Pacific for which long records of good quality are available. The "observed" sea level curves displayed similar behavior. Computed trends (for the thermosteric and observed sea level curves) did not differ by more than 0.2 mm/year.
12. See www.pol.ac.uk/psmsl.
13. All records were corrected for the inverted barometer response of sea level to atmospheric forcing with pressure data from the National Centers for Environmental Prediction 40-year reanalysis (23). The equivalent sea level trend associated with this correction is less than 0.2 mm/year. Land motion due to glacial isostatic adjustment was also corrected for with the ICE-4G VM2 model (10).
14. B. G. Hong, W. Sturges, A. J. Clarke, *J. Phys. Oceanogr.* **30**, 2088 (2000).
15. J. I. Antonov, S. Levitus, T. P. Boyer, in preparation.
16. GLOSS, *Global Sea Level Observing System Implementation Plan 1997* (Technical Series 50, Intergovernmental Oceanographic Commission, Pasadena, CA, 1998).
17. We used PSMSL data from 209 tide gauges of the GLOSS network (excluding the Mediterranean and Black seas) to compute a pseudo global mean sea level rate of change for 1993–98. Inverted barometer and postglacial rebound corrections were applied. The estimated rate is 5 ± 1 mm/year, compared with the 3.2 ± 0.2 mm/year rate based on Topex/Poseidon. A majority of these tide gauges are located in positive trend regions, which causes this higher rate estimate. Anyway, inspection of Fig. 2A shows that the present configuration of continental coastlines and ocean islands simply fails to correctly sample spatial variations of the sea level change.
18. G. T. Mitchum, *Mar. Geod.* **23**, 145 (2000).
19. A. Cazenave, F. Remy, K. Dominh, H. Douville, *Geophys. Res. Lett.* **27**, 3755 (2000).
20. A. E. Gill, *Atmosphere-Ocean Dynamics*, vol. 30 of *International Geophysics Series* (Academic Press, London, 1982).
21. R. S. Nerem, G. T. Mitchum, in (1), pp. 329–349.
22. B. C. Douglas, *Surv. Geophys.* **18**, 278 (1997).
23. E. Kalnay et al., *Bull. Am. Meteorol. Soc.* **77**, 437 (1996).
24. We thank S. Levitus and his co-workers for having made available their global gridded temperature data set. We also thank J. Antonov for useful discussions on the role of salinity. This work was supported by CNES, CNRS, and the French ministry of research.

18 June 2001; accepted 13 September 2001

Diversity and Productivity in a Long-Term Grassland Experiment

David Tilman,^{1*} Peter B. Reich,² Johannes Knops,³ David Wedin,⁴ Troy Mielke,¹ Clarence Lehman¹

Plant diversity and niche complementarity had progressively stronger effects on ecosystem functioning during a 7-year experiment, with 16-species plots attaining 2.7 times greater biomass than monocultures. Diversity effects were neither transients nor explained solely by a few productive or unviable species. Rather, many higher-diversity plots outperformed the best monoculture. These results help resolve debate over biodiversity and ecosystem functioning, show effects at higher than expected diversity levels, and demonstrate, for these ecosystems, that even the best-chosen monocultures cannot achieve greater productivity or carbon stores than higher-diversity sites.

Recent demonstrations that greater plant diversity can lead to greater productivity (1–5) have generated considerable debate (6–14). This has been fueled by uncertainty about which of many alternative hypotheses (6, 9, 15–20) is operating in nature and about the number of species required to maintain ecosystem functioning (6, 7, 12, 14). We report results of a long-term experiment that allows tests of these alternative hypotheses.

It has been hypothesized that productivity may be greater at higher diversity because of “niche complementarity” among particular combinations of species and the greater chance of occurrence of such combinations at higher diversity (15–20). Niche complementarity, which results from interspecific differences in resource requirements and in spatial and temporal resource and habitat use, or from positive interactions (21), is predicted to allow stable multispecies coexistence and sustainably greater productivity at higher diversity (17, 18). Alternatively, it has been hypothesized that reported diversity effects might be short-lived transients caused solely by the presence of some species with high growth rates (6); be experimental artifacts resulting solely from species pools containing some low-viability species (6); or result from the most productive species being the best competitor (6, 9, 17). These “sampling effects” all result from the greater chance of any given species being present at higher diversity and from dynamics that cause a single species to dominate and determine

ecosystem functioning (6, 9, 17, 19).

Sampling and complementarity have different signatures (6, 9, 12, 17–20). Sampling effects limit the maximal productivity of higher-diversity plots to that of the best monoculture, giving an upper bound of variation in community performance that is independent of diversity. With niche complementarity, the upper bound increases with diversity because no monoculture is as productive as some combinations of two species and no combination of N species is as productive as some combinations of $N + 1$ species.

In a 7-year experiment [(4), supplement A (22)], we controlled one component of diversity, the number of plant species, in 168 plots, each 9 m by 9 m. We seeded the plots, in May 1994, to have 1, 2, 4, 8, or 16 species, with 39, 35, 29, 30, and 35 replicates, respectively. The species composition of each plot was chosen by random draw from a pool of 18 grassland perennials that included four C4 (warm-season) grasses, four C3 (cool-season) grasses, four legumes, four nonlegume forbs, and two woody species. All species occurred in monoculture, and all but three were in at least two monoculture plots, allowing comparison of responses of each species in monoculture to higher-diversity combinations of these same species. We do not use 76 additional plots that had functional group compositions drawn from an augmented species pool or 46 plots planted to 32 species (4) because the additional species were not grown in monoculture, and combining results from different species pools could introduce bias. We focus analyses on species number, because it was directly controlled, and functional group composition (23), because of its hypothesized importance. Other measures of diversity, including number of species per functional group and the presence or absence of species or functional groups, are highly correlated with species number and show similar responses.

In our grasslands, plant aboveground living biomass, because it is all produced within a growing season, is an index of primary productivity. In contrast, total biomass (aboveground plus belowground plant biomass) measures carbon accumulated in living tissues. Both aboveground and total biomass increased highly significantly with species number each year (Fig. 1, A and B, and Table 1), and functional group composition explained a highly significant amount of the residual variation (Table 1). Moreover, when the effect of each variable was determined after controlling for effects of the other (type III regressions), effects of functional group composition predominated in the early years [as for (24, 25)], but species number had highly significant positive effects on both aboveground and total biomass by 1999 and 2000, showing the simultaneous importance of species number and functional group composition in the long-term (Table 1).

The initial saturating dependence of aboveground and total biomass on species number (Fig. 1, A and B) became, by 2000, a linear increase for species number ≥ 2 . In 2000, 16-species plots had 22% greater aboveground biomass total and 27% greater total biomass than 8-species plots (both differences significant; t tests: $P = 0.018$, $P = 0.002$, respectively). The dependence of biomass on species number and functional group composition became progressively stronger, explaining about one-third of variance in 1997 and two-thirds in 2000 (Table 1). This strengthening of the effect of diversity and the increasingly steep and linear trends (Fig. 1, A and B) fail to support the hypothesis (6) that diversity effects were short-lived transients. Comparable and significant ($P < 0.01$) dependences of total and aboveground biomass on diversity and composition were observed when analyses used the actual number of planted species observed in each plot (Fig. 1C) or the Shannon diversity index [supplement B (22)].

We tested the low-viability sampling hypothesis by identifying the five species that attained least total biomass in monoculture in 2000 and excluding from analysis plots containing any combinations of just these species. Total biomass was still significantly dependent on species number and functional group composition in the remaining 131 plots [general linear model (GLM) type III regression: $F_{\text{OVERALL}} = 8.39$, $P < 0.001$; $F_{\text{DIV}} = 11.6$, $F_{\text{COMP}} = 4.36$, $P < 0.001$ for each]. Similar results occurred when we excluded from analysis of aboveground biomass plots containing any combinations of the five species with least aboveground biomass in monoculture (GLM type III regression: $F_{\text{OVERALL}} = 7.08$, $P < 0.001$; $F_{\text{DIV}} = 10.6$, $P = 0.0014$; $F_{\text{COMP}} = 3.84$, $P < 0.001$). In another analysis, we excluded the 30 plots with the lowest total biomass in 2000 (total biomass < 400 g

¹Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, MN 55108, USA. ²Department of Forest Resources, University of Minnesota, St. Paul, MN 55108, USA. ³School of Biological Sciences, University of Nebraska, Lincoln, NE 68588, USA. ⁴School of Natural Resource Sciences, University of Nebraska, Lincoln, NE 68583, USA.

*To whom correspondence should be addressed. E-mail: tilman@umn.edu

m^{-2}). Species number and composition still had highly significant effects on total biomass in the remaining plots (GLM type III regression: $F_{\text{OVERALL}} = 5.05$, $P < 0.001$; $F_{\text{DIV}} = 12.3$, $P < 0.001$; $F_{\text{COMP}} = 2.82$, $P < 0.001$). Similarly, when we excluded the 31 plots with lowest aboveground biomass ($< 100 \text{ g m}^{-2}$) in 2000, effects of composition and species number remained highly significant (GLM type III regression: $F_{\text{OVERALL}} = 3.77$, $P < 0.001$; $F_{\text{DIV}} = 10.5$, $P < 0.0016$; $F_{\text{COMP}} = 2.28$, $P = 0.0015$). Similar results occurred with lower and higher cutoffs (including 50% higher) for aboveground and total biomass. In total, the dependence of biomass on species number and composition was not explained solely by sampling effects for a species pool containing some poorly performing species.

We tested the sampling hypothesis that the most productive species determined the effects of diversity (6, 9, 17) by retaining in analyses of year 2000 results only plots containing at least one of the nine species with the highest monoculture total biomass in 2000. Total biomass remained significantly dependent on species number and functional group composition in these 145 plots, and in the subset of 95 plots that contained at least two of these nine species [type III regressions; supplement C (22)]. Similar results occurred for aboveground biomass in 2000 [supplement C (22)]. These analyses fail to support the sampling hypothesis. Another test comes from examining performance of higher-diversity plots relative to the best monoculture (Fig. 2). In 1999 and 2000, many higher-diversity plots had greater aboveground and total biomass than the single best-performing monoculture (Fig. 2). The percentage of such plots was an increasing function of diversity, on average for 1999 and 2000, with about half of the 16-species plots having greater aboveground or total biomass than the best monocultures (Fig. 1D). The strength and repeatability of this increasing upper bound in both aboveground and total biomass support the importance of niche effects and refute the hypothesis that sampling effects were the sole explanation for the long-term effects of diversity. The coexistence of most species, with about 12 planted species per 2 m^2 persisting in each 16-species plot (Fig. 1C), further supports niche complementarity. However, in earlier years, such as 1997, few high-diversity plots had greater biomass than top monocultures, and the percentage was independent of species number (Fig. 1D), which is consistent with the sampling hypothesis (6, 7, 12).

The increasing importance of complementarity and the increasingly linear effects of species number raise another question. Did complementarity occur among most species—i.e., did most species contribute to increasing community biomass—or is there a smaller set of species with complementary

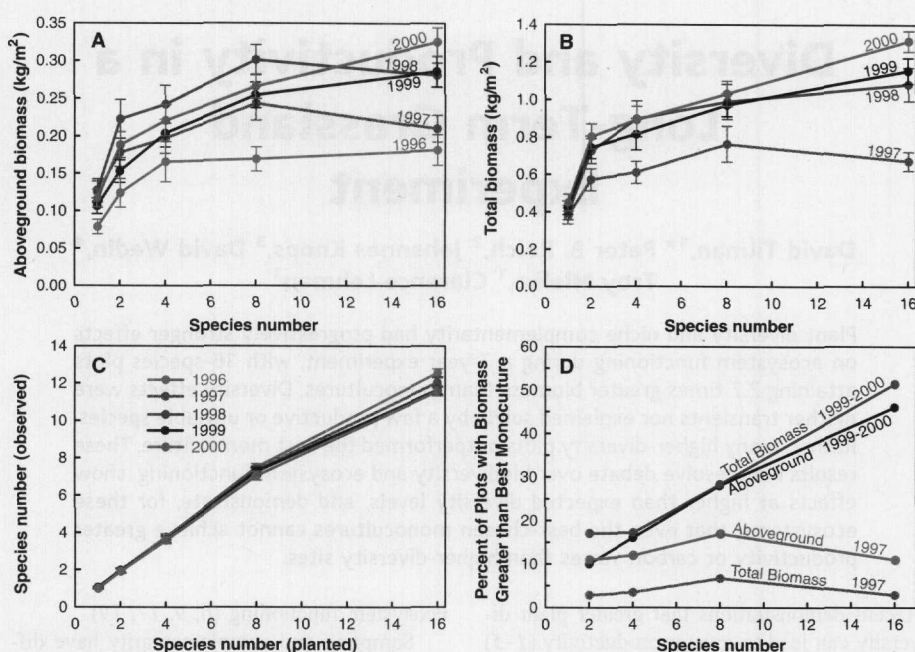


Fig. 1. The dependence of (A) plant aboveground biomass and (B) total biomass (aboveground plus belowground living plant mass) on the number of planted species. Data are shown as the mean \pm SE. (C) The relation between the number of species planted in a plot and the actual number (mean \pm SE) of planted species visually observed in a 2 m^{-2} area of each plot. (D) The percentage of all plots of a given planted diversity level, on average for 1999 and 2000 combined, or on average for 1997, that had greater biomass than the single monoculture plot with the greatest biomass.

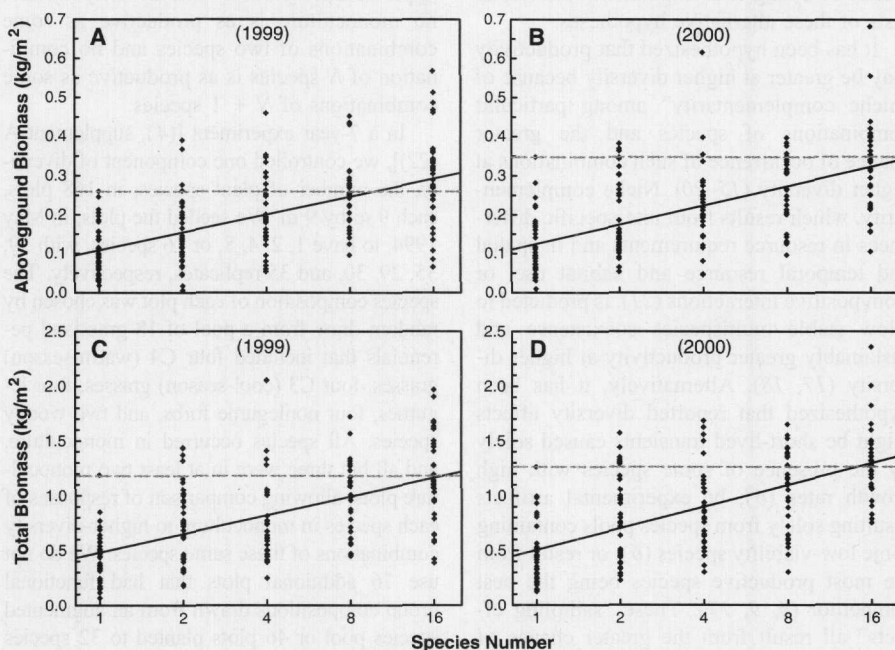


Fig. 2. The dependence of aboveground (A and B) and of total (C and D) biomass of each plot on planted species number for 1999 and 2000. The broken line shows the biomass of the top monoculture for a given year. The solid line is a regression of biomass on the logarithm of species number. Logarithm of species number was used in the figure because it gave slightly better fits, but was not used in Table 1 because it often gave slightly lower R^2 values than species number.

interactions, with this set being increasingly likely to co-occur at higher diversity (19)?

We used analysis of variance (ANOVA) to determine the simultaneous effects of the presence or absence of each species (entered as main effects) on aboveground or total biomass

(one test per year, with 4 years for total biomass, 5 years for aboveground biomass). Three or four species had significant ($P < 0.05$) positive effects on aboveground or total biomass in most years. Among legumes, *Lupinus perennis* had significant effects in all nine tests, *Lespe-*

Table 1. Analyses, using general linear models, of the effects of number of planted species (continuous variable; entered first using type I SS) and of functional group composition (categorical variable; entered second) on total biomass and on aboveground biomass, showing results for each year

Variable analyzed	Year	Overall			Species number (entered first)		Functional group comp. (entered second)		Species number (entered second)	
		R ²	F value	P	F value	P	F value	P	F value	P
Total biomass	1997	0.32	2.26	0.001	9.80	0.002	2.06	0.004	3.88	0.051
Total biomass	1998	0.47	4.37	<0.001	43.8	<0.001	2.91	<0.001	2.38	0.13
Total biomass	1999	0.60	7.31	<0.001	94.2	<0.001	4.09	<0.001	7.11	0.009
Total biomass	2000	0.68	10.5	<0.001	152.	<0.001	5.27	<0.001	12.3	<0.001
Aboveground biomass	1996	0.41	3.28	<0.001	15.5	<0.001	2.83	<0.001	3.80	0.053
Aboveground biomass	1997	0.39	3.02	<0.001	12.3	<0.001	2.60	<0.001	0.52	0.47
Aboveground biomass	1998	0.49	4.80	<0.001	31.8	<0.001	3.81	<0.001	2.56	0.11
Aboveground biomass	1999	0.56	6.27	<0.001	90.7	<0.001	3.15	<0.001	14.9	<0.001
Aboveground biomass	2000	0.61	7.80	<0.001	111.	<0.001	3.97	<0.001	10.7	<0.001

deza capitata in six tests, and *Petalostemum purpureum* in two tests. *Schizachyrium scoparium* and *Sorghastrum nutans*, both C4 grasses, were significant in five tests each. These are five of the six most abundant species in mixtures. A rarer forb also had a significant effect. Similarly, when plots were characterized by the presence or absence of functional groups in ANOVAs, in 2000 there were significant positive effects of legumes ($P < 0.001$), forbs ($P < 0.05$), and C4 grasses ($P < 0.01$) on aboveground biomass, and significant positive effects of legumes ($P < 0.001$) and C4 grasses ($P < 0.001$) on total biomass. For aboveground biomass, only the legume \times C4 grass interaction was significant ($P < 0.05$), and it was positive. For total biomass, the legume \times C4 grass interaction was marginally significant ($P = 0.068$) and biased toward positive (26), suggesting complementarity or facilitation among legumes and C4 grasses (4). However, even after controlling for the presence or absence of all functional groups, there were positive ($P < 0.02$) effects of species number on both aboveground and total biomass in 2000, indicating that biomass also depended on species number rather than on just the presence of functional groups.

Although these analyses suggest that the presence of about five dominant species might explain much of the effects of diversity, there may be small but additive effects of rarer species. To test this possibility, we ranked all species on the basis of their abundance (percentage cover) in 16-species plots in 2000, and created 17 new diversity indices. Each index states how many of the N most abundant species ($N = 2, 3, \dots, 18$) had been planted in each plot. We then determined which diversity index (log transformed) explained the most variance in aboveground or total biomass in 2000. Aboveground biomass was most dependent on how many of the 9 to 13 most abundant species were planted in each plot [supplement D (22)], showing that many rarer species contributed detectably to

aboveground biomass. However, total biomass in 2000 was most dependent on how many of the four most abundant species were planted [supplement D (22)], likely because three of these four are C4 grasses, the species that accumulate the greatest root mass.

In summary, diversity effects were neither transients, nor explained in the long-term solely by other sampling-effect hypotheses, nor solely by the presence of legumes on a low- N soil. Rather, niche complementarity contributed significantly. Plant species number [as in (1–5, 20, 21)] and functional group composition [as in (4, 5, 24, 25)] became simultaneously and approximately equally important in our long-term experiment. Compared with the average of the single best species in monoculture, our 16-species plots had 39% greater aboveground biomass and 42% greater total biomass on average for 1999 and 2000. Moreover, 16-species plots in 1999 and 2000 had 2.7 to 2.9 times greater aboveground and total biomass than the average for all species in monoculture (Fig. 1A). The nonsaturating increase in aboveground biomass with diversity likely reflects niche effects among about 9 to 13 species and their greater chance of co-occurrence at higher diversity (19), whereas such effects among about four species seem to account for total biomass responses.

The demonstration that diversity effects strengthened through time and were not the result solely of sampling effects or functional group composition should resolve aspects of the biodiversity debate (6–14). Moreover, our results suggest, for ecosystems in which niche complementarity occurs, that even with the wisest choices, monocultures may be less productive and accumulate less living carbon than many higher-diversity species combinations. Our results show that ecosystem processes are simultaneously influenced by diversity and composition, but long-term work in additional ecosystems is needed to determine the generality of our results, to better understand the effects of nonrandom community assembly and disassembly, and to better

determine the implications of biodiversity for ecosystem management.

References and Notes

1. S. Naeem, L. Thompson, S. Lawler, J. Lawton, R. Woodfin, *Nature* **368**, 734 (1994).
2. ———, *Philos. Trans. R. Soc. London Ser. B* **347**, 249 (1995).
3. D. Tilman, D. Wedin, J. Knops, *Nature* **379**, 718 (1996).
4. D. Tilman, J. Knops, D. Wedin, P. Reich, M. Ritchie, E. Siemann, *Science* **277**, 1300 (1997).
5. A. Hector et al., *Science* **286**, 1123 (1999).
6. M. Huston, *Oecologia* **110**, 449 (1997).
7. J. Grime, *Science* **277**, 1260 (1997).
8. E. Garnier, M.-L. Navas, M. Austin, J. Lilley, R. Gifford, *Acta Oecol.* **18**, 657 (1997).
9. L. Aarssen, *Oikos* **80**, 183 (1997).
10. J. Hodgson, K. Thompson, P. Wilson, A. Bogaard, *Funct. Ecol.* **12**, 843 (1998).
11. D. Wardle, *Oikos* **87**, 403 (1999).
12. D. Wardle et al., *Bull. Ecol. Soc. Am.* **81**, 235 (2000).
13. J. Kaiser, *Science* **289**, 1282 (2000).
14. L. Guterman, *The Chronicle of Higher Education* (13 October 2000), A24.
15. J. Vandermeer, *The Ecology of Intercropping* (Cambridge Univ. Press, New York, 1989).
16. M. Swift, J. Anderson, in *Biodiversity and Ecosystem Function*, E.-D. Schulze, H. Mooney, Eds. (Springer Verlag, Berlin, Germany, 1993), pp. 15–41.
17. D. Tilman, C. Lehman, K. Thomson, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1857 (1997).
18. C. Lehman, D. Tilman, *Am. Nat.* **156**, 534 (2000).
19. M. Loreau, *Oikos* **91**, 3 (2000).
20. M. Loreau, A. Hector, *Nature* **412**, 72 (2001).
21. C. P. H. Mulder, D. D. Uliassi, D. F. Doak, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6704 (2001).
22. Supplementary Web material is available on Science Online at www.sciencemag.org/cgi/content/full/294/5543/843/DC1.
23. There are 31 possible combinations of five functional groups taken 1 to 5 at a time ($2^5 - 1$), of which 28 combinations occurred in the 168 plots. The combination that occurred in a plot is referred to as its functional group composition.
24. D. Hooper, P. Vitousek, *Science* **277**, 1302 (1997).
25. ———, *Ecol. Monogr.* **68**, 121 (1998).
26. Adjusted mean total biomass in 2000, from ANOVA: 342 g m⁻² for plots with neither C4 grasses nor legumes; 832 g m⁻² for C4's but not legumes; 564 g m⁻² for legumes but not C4's; 1234 g m⁻² for both.
27. We thank the National Science Foundation (NSF/DEB 0080382 and NSF/DEB 9629566) and the Andrew Mellon Foundation for support; J. Fargione, S. Pacala, S. Levin, A. Dobson, and J. Reichman for comments; and N. Larson, C. Bristow, and L. Johnson for assistance.

5 March 2001; accepted 20 August 2001