Quantifying internally generated and externally forced climate signals at regional scales in CMIP5 models

Kewei Lyu1,2, Xuebin Zhang2, John A. Church2, and Jianyu Hu1

1State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen, China, 2CSIRO Oceans and Atmosphere Flagship, Hobart, Tasmania, Australia

Abstract The Earth’s climate evolves because of both internal variability and external forcings. Using Coupled Model Intercomparison Project Phase 5 (CMIP5) models, here we quantify the ratio of externally forced variance to total variance on interannual and longer time scales for regional surface air temperature (SAT) and sea level, which depends on the relative strength of externally forced signal compared to internal variability. The highest ratios are found in tropical areas for SAT but at high latitudes for sea level over the historical period when ocean dynamics and global mean thermosteric contributions are considered. Averaged globally, the ratios over a fixed time interval (e.g., 30 years) are projected to increase during the 21st century under the business-as-usual scenario (RCP8.5). In contrast, under two mitigation scenarios (RCP2.6 and RCP4.5), the ratio declines sharply by the end of the 21st century for SAT, but only declines slightly or stabilizes for sea level, indicating a slower response of sea level to climate mitigation.

1. Introduction

Changes in the Earth’s climate arise from two main factors: (1) unforced variability internally generated within the climate system on various spatial and temporal scales, e.g., the El Niño–Southern Oscillation (ENSO) on interannual time scale and (2) the responses to external forcings such as solar activity, volcanic eruptions, and human-induced changes of greenhouse gases and aerosols. Since internal climate variability can temporarily add to or counteract the forced response, it is difficult to distinguish the responses to external forcings from internal climate variability. A good example is the recent global warming hiatus event, i.e., the slowdown in the rate of global mean surface air temperature (SAT) rise since 1998 despite the continuing increase in greenhouse gas emissions. It has been suggested that internal variability is one important factor modulating the short-term warming rate of Earth’s average SAT [e.g., Marotzke and Forster, 2015].

The spatial distributions of either forced or internal climate signals are generally not spatially uniform, and as a result, the regional climate is usually much more variable than the global average. Thus, the detection and attribution of climate changes at regional scales is challenging, especially when the observational data are available for only a short time period. Using sea level as an example, the regional sea level change pattern over the last two decades in the Pacific measured by satellite altimeters is related, at least in part, to the basin-scale interdecadal climate variability [e.g., Zhang and Church, 2012]. When the analysis period is extended to longer time scales (e.g., centennial), the effect of internal climate variability on regional sea level trends becomes less dominant but remains large at high latitudes [Bordbar et al., 2015; Carson et al., 2015], where there is large low-frequency sea level variability [Monselesan et al., 2015].

State-of-the-art climate models have been used extensively to project future climate. Uncertainties in climate projection primarily result from three factors: (1) the uncertainty in future emission scenarios (which depends on societal choices), (2) intermodel differences in simulating responses to a given forcing, and (3) internal variability of the climate system. The internal variability could be the dominant source of uncertainty in climate projections for the next few decades [Hawkins and Sutton, 2009, 2011; Hingray and Said, 2014]. For example, for regional sea level projections, the uncertainty arising from internal variability is larger than that from scenario differences over most ocean areas during the first half of the 21st century [Little et al., 2015]. Recently, several studies have emphasized the uncertainties of climate projections that are the result of internal climate variability using the Community Climate System Model version 3 with 40 ensemble runs [Deser et al., 2012a, 2012b, 2014; Hu and Deser, 2013].
It is challenging to distinguish between externally forced and internally generated climate signals because both historical and future simulations, as well as observations, reflect a combination of forced responses together with the unforced variability [Solomon et al., 2011; Franzke et al., 2015]. To understand the observed past climate and to predict future climate, it is necessary to quantify the relative magnitude of the externally forced climate signals compared to the internally generated variability. The amplitude of internal variability is usually derived from preindustrial control runs of climate models, with fixed forcing at preindustrial levels, which implicitly assumes that internal variability does not change under external forcing, even though internal variability may also experience some changes [Church et al., 2013a; Huntingford et al., 2013; Kim et al., 2014].

The separation of forced climate signals from internal variability can be addressed using multiple ensemble runs of the same climate model, under the same prescribed forcing but with different initial conditions (i.e., different phases of internal variability). When the ensemble size is large enough (e.g., 30–40 members), averaging across all ensemble runs can approximately reveal the model’s “true” response to external forcings, and the contribution from internal variability can be defined as departures from the ensemble average [Hu et al., 2012; Deser et al., 2012a, 2012b, 2014; Hu and Deser, 2013]. However, for most climate models, the ensemble size is relatively small or even only one simulation is available. Ting et al. [2009] applied a variance analysis method, which was originally developed by Harzallah and Sadourny [1995], to separate externally forced and internally generated variance from a small number of ensemble runs. Using climate models with multiple but limited ensemble runs, we extend the work of Ting et al. [2009] to quantify the ratio of externally forced variance to total variance for two important climate variables: SAT and sea level. In contrast to Ting et al. [2009] who focused on decadal time scale variations by applying a low-pass filter, we consider variations on interannual and longer time scales by annually averaging the data before analysis. We examine how this ratio evolves in the future under different emission scenarios, for different climate variables and for different geographic locations. We consider three future Representative Concentration Pathway (RCP) scenarios [Van Vuuren et al., 2011]: the aggressive mitigation scenario RCP2.6, the medium scenario RCP4.5, and the business-as-usual scenario RCP8.5.

2. Data and Methodology

The new generation of climate models from the Coupled Model Intercomparison Project Phase 5 (CMIP5) provides the data source for our study [Taylor et al., 2012]. We analyze five CMIP5 models which have at least four ensemble runs of historical, RCP2.6, RCP4.5, and RCP8.5 simulations. The CMIP5 models used are as follows: CanESM2 [Arora et al., 2011], CCSM4 [Gent et al., 2011], CSIRO-MK3-6-0 [Rottsteyn et al., 2012], HadGEM2-ES [Martin et al., 2011], and IPSL-CM5A-LR [Dufresne et al., 2013], with the ensemble size of 5, 5, 10, 4, and 4, respectively. The monthly data are first annually averaged, implying that the internal variability discussed in this study refers to year-to-year variability on interannual and longer time scales. Sea level simulations in climate models consist of two parts: one is dynamic sea level (DSL, defined as regional deviations from the global mean), which represents the ocean dynamical adjustment to the changing climate; the other is the global mean thermosteric sea level (GMTSL) due to global ocean thermal expansion, which is one of the major contributions to the global mean sea level rise [Church et al., 2011; Church et al., 2013b]. The drifts in CMIP5 historical and future experiments are removed by subtracting the cubic fit to the sea level simulations at each grid point in the corresponding preindustrial control runs [Sen Gupta et al., 2013].

In observations or each ensemble run of the climate models, the total variance \( v \) of the climate variable \( x \) at a given location can be regarded as the sum of the forced signal variance \( v_F \) due to the external forcings and the internal signal variance \( v_I \) generated internally in the climate system:

\[
v = v_F + v_I. \tag{1}
\]

For a model with a large number of ensemble runs, the forced and internal variances (\( v_F \) and \( v_I \)) can be estimated from the ensemble average and departures from the ensemble average, respectively. Following Ting et al. [2009], the formulas are

\[
\begin{align*}
v_F &= \frac{1}{M} \sum_m \left( \frac{1}{N} \sum_n x_{mn} - \frac{1}{MN} \sum_m \sum_n x_{mn} \right)^2, \tag{2} \\
v_I &= \frac{1}{NM} \sum_n \sum_m \left( x_{mn} - \frac{1}{N} \sum_n x_{mn} \right)^2. \tag{3}
\end{align*}
\]
where \( M \) is the length of the time interval over which the variances are estimated and \( N \) is the ensemble number. Note that, according to the definition of the forced variance (equation (2)), the forced signal is the ensemble-averaged change during the time interval of \( M \) years, rather than relative to an earlier reference period.

When using a relatively small number of ensemble simulations, the internal variability cannot be fully removed by averaging across the ensemble, which systematically increases the variability of the ensemble average as the ensemble average contains both the forced signal and some elements of the internal variability. Thus, the estimate of forced variance in equation (2) is biased high. The unbiased estimates of forced and internal variances \((v_F \text{ and } v_I)\) [Harzallah and Sadourny, 1995] are the following:

\[
v_F = v_F \frac{1}{N-1} v_i,
\]

\[
v_I = N \frac{N-1}{N} v_i.
\]

Sensitivity tests using 10 ensemble runs of CSIRO-MK3-6-0 and 30 ensemble runs from the Community Earth System Model large ensemble project [Kay et al., 2015] confirm that after correction of the bias, these unbiased estimates give robust results that are no longer sensitive to the ensemble size (see Figures S1 and S2 in the supporting information). In the case of a small ensemble size and large internal variability, the resulting bias in the ensemble average (second term on the right side of equation (4)) could potentially be larger than the ensemble average (first term on the right side of equation (4)), which leads to forced variance estimates being less than zero. Under this overcorrected situation, the forced variance is set to zero [Harzallah and Sadourny, 1995]. Then the ratio \((R)\) of forced variance to total variance can be calculated using the unbiased variance estimates:

\[
R = \frac{v_F}{v_F + v_I}.
\]

These variance calculations could be applied to a large ensemble of all available ensemble runs from different models. While the multimodel ensemble mean provides the best estimate of the forced climate signal [e.g., Collins et al., 2013] and thus the forced variance, the internal variance estimates could contain a spurious contribution from the intermodel uncertainty [Hawkins and Sutton, 2009]. Therefore, in this study, the internal variance \((v_i)\), forced variance \((v_F)\), and the ratio \((R)\) are calculated at each grid point using each model’s multiple ensemble runs first and the multimodel mean estimates are derived afterward.

3. Results

3.1. Spatial Patterns of Forced Variance Ratio

We examine the multimodel mean spatial patterns of forced variance ratio for the historical period (1860–2005) (Figure 1, see Figures S3–S5 in the supporting information for ratio patterns over the projection period). For the DSL (Figure 1a), most ocean areas indicate a small ratio, indicating the overwhelming dominance of internal DSL variability relative to the forced DSL change. The most striking features occur at high latitudes of the Southern Hemisphere, where there are two zonal bands of relatively large ratios on both sides of 50°S. Over the historical period, the two maxima of zonal average ratio are 15% around 40°S and 23% around 67.5°S, respectively (see Figure S6 in the supporting information for zonal average results). It is important to note that changes of the DSL along these two bands are of the opposite signs: rising (falling) in the north (south) [e.g., Yin, 2012; Lyu et al., 2014], which cannot be recognized from the variance patterns. This dipole-like meridional structure of DSL change is mainly induced by a strengthening and poleward shift of the westerly winds, as suggested by both low-resolution climate models and eddy-permitting ocean models [Bouttes et al., 2012; Frankcombe et al., 2013]. Another region of interest lies in the northwestern Atlantic, where ocean dynamics (e.g., the slowdown of the Atlantic meridional overturning circulation) contribute to the regional sea level rise [Yin et al., 2009; Yin and Goddard, 2013] and the forced DSL change could account for as much as 20% of the total variance.

The model-simulated GMTSL change is mainly driven by anthropogenic and natural forcings with small internal variability (only on the order of millimeter) [Slagen et al., 2014b]. Thus, including GMTSL in the calculations has
little impact on the internal variance estimates but alters (often increases) the magnitudes of the forced variance (see Figure S7 in the supporting information for spatial patterns of internal and forced variances). Large internal DSL variability generally results in regions of small forced variance ratios, e.g., the tropical Indo-Pacific Ocean and the Kuroshio Extension in the northwestern Pacific (Figure 1b). In the Southern Ocean, the falling DSL largely offsets the effect of the GMTSL, which leads to small ratios. In contrast, to the north, rising DSL together with GMTSL rise leads to larger ratios with the zonal mean value of 40% at 40°S over the historical period. Large forced ratios are also found in the tropical and southern Atlantic. The ratios at low latitudes increase largely over the projection period (see Figures S3–S5 in the supporting information).

For SAT, the forced variance ratios are largest in tropical areas, except for the regions that are heavily affected by the ENSO variability, e.g., the central and eastern tropical Pacific (Figure 1c). The zonal average ratio in the tropics is larger than 40% over the historical period. Consistent with our results based on climate models,
robust warming trends relative to the background internal variability have been detected based on observations over the tropical Indian Ocean [Du and Xie, 2008], the tropical Atlantic Ocean [Tokinaga and Xie, 2011], the western tropical Pacific [Wang et al., 2015], and land at low latitudes [Mahlstein et al., 2012]. At higher latitudes, although the forced variance is larger, the internal variability there is also larger (see Figure S7 in the supporting information), resulting in the ratio being smaller than that in the tropics.

The ratio patterns described above resemble those for the time of emergence, which is defined as the time when the climate change signals emerge from (become larger than) the internal variability [Diffenbaugh and Scherer, 2011; Mahlstein et al., 2011; Hawkins and Sutton, 2012; Mora et al., 2013; Muir et al., 2013; Lyu et al., 2014; Richter and Marzeion, 2014; Bilbao et al., 2015]. Besides the variance ratio used in this study or the widely used signal-to-noise ratio [e.g., Mahlstein et al., 2011], the time of emergence is another important metric used to compare the climate change signal with internal variability but from a different perspective. The larger ratio of forced variance to total variance generally corresponds to earlier emergence of forced signals from internal variability.

### 3.2. Globally Averaged Forced Variance Ratios

The evolution of globally averaged ratio with time is also examined (Figure 2), with the concatenated historical simulations (1860–2005) and future projections (2006–2100) under three emission scenarios. The variances are first calculated using a fixed time interval. For a 30 year running window, the forced variance in DSL only accounts for a small part of total variance throughout the 21st century (Figures 2a–2c). The globally averaged ratio increases slightly to 9%–15% under RCP8.5, and to around 5% under both RCP2.6 and RCP4.5 by 2100. For SAT and DSL plus GMTSL, the globally averaged ratios have larger increases with time under RCP8.5 (Figure 2c), corresponding to the continuous increase of radiative forcing. Over the last 30 years of the 21st century (i.e., 2071–2100), the forced signals are projected to account for a large part of the total variance (globally 43%–56% for SAT and 53%–67% for DSL plus GMTSL).

Under RCP2.6 (Figure 2a), the globally averaged variance ratio for SAT is projected to peak around the beginning of this century and then declines to less than 5% after around 2050. Similar but less sharp decline

---

**Figure 2.** The globally averaged ratio of forced variance to total variance computed over (a–c) 30 year, (d–f) 50 year running windows, and (g–i) the period starting in 1860 and ending from 1870 to 2100. Results from five individual models are shown. Dynamic sea level is indicated by green lines, dynamic sea level plus global mean thermosteric sea level (GMTSL) by blue lines, and surface air temperature by red lines.
occurs for RCP4.5 with less aggressive mitigation (Figure 2b). In contrast, the globally averaged variance ratio for
DSL plus GMTSL exhibits much smaller decrease than that for SAT under RCP2.6, and only becomes flat under
RCP4.5, ending at 25%–40% by the end of the 21st century. This contrast under two mitigation scenarios
reflects that these two climate variables exhibit very different responses to the same mitigation scenario.
Compared with the SAT, the DSL plus GMTSL, which represents the whole depth-integrated change of sea
water properties, takes a much longer time to respond to climate change mitigation.

The changing rates of global mean SAT and GMTSL have similar temporal evolutions to the variance ratios
(not shown). Under RCP2.6 and RCP4.5 with a stabilization of radiative forcing, the global mean SAT is
projected to be nearly stabilized with very small warming rate by the end of the 21st century [Collins et al.,
2013; Smith et al., 2015]. However, the rate of GMTSL rise only decreases slightly under RCP2.6 and becomes
nearly constant under RCP4.5 by the end of the 21st century [Church et al., 2013a; Lyu et al., 2014]. Various
climate model experiments have shown that sea level would continue to rise due to thermal expansion after
the global mean SAT is stabilized by aggressive mitigations [e.g., Bouttes et al., 2013]. Several studies have
suggested that the ongoing heat uptake of the deep ocean is an important contribution to the continuing
sea level rise [e.g., Stouffer and Manabe, 1999; Meethl et al., 2012].

When longer running windows are used (e.g., 50 years), the evolutions of globally averaged ratio closely follow
the same tracks of 30 year running windows except that the ratios are generally larger (Figures 2d–2f).
This implies that with the same forcing, it is easier to detect forced signals by extending the analysis period.
The globally averaged ratios over both time windows increase rapidly starting from about 1950, corresponding
to the rapid increase of radiative forcing from anthropogenic emissions [Myhre et al., 2013].

Note that in the above two cases with a fixed length time window, the forced variance only measures the
magnitude of forced change over the time interval rather than the forced change signals since the start of
the simulation. For example, although the forced variance ratio and warming rate for SAT over the last 30 or
50 years of the 21st century are projected to be small under RCP2.6 and RCP4.5, the global SAT by the end
of 21st century will be well above the preindustrial or current level, e.g., 1.1°C to 2.6°C (5% to 95% range) for
2081–2100 relative to 1986–2005 under RCP4.5 [Collins et al., 2013]. In order to capture this information, we
apply the same variance analysis to time intervals with the starting time being fixed at 1860 and the end time
increasing from 1870 to 2100. In this case, the globally averaged ratios of forced variance continue to increase
with time for all three variables under all three scenarios, reflecting the cumulative effect of externally forced
climate change (Figures 2g–2i).

Changes of the variance ratio with time could be due to changes in the forced variance, the internal variance
or both. We found that changes of the forced variance ratio are mainly a result of changes of the forced
variance (see Figure S8 in the supporting information). The magnitudes of the internal variance calculated
over different time intervals (e.g., 30 years and 50 years) are very similar. The change of the internal variance
with time is relatively small compared to that of the forced variance, especially when averaged at global
scales, consistent with the finding by Hu et al. [2012] based on one single model. Note that the internal variance estimated according to equation (5) from multiple ensemble runs of the climate model is comparable
with that calculated directly from the preindustrial control run of the same model (not shown).

### 3.3. Regional Analyses

We further calculate the forced variance ratio over two representative regions, a tropical region and a
high-latitude region, to illustrate the regional differences. Note that different regions are chosen for
SAT and sea level (see Figure 1 for their locations). The same variance analysis method is applied to the running
50 year window of regional mean time series from concatenated historical and RCP8.5 simulations. The ratios
tend to be smaller with a shorter running window.

Two regions are chosen for SAT: the Maritime Continent (108–128°E, 10°S–10°N) and the North American
Continent (104–84°W, 34–54°N) (Figure 1c). All five models indicate larger ratios of forced variance over
the Maritime Continent than over the North American Continent (Figure 3a). Over the North American
Continent, the forced variance is greater than that over the Maritime Continent, but the internal variance
over the North American Continent is 10 times larger than that over the Maritime Continent (see Figure S9
in the supporting information). Around the beginning of the 20th century, all models show a large, temporary
increase of forced variance ratio over the Maritime Continent. It could possibly be a result of the inclusion of
large tropical volcanic eruptions over that period (Krakatau in 1883 and Santa Maria in 1902), which lead to abrupt changes of the radiative forcing [Myhre et al., 2013], the tropical climate [Maher et al., 2015] and thus the forced variances.

For DSL, we choose two regions in the northwestern tropical Pacific (130–150°E, 0–20°N) and the northwestern Atlantic (60–40°W, 42–62°N) (Figure 1a). In the northwestern Atlantic, three models show large ratios of forced variance since the early 21st century, but the other two models have much smaller ratios (Figure 3b), indicating the large intermodel uncertainty of DSL projections in this region [Yin, 2012; Little et al., 2015]. The internal variance estimates over a running 50 year window (see Figure S9 in the supporting information) are not necessarily stable with time due to modulations of the low-frequency variability (e.g., multidecadal and centennial). One model (IPSL-CM5A-LR), in particular, shows much larger fluctuations than other models in internal variance estimates of the northwestern Atlantic sea level, because this model exhibits energetic low-frequency signals on multidecadal and longer time scales and no coherent long-term change signal can be found in the four ensemble runs. In the northwestern tropical Pacific, all models indicate that the forced variance in DSL only accounts for less than 5% of the total variance over a 50 year period (Figure 3b). The ratio is projected to increase in the future under RCP8.5, but the intermodel spread also increases with time.

4. Summary and Discussion

In this study, we have investigated the ratio of externally forced climate variance to total variance at regional scales due to either natural or anthropogenic forcings, with focus on SAT and sea level. The variance analysis method applied in this study provides a better partitioning of the total variance into internally generated and externally forced components from limited ensemble of climate model simulations.

For SAT, the forced variance ratios are found to be large in tropical areas, where the forced climate change signal is more evident above the background of small internal variability. In contrast, the lower ratios in extratropical areas result from the larger internal variability. For sea level with ocean dynamics and global mean thermosteric contributions included, large ratios are found at high latitudes in the Southern Hemisphere (~40°S) and the tropical Atlantic over the historical period. It is important to note that adding other contributions (e.g., land ice loss) would enhance the sea level rise signal in most ocean areas [Church et al., 2013a; Lyu et al., 2014; Slangen et al., 2014a], further enhancing the ratios of forced variance to total variance.

When only the contribution of ocean dynamics is considered for regional sea level change, the forced variance ratios are generally small except at high latitudes. Recently, there has been a debate as to whether the rapid sea
level rise in the western tropical Pacific observed by altimetry over the past two decades, with the global mean removed, contains any anthropogenic signal or is dominated by internal variability [Zhang and Church, 2012; Meyssignac et al., 2012; Hamlington et al., 2014; Bilbao et al., 2015; Carson et al., 2015; Palanisamy et al., 2015]. This issue is not fully resolved as yet. Based on our estimates of forced variance ratio, it would be hard to detect any forced dynamic sea level change signal, since the ratio of forced variance to total variance is quite low in that region (less than 5% over a 50 year window).

Averaged globally, the forced variance ratios increase continuously with time during the 21st century under all three scenarios when calculated with a fixed starting time (e.g., the preindustrial level). When calculated over a fixed time interval (e.g., 30 or 50 years), the continuous increase of ratio only exists under RCP8.5. In this case, the future increases of the globally averaged ratio for SAT and sea level closely follow each other. However, over a fixed time interval under RCP2.6 and RCP4.5, the globally averaged ratios for these two variables exhibit quite different changes after reaching the peak: declining for SAT but only becoming nearly flat or declining slightly for sea level by the end of the 21st century. This contrast clearly reflects that the sea level takes a longer time than SAT to respond to the same mitigation scenario.

We suggest that for the purpose of addressing the effects of internal variability, modeling groups should provide multiple realizations of the same model experiment if possible. However, large ensembles are not generally available as they are computationally expensive. While this study is an initial attempt, more methods should be developed to partition the internal and forced climate components for ensembles of small size. We also note that there are still considerable intermodel spreads in our estimates of forced variance ratio, especially for sea level. Most climate models suffer from mean state biases at regional scales [e.g., Wang et al., 2014], which raise large uncertainties in both the simulation of responses to external forcings and the representation of internal variability. For example, the biases of tropical mean states in climate models could influence simulations of the Indian Ocean Dipole [Cai and Cowan, 2013] and ENSO [Huang and Kug, 2015] as well as projections of future changes [Huang and Ying, 2015]. Any biases or errors in model simulation can inevitably affect our estimate of the ratio of forced variance to total variance, but it is hard to quantify such impact especially in a global framework, since many biases are regional. Potentially, this part of the uncertainty could be narrowed with further improvement of climate models [Hawkins and Sutton, 2009].

Acknowledgments
We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP, the U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We also thank Didier Monselesan and Jaclyn Brown for their detailed comments on the draft, as well as three anonymous reviewers for their comments and helpful suggestions. J.H. and K.L. were supported by the National Basic Research Program of China (2015CB954004). J.A.C. and X.Z. were supported by the Australian Climate Change Science Program (ACCCSP). K.L.’s visiting at CSIRO was funded by the China Scholarship Council (201306310079).

References


