Responses to Reviewer 1

We thank the reviewer for her/his supportive review of our manuscript. Her/his suggestions will help us to strengthen the revised manuscript. In the response below, we address her/his concerns sequentially, with our responses indicated in "Arial" font, whereas her/his comments are indicated in "Times New Roman" font.

This study explores different contributors to the increase in the atmospheric CO_2 seasonal amplitude, as predicted by the CESM in simulations that span 1950—2300. I am generally supportive of this paper. Clearly an impressive effort went into it and it is well organized and written. However, I have some major concerns, listed in order of decreasing priority:

1) There are some steps in the methodology that need more detail and justification – some of them could/should really be stand-alone papers. These include a) The pulse-response method. b) The documentation of mid-latitude trends in observed CO_2 amplitude

We have added additional documentation, detailed below, to demonstrate the pulse-response approach we used to calculate atmospheric CO_2 mole fraction data. Because the focus of this paper is not on attributing drivers of the observed change in the amplitude, but rather exploring how seasonality changes in the future in a prognostic ESM, we prefer to minimize the discussion of observed mid-latitude trends.

2) The CESM does a poor job of reproducing the current CO_2 amplitude and the historical observed amplitude trends, which undermines confidence in the results presented here. Although I think the exercise is still worthwhile, some sort of well thought out rationale or statement is needed to explain why readers should believe or pay any heed to the future model results going out to 2300, e.g., are there certain results that are robust and insightful despite the model's poor present-day performance?

The exercise of predicting carbon-climate coupling in a fully prognostic model is still relatively new. Although CESM shows significant deficits in the simulated mean annual cycle and its trend, the model includes parameterizations for many of the processes that may be important in controlling its change with time, and we note that CESM qualitatively captures the northward increase in the NH atmospheric CO_2 seasonal amplitude as well as the increasing trend in the annual seasonal cycle amplitude. These suggest that the parameterizations included in CESM can be used to examine how the seasonal amplitude might evolve when subject to the radiative and fertilization effects of increased atmospheric CO_2 concentration, and also to identify deficiencies in current model parameterizations. While we do not expect that the simulation provides an accurate description of either carbon cycling or physical climate out to 2300, the results of the simulation do allow us to (1) understand the balance of major drivers, (2) identify deficiencies that may need to be addressed in future model development.

We will add the following statement to the discussion of the revised manuscript:

"Although CESM does not quantitatively reproduce the contemporary mean annual cycle amplitude or its trend over the last 50 years, parameterizations in the model qualitatively reproduce diagnostics such as the increase in both the mean annual cycle and its multi-decadal trend. Thus, we can use the model to understand partitioning of the long-term response to climate change or to fertilization, with an eye toward identifying areas for future model improvement."

Expanding on 1a) The pulse-response method. This could really be a stand- alone paper (see, e.g., Nevison, C.D., D.F. Baker, and K.R. Gurney, A methodology for estimating seasonal cycles of atmospheric CO_2 resulting from terrestrial net ecosystem exchange (NEE) fluxes using the Transcom T3L2 pulse-response functions, Geosci. Model Dev. Discuss., 5, 2789-2809, 2012, www.geosci-model-dev- discuss.net/5/2789/2012/ doi:10.5194/gmdd-5-2789-2012, 2012.)

While I support the method and realize that it would be prohibitively expensive computationally to break down the contributions to CO_2 amplitude change from different regions and mechanisms without some sort of shortcut approach like the Pulse Response method, I think it needs more than a 1 paragraph explanation. For example:

i) Is there any IAV in the meteorology used to create the pulse fields? Also, what is the consequence of assuming those met fields will still apply in 2300?

We appreciate that the reviewer recognizes that the pulse-response method is a necessary computational shortcut to examine the regional contributions to atmospheric CO_2, and will provide more details about the method in the revised manuscript. The pulse response approach does contain interannual variability in the met fields, but as the reviewer identifies, there are substantial consequences in assuming those met fields will still apply in 2300. We note that in our manuscript Fig. 2c, mismatches grow from 2 ppm to 3 ppm in the mid- and high latitudes when the land CO_2 tracer in CESM (4-d) is sampled at the sites we use in Fig. 1 vs when NEE (which embodies all the land processes that influence the land CO_2 tracer) is convolved with the pulse-response function. This deficiency should be identified in the paper, and we plan to add the following text to our description of the pulse-response method:

"Although the CESM simulated the three-dimensional structure of atmospheric CO_2, we used a pulse-response transport operator to separate the imprints of CO_2 fluxes from different regions on the hemispheric CO_2 patterns. The transport operator was developed using the GEOS-Chem transport model (version 9.1.2, Nassar et al. (2010)). GEOS-Chem was configured as in Keppel-Aleks et al. (2013) on a 4° × 5° horizontal grid with 47 vertical layers, and forced with meteorology fields from the 3–6-hourly Modern Era Retrospective-Analysis for Research and Applications (MERRA) reanalysis dataset (Rienecker et al., 2011). A tagged 1 Pg C month^{-1} pulse was released for each of the 20 terrestrial source regions in Fig. 1 for each calendar month. Each 1 Pg C month^{-1} pulse was distributed spatially according to monthly fluxes from the Carnegie-Ames-Stanford Approach (CASA) fluxes from Olsen and Randerson (2004). At a given location, the magnitude and phasing of the atmospheric CO_2 response of the pulse depends on the characteristics of atmospheric transport. For example, at Barrow (BRW) in Northern Alaska, a 1 Pg pulse released in Boreal North America has a large impact on atmospheric CO_2 (2 ppm, Fig. SA1a) during the first 1-3 months after a pulse is released. In contrast, when the pulse is released from temperate North America, there is a phase lag of 1 month (Fig. SA1b,c), and when the pulse is released from the Amazon, there is a delay in the peak response at BRW of 3 months (Fig. SA1d). The magnitude of the response is also smaller (e.g., 0.02 ppm for a 1 Pg pulse released in the Amazon vs 2 ppm for boreal North America; Fig. SA1), since the pulse has already diffused over much of the globe. We also note that seasonal patterns in atmospheric transport affect the imprint a pulse leaves on atmospheric CO_2 (Fig. SA2). For example, a 1 Pg pulse from the boreal region leaves a 2 ppm contribution on CO_2 at Barrow during the winter months, but more rigorous vertical mixing in the summer months reduces the imprint to 0.5 ppm. Following the twelve month period in which pulses were released, the signals were allowed to decay for 60 subsequent months, at which point CO_2 was well-mixed in the atmosphere (Fig. SA2a-d).

We then sampled GEOS-Chem at the locations of 41 NOAA cooperative CO_2 flask sample sites (Dlugokencky et al. (2013); Table 1, Fig. 1) for each month simulated. This resulted in a CO_2 transport operator matrix with dimensions $N_{reg.} \times N_{obs.} \times N_{mon.}$. We aggregated NEP fluxes from CLM4 to the spatial scale of the 20 source regions (Fig. 1), and used matrix multiplication to propagate these fluxes to atmospheric CO_2 space. We calculated the monthly mean CO_2 mole fraction at the observation sites by summing over the regional contributions to get a CO_2 response matrix with dimensions (N_{obs} x N_{mon}).

We analyzed both the CO_2 fields from global fluxes and CO_2 patterns influenced only by larger regions representing Arctic, boreal, temperate, subtropical, tropical, and Southern Hemisphere (SH) ecosystems. We calculated the CO_2 annual cycle amplitude values as the peak-to-trough differences in CO_2 summed over each component region (e.g., the CO_2 annual cycle amplitude at a given station from pulses emitted from the Arctic was calculated as the peak-to-trough difference in the sum of CO_2 from pulses emitted by the blue regions in Fig. 1). We note that our analysis focuses on surface observations of atmospheric CO_2, and does not include aircraft measurements.

The advantage of this method is that we can efficiently compute the regional contribution to changes in atmospheric CO_2; it would be prohibitively expensive to run a full atmospheric transport model for each of the regions separately for 350 years. To evaluate this method, we show a comparison in which we have generated CO_2 using NEE, since the land CO_2 tracer in the CAM4 is derived from NEE (despite that we use NEP for subsequent analyses). The magnitudes generally differed by less than 2 ppm due to different model boundary layer schemes and atmospheric transport (Fig. 2c). We note that the largest differences were during the last century of the simulation, which we hypothesize was due to shifts in atmospheric transport in response to the dramatic climate change in the CAM4. The fact that long-term trends in transport are not simulated by the pulse-response approach is one of the major sources of bias. By neglecting long-term trends in transport, we induce a bias into atmospheric CO_2 that increases with time (Fig. 2c). In a site-by-site comparison (Fig. SA3), the mismatch in time appears to be due to amplification of existing biases in the pulse-CO_2 compared to the full transport-CO_2. A second source of uncertainty is that the spatial distribution of fluxes within each region is different in CESM compared to CASA. We expect that this has a minimal impact based on results from Nevison et al. (2012), who showed that a similar pulse response code using different transport models did a reasonable job (r^2=0.8) of simulating the fossil fuel influence on CO_2 despite that CO_2 has a vastly different spatial configuration than do ecosystem fluxes.

We also assessed the validity of the assumption to model only the land contributions to trends in the mean annual cycle of CO_2 by calculating the CO_2 amplitudes in the CAM land and ocean tracers. We found that the contemporary peak-to-trough amplitude in the ocean tracer averaged across our high latitude stations was 2 ppm (in contrast to 10 ppm in the land tracer). Although both the land and ocean amplitudes grow with time, by 2300, the high latitude ocean tracer had an amplitude of 3 ppm, only 18% of the land amplitude for this time period. Ocean carbon uptake was found to change significantly in CESM through 2300 (Randerson et al., 2015), but based on these numbers, ocean CO_2 still had a smaller imprint on the atmospheric annual cycle."

ii) How are the 60-month decaying pulses combined to create a model atmospheric CO_2 cycle?

We have addressed the reviewer's question in the revised text, above, and created two additional figures (Fig. SA1 and Fig. SA2) to show this process graphically.

iii) In figure 2, the pulse-response amplitudes at midlatitudes are 3 ppm or more smaller than the fully prognostic tracer. This doesn't seem "broadly similar" and undermines confidence that this methodology can detect subtle trends, esp. in the midlatitudes.

To provide better validation for the reader to assess the bias induced by the pulse-response method, we have prepared Fig. SA3, which shows the mean seasonal cycle at a high-latitude (BRW), mid-latitude (SHM), subtropical (KEY), and tropical (MLO) NH site. These sites were selected since they have observational records dating to the 1980s (gray circles shown in Fig. 1). We plot both the CESM land CO_2 tracer and the pulse-response CO_2 for four periods for each site: 1990—1999, 2090—2099, 2190—2199, and 2290—2299.

The site-by-site comparison shows that (1) the biggest mismatches in 2300 between the full-transport and the pulse-response CO_2 owe to persistent biases that exist for the present, e.g., the high January bias at Barrow (BRW) and the one-month phase shift in the summer minimum at KEY. This would suggest that changes in transport patterns due to climate change induce a smaller mismatch than present-day biases in the method. (2) the method is able to capture fairly subtle variations in the mean annual cycle, such as the "W" shape that the mean annual cycle at SHM develops over time.

iv) The GMD Discussions paper above was never accepted for final publication, due to reviewers who thought adjoint methods were superior. While the current method is superior in that it divides land into a larger number of regions (20 v. 11), the GMDD paper on the other hand was applying the method to estimate mean seasonal cycles, which are easier to get right than the more subtle trends in amplitude over time examined here.

We agree with the reviewer that not only are the improved resolution of land areas an advantage of our pulse-response code over the Transcom regions, but also the fact that these land regions were determined based on similarity in annual mean NPP and its seasonality. We agree that comparison of mean annual CO_2 cycle, rather than its trend, places a lower burden on the code. For this application, however, the mean annual CO_2 amplitude changes by up to 10.6 ppm by 2300. Thus, the relative error, assuming a 3 ppm difference, is still only ~28% of the total trend.

Expanding on 1b) I'm not sure there is any evidence that CO_2 seasonal amplitude is increasing at midlatitude sites such as NWR or UUM, KZM/D. In fact, if anything, they may be decreasing – possibly due to drought effects. The most robust effects are seen at BRW, with the amplitude increase at MLO less than half that of BRW. I don't think Zeng et al. (2014) is an adequate reference to prove that midlatitude CO_2 amplitude is increasing, since they don't actually show this.

We thank the reviewer for this comment. One of the reasons we aggregate the sites depicted in Fig. 1 into high-, mid-, subtropical, and tropical latitude belts is to minimize local effects that may be present at the sites we have chosen and to instead focus on a more large-scale pattern of variation. A challenge to this approach is that there are relatively few ESRL sites with records dating to the 1980s or earlier. We will include this rationale in our methods discussion by including the text:

"In our analysis, we aggregate the sites into high-, mid-, subtropical, and tropical latitude belts to minimize local effects at individual sites and instead to focus on large-scale trends owing to broad patterns of climate change."

We will also thoroughly check our referencing in the revised paper and be sure to reference observationally based papers such as Randerson et al. (1997) and Graven et al. (2013) rather than modeling-derived studies.

Minor comments: p.1, L8, The term "changing atmospheric composition" to encompass CO_2 fertilization and N deposition is confusing. These two don't really belong in the same category, in my opinion, since the N deposition is relevant mainly after it deposits on the soil, i.e., the authors are not looking at some sort of physiological response of plants to increased atmosphere NOx or NH3 concentration.

We agree with the reviewer that CO_2 fertilization and N-deposition represent two distinct forcings on ecosystem carbon exchange. Unfortunately, the simulations which were conducted as part of the CESM Biogeochemistry Working Group did not separate these two changes, thus we cannot resolve the specific forcing from the model output available since there is no analogous CESM ECP simulation that excludes N-deposition and the CO_2 radiative effect. In each of the ECP simulations, reactive nitrogen deposition was kept constant at 2100 values (Randerson et al., 2015), so future trends in the mean annual cycle amplification were due to nitrogen deposition levels at 2100 interacting with trends in CO_2.

Devaraju et al. (2016) did perform experiments looking at the individual and combined effects of CO_2

fertilization, N-deposition, climate change, and LUC on historical NPP trends using the CESM1(BGC). They found that CO_2 fertilization and N-deposition contributed 2.3 and 2 PgC yr^{-1} to the 4 PgC yr^{-1} historical increase in global terrestrial NPP. Given the conditions of the experiments and the fact that CO_2 fertilization and N-deposition contribute similarly to global and historical NPP trends in the CESM, we present our results based on the combined effects of CO_2 fertilization and N-deposition.

We will reference this paper in the text, and include the following statement:

"Results from Devaraju et al. (2016) suggest that global NPP is influenced equably by CO_2 fertilization and nitrogen deposition over the historical period in CESM, so trends in the mean annual cycle amplitude were likely influenced by this enhanced NPP. In these simulations, nitrogen deposition was held fixed after 2100, so trends in the amplitude were influenced by anthropogenic nitrogen deposition but not forced by transient deposition."

We will also use clearer language to describe that these two effects are included in the simulations by replacing "changing atmospheric composition" with "CO_2 fertilization and N-deposition" throughout the revised manuscript.

p.1, L12 is confusing as written – in one case we have the end time (2300) and in the other we have the start time (after 2100). Please rewrite to clarify start and end times for both effects

We will revise the sentence to read "CO_2 fertilization and N-deposition in NH boreal and temperate ecosystems were the largest contributors to mean annual cycle amplification over the midlatitudes for the duration of the simulation (1950—2300) and for the Arctic from 2100—2300."

p.1, L15 "rather than the strength of the terrestrial carbon sink" please explain more clearly what is meant here.

We will clarify this sentence to read "Greater terrestrial productivity during the growing season was the largest contributor to the annual cycle amplification throughout the Northern Hemisphere."

p.1, L17, suggest replacing "is not predicated on" with "does not necessarily imply" p.1, L20 I think it's more accurate to say "at some NH sites" rather than "over the NH" (see my comments above about midlatitude trends).

We will change the sentence to read "Prior to 2100, CO_2 annual cycle amplification occurred in conjunction with an increase in the NH land carbon sink, but these trends decoupled after 2100, underscoring that an increasing atmospheric CO_2 annual cycle amplitude does not necessarily imply a strengthened terrestrial carbon sink."

p.2, L31 missing AND between citations.

We will add the "and" between McDonald et al. (2004), and Barichivich et al. (2013).

p.2, L35 suggest saying, "Model evidence suggests that the combined effects . . ." and delete "in simulations."

We will revise the sentence to read, "Model evidence suggests that the combined effects of climate change and shifts in vegetation cover can also enhance GPP."

P2., L20 and p.3, L17 again I find the catch-all term "changes in atmospheric composition" confusing.

We will refer to “changing atmospheric composition” as “CO_2 fertilization and N-deposition”.

p. 6, L30. It seems like a stretch to call 425 ppm and 391 ppm “roughly equivalent”

We agree with the reviewer and will modify the text in the revised manuscript to state "We note that the drivers of the amplitude increase during 1985—2013 were simulated to different levels of fidelity: the NH atmospheric temperature increase over land was roughly equivalent (1.02 K vs 0.95 K in the NCEP-NCAR reanalysis [Kalnay et al., 1996]), but the NH atmospheric CO_2 mole fraction in CESM was too high (425 ppm vs 391 ppm). Previous analysis of the CESM shows that the high CO_2 bias is attributable to persistent weak uptake in both land and ocean (Keppel-Aleks et al., 2013; Long et al., 2013)."

p.7, L5 Please provide a reference for the observed mid-latitude trend of 0.04 ppm yr^{-1}.

We thank the reviewer for calling this detail to our attention. We calculated the 0.04 ppm yr^{-1} midlatitude trend from the 1985—2013 monthly observations at Shemya Island, Alaska (SHM), which we selected to represent the midlatitudes (40°N—60°N; Table 1) based on its sufficiently long period of record.

We will revise the text to clarify "Both the modeled and observed trends in the CO_2 annual cycle amplitude were calculated from individual sites whose records date to 1985 (gray circles in Fig. 1). The modeled trend in the CO_2 annual cycle amplitude..."

P8, L19 Please explain further. Why is this consistent with effects being proportional to GPP?

Regional GPP is smaller in the Arctic than in the other regions we analyze in the paper, thus fertilization acts as a knob on a smaller gross flux term. We will revise the statement to read, "CO_2 fertilization and N-deposition effects were smallest in the Arctic, the region with the smallest GPP for the contemporary period. In the CESM, the impact of CO_2 fertilization on the amplitude trend roughly scales with to the magnitude of overall GPP, consistent with hypotheses from Tans et al., (1990) and Schimel et al (2015) that the fertilization effect on the land carbon sink is proportional to productivity."

P9, L12, to avoid confusion, would suggest splitting into 2 sentences: “..simulation. These latter influences added 4.7 ppm . . .”

We will split these two sentences in the revised manuscript.

P9, L27 The Zeng et al reference, in my reading, does not actually demonstrate that the spatial distribution of where atmospheric CO_2 amplitude increases are seen (mainly at high latitudes) are consistent with agriculture, which is large at mid-latitudes.

We agree with the reviewer that the Zeng et al. (2014) reference does not explicitly calculate the impact of midlatitude agricultural fluxes on the high latitude mean annual cycle amplitude, where the trend is largest. However, our results (Fig. 7) show that temperate ecosystems leave a large imprint on the mean annual cycle amplitude at high latitudes. Thus, if crops were included in the CESM, the model would show the imprint at high latitudes. We therefore prefer to leave the statement unchanged.

P10, L23, “perhaps indicating . . .” Please explain further.

Based on comments from both reviewers, we have decided to revise the text to remove this statement. Instead, we will include a statement that this finding demonstrates the importance of considering latitudinally resolved CO_2 in models for diagnosing compensating errors. As described in the response to Reviewer 2, one difference between our paper and other papers on the mean annual cycle is that we explicitly consider how fluxes propagate to atmospheric CO_2 rather than simply aggregating hemispheric fluxes.

We will revise the text to read: "This result underscores the importance of considering meridionally resolved atmospheric CO_2 data that explicitly considers the role of transport, since a Northern Hemisphere average masks incorrect spatial patterns in the CESM."

Additional References

Devaraju, N., Bala, G., Caldeira, K., and Nemani, R.: A Model Based Investigation of the Relative Importance of CO_2-Fertilization, Climate Warming, Nitrogen Deposition and Land Use Change on the Global Terrestrial Carbon Uptake in the Historical Period, Clim. Dynam., 47, 173—190, doi:10.1007/s00382-015-2830-8, 2016.

Nevison, C. D., Baker, D. F., and Gurney, K. R.: A methodology for estimating seasonal cycles of atmospheric CO2 resulting from terrestrial net ecosystem exchange (NEE) fluxes using the Transcom T3L2 pulse-response functions, Geosci. Model Dev. Discuss., 5, 2789–2809, doi:10.5194/gmdd-5-2789-2012, 2012.

Tans, P. P., Fung, I. Y., and Takahashi, T: Observation constraints on the global atmospheric CO_2 budget, Science, 247, 1431–1438, 1990.

Olsen, S. C., and J. T. Randerson: Differences between surface and column atmospheric CO_2 and implications for carbon cycle research, J. Geophys. Res., 109, D02301, doi:10.1029/2003JD003968, 2004.

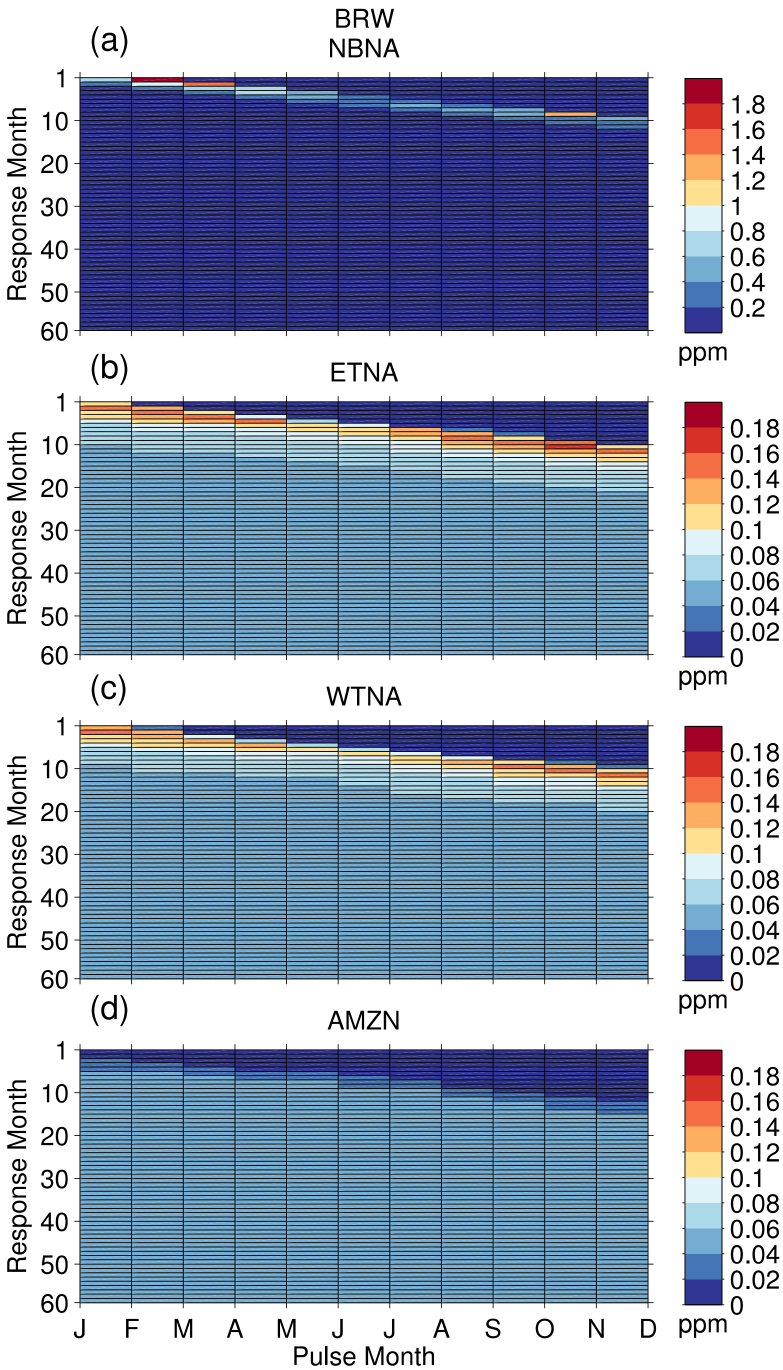

Fig. SA1: The imprints of 1 Pg pulses emitted in 12 successive months (x-axis) from (a) NBNA, (b) ETNA, (c) WTNA, and (d) AMZN on the atmosphere sampled at BRW over a 60-month period (y-axis).

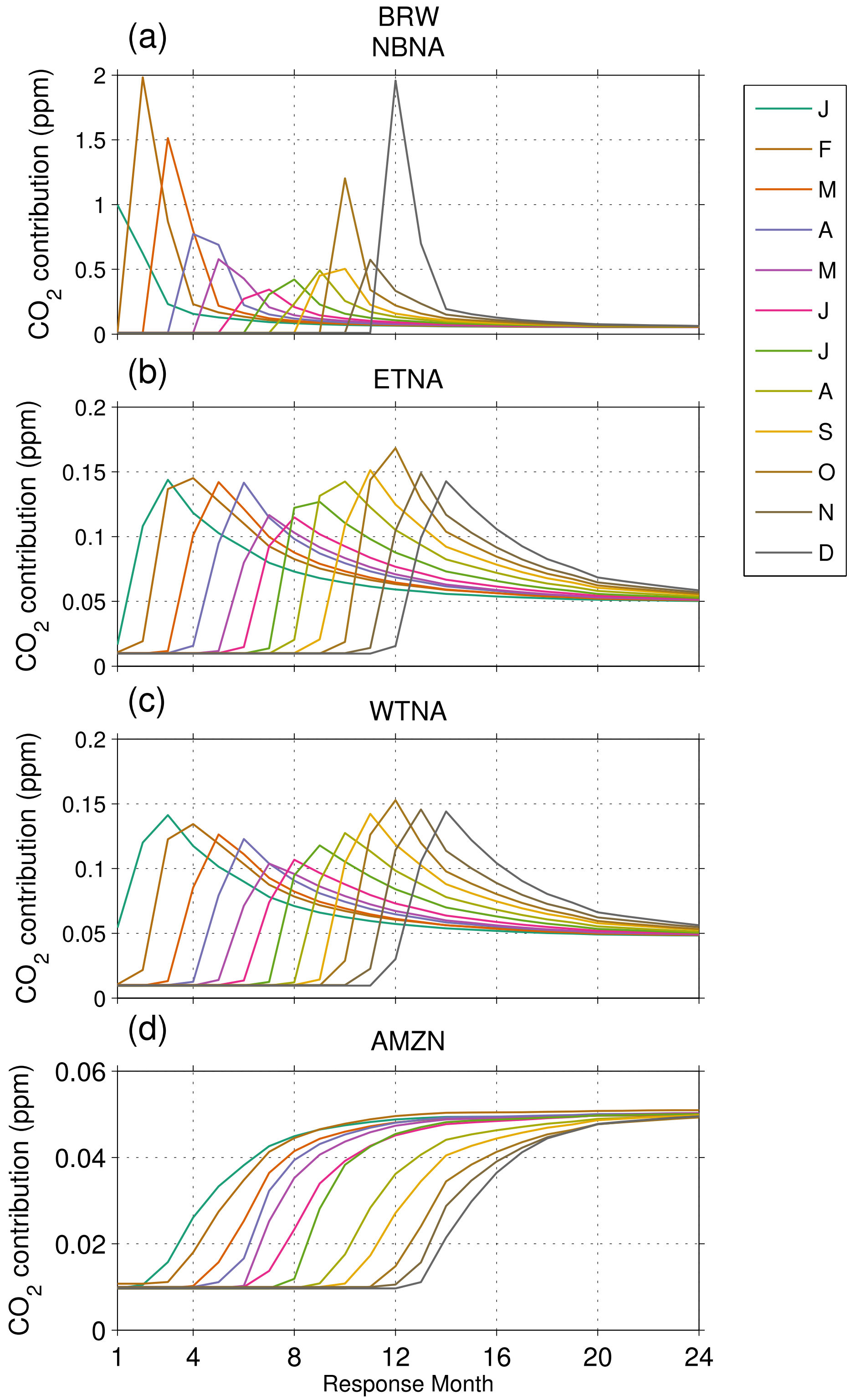

Fig. SA2: The imprints of 1 Pg pulses emitted from (a) NBNA, (b) ETNA, (c) WTNA, and (d) AMZN in each month (contours) on the atmosphere sampled at BRW.

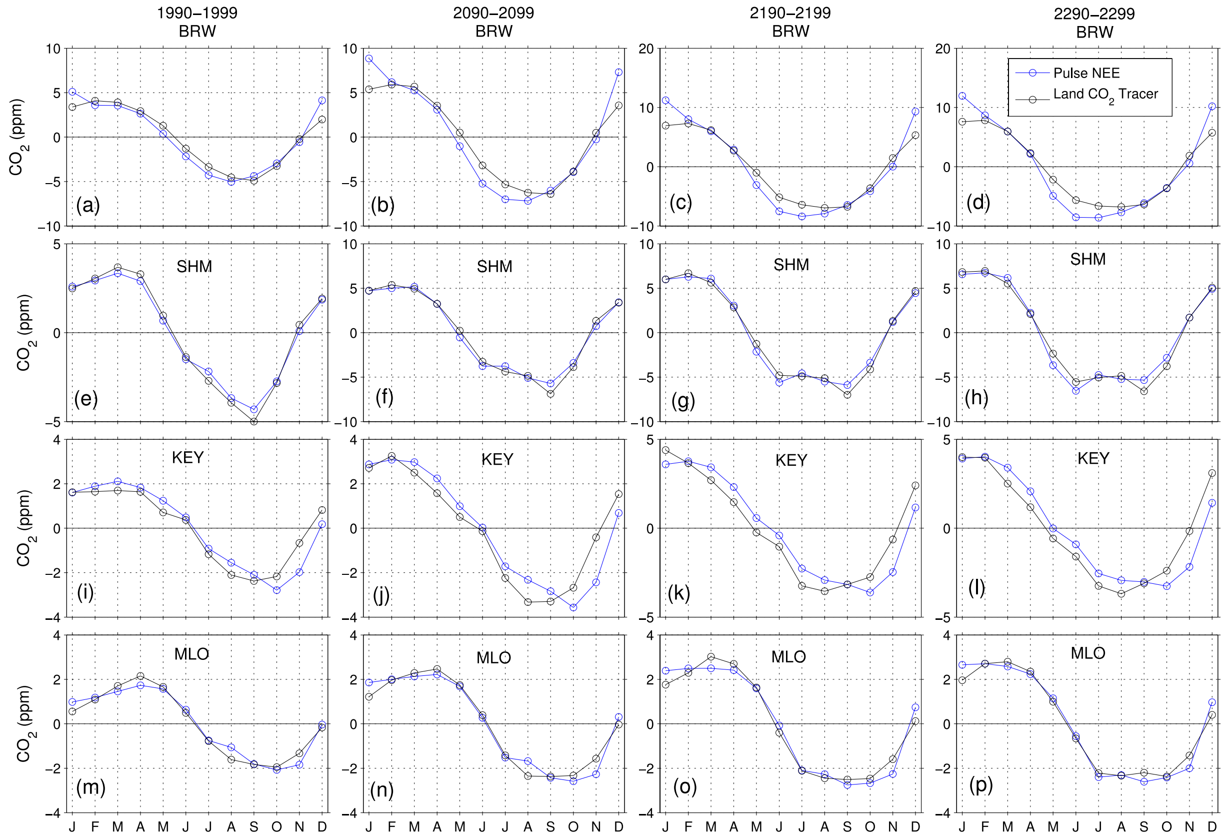

Fig. SA3*:* Mean annual cycles of atmospheric CO_2 derived from (blue curves) NEE run through the pulse response function and (black curves) the CESM land CO_2 tracer for (a—d) BRW, (e—h) THD, (i—l) KEY, and (m—p) MLO in 1990—1999, 2090—2099, 2190—2199, and 2290—2299.