The authors thank both reviewers for their helpful comments. In the following, our responses are in **bold**, and corrections to the text are in ***bold italics.***

Reviewer #1 comments:

1. Variability of ECs between fires:
I think that only listing the standard error of the fit in Table 2 (and in Figure 8) could be confusing for readers who might not have a full understanding of just how variable fires within a given biome or ecoregion can be. In order to model fires correctly (or to interpret model results) it is important to not just understand the mean emissions from the fire, but how much they can vary from event to event. I appreciated the discussion, starting at Line 3 of Page 21677, and Figure 2 showing the variability for all fires with high FRP (> 5000 MJ/s), but it is not clear if the factor-of-two geometric standard deviation given there is supposed to apply to all biomes and ecoregions, or if some are more variable than others. In addition, Figure 2 only has data for the high FRP events, and so it is not clear if that variability is consistent with the data in Table 2 and the supplemental tables. I thus think it is important to list the natural variability (perhaps as geometric standard deviation) in Table 2 and the supplementary tables. I would also suggest that the variability, not the standard error of the fit, be used in Figure 8 as that is more consistent with the approach used in the literature reviews presented there for comparison (as noted by the authors in L1-4 of page 21684). And I would suggest adding a quantitative estimate of the variability in individual fire ECs to the conclusions (L13, Page 21686) and the abstract.

We respect the reviewer's comments regarding the importance of recognizing the variability in individual fires, and indeed we tried to highlight the importance of that variability and the difference between the values we present here and those in the traditional literature, in our discussion of error and variability. However, we have kept the use of the standard error for the following reasons. First, the strength of our analysis is in the large number of observations, which enable a reduction in uncertainty in the mean values relative to in situ studies that cannot achieve similar statistical sampling. The estimates of standard error are thus an important component of our analysis. Secondly, our analysis convolves measurement uncertainties with fire-to-fire variability and thus high variability is not solely indicative of fire variability. Finally, characterization of the variability is of interest for assessing predictability of emissions from individual fires.

We have rewritten our discussion of variability and uncertainty to address both reviewers' comments regarding uncertainty and variability:

P21676 L12: "…calculated EC (grasslands).

NO_x EFs for wildfires are inherently variable, and thus it is important to distinguish between this variability and the confidence in the mean estimate of an EF for a group of fires. Traditional in situ measurement campaigns generally include the standard deviation of any calculated EF to provide an estimate of the variability between fires. The standard deviation conflates both inherent variability and also

measurement uncertainty; however, for in situ measurements of fire emissions the measurement uncertainty tends to be low and thus the standard deviation primarily reflects the variability. These standard deviations tend to be large relative to the value of the EF due to this inherent large variability. An alternative metric, the standard error, indicates the confidence in the mean value rather than this inherent variability, and is valuable when attempting to assess uncertainties associated with emission estimates for aggregates of several fires. The standard error is related to the standard deviation but also depends on the number of observations—as the sample size of observations increases, the standard error decreases while the standard deviation does not. A major strength of our analysis is in the large number of observations we are able to analyze; however, each individual measurement in our analysis is associated with a higher uncertainty than in situ measurements. This uncertainty in individual measurements is large enough that estimates of variability in our analysis necessarily conflate the inherent natural variability and also the measurement uncertainty. Therefore, for the biome- and ecoregion-scale ECs reported in this work, we provide an estimate of the standard error only, noting that there remains a large variability in ECs for individual fires.

For an example of this inherent **variability (with contributions from individual measurement uncertainty),** ***we calculate*** **ECs directly…"**

P21677 L14: "…geometric mean. ***Unfortunately, this kind of analysis requires a large number of high-FRP fires so it cannot be used to compare variability between biomes and ecoregions.*** **Another way…"**

P21677 L19: "…standard error, ***or estimate of the uncertainty in the mean EC*** **(15% or lower).** ***As previously mentioned,*** **this is because of the large number of observations that factor into each EC.** ***One advantage of the nonparametric bootstrap resampling method we use to calculate each EC*** **is that it…"**

2. Quantitative details on potential biases and errors in the methods used:
There are several points in the paper where I feel that I needed more information to evaluate the potential biases and errors in the method used to determine the NOx ECs:

L2-4, Page 21670: What is the lowest spatial resolution of the 40 pixels you use? Is there any potential or observed bias between near-nadir pixels and the most off-nadir pixels you are using?

There is no expected or observed difference between pixels in different across-track positions.

P21670 L4: "…swath; ***resolution at the largest remaining pixels is approximately 15.5 × 42 km^2.*** **We also…"**

L4-7, Page 21670: Is this cloud-clearing criterion sufficient? Do the results change significantly if you adopt a stricter or looser value? Is there also a filter for cloud optical depth?

A large number of previous studies using OMI NO_2 observations to examine surface emissions have used this criterion or similar ones (0.2-0.3 cloud fraction, or 0.5-0.7 cloud radiance fraction). In our work, we tested the effects of the same range on a subset of the global data set and found small changes to this criterion have little impact on the results; generally, decreasing the cloud fraction accepted as valid reduces the number of points available for analysis and thus increases uncertainties, while increasing the value has limited effect because the number of detected fires for which there is a high cloud fraction is very small. There was no observable effect on the mean. The OMI cloud product does not provide an estimate of cloud optical depth (in fact the effective cloud fraction that we use is a convolution of the geometric cloud fraction and the cloud optical depth) and thus we do not provide any factors specific to that property.

P21670 L7: "…row anomaly. ***Restrictions to cloud fraction (20-30% cloud fraction, or 50-70% radiance fraction) are generally accepted as good practice in studies of OMI NO_2 columns when looking at surface emissions. Small differences in this value do not significantly affect the results presented here. …*** **"**

L3-4, Page 21671: Can you be quantitative about this potential low bias? Can it explain the factor of 2-5 differences seen in Mebust et al. (2011), or is it a more minor effect?

Quantitative estimation of this effect is a study of sufficient magnitude that it would be beyond the scope of this work to provide it. As described in Mebust et al. (2011), there has been limited other study of this effect and no peer reviewed assessment of its magnitude. There is currently a discussion paper in Atmospheric Measurement Techniques Discussions by Bousserez that explores this effect, and results from that paper suggest that it is more than sufficient to explain the difference.

P21671 L3: "…and aerosol. ***The potential magnitude of this effect is also unknown and estimating this magnitude is beyond the scope of this work.*** **Regardless…"**

L1-6, Page 21672: I feel that this discussion of potential biases in the MODIS FRP is too short and too qualitative. I would include some of the more quantitative bias estimates from Mebust et al. (2011) here to give the reader a better understanding of the potential magnitude of the biases.

We intentionally limited this discussion of MODIS FRP biases to biases that might affect the relative differences between biomes, rather than the broader biases that might impact the absolute values of our results. We have added some text including a summary of the absolute biases discussed in Mebust et al. (2011) including the MODIS FRP biases.

P21671 L7: "…reference EF). ***We performed a comprehensive quantitative assessment of potential error sources, including the assumed NO_x lifetime, assumed NO_2:NO_x ratio, value used to convert from energy released to biomass burned, diurnal behavior of FRP, choice of wind height, screening of FRP by clouds and/or canopy cover, and possible diurnal behavior of emission factors. We found that, with the exception of any bias in OMI NO_2 for which we could not provide a quantitative estimate, the sources with an expected positive bias were balanced by sources with an expected negative bias. Thus*** **the source of** ***the observed*** **discrepancy…"**

P21671 L10: "…reliable. ***A more comprehensive and quantitative assessment of errors can be found in Mebust et al. (2011).*****"**

P21672 L3: "…likely small. ***In Mebust et al. (2011), we determined that the magnitude of this bias was on the order of 15-30% and that its effect on the derived ECs is counteracted by other biases. Our*** **analysis…"**

L8-10, Page 21672: Is this a valid assumption? How different were the data for the years 2009 and 2010? Why didn't you limit your analysis to 2005-2010? I think you need to justify this more.

P21672 L10: "…for years 2005-2010. ***Using comparisons of 2009 and 2010 land cover, we estimate that only about 10% of 2011 observations will be assigned an incorrect biome due to this assumption, which amounts to about 0.5% of all observations.*** **The IGBP…"**

L1-7, Page 21673: Why is there no discussion here about the uncertainty or bias in the wind fields? Again, a summary of the discussion from Mebust et al. (2011) could be included here.

We have included the wind biases in our comments responding to the reviewer's criticisms about our discussion of potential biases MODIS FRP.

L23-26, Page 21673: I'm not clear why this screening was applied, or how it might affect the results. Wouldn't eliminating large fire events bias the results? What percentage of the fire events were removed by this screening?

This screening is not intended to eliminate large fires, but instead to eliminate small fires that are close enough together that they are observed by adjacent OMI pixels, but spread out enough and small enough individually that they cannot be treated as a single fire. Our analysis aggregates the entire region together and treats it as a single fire—in some cases analyzing "individual fires" that span hundreds of kilometers. This type of aggregate violates our assumption that there is a "center" to

the fire and a single smoke plume rather than several aggregated smoke plumes. The number of situations where this is the case is small enough that we were

We qualitatively examined several of these aggregated regions and found that nearly all aggregates larger than 2 OMI pixels in the across track direction or 3 OMI pixels in the along track direction fell into this category of widespread low intensity fire.

"…across-track direction. ***Events greater than this size were typically aggregates of multiple small fires (e.g. several nearby agricultural fires) rather than individual fires. This restriction removed fewer than 1% of observations. …*****"**

L8-14, Page 21684: Isn't this hypothesis inconsistent with the results from Mebust et al. (2011) referenced on L5-7 of Page 21671? In Mebust et al (2011) your EFs were factors of 2-5 lower than the reference values, implying either a underestimate of NOx emission or an overestimate of FRP, but here you are saying that for forests you are higher than the reference values, and suggest an underestimate of FRP as a potential explanation. I may just be confused, but I think you need to clarify what you think is responsible for the differences in both cases.

As we state several times on P21684, the values for forests are higher than reference values in a relative sense, not an absolute sense. In other words, the ratio between forest and grass ECs that we calculate is higher than the ratio between forest and grass EFs from Akagi et al. (2011). Any bias in FRP that would affect the absolute difference between ECs and EFs (i.e. the difference mentioned on P21671) would be a uniform bias, while the bias that would cause relatively higher values in forests is a biome-specific bias. We hope that our additional paragraph regarding potential absolute biases in the analysis is helpful in clarifying the potential for absolute bias from MODIS FRP.

P21684 L14: "…forested regions. ***Lower FRP in forest biomes due to canopy screening effects would elevate the ECs in those biomes relative to the other biomes.*** **We estimate…"**

3. "Fuel Type" versus "Biome"
At L21, Page 21666 and elsewhere, I don't think you should use "fuel type-specific" here – I think "biome-specific" would capture your meaning better. I'd suggest similar changes wherever "fuel type" is used as a synonym for "biome," to avoid confusion. Fuel type implies knowledge of the specific fuels (species, litter vs. grass vs. wood, average diameters, moisture content, etc.) that you don't have available. Biome more accurately captures your meaning.

Instead of changing this phrase throughout the manuscript, because it is used in several places and it is not quite identical to our use of the word biome, we have

instead added a clear definition of "fuel type" as used in this manuscript. We included this after the introduction because in the introduction we distinguish "fuel type" from characteristics such as moisture or structure directly in the text.

P21675 L14: "…"ecoregion-scale" EC). ***We note that the phrase "fuel type" as used throughout this work is intended to distinguish across-biome variability in fuel composition (e.g. forest vs. grassland fuels), not within-biome variability in characteristics such as fuel moisture, relative contributions of leaf litter vs. woody materials, etc.*** **Fire biomes…"**

Other Comments:
L7, Page 21667: "empirically measured" – as opposed to theoretically measured? More seriously, I'd suggest a short discussion here how NOx EFs are calculated based on in situ observations (and giving the units of EFs, g NO/kg DM burned).

We have clarified the language in this sentence. In the interest of space, we have neglected to add a description of how EFs are calculated from in situ observations because this information is not required to understand the manuscript.

P21667 L7: "…an emission factor (EF ***in g kg^{-1} biomass burned*****) derived from in situ measurements of smoke is multiplied…"**

L19-20, Page 21667: I am not clear if you are arguing that, all other things being equal, NOx EFs are not necessarily correlated with MCE and fuel N content, or just that all other things are rarely equal, and so it is hard to observe the correlation?

We are not trying to argue that EFs are not correlated with MCE and fuel N content. We are expressing the lack of quantification and clarity around these relationships in the scientific record.

P21667 L25: "…limited*, leading to a lack of understanding regarding the extent of these relationships.* Observed…"

L28, Page 21668: Do you really mean "simultaneous" estimations of energy and pollutant emissions, or near-simultaneous? How near? How far apart in space and time do the OMI NO2 and MODIS FRP observations need to be for your EC method to no longer work?

The OMI and MODIS pixels are collocated in space. The difference in time of observation between OMI and MODIS observations used in this analysis ranges from 7 to 22 minutes. MODIS observations are taken first, so under typical wind speeds the OMI observations will include smoke emitted at the time of the MODIS observation. Given the spatial resolution of the OMI pixels, the time difference of

the two observations (which is forced to remain within a narrow window by the satellite paths) is unlikely to play a role in our ability to measure these emissions.

P21668 L28: "…allowing for *near*-simultaneous estimation…"

L5-7, Page 21671: As this paragraph conflates biases in both the OMI NO2 columns and the MODIS FRP, maybe it belongs more in Section 3?

This paragraph is intended to indicate that a potential bias in the OMI NO_2 columns is the most likely source of the discrepancy. The discussion we have added in response to previous comments that mentions the other biases clarifies this issue so we will not make any further modifications.

L15-16, Page 21671: Please include the equatorial overpass time of Aqua, and be more specific on the "near-coincident" measurements – how near in space and time do they have to be?

P21671 L15: "…column densities. *The Aqua equatorial overpass time is ~1:30pm (local time), meaning OMI and MODIS measurements are typically made ~15 minutes apart.* Fires are…"

L27-28, Page 21675: Were these three points unique in any way? Can you say why they were outliers other than their impacts on the fit?

The reasons for these outliers, beyond statistical randomness, were not apparent.

L2, Page 21679 and elsewhere: How do you explain the low Rˆ2 values of ~0.1? Does this suggest your approach to assigning ecoregions is conflating two different fire types, or do you believe that there are cases where FRP is uncorrelated with NOx emissions?

The extremely low (<0.2) R^2 values occur in ecoregions where the fires cover a narrow range of low FRP values. In these cases, the spread in x is small relative to the high uncertainties in y, obscuring the correlation.

L1-2, Page 21680: Why do the shrubs show higher FRP? They have more fuel loading than grasslands, but why are they higher than forests? Is this evidence for canopy shielding of FRP, or do the forests just burn more slowly, or what?

This phenomenon has not been studied. In relation to forests, shrub fires may burn at a hotter temperature, or may spread more quickly thus covering a larger area,

both of which could produce a higher FRP. Canopy shielding for forest fires may also play a role.

P21680 L2: "…forest fires). ***The reasons for higher FRP in shrub biomes is not clear, but may be related to the higher fuel loading as compared to grass fires, and the potentially higher temperature combustion and/or faster rate of spread as compared to forest fires.*****"**

L9-11, Page 21683: Why do the shrub fires emit more NOx per unit energy? Higher fuel N content? Different combustion conditions?

P21683 L11: "…forest fires. ***The reason for higher shrub NO_x emissions is unclear, but may stem from higher MCE, or from N deposition due to nearby anthropogenic emissions, an effect that is lessened in our analysis due to our removal of points with high NO_x background (Laursen et al., 1992).*** **In both…"**

Typos and Minor Edits:

L17, Page 21666: "mean fire emissions of NOx per unit energy" would be clearer.

Change made.

L18-21, Page 21666: These two phrases appear to say the same thing, possibly because it's not clear what "this range" in the second phrase refers to. Did you mean the range of biome mean ECs (0.250-0.362 g NO/MJ)?

P21666 L20: "…outside ***the range of all mean biome ECs*****, implying…"**

L26, Page 21666: I understand the focus of this paper is NOx, but since you mention aerosols as a direct climate impact, I think you should also mention their indirect impacts via clouds.

P21666 L26: "…influencing the chemistry ***or physics*** **of climate forcers, e.g. nitrogen oxides (NO_x = NO + NO_2) and CO acting as ozone (O_3) precursors** ***or aerosol indirect impacts on clouds*** **(Bowman et al., …"**

L2, Page 21667: The US EPA air quality standard for NO2 is pretty high (100 ppb for 1-hour, 53 ppb annual average), so I think it's not quite right to say the NO2 from fires has a direct health impact – if you were breathing smoke at that concentration, your health would likely be impacted, but I think NO2 exposure would be the least of your worries. I'd leave it at O3 and aerosols.

Change made.

L5, Page 21667: remove comma after "emissions."

We opted to change "and" to "as well as" rather than remove the comma.

L9-10, Page 21667: I'm not sure it's that they are "simpler to measure or estimate", it's more that you can measure burn area, and by combining that with an estimate of fuel loading and combustion completeness you can get an estimate of biomass burned. I'd suggest rephrasing this.

We removed the "simpler to measure or estimate" clause.

L12, Page 21668: "span the globe, have full annual coverage" is not strictly true if we interpret it to mean global, continuous coverage. Can you rephrase this to be more narrow and accurate, like "high spatial and temporal coverage" or "daily observations at high spatial resolution?"

P21668 L12: "…globe, *sample throughout the year, and include* many fires"

L13-15, Page 21669: I don't see what these two sentences add to the paper, so I'd suggest cutting them.

While the first sentence is not strictly necessary, it frames the coming paragraph. The second sentence is necessary to highlight the time coverage of our analysis. We have left these in.

L22, Page 21670: "few NOx sources" – did you mean few anthropogenic sources, few continuous sources, few large sources, or what?

P21670 L22: "…few *other* NO_x emission sources"

L15, Page 21672: Add refs for "common EF schemes"

P21672 L15: "…tropical, temperate, and *occasionally* boreal forests *(Andreae and Merlet, 2001; Hoelzemann et al., 2004; van der Werf et al., 2010; Akagi et al., 2011)*;"

L13, Page 21673: Coincident or nearly coincident?

P21673 L13: "…coincident *in space and from the same orbit (ranging from 7-22 minutes apart).* OMI…"

L23, Page 21675: Add ref for bootstrap method.

We have added the following reference:
Efron, B., and Tibshirani, R.: Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy, Stat. Sci., 1, 1, 54-75, 1986.

L1-2, Page 21676: Seems like this sentence fits better after Line 25 of page 21674.

Change made.

L4-5, Page 21678: You might consider rewording this, as it requires several readings to make sense of right now (I thought it was a typo on my first read through). Perhaps start by saying that 3 forest ecoregions spanned two biomes, and so you end up with 45 ECs for the 42 ecoregions.

P21678 L4: "…distinct ecoregions. ***Our analysis identified 42 separate ecoregions. Three of the forest ecoregions spanned multiple biomes (e.g. tropical and temperate forests) so we calculate two ECs for each of these three ecoregions; thus we calculate 45 ecoregion-scale ECs in all.*** **Maps of…"**

L12-13, Page 21678: I know what you are saying, but the phrase "statistically significantly different" is very awkward (and statistically is misspelled the second time). I think, since you say "at the 0.05 level" later, you can cut the word "statistically" in both places.

Change made.

L18-24, Page 21680: "These ECs" on L18 – it is not clear to me if you are referring to the biome or ecoregion ECs, or both. It is also not clear what "range" you are referring to on L19. Finally, on L24, please clarify which EC you mean when you say "within the standard error of the EC."

P21680 L18: "…biome-scale EC. ***These "biome-similar" ecoregion ECs fall within 32% of the mean biome EC for all biomes except agriculture; within some biomes, the difference between the largest and smallest "biome-similar" ecoregion ECs is as low as a few percent.*** **Differences in…"**

P21680 L25: "…range, and ***for*** **9 more** ***ecoregion-scale ECs the difference between the EC and this range is less than the standard error of the ecoregion-scale EC.*** **However, …"**

L1-2, Page 21681: I think it would be clearer if you used the phrase "standard error of the fit" or "standard error of the mean EC" instead of just "standard error" both here and throughout the paper.

P21681 L1: "…range within the ***ecoregion*** **standard error** ***of the EC*****, or even twice the** ***ecoregion*** **standard error** ***of the EC*****…."**

L9, Page 21681: I'd suggest adding "if the mean is used" to the end of this sentence after "for these specific regions."

P21681 L9: "…regions ***if the mean EC is used.*** **The…"**

L16, Page 21682: I'd suggest adding "using the updated method" to the end of this sentence after "indicated above." This would make clear that you are not just reproducing a plot from Mebust et al. (2011), but that this is a revised analysis.

Change made.

L24, Page 21684: Remove "e.g."

Change made.

Table 2: In either the caption or as a footnote to the table, make clear that the uncertainty in the second column is the standard error of the fit for the mean EC, NOT the variability between fires, and add an estimate of the variability between fires to the table.

We added the following sentence to the table. For the reasons given above, we have not included an estimate of the variability.

"*Uncertainty for ECs and EFs is given as the standard error of the EC (σ).*"

Figure 1: Describe what is meant by "other" in the figure caption, as you do in the text.

Figure 1 caption: "…"mixed fuels". *Fires are labeled as "other" if at least 75% of measured FRP came from fire pixels not assigned a biome type.*"

Figure 2: In both the caption and on the x-axis of the figure, make clear these are ECs of NOx expressed as g NO. Right now if this figure were cut from the paper it would not be clear what species is being estimated.

Figure 2 caption: "…ECs *of NO_x (as NO)* measured…"

The x-axis label of the figure was also changed. Please see new figures at the end of this document.

Figures 3-6: In the captions, where you say "range of mean biome ECs", you mean the standard error of the fit for the mean biome EC, correct? If so, I'd use that more specific language to avoid confusion between standard error of the mean and variability between fires.

We have added lines to each figure to indicate the mean biome EC for that particular biome. Please see the updated figures at the end of this response. The grey area refers to the range of all mean biome ECs. We have clarified this text in the Figure captions as well:

Figure 3 caption: "…(e.g. Region E). ***The tropical (solid), temperate (dashed) and boreal (dotted) forest mean biome ECs are indicated by the black horizontal lines.*** **The range of** ***all*** **mean biome ECs…"**

Figure 4 caption: "…In (b), ***the mean grassland biome EC is indicated by the black line, while*** **the range of** ***all*** **mean biome ECs…"**

Captions for Figures 5 and 6 were changed in a similar manner to Figure 4.

Reviewer #2:

p21666, Line 20-21: "and a similar number of ecoregion ECs falls outside this range". I do not understand what this means. What range are you referring to?

In response to Reviewer 1 we have corrected this statement to clarify: "the range of all mean biome ECs".

Section 4.1: This section needs to be rewritten to improve clarity. There is an attempt to explain the difference between standard deviation and standard deviation of the mean, which in my opinion should be more concise (maybe include some references here?).

We have substantially rewritten this section to address comments by both reviewers. Please see the first response to Reviewer #1.

p21677 line 5, "we note the standard deviation is the square root of the variance" is not necessary, since all readers are familiar with this basic concept.

We have removed this clause.

Also, p21677 line 20-23 is a repetition of p21676 line 20-25

We have rewritten this as part of the above rewrite and now the explanation of the standard error's dependence on the number of observations is only included once.

Section 4.2: The authors discuss the comparison between specific ecoregion ECs and the mean ECs for the corresponding biome, but in Figure 3 the grey band indicate the range for all mean biome ECs (it is indeed the same for all Figures 3-6). It is a bit odd to present results this way. The mean ECs for each type of biome should also appear on Figure 3-6, since that is what is discussed here.

We have added a line to indicate the mean EC for each biome (see also our response to Reviewer 1). The updated figures can be found at the end of this response.

Technical corrections:

p21666, Line 19: p has not been introduced before, and you should avoid to use it in the abstract.

P21666 L19: "…significantly (*with 95% confidence*) from…"

p21676, Line 10: R needs to be defined.

"…(N), and *the coefficient of determination (R^2).* Calculated…"

p21678, line 9: p should be defined

P21678 L9: "…p values, the probability that each ecoregion EC is the same as the mean biome EC, for statistical…"

p21678, line 13: statistically

We appreciate the technical spelling comment; we removed this word in response to comments from Reviewer 1.

p21681, line 6: "moderate number of observation rather than a small number". "Moderate" and "small" are subjective. Rather say something like: "Given the number of observations used for ecoregions that deviate from the mean ECs, these results can be considered statistically robust".

P21681 L6: "…type-specific. *The number of observations within each ecoregion that deviates significantly from the mean biome ECs is large enough to confirm that these differences are statistically robust and not the product of anomalously small sample size.* The large…"