Commenter I

Nitrogen Management Project Protocol
Minimum Data Standard Public Comment Period

Reserve Technical Questions 2:

(a) Relatively recent published studies/meta-analyses of potential interest include:

- Philibert et al. (2012). Quantifying Uncertainties in N_2O Emission Due to N Fertilizer Application in Cultivated Areas. PLoS ONE 7(11):e50950. doi:10.1371/journal.pone.0050950.
- Kim et al. (2013). Linear and nonlinear dependency of direct nitrous oxide emissions on fertilizer nitrogen input: A meta-analysis. Agriculture, Ecosystems and Environment 168 (2013) 53– 65.
- Parkin et al. (2012). Calculating the Detection Limits of Chamber-based Soil Greenhouse Gas Flux Measurements. J. Environ. Qual. 41:705–715.

(b) Researchers at MSU-KBS (I. Shcherbak, N. Millar, and G. Philip Robertson) have conducted a global meta-analysis to provide a first quantitative comparison of N_2O emissions for studies of multiple N rates (at least two non-zero) otherwise identical. This includes 78 studies containing 231 site years of data, the largest such data-set assembled. Results show that a nonlinear emission factor better represents with lower uncertainty global emission patterns. The manuscript was submitted in early January 2014.

Reserve Technical Questions 3:

(a) No, they should not be considered equal. Of course data-sets in peer-reviewed journals are of varying quality too, but in general they will have passed a substantially higher bar for publication in the public domain. An alternative to peer-review publication may be peer-review by an expert panel that includes soil trace gas experts and a statistician, preferably university based. It is impossible to be completely prescriptive with respect to sampling diverse ecosystems - review must often fall back to expert judgment.

(b) No, because of the difficulty of generalizing what is high quality data, other than through publication following scientific peer review, or review by expert panel as noted above.

I appreciate that the Reserve are attempting to ensure a high bar for data acceptance. However, I would ask the question why, if the data is of high quality, has it not been submitted to or published in a scientific peer-reviewed journal, and therefore made available for the wider scientific good. Examples the Reserve mentions (USDA and ARB funded studies) have or will likely be published in peer-reviewed journals.

Given the Reserve’s current rigorous data standard requirements for field experimentation, I suggest that organizations other than universities or those working in direct association with universities will find it burdensome to generate acceptable data-sets. In fact a number of studies by experts in the field may not be completely compliant with these criteria.

(c) As noted above, statistical rigor as certified by a peer review panel that includes a statistician is vitally important. Also, the reporting of negative results to remove bias – e.g. when/where a nitrification inhibitor did not work compared to when/where it did.

Other criteria (as required with peer reviewed publications) should potentially be used for screening. This could include identification of personnel carrying out the work, ensuring data has not been previously published, identifying conflicts of interest (real or perceived) and requiring transparency in relation to the individuals and organizations' source of funding.

Reserve Technical Questions 8:

Regarding the appropriateness of a "given region"; USDA Land Resource Regions (LRRs) are designed "as a vehicle for extrapolating research results across political boundaries" (USDA Handbook 296, 2006). LRRs are a group of geographically associated major land resource areas; locations within share common physiography, geology, climate, water, soils, biological resources, and land uses. This logic underpinned the acceptance of the extrapolation of the results from the Michigan based field study (Hoben et al. 2011) to the North Central Region.

It may therefore be valid and appropriate to extrapolate results (such as emissions factor equations) from a single study at a small number of locations in a LRR to that entire LRR, given that the study investigates representative management practice, includes representative crop(s)/rotations and conforms to other criteria.

Given the paucity of data, it may also be appropriate to consider a (semi) agnostic approach to crop type, by identifying major crop groupings (e.g., forage, small grain, N-fixers, etc.) and allowing emissions data from an individual crop within each group to be representative of all crops within the group within a LRR.

There is no scientific justification for breaking eligibility into specific crops x LRRs. A recent analysis of the Stehfest and Bouwmann data set conducted during the ACR validation of the MSU-EPRI methodology showed that crop type does not matter when defining a minimum emissions factor. To restrict an EF to an individual crop within an LRR would introduce a large degree of stochasticity that would reduce overall confidence.

(a) I suggest one study over two years with three sites to represent an entire LRR.

(b) I believe LRRs are the most appropriate region at this time. Given that these are the largest region discussed; when they are characterized by a suitable study these numbers would be appropriate for smaller regions within the LRR (e.g., MLRAs).

(c) This will depend on the design of the study. The question highlights why a statistician should be an evaluator on the expert panel.

Reserve Technical Questions 15:

On a more general note, for clarity I would suggest some minor revisions to the language used to describe the different outlier types.

Type 1. "A N_2O concentration value from an individual gas sample collected during a chamber closure period and used along with other gas samples collected from the same chamber during the same closure period to estimate N_2O flux rate from that chamber", or similar.

Type 2. "N_2O flux rate data calculated from the change in N_2O concentration over time ($\delta c/\delta t$) during a single chamber closure period", or similar.

Type 3. N_2O flux rate data calculated as an average across multiple chambers in a single plot or field or individual chambers in multiple plots that represent a treatment", or similar.

Extending this principle, you could imagine a Type 4, in which N_2O fluxes from all treatments measured at a particular time can be considered 'outliers'. This could potentially be due to technical issues or anomalous environmental conditions on that day that raised or lowered fluxes wholesale.

(a) Here I assume you are primarily referring to current outlier types 2 and 3 in D2.3? If so, I agree with the current thinking, i.e., to NOT allow outlier exclusion based purely on a statistical test (e.g., 3 standard variations from the mean). More contextual evidence should be provided to justify the exclusion(s).

(b) I am uncertain if this flux rate change refers to a single chamber/replicate/treatment at a single time (hour/day) and/or across an extended time period (growing season/year).

However, in general the exclusion of a flux measurement that cannot be attributed to technical issues, and that is based on a single arbitrary value seems counter to the Reserve's current approach of inclusion. Ideally, I believe all justified values should be included, irrespective of the flux change (absolute or %) consequences of their inclusion. The relevant offset methodology should contain calculation tools that accommodate this variability, even if this means less credit reward. This is the result of inherent natural system variability that unfortunately is for the most part beyond management control.

I would also suggest that if exclusions are allowed, that in effect they will remove an anomalously high, not low, flux, resulting in a reduction in the estimated flux rate, as opposed to an increase. Current common methods of gas concentration determination also do not favor the consistent reporting of very low positive or negative fluxes. These potential flux reductions need to be treated carefully. An arbitrary threshold value may allow inclusion of a high flux from a 'control' treatment, but disallow a similar anomalously high flux from a management practice treatment, resulting in an arbitrarily high flux difference between control and treatment and potentially generating 'extra' offset credits. Of course the opposite could also occur, with the generation of artificially low offset credit numbers. These issues may be minimized if an appropriate variability tool is in place that reduces offset credits based on increased variability.

(c) Please see above answer. Again an arbitrary value may be best avoided in favor of quantification tools in the methodology that accommodate variability. If a value must be put in place, then I would suggest one that is (much) higher than +/- 10% of the annual N_2O flux. As an example of the alteration of average daily N_2O emissions calculated over multiple years based on the exclusion/inclusion of one daily flux estimate during this period, please see Robertson et al. (2000) Science 289, 1922-1925; Fig. 1. Here, the removal of a single day of anomalously high fluxes in the no-till and low-input systems in 1999 and 1991, respectively, reduced the average daily fluxes from those treatments measured across nine years (1991 – 1999) by ~ 76% and 44%, respectively (See Supplemental Table 1, footnote 1 and 2).