Supplementary information for Reply to Wassmann et al:

More data at high sampling intensity from medium- and intense-intermittently flooded rice farms is crucial

Our fundamental message (1) is that under intense forms of intermittent flooding, the focus of mitigation should shift from methane (CH\(_4\)) to nitrous oxide (N\(_2\)O) which is a long-term climate forcer. This must be taken seriously given that alternate wetting & drying (AWD) is being promoted to reduce CH\(_4\) (see SI Figure S1).

SI Figure S1

Legend to SI Figure S1. General understanding of climate impacts of rice cultivation under continuous flooding and AWD (also called multiple aeration) (FAO’s Climate smart practice brief on rice, R&S, 2014) (2) compared with results from two studies (1, 3) showing high nitrous oxide emissions under multiple aeration. These new results highlight that N\(_2\)O, not CH\(_4\), can be the dominant greenhouse gas emitted by rice farms under intense forms of intermittent flooding.

Our original study (1) was recently critiqued by a group of scientists from International Rice Research Institute (4) who have been instrumental in advocating AWD over the past decade.

We posit that many farms using AWD might be under mild-intermittent flooding regimes that we have proposed would minimize both methane and rice-N\(_2\)O. However, the global community cannot assume that farms practicing AWD are necessarily under mild-intermittent
regimes and more research is needed to map flooding regimes and measure rice-N$_2$O.

We are glad to note several areas of agreement between the perspectives presented by Wassmann et al and our conclusions.

- A neglect of N$_2$O emissions will distort estimates of overall climate impact of rice cultivation.
- We also welcome inclusion of daily water levels from rice fields as a part of Measurement, Reporting and Verification (MRV) in future mitigation programs.
- The global community should take up initiatives that allow for comprehensive observations of water levels by advanced sensor and/or remote sensing technologies.
- We appreciate the call by Wassmann et al to improve GHG calculation tools such that they include rice-N$_2$O emissions in their calculations. While China, India and Indonesia (5-7) mention that rice farms can produce N$_2$O and/or refer vaguely to using country specific EFs they do not provide EFs for rice or clarify what fraction of soil N$_2$O, if any, came from rice. If these countries used an average rice-N$_2$O EF of 0.3% from IPCC guidelines (8), it is tens of times smaller than our highest EF (15-31%). This just reinforces our assessment that the potential of high rice-N$_2$O should be a consideration in the inventories as well as GHG emission rate estimation pathways (e.g., Cool Farm Tool).
- We are also in agreement with Wassman et al that GIS-based suitability mapping for AWD is a good place to integrate the analysis of high rice N$_2$O in a given region.

This supplement aims to respond at length to specific technical issues raised in the supplement (available through this page) to the critique by Wassmann et al. (4). The original critique by Wassmann et al is presented below in italics with our responses following.

**Field design**

*Wassmann et al.: In the first section in the supplement, Wassmann et al state that “….while non-standardized treatments differing across farms seems straight-forward for capturing business-as-usual (baseline) conditions in rice farming, it is unclear how this field design can lead to statements on recommended mitigation practices, namely the stimulating effect of alternate wetting and drying (AWD) on N$_2$O emissions.”*

As pointed out by Wassmann et al themselves (see their Reference S10), our study is not the first one to indicate that AWD (or intermittent flooding) stimulates N$_2$O emissions. Many studies including many coauthors of Wassmann et al have shown that AWD conditions trigger N$_2$O while decreasing methane emissions. Our study has simply increased the total range of N$_2$O emissions possible under intermittently flooding regimes. As mentioned in the main text, it is not necessary to compare intermittently flooded farms to continuously flooded farms as controls to show co-relationships between different kinds of intermittent flooding and rice-N$_2$O.
The notion that these water regimes correspond to the currently promoted mitigation strategies, namely AWD, deserves caution. By definition, AWD is an intentionally managed practice in which “irrigation water is applied a few days after the disappearance of the ponded water” (S1). In many parts of the world it is difficult for farmers to maintain a rice fields under flooded conditions. This could be for various reasons including high percolation of soils, inadequate rainfall, limited access to irrigation water, etc. But in any case, these inherent fluctuations in floodwater tables should not be confused with AWD.

We understand that Wassmann et al. are defining alternate wetting and drying (AWD) as a specific mitigation strategy whereas we have defined AWD as intermittent flooding that includes multiple aeration, i.e., alternating dry and wet periods regardless of the cause (active drainage and/or percolation). However, we do not think that our results are high and anomalous because of low clay content in soils at our study sites. A previous study has shown high rice-N\textsubscript{2}O under deliberate AWD in clay-rich soils (3). Our conclusion is that regardless of how one defines AWD, unless we have very precise maps of floodwater levels, the global community is at risk of ignoring high N\textsubscript{2}O emissions. This is because multiple organizations are advising farmers in many parts of Asia to follow AWD. Our field experience suggests that AWD might be getting promoted without complete or uniform understanding of both definition of AWD suitability and N\textsubscript{2}O tradeoffs.

The high emission rates in this study are associated with high application rates of N fertilizer (>200 kg N per ha) which is clearly above recommended values. N\textsubscript{2}O emissions were substantially reduced as long as N applications have been lowered or omitted. Although these emissions are still high (2.5 – 11 kg N\textsubscript{2}O/ha) as compared to previous studies, the relative decrease corroborates the strategy of combining recommendations on water management with those on fertilizer management – an approach that is integral part of mitigation efforts in the form of “AWD+”.

Recommended or not, our baseline nitrogen use rates are based on the actual farmer surveys. Most crucially, we have already pointed out (9) that while nitrogen availability is necessary for N\textsubscript{2}O production, high N rate was not central to high rice-N\textsubscript{2}O. Rice-N\textsubscript{2}O was very low at two farms where N rate was very high (Farms 2 and 4). When farms are flooded and soil-oxygen content is low, either ammonia doesn’t nitrify (and no substrate is available for denitrification) or N\textsubscript{2}O converts into N\textsubscript{2} (complete denitrification).

The data on soil physical properties and water input indicates that most of the sites selected for this study were very leaky in nature. The clay content of all five farms is 6- 25%. Four out of five sites have a clay content of less than 20%, with an exceptionally low clay content (6%) in Farm 2. The leaky nature of the soil is also reflected on page 6 of the main manuscript where it is mentioned that “The daily water levels represent a snapshot because they dropped quickly (4–15 cm within 24 h) after irrigation”. In general, paddy growing soils (puddled) have a hard pan and percolation rate varies from 0.25 mm/day (S2) to 3.9 cm/day (S3).
We pointed out above that we are including situations with water loss whether through percolation or surface drainage. Farmers often grow rice in regions that are not optimum for rice cultivation which supports our case for more precise measurement of water levels/regimes to improve our understanding of Net GHG emissions from global rice cultivation.

_Wassmann et al._: The supplementary material contains conflicting statements- on page 49, the water use for farm 2 is unbelievably high (9543 mm) for baseline practice while on page 9-10, under Text section 7, it is mentioned that “for farms like Farm 2.... where water likely doesn’t percolate down quickly”. If the percolation rate was low in Farm 2, the water input should not be so high.

Thank you for pointing out this inconsistency. It is correct that Farm 2 actually did not have slow percolation rate. After our experiences at Farm 1 in year 2012 and Farm 3 in 2012-2013, we had indeed used extraordinary amounts of water to try to keep Farms 2 and 4 flooded. Farm 2 BP plots and Farm 4 baseline (BP) plots were irrigated a total of 56 and 55 times, respectively, during one season. We have removed the phrase “where water likely doesn’t percolate down quickly” from our online recommendations.

_Wassmann et al._: The authors state that AWD “includes allowing water to drop down to 15 cm below soil level and roughly corresponds to our medium-intermittent flooding regimes.” However their definition of medium-intermittent flooding is broad; ranges from 0 to 5.5 flood events per rice season. Moreover, in the section “The risk of enhanced rice N₂O....” the authors infer that any form of intermittent flooding is being promoted as an approach to reduce CH₄. However, we are not aware of methods to reduce CH₄ that correspond to “intense-intermittent flooding” as described by the authors.

We agree that there are a large range of water indices and flood events for each intermittent flooding category. As stated in the main text, many AWD farms might be under mild-intermittent flooding regimes that minimize both methane and rice-N₂O. See SI Figure S2 for multiple ways in which water management can lead to mild-intermittent flooding regimes. While research organizations might not recommend intense intermittent flooding, it is clear to us that rice farms can end up having intense intermittent flooding conditions because of practical issues. Thus, it is crucial to map flood regimes at farmer-managed farms and measure N₂O emissions rates at a variety of flooding regimes.
Legend to SI Figure S2: Different mild-intermittent flooding regimes: Mild-intermittent flooding regimes can look very different in diverse geographies with different percolation rates as long as other defining criteria (water index and number of flooding events) are met.

**Wassmann et al.:** In the two farms with high N₂O emissions, fields have been without any standing water for at least half of the season (Supplement p. 26; p. 28/29). Especially in the 2013 experiment of Farm 3 (Supplement/p. 29), it seems implausible that the depth of the soil water did not trigger any yield penalty.

We have clearly noted yield penalties in our Table 1. We have also already pointed out that more research is needed to minimize climate impacts per unit yield (see also SI Fig. S38 in the initial study(1)).

**Wassmann et al.:** In the farms with very low emissions, the term ‘intermittent irrigation’ is misleading as water levels have only been replenished once those have subsided to the soil surface. It seems questionable if the term ‘intermittent irrigation’ is appropriate for this practice. Soil conditions under such treatment without flooding will be comparable to ground cover rice production systems – a practice that is known to enhance N₂O emissions while overall emissions will still be lowered as compared to continuous flooding (S4).

We have never used the phrase “intermittent irrigation” in our paper. We assume the comment is meant to reference “intermittent flooding” and as mentioned above, regardless of how one defines intermittent flooding, when farmers grow rice under non-continuously flooded conditions, high N₂O fluxes are possible under intense forms of intermittent flooding. We do not agree that the increase in GWP due to increase in N₂O will always be less than decrease in GWP caused by
decrease in CH$_4$ emissions. In other words, net GWP due to intense forms of intermittent flooding practices can be higher than that of baseline practices that involve continuous or mild-intermittent flooding.

Wassmann et al.: Important considerations of AWD or intermittent wet and dry practices are intentionally managed practices that take into consideration N fertilization, crop stage and other factors. (NOTE: In spite – or because of – the very voluminous supplementary material, we could not find the dates of N fertilization in relation to flooding events.) In the field experiments of this study, it seems that water levels have only been replenished once those have subsided to the soil surface. This water regime was apparently caused by soil conditions (low clay content) and cannot be deemed an advanced irrigation practice such as AWD.

Timing of addition of fertilizers (organic or inorganic) is mentioned in Tables S4-S9 and is also communicated through solid red lines in SI figures 2 and 3. This comment is related to the different definitions of alternative wetting and drying. As noted above, we are including drying through percolation in our definition of AWD and the clay content is immaterial.

**Sampling frequency**

Wassmann et al.: The study suggests – at least implicitly -- that previous studies have missed recording high N$_2$O emissions due to the insufficient sampling frequency. This argument is unjustified given the high number of field observations and the diversity of sampling strategies applied. Several studies with manual sampling comprise high frequencies, e.g. in daily intervals (2). Moreover, automated measurements provide continuous measurements of N$_2$O emissions in sub-daily intervals from rice fields including those in tropical Asia (3, 4, 5, 6).

These research-station based studies indeed do have high sampling frequencies but they are not relevant because they were performed under conditions which should not trigger high rice-N$_2$O, i.e., continuous (10, 11) or lowland flooding (12), no waterlogging (post-rice-harvest dry period) (13) or mid-season drainage (a form of mild-intermittent flooding) (14). In contrast, we have called for high frequency measurements from medium- or intense- intermittent flooding regimes, which could be common in farmer-managed farms.

Wassmann et al.: Field sampling was done by closed chamber measurements covering 40-60% of all days of a season for a 45-minute deployment period per measuring day (NOTE: There is a discrepancy in the study between GHG sampling density of the sites with high N$_2$O emissions (33-44% of all days) and the other farm sites (49-64% of all days)). While the authors call their method “high intensity measurements” this seems only justifiable in comparison to the studies cited in the text, but not against the backdrop of a large number of other studies.

We would like to clarify that we did not ever have a chamber deployment period of 45 minutes
but only 30 minutes. Please refer to Tiwari et al (15) for all details of our methodology. Wasserman et al make a good point that that there are other high frequency measurements in the literature but as previously noted, these do not include flooding regimes which would be expected to promote N₂O emissions. Our point is that high frequency measurements are needed to fully evaluate net GHG emissions under the range of water management regimes likely to be used by farmers. Regardless of sampling frequency, we did not use broad peaks (see below).

**Wassmann et al.:** Our letter mentions only a few of these studies and this list could easily be expanded by manual sampling studies in a similar intensity as in this new study (e.g. S5) as well as automated field records from different parts of the world, e.g. Japan (S6, S7, S8). Automated measurements have also been conducted in rice fields with unstable floodwater levels (6). In spite of well-documented records of short-term emission spikes in the literature, none of these measurements have ever recorded cumulated emissions in the magnitude as reported in the study by Kritee et al. Continuous and high-resolution measurements with automated systems clearly show the dynamic nature of N₂O emissions. Due to the short-term nature of these emission spikes, however, these events did not lead to large enhancements in the overall emissions.

The additional studies cited in the supplement (S5 to S9) do not include flooding regimes where high N₂O are expected. Our study points out the risk of high nitrous oxide emissions from medium- or intense-intermittently flooded rice farms.

As noted previously as well as in the main text, Tariq et al (16) (Reference S5 in Wassman et al’s supplement) present data from farms with mid-season drainage. As clarified in our initial study on rice-N₂O (1), mid-season drainage is a form of mild intermittent flooding. The earlier study by Nishimura et al (17) is a method development study without clear results from any one crop or season. While the later study led by the same author, Nishimura et al (18). Shows very clearly that their farms had fluctuating above-zero water levels, it seems that the water levels stayed below zero no more than 1-3 days. Overall, the water indices of this study seem very high and will likely qualify only as continuous or mild-intermittent flooding. Study by Minamikawa et al (19) does not report emissions of N₂O but rather N₂O concentrations in drained water. In other words, Minamikawa et al do not address direct N₂O emissions. They are addressing indirect N₂O emissions which are not a focus on our study.

We note that Reference 6 of Wassmann et al (Abao et al) does not present water level data. While the authors mention the word rainfed, our understanding is that they only measured rice during “wet” season which implies lowland rainfed conditions which are not susceptible to as much fluctuating aeration as upland rainfed farms.

**Interpolation**

**Wassmann et al.:** The new study used a non-linear interpolation method, but individual emission spikes are still reflected as broad peaks. The impacts of such observation gaps in N₂O records
have clearly been shown by comparing manual vs. automated records (3) as well as distinct permutations of sampling intervals based on automated N₂O records (8). These inherent uncertainties in manual sampling should not lead to discarding the computed seasonal emission rates, but questions their use as benchmark for assessing accuracies of other records.

While the new study rightfully finds that N₂O emissions are dominated by short-term spikes, the dynamic nature of this pattern is not reflected in the interpolation method. The authors elaborate that their interpolation method was non-linear, but the interpolation graphs in the supplement show fairly straight lines between two data points even if the sampling intervals were 3-4 days (and in some cases up to 6 days). While these episodes of broad peaks are the actual cause of the resulting high seasonal emissions, this approach is attached to considerable uncertainties and prone to overestimation of cumulated emissions (3) that are not properly discussed in the article or its supplement.

The impacts of such observation gaps in N₂O records have clearly been shown comparing distinct ‘permutations’ of sampling frequencies based on their automated N₂O records (8). Obviously, such systematic errors will also be innate in other manual sampling studies, but the sampling frequency in the new study is by no means sufficient to eradicate these uncertainties. Thus, there is no basis for claiming that previous studies have underestimated cumulated emissions whereas the values by Kritee et al represent a uniformly acceptable reference. Episodic N₂O fluxes (as observed in this study) will require daily measurements to achieve 10% accuracy when calculating annual means (8).

We are happy to clarify that we did not use broad peaks for interpolation. We instead used an exponential curve interpolation method described earlier (15), because we recognized errors introduced using other interpolation methods. Our supporting figures show broad peaks because they present linear interpolation of raw datasets and not the plots based on exponential decay of peaks which were used to estimate seasonal N₂O emissions.

Model development

Wassmann et al.: Model development in this study is limited to the initial step of multiple regression, but omits the decisive step of model ‘validation’ with an independent data set. Multiple regression alone can be done with almost any given data set, so this will not automatically entail more reliable extrapolations of N₂O emissions. The data set from the five farms was used to develop a model for simulating GHG emissions through multiple regression and selecting the “best-fit” model for N₂O and CH₄. However, the approach for model development presented in this study can only be deemed as an initial step and not as a completed model development. What the authors have done is generally described as ‘parametrization’ of a new model which is based on a given data set. The study fails to conduct the decisive step of model ‘validation’ which has to be done by testing the regression model with an independent data set.
Our sample size is small and could not be broken down into parametrization/calibration and validation datasets. Having said that, our model is constrained to simple linear functions of known mechanistic drivers of fluxes and the fit is excellent. We are also not advocating our model against any other approach. As we pointed out in our main response, DNDC based model might have given similar results as our data if they had used higher N rates, incorporated measured high nitrous oxide rates and medium to intense intermittent flooding in their study.

_Wassmann et al.: Multiple regression alone can only be seen as a very limited accomplishment and can be done with almost any given data set. The material provided in the supplement also shows that the model has a poor fit for the 3 farms that have lower but somehow more expected N\textsubscript{2}O emissions. For these farms the R\textsuperscript{2} value is essentially 0.0. Assuming that farm 1 and 3 are outliers - for whatever reason - and farm 2,4 and 5 represent more common conditions this questions the use of this model for up-scaling._

We do not see how our fit is bad for three farms. Moreover, R\textsuperscript{2} values are not calculated for a few data points [or farms in this case] but rather all the farms considered for a model. The overall fit is excellent with a correlation of 0.86 and R\textsuperscript{2} of 0.8. In case it has been missed, we would like to call attention to the figure (SI Figure S3) below which was a part of our supporting information.

**SI Figure S3**

Legend to SI Figure S3: Plot of fitted vs measured N\textsubscript{2}O emissions, using the multivariate regression model that includes water index, continuous flooding events and input of inorganic N. The water index captures the cumulative flooding conditions at each farm but the number of continuous flooding events reflects the temporal pattern that gave rise to the flooding conditions at a specific Farm. Water index, periods of continuous flooding and inorganic N explain 70%, 10% and 4% of the variance in the data, respectively. Even though periods of continuous flooding and inorganic N input explain a small fraction of the total variance when compared to water index, their addition to the model is statistically significant.
**Wassmann et al.:** It is unclear why this non-validated regression model should provide more reliable extrapolations of regional N\textsubscript{2}O emissions in the Indian subcontinent than methodologies based on meta-analysis of emission rates such as the IPCC methodology which is currently updated for the new guidelines of GHG emission reporting. Likewise, such extrapolations could be based on simulation models that have been validated with a wealth of continuous N\textsubscript{2}O records from rice fields (e.g. S9).

First, we are not advocating our model against any other approach. As explained in our initial publication, biogeochemical (DNDC) model gives results similar to ours under mild-intermittent scenario. As explained in the paper, this model might have given risks similar to ours at intensely intermittent flooding scenarios if more updated fertilizer, empirically found high N\textsubscript{2}O rates and multiple intermittent-flooding regimes were incorporated (1). Second, we could not use data from any other studies to validate our model because no study presents water index and flooding frequency information.

**Wassmann et al.:** As stated above, the selected sites in this study seem to have some characteristics that may impede a generalization of findings to entire South Asia. The sites investigated are either in the arid or semi-arid agro-ecological region. The findings therefore have limited diagnostic value for the conditions found in major rice growing areas of South Asia which are dominated by tropical lowland intensive rice production.

We are not sure if the rice growing regions in the world indeed are lowland and will not face intense forms of intermittent flooding. We note that multiple aeration events similar to what we observed at our study farms are common in both irrigated and rainfed rice farms in India(20, 21), Pakistan(21), Nepal(22), Bangladesh(23), China(24) and South America as a result of high evapo-transpiration rates, unreliable water/electricity supply, rainfall regimes, soil characteristics, and/or topography(25).

We call for precise mapping of flooding regimes.

**Interpretation of ‘risks’**

**Wassmann et al.:** The study concludes that the newly recorded emission rates translate into a high risk of underestimating N\textsubscript{2}O emissions. While this logic appears sound, this finding remains weak as long as there is no concomitant information on the likelihood of such a risk. On the same grounds as arguing in favor of increasing regional and global estimates, these individual field records of high emissions could also be interpreted as statistical outliers or anomalies.

Our extrapolation under the continuous-flooding scenario suggests negligible rice-N\textsubscript{2}O and our extrapolation under mild-intermittent scenarios is similar to estimates by other groups. Hence, we have never claimed that all scenarios result in high N\textsubscript{2}O emissions. We do caution that precise
management condition maps as well as high sampling intensity measurements at a range of flooding regimes are needed to understand N₂O tradeoffs with intermittent-flooding regimes designed for reducing methane emissions.

_Wassmann et al.:_ The risk of increased N₂O emissions following the application of AWD has been well-established in several publications including studies with genuine ‘high intensity sampling’ (e.g. S10). However, the meta-analyses of published data clearly reveal that individual records of such high emissions are juxtaposed by a body of evidence that N₂O emissions only slightly increase by AWD. This type of meta-analysis has also been used for developing the IPCC 2006 methodology. On the basis of the data presented, the only new information that can logically be derived from those findings is larger uncertainty in the estimates – and not necessarily higher overall estimates as claimed in the title.

It will be a misrepresentation of our study to imagine that we are arguing that nitrous oxide emissions from rice farms across the world is necessarily very high. One study cannot determine either the average business-as-usual climate impacts of rice cultivation or the average mitigation potential of any farming technique(s). Our aspiration has been to call attention to the risks of neglecting a potentially large rice-N₂O problem. We invite mapping of flooding regimes at farmer-managed farms (as opposed to research-station plots that have consistent access to water/electricity) and more studies at high sampling intensity at a range of intermittently flooded farms.

**Conclusion**

Overall, we believe that more data is necessary to better assess both the conditions that trigger high rice-N₂O and the global spatio-temporal variability of those conditions. We ask the global community to produce more data that can be used to validate/improve our empirical model, and to incorporate our results into existing biogeochemical models (e.g., DNDC and Daycent). Weekly measurements are clearly insufficient to capture rice-N₂O and we encourage a sampling intensity of at least once every two days.

**References**


