June 4, 2015

Greg Mayeur, Ph.D.

California Air Resources Board

1001 I Street

Sacramento, California 95814

RE: Comments on the Compliance Offset Protocol Rice Cultivation Projects

Dear Dr. Mayeur:

Thank you for the opportunity to comment on the proposed Rice Cultivation Projects Compliance Offset Protocol (“Protocol”).

The questions and comments we raise below not only relate to whether this protocol meets the standards laid out for all tradable credits under AB 32, but also to making explicit the criteria and procedures used to develop this protocol, and the precedent this protocol may serve to other future protocols in the agriculture sector. In our comments below, we raise two substantive concerns regarding the evidentiary basis of the Protocol and the quality standards used by ARB in developing this protocol.

First, we note that the DeNitrificationDeComposition (DNDC) model used in the Protocol to estimate emissions reductions was been validated using a static old and incorrect atmospheric CO2 concentration of 350ppm for site-years of data from the late 1980s to 2012. These 87 site-years of data are the basis on which DNDC has been validated for rice cultivation in California and the Mid-South. When we redid the validation of DNDC using appropriate atmospheric CO2 concentrations for each of the 87 site years of data, we found that the Protocol as written will not achieve the level of conservativeness intended by the ARB unless the uncertainty deduction factor is updated.

Second, for over a year we have asked ARB staff a very basic question about the scientific basis of the Protocol—is the model used to estimate emissions reductions by projects under the Protocol adequately validated for each of the new practices being credited?—without yet receiving an answer for alternative wetting and drying (AWD) projects. We have asked many times for the basis on which the DNDC model has been validated for alternative wetting and drying (AWD) projects in the Mid-South. Such a question is relevant given that validation of the model is based on only one site-year for a field employing AWD practice, and given that the emissions release processes are different for fields employing this practice than for fields employing continuous flooding. We are surprised and concerned that ARB is moving forward with the final version of the Protocol before its staff has a satisfactory answer to this important question.

These two issues make us concerned that ARB has abdicated quality control responsibility to contractors without sufficient oversight, and without being responsive to stakeholders who identify potential quality control issues. We identified and analyzed these substantive concerns with a relatively small amount of effort; it is clear to us from these two examples that ARB should not adopt the DNDC-based Protocol until staff has undertaken more thorough validation. We stress that we have been good faith participants in the Technical Working Group process since its inception in 2013, and that we have repeatedly engaged with ARB staff about our concerns, in person, on the phone, and over email, for over two years.

We also stress an important procedural concern related to the responsiveness to staff to stakeholder questions related to the scientific basis on which the Protocol has been developed and is recommended for adoption. We are surprised that ARB released a final draft of the Protocol without answering our question about the basis on which DNDC has been deemed validated for AWD projects, which we have repeatedly asked and to which we have been promised an answer. First, we stress the importance of peer review. Accurately estimating emissions reductions from offsets projects, especially when they involve land use change, is complex. Input from stakeholders is essential to the process of developing a protocol that is scientifically sound. Answering relevant stakeholder questions about the scientific and factual basis of the protocols ARB develops is necessary for peer review. Answering such stakeholder questions is also an essential part of notice and comment rulemaking. A public review period is ineffective if the public is not provided the information needed to review the protocol being put forward for public comment or if the agency fails to respond to material public comments. This is also essential from the perspective of instilling public trust in ARB’s offset program. Answering public questions about the scientific basis of ARB’s protocols is especially important for offsets, the use of which will allow continued or increased operation of facilities such as refineries in communities suffering disproportionate environmental and health impacts.

Finally, we raise a last substantive concern about the evidentiary basis for project types being found to pass the Performance Standard test for additionality. Given the problems associated with data transparency, response to stakeholder concerns, and quality control, we stress that ARB has not provided sufficient recent evidence for the decisions that it has made regarding additionality assessments. Most data cited are older than ten years.

We are looking forward to receiving answers to the follow questions about the proposed Protocol:

***1. DNDC should be validated using corrected values for atmospheric CO2 concentrations appropriate for each site-year of data used in the validation; doing so may result in the need to generate a different uncertainty deduction factors for each rice growing region.***

***The following comments are submitted in relation to the modification of Equation 5.4 to use a fixed structural uncertainty deduction factor, rather than a variable value, and to remove the use of separate uncertainty deduction factors for each rice-growing region.***

Previous versions of the Protocol included the use of separate structural uncertainty deduction factors for each rice growing region. In those versions, separate structural uncertainty deduction factors were calculated, based on the site-years of available data from each rice growing region, with the *a priori* assumption that climatic and soil variability between regions might lead to identifiable and consistent differences in the uncertainty of DNDC’s estimations of methane (CH4) emissions and reductions between the rice growing regions. In Equation 5.4 in the current Protocol draft, these separate uncertainty deduction factors have been removed.

While the rationale for the removal of these separate factors was not directly stated in the Summary of Proposed Modifications (only the choice to use a single fixed value of 0.128 MTCO2e/ha rather than a variable value based on the number of acres enrolled in projects was explained), it is our understanding that the evidentiary basis for the change lies in a statistical analysis that concluded there was no difference in the performance of the DNDC model between states and that the best-fit statistical model was a mixed-effects model that included a random term for site-year and was selected using an information-theoretic approach. This analysis was performed by ARB contractor, Dr. William Salas of Applied GeoSolutions, and sent to us via email on February 18th, 2015. The text of that analysis, which provides the rationale for using a single uncertainty deduction factor for the DNDC model in the context of a carbon market in the United States is appended to this comment as Appendix A and Appendix B.

*Is the statistical analysis performed by Dr. William Salas included in this appendix the evidentiary basis for the choice by ARB to use a single structural uncertainty deduction factor, rather than a separate one for each rice growing region?*

We assume that it is. Our assumption is based on the following: Dr. Salas has repeatedly presented information on DNDC validation to the Rice Cultivation Projects Technical Working Group. Information from Dr. Salas, including the graphical representation of the structural uncertainty deduction, was presented by ARB Staff in their Workshop presentation on February 20th, 2015. Dr. Salas has spoken to us on the phone along with ARB Staff about these issues relating to DNDC validation on March 19th, 2015, and Board Staff have repeatedly directed us to Dr. Salas, when we approached ARB staff as stakeholders with questions and concerns about DNDC validation.

We do not have any concerns about the scientific approach to statistical analysis that was performed and outlined in the appended documents, Appendix A and Appendix B. Our concern lies instead with the data on which these analyses were performed. Based on our analysis, we believe that there is a substantive error in the values of the modeled methane (CH4) fluxes used in the validation assessments.

DNDC model validation for CH4, the calculation of the structural uncertainty and deduction factors, and the statistical analyses to determine whether separate factors were needed for different regions were all performed using a set of 87 site-years of measured CH4 fluxes. These data were presented graphically by Board staff during a Workshop on February 20, 2015. The measured CH4 fluxes are rice fields in the United States from four states Louisiana, Arkansas, Texas and California, all made between1989-2012. All of these data were published in peer-reviewed literature, andlist of these publications was made available to the public by the Board Staff.

The modeled CH4 fluxes used in the analysis to assess DNDC model performance were based on model runs performed by Dr. Chengsheng Li at the University of New Hampshire. As stakeholders interested in the integrity of the Protocol and the validation of the DNDC model for use in estimating emissions, we wanted a chance to review the model input files and parameters which were used to estimate emissions for the 87 site-years of field measured data. This is because the output of these model runs forms the entire basis for the assertion that the DNDC model is an accurate representation of methane emissions from rice cultivation in the United States and is thus a critical component of the rationale for the use of the model in the first place. The input values and model parameters from these model runs were not made publically available during the Technical Working Group Process. The only information during this process that was made publically available was the graphical representation comparing modeled and measured CH4 fluxes and a list of the peer-reviewed publications from which the measured CH4 fluxes were taken. No information on the details of the model runs was provided to the public at this time.

After repeatedly asking ARB staff for access to the original model input files by email, in submitted comments, over the phone and in person, on March 19th, 2015, on a phone call with Dr. Greg Mayeur and Dr. William Salas, Dr. Salas offered to send us a .rar file which included the DNDC input files used in model validation and the original spreadsheet of data from Dr. Li on which subsequent statistical analyses of model uncertainty had been performed. These files were sent to us as well as to ARB Staff via DropBox on March 22nd, 2015, for which we are very thankful. Having access to these critical data is essential to assessing the basis on which ARB Staff have proposed using the DNDC model. These files included input files for 84 of the 87 site-years used in analysis. Three files were missing (for Brennan 2010 California Mid-Season drainage, RES 2011 California M206, and for Rogers, Stuttgart 2011 California Between-rows).

Since receiving the input files electronically, we have performed several analyses We double-checked that the information on soil type, crop, flooding, irrigation, and fertilizer was correctly transferred from the information in the peer-reviewed literature to the DNDC input file. In all cases that we checked, this was done accurately, with only one minor issue: the climate data in the input file for four varieties in the “Merle, Stuttgart” files from Arkansas was from 2011, but the methane measurements were made in 2012. But, updating these climate data files to 2012 climate data from the same site, and re-running the model had only a modest effect on the output results. Using the DNDC95.exe program, we then re-ran the model using the input files sent to us in order to re-create the reported “modeled” CH4 emissions file in the spreadsheet, and were able to achieve the same results as reported and presented to the Technical Working Group.

However, we did note an issue of major substantive concern with how the DNDC model was run for model validation. For 80 of the 84 input files that we were sent, representing field measurements from 1989 to 2012, the background atmospheric CO2 concentration used in the model was set to the DNDC default value of 350ppm (For reference, the 15-day draft Protocol uses a default value of 400ppm). For 4 of the 84 files, representing field measurements in 2012, the background concentration was listed as 370ppm.

Global average CO2 concentrations have not been at 350ppm since the mid-1980s. In 2014, global average atmospheric concentrations were 398.6ppm. The table below shows the average atmospheric CO2 concentration, globally, for the years included in the validation dataset. The data are provided by the National Oceanic and Atmospheric Administration (NOAA) as a global average:

 **Year CO2 (ppm)**

|  |  |
| --- | --- |
| **1989** | **352.91** |
| **1990** | **354.19** |
| **1991** | **355.59** |
| **1992** | **356.37** |
| **1993** | **357.04** |
| **1994** | **358.89** |
| **1995** | **360.88** |
| **1996** | **362.64** |
| **1997** | **363.76** |
| **1998** | **366.63** |
| **1999** | **368.31** |
| **2000** | **369.48** |
| **2001** | **371.02** |
| **2002** | **373.10** |
| **2003** | **375.64** |
| **2004** | **377.36** |
| **2005** | **379.63** |
| **2006** | **381.81** |
| **2007** | **383.59** |
| **2008** | **385.45** |
| **2009** | **387.36** |
| **2010** | **389.90** |
| **2011** | **391.65** |
| **2012** | **393.88** |
| **2013** | **396.52** |
| **2014** | **398.60** |

The use of 350ppm instead of the actual atmospheric concentration for that year is a major cause for concern because these default values in the validation runs of the model are too low for the years in which the measurements were made. While *in* *situ* atmospheric concentrations of CO2 can vary daily and seasonally due to local conditions, the background value used in the model is intended to represent the concentration for the entire year that is being run.

The atmospheric concentration of CO2 is a critical value in calculating methane emissions from rice cultivation because higher atmospheric CO2 values lead to greater CO2 fixation by crops and ultimately a larger store of organic carbon available for methanogenesis. Thus, the fact that these concentration values were not updated in the input files used for the validation data set means that the ***modeled CH4 fluxes used for assessing DNDC uncertainty are not correct.***

After realizing this error, we reran the DNDC model using the corrected atmospheric CO2 concentrations found in the table above. Doing so had an average effect of increasing the model-estimated annual CH4 emissions by 8.4±7.1 kg CH4-C ha-1 y-1, with a max effect of a 48.4 kg CH4-C ha-1 y-1 increase, representing a greater than 13% increase in the modeled emissions value from this single input factor correction. The effect of this error is more pronounced for more recent years, as atmospheric concentrations were much farther from the default 350ppm in 2012 than in 1989.

The effect of our correcting the modeled outputs is not insignificant. The graph below shows the shift in the graphical representation of the modeled vs. measured outputs and the corresponding shift in the best-fit model (represented by the lines) from the one provided by Dr. Salas. (Grey points are the original model, colored points are the updated values, the darker line reflects the shifted best-fit model using the information-theoretic approach used by Dr. Salas). The ultimate result is that, because updating the CO2 concentrations has the effect of increasing all modeled values (to varying degrees), the entire correlation shifts to the right in the graph.



There are several important implications of this shift in the model-output data. The first is that the value of the structural uncertainty deduction used in Equation 5.4 (0.128 MTCO2e/ha) needs to be recalculated.

The structural uncertainty deduction factor is a function of the modeled relationship between measured and modeled CH4 fluxes and the number of participating acres in the aggregate to which it is being applied. Calculating the factor results in an amount of tCO2e per hectare that should be deducted from calculated emissions reductions due to the structural uncertainty of the model’s performance.

That value of 0.128, according to the Summary of Proposed Modifications, is the “plateau” value that is achieved at around 5,000 project acres total. (Previous versions of the Protocol had updated the uncertainty deduction values based on the number of acres of rice participating each year, but it was found that the amount of anticipated projects passes the asymptotic threshold so a set value could be used, which makes sense).

Following Appendix A, the structural uncertainty deduction is calculated from the following equation:

Where:

|  |  |
| --- | --- |
| Symbol | Definition |
| u | The uncertainty deduction for the aggregate; the amount to be deducted from  |
|   | The total modeled offset for the project |
|   | The fixed-effect slope associated with x in the regression model |
|   | The value of Student’s t with error rate . k is the number of baseline-project pairs in the validation data set. |
|   | The total variance of the error term in the regression model, including all random effects as well as the true residual |
|   | The correlation between errors for a baseline-project pair observed at the same site and year |
|   | Area (ha) associated with an individual field in the aggregate. Under current assumptions, is treated as equal to 1, and n is the total area (ha) of the aggregate. |
|   | DNDC-predicted flux (kg/ha) associated with an individual field, using the project treatment |
|   | Corresponding DNDC-predicted flux associated with an individual field, under the baseline treatment |

The following parameters in this equation are calculated from the summary of the variance of the “best-fit” statistical model comparing measured and modeled emissions:

As described in Appendix A, from all the information that has been made available to us by Dr. Salas, it is our understanding that the original best-fit statistical model was selected using an information-theoretic approach based on Akaike Information Criterion (AIC), with a mixed-effects model with a random effect for site-year, implemented in the *nlme* package in the R statistical program. Appendix A shows how the parameters for the uncertainty equation are calculated from the information on the variance of the model.

We have re-run the exact same code in the R statistical program (the *nlme* package) as Dr. Salas. Using the data originally sent to us on March 22nd, 2015, we were able to obtain the exact same parameter numbers for ,and that are given in Appendix A, so we have confidence in our performance of the statistical analysis. As Dr. Salas had asserted in Appendix B, in our re-creation of his analysis using the same data and approach described, the best-fit model was the one reported and used in calculating structural uncertainty that was used in the Protocol.

**We have now re-run this statistical analysis to pick the best-fit model using the *nlme* package in R using our corrected model output values using correct values for atmospheric CO2 concentration**. The summary of the nlme new best fit model, based on AIC criterion is given here (this is the same summary form as was provided in Appendix A).

Linear mixed-effects model fit by REML

 Data: data

 AIC BIC logLik

 949.0669 958.8375 -470.5335

Random effects:

 Formula: ~0 + x | State

 x Residual

StdDev: 0.2013773 0.4270917

Variance function:

 Structure: Power of variance covariate

 Formula: ~x

 Parameter estimates:

power

 1

Fixed effects: y ~ x2

 Value Std.Error DF t-value p-value

(Intercept) -3.482680 0.7602852 82 -4.580754 0

x2 0.916452 0.0978295 82 9.367854 0

 Correlation:

 (Intr)

x2 -0.124

Standardized Within-Group Residuals:

 Min Q1 Med Q3 Max

-1.85885372 -0.55547852 -0.06493635 0.29953547 4.22015353

Number of Observations: 87

Number of Groups: 4

Using this new model has several effects. First and foremost, it changes several of the terms used in the structural uncertainty deduction calculation. These changes are given in the table below. They have the effect of slightly increasing the uncertainty deduction factor. Given that this model is the basis for the choice of a uniform number (0.128 MTCO2e/ha) in the Protocol, our results demonstrate the need to correct the structural uncertainty deduction to reflect the updated modeled results based on the updated best-fit mixed-effects model. This new factor still plateaus at a given number of acres, but the number at which it plateaus is different from that reported and is slightly higher.

|  |  |  |
| --- | --- | --- |
|  | Original Model | Corrected Model |
|  | 1.0859 | 0.9165 |
| var() | 0.00516 | 0.009570611 |
|  | 0.211 | 0.222960137 |
|  | 0.2164 | 0.18188371 |

Secondly, and more importantly, the conclusion of Dr. Salas in Appendix B that the *“*that the data do not support the need for a separate calculation *(of structural uncertainty deduction factors by region)”*, no longer holds with the corrected data. By AIC criterion, using the exact same statistical package used previously, the best-fit model is now one that includes a random effect of the rice-growing state, and not for site-year.

The AIC values and results of the two model runs (one with the same form as the one used by Dr. Salas but with updated values, the other the best choice based on AIC value) are shown below.

AIC for previous best-fit model and random effect of site year: 949.429

AIC for previous best-fit model and random effect of state: 950.542

Choose: model with random effect of site-year (this is the one reported in Appendix A).

AIC for model with corrected data and random effect of site year: 950.141

AIC for model with corrected data and random effect of state: 949.0669.

Choose: Model with random effect of state.

The lower the AIC, the more likely the model is supported by the data.

Based on the model output data that were calculated using 350ppm, the single uncertainty deduction factor for all regions was strongly supported by the data. However, with the updated data using the actual atmospheric CO2 for each year to run the model, the data support a separate uncertainty deduction factor for each region. This is driven by a difference in model performance for Louisiana site-years compared with other states.

It is important to emphasize that none of these new results suggest that the DNDC model is not appropriate for use in estimating emissions reductions. They do, however, suggest the need for a more careful examination of the site-years of data and the input values used for the model runs in the validation dataset.

The above analysis does show substantial reason for concern that the process used to validate the DNDC model, and to assess its uncertainty and bias, merits another look and careful quality control. With access to the original data and input files, our quick look “under the hood” found an error that changed one of the key statistical conclusions about the model’s performance, suggesting a revisitation of several of the modifications in the 15-day draft. We have not, however, performed and exhaustive audit or peer-review of the model validation.

We humbly suggest that the ARB take the time necessary to assess all the available information about DNDC performance to make sure the numbers and assertions in the Protocol are, in fact, as good as they can be. In our opinion, this need not be a lengthy delay, but more time is clearly required to get the numbers right, as evidenced by the information provided above.

We note that all our analyses are based on the 87 site-years of data that have been presented in public presentations of model validation and that were shared with us upon request. There may have been additional site-years of data used in the analyses that formed the rationale for use of the model in the Protocol, though we have no information to this effect. However, even if there were additional site-years of data used in the final dataset that was used to calculate the uncertainty deduction factor used in the Protocol, the substantial issue remains that model validation was performed using a default value for atmospheric CO2 concentration that was not updated to reflect actual on-the-ground conditions, nor in a manner consistent with the default values in the Protocol itself.

We request ARB to provide a full documentation of the rationale and evidentiary basis for its assertion that the DNDC model accurately estimates emissions of CH4 from rice fields in the United States, and for the choice of uncertainty deduction factor in Equation 5.4. We have performed the above analyses using all the publically available information, in addition to information provided to us in private communication by ARB staff and by contractors working with ARB staff.

***2. On what basis has DNDC been validated for AWD projects in the Mid-South?***

***The following comments are submitted in relation to the modification of Equation 5.4 to use a fixed structural uncertainty deduction factor, and its implications for DNDC model validation****,* ***specifically for AWD projects****.*

In the Staff Report accompanying the 45-day draft of the Rice Protocol, released on October 28, 2014, the report explains why the Alternate Wetting and Drying (AWD) project-type is only eligible for crediting in the Mid-South:

*“However, alternate wetting and drying is only an eligible project activity in the Mid-South Rice Growing Region under the proposed Rice Cultivation Protocol because the DNDC model has not been validated for this activity in the California Rice Growing Region. If validation of the DNDC model for alternate wetting and drying occurs in the future, ARB will consider adding it as an eligible project activity.”*

This rationale strongly implies that the DNDC model has been validated for the AWD project activity in the Mid-South Rice Growing Region.

In Appendix B to this document, which outlines the evidentiary basis for using a single uncertainty deduction factor in the Protocol, a modification found in the 15-day draft, provided to us by Dr. William Salas, it is carefully explained that there is no statistical difference in model performance between the rice-growing regions or between individual states, even though *a priori* one might assume that the soil and climatic variability between the regions might suggest a basis for a difference in performance. However, this documentation does not mention an analysis of DNDC model performance by separate project activity types, despite the implication that a validation assessment (comparing the relationship of measured to modeled emissions) for this particular project type had been performed.

We thus had several questions about AWD project validation:

*Has the DNDC model been validated for AWD project activities in the Mid-South region? Has this validation been distinct from the overall validation dataset involving 87 site-years that has been presented by Board staff? On what basis was this validation for AWD done?*

How the DNDC model performs estimating methane emissions from a rice field in which an Alternate Wetting and Drying practice has been implemented may not be the same as how the model performs estimating methane emissions from fields not doing AWD. In fact, the peer-reviewed literature suggests that repeated spikes in emissions can occur due to wetting and drying, as biogeochemical conditions shift from anaerobic to aerobic and back again (Linquist et al. 2015). Thus, an AWD project is treating accumulated methane from methanogenesis and its emission very differently than other project types do. It is thus critical to ascertain whether the DNDC model estimates emissions for AWD projects with the same uncertainty as for other field conditions.

Because of this critical need, and the language in the October 28, 2014 Staff Report implying that DNDC’s ability to accurately estimate AWD emissions had been validated for the Mid-South, but not for California, we spoke with ARB Staff Dr. Greg Mayeur and with ARB contractor Dr. Bill Salas of Applied GeoSolutions on March 19th, 2015 by phone to express our concern. During that phone call, all parties agreed that a separate analysis of the uncertainty of model performance for AWD projects that could be compared with the overall structural uncertainty calculated from all site-years in the validation dataset made sense. Mimicking the analysis done for individual regions and states outlined in Appendix B, this analysis would allow us to determine if the uncertainty in model performance for AWD project CH4 emission estimation was different from uncertainty for non-AWD projects.

At the time, Dr. Salas indicated that he would provide statistical analysis comparing the uncertainty for AWD projects with that of all site-years in the data set, as had been done for the rice-growing regions (see Appendix B)

As of June 2nd, 2015, after several follow up emails and phone calls to Dr. Mayeur, we had not seen this additional, critical analysis.

Because we have not seen the results of an analysis of structural uncertainty of the DNDC model for CH4 emissions on a project-basis, and because there are *a priori* reasons, based on data presented in peer-reviewed literature to suspect that the model may perform differently for AWD projects than for other projects, we are concerned about the use of the DNDC model for estimating emissions reductions from AWD projects.

Thus, these key questions remain unanswered:

*Is the use of a single structural uncertainty deduction factor, based on the site-years of data that are available, justified for all project types, including AWD? If so, on what basis is it justified? Is there any difference in structural uncertainty of the DNDC model for AWD projects when compared with the data set of all available site-years from the United States? On what basis do we know that DNDC’s structural uncertainty in the estimation of CH4 emissions and emissions reductions is the same for AWD projects as for other projects?*

As such,we are unable to provide adequate comment on this protocol because ARB has failed to provide key elements of the evidentiary basis for their decision.

In particular, our review of the 87 site years of data reveals that only 1 site-year includes an AWD project (in Arkansas -- though there may be more recent ones that were published in 2015). Thus, with the evidentiary basis that is currently available for public review, we can conclude that it is only on the basis of this single data point, that ARB has determined that the DNDC model can accurately estimate emissions of CH4 from a field implementing AWD in the Mid-South. The one-data point had a measured flux of 5.00 kg CH4-C ha-1 y-1, compared with an estimated flux of 11.84 kg CH4-C ha-1 y-1. This is not enough information on which to make an assessment of model performance for AWD. It does strongly suggest that the DNDC model *likely* will do a good job estimating emissions reductions from AWD, but we simply don’t know and cannot know with the available information.

It is important re-emphasize that a project-type specific assessment of model validation was referred to by board staff in their Staff Report from October 2014. That assessment has not been made available, and the available evidence suggests that the assessment was made on the basis of a single data point. If this is not the case, and additional site-years of information were used to perform a statistical analysis, we have not seen that analysis and therefore we are unable to provide adequate comment on this protocol because ARB has failed to provide key elements of the evidentiary basis for their decision.

We want to emphasize that our concern is not as much with any evidence that DNDC is *not* performing well enough, but that the information that is available is not adequate to make a positive assertion that DNDC is validated for this project-type. If there are additional lines of information that lead ARB to this conclusion, we look forward to seeing them.

***3. On what basis is the Protocol considered to meet the requirement that the credits generated under the Protocol must be real and conservatively estimated with regard to the effect of N2O emissions on the number of credits generated?***

***The following comments are submitted in relation to the modifications of terms in Equations 5.2.1., 5.3.1, 5.3.2 and then the subsequent use of these calculated values in Equation 5.4.1, specifically as they relate to the calculation of nitrous oxide (N2O) emissions*.**

We understand that the Protocol does not credit reductions in nitrous oxide (N2O) emissions from projects, and only allows for a debit based on increases of N2Oemissions that may result from project implementation (this is reflected in the “MIN” term in Equation 5.4.1.)

However, we are not aware of any presentation of information demonstrating that the DNDC model has been validated for estimating N2O emissions. All information presented to the Technical Working Group related only to validation of CH4 emission estimation.

*Is it true that the DNDC model been not been validated for N2O emissions from rice cultivation in the United States? And if so, on what basis is the Protocol considered to meet the requirement that the credits generated under the Protocol must be real and conservatively estimated with regard to the effect of N2O emissions on the number of credits generated?*

Based on our phone conversation with ARB Staff Dr. Greg Mayeur on March 19th, 2015, we understand that the only project types for which increases in N2O emissions might be anticipated are for AWD projects, and that, based on the peer-reviewed literature (i.e. Linquist et al. (2015)), it is expected that such increases will only represent a small fraction of the overall net reduction in GHG emissions from the reduction in CH4 emissions from such projects. It is also our understanding from conversations with Dr. Bruce Linquist that the only circumstances where large amounts of N2O might be released from an AWD project would be if drainage occurred immediately following fertilizer application. Such an activity would be unlikely, because of the loss in fertilizer that would result if a farmer were to drain shortly after fertilizer application before the fertilizer has been taking up by crops. We confirmed this assessment on the phone with Dr. Greg Mayeur on March 19th, 2015.

Thus, our concern is not that there is evidence that the inclusion of N2O termsin Equations 5.2.1., 5.3.1, 5.3.2, and 5.4.1 will result in substantial over-crediting. Our concern is primarily focused on the precedent set by adopting a Protocol which uses of a process-based biogeochemical model that has not been validated to estimate the flux of a greenhouse gas from land use activity. In order to address such concerns about precedent and evidentiary basis for use of a model, in a Protocol which is using such a model for the first time in the Compliance Offset Program, we ask ARB to be explicit and public about how they reached the conclusion that DNDC outputted values of N2O were sufficiently validated for inclusion in the Protocol.

***4. We request more clarification of the basis on which each of the project types is considered additional under the Performance Test Standard.***

The current practice data cited in ARB’s October 28, 2014 Staff Report and Rice Cultivation Projects Compliance Offset Protocol for each of the practice types is outdated; almost all data are from ten years ago or more. Dry seeding in California, and AWD and early drainage in the Mid-South have been promoted since then due to their water cost and weed control benefits. More current practice rates are important to assess whether these practices are currently common practice. They are also important to enable ex-post assessment of the effects of the Protocol on emissions reductions such as net-to-gross analyses. Without better documentation of current practice, it will be difficult for ARB and outside observers to assess the effectiveness of the Protocol in changing farmer practice.

1. *Dry seeding in CA: We understand from Cline (2003) and Carol (2009), cited in October 28, 2014 release of the Protocol and Staff Report, that dry seeding is being encouraged in California due to its advantages for herbicide resistant weeds associated with wet seeding. How much has dry seeding increased since 2009? Is it still being used by a very small percentage of farmers given its benefits?*
2. *AWD in the Mid-South: What rates are farmers currently performing AWD in the Mid-South? We understand that very few were employing this practice in 2006, but that it has large advantages for water use reduction. But we understand from Hardke 2014, cited in the October 2014 ISOR, that “Intermittent flood or alternate wetting and drying (AWD) is another water management practice that is not currently widely practiced in Arkansas but has been practiced with success in nearby states...” What is current practice?*
3. *Early drainage the Mid-South: The Counce 2009 article, cited in the October 2014 ISOR, states that farmers are reluctant to drain early for fear of harming the rice but that there would be significant water cost savings. What is current practice for early drainage in the Mid-South where a model is used to determine the baseline drainage date?*

Sincerely,

Aaron Strong, Ph.D. Candidate

Environment & Resources (E-IPER)

Stanford University

alstrong@stanford.edu

Barbara Haya, Research Fellow

Stanford University

bhaya@berkeley.edu

**Citations**

Linquist, B. A., Anders, M. M., Adviento‐Borbe, M. A. A., Chaney, R. L., Nalley, L. L., da Rosa, E. F., & Kessel, C. (2015). Reducing greenhouse gas emissions, water use, and grain arsenic levels in rice systems. *Global change biology*, *21*(1), 407-417.

**APPENDIX A. (PROVIDED BY DR. BILL SALAS)**

**UNCERTAINTY DEDUCTION FOR DNDC-PREDICTED CH4**

**Background Work.**

The first step in developing the uncertainty deduction calculation is to assemble a validation data set including a relevant set of treatments at multiple representative locations within the region or regions under consideration. The data employed here are from 87 site-treatment combinations in California, Texas, Louisiana, and Arkansas. The data were used in an information-theoretic approach using AIC to identify whether the underlying error structure was additive or multiplicative, and to screen alternative models representing different assumptions about the random effects (site, state, region, and year) influencing DNDC prediction errors. Follow-on testing allows for the assessment of the presence of bias (using equivalence testing) and homogeneity of variances between baseline and project treatments. These steps are all conducted using maximum likelihood estimation. Once the final model and its error structure have been identified, the model is re-fit using restricted maximum likelihood estimation (REML), to provide more efficient estimates of the variance components of the model. The variance components are critical to the uncertainty deduction calculation.

**Fitting The Final Model.**

In this case, all of the statistical modeling was conducted within the R statistics program, using the nlme package for mixed-effects modeling. The following line loads the data into a data frame called CH4. The DNDC-predicted fluxes are in the variable x, and the field observed fluxes are in the variable y:

CH4<-read.csv('CH4\_data\_updated.csv')

Next, we fit the final model using REML:

library('nlme')

bestmodel<-lme(y~x, random=~0+x|site\_year,

 weights=varPower(1, form=~x, fixed=1),

 data=CH4, method="REML")

summary(bestmodel)

And we see the following as output:

Linear mixed-effects model fit by REML

 Data: CH4

 AIC BIC logLik

 949.429 959.1996 -470.7145

Random effects:

 Formula: ~0 + x | site\_year

 x Residual

StdDev: 0.2136662 0.4065882

Variance function:

 Structure: Power of variance covariate

 Formula: ~x

 Parameter estimates:

power

 1

Fixed effects: y ~ x

 Value Std.Error DF t-value p-value

(Intercept) -2.747437 0.9716754 65 -2.827526 0.0062

x 1.085872 0.0717988 65 15.123816 0.0000

 Correlation:

 (Intr)

x -0.367

Standardized Within-Group Residuals:

 Min Q1 Med Q3 Max

-1.5441112 -0.4667372 -0.1198119 0.1223516 4.3950235

Number of Observations: 87

Number of Groups: 21

**The Uncertainty Calculation.**

The CH4 data reflect a multiplicative error structure, with equal variance for baseline and project treatments, but with possible bias in the model. The formula for the uncertainty deduction is

with the symbols defined as:

|  |  |
| --- | --- |
| Symbol | Definition |
| u | The uncertainty deduction for the aggregate; the amount to be deducted from  |
|   | The total modeled offset for the project |
|   | The fixed-effect slope associated with x in the regression model |
|   | The value of Student’s t with error rate . k is the number of baseline-project pairs in the validation data set. |
|   | The total variance of the error term in the regression model, including all random effects as well as the true residual |
|   | The correlation between errors for a baseline-project pair observed at the same site and year |
|   | Area (ha) associated with an individual field in the aggregate. Under current assumptions, is treated as equal to 1, and n is the total area (ha) of the aggregate. |
|   | DNDC-predicted flux (kg/ha) associated with an individual field, using the project treatment |
|   | Corresponding DNDC-predicted flux associated with an individual field, under the baseline treatment |

By examination of the model outputs, and some simple calculations, we can identify , , and :

* is the fixed-effect “slope” of the model, or 1.0859. Its variance, , equals the square of its standard error, or 0.07182=0.00516.

* is the sum of the variances of the site-by-year random effect, and the true residuals. The model output reports the standard deviations, so we square these to get their variances. Thus we have .

* is the variance of the site-by-year random effect, as this would be a shared error term for any project-baseline pair, divided by . Thus .

For this data set, we also have t=1.663. This completes the set of parameters we can recover from the model alone. Substituting in these factors, we have

For the full uncertainty deduction calculation, we also require information about the aggregate or portfolio being discounted. At the simplest level, we need to know the proportion of the total area represented by a set of modeled fields or strata within the aggregate, the predicted emissions under the baseline scenario for each, and the predicted emissions under the project scenario for each. The difference between project and baseline emissions is the predicted offset. For example, consider the following simplified aggregate:

|  |  |  |  |
| --- | --- | --- | --- |
| Proportion of Total Area | xbi, kg C/ha/yr | xpi, kg C/ha/yr | Modeled Offset, kg C/ha/yr |
| 0.25 | 50 | 15 | 35 |
| 0.25 | 150 | 50 | 100 |
| 0.25 | 250 | 125 | 125 |
| 0.25 | 500 | 250 | 250 |

For this portfolio, we can compute the average modeled offset per unit area as:

We can also simplify dealing with the summand within the square root, by first calculating a series of areal averages:

Finally, we may compute the uncertainty deduction for this aggregate as:

and substituting in the total area of the aggregate completes the calculation.

**APPENDIX B. (PROVIDED BY DR. BILL SALAS)**

**Do We Need A Region-Specific Uncertainty Discount Calculation for DNDC?**

Given the differences in climate, soils, and other biophysical factors between California and the U.S. mid-south (Texas, Louisiana, and Arkansas), it is reasonable to suspect that DNDC might perform differently in estimating CH4 emissions and offsets. Thus, a separate uncertainty discount calculation would be warranted. On the other hand, if a single calculation could be applied, it would simplify accounting for uncertainty within the market framework.

Our analysis (based on a validation data set of 87 site-treatment combinations) indicates that the data do not support the need for a separate calculation. This data-driven conclusion arises from several lines of statistical evidence:

**1. The error structure of the best predictive model, as selected using information-theoretic criteria.** The uncertainty deduction calculation is strongly driven by an underlying regression model, which for a multiplicative error takes the form:

where y is the observed flux in the validation data, x is the predicted flux, and are an intercept and slope that must be estimated from the data, and is an error term. Unlike an ordinary regression model, the error term can include components that vary between sites, between regions, or between years of observation. These components imply correlations between different observation/prediction errors. If there were strong regional differences in model performance, then the information-theoretic model selection procedure should choose a model for the error term that includes a component related to region or state.

In fact, this does not occur with the CH4 validation data. The best predictive model includes a random component related to the combination of site and year: at a particular site in a particular year, all model predictions may tend to run slightly high or slightly low, but this is not consistent across sites within the same year, or across different years within the same site. No model that included a consistent term related to state or region was strongly supported by the data. This implies that the uncertainty deduction can be driven by the same underlying model for both regions of the U.S.

**2. Alternative model evaluations using traditional frequentist hypothesis testing.** Another approach would be to ask whether there are significant differences between states, or between regions in the performance of the model. Because the underlying regression model is linear, this can be evaluated using a set of hypothesis tests closely related to what would be called Analysis of Covariance (ANCOVA) if the model had a simple error structure. In this case, the only difference is that the hypothesis testing must be conducted while accounting for the multiplicative nature of the error, and the presence of a site-by-year random effect. The results of these hypothesis tests are clear. No attempt to introduce state or region as an influence on the slope or intercept of the regression model is statistically significant, or even marginally so (p>0.10 for all tests conducted). From a hypothesis testing perspective, we find no evidence to reject the null hypothesis that DNDC performance in the different states or the two regions is essentially the same.

**3. Further evaluation of the variance structure.** It is not just the slope and intercept of the regression that matter for the uncertainty deduction, it is also the magnitude of the residual variance. As a final check, we tested whether the magnitude of the residual variance was different between regions, using an Ansari-Bradley test. The test found no significant difference in the residual variance between regions (p=0.53).

**Conclusion.** The validation data provide no compelling evidence that separate models are needed to describe DNDC performance between California and the mid-South, despite efforts to uncover such differences using multiple approaches. This does not mean that a separate regression model might not be needed for some other, new region. From a purely statistical perspective, validation data will always be desirable to test whether a new regression model is needed and, if so, to identify the bias and error structure of DNDC predictions. However, the consistent performance of DNDC across these regions, with their differences in conditions, does suggest that DNDC captures salient biophysical features in a reliable fashion.

