To: Patrick Barickman, Whitney Oswald, Brock LeBaron
From: Jon Wilkey
Date: March 31, 2016
RE: March 2016 monthly report on integrating equipment-based emission factors into the R model

During the month of March, work on the R model has focused on analyzing the information contained in each of the OGEI database tables. The goal of this data analysis work was to produce all of the information necessary to enable a Monte-Carlo simulation of the inputs required for calculating emissions from the equipment types in UDAQ's emissions inventory workbooks. That task was completed, and a full report on the results of that analysis is attached in the file "OGEI Data Analysis" (as both an *.html and *.pdf file).

To summarize the contents of that report, a method has been developed to analyze all of the unique combinations of the input variables contained in the OGEI database by:

1. Finding all unique combinations of the necessary input variables in each table
2. Counting the frequency with which each combination is repeated
3. Determining the cumulative probability of each combination occurring

The resulting cumulative probability table (CPT) contains one row for each combination of inputs along with the cumulative probability of that combination occurring. The CPT can then be combined with the existing Monte-Carlo structure of the R model (randomly picking a number between 0 and 1) to select a set of inputs for calculating equipment-based emissions. Like other parts of the R model, the script for generating the CPTs (and the OGEI Data Analysis report itself) are fully automated, and can be rerun as the OGEI database is updated.

With the data analysis work complete, I believe I am on track for delivery of the finished R model to UDAQ by the end of the month of April. Over the coming month I'm planning on building the necessary Monte-Carlo simulation functions for randomly drawing from the CPT tables. After connecting those functions to the previously written functions for calculating equipment-based emissions, this project will be completed. The User Manual documentation will be updated to reflect the changes made over the October 2015 version of the model. Please let me know if UDAQ staff would like to have an in-person training session to cover these changes.

In addition to the OGEI database analysis work, I also devoted some time to engaging Doug Henderer (Newfield Corporate Air Quality Manager) regarding model predictions for drilling and production over the next five years. I updated and ran the model to produce predictions for the 2016 – 2020 time period. Doug and I disagreed over the model's projections about future drilling activity over that time period. Ultimately, I believe we concluded that we're each tracking different things. Doug was tracking how many new news had been started (or spudded), while in the R model I track when wells first start producing. We agreed that there is a time delay between these two events, and I performed a regression analysis to determine what the best time delay was between energy prices and when a well first produces. At present, the model estimates the number of wells drilled using the equation:

$$W_t = a \cdot OP_{t-1} + b$$

where *W* is the number of wells that first produce in month *t*, *OP* is oil price, and *a* and *b* are fitted coefficients. However stated more generally, the model is:

$$W_t = a \cdot OP_{t-x} + b$$

where *x* is the time delay between past oil price and the number of wells that produce for the first time. In the regression analysis, I searched for the optimal value of *x* over the interval of 1 to 24 months, finding that the best fit (as measured by R^2 values) was achieved with $x = 3$ months. This raised the R^2 value of the oil price drilling model from ~0.70 to 0.74. However the revised fit does not substantially change the predicted drilling activity over the 2016 – 2020 period (which is ultimately driven by EPA expectations of returning to ~$60/bbl oil within the next few years). If requested I'd be happy to investigate this particular issue further.