Response to Charge Question on

Assessment of the Median Nitrate Concentration model for the New Jersey Highlands in Subwatersheds and Land Use Capability Zones

Report

of the

NJDEP Science Advisory Board

Judith Weis, Ph.D. (Chair)
Clinton J. Andrews, Ph.D., P.E.
John E. Dyksen, M.S., P.E.
Raymond A. Ferrara, Ph.D.
John T. Gannon, Ph.D.
Tavit Najarian, Ph.D.
Robert J. Laumbach, M.D., MPH
Peter B. Lederman, Ph.D., P.E.
Paul J. Lioy, Ph.D.
Robert J. Lippencott, Ph.D.
Nancy C. Rothman, Ph.D.
Michael Weinstein, Ph.D.
Anthony J. Broccoli, Ph.D.
Mark G. Robson, Ph.D.
David A. Vaccari, Ph.D., P.E.
Lily Young, Ph.D.

August 29, 2014

Final Report

Assessment of the Median Nitrate Concentration model for the New Jersey Highlands in Subwatersheds and Land Use Capability Zones

By the Water Quality and Quantity Standing Committee of the NJDEP Science Advisory Board

Dr. David A. Vaccari, Stevens Institute of Technology (Chair)
Mr. Thomas W. Amidon, Kleinfelder
Dr. Brian T. Buckley, Rutgers University EOHSI
Dr. Keith R. Cooper, Rutgers University
Dr. William J. Cromartie, Richard Stockton College of NJ
Dr. Josh T. Kohut, Rutgers, The State University of New Jersey
Dr. Xiaoguang Meng, Stevens Institute of Technology
Dr. Anthony S. Navoy, United States Geological Survey
Dr. Orren D. Schneider, American Water
Dr. Pen Tao, North Jersey Water Supply Commission
Dr. Christopher G. Uchrin, Rutgers University

August 29, 2014

The Water Quantity and Quality standing committee (WQQC) of the New Jersey Science Advisory Board (SAB) met by teleconference on June 18, 2014 and August 18, 2014 to discuss the questions posed by NJDEP:

1. Is the median nitrate concentration model appropriate for estimating median nitrate concentrations for planning purposes at a regional scale?
2. Are the input parameters to the model appropriate?
3. Are the model's results appropriate for use by land use capability zone?

It should be noted that the WQQC focused exclusively on the validity of the model[1] to predict median nitrate concentration, and not on how the model predictions might be used to regulate septic tank density.

The model estimates median groundwater nitrate concentrations in 2000-foot-square cells of the Highlands Planning and Preservation Areas of New Jersey as they depend upon certain local characteristics (New Jersey Highlands Council, 2008) , namely:

- Percent urban land use
- Percent agricultural land use
- Length of streams in the quadrant
- Number of known contaminated sites
- Septic tank density

The median nitrate concentrations would then be used to set the allowable domestic septic tank density for new development to satisfy anti-degradation requirements using the Nitrate Dilution Model (NDM). The NDM has previously been found to be appropriate for estimating the effect of nitrate loading on downgradient groundwater quality at the medium-field scale by the SAB (NJDEP Science Advisory Board, 2011).

The model was based on two sources. Data were obtained from 352 wells in the USGS National Water Information System (NWIS). These data have been well qualified by QA/QC procedures, but do not sufficiently cover the geographical area and/or the range of land use characteristics. To supplement these data, results from an additional over 19,000 wells were obtained from the Private Well Testing Act (PWTA) program. These were aggregated into 4,379 2000-foot-square cells and combined with the USGS data. The analyses in the PWTA dataset were conducted by NJDEP-certified laboratories, although they were not as fully qualified as were the NWIS data.

[1] Primary model assumptions can be found in Appendix A of this document

The modeling effort consisted of the following: A set of multivariate logistic equations was developed, each to predict the probability of exceeding a target nitrate concentration ranging from 0.05 mg NO_3-N/L to 10.0 mg/L NO_3-N/L (written communication, Ron Baker, 2014)

> *"In contrast to the more common use of logistic regression (such as determining the probability of exceeding a target value of 5.0 mg/L NO_3-), the objective of this analysis is to determine the median concentration of nitrate in areas of varying scale in the Highlands. The median nitrate concentration is found by identifying the logistic equation for which p=0.5, which occurs where half the estimated nitrate concentrations are greater than the target concentration, C_t. To accomplish this, a set of all logistic equations with C_t values of 0-10 mg/L in increments of 0.05 mg/L up to 1.0 mg/L, and increments of 0.1 mg/L up to 10 mg/L were developed. The p-value of each equation was calculated for each grid square, and the two equations with p-values closest to 0.5 (one slightly greater than and one slightly less than 0.5) were identified. The estimated nitrate concentration in the grid cell was determined by linear interpolation."*

The results were used to estimate the distribution of median nitrate concentrations at the HUC 14 subwatershed level. Comparison with the measured distribution showed a strong linear relation.

The WQQC requested validation of the model by applying it to only the well-validated individual NWIS data. The linear correlation between the data and the predictions was about 22%. Although noisy, this was judged to be statistically significant. This is especially important considering that this validation was done at the level of the individual well, whereas the model will be applied at the regional level.

The WQQC also observed that not all the input parameters were statistically significant for all of the logistic models. For example, the number of contaminated sites was only significant for models with target concentrations less than or equal to 1.0 mg NO_3-N/L. However, given that the validation results were reasonably good, the WQQC agreed that including these variables in the model would not significantly adversely affect its predictions.

Findings

The WQQC recommends the SAB adopt the following findings for the charge questions:

1. Is the median nitrate concentration model appropriate for estimating median nitrate concentrations for planning purposes at a regional scale?

Yes, the model significantly accounts for the effect of the input parameters on median nitrate concentrations in groundwater at the regional level.

2. Are the input parameters to the model appropriate?

Each of the variables is significant to at least some of the logistic models, and do not seem to adversely affect the overall predictions. Thus it is appropriate to include them in the modeling effort.

3. Are the model's results appropriate for use by land use capability zone?

Yes. The combination of choice of input parameters and the modeling approach is able to generate estimates of median nitrate concentrations that take into account land use effects at the regional level that are improvements over estimates that could be made in the absence of this approach.

References:

Hoffman, J.L. and A. Petriman, (2014) "Nitrate Concentrations in Groundwater of New Jersey's Highlands Region, NJDEP Technical Memorandum 14-1

New Jersey Department of Environmental Protection Science Advisory Board, (2011) "Response to Charge Questions on the Nitrate Dilution Model, Summary Report" (http://www.state.nj.us/dep/sab/nitrate-dilution-model-report.pdf)

New Jersey Highlands Council, (2008), New Jersey Highlands Regional Master Plan, Technical Report: Water Resources Volume 1, Watersheds and Water Quality, "Nitrate concentrations and septic system density of the Highlands Region", p. 114-173.
Website: http://www.highlands.state.nj.us/njhighlands/master/tr_water_res_vol_1.pdf

Appendix A:

Highlands Logistic Regression Model Assumptions

The following are some of the model assumptions for the 2012 Highlands Logistic Regression Model that includes the Private Well Testing Act (PWTA) data. The 2012 model is based on a 2008 Logistic Regression Model used by the Highlands Council and discussed in their 2008 Technical Report.

1. The model assumes that the optimal set of predictive variables, determined during the analysis of the original 350 NWIS wells in 2008 (and reported in the Highlands Council 2008 Technical Report), are the same variables that are significant when using the PWTA data set (e.g. percentage of urban land use, septic system density, percentage of agricultural land use, lengths of streams and number of known contamination sites).

2. The model assumes that the laboratory reported values for nitrate concentration that are sampled from several localized sites, are taken from a well-mixed groundwater system and represent the overall ground water quality of that area.

3. The model assumes that by adjusting the non-detect values reported by the lab up to the reporting limit, that these values will not bias the resultant median values because of the inherent nature of a median value.

4. The logistic regression model assumes that the predictor variables are independent of each other.

5. The model assumes the wells used to obtain the PWTA nitrate concentration values are similar to those used in the 2008 NWIS evaluation.

6. The model assumes that the nitrate concentration is equal to the lab reported Nitrate + Nitrite, assuming that the nitrite is negligible.

7. The model assumes that the water quality (including nitrate concentration) underlying each grid cell is directly affected by surface activities in that grid cell