ANNUAL REPORT:

Improving BASINS/HSPF predictions of nitrogen export to improve TMDL accuracy using NASA imagery

Investigators:
Philip A. Townsend, University of Wisconsin
Angélica L. Gutiérrez-Magness, University of Maryland
Keith N. Eshleman, University of Maryland Center for Environmental Science
Brenden E. McNeil, West Virginia University

Project funded by NASA under:
Decision Support through Earth Science Research Results
Solicitation NNH08ZDA001N-Decisions

NASA Award #:
NNX09AO15G

Project Overview

Mathematical models are a critical element in predicting the effects of human activities and natural processes on the quality of water in lakes and streams. Water quality decision support systems such as BASINS (EPA software based on the hydrological simulation model known as HSPF) are presently being used to simulate pollutant loads from large complex multi-use watersheds in order to partition the total load of various constituents into more management-relevant components, thus improving the efficacy of nutrient management efforts. Such partitioning is an important aspect of establishing and addressing total maximum daily loads (TMDLs) of various pollutants—a task presently mandated by USEPA under the Clean Water Act regulations. Some of the more useful applications of BASINS in this regard include (1) categorizing non-point source loads as either controllable or non-controllable to facilitate establishing meaningful tributary nutrient reduction goals for Chesapeake Bay subwatersheds; and (2) estimating pollutant loads by major land use category (e.g., forest, agriculture, urban, etc.). Remote sensing imagery can play a significant role in (1) partitioning pollutant loads among different land cover types; (2) aiding in the detection of changes in nutrient retention due to land cover conversions (i.e., the permanent change from one land cover type to another); and (3) incorporating landscape dynamics (i.e., seasonal changes in vegetation density or properties) into simulations of nutrient loads over time. **The primary activities of the project are: (1) derivation of annual image-derived metrics of landscape condition; (2) incorporation of the derived metrics into the EPA/BASINS software to demonstrate its use for predicting surface water N concentrations and loads; and (3) interacting with end-users to incorporate the improvements into the decision-making process.**

Project Schedule – Year 1

This report is being provided at the nine-month point of Year 1 of the overall project. This report reflects accomplishments for the partial first year of the project. As indicated in the proposal, Year 1 objectives were focused on "model/data development and preliminary testing," with the following specific goals (year-by-year objectives for all years are listed in Appendix 2):

Remote Sensing (Townsend): Development of remote sensing indices of disturbance and testing against stream nutrient data. Key milestone: Development of the remote sensing index for assessing spatial variability in N loads from forests across years.
Modeling/HSPF-BASINS (Gutiérrez-Magness): Adaptation and modification of HSPF to accept spatially varying (by basin or landscape segment) measures of forest N load, development of simulation runs to test both baseline (existing) and modified model. Key milestone: Model modifications prepared, initial simulations prepared.
Hydrology (Eshleman): Development of hydrology data sets, both for assessing remote sensing imagery (Townsend) and for simulation data sets (Gutiérrez-Magness). Key milestone: Hydrology data sets assimilated, with preliminary analyses.
Geospatial analysis (McNeil): Assimilation of geospatial data sets (in conjunction with other PI's) to test differences in forest N export across physiographic/biogeographic/edaphic/climatic boundaries. Key milestone: Geospatial data sets assimilated, preliminary analyses with respect to hydrology data.

Year 1 Accomplishments

Remote Sensing (Townsend, University of Wisconsin)
The remote sensing team has been developing MODIS data sets and conducting statistical analyses to generate the image-derived parameters needed to modify nutrient export predictions from forested and mixed-use watersheds. To do this, we are using two extensive datasets of stream water quality including total dissolved nitrogen (TDN), nitrate-N and phosphorus (P) collected in 324 watersheds between 2001-2004 (Figure 1, Robertson et al. 2006, Stanley and Maxted 2008). The watersheds cover a range of forest composition and proportion of forest cover. Our objective was to use "off-the-shelf" (or nearly off-the-shelf) MODIS data to predict nutrient concentrations. Watershed-based MODIS coefficients for forests would then be provided as input factors for the HSPF modeling.

Our analytical design included the acquisition of MODIS data from each year of sampling (and the year previous to a samole) and the development of empirical relationships between MODIS parameters and water quality variables. The MODIS parameters we use are derived for spring summer and fall (i.e., three per year) and include vegetation indices (NDVI, EVI), gross primary productivity (GPP), the tasseled cap index (following Lobser and Cohen 2007) and related disturbance index (Healy et al. 2005), and estimates of fractional cover (green vegetation, non-photosynthetic vegetation [NPV] and soil) derived from spectral mixture analysis. Key concepts of the effort include:
(1) MODIS parameters are derived for an entire watershed;
(2) MODIS parameters include means and standard deviation per watershed. Use of means alone facilitates pixel-wise estimates of nutrient export, whereas use of means and standard deviations facilitates development of stronger models, but predictions can be applied only on a watershed basis;
(3) Only MODIS variables ***from the year prior*** to a stream measurement are employed in the empirical model, although the statistical model is developed using all stream data, regardless of year of sampling. That is, if Stream A was sampled in 2001, its NDVI values would be from 2000, while Stream B (sampled in 2003, for example), would use NDVI values from 2002. Both Stream A and Stream B would be used for the empirical model;
(4) Statistical modeling to identify important MODIS variables involves a new implementation of partial least-squares regression that is under development by Ph.D. student Aditya Singh. Briefly, the new method optimizes model development and variable selection to limit the variables used for predicting (in our case) nitrate-N export as a function of a relatively small number of interpretable variables compared to the large population of potential image variables;

(5) Model predictions can be applied to any year for which MODIS data are available, including years for which no water quality data exist. *This was a major objective of the research, and our results for Wisconsin indicate that the models are robust across years;*
(6) Model evaluation has followed two approaches at this point. First, we have developed separate models using either the Robertson et al. (2006) or Stanley and Maxted (2008) data, and then validated against other data set. Second, we have split the samples by year – developing the model using 3 years of data and cross-validating against a fourth. All results have been stable.
This initial effort for Wisconsin is nearing completion, with a full manuscript currently being edited by co-authors[1].

At present, our analyses indicate that the most important variables for predicting nitrate-N export from watersheds include spring EVI, spring GPP and summer green vegetation fraction. NPV fraction and standard deviation of soil fraction are also important. We found that we can predict nitrate-N across all years with an $R^2 = 0.8005$. Cross-validation of the model derived from Stanley and Maxted (2008) using Robertson et al. (2006) yielded $R^2 = 0.635$, while cross-validation of the model derived from Robertson et al. data $R^2 = 0.789$. The modeling effort employed all watersheds, regardless of percent forest. Because our work is focused specifically on improving predictions from forested watersheds, we also analyzed model predictions according to forest cover. ***Cross-validated model predictions for watersheds >50% forest always exceeded $R^2 = 0.672$.*** An example of model predictions of nitrate-N export for Wisconsin watersheds is provided in Figure 2.

While this portion of the work is being finalized, we will soon commence with the development of predictions for study watersheds in the Chesapeake Bay Watershed.

<u>Modeling/HSPF-BASINS (Gutiérrez-Magness, University of Maryland)</u>
1. We have assembled all of the GIS layers and meteorological data using TRMM-NASA for basins located in Wisconsin. We are now preparing run models using the simplified HSPF.
2. We have completed the GIS work and development of meteorological data for the Big Run watershed (Chesapeake Bay).
3. Current work by Gutiérrez-Magness is focused on the implementation of previous work by Eshleman and others (Eshleman et al. 2000) for nitrogen into the Phase 5.0 of the Chesapeake Bay Watershed model. This has involved FORTRAN Code modification to allow the Chesapeake Bay Model to accept information on forest disturbance to the model. Since March, 2010, this effort includes collaboration on a weekly basis with Ping Wang at the Chesapeake Bay Program Office (CBPO) to complete this work.

<u>Hydrology (Eshleman, UMCES Appalachian Laboratory)</u>
Since the beginning of the project in 2009, the Appalachian Laboratory group has been working to assemble datasets from multiple small forested watersheds in the Chesapeake Bay basin for use in the water quality modeling effort. We have already produced water quality (i.e., nitrate-N concentrations) and continuous hourly discharge datasets for the Upper Big Run (MD) watershed for the period 1995-2008. We have made these data available for testing by the modeling group; these datasets expand the database that was used by Eshleman *et al.* (2008) in an analysis of long-term trends. We are presently involved in producing comparable datasets for several other watersheds for which we have relatively long records, including: Black Lick, Deep Run, and Terrapin Run (all in western MD). After

[1] A. Singh, A. Jakobowski, I. Chidster and P. Townsend, A remote sensing approach to predicting stream water quality in Wisconsin, for submission to *Ecological Applications*.

preliminary model testing has been completed, we will be working closely with the modeling group to analyze the results in the context of understanding the primary factors—and their representation using remote-sensing imagery—governing nitrate-N leakage from the study watersheds.

Geospatial Analysis (McNeil, West Virginia University)
Year 1 activities have included the development of a new disturbance history for the Western Maryland watersheds (Fifteenmile Creek and Savage River) by graduate student Lindsay Deel. Deel is considering working with Gutiérrez-Magness and the CBPO personnel in Annapolis for her Ph.D. dissertation. The West Virginia University team has begun compiling the geospatial data in conjunction with streamwater, lysimeter, and disturbance history data for Western Maryland that will be used to generate models characterizing N export with respect to spatial variability in forest species composition and disturbance history.

Coordination with Agencies

Co-Investigator Angélica Gutiérrez-Magness is our chief liaison to management agencies. She is already working extensively with Ping Wang at the CBPO to implement our proposed changes in the operational version of the Chesapeake Bay Model (based on HSPF). Because of this interaction, our plan to link our work with EPA (proposed for Years 3 and 4) is already underway. Moreover, WVU Ph.D. student is considering working directly out of the CBPO to provide a direct on-site link between the geospatial work and the modelers.

Year 2 Plan

While the Year 1 effort has been mostly focused on dataset development and generation of parameters from remote sensing data, Year 2 will be directed more specifically at incorporation of the new data sets and forest parameters into the water quality modeling using HSPF.

Remote Sensing (Townsend): Finalize remote sensing-nutrient export relations, as modified by geospatial variables. Key milestone: Provide remote sensing measures for simulation runs.
Modeling/HSPF-BASINS (Gutiérrez-Magness): Simulation runs using remote sensing measures. Key milestone: Simulations run and compared against baseline.
Hydrology (Eshleman): Finalize analyses of watershed nutrient export. Key milestone: Hydrological relations tested and validated, in conjunction with modeling.
Geospatial analysis (McNeil): Finalize geospatial nutrient export functions from forests. Key milestone: With Townsend, validate geospatial relationships, including remote sensing measures.

References

Eshleman, K. N., R. H. Gardner, S. W. Seagle, N. M. Castro, D. A. Fiscus, J. R. Webb, J. N. Galloway, F. A. Deviney, and A. T. Herlihy. 2000. Effects of disturbance on nitrogen export from forested lands of the Chesapeake Bay watershed. *Environmental Monitoring and Assessment* 63:187-197.

Eshleman, K.N., K.M. Kline, R.P. Morgan II, N.M. Castro, and T.L. Negley. 2008. Contemporary trends in the acid-base status of two acid-sensitive streams in western Maryland. *Environmental Science and Technology* 42:56-61.

Robertson, D. M., Graczyk, D. J., Garrison, P. J., Wang, L., Laliberte, G., Bannerman, R. 2006. Nutrient Concentrations and their relations to biotic integrity of wadeable streams in Wisconsin. United States Geological Survey Professional Paper 1722.

Stanley, E. H., Maxted, J. T. 2008. Changes in the dissolved nitrogen pool across land cover gradients in Wisconsin streams. *Ecological Applications* 18(7):1579-1590.

Figure 1. Watersheds used for the MODIS testing, provided by (A) Robertson et al. (2007) and (B) Stanley and Maxted (2008).

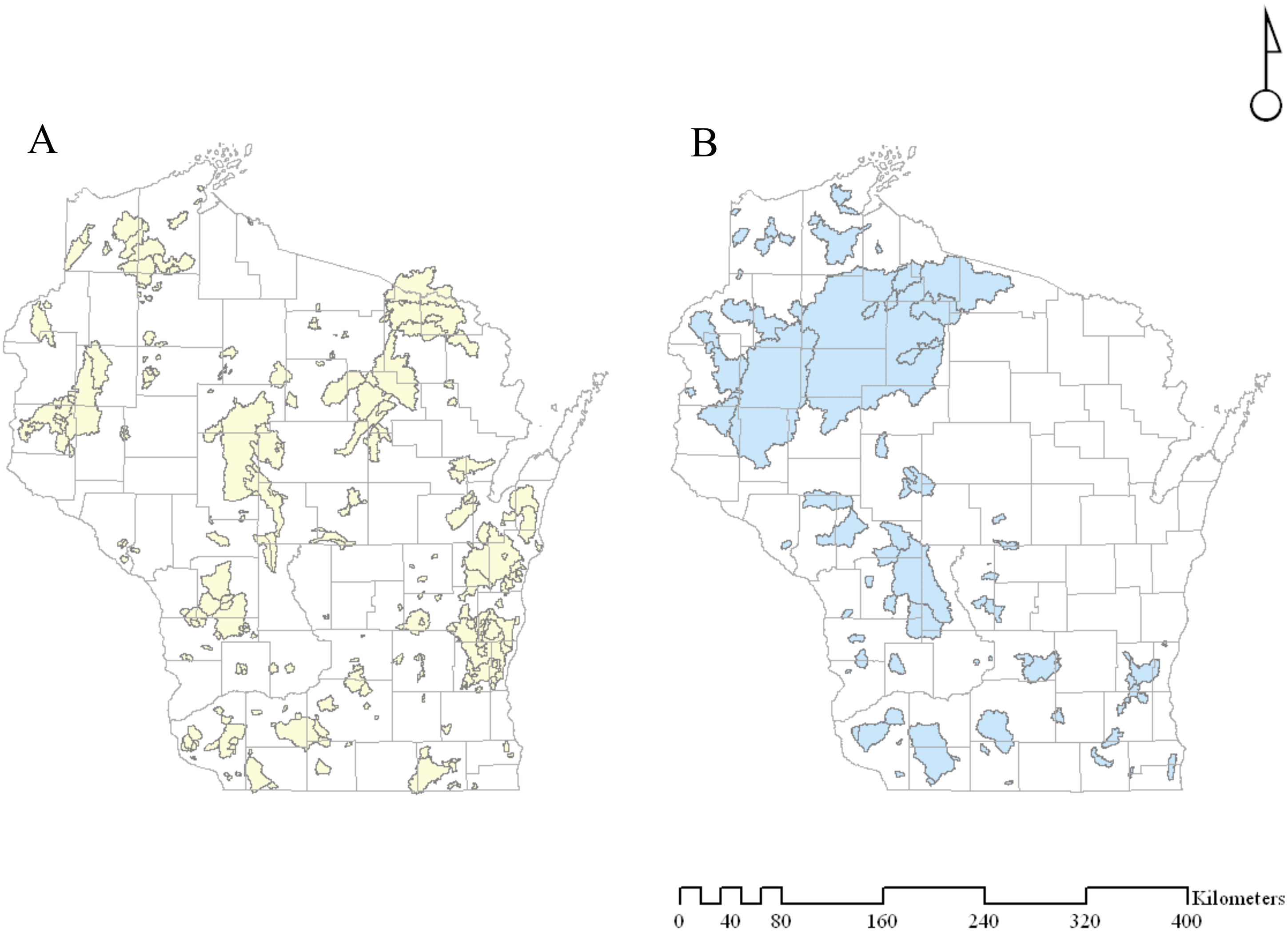
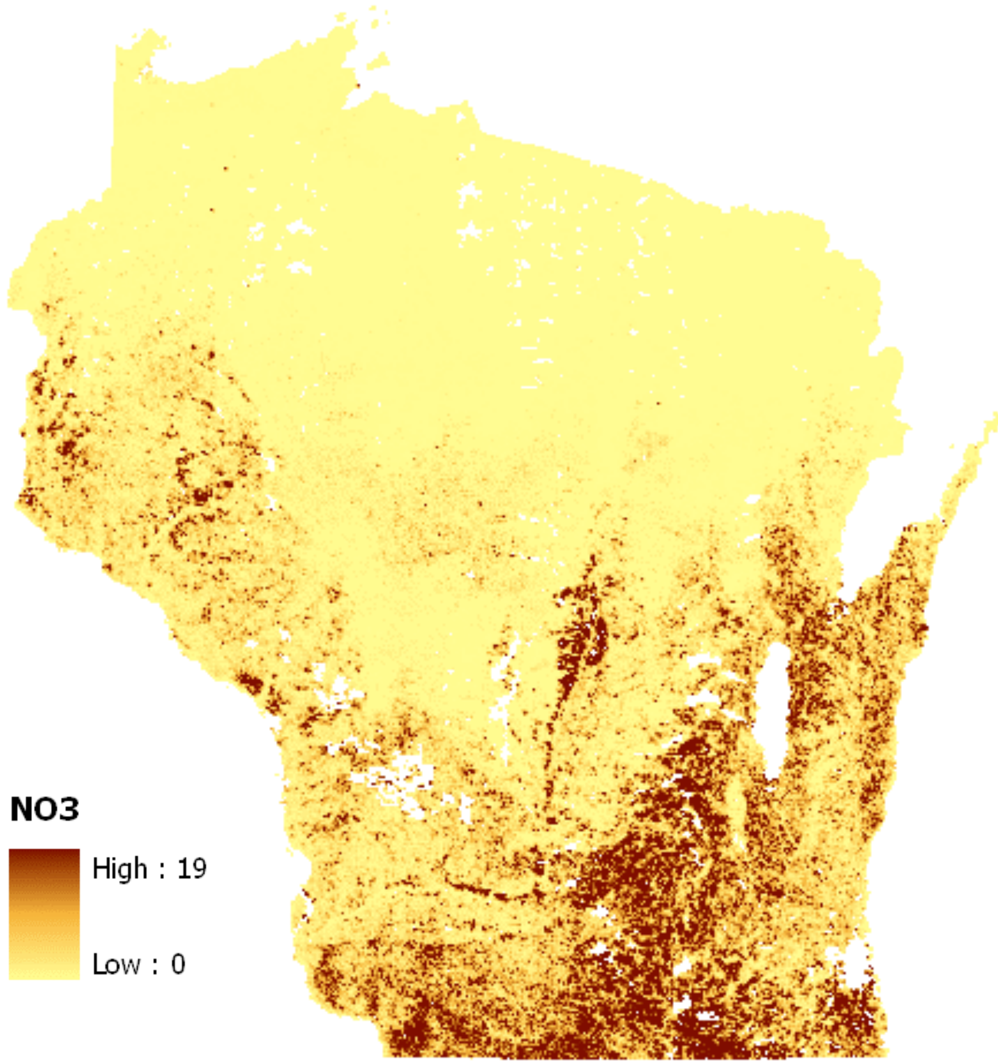


Figure 2. MODIS-derived predictions of nitrate-N export for Wisconsin, 2004. This model is generated using MODIS data from 2003.



APPENDIX A – Project Schedule from the Original Proposal

Year 1 – Model/Data Development and Preliminary Testing

Remote Sensing (Townsend): Development of remote sensing indices of disturbance and testing against stream nutrient data. Key milestone: Development of the remote sensing index for assessing spatial variability in N loads from forests across years.

Modeling/HSPF-BASINS (Gutiérrez-Magness): Adaptation and modification of HSPF to accept spatially varying (by basin or landscape segment) measures of forest N load, development of simulation runs to test both baseline (existing) and modified model. Key milestone: Model modifications prepared, initial simulations prepared.

Hydrology (Eshleman): Development of hydrology data sets, both for assessing remote sensing imagery (Townsend) and for simulation data sets (Gutiérrez-Magness). Key milestone: Hydrology data sets assimilated, with preliminary analyses.

Geospatial analysis (McNeil): Assimilation of geospatial data sets (in conjunction with other PI's) to test differences in forest N export across physiographic/biogeographic/edaphic/climatic boundaries. Key milestone: Geospatial data sets assimilated, preliminary analyses with respect to hydrology data.

Year 2 – Simulation Modeling and Comparison

Remote Sensing (Townsend): Finalize remote sensing-nutrient export relations, as modified by geospatial variables. Key milestone: Provide remote sensing measures for simulation runs.

Modeling/HSPF-BASINS (Gutiérrez-Magness): Simulation runs using remote sensing measures. Key milestone: Simulations run and compared against baseline.

Hydrology (Eshleman): Finalize analyses of watershed nutrient export. Key milestone: Hydrological relations tested and validated, in conjunction with modeling.

Geospatial analysis (McNeil): Finalize geospatial nutrient export functions from forests. Key milestone: With Townsend, validate geospatial relationships, including remote sensing measures.

Year 3 – Technology Transfer and Model Implementation

Although we will coordinate with our collaborators in all years of the project, in Year 3, we will work directly with EPA, MDE and CBPO to transfer the methodological modifications to our cooperators. We will particularly focus on the automation of parameter development from remote sensing measures. Then, the primary effort in Year 3 will be to implement simulation runs *in conjunction with* the cooperators to evaluate model performance under simulation scenarios of their choosing. All PI's will continue their efforts under the roles already defined: Townsend (remote sensing), Gutiérrez-Magness (modeling), Eshleman (hydrology) and McNeil (geospatial analysis). Key milestone: transfer of model enhancements to collaborating agencies, development and review of documentation.

Year 4 – Agency Implementation

The main efforts of the project will be completed during Year 4. In this year, we will consult with the cooperators to modify and streamline model enhancements, and test new NASA inputs, should data from new satellites be available (2012-2013).