Mapping Amyotrophic lateral sclerosis lake risk factors across northern New England

Nathan Torbick: ntorbick@appliedgeosolutions.com
Applied Geosolutions, LLC, Newmarket, New Hampshire, USA
Elijah Stommel, Tracie Caller, Dartmouth Hitchcock Medical Center
Sarah Hession, Michigan State University

Overview

Concern over water quality and human health is rising as anthropogenic eutrophication and climate change are increasing the frequency, magnitude, and intensity of harmful algal blooms. The cyanotoxin, Beta-Methylamino L-Alanine (BMAA), has been identified as a risk factor for Amyotrophic Lateral Sclerosis (ALS). ALS is a progressive, fatal neurodegenerative disease with a lifetime risk of 1 in 700. Despite many recent discoveries about the genetics of ALS, the etiology of sALS remains largely unknown. There is a broad scientific consensus that ALS is caused by gene-environment interactions. Evidence has shown potential linkages between water quality, cyanobacteria, and high ALS incidence. The goal of this project is the **development of an eco-epidemiological model that uses satellite remote sensing water quality and cyanobacteria exposure maps to investigate lake water quality and ALS incidence**.

Figure: NASA EOS Terra MODIS image. Lake Erie western basin, Microcystis aeruginosa bloom

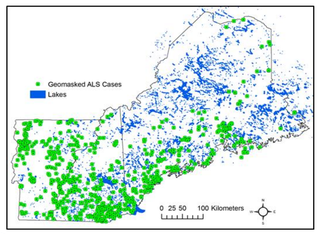
Data

❑ Approach relies on multiple satellites:

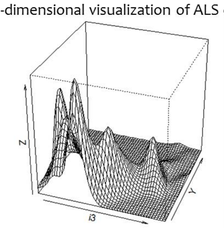
- Landsat 5 TM & 7ETM+, and LDCM scale imagery
  - ❖ 8-day intervals, 30m spatial, 180km swath, broad vis-nir channels
- MERIS
  - ❖ ESA platform, ~weekly Full Resolution, ~240m spatial, 15 narrow bands across vis-nir domain
- MODIS
  - ❖ high temporal frequency, 250m -1km spatial, large area, 2001- present

❑ Human health ALS database

- Dartmouth Hitchcock Medical Center is compiling an ALS database with more than 800 cases across northern New England. Records from DHMC, the Muscular Dystrophy Association of Northern New England, and surveys were searched to identify cases of ALS diagnosed between January 1997 and October 2009. When possible, we confirmed accuracy of diagnosis, year of diagnosis, and demographic history of patients identified by review of medical records, the Social Security Death Index, obituaries, and data supplemented from questionnaires.
- Strict adherence to privacy and high standards for handling human health data



Three-dimensional visualization of ALS cases



Approach & Results

- ❑ Operationalize Landsat maps of lakes >8 hectares for wall-to-wall mapping of Chl-a, TN, and SD; band ratio empirical regression using MODTRAN corrected water leaving radiances and in situ observations
- ❑ Operationalize cyanobacteria metrics for lakes over >100 hectares with MERIS using spectra derivative shape filters

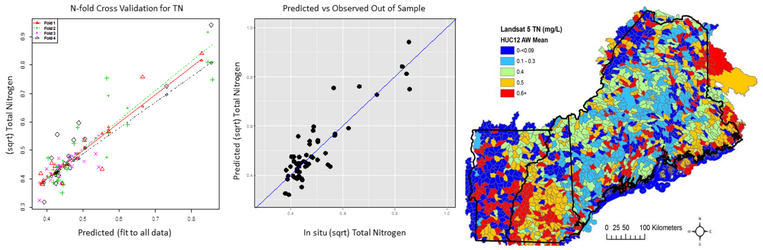


Figure: Cross validation (left) of Landsat Total Nitrogen algorithm, out of sample evaluation statistics of modeled vs observed (middle), and map of ake area weighted sqrt TN by watershed (HUC12) scale (right).

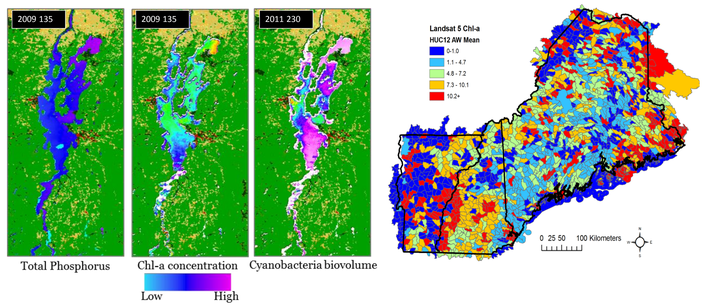


Figure: Large lakes (e.g., Champlain) using MERIS FR1b to map TP, Chl-a, cyanobacteria biovolume (left) and Landsat 5 TM water leaving radiance for mapping maximum lake area weighted (AW) mean Chl-a at watershed (HUC12) scales (right).

❑ Spatial clustering using expected counts and LISA: $c_i = \frac{(o_i - e_i)}{e_i}$

where oi represents the number of observed ALS cases per census tract and ei represents the expected number of cases normalized for age, sex, and population

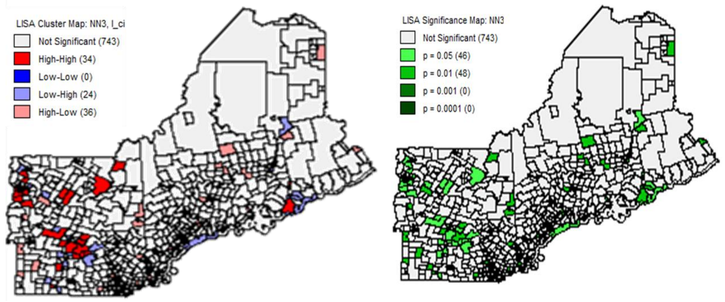


Figure: Local Indicator of Spatial Autocorrelation Map with a tract-level aggregation of ALS cases represented as normalized excess case counts (left). A map of the levels of statistical significance where clusters are present is shown on the right.

Results Cont.

❑ **Logistic regression**: $F(Y_i)=\beta_0+\beta_1 X_{1i}+ \beta_2 X_{2i}+ +\cdots+\beta_k X_{ki}+\varepsilon_i$

where Yi represents the dichotomous dependent variable (1 = belongs to a hot spot, 0 = does not belong to a hot spot), X_1,…,X_k represent the independent variables or potential risk factors, F(Y_i) is a logit link function to transform the binary dependent variable to the appropriate scale for estimation of regression coefficients (β_1,…,β_k) and εi represents random residual errors

❑ **Cluster membership variogram**

❑ **Logistic regression residuals variogram**

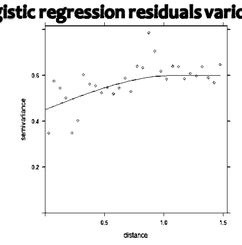


Table: Summary of selected logistic regression model for hot spot membership. p-values are interpreted as follows: < 0.05 significant at the 95% significance level; Odds Ratio (OR) represents the odds of belonging to a hot spot; OR > 1 indicates increasing odds of belonging to a cluster for each unit increase in the independent variable, OR <1 indicates decreasing odds of belonging to a cluster for each unit increase in the independent variable. Lower confidence limits (LCL) and upper confidence limits (UCLs) are provided.

						95% Confidence Interval	
Significant risk factors	Coefficient	Std. Error	z value	p-value	OR	LCL	UCL
(Intercept)	-0.066	0.972	-0.07	0.9461	0.94	0.142	6.498
SD30km	-0.921	0.231	-3.99	0.0001	0.40	0.249	0.614
TN30km	0.884	0.262	3.38	0.0007	2.42	1.460	4.124
CHLA10km	0.045	0.024	1.83	0.0678	1.05	0.996	1.098

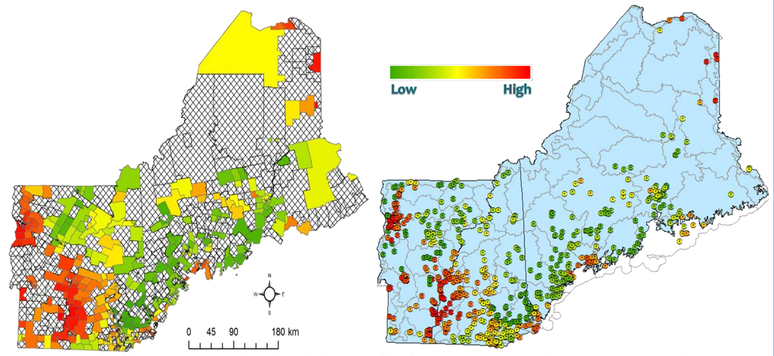


Figure: Risk of belonging to a localized cluster of higher than expected ALS counts based on eco-epidemiological model by census tract (left) and ALS patient location (right) scales. Risk displayed as low (green) to high (red).

Conclusions

- ❑ Strengths of satellite remote sensing information can help overcome traditional field limitations
- ❑ Broadly, poorer lake water quality was significantly associated with increased odds of belonging to an ALS cluster in the region
- ❑ Findings support the hypothesis that sALS can, in part, be driven by environmental water-quality lake conditions that promote harmful algal blooms
- ❑ Geographic scale needs to be diligently considered when evaluating relationships between ecological processes, risk factors, and human health outcomes