Frontiers in Ecology and the Environment

A century of changing land-use and water-quality relationships in the continental US

Whitney Broussard and R Eugene Turner

***Front Ecol Environ* 2009; doi:10.1890/080085**

This article is citable (as shown above) and is released from embargo once it is posted to the *Frontiers* e-View site (www.frontiersinecology.org).

A century of changing land-use and water-quality relationships in the continental US

Whitney Broussard[1,2*] and R Eugene Turner[1]

We quantify the relationships between riverine nitrate–nitrogen (NN) concentration and agricultural land use in the continental United States – from the early 1900s through the end of the last century – on spatial scales ranging from the entire Mississippi River Basin to 1000 km^2 watersheds. Cropland cover is linearly related to the NN concentration that exits a watershed at both the beginning and end of the 20th century. In addition, the slope of the relationship is higher at the end of the century, and the intercept of the regression analysis is not different from zero. These findings imply that agriculture was already affecting NN export by the early 1900s, that intensive management practices in modern agriculture have significantly increased the NN export per hectare of cropland, and that the baseline of exported NN has not shifted. We identify agricultural practices, principally associated with corn cultivation, that contribute substantially to NN concentrations and suggest that increasing cropland diversity and perennial plant cover can reduce NN loading.

Front Ecol Environ 2009; doi:10.1890/080085

The global production rate of biologically available nitrogen (N) in terrestrial ecosystems has doubled since 1960, primarily resulting from human activity that consumed nearly eight times more nitrogenous fertilizer in 2003 than in 1960 (MA 2005). This human-derived N fixation now produces more biologically available N than all other natural sources combined (Galloway *et al*. 1995). Food supplies increased substantially with the industrial production of N-based fertilizer, so that now, one-third of humanity's protein consumption depends on synthetic N fertilizer (Smil 1997). However, excessive nutrient loading has become an important driver of ecosystem change in terrestrial and aquatic ecosystems and is projected to increase in the 21st century (MA 2005).

Previous studies have identified agricultural landscapes as sources of riverine N (Howarth *et al*. 1996; Goolsby *et al*. 1999; McIsaac *et al*. 2001; Boyer *et al*. 2002; Donner 2003; Turner and Rabalais 2003) and have demonstrated how N loading contributes to the formation of oxygen-depleted bottom waters in coastal systems (Rabalais *et al*. 1999 2007; Turner *et al*. 2006). Nevertheless, there are still uncertainties about the baseline values necessary for restoration of riverine water quality (Smith *et al*. 2003).

This study examines the relationships between various land-use practices and riverine N at the beginning and end of the 20th century in the continental United States. We studied 56 watersheds – ranging in size from the Cache River Basin (976 km^2) to the Mississippi River Basin (2 937 502 km^2) – to test the hypothesis that there is a statistically significant, quantifiable association between agricultural land-use practices and riverine nitrate-nitrogen (NN) concentrations. The total N flux from the Mississippi River Basin over the period of 1980–1999 increased nearly threefold as compared with that between 1955–1970, due almost entirely to increases in nitrate, which is the most abundant form of N (62%) at the mouth of the Mississippi River (Goolsby and Battaglin 2001). We focused on the variability of NN concentrations by watershed because (a) NN is the most abundant dissolved N ion in river water and (b) data on NN concentrations are available for the first part of the 20th century.

[1]*Department of Oceanography and Coastal Resources, Louisiana State University, Baton Rouge, LA;* [2]*current address: Department of Renewable Resources, University of Louisiana at Lafayette, Lafayette, LA* [*]*(wbroussard@louisiana.edu)*

Methods

Historical water-quality data

We collected two basic types of data (land-use records and water-quality records for each study watershed, from the beginning and end of the 20th century) to test the hypothesis that land use and water quality are related. Data on the NN concentration in rivers in the early 1900s are from Clarke (1924). Contemporary water-quality data are from the US Geological Survey (USGS) National Water Information System (USGS 2007). A 5-year grand mean was calculated from the annual means in order to match the 5-year intervals between the Census of Agriculture data described below. We identified current USGS monitoring stations located on the same river reaches as the Clarke study stations, to compare the historical and contemporary water-quality data. The drainage basins (watersheds) for each station were digitally delineated using the National Hydrography Dataset Plus (USEPA and USGS 2005). We identified the study watersheds that are not encompassed by the areal extent of another watershed and present the results only from these independent watersheds.

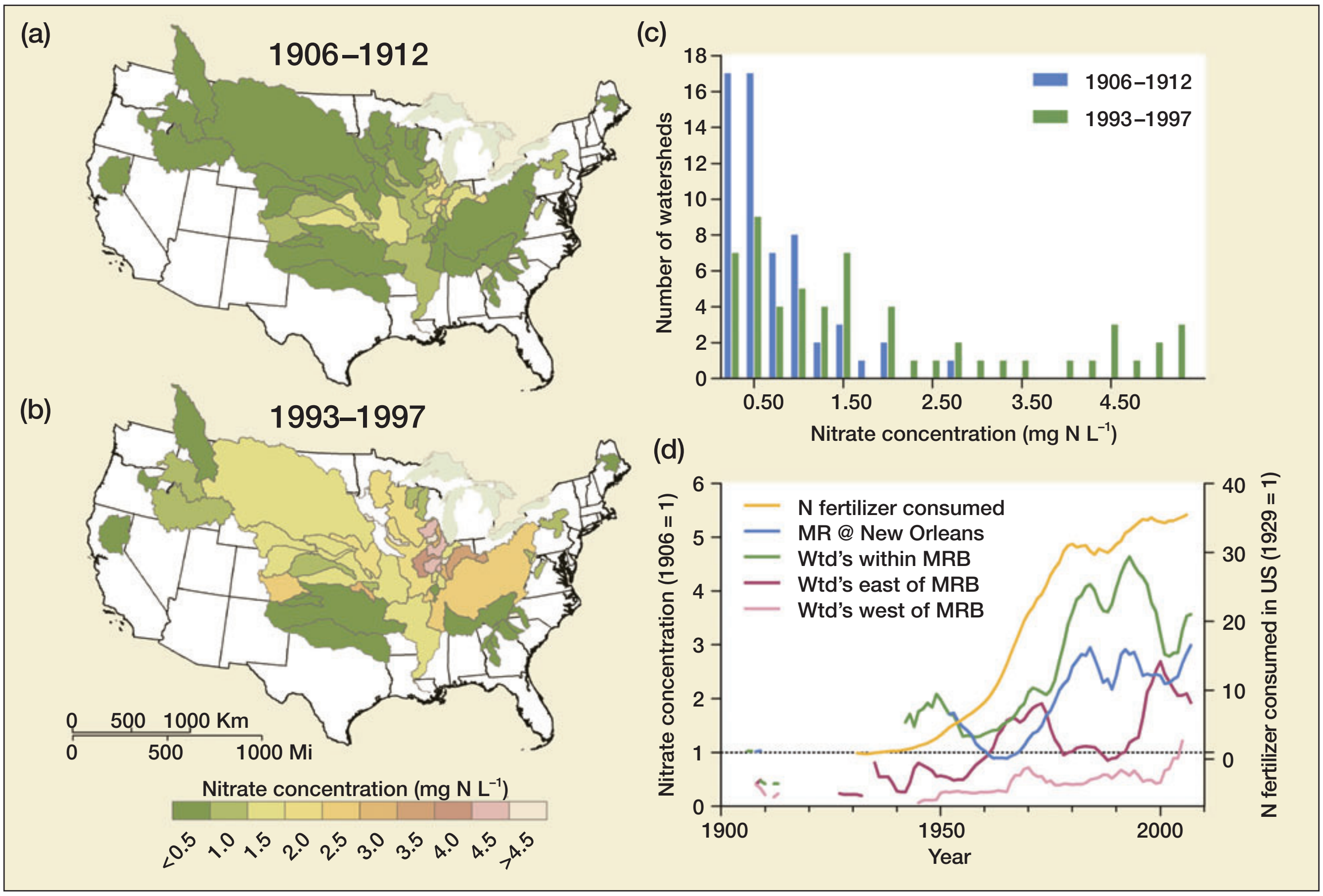

Figure 1. *The magnitude and spatial distribution of 5-year averaged nitrate–nitrogen (NN) concentrations by watershed (wtd) for (a) 1906–1912 and (b) 1993–1997. (c) NN concentration frequency histogram for the same watersheds in 1906–1912 and 1993–1997. (d) A 5-year running average, normalized time series (Mississippi River [MR] at New Orleans, 1906 = 1) of NN concentrations for the same watersheds averaged by region and total US N-fertilizer consumption (1929 = 1). MRB = Mississippi River Basin.*

Historical land-use data

We developed a GIS (Geographic Information System) database using county-level Census of Agriculture data, to quantify changes in land use over the past 100 years. Census of Agriculture data from 1850 to 1910 are from public document ICPSR 2896 (Haines and ICPSR 2004), whereas those from 1954 to 2002 were organized into datafiles from published data records by M Haines and colleagues at Colgate University (USDA 1979, 2005; US Bureau of the Census 1990, 2001). All Census of Agriculture data were spatially referenced by attributing county boundaries with the county-level data tables (Earl *et al.* 1999). We summed the county-level land-use data by watershed, using techniques similar to those described by Boyer *et al.* (2002). WebPanel 1 provides a detailed explanation of the database preparation and area calculations.

Results and discussion

Water-quality changes

The grand mean NN concentration for all study watersheds increased threefold from 1906–1912 to 1993–1997 (0.60 ± 0.07 to 1.79 ± 0.24 mg N L^{-1}, $\mu \pm$ SE, $t = -5.90$, $P < 0.0001$). The minimum and maximum values were 0.03 and 2.71 mg N L^{-1} over the period 1906–1912, and 0.06 and 8.54 mg N L^{-1} over the period 1993–1997, respectively. The NN concentrations increased more than tenfold in the Iowa, Minnesota, and Des Moines Rivers.

The spatial distribution of the NN variation among watersheds is similar in both time periods, with the higher concentrations in the Midwest and the lower concentrations in the Northwest, southern Great Plains, and the Piedmont (Figure 1a, b). In 1906–1912, only one watershed (out of 56) reported a NN concentration higher than 2 mg N L^{-1}, whereas 18 watersheds reported NN concentrations higher than 2 mg N L^{-1} in 1993–1997. Thirty-three watersheds (out of 56) in 1906–1912 and 16 watersheds in 1993–1997 had NN concentrations below 0.5 mg N L^{-1}. Figure 1c is a frequency histogram of the NN distribution in all the watersheds for both time periods that shows increasing proportion of watersheds with higher NN concentrations over time. The three lowest NN averages from 1906–1912 were from the Penobscot River, Maine, Columbia River, Washington, and Wateree River, South Carolina. The three lowest NN averages from 1993–1997 were from the John Day River, Oregon, Columbia River, Washington, and Penobscot River, Maine. The three highest NN averages from 1906–1912 were from the Vermillion River, Illinois, Miami River, Ohio, and the Illinois River, Illinois. The three highest NN averages from 1993–1997 were from the Vermillion River, Illinois, Minnesota River, Minnesota, and Iowa River, Iowa. Note that two rivers (the Columbia and Red rivers) continued to have the lowest concentrations of NN during both time periods and that one river (the Vermillion River) had one of the highest concentrations during both time periods. This suggests that the anthropogenic activities driving the NN concentrations at the beginning of the 20th century are similar to those of the present day.

The water-quality data for the different watersheds were separated into three geographical regions: the Mississippi River Basin (MRB), west of the MRB, and east of the MRB. Figure 1d is a 5-year running average time series of the available NN data, normalized to values reported in 1906, showing the increasing concentrations and range of values that occurred toward the end of the century. The concurrent rise in fertilizer use in the US (USDA 2008; Figure 1d) is evidence of the relationship between intensive agricultural land-use and water-quality degradation. Other studies have shown similar results, indicating that the current high NN values were not widespread until after World War II and the start of modern agriculture in the 1950s and 1960s (Goolsby and Battaglin 2001). The rise in NN concentration in the late 1940s (Figure 1d) is probably due to the start of intensive mechanical tillage and drainage.

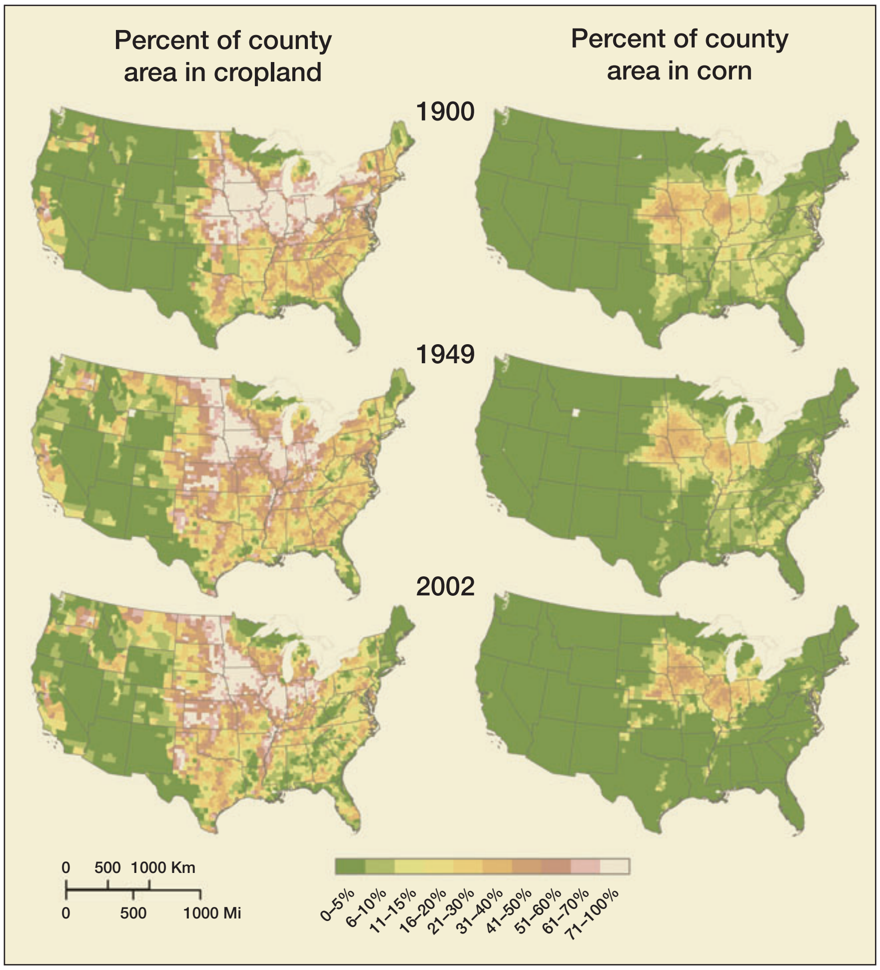

Figure 2. *The percent of land area in each county that is cropland (left) and corn harvested for grain (right). Cropland is defined as the "Number of improved acres" in the 1900 Census of Agriculture and "Total cropland" in 1949 and 2002.*

Land-use changes

The 20th century was marked by specialization and consolidation in American agriculture (Figure 2; Table 1). Although the total area covered by farmland increased by 11.5% from 1900 to 2002, the number of farms fell by almost 63%, the number of farms with more than 405 ha (1000 acres) increased by 65%, and the average farm size doubled. Farms larger than 405 ha occupied nearly 67% of the total US farmland in 2002.

Corn harvested for grain, which represented 91% of the 2002 US corn acreage, occupied nearly 23% of US croplands in 1900 and 17% in 2002 (Figure 2; Table 1). More importantly, there was a spatial redistribution of corn production; whereas corn was grown throughout the midwestern and southeastern US in the early 1900s, today, these efforts are primarily concentrated in the Midwest (Figure 2). This pattern is evident in the Piedmont region of the Southeast where, for example, between 1900 and 2002, Georgia and Alabama lost 56% and 57% of their cropland, respectively. Corn plantings also dropped by 1 298 219 ha (92.2%) and 1 033 972 ha (93.1%) in these two states in the same period. In the heart of the "Corn Belt", corn cultivation in Iowa and Illinois increased by 1 125 314 ha (28.4%) and 497 153 ha (12.0%), respectively, during the same period. The substantial increase in corn plantings in Iowa and Illinois appears to account for most of the spatial redistribution of corn agriculture in the US over the past century.

Land-use and water-quality relationships

There is a direct relationship between the fraction of a watershed's total area in cropland or cropland planted in corn with the concentration of NN exiting the watershed (Figure 3). Best-fit, simple linear regressions are shown in Table 2, where NN is the nitrate concentration exiting a watershed, in mg N L^{-1}, and $F_{cropland}$ and F_{corn} are the frac-

Table 1. Summary statistics of agriculture in the US at the beginning and end of the 20th century

Variable (unit)	*A* *1900*	*B* *2002*	*C* *Absolute change [column B – A]*	*D* *Relative change [column C/A]*
1. Number of farms (farms)	5 700 000	2 100 000	–3 600 000	–63%
2. Farms > 405 ha (farms)	29 000	47 000	+19 000	+65%
3. % all farms as large farms [Row 2/Row 1]	0.8	8.3	na	na
4. Farmland (M ha)	340	380	+39	+12%
5. Farms > 405 ha (M ha)	nd	250	na	na
6. % of farmland in large farms [Row 5/Row 4]	nd	67	na	na
7. Average farm size (ha)	59	180	+120	+200%
8. Cropland (M ha)	170	180	+7.8	+5%
9. Corn for grain (M ha)	38	27	–11	–28%

Notes: Column A: the absolute change of the variable from 1900 to 2002; column B: the fractional change. Data are rounded to two significant digits. Columns C and D reflect calculations of raw data. One hectare = 2.471 acres; M ha = million hectares; nd = no data; na = not applicable.

tions of a watershed's total area being used as cropland or corn harvested for grain, respectively. Although the study watersheds included urban and industrial areas, as well as forestry and other land uses that potentially impact N export, the evidence presented here indicates a strong association between agricultural production and riverine NN concentration.

An analysis of covariance performed for the cropland and corn harvested for grain variables yields similar results: the regressions for 1906–1912 and 1993–1997 have significantly different slopes ($P < 0.0001$), but the intercepts are not significantly different from zero or from each other.

Surprisingly, the NN concentrations in the early 20th century are highly correlated with the extent of agricultural land use at that time. Manures were the primary source of fertilizer for these farm operations, drainage practices were minimal, and the primary work was powered by animals. As a result, the net loss of NN from the fields into streams and rivers was lower than that of today (Figures 1 and 3) and was probably driven by soil mineralization and flushing of virgin soil organic matter with the introduction of agricultural tillage. The significant linear regressions between land use and NN in both 1906–1912 and 1993–1997 indicate that certain agricultural practices drive the variations in NN concentrations among watersheds and across timescales. However, the increase in the slopes of both relationships between 1906–1912 and 1993–1997 indicates a shift in the observed relationships and reveals the effects of changing management practices. The current use of mechanical tillage, tile drainage, and modern fertilizers, for example, could explain these changing relationships. Interestingly, watersheds with over 60% cropland or 25% corn have the highest NN concentrations and seem to be driving the observed rise in NN yield in the latter half of the 20th century.

There was no change in the NN concentration for watersheds with 0–30% cropland cover for four intervals from 1906–1912 to 1993–1997 (Figure 3c). This could be explained by the buffering capacity of soil ecosystems and suggests that soil systems in minimally impacted watersheds (< 30% cropland) exported similar NN concentrations at both the beginning and end of the past century. However, there was a significant change in riverine NN concentrations be-

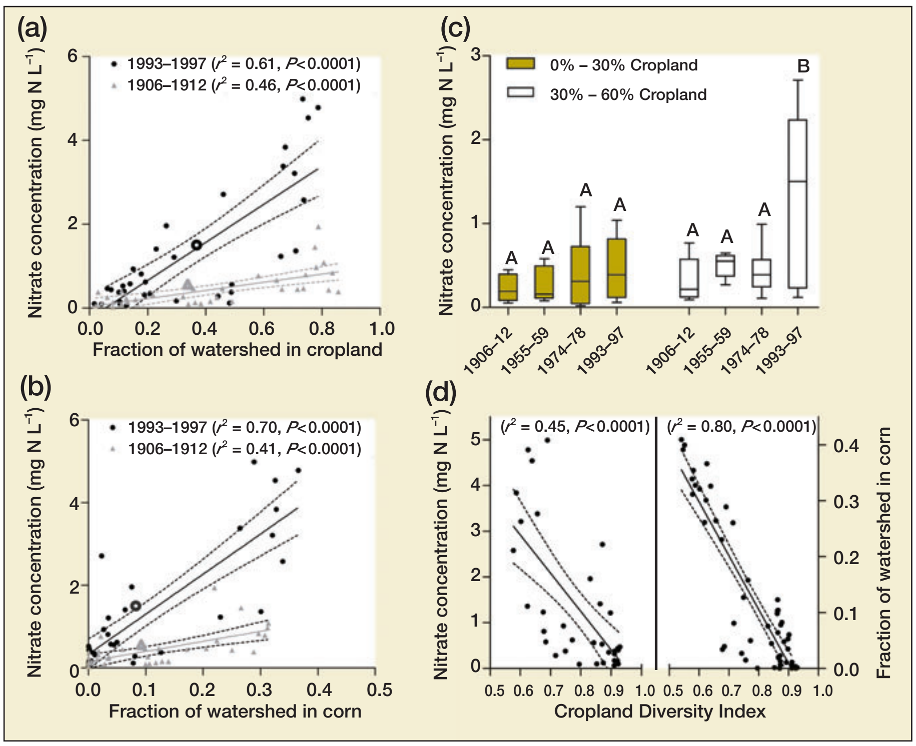

Figure 3. *A linear regression between (a) the fraction of cropland cover or (b) harvested corn and the concentration of nitrate–nitrogen (NN) by watershed at the beginning and end of the 20th century. Data from New Orleans (not included in the analysis) are indicated as a bold triangle or circle, for 1906–1912 and 1993–1997, respectively. (c) The average NN concentration for watersheds with 0–30% and 30–60% cropland cover. The features are the minimum and maximum values, 25th and 75th percentiles, and the median (alpha = 0.01, separate analyses by group). (d) A linear regression between a watershed's Cropland Diversity Index and the NN concentration in 1993–1997 (left) or the fraction of the watershed in corn (right).*

tween 1974–1976 and 1993–1997 in watersheds with 30–60% cropland cover (Figure 3d). This suggests that the higher riverine NN concentrations observed in intensively managed watersheds (>30% cropland) are a functional response to agricultural N inputs, and that the rise in nitrogen loading was most significant after the mid-1970s, when fertilizer consumption was about half that of present-day values (Figure 3d). It was in the mid-1970s that the formation of the hypoxic zone in the waters at the mouth of the Mississippi River became a regular summertime event (Turner *et al.* 2008).

Cropland diversity

We examined the GIS database for those land-use practices inversely associated with NN export. We used a simple Cropland Diversity Index (CDI) – modified from the biological standard Simpson's Diversity Index (Simpson 1949) for estimating species diversity within a known population – to measure the probability that two hectares of cropland randomly chosen from a watershed are planted with distinct crops (Figure 3d). The equation used to determine the CDI is

$$CDI = 1 - \sum_{i=1}^{9} \frac{crop_i^2}{cropland^2}$$

where *CDI* is the Cropland Diversity Index for a given watershed, *crop* is the number of hectares of a particular crop type *i* within that watershed, and *cropland* is the number of hectares of cropland within that watershed.

To generate the CDI, we included reported harvested areas of nine crops: barley, corn, cotton, hay, oats, rice, sorghum, soybeans, and wheat. A low CDI score indicates a low probability that two random hectares are planted with different crops. A CDI score of 0.5, for example, represents a 50% chance that two randomly chosen hectares of cropland belong to distinct crops (eg cropland covered exclusively with a corn–soybean rotation system). A surprising result showed that, in 1993–1997, cropland diversity was inversely related to riverine NN concentrations (r^2 = 0.45, $P < 0.0001$; Figure 3d). The diversity of croplands in these watersheds is also directly related to the fraction of a watershed harvested in corn (r^2 = 0.80, $P < 0.001$; Figure 3d). Watersheds with a larger fractional area planted with corn showed lower cropland diversity and higher NN concentrations. This result, coupled with those from Figure 3 a–c, suggests that the rise in intensive, homogeneous croplands, especially corn production systems, is responsible for a substantial proportion of the increase in NN concentration values observed in the latter half of the 20th century. As with more recent model predictions (Donner and Kucharik 2008), the results of this analysis suggest a potential increase in NN loading as the amount of corn agriculture within watersheds increases to meet the demand for biofuel production.

Table 2. Regression models and supporting statistics describing land-use and water-quality relationships at the beginning and end of the 20th century

Year	*Regression model*	r^2	P *value*
1906–1912	NN = 0.03 + (0.97)($F_{cropland}$)	0.46	< 0.0001
1993–1997	NN = –0.23 + (4.54)($F_{cropland}$)	0.61	< 0.0001
1906–1912	NN = 0.16 + (2.43)(F_{corn})	0.41	< 0.0001
1993–1997	NN = 0.33 + (9.70)(F_{corn})	0.70	< 0.0001

Conclusions

The average NN concentrations in the 63 rivers monitored over the past century were three to four times higher at the end of the century as compared with those at the beginning. The magnitude of this change was especially noticeable in the Mississippi River Basin and the Midwest region. At the beginning of the 20th century, a linear relationship existed between NN concentrations and the extent of agricultural cropland, and particularly with the area of harvested corn. Agricultural cropland planted with corn was the most important agricultural driver of riverine NN concentrations, according to the factors considered here. In addition, the slopes of the regression models have increased over the past 100 years, indicating that management practices – such as commercial fertilizer application, mechanical tillage, and intensive drainage – are responsible for the increase in NN export per hectare of cropland in the latter half of the century. We recognize that the national specialization and intensification of corn agriculture in the midwestern states have met social goals to produce high-yielding grain crops in large quantities. We also recognize that complex socioeconomic structures determine regional land-use patterns. However, our results indicate that the continued expansion of modern corn agriculture will likely increase the NN concentration in rivers and streams, particularly in those watersheds where corn cropland occupies more than 25% of the total area.

The NN concentrations in low-disturbance watersheds, on the other hand, have not shifted in the past 100 years. That is to say, the intercepts for all land-use and water-quality relationships examined in this study are not significantly different from zero at the beginning and end of the century. This implies that water quality may be rehabilitated if there are changes in land-use practices, such as a reduction in cropland area dedicated to corn agriculture, increases in cropland diversity, or changes in intensive land management. Increasing the area of perennial crop systems could meet these objectives and could lead to major improvements in water quality, ecosystem services, and the production of the farming operation (Cox *et al.* 2006; Glover *et al.* 2007; USEPA 2008). Policy initiatives affecting these land-use practices will have a direct effect on the water quality of both higher-order streams and coastal areas, by improving soil quality as well as benefiting farms. A full assessment of the direct and indirect impacts of these improvements, taking into account complex social, political, and economic systems, will require large-scale (5000 km^2) and long-term (decades) watershed-level empirical studies (Jordan *et al.* 2007).

Acknowledgments

We thank D Keeney for comments on the manuscript, M Haines for providing the historical land-use data, and H Peele for assistance with GIS applications. Support was provided by a Louisiana Board of Regents Fellowship, a Graduate Research Fellowship from The Land Institute, a Louisiana State University Graduate School Dissertation Fellowship, and a National Oceanic and Atmospheric Administration Coastal Ocean Program MULTISTRESS Award number NA16OP2670.

References

Boyer EW, Goodale CL, Jaworski NA, and Howarth RW. 2002. Anthropogenic nitrogen sources and relationships to riverine nitrogen export in the northeastern USA. *Biogeochem* **57/58**: 137–69.

Clarke FW. 1924. The composition of the river and lake waters of the United States. Washington, DC: Government Printing Office, US Geological Survey.

Cox TS, Glover JD, van Tassel DL, *et al.* 2006. Prospects for developing perennial grains. *BioScience* **56**: 649–59.

Donner SD. 2003. The impact of cropland cover on river nutrient levels in the Mississippi River basin. *Global Ecol Biogeogr* **12**: 341–55.

Donner SD and Kucharik CJ. 2008. Corn-based ethanol production compromises goal of reducing nitrogen export by the Mississippi River. *Proc Natl Acad Sci USA* **105**: 4513–18.

Earl C, Otterstrom S, and Heppen J. 1999. HUSCO 1790–1999: historical United States county boundary files [on CD-ROM]. Baton Rouge, LA: Geoscience Publications, Department of Geography and Anthropology, Louisiana State University.

Galloway JN, Schlesinger WH, Levy H, *et al.* 1995. Nitrogen-fixation – anthropogenic enhancement – environmental response. *Global Biogeochem Cycles* **9**: 235–52.

Glover JP, Cox CM, and Reganold JP. 2007. Future farming: a return to roots? *Am Sci* **297**: 81–89.

Goolsby DA, Battaglin WA, Lawrence BT, *et al.* 1999. Flux and sources of nutrients in the Mississippi–Atchafalaya River basin: topic 3 report of the integrated assessment on hypoxia in the Gulf of Mexico. Coastal Ocean Program Decision Analysis Series Number 17. Silver Spring, MD: National Oceanic and Atmospheric Administration Coastal Ocean Program.

Goolsby DA and Battaglin WA. 2001. Long-term changes in concentrations and flux of nitrogen in the Mississippi River Basin, USA. *Hydro Proc* **15**: 1209–26.

Haines MR and ICPSR (The Inter-university Consortium for Political and Social Research). 2004. Historical, demographic, economic, and social data: the United States, 1790–2000. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Howarth RW, Billen G, Swaney D, *et al.* 1996. Regional nitrogen budgets and riverine N and P fluxes for the drainages to the North Atlantic Ocean: natural and human influences. *Biogeochem* **79**: 163–86.

Jordan N, Boody G, Broussard W, *et al.* 2007. Sustainable development of the agricultural bio-economy. *Science* **316**: 1570–71.

McIsaac GF, David MB, Gertner GZ, and Goolsby DA. 2001. Eutrophication: nitrate flux in the Mississippi River. *Nature* **414**: 166–67.

MA (Millennium Ecosystem Assessment). 2005. Ecosystems and human well-being: synthesis. Washington, DC: Island Press.

Rabalais NN, Turner RE, Justiç D, *et al.* 1999. Characterization of hypoxia: topic 1 report for the integrated assessment of hypoxia in the Gulf of Mexico. Silver Spring, MD: National Oceanic and Atmospheric Administration Coastal Ocean Program.

Rabalais NN, Turner RE, Sen Gupta BK, *et al.* 2007. Hypoxia in the northern Gulf of Mexico: does the science support the plan to reduce, mitigate, and control hypoxia? *Estuaries Coasts* **30**: 753–72.

Simpson EH. 1949. Measurement of diversity. *Nature* **163**: 688.

Smil V. 1997. Global population and the nitrogen cycle. *Sci Am* **277**: 76–81.

Smith RA, Alexander RB, and Smith GE. 2003. Natural background concentrations of nutrients in streams and rivers of the conterminous United States. *Environ Sci Technol* **37**: 3039–47.

Turner RE and Rabalais NN. 2003. Linking landscape and water quality in the Mississippi River Basin for 200 years. *BioScience* **53**: 563–72.

Turner RE, Rabalais NN, Justiç D, and Dortch Q. 2006. Predicting summer hypoxia in the Northern Gulf of Mexico: riverine N, P and Si loading. *Mar Pol Bull* **52**: 139–48.

Turner RE, Rabalais NN, and Justiç D. 2008. Gulf of Mexico hypoxia: alternate states and a legacy. *Environ Sci Technol* **42**: 2323–27.

US Bureau of the Census. 1990. Census of agriculture, 1987. Washington, DC: US Bureau of the Census.

US Bureau of the Census. 2001. Census of agriculture, 1999. Washington, DC: National Agricultural Statistical Service.

USDA (US Department of Agriculture). 1979. Census of agriculture, 1949–1974, cumulative file. Washington, DC: Economic Research Service.

USDA (US Department of Agriculture). 2005. Census of agriculture, 2002. Washington, DC: National Agricultural Statistical Service.

USDA (US Department of Agriculture). 2008. Economic Research Service data sets: US fertilizer use and price. www.ers.usda.gov/data/fertilizer use/. Viewed 3 Mar 2008.

USEPA (US Environmental Protection Agency). 2008. Hypoxia in the northern Gulf of Mexico: an update by the EPA Science Advisory Board. EPA Science Advisory Board report 08-003. www.epa.gov/sab. Viewed 29 May 2008.

USEPA (US Environmental Protection Agency) and USGS (US Geological Survey). 2005. National hydrography dataset plus – NHDPlus. www.horizonsystems.com/nhdplus/index.php. Viewed 20 Sept 2007.

USGS (US Geological Survey). 2007. National Water Information System. http://waterdata.usgs.gov/nwis. Viewed Dec 2007.

WebPanel 1. Database preparation and area calculations

NN data for the early 1900s are from *The composition of the river and lake waters of the United States* (Clarke 1924). The "Clarke" dataset has 192 tables of semi-monthly water sample analyses from 190 monitoring stations on 156 rivers and lakes across the continental US, with one station in Alaska, from 1905 to 1921. We digitized the tabular data by hand, calculated an annual mean from the hand-entered data, and compared them against the published value in the original document for quality control.

The Clarke water-quality dataset was the culmination of a previous project, *The quality of surface waters in the United States, part 1: analyses of surface waters east of the one hundredth meridian* (Dole 1909), that ended prematurely with the death of the lead author, RB Dole. Dole explained that they used the phenolsulphonic acid method to determine nitrate concentrations, as prescribed in *American Health Papers and Reports* at the time. He cautioned that, "Practical considerations make it impossible to perform the test less than 10 days after some of the samples had been collected" (Dole 1909). He warned that the reported nitrate values could be problematic, due to this time delay from field sampling to laboratory processing. We tested whether the use of data from a delayed sample analysis compromised the Clarke dataset by comparing the Clarke nitrate values with those reported in *Report of chemical survey of the waters of Illinois: report for the years 1897–1902* (Palmer 1903). Palmer reported in-stream nitrate concentrations for 514 samples in Illinois rivers and streams, from 1897 to 1902. The Palmer dataset had seven stations that matched the location of the Clarke stations. Palmer specified that the samples were analysed in a timely manner: "The sample of water should be collected immediately before shipping by express, so that the shortest possible time shall intervene between the collection of the sample and its examination" (Palmer 1903). These efforts resulted in a 1–3-day delay between field collection and laboratory analysis. For these reasons, we consider the Palmer data a dependable dataset to compare with the Clarke dataset. We performed a simple, two-tailed *t*-test for differences between the semi-monthly values of nitrate concentrations at six of the seven stations. We found significant differences in nitrate concentrations at three stations and no significant difference at three stations (alpha = 0.01). We then performed a paired *t*-test for differences among the grand mean nitrate concentrations by station and found no significant difference ($P = 0.70$, n = 7) between the two datasets. We concluded that the nitrate values reported in Clarke are reasonably accurate and that use of the Clarke dataset is acceptable for analyses on large spatial and temporal scales. In addition, the Clarke dataset has been used by the US Geological Survey (USGS) in official publications, implying their confidence in the dataset (Goolsby *et al.* 2001). We do not know of any other comparable dataset for water quality on the national scale for this time period.

We accessed the USGS National Water Information System, as described in the methods section, to identify current USGS monitoring stations located on the same river reaches as those in the Clarke study. If there were no stations that met these criteria, then a station was chosen that was in the same 8-digit hydrologic unit code (HUC). A station was removed from the analysis if contemporary data could not be obtained. Of the original 157 watersheds in the Clarke dataset, there were 122 modern stations that measured NN between 1940 and 2002. Sixty-three stations had water-quality data between 1993 and 1997. If NN data were unavailable, then the NN concentration was estimated from nitrate–nitrite as nitrogen values, assuming that NN was 99.85% of nitrate–nitrite as nitrogen based on grand averages estimated from the contemporary USGS data. The water-quality data were then spatially referenced by watershed, as described in the methods section and compiled into a GIS (geographic information system) database, using ArcGIS 9.1 (ESRI 2008) for spatial analysis.

All available Census of Agriculture data (see methods section) were similarly compiled and included in the database. We spatially referenced the raw land-use data files with regionally accurate county boundary shapefiles for each respective year using *Historical United States county boundary shapefiles* (Earl *et al.* 1999). Because county boundaries change with time, it was imperative that we used accurate boundaries for a given year when associating the statistical land-use data to a specific spatial extent.

We followed techniques similar to Boyer *et al.* (2002) to sum the county-level agriculture data by watershed for a single time period using an Albers Equal Area projection and the following equation

$$F_{LU} = \frac{\left(\sum_{i=1}^{n} L_i * C_i\right)}{W}$$

where F_{LU} is the fraction of a watershed's total area in a particular land-use type, n is the number of counties within a watershed, L is the reported total area (hectares) of that land-use practice in a specific county, C is the fraction of the county area that lies within the watershed, and W is the total watershed area (hectares).

The maximum resolution of the analysis is the county level because the Census of Agriculture only reports information at that level of detail. We assumed that there was homogeneous land cover at the county level in order to split the agricultural data values between two watersheds if a watershed boundary divides a county, as it usually does. We think that the technique represents an improvement over state-level analyses, and is a robust technique to analyze land-use and water-quality relationships on larger spatial and temporal scales.

We used SAS statistical software (SAS Institute Inc 2003) for database management and Prism statistical software (GraphPad Software Inc 2007) to develop regression models.

Continued

WebPanel 1. ***(Continued)***

Web only references cited

Boyer EW, Goodale CL, Jaworski NA, and Howarth RW. 2002. Anthropogenic nitrogen sources and relationships to riverine nitrogen export in the northeastern USA. *Biogeochem* **57/58**: 137–16.

Clarke FW. 1924. The composition of the river and lake waters of the United States. Washington, DC: Government Printing Office, US Geological Survey.

Dole RB. 1909. Quality of surface waters in the United States: part I – analyses of waters east of the one hundredth meridian. Washington, DC: Government Printing Office, US Geological Survey.

Earl C, Otterstrom S, and Heppen J. 1999. HUSCO 1790–1999: historical United States county boundary files [on CD-ROM]. Baton Rouge, LA: Geoscience Publications, Department of Geography and Anthropology, Louisiana State University.

ESRI (Environmental Systems Research Institute Inc.). 2005. ArcGIS 9.1. Redlands, CA: Environmental Systems Research Institute Inc.

Goolsby DA, Battaglin WA, Allenbach BT, and Hooper RP. 2001. Nitrogen input to the Gulf of Mexico. *J Environ Qual* **30**: 329–36.

GraphPad Software Inc. 2007. *Prism5 for Windows*. San Diego, CA: GraphPad Software Inc.

Palmer AW. 1903. Chemical survey of the waters of Illinois: report for the years 1897–1902. Urbana, IL: University of Illinois.

SAS Institute Inc. 2003. SAS/STAT version 9. Cary, NC: SAS Institute Inc.

