

NRC Publications Archive Archives des publications du CNRC

Geospatial data for developing nutrient SPARROW models for the Midcontinental region of Canada and the United States

Vouk, Ivana; Burcher, Richard S.; Johnston, Craig M.; Jenkinson, R. Wayne;
Saad, David A.; Gaiot, John S.; Benoy, Glenn A.; Robertson, Dale M.; Laitta,
Michael

For the publisher's version, please access the DOI link below./ Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

Publisher's version / Version de l'éditeur:

<https://doi.org/10.4224/23004810>

Technical Report (National Research Council of Canada. Ocean, Coastal and River Engineering); no. OCRE-TR-2018-014, 2018-12-21

NRC Publications Archive Record / Notice des Archives des publications du CNRC :

<https://nrc-publications.canada.ca/eng/view/object/?id=0e5ed829-58a1-4742-a198-8322dbdc327c>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=0e5ed829-58a1-4742-a198-8322dbdc327c>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

NRC-CMRC

Geospatial Data for Developing Nutrient SPARROW Models for the Midcontinental Region of Canada and the United States

Report No.: OCRE-TR-2018-014

Date: December 21, 2018

Authors: Ivana Vouk¹, Richard S. Burcher¹, Craig M. Johnston², R. Wayne Jenkinson³, David A. Saad², John S. Gaiot⁴, Glenn A. Benoy³, Dale M. Robertson², and Michael Laitta³

¹ National Research Council of Canada – Ocean, Coastal, River Engineering Research Centre

² U. S. Geological Survey

³ International Joint Commission

⁴ Ontario Ministry of Natural Resources and Forestry



© 2018 Her Majesty the Queen in Right of Canada, as represented by the National Research Council Canada.

Cat. No. NR14-3/2018-14E-PDF

ISBN 978-0-660-29094-2

Series: Technical Report (National Research Council Canada. Ocean, Coastal and River Engineering).

Report No.: OCRE-TR-2018-014

Second printing. Includes revisions of June 2020.

EXECUTIVE SUMMARY

Through the International Watersheds Initiative of the International Joint Commission (IJC), the Spatially-Referenced Regressions on Watershed attributes (SPARROW) model developed by the U.S. Geological Survey (USGS) is being applied to the Great Lakes, Rainy River – Lake of the Woods and Red-Assiniboine basins. The objective of this binational application of the SPARROW model is to better understand and quantify the sources of phosphorus (P) and nitrogen (N) that contribute to regional water-quality issues like algal blooms and eutrophication in Lake Erie and other parts of the Great Lakes, as well as Lake of the Woods. Led by the IJC, a team of researchers from the National Research Council of Canada – Ocean, Coastal and River Engineering Research Centre, USGS, and IJC are extending the SPARROW modelling work previously completed for the Red-Assiniboine basin and the U.S. portions of the Great Lakes, Ohio, Upper Mississippi, and Souris-Red-Rainy river basins to cover all of the Great Lakes, Rainy River – Lake of the Woods and Red-Assiniboine basins. The current effort is termed the Midcontinent SPARROW modelling study.

This report describes the data used to develop the Midcontinent SPARROW models, specifically the sources of original data, assembling the data, and the processing and harmonization required between the U.S. and Canada data needed to produce these models. Details provided include the:

- development of a digital stream network and related catchments – most significantly in the Canadian regions of the Great Lakes and Rainy River – Lake of the Woods basins where these data were not available to create a seamless binational network across the model domain;
- calculation of variables to aid in the determination of in-stream and in-reservoir decay of P and N;
- quantification of Canadian diversions within the Midcontinental region (i.e., Lake St. Joseph, Long Lake and the Ogoki Reservoir);
- development of binational input nutrient sources considered for model development (i.e., land cover, inorganic farm fertilizer, manure, atmospheric deposition, point-source pollution from wastewater-treatment plants and contribution from non-modelled watersheds); and
- development of delivery variables considered to be most predominant (i.e., temperature, precipitation and ensuing runoff, soil permeability and clay content, slope of the catchments, and tile drainage).

The majority of the geospatial data collection and processing was required for Canadian datasets because many of the U.S. datasets were already assembled for previous SPARROW model applications in the U.S. The task of harmonizing data between the U.S. and Canada was important to ensure consistency of the datasets used in the models. The harmonized digital stream network, delineated catchments and input data for each catchment (i.e., source and delivery variables), created for the Midcontinent SPARROW models, are available for download at url: <https://doi.org/10.4224/300.0001>.

Contents

EXECUTIVE SUMMARY	i
Contents	iii
Figures	v
Glossary	vi
Acronyms	vii
Acknowledgements	x
1 Introduction	1
2 Background	2
2.1 The SPARROW model	3
2.2 Midcontinent SPARROW Model domain	3
2.3 Integration and harmonization of geospatial datasets	5
3 Stream network	6
3.1 Stream network data sources	6
3.2 Stream network development and refinement	6
3.3 Preserving location integrity of calibration points	10
3.4 Lakes data and shoreline segments	11
4 Catchment delineation	12
5 Input variables to calculate decay of modelled constituents	14
5.1 Flow and drainage area estimates	14
5.2 Velocity estimates	16
5.3 Hydraulic load estimates	17
6 Diversions	18
7 Nutrient source and delivery variables	19
7.1 Input sources	19
7.1.1 Land cover	20
7.1.2 Fertilizer	21
7.1.3 Manure	23
7.1.4 Atmospheric deposition	25
7.1.5 Wastewater treatment plant inputs	26
7.1.6 Non-modeled contributing areas	27
7.2 Delivery variables	28
7.2.1 Climatic data	28
7.2.2 Soil characteristics	31
7.2.3 Mean catchment slope	33
7.2.4 Tile drainage	34
8 Summary	36

9	References.....	37
	Appendix A – Coordinate System	A-1
	Appendix B – Land cover	B-1
	Appendix C – Background to Fertilizer and Manure development.....	C-1
	Fertilizer	C-1
	Manure	C-2

Figures

Figure 1: States and provinces comprising the Midcontinental region.....	2
Figure 2: The Midcontinent SPARROW models domain. The inset identifies the locations of Brokenhead, Whitemouth and Whiteshell Rivers basins.....	4
Figure 3: Final Midcontinent SPARROW model digital stream network – colour coded to indicate the original data source(s).....	9
Figure 4: Calibration points (monitoring stations) used to preserve stream segments during thresholding and generalization.....	10
Figure 5: Midcontinent SPARROW model catchments. Enlarged view of the catchments is provided in the inset as an example, indicated by the yellow rectangle the red arrow is pointing to on the larger map.....	13
Figure 6: Validation results for estimated versus observed flows (left) and drainage areas (right) in the Great Lakes basin. The 1:1 reference line is indicated with a red line.	15
Figure 7: Validation results for estimated versus observed flows (left) and drainage areas (right) in the Winnipeg River basin. The 1:1 reference line is indicated with a red line.....	15
Figure 8: Estimated velocity in streams throughout the Midcontinental region.	16
Figure 9: Estimated inverse of hydraulic load throughout the Midcontinental region.	17
Figure 10: Stream diversions in Canada (with respect to the extents of the Midcontinental region). The inset provides a close-up view of Lake St. Joseph, Ogoki Reservoir and Long Lake, with red arrows indicating the direction of flow diversion. Lake Nipigon is identified in the inset.	18
Figure 11: Harmonized land cover throughout the Midcontinental region.....	20
Figure 12: Total phosphorus (above) and total nitrogen (below) input rates, by catchment, throughout the Midcontinental region.....	22
Figure 13: Total phosphorus (above) and total nitrogen (below) input rates, by catchment, from manure spread on agricultural and pasture lands throughout the Midcontinental region.	24
Figure 14: Atmospheric deposition rates of total nitrogen from CMAQ, by catchment, throughout the Midcontinental region.....	25
Figure 15: Wastewater-treatment plants used in the Midcontinent SPARROW models.....	27
Figure 16: Mean annual air temperature, by catchment, throughout the Midcontinental region.	28
Figure 17: Mean annual precipitation by catchment, throughout the Midcontinental region.	29
Figure 18: Mean annual runoff by catchment, throughout the Midcontinental region.	30
Figure 19: Soil permeability (above) and clay content (below) by catchment, throughout the Midcontinental region.	32
Figure 20: Mean catchment slopes by catchment, throughout the Midcontinental region.	33
Figure 21: Percentage of the catchment underlain with tile drains, by catchment, throughout the Midcontinental region.	35

Glossary

high-resolution NHD	1:24,000-scale national hydrographic dataset
HYDAT	Environment and Climate Change Canada HYdrological DATabase that contains computed data from Water Survey of Canada's hydrometric gauges
medium-resolution NHDPlus	1:100,000-scale national hydrographic dataset
SPARROW	SPAtially Referenced Regression On Watershed attributes model is a spatially-referenced regression-based model developed by the U.S. Geological Survey that relates patterns in loads of water-quality parameters to human activities, climate, hydrology, geology and physiography

Acronyms

3DEP	3D Elevation Program (U.S.)
AAFC	Agriculture and Agri-Food Canada
CaPA	Canadian Precipitation Analysis
CCS	Census consolidated subdivisions (Canada)
CD	Census divisions (Canada)
CDED	Canadian Digital Elevation Dataset
CFS	Canadian Forest Service
CMAQ	Community Multiscale Air Quality Modeling System (U.S.)
DEM	Digital elevation model
ECCC	Environment and Climate Change Canada
EWC	Enhanced Watercourse (Ontario)
GIS	Geographic information system
GL	Great Lakes
HUC12	12-digit Hydrologic Unit Code (U.S.)
IWI	International Watersheds Initiative (IJC)
MCWS	Manitoba Conservation and Water Stewardship
MOECC	Ministry of Ontario Environment and Climate Change
MSD	Manitoba Sustainable Development
NAWQA	National Water-Quality Assessment (U.S.)
NED	National Elevation Dataset (U.S.)
NHD	National Hydrographic Dataset (U.S.)
NHDPlus V1	National Hydrography Dataset Plus Version 1.0 (U.S.)
NHDPlus V2	National Hydrography Dataset Plus Version 2.0 (U.S.)
NHN	National Hydro Network (Canada)
NLCD	National Land Cover Dataset (U.S.)
NRCan	Natural Resources Canada
NRC-OCRE	National Research Council of Canada – Ocean, Coastal and River Engineering

NTDB	National Topographic Database (Canada)
OIHD	Ontario Integrated Hydrology Data
OHN	Ontario Hydro Network
OMAFRA	Ontario Ministry of Agriculture, Food and Rural Affairs
OMNRF	Ontario Ministry of Natural Resources and Forestry
PRISM	Parameter-elevation Regressions on Independent Slopes Model
RA	Red-Assiniboine River
SAS	Statistical Analysis Software
SPARROW	SPATIally Referenced Regressions On Watershed attributes
SSDA	Sub-sub-drainage area
STATSGO	STATe Soil GeOgraphic database (U.S.)
TN	Total Nitrogen
TP	Total Phosphorus
USEPA	U.S. Environmental Protection Agency
USGS	U.S. Geologic Survey
WBD	Watershed Boundary Dataset
WR	Winnipeg River
WSC	Water Survey of Canada
WWTP	WasteWater-Treatment Plant

This report is dedicated to the memory of Craig Johnston (1969–2018). A GIS specialist with the U.S. Geological Survey, Craig was an invaluable member of the binational SPARROW modelling team who helped introduce the International Joint Commission to the SPARROW approach. He was our leading light on the development of binational stream networks – the true backbones of the models. Craig had a wonderful spirit and is dearly missed.

Acknowledgements

The authors are grateful to the National Research Council's staff: Vahid Pilechi, Sean Ferguson and student Isabelle Cheff for their efforts in processing much of the Canadian source and delivery variables, and to Anne Collins for processing the initial Great Lakes stream network. The authors would like to thank Glenn Lelyk of Agriculture and Agri-Food Canada (who had devoted a number of hours to provide a product relevant to the effort) for processing and providing Canadian soils information. The authors would also like to thank Jason Vanrobaeys of Agriculture and Agri-Food Canada and Mark Henry of Statistics Canada for guidance on Canadian agricultural outputs. Finally, the authors would like to acknowledge other researchers, scientists and water resource managers who answered our phone calls and replied to our emails: Elaine Page, Geoff Reimer and Tara Wiess (Manitoba Sustainable Development), Andy Kester (Ontario Ministry of Agriculture, Food and Rural Affairs), Doug Johnson (Saskatchewan Water Security Agency), Antonette Arvai (International Joint Commission), Bruce Davison (Environment and Climate Change Canada), Pam Minifie (Saskatchewan Ministry of Environment) and Amanda Giamberardino (Fertilizer Canada). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the Canadian and U.S. Governments.

1 Introduction

Many watersheds straddle the border between Canada and the United States (U.S.). Through the International Watersheds Initiative (IWI) of the International Joint Commission (IJC)¹, the Spatially Referenced Regressions on Watershed attributes (SPARROW) model, developed by the U.S. Geological Survey (USGS), is being applied to the Midcontinental region of North America (Figure 1). This region consists of the Great Lakes (GL) basin, Winnipeg River (WR) basin, Red-Assiniboine River (RA) basin, Upper Mississippi River basin and Ohio River basin, including the basins between the WR and RA (i.e., Brokenhead River, Whiteshell River and Whitemouth River basins that for the purpose of this study are considered part of the WR basin) (Figure 2). This large area spans parts of Saskatchewan, Manitoba, and Ontario in Canada, and the upper midwest and Great Lakes states of the U.S., as illustrated in Figure 1. The motivation for the study is to better understand and quantify the sources of phosphorus and nitrogen that contribute to regional water-quality issues like algal blooms and eutrophication in Lake Erie and other parts of the Great Lakes, as well as Lake of the Woods basins.

Led by the IJC, an international team was assembled for the project, including researchers from the National Research Council of Canada – Ocean, Coastal and River Engineering (NRC-OCRE) Research Centre, USGS and the Ontario Ministry of Natural Resources and Forestry (OMNRF). Successful development of the geospatial dataset required significant contribution from the following additional Canadian federal and provincial collaborators: Environment and Climate Change Canada (ECCC), Agriculture & Agri-Food Canada (AAFC), Manitoba Sustainable Development (MSD) formerly known as the Manitoba Conservation and Water Stewardship (MCWS), and Ministry of Ontario Environment and Climate Change (MOECC). This model builds on the experience gained from the USGS application of the SPARROW model for the National Water-Quality Assessment (NAWQA) Project studying the U.S. portions of the Great Lakes, Ohio, Upper Mississippi, and Souris-Red-Rainy river basins (Robertson and Saad, 2011) and the SPARROW modelling work previously completed on the RA basin (i.e., RA SPARROW models) (Benoy *et al.*, 2016). It also takes advantage of the IWI Data Harmonization Project, which pioneered the development of interoperable hydrographic and geospatial datasets for transboundary basins along the international border (Laitta, 2010).

¹ One of the functions of the IJC is to prevent and resolve disputes regarding waters shared by Canada and the U.S. The IWI is a mechanism by which the IJC supports projects that apply approaches and tools that may be used to understand and address a variety of water-related issues.

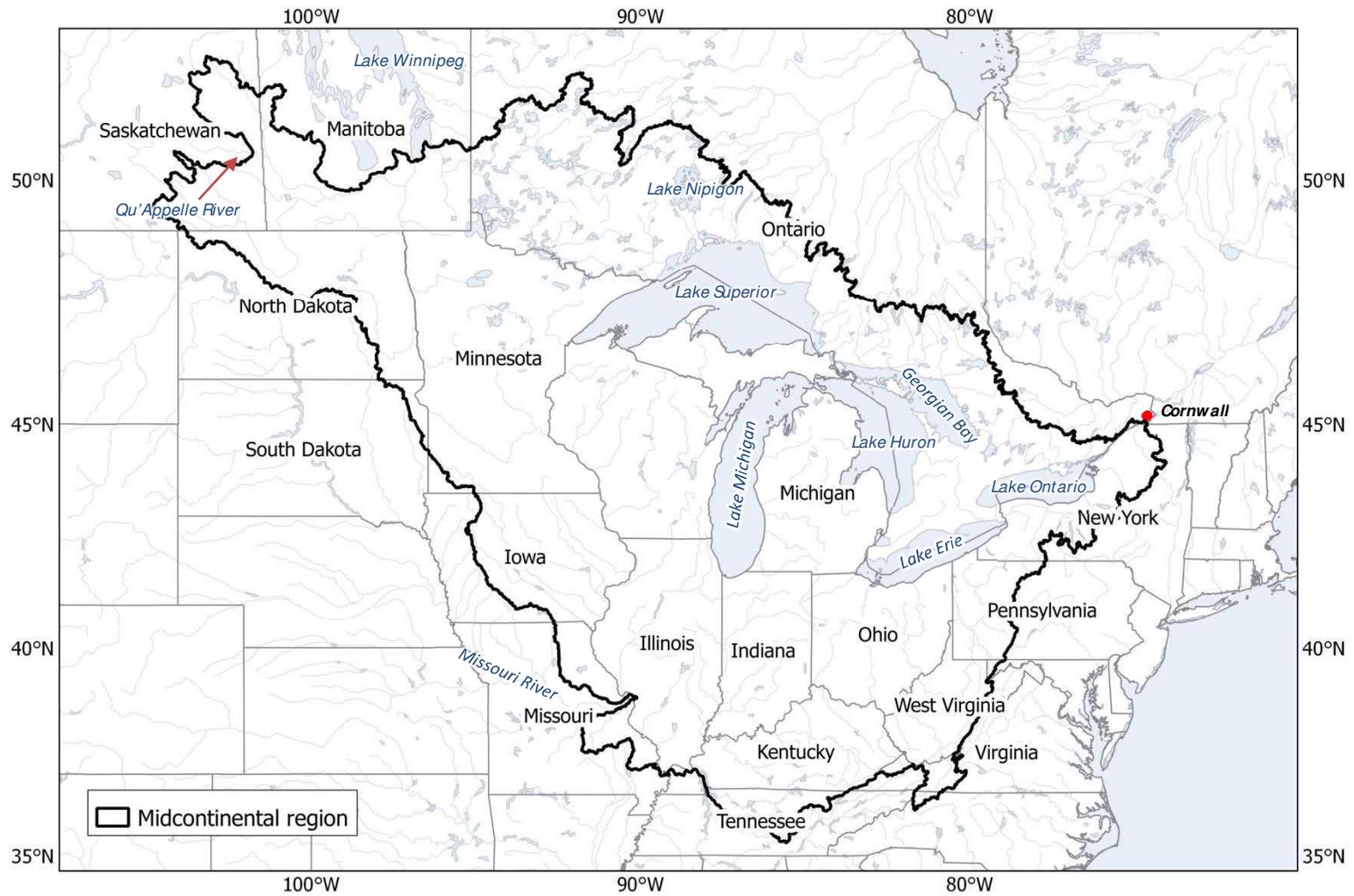


Figure 1: States and provinces comprising the Midcontinental region.

2 Background

2.1 The SPARROW model

SPARROW is an empirically based spatially-referenced regression modelling approach developed by the USGS that relates patterns in loads of water-quality parameters to human activities, climate, hydrology, geology and physiography (Schwarz *et al.*, 2006). The SPARROW model has been applied in many regions including the entire continental U.S., as well as New Zealand. The current application of the SPARROW model to the area termed the *Midcontinent* represents the second binational implementation of this model, after its successful first implementation for the RA basin.

SPARROW models simulate long-term mean-annual constituent loads in streams and rivers over large geographic areas (Preston *et al.*, 2009). The models can be used to help understand the origin and transport of nutrients or other types of contaminants from their major sources. Four types of data are used to “build” SPARROW models: stream network information to define stream reaches and catchments; loading information for many sites within the model boundaries (dependent variables); information describing all of the input sources of the modelled constituents (independent variables); and information describing the natural environment of the modelled area that results in statistically significant variability in the delivery and decay of the modelled constituent (independent variables).

The Midcontinent SPARROW models were developed to simulate nutrient loads in streams throughout the area. The dependent variable in these models is long-term mean-annual loads (total phosphorus or total nitrogen) normalized to a specific base year (Preston *et al.*, 2009). The base year for the Midcontinent SPARROW models is 2002, which was selected so that estimated loads would coincide with available geospatial datasets describing input data sources. In this report, we describe the stream network used to define the reaches and catchments, estimate the input sources of nutrients (phosphorus [P] and nitrogen [N]), and assess potential delivery and decay mechanisms for each of the Midcontinent SPARROW catchments. Saad *et al.* (2018) summarized all of the loads used to build these Midcontinent SPARROW models. Though customized for this SPARROW application, the binational datasets described here may prove useful for any of a number of other projects in the Midcontinental region.

2.2 Midcontinent SPARROW model domain

The target domain of the Midcontinent SPARROW models for prediction and analysis of nutrients encompasses the Rainy River – Lake of the Woods and the GL basins as far downstream as approximately Cornwall, Ontario. These basins are of primary interest to the IJC and identified in Figure 2. In order to maximize utility of the available water-quality monitoring station data and improve model calibration, the Midcontinent SPARROW models are being developed including the domains of the previously developed U.S. portions of the Great Lakes, Ohio, Upper Mississippi, and Souris-Red-Rainy river basins (Robertson and Saad, 2011) and the RA basin (Jenkinson and Benoy, 2015), and the entire WR basin, including the Brokenhead, Whiteshell and Whitemouth basins which bridge the gap between RA and WR basins. The map in Figure 2 also illustrates the extent of the Midcontinent SPARROW models, the Canada-U.S border and the basins comprising the models.

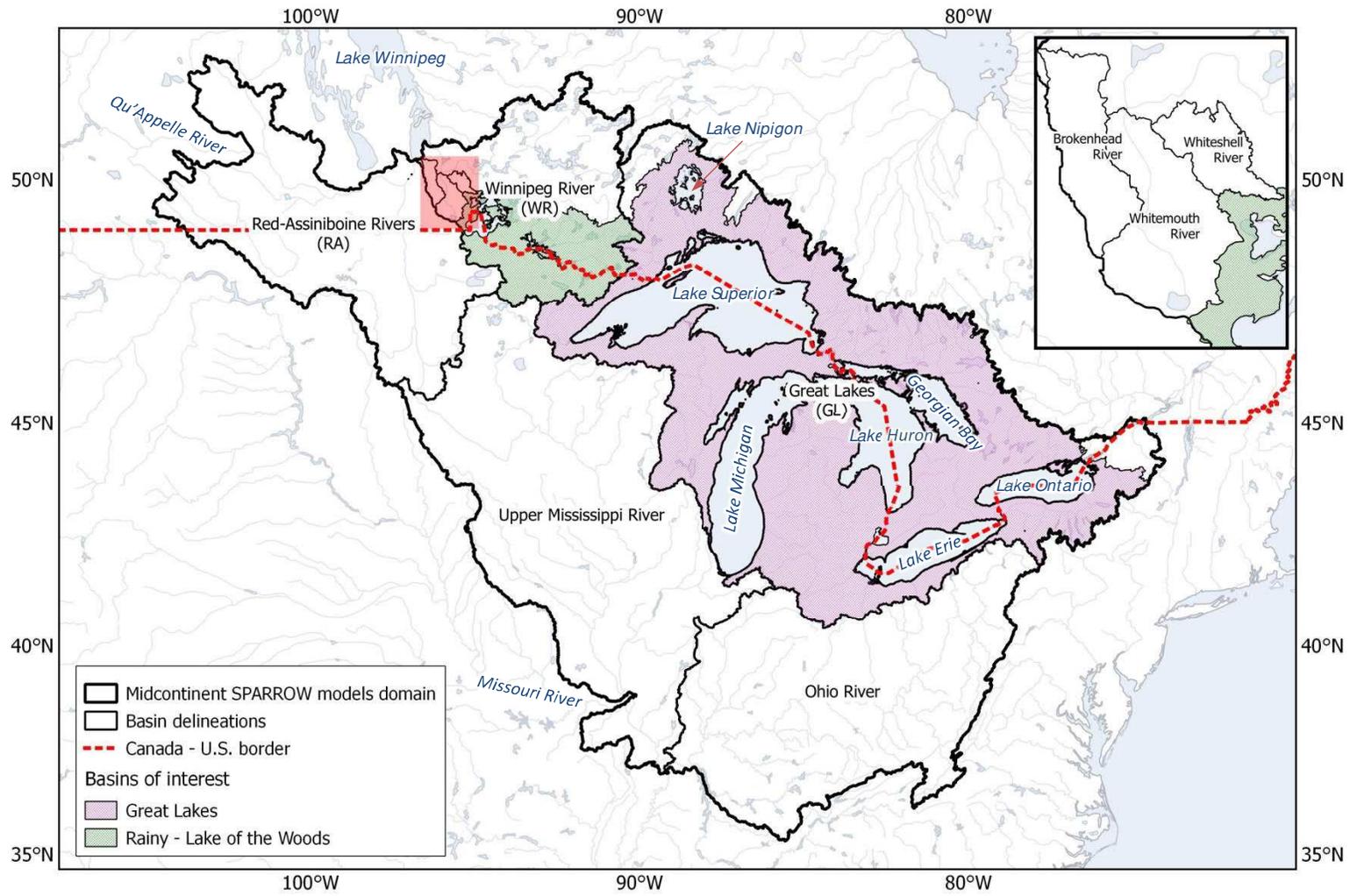


Figure 2: The Midcontinent SPARROW models domain. The inset identifies the locations of Brokenhead, Whitemouth and Whiteshell Rivers basins.

2.3 Integration and harmonization of geospatial datasets

As SPARROW is a spatially-referenced model, it requires extensive and contiguous geospatial datasets for its construction, calibration and operation. As these Midcontinent SPARROW models are meant to be binational, they require harmonization (i.e. the process of bringing together data of varying formats, delineations, naming conventions and transforming it into one cohesive dataset) of various datasets both between the U.S. and, Canada, as well as amongst provinces and states. The harmonization tasks extend the work of the IJC data harmonization project whereby certain physiographic and hydrographic datasets (e.g., water courses and drainage areas) in border catchments were adjusted and modified in order to gain geospatial consistency between Canadian and the U.S. jurisdictions (Laitta, 2010). The harmonization approaches adopted followed methodologies developed as part of the RA SPARROW modelling effort for harmonizing geophysical (i.e., land cover) and agricultural datasets. The following sections (i.e., 3 through 7) describe the origin of data and methods used for developing input datasets for the Midcontinent SPARROW models. Although automation was used in some cases, the preparation of these datasets was primarily achieved through a series of manual adjustments to ensure integrity of the input data.

The harmonized digital stream network, delineated catchments and input data for each catchment (i.e., source and delivery variables) created for the Midcontinent SPARROW models are available for download at url: <https://doi.org/10.4224/300.0001>. The README.txt file included with the datasets provides the list of all data required to run the Midcontinent SPARROW models for 2002 base year. The datasets have been finalized (i.e., verified and adjusted where necessary) for the use in the Midcontinent SPARROW models. The details of the coordinate system used in the Midcontinent SPARROW models are provided in Appendix A.

3 Stream network

A SPARROW stream network is a digital geospatial dataset that describes the geographic path and connectivity of water bodies in a particular domain or watershed. The model requires a detailed and contiguous stream network that includes network segment connectivity so that flow paths and nutrient transport can be accurately assessed. The network is a combination of streams and catchments. The stream data also maintain connectivity and delineate the transport through the network. Characteristics of the catchments are used to describe the physical properties of the area each stream segment drains (Brakebill *et al.*, 2011), which are further discussed in Section 4. The Midcontinent SPARROW stream network² is comprised of digital stream networks spanning five different watersheds: the RA basin, GL basin, WR basin (all of which straddle the U.S.–Canada border) as well as the Upper Mississippi River basin and the Ohio River basin that are fully situated within the U.S. The Qu'Appelle River in Saskatchewan and Missouri River in central U.S. were also considered but due to their associated limitations were excluded from the stream network and simply included as contributing point sources (further discussed in Section 6).

3.1 Stream network data sources

The digital stream network for Midcontinent SPARROW models was developed from multiple sources. The spatial extent of data from each source is shown in Figure 3 and described in detail below. The digital stream network within the U.S. portion of the GL basin – for the Upper Mississippi River and the Ohio River basins – is based on the 1:100,000-scale National Hydrography Dataset Plus Version 2.0 (NHDPlus V2) by Moore and Dewald (2016). The NHDPlus is a geospatial, hydrologic framework dataset constructed from the NHD stream network (U.S. Geological Survey, 1999), which includes catchments and a suite of network attributes for network navigation. The first version of NHDPlus (NHDPlus V1) was released in 2006 (U.S. Geological Survey and U.S. Environmental Protection Agency, 2006). NHDPlus V2, released in 2011, is an improvement over V1 and includes some network enhancements and corrections to the previous stream network.

The digital stream network spanning the Canadian portion of the GL basin was constructed using the Ontario Integrated Hydrology Data (OIHD) Enhanced Watercourse (EWC) dataset, which is a subset of the Ontario Hydro Network (OHN) developed by the Ontario Ministry of Natural Resources and Forestry (2012). The Canadian National Hydro Network (NHN) developed by Natural Resources Canada (NRCan) (Canadian Council on Geomatics, 2009) was also considered, but its Ontario portion was largely derived from an older version of OHN. The RA SPARROW model digital stream network was developed by Jenkinson and Benoy (2015) using a variety of data sources available across the region. The medium-resolution (1:100,000-scale) NHDPlus V1 stream network was used for the portion of the Red River basin immediately downstream of the Park River confluence. The stream network for the entire Souris River basin and the remainder of the Red River basin were incorporated with NHDPlus V1 using the high-resolution (1:24,000-scale) National Hydrography Dataset (NHD) harmonized with the Canadian (1:50,000-scale) NHN (Laita, 2010). The WR basin digital stream network was constructed using the high-resolution NHD harmonized with data from the NHN. Whitemouth, Whiteshell and Brokenhead Rivers basins in southeast Manitoba³ were incorporated into the overall stream network to fill the gap between the RA and WR basins.

3.2 Stream network development and refinement

Care was taken such that the newly created digital stream network aligned with the hydrography represented in the medium-resolution NHDPlus V2 in terms of regional stream density and structure. Existing digital stream data products were merged into a single, seamless stream network for all binational basins to develop transboundary and harmonized stream networks, completing the areal extent. The resulting Canadian stream datasets in the GL and WR basins along with the high-resolution NHD in the U.S. portion of the WR basin were much denser than the medium-resolution NHDPlus V2 network. For example, within the Canadian GL basin alone, the EWC contained over 1.7 million reaches. A target upper limit for the entire Midcontinental region was set to approximately one million reaches to expedite the execution and calibration of the SPARROW models

² The Midcontinent SPARROW digital stream network was developed with the source data available to the developers at the time. Data providers have continued to improve their datasets, which will not necessarily be reflected in the digital stream network for this set of Midcontinent SPARROW models.

³ For simplicity the term *WR basin* will include the Whitemouth, Whiteshell and Brokenhead rivers basins hereafter in the report.

and to support the feasibility of online tools to present model results. Therefore, refinement of the Canadian and high-resolution NHD data sources was required to make the digital stream networks on both sides of the border consistent in density.

Initial digital stream network refinement involved the removal of isolated stream segments and constructed streams (i.e., man-made ditches and drains). In areas where constructed streams were identified, the thinning was straightforward. However, this proved to be an onerous effort in portions of the WR basin digital stream network (i.e., in Manitoba), which lacked consistent and reliable information (i.e., missing or inconsistently identified attributes). Thus, for the WR basin digital stream network a number of constructed streams were retained unlike the digital stream network in other basins of the Midcontinent SPARROW models.

Next steps in refinement of the digital stream network required further thinning and generalization. Further thinning involved the removal of headwater streams below a specified areal drainage area⁴ threshold. Generalization involved the simplification of the stream network suitable for the purpose of the SPARROW modelling effort. Using the upstream drainage area of each unique stream segment, a thresholding metric was employed. In this way all stream segments below a defined drainage area threshold were easily removed, providing a consistent and physically-based method for controlling the network density across the domain. The EWC did not have drainage area as an attribute. However, a flow-accumulation raster (i.e., a grid identifying the upstream flow accumulation for each cell in the grid) was available allowing the mapping of the drainage areas back to the stream network. Although generally successful, inconsistencies between the scale of the mapped stream segments and the raster layer (30-m horizontal resolution) posed challenges such as: assigning large values to some very small drainage areas resulting in erroneous retention of these data; and producing disconnected stream networks when stream segments did not honour the flow-accumulation grid values, which erroneously excluded entire sections upstream of the disconnection from the network.

To account for these issues, special quality-control scripts were developed in Python (Python Software Foundation, 2018) to recursively track through the digital stream network and look for physical inconsistencies in the assigned drainage area values and correct or adjust them. For instance, as the network progresses upstream, the drainage area assigned to each stream segment should monotonically decrease. If that was not observed, drainage area values were corrected if possible or flagged for manual intervention.

With a consistent drainage area attribute assigned to the EWC, generalization was then possible. The drainage area statistics of the existing medium-resolution NHDPlus V2 network were examined to determine an appropriate drainage area threshold. A reasonable lower drainage area threshold for the medium-resolution NHDPlusV2 network was between 1.0 km² and 2.5 km². Various thresholds were experimented with for the EWC and were adjusted to optimize the number of unique stream segments and to ensure that visually spatial consistencies existed across the border with respect to network density. The 1.0-km² drainage area threshold was adequate for digital stream network generalization and was, therefore, applied to the EWC.

Methods derived during the construction of the digital stream networks for the RA basin and the Canadian part of the GL basin were used to develop the WR basin digital stream network. Unlike for the GL basin stream network, a single flow-accumulation raster for the full WR basin was not available to aid in the refinement of its digital stream network. Several digital elevation model (DEM) sources were reviewed, which if used would have created a patchwork dataset requiring harmonization between inter-provincial and international borders. Instead, the DEM used in the WR basin originated from the 30-meter USGS product National Elevation Dataset (NED) (Gesch *et al.*, 2002) that includes Canadian elevation data from the National Topographic Database (NTDB) (Natural Resources Canada, 2002). NED has since been incorporated into the 3D Elevation Program (3DEP) (Sugarbaker *et al.*, 2016). Using the NED dataset ensured that cell sizes were consistent, and harmonized products were applied throughout the basin. A series of computer scripts were created in Python (Python Software Foundation, 2018) to generate a flow direction grid, which in turn was used to derive a flow accumulation raster. The 1.0-km² drainage area threshold was applied to the WR basin stream network similar to that used to develop the Canadian GL basin stream network.

⁴ Drainage area refers to a geophysical area that drains runoff to a common location. The drainage area was calculated using runoff maps further discussed in Section 7.2.1.

The final digital stream networks were verified for potential connectivity, topology and routing issues using Esri's geographic information system (GIS) software (Esri, 2010). For example, disconnected streams and network loops (i.e., where multiple flow directions existed) were identified and resolved. Once finalized, the basins were provided a *hydrologic sequencing code* – a numeric code applied to individual stream segments allowing SPARROW to recognize the connectivity or stream segment order from headwater to terminal reaches. SPARROW uses the hydrologic sequencing code to sort data records and inherently allow the accumulation of constituent mass for model calibration and prediction.

Refining and processing of the original stream networks led to networks of size appropriate for the SPARROW model. The GL digital stream network decreased to approximately 11% (reduction of 89%) and the WR to approximately 28% (reduction of 72%) of the original streams.

The stream network for the Midcontinental region is shown in Figure 3, with each of the original sources identified in different colours.

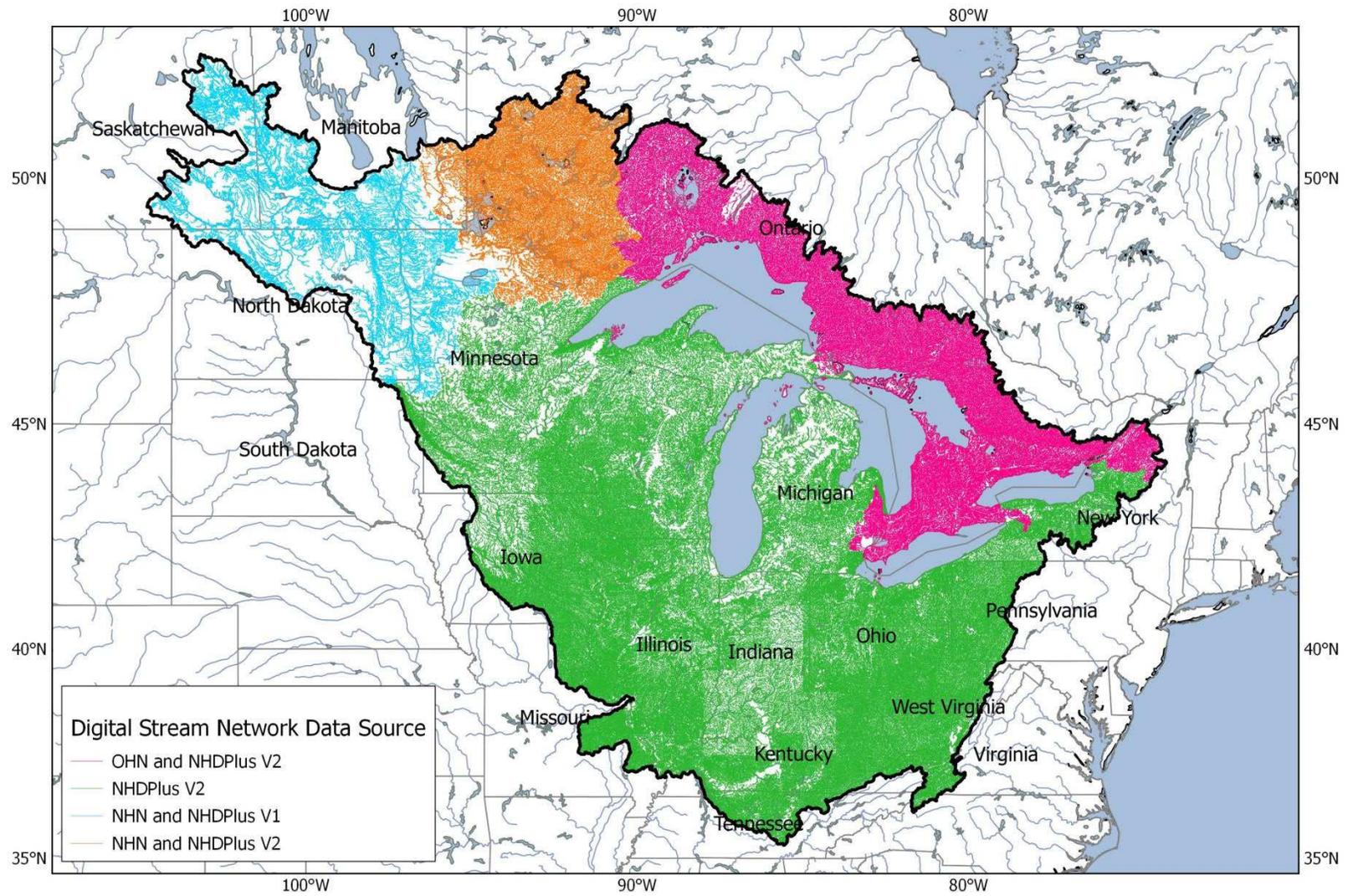


Figure 3: Final Midcontinent SPARROW model digital stream network – colour coded to indicate the original data source(s).

3.3 Preserving location integrity of calibration points

In order to calibrate SPARROW models, monitoring stations with estimated loads (calibration points) are needed (Saad *et al.*, 2018). All monitoring stations were assigned to the digital network prior to thinning for the network simplification (of Section 3.2). All stations not directly associated with a stream segment were paired to the nearest stream segment within a 50-m radius. The paired stream segment – monitoring station were inspected for accuracy and corrected where necessary. Stream segments containing monitoring stations acting as calibration points may have been unintentionally removed during the network refinements because of the defined thresholds. To overcome this issue, stream segments with monitoring stations were identified and preserved during refinement. The calibration points used in the model are shown in Figure 4 for both the U.S. and Canada.

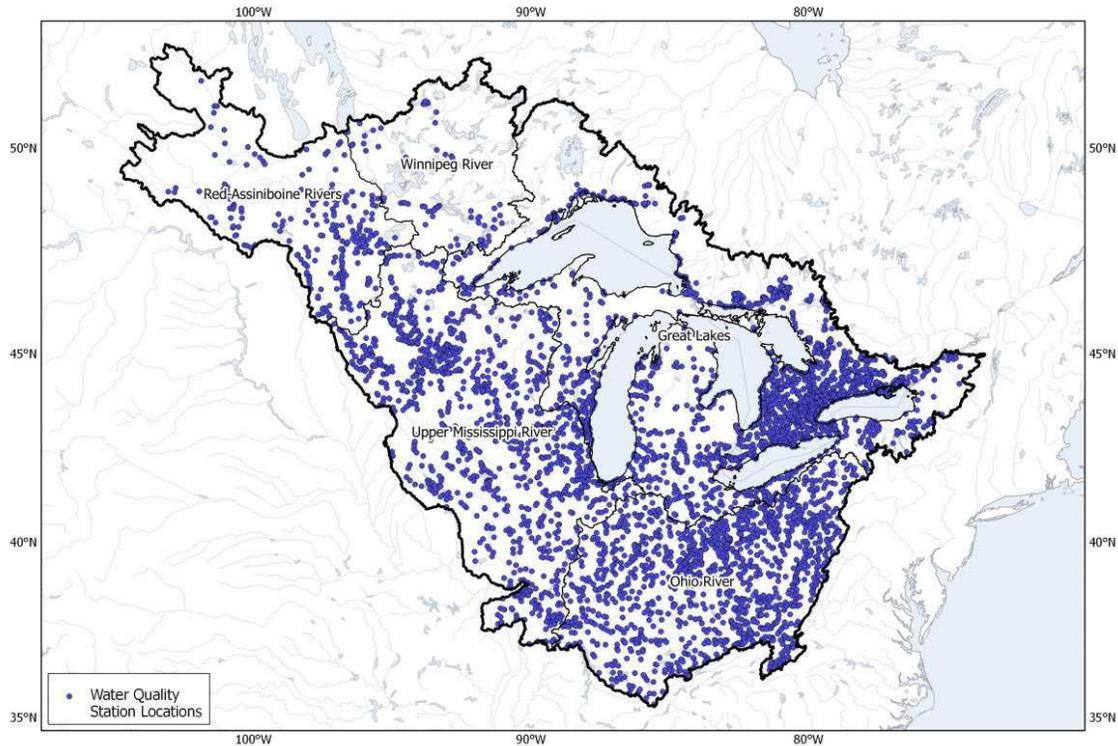


Figure 4: Calibration points (monitoring stations) used to preserve stream segments during thresholding and generalization.

3.4 Lakes data and shoreline segments

Waterbodies, such as lakes and reservoirs, are explicitly defined in SPARROW models to incorporate the increase in retention time imposed by these factors on the streamflow, which may affect nutrient decay processes (Schwarz *et al.*, 2006). Waterbodies in the SPARROW model are hydraulically connected to the digital stream network through artificial flow-path features (see Section 5.3).

For the U.S., waterbodies from the NHDPlus V2 data were used to calculate hydraulic load (see Section 5.3) for each reach segment associated with a waterbody. Additional waterbody information from the National Inventory of Dams (U.S. Army Corp of Engineers, 2016) supplemented the NHDPlus V2 dataset.

For the Canadian GL basin stream network, a subset of the OIHD, the Integrated Waterbody dataset, was initially used as the preferential data source. However, these data included too many waterbody features including streams, canals and ditches, which were unnecessary for the SPARROW model. Furthermore, these differences in features were not distinguished within the given dataset. This issue was overcome by using the CanVec dataset (Centre for Topographic Information, 2012), which was very similar to the OIHD Integrated Waterbody dataset. The CanVec dataset possessed necessary definitions of the waterbody features to aid in the removal of non-lake waterbodies. A minimum lake area threshold of 0.007 km², consistent with the medium-resolution NHDPlus V2 network criteria for lake inclusion, was applied to the Canadian dataset. Lakes that are geographically isolated from the stream network were removed. The final dataset included approximately 63,000 defined waterbodies.

In addition to stream segments, the Midcontinent SPARROW models also require shoreline segments to be defined that contribute direct runoff to the Great Lakes. As such, the GL basin stream network required delineation and identification of each shoreline segment along the Great Lakes coast between every river and stream draining into the lakes. Direct shoreline segments were obtained directly from the NHDPlus dataset. The OMNRF provided the shoreline for the Canadian coast of the Great Lakes in a separate single segment dataset. This shoreline was partitioned at locations where streams discharged along the shoreline. Additionally, some manual edits were required to incorporate island data (particularly in the U.S.) or inlet and bay data that were not properly captured in the datasets.

The OMNRF also provided a harmonized lake dataset for the WR basin between Ontario and the U.S. Review of this dataset revealed similar issues with non-lake waterbodies not being differentiated in the Ontario portion of the basin. In contrast, the harmonized NHD/NHN dataset was in good agreement along the border and thus used for the WR basin lakes dataset. Shoreline segments for Lake Winnipeg were created manually (i.e., digitized and edited) to capture the full connectivity between the WR and RA basins.

4 Catchment delineation

SPARROW models are developed using information derived from a routed, digital stream network and associated catchments. Catchment refers to an area whose local surface-water and groundwater flow drains directly to a stream segment in the network. SPARROW model input – describing nutrient input on the landscape – is defined for each catchment. Catchments are also used to account for watershed characteristics that enhance or inhibit nutrient delivery to the stream.

Catchment delineation was done using several methods depending on the available sources of data. Using the existing NHD product, the USGS, in cooperation with the U.S. Environmental Protection Agency (USEPA), developed a method to generate a hydro-enforced DEM (i.e., a DEM with embedded streamflow directions) for catchment delineation. This method was compared to other similar catchment delineation techniques (Johnston *et al.*, 2009) and deemed to be the most accurate. Hydro-enforcing the DEM aligns the DEM flow paths with the digital stream network so that the two datasets are in agreement and the streamflow direction of the network is preserved. The hydro-enforcement process uses manually delineated drainage divides from existing topographic maps or higher resolution data. Enforcing existing drainage divides ensures that delineated catchments agree with existing recognized watershed units used by water resource agencies. Additionally, mapped waterbodies from the NHD are enforced in the DEM through creation of an artificial reach on the digital network (i.e., a sloping gradient from the lake shoreline to the artificial stream) within the waterbody. Enforcing the waterbody shoreline provides better catchment delineations for the reaches of the stream network within and connecting to these waterbodies. The hydro-enforced DEMs are consequently used to compute flow direction and determine the catchments for each unique stream segment.

NHDPlus V2 delineated catchments were used for the Upper Mississippi River, Ohio River and the U.S. part of the GL basins. Using modified NHDPlus processing scripts, the same principles were applied to define the WR catchments. Initially, the NED DEM (Gesch *et al.*, 2002) was processed to create a WR basin hydro-enforced DEM. The Watershed Boundary Dataset (WBD) 12-digit Hydrologic Unit Codes (HUC12) (U.S. Department of Agriculture and Natural Resources Conservation Service, 2004) was used to enforce the drainage divides in the U.S. and along the international border where the hydrologic units were harmonized with comparable Canadian units. The verified binational stream networks for the WR basin (described in Section 3) were used as hydro-enforcement to ensure that flow paths of the DEM matched that of the digital stream network. Waterbodies from the harmonized NHD with NHN data were used to enforce the shorelines of networked lakes in the WR hydro-enforced DEM. A flow direction grid computed from the hydro-enforced DEM was used to generate the catchments for the WR basin. Development of the RA basin catchments is described by Jenkinson and Benoy (2015), which is very similar to the process described here but uses NHDPlus V1 for the southern portion of the Red River basin. Adjustments were made to the catchments of the RA basin to include an additional contributing drainage area (from HUC12). The catchments within the Canadian GL basin were computed using the OMNRF enhanced flow direction grid, which is a component of the OIHD. Along the international border, delineated catchments required manual edge-matching. For the transboundary waterways (e.g., Niagara River, St. Marys River, Detroit River, etc.) the harmonized stream network (NHDPlus V2 with the OMNRF's EWC) was reprocessed with the NED data to create a hydro-enforced DEM and a flow direction grid to generate catchments. These catchments were integrated with the surrounding catchments from NHDPlus V2 and the data derived from the Ontario datasets. Because of the varying data sources used to delineate catchments, gaps and overlaps occurred along basin borders. Therefore, additional processing was required to align the catchments seamlessly. For the Canadian part of the GL basin, precedence was given to either the NHDPlus V2 derived catchments or the OIHD derived catchments for areas of overlap. Small gaps were identified and through a series of GIS-based tools and manual manipulations were affixed to the most appropriate adjoining catchment.

The average size of the final catchments for the Midcontinent SPARROW models is 2.3 km². Catchments are shown in Figure 5.

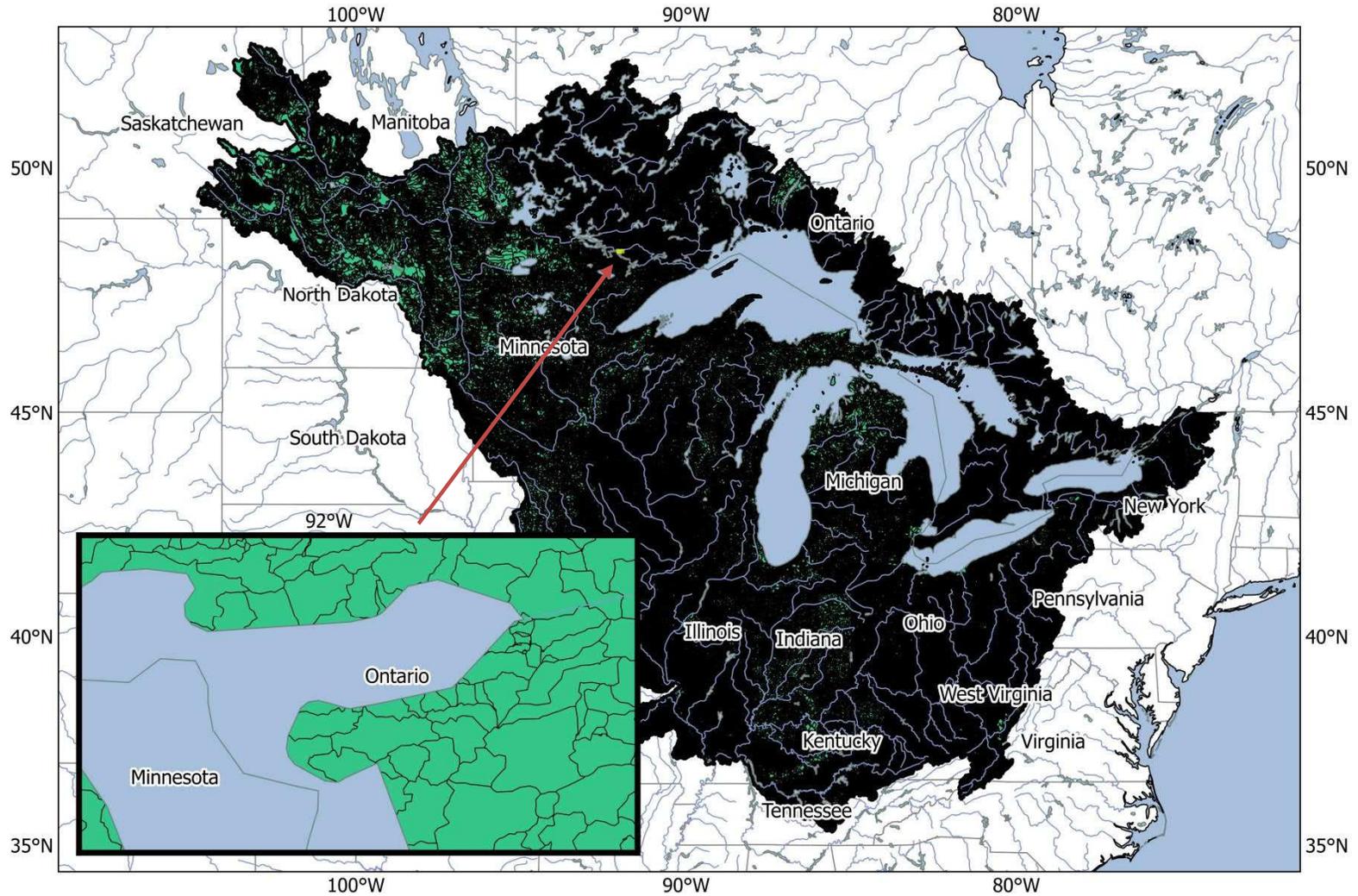


Figure 5: Midcontinent SPARROW model catchments. Enlarged view of the catchments is provided in the inset as an example, indicated by the yellow rectangle the red arrow is pointing to on the larger map.

5 Input variables to calculate decay of modelled constituents

The SPARROW model has the ability to quantify the loss (attenuation) of the modelled constituents caused by in-stream and in-reservoir decay. Stream decay rates are typically quantified for each stream segment using stream reach time of travel (reach length divided by stream velocity) usually grouped by various flow classes (Schwarz *et al.*, 2006). Reservoir decay rates are evaluated for each reservoir and waterbody on the stream network and are typically based on the inverse of hydraulic load (mean streamflow divided by water body surface area). Estimates of streamflow, velocity and inverse hydraulic load estimates are described below.

5.1 Flow and drainage area estimates

Within the SPARROW model, flows of each stream segment of the network are often used to group potential decay variables for each SPARROW catchment by flow class. Flows for the Midcontinent SPARROW model were computed using estimates of annual runoff (McCabe and Wolock, 2011; Wolock and McCabe, 2018) and accumulated throughout the stream network. The results at representative points were validated to available gauging station data. Additionally, the surface area of each unique SPARROW catchment was estimated and accumulated throughout the stream network. The accumulated areas were compared to observed drainage areas to enhance the validation of the stream network. Observed flows and drainage areas were collected from selected Water Survey of Canada gauging stations data extracted from ECCC's HYDAT.mdb, released on [Jan. 26, 2015]⁵. Selection criteria were based on data availability (i.e., continuous flows and representation of 2001- 2002 time period⁶) and proximity to the developed stream network.

A total of 279 gauging stations were selected for the GL basin stream network and 34 gauging stations for the WR basin stream network. Scatterplots were used to compare estimated and observed flows and drainage areas (Figure 6). The results for the GL basin stream network shown in Figure 6 indicate that estimated flows based on runoff maps developed from 30-year normal represented the observed flows well, where the observed flows range in data availability (i.e., anywhere between 2 and 30 years). A similar validation procedure for the WR basin stream network was conducted (Figure 7). This procedure demonstrated that estimated flows were approximately 18% lower than the measured flows in HYDAT. Therefore, the estimated flows were increased 18%.

⁵ Hereafter, the term HYDAT will be used to refer to ECCC's HYDAT.mdb data.

⁶ Criteria were limited to years between 1970 and 2012.

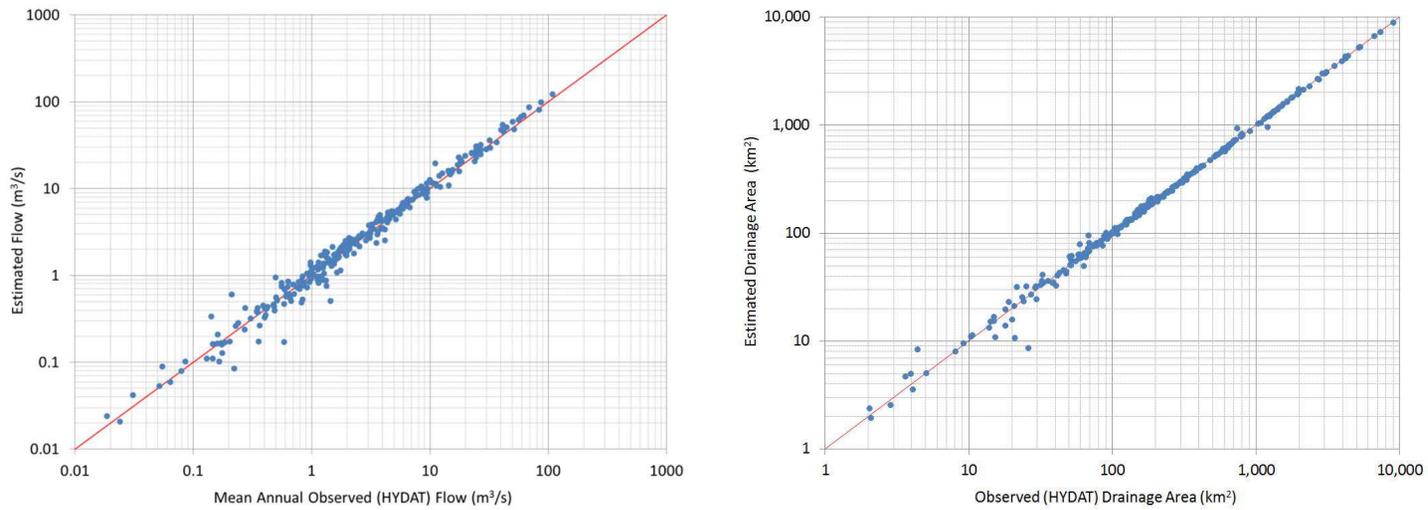


Figure 6: Validation results for estimated versus observed flows (left) and drainage areas (right) in the Great Lakes basin. The 1:1 reference line is indicated with a red line.

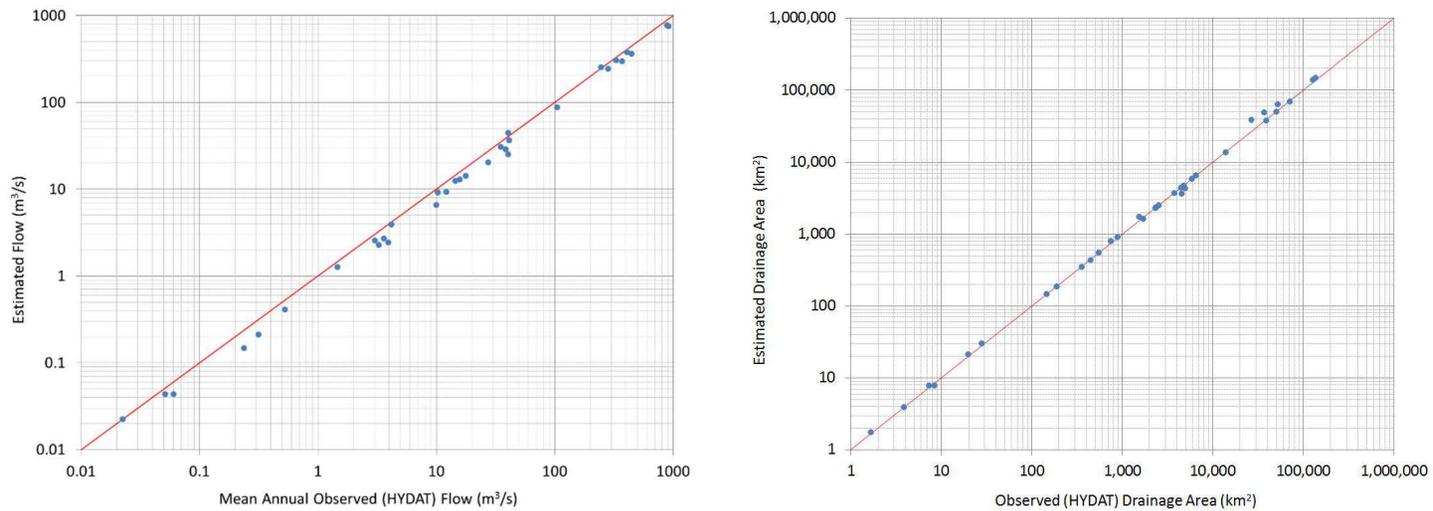


Figure 7: Validation results for estimated versus observed flows (left) and drainage areas (right) in the Winnipeg River basin. The 1:1 reference line is indicated with a red line.

5.2 Velocity estimates

Flow velocities for all individual stream segments in the U.S. were obtained from NHDPlus. Flow velocities for Canadian stream segments were calculated using the Jobson equation (Jobson, 1997). The Jobson equation estimates velocity from streamflow, drainage area and channel slope. Detailed procedure for calculating velocities is available in Jenkinson and Benoy (2015). For the GL and WR basins, the estimated flow and drainage area data were used with channel slopes calculated using available DEMs (see Section 7.2.3) to estimate velocities. Velocity estimates for the full stream network are illustrated in Figure 8. Small channel slopes in western Ontario – in the area concentrated with lakes – resulted in very low estimated stream velocities. Because most stream segments have reservoirs in this area, most of these velocities would not be used in SPARROW models.

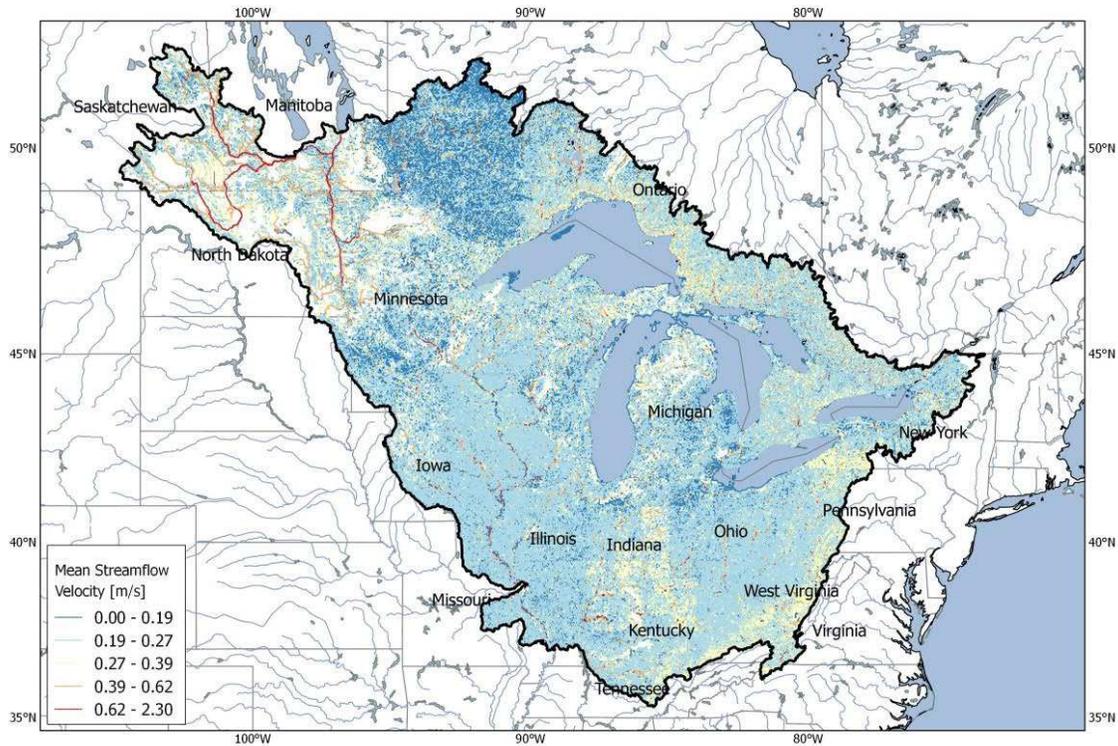


Figure 8: Estimated velocity in streams throughout the Midcontinental region.

5.3 Hydraulic load estimates

The reservoir/lake (Section 3.4) and streamflow (Section 5.1) datasets were used for the calculation of inverse hydraulic loads and examination of reservoir attenuation. Detailed procedure for this calculation is available in Jenkinson and Benoy (2015). Inverse hydraulic load estimates (computed from mean streamflow divided by water body surface area) for the full stream network are illustrated in Figure 9. The dense population of reservoirs and relatively low streamflow resulted in very low inverse of hydraulic load estimates for the stream network throughout Ontario.

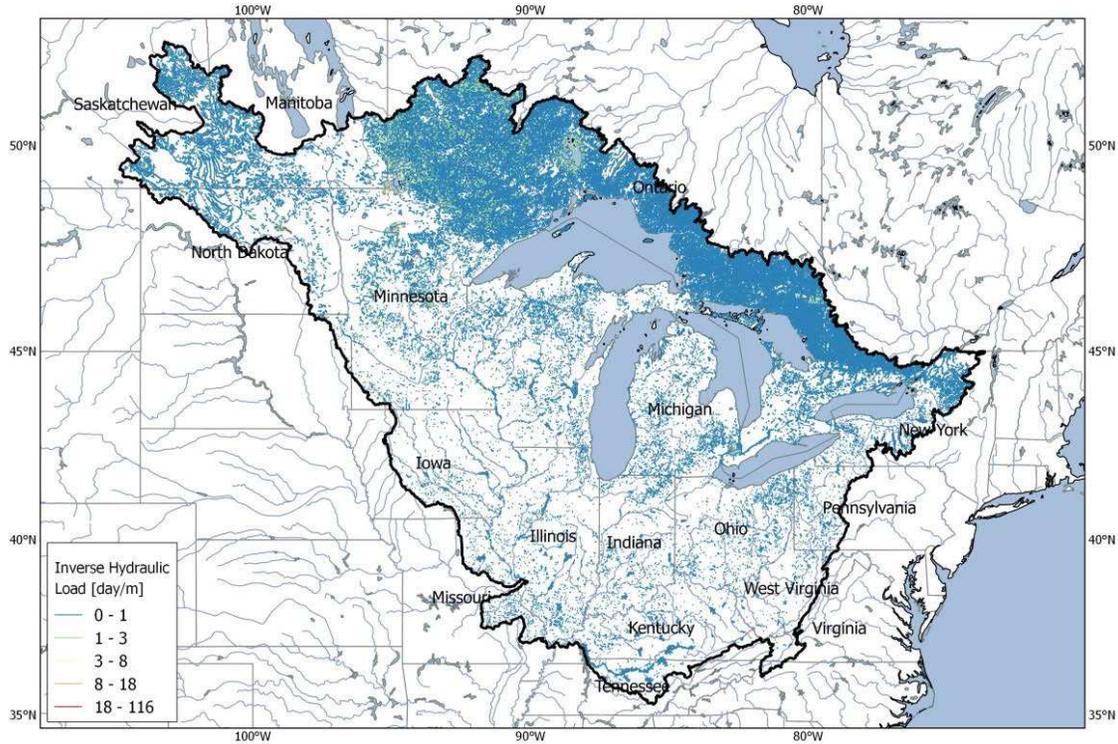


Figure 9: Estimated inverse of hydraulic load throughout the Midcontinental region.

6 Diversions

Structures that divert streamflow from the main channel to one or more alternate downstream channels (or flow paths) are called diversions. Diversions are represented in the SPARROW digital stream network using an attribute (*frac*) that describes the fraction of total flow diverted to alternative flow paths. If part of the flow is diverted out of the basin, then the *frac*s do not sum to 1.0. NHDPlus V2 includes numerous identified diversions in the U.S. (McKay *et al.*, 2012). In Canada, diversions were not readily available in the digital stream networks. However, several notable diversions in Canada that represent a significant amount of diverted streamflow were incorporated into the digital stream network. These included diversions at Lake St. Joseph, Long Lake and the Ogoki Reservoir.

Outflow from Lake St. Joseph, in northwest Ontario, originally drained north toward James Bay except during high flows when it also drained into the WR basin. In 1958, a structure was constructed diverting most of the Lake St. Joseph flow into the WR basin (Lake of the Woods Control Board, 2018). Outflow from Long Lake and the Ogoki Reservoir, north of Lake Superior in Ontario, discharged north toward James Bay. In 1939, most of the flow from Long Lake was diverted south toward Lake Superior. Outflow from the Ogoki Reservoir was diverted in 1943, so that most of the flow was routed south into Lake Nipigon and eventually Lake Superior (Peet and Day, 1980; Lakehead University, 2016). All three diversions are illustrated in Figure 10.

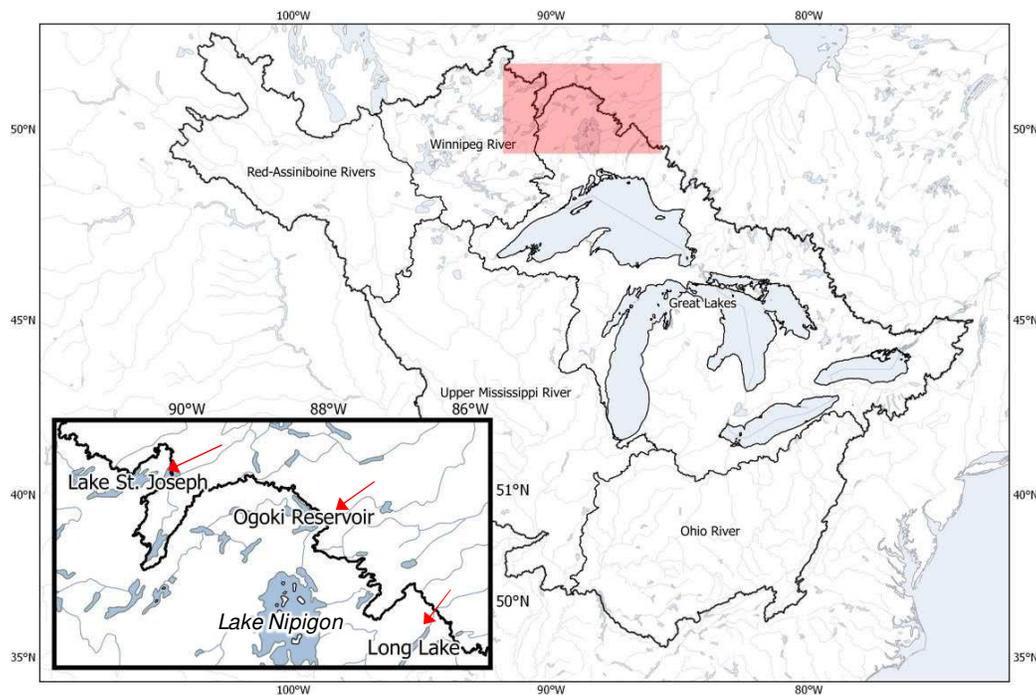


Figure 10: Stream diversions in Canada (with respect to the extents of the Midcontinent region). The inset provides a close-up view of Lake St. Joseph, Ogoki Reservoir and Long Lake, with red arrows indicating the direction of flow diversion. Lake Nipigon is identified in the inset.

The average flow diverted into the Midcontinent region was estimated using HYDAT flow data. Historical discharges for the period from 1970 to 1994 were used to estimate the average diversion fractions for flow kept in the study area (Lake St. Joseph, 0.87; Long Lake, 0.92; Ogoki Reservoir, 0.91). Diversion fractions were added to the digital stream network without modifying the original stream routing attributes.

7 Nutrient source and delivery variables

Sources in the SPARROW model describe the amount of the modelled constituent (in this case total phosphorus [TP] and total nitrogen [TN]) input throughout the entire region. Delivery variables explain or describe variability in the amount of the nutrients delivered to the stream, in other words, enhanced or reduced transport to streams. Values for all sources and delivery variables were attributed to each catchment throughout the Midcontinental region using zonal statistics (Esri, 2010) resampled to a 30-m resolution, except for wastewater-treatment plant inputs.

7.1 Input sources

Inputs of TP and TN from both naturally derived and anthropogenic sources throughout the entire Midcontinental region were considered. The following TP and TN sources were included for possible inclusion in the Midcontinent SPARROW models:

- various land-cover categories,
- inorganic farm fertilizers,
- manure,
- atmospheric deposition,
- wastewater-treatment plant effluent, and
- non-modeled watersheds that contribute nutrients to the model area (i.e., Qu'Appelle River and Missouri River watersheds)

7.1.1 Land cover

Land cover may be used as a source variable of TP or TN in a SPARROW model when specific inputs in certain areas are difficult to quantify. Within SPARROW, inputs from specific land covers are quantified as export coefficients during model calibration. For example, export from urban areas can be extracted from SPARROW models (Robertson and Saad, 2011). In order to be considered in the Midcontinent SPARROW models, a consistent land-cover dataset was required using data products from both U.S. and Canada. Land-cover data provided for the Canadian portion of the Midcontinental region were data circa 2000, at a 30-m resolution, from NRCan's Geobase (Centre for Topographic Information, 2009). Land cover data for the U.S. portion were obtained from the U.S. National Land Cover Dataset (NLCD) circa 2001 (Homer *et al.*, 2007), also at a 30-m resolution. The two data products were synchronized using a set of cross-walk tables (provided in Appendix B) to produce a common harmonized land cover product over the domain (Figure 11).

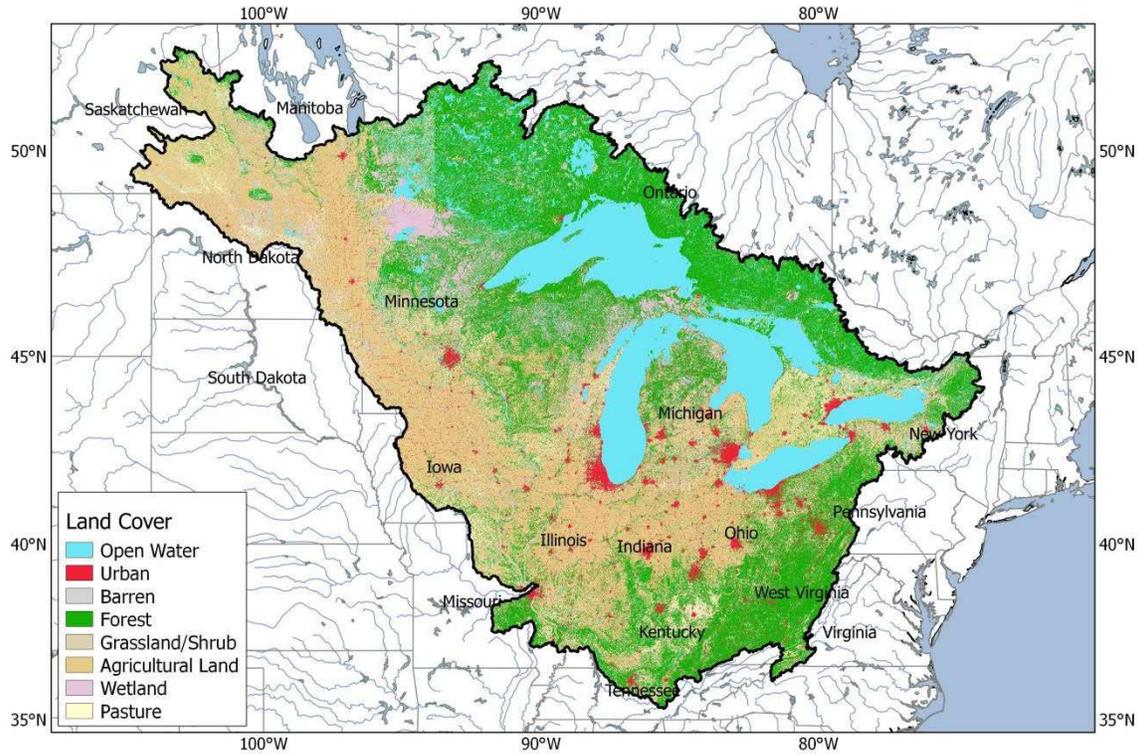


Figure 11: Harmonized land cover throughout the Midcontinental region.

7.1.2 Fertilizer

Inorganic farm fertilizers (fertilizer hereafter) inputs in the U.S. were based on county-level estimates from Ruddy *et al.* (2006). TP and TN inputs from farm fertilizers were derived from 2002 county sales aggregated by state and allocated back to the individual counties based on their respective fertilizer expenditures as reported by the Association of American Plant Food Control Officials and the U.S. Census of Agriculture (Ruddy *et al.*, 2006; National Agriculture Statistics Service, 2004). The total county fertilizer input was then allocated to each SPARROW catchment based on the fraction of total agricultural land of the county that is in each catchment (i.e., *agricultural land* class as defined in the land-cover layer).

Fertilizer inputs in Canada were not available; therefore, fertilizer inputs needed to be estimated for this area. In the process of quantifying fertilizer inputs for the RA SPARROW model, Jenkinson and Benoy (2015) worked with AAFC staff to obtain the necessary data. A similar approach was used to estimate fertilizer inputs to the catchments in Canadian GL and WR basins from 2001 Census of Agriculture (Agriculture and Agri-Food Canada, 2002; 2007) data by census divisions (CD) and census consolidated subdivisions (CCS) that are a subset of CD. The CCS data provided greater resolution and were used where available, otherwise CD data were used. Appendix C provides further detail for calculating fertilizer loadings using these data. Once calculated, as with the U.S. county data, Canadian data were allocated to each SPARROW catchment by the fraction of the total agricultural land in the CD or CCS (Figure 12).

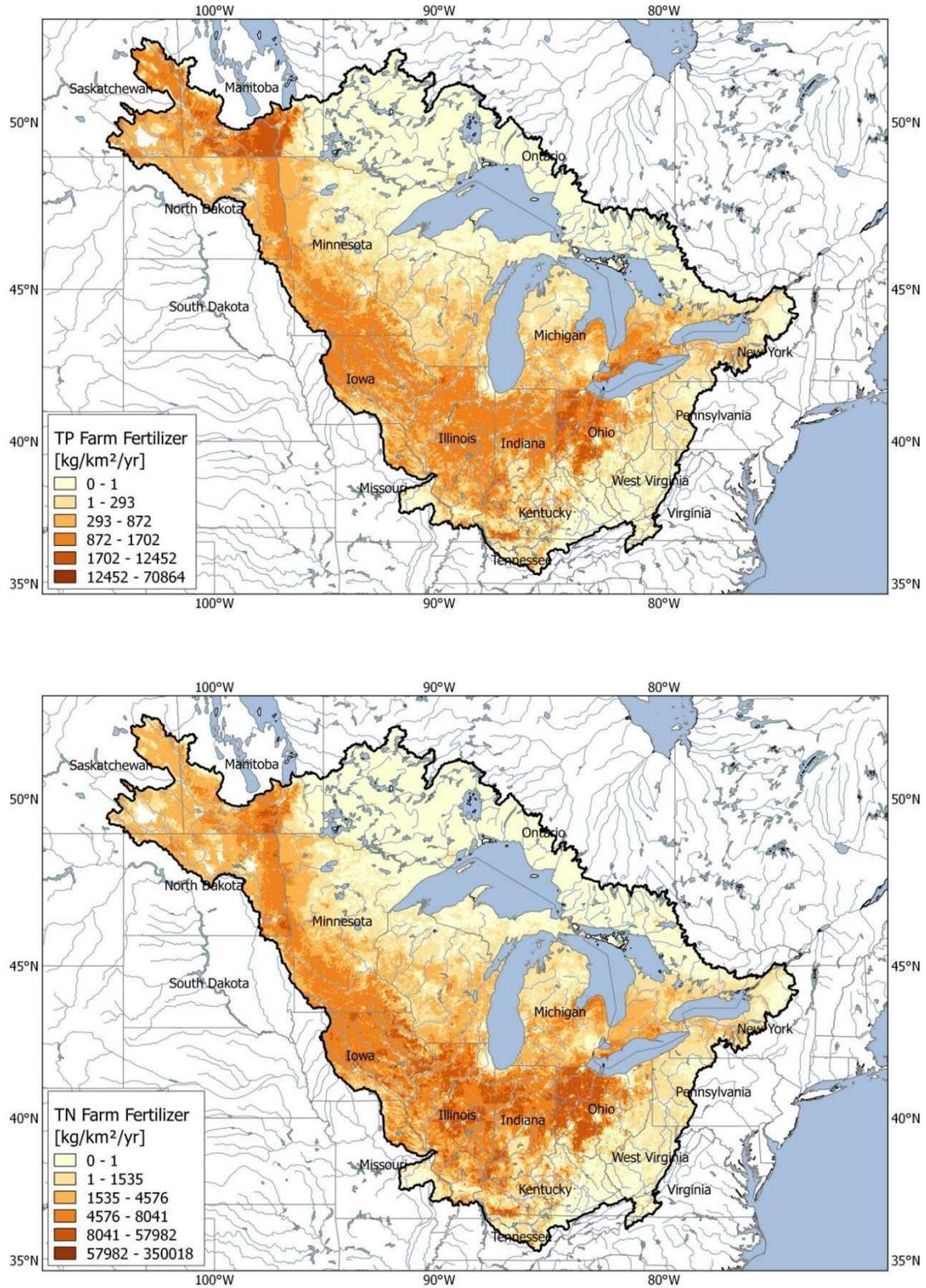


Figure 12: Total phosphorus (above) and total nitrogen (below) input rates, by catchment, throughout the Midcontinental region.

7.1.3 Manure

Livestock head counts are collected in both U.S. and Canada. A cross-walk table was devised to harmonize common classes of livestock between the U.S. and Canada data following Jenkinson and Benoy (2015) (provided in Table C-3 Appendix C). Livestock head counts were then multiplied by P and N coefficients for each animal type based on production rates (Ruddy *et al.*, 2006) (provided in Table C-4 Appendix C). Total calculated manure was allocated to each SPARROW catchment based on the fraction of the agricultural and pasture land in the catchment (i.e., *agricultural land* and *pasture* class as defined in the land-cover layer). Livestock head counts by county in the U.S. and by sub-sub drainage area (SSDA) from the Interpolated Census of Agriculture (Agriculture and Agri-Food Canada, 2001) in Canada were used to allocate the P and N from manure. Where SSDA data were suppressed, as explained in Appendix C, 2001 Census of Agriculture (Agriculture and Agri-Food Canada, 2002; 2007) data by CD were substituted. For more information refer to Table C-5 Appendix C. County-level head counts from the 2002 Census of Agriculture (National Agriculture Statistics Service, 2004; Mueller and Gronberg, 2013) were used for the U.S portion of the model area. The N and P input rates from manure, by catchment, throughout the Midcontinental region are shown in Figure 13.

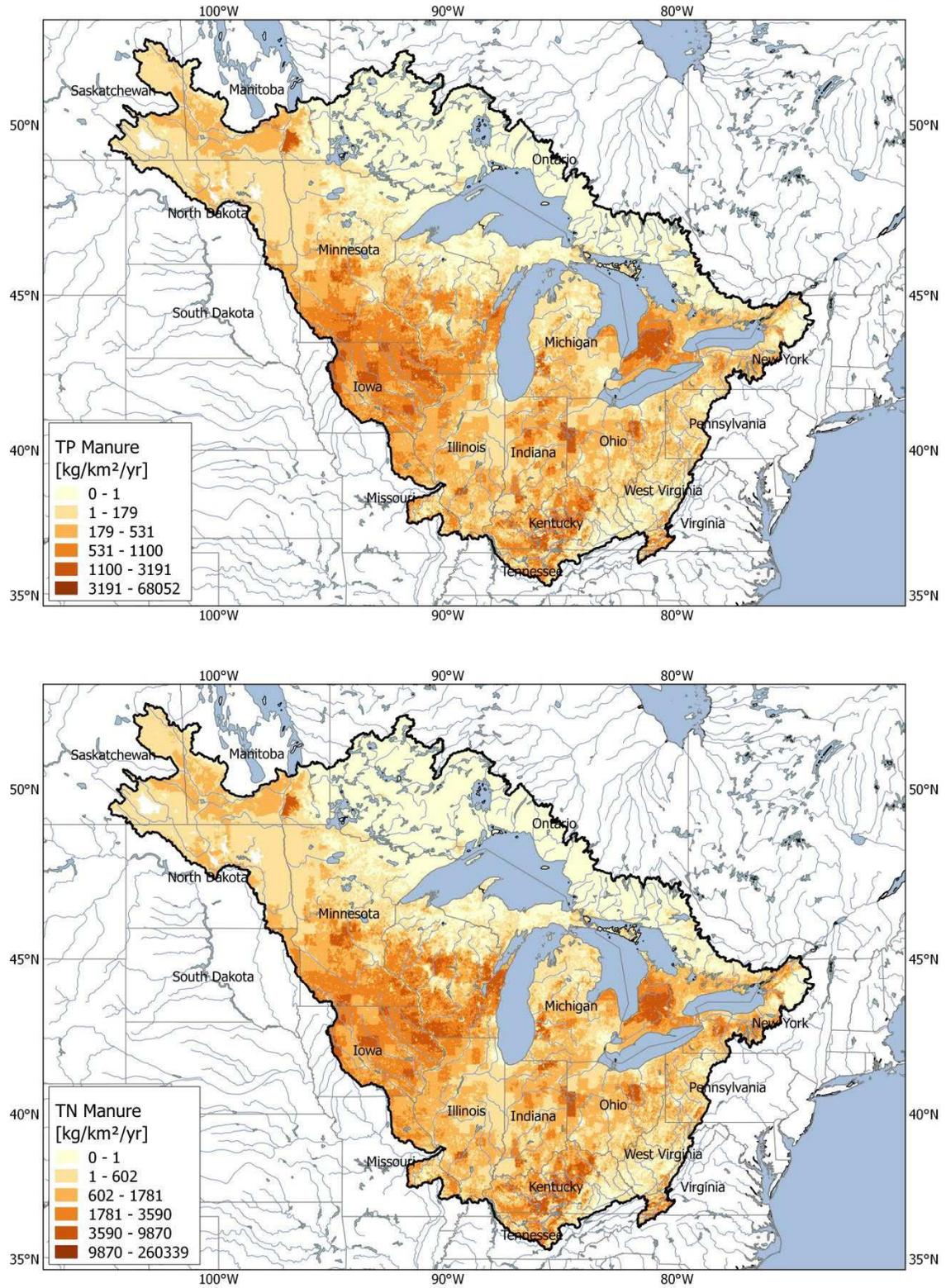


Figure 13: Total phosphorus (above) and total nitrogen (below) input rates, by catchment, from manure spread on agricultural and pasture lands throughout the Midcontinental region.

7.1.4 Atmospheric deposition

In previous SPARROW models developed for the U.S., atmospheric deposition was a significant contributing factor to TN loading (Robertson and Saad, 2011). For the Midcontinent SPARROW models, atmospheric deposition of N (a sum of wet and dry deposition) were estimated from the Community Multiscale Air Quality Modeling System (CMAQ) annual deposition rates for 2002 (Schwede *et al.*, 2009; Hong *et al.*, 2011; U.S. Environmental Protection Agency, 2018). Two different resolutions were employed (12 km and 36 km); 12-km resolution was used where available. The 36-km resolution region data were used to supplement any missing regions, primarily in the northern reaches of the Midcontinental region. Atmospheric deposition rates of N throughout the Midcontinental region are shown in Figure 14.

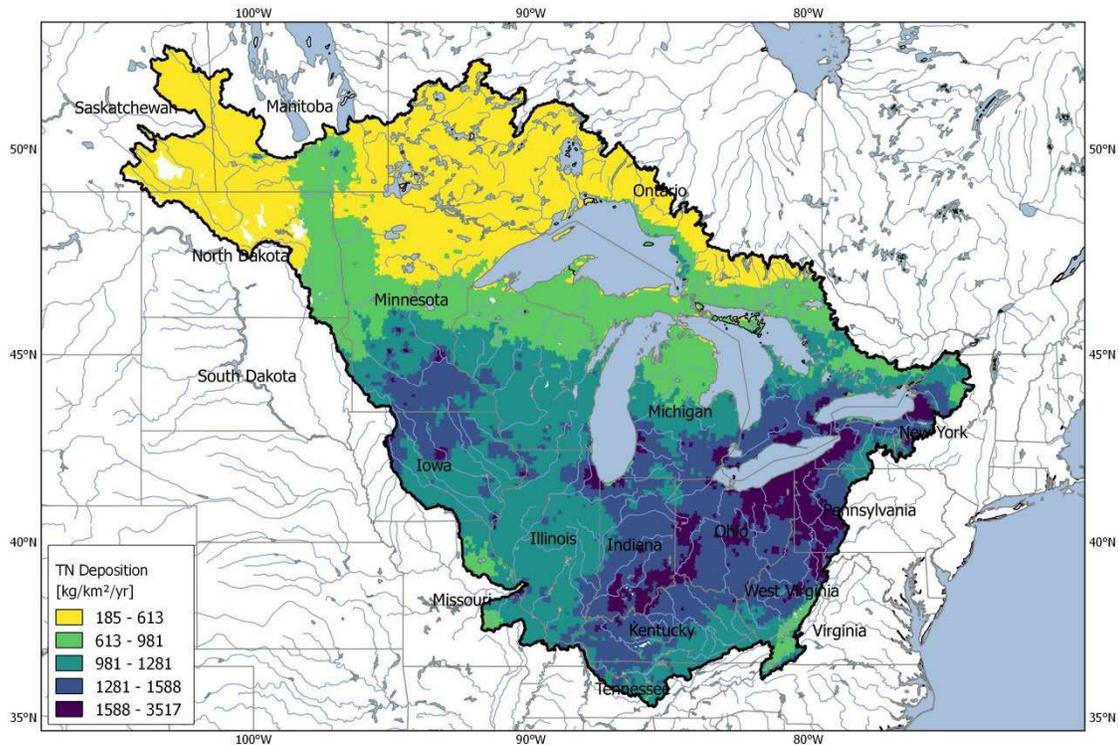


Figure 14: Atmospheric deposition rates of total nitrogen from CMAQ, by catchment, throughout the Midcontinental region.

7.1.5 Wastewater-treatment plant inputs

P and N inputs from wastewater-treatment plants (WWTPs) throughout the Midcontinental region were based on effluent data for the U.S. and Ontario, and population estimates for Manitoba and Saskatchewan. For the U.S., 2002 annual WWTP effluent loads were calculated for each facility using monthly effluent flow and constituent concentration data from the USEPA Permit Compliance System national database (Robertson and Saad, 2011; Maupin and Ivahnenko, 2011). Effluent flow data were readily available for each facility; however, TN and TP concentration data were often unavailable (TP concentration data were more readily available than TN data). In cases where effluent concentration data were not available, effluent loads were estimated using typical pollutant concentrations (TPCs) based on the type and size (relative to effluent flow) of a facility, following methods described by Maupin and Ivahnenko (2011).

For Ontario, 2002 annual WWTP effluent loads were calculated using monthly flow and concentration data provided by the Ontario Clean Water Agency and Ministry of the Environment and Climate Change (personal communication, Antonette Arvai, 2013). Similar to the U.S., TP concentrations were occasionally and TN concentration data were frequently not available for some facilities in Canada. In such cases, TP and TN TPCs were generated based on either median concentration by facility treatment level (primary, secondary, tertiary or lagoon), or overall median concentration when facility treatment level information was missing (Table 1). Median concentrations were based on available data for 2002 through 2011.

Table 1: Median TP and TN WWTP effluent concentrations for Ontario, 2002 through 2011

Constituent	Overall	Primary	Secondary	Tertiary	Lagoon
Median Concentration, mg/L (overall and by treatment level)					
TP	0.3	0.76	0.37	0.13	0.26
TN	13.6	16.4	14.3	16.0	6.7

For WWTPs in Ontario, effluent records were identified as *continuous* or *seasonal* in the database. Facilities identified as continuous occasionally did not have all 12 months of flow and concentration data. For these sites, monthly loads were first calculated using the available monthly flow and concentration data and summed. Annual loadings were then computed by scaling up measured loads by the fraction of the year represented by available data. For example, if a facility had 9 months of available load data, the annual load was calculated as a sum of monthly loads multiplied by 1.33 (or 12/9). For seasonal facilities with missing monthly data, annual loads were calculated as the sum of available monthly loads.

Estimates of TP and TN loads from WWTPs in Manitoba and Saskatchewan were obtained from the MCWS, the Saskatchewan Ministry of Environment and the Saskatchewan Water Security Agency (personal communication, Elaine Paige, 2013). Data in the RA basin were compiled for the RA SPARROW model area by Jenkinson and Benoy (2015). Data in the Manitoba region of the WR basin were compiled by Bourne *et al.* (2002). Effluent loads were calculated using 2008 population estimates and removal efficiencies based on facility treatment levels (Chambers *et al.*, 2001). Influent loads were calculated from the number of people served by the facility multiplied by 1.23 kg per person per year for TP and 3.65 kg per person per year for TN (Chambers *et al.*, 2001). Effluent loads were calculated as the influent load multiplied by 1 minus the removal efficiency, which ranged from 0.59 to 0.66, depending on the facility treatment level.

Locations of the WWTP facilities are shown in Figure 15. The locations of the all facilities in the U.S. and Canada were visually inspected using aerial photographs and confirmed where possible. WWTP inputs were summed for each SPARROW catchment based on the facility locations.

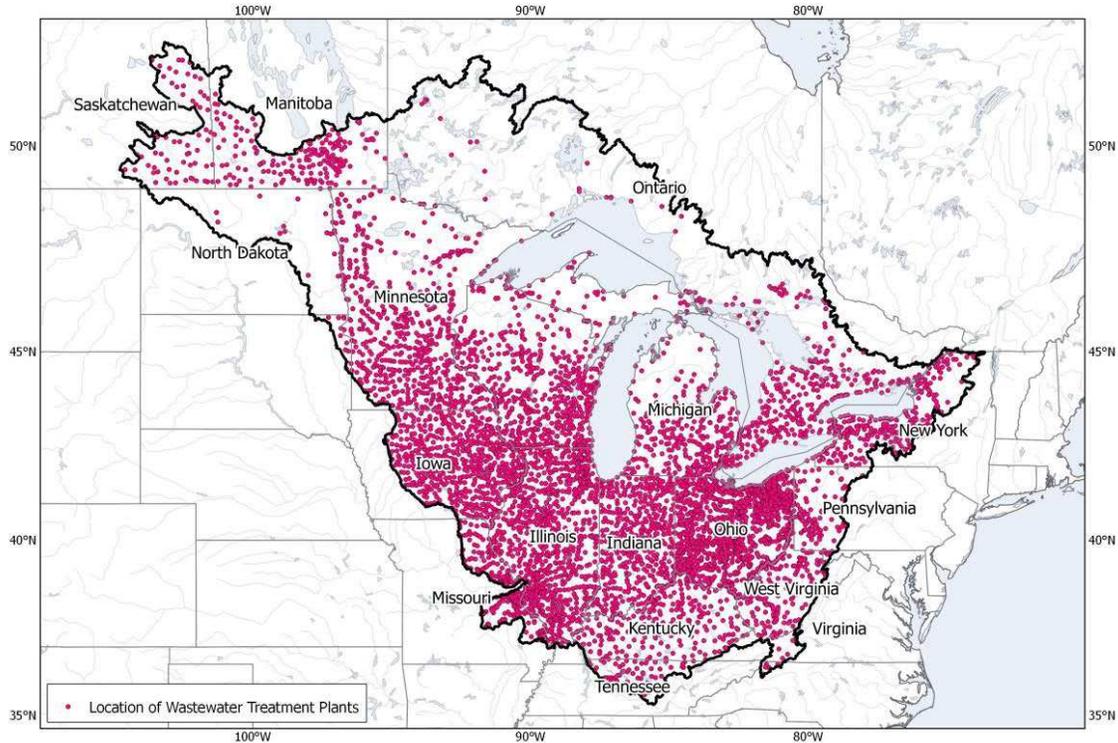


Figure 15: Wastewater-treatment plants used in the Midcontinent SPARROW models.

7.1.6 Non-modeled contributing areas

The Qu'Appelle River in Saskatchewan and Missouri River in central U.S. contribute P and N to the Midcontinental region; however, neither basin is explicitly included in the region. Difficulties in developing a digital stream network for the Qu'Appelle River prevented inclusion of its area. The Missouri River was excluded in an effort to limit the size of the modelled area. To account for the P and N inputs from these rivers, the estimated long-term mean annual TP and TN loads from these areas were input as point sources at the location in the digital stream network that represents the confluence with the non-modeled areas. The long-term mean annual loads for the Missouri River (Station ID 06934500; TP=28,220,000 kg/yr; TN=195,900,000 kg/yr) were estimated using the computer program Fluxmaster (Schwarz *et al.*, 2006), similar to the loads used for the regional SPARROW models (Saad, *et al.*, 2011). The long-term mean annual loads for the Qu'Appelle River (Station ID SA05JM0014, TP=67,880 kg/yr; TN=472,700 kg/yr) were estimated using a newer version of the computer program Fluxmaster following methods described in Saad *et al.* (2018).

7.2 Delivery variables

Delivery variables in SPARROW describe properties of the landscape that relate to climatic, natural or anthropogenic terrestrial processes affecting contaminant transport from the land to the stream. The following delivery variables were examined for inclusion in the Midcontinent SPARROW models:

- climate (i.e., precipitation, air temperature and runoff),
- soil characteristics,
- mean catchment slope, and
- tile drainage.

7.2.1 Climatic data

Climatic data, such as air temperature and precipitation, have been shown to be significant variables in earlier SPARROW modelling studies (Robertson and Saad, 2011; 2013). Runoff was another variable considered for this category because it represents a direct delivery method from land to water relying on climatic data, such as air temperature and precipitation.

Air Temperature

For the U.S. portion of the Midcontinental region, mean annual air temperature data (30-year normals from 1971 to 2000) were obtained from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) Climate Group (PRISM Climate Group, 2006). PRISM data are comprised of average values at the end of each decade over a 30-year period. For the Canadian portion of the region, air temperature data collected by the Canadian Forest Service (CFS) (McKenney *et al.*, 2006) were used. CFS temperature data consist of minimum and maximum temperature grids, whose values were reprocessed as defined in (McKay *et al.*, 2012) to ensure congruity between U.S. and Canada. Mean annual temperatures throughout the Midcontinental region are shown in Figure 16.

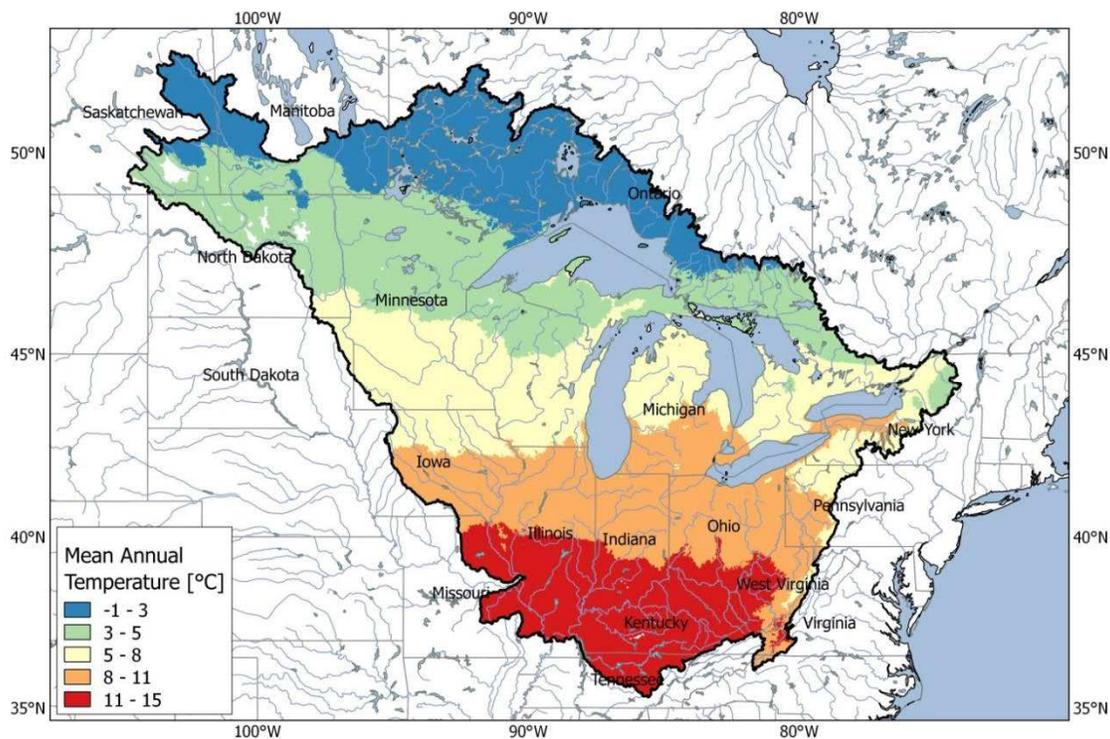


Figure 16: Mean annual air temperature, by catchment, throughout the Midcontinental region.

Precipitation

Mean annual precipitation rates for the entire region were obtained from Canadian Precipitation Analysis (CaPA) (National High Impact Weather Laboratory, 2014) data. Jenkinson and Benoy (2015) demonstrated that CaPA data were seamless across the U.S. – Canada border. The CaPA data for the Midcontinent SPARROW models were obtained from ECCC (personal communication, Bruce Davison, March 27, 2015). Data were provided in 6-hour time-steps, covering years 2002 to 2010, which were accumulated on an annual basis and calculated to 9-year mean annual values. The original grid cell resolution was provided in 10 km, with precipitation provided in units of metres. Mean annual precipitation throughout the Midcontinental region is shown in Figure 17.

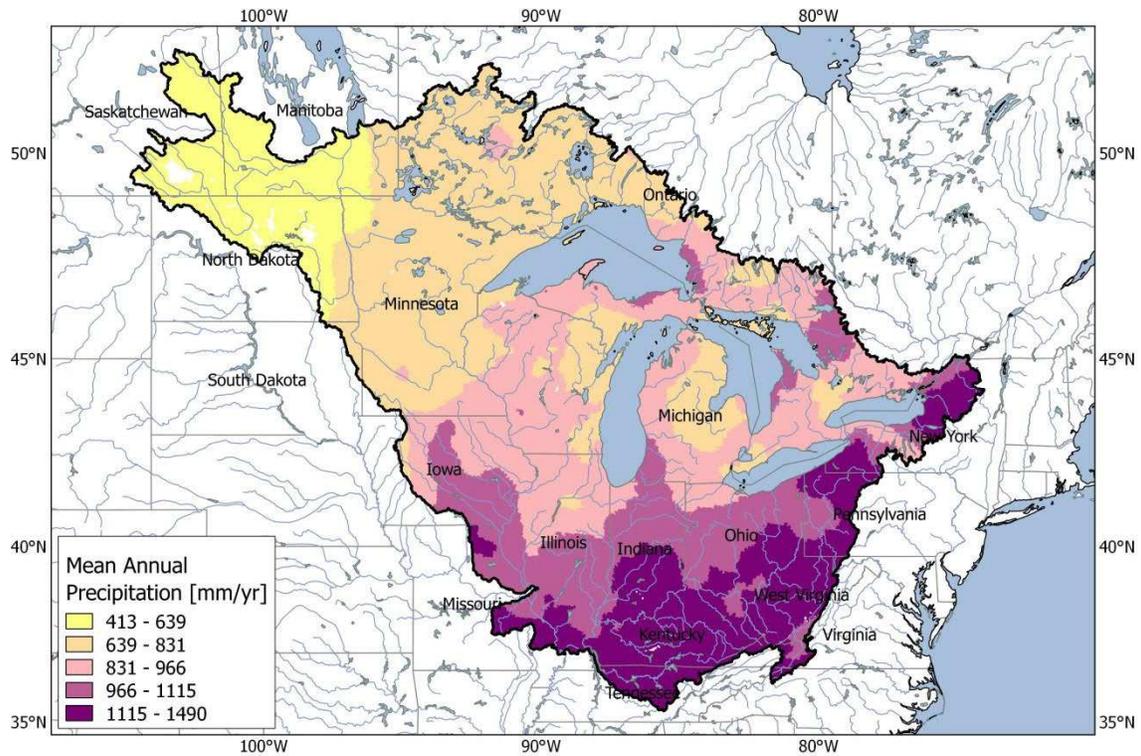


Figure 17: Mean annual precipitation by catchment, throughout the Midcontinental region.

Runoff

Runoff is controlled by climatic variables, such as precipitation and air temperature and by basin characteristics, such as slope and soils. Mean annual runoff covering the entire Midcontinental region was estimated using the Wolock and McCabe (1999) water balance model. U.S. data was obtained from Wolock and McCabe (2018). Canadian runoff data were computed from the PRISM and CFS data described in the Air Temperature discussion above. Mean annual runoff throughout the Midcontinental region is shown in Figure 18.

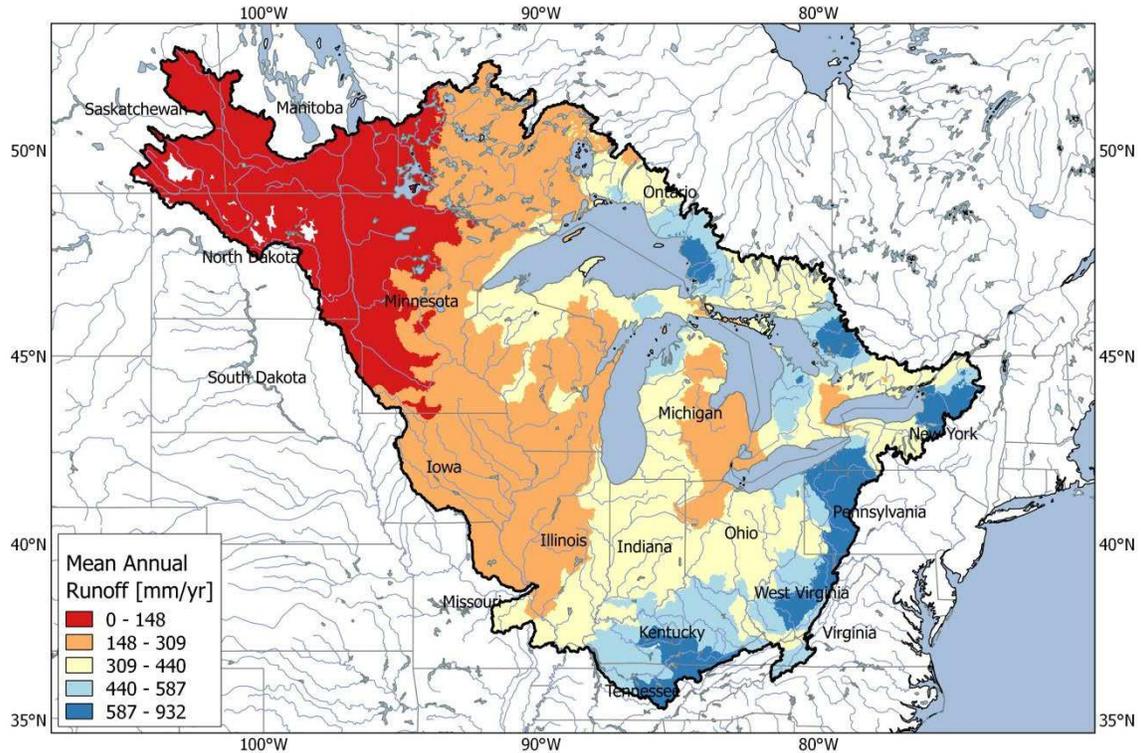


Figure 18: Mean annual runoff by catchment, throughout the Midcontinental region.

7.2.2 Soil characteristics

Variability in soil characteristics throughout the Midcontinental region were described using soil permeability and clay content. For the U.S. portion of the Midcontinental region, these datasets came from the Natural Resources Conservation Service State Soil Geographic (STATSGO) database (Schwarz and Alexander, 1995; U.S. Department of Agriculture, 1994; Wolock, 1997; Wieczorek and LaMotte, 2010). Data for the Canadian portion of the region were estimated with assistance by AAFC. Unlike for the provinces of Saskatchewan and Manitoba, soil permeability data for the province of Ontario were incomplete. Soil Landscape of Canada data version 2.2 (Agriculture and Agri-Food Canada, 1996) were reprocessed by AAFC (personal communication, Glenn Lelyk, 2015–2017) to be consistent with the U.S. data. To harmonize U.S. and Canadian data, the Canadian data were adjusted to the STATSGO soil permeability scale (0.06, 0.13, 0.4, 1.3, 4, 13, and 20 in/hr) where available. In areas where data were unavailable (primarily north of Lake Superior, Lake Huron and Georgian Bay), soil characteristics were estimated based on neighbouring data and a basic understanding of the geology in the area. This method provided less spatial detail than data used in the SPARROW models by Jenkinson and Benoy (2015) and Robertson and Saad (2011), but resulted in consistency over the full Midcontinental region. The final soil permeability and clay content distributions are shown in Figure 19.

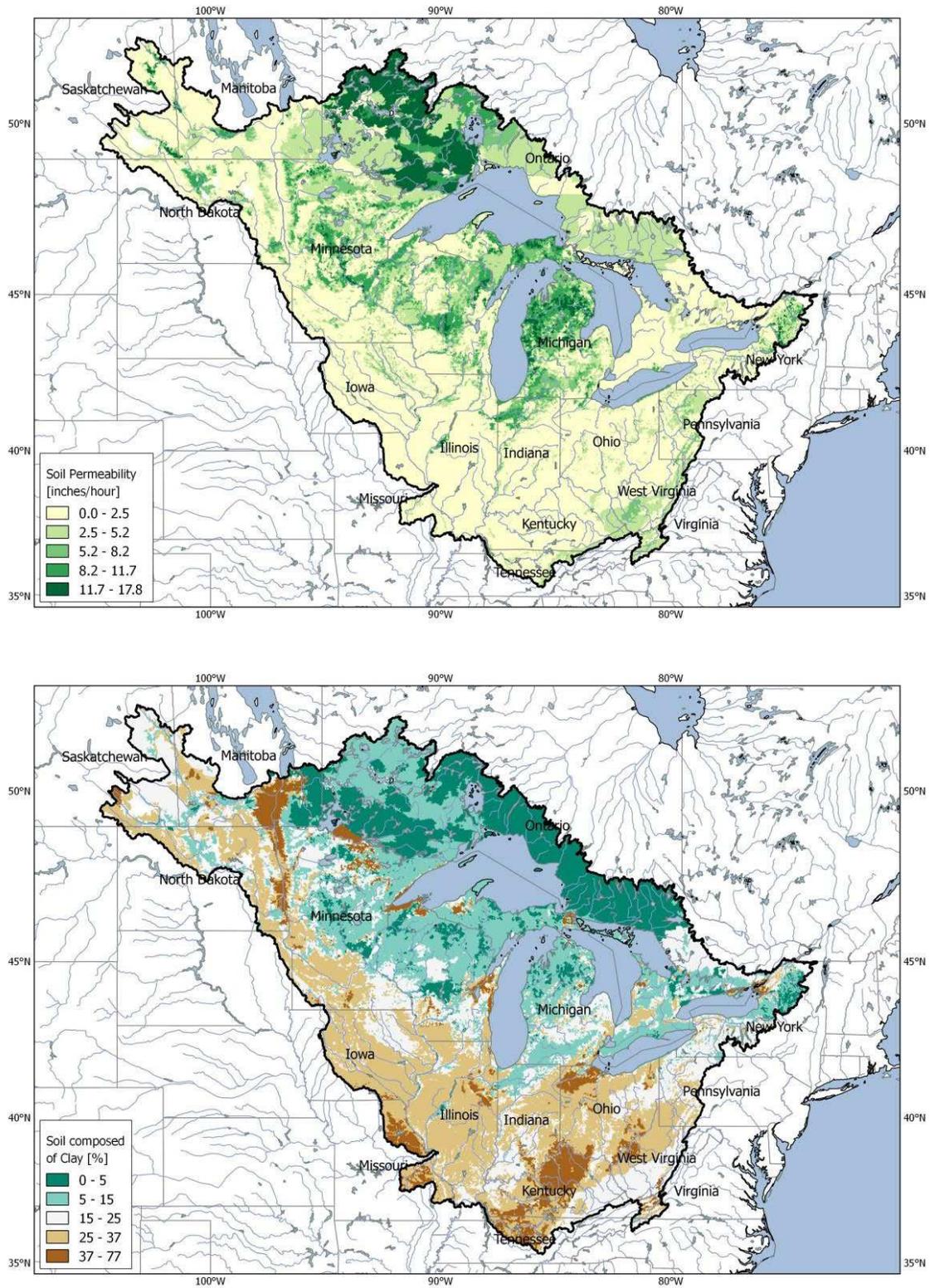


Figure 19: Soil permeability (above) and clay content (below) by catchment, throughout the Midcontinental region.

7.2.3 Mean catchment slope

Three different DEMs were used to calculate mean catchment slopes throughout the Midcontinental region. The NED DEM (Gesch *et al.*, 2002) was used for the U.S. portion of the region and the WR basin. The Canadian Digital Elevation Dataset (CDED) was used for the RA basin (Jenkinson and Benoy, 2015). The OIHD DEM (Ontario Ministry of Natural Resources and Forestry, 2012) was used for the Canadian portion of the GL basin. The NED and OIHD DEMs had a horizontal resolution of 30 m, while the CDED (Centre for Topographic Information, 2000) had a horizontal resolution of 15 m. For those catchments where calculations resulted in zero or negative slopes, a value of 0.00001 was enforced to avoid calculation errors in the SPARROW model. Mean catchment slopes throughout the Midcontinental region are shown in Figure 20.

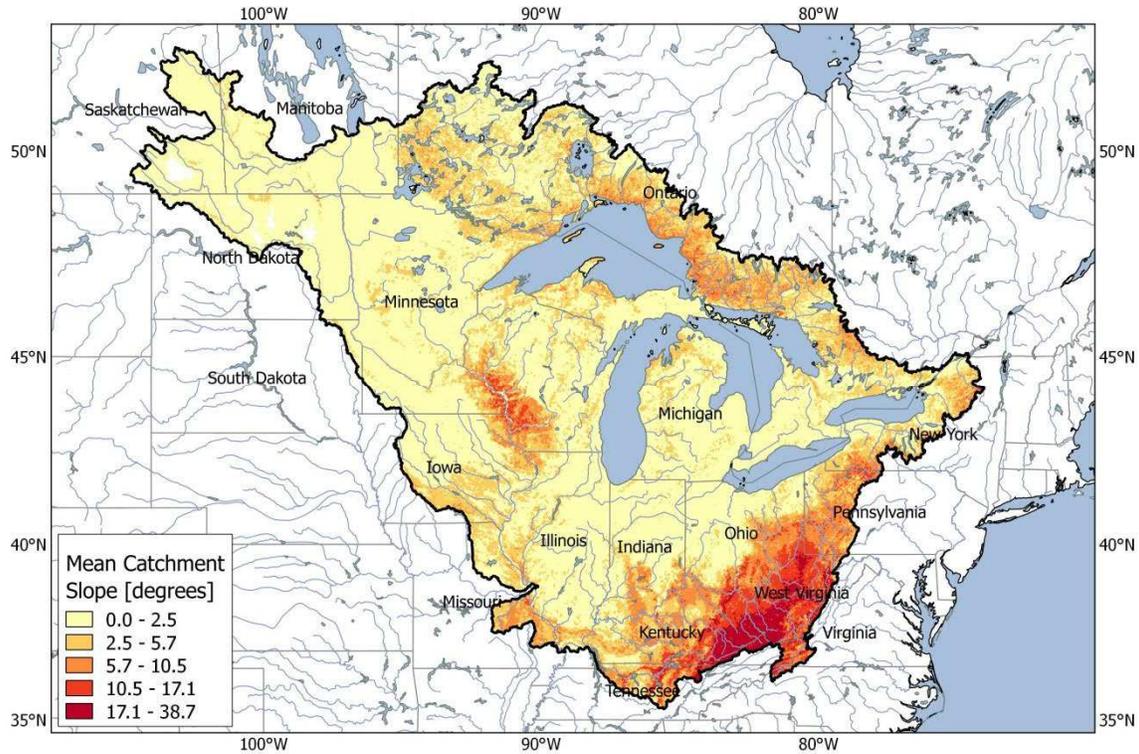


Figure 20: Mean catchment slopes by catchment, throughout the Midcontinental region.

7.2.4 Tile drainage

Robertson and Saad (2011) have shown, using SPARROW models, that the amount of area drained by tiles may affect the delivery of P and N to streams. Tile drainage data for the U.S. portion of the region were based on early 1990s information compiled by Nakagaki *et al.* (2016) following methods described by Sugg (2007). This information was processed to NHDPlus catchments using methods described by Wieczorek *et al.* (2016). For the Canadian portion of the region, tile drainage data of various forms were available at a provincial level. The Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) collects data describing tile drainage within the province and provides the information in the dataset *Tile Drainage Area* (Ontario Ministry of Agriculture Food and Rural Affairs, 2015). This information is available at a horizontal resolution of approximately 500 m and does not provide precise representations of the area drained. The assembled information provides the location and, in some cases, the size of installed tile drains. Areas with tile information were well represented as of 1983, when reporting and data collection commenced. Manitoba Sustainable Development (MSD) maintains a database describing tile drainage throughout the province based solely on location (personal communication, Tara Wiess, Manitoba Sustainable Development, 2016). Saskatchewan has historically been relatively arid, with only recent annual precipitation exceeding 500 mm. As a result of these arid conditions, tile drains have rarely been used in the province. Consequently, relevant information is deemed too insignificant to collect from Saskatchewan (personal communication, Doug Johnson, 2016); therefore, tile drains were assumed to not be present throughout this province.

Tile data for Ontario were processed for locations installed prior to and including 2002 (which resulted in approximately 74,000 records). Within this dataset, all of the records were provided as polygon features, but only 720 records had attribute information that had both polygon size and the area with tile drains. Therefore, to assess tile drainage in Ontario, installed drain-size area data from the 720 records with detailed data were used to obtain a relationship between polygon size and the area with tile drains. This relation was then used with polygon surface area to obtain information for all 74,000 areas with tile drains.

Tile drain locations in Manitoba were provided as a point feature by MSD. A total of 912 records existed, but only 28 had attribute information describing the area with tile drains. A dataset describing tile drainage areas in North Dakota (Finocchiaro, 2016) was used to improve the usefulness of the Manitoba data. The median tiled area for each location with North Dakota's tile data was calculated (0.615 km²) and applied as a factor to each Manitoba point location already provided. For both Ontario and Manitoba, the amount of tiled area in each catchment was summed and the percentage of each catchment underlain by tiles was calculated resulting in data similar to that compiled for the U.S. The percentages of each catchment underlain with tile drains are shown in Figure 21.

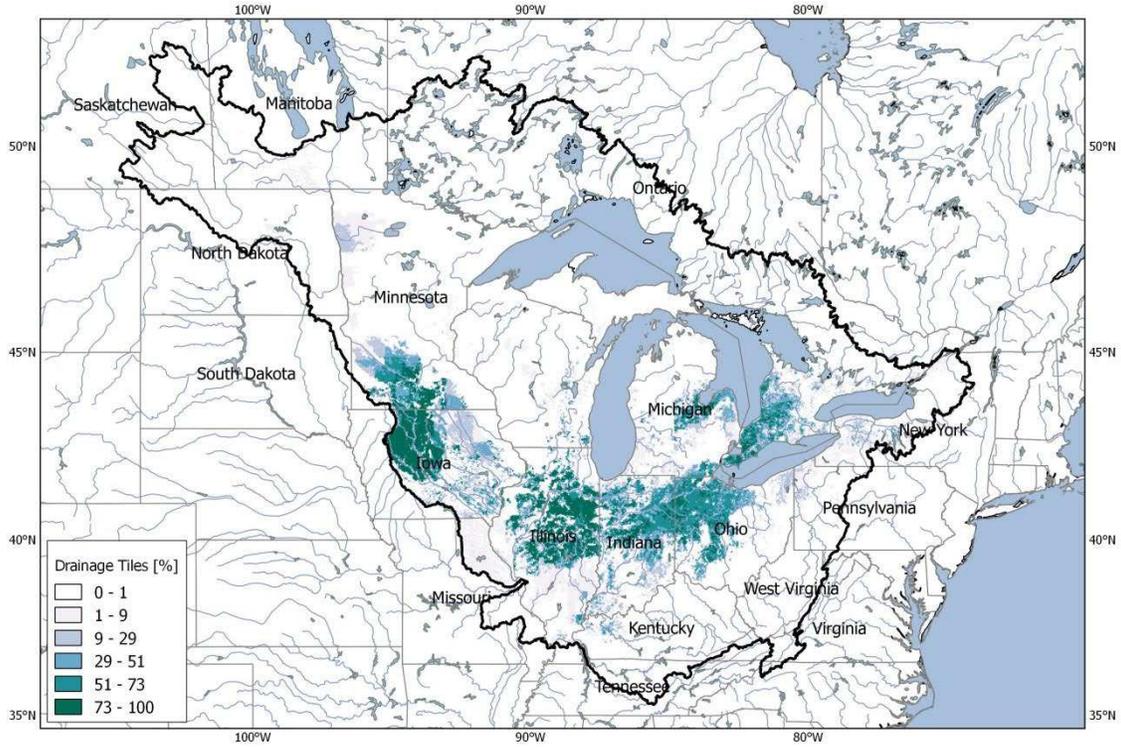


Figure 21: Percentage of the catchment underlain with tile drains, by catchment, throughout the Midcontinental region.

8 Summary

SPARROW models require extensive and contiguous geospatial datasets describing the stream network, catchments, nutrient inputs, and environmental variables affecting the transport of nutrients to the streams, and transport throughout the stream network. Using the data and procedures described in this report, a harmonized digital stream network was developed with information describing waterbodies and streamflow along the network. The stream network was used to create catchments (~2.5 km²) throughout the Midcontinental region of North America. Various geospatial datasets describing phosphorus and nitrogen inputs to each of the catchments (for conditions similar to 2002) and landscape characteristics that potentially describe variability in the delivery of nutrients to streams were assembled and harmonized across the U.S. – Canada border. The phosphorus and nitrogen inputs and landscape characteristics were then geospatially allocated to each of the catchments throughout the Midcontinental region. All of the nutrient input data, delivery data, and stream and reservoir data can be used to develop total phosphorus and total nitrogen SPARROW models for the Midcontinental region of North America. All of these data are available for download at url: <https://doi.org/10.4224/300.0001>.

9 References

- Agriculture and Agri-Food Canada. (1996). *Soil Landscape of Canada version 2.2*. accessed: January 19, 2018, url: [<http://sis.agr.gc.ca/cansis/nsdb/slc/v2.2/index.html>]
- Agriculture and Agri-Food Canada. (2001). Interpolated Census of Agriculture to Soil Landscapes, Ecological Frameworks, and Drainage Areas of Canada. pp. 32.
- Agriculture and Agri-Food Canada. (2002). *2001 Census of Agriculture*. accessed: March 15, 2016, url: [<http://www.statcan.gc.ca/pub/95f0301x/>]
- Agriculture and Agri-Food Canada. (2007). *2001 Census of Agriculture: Data tables*. accessed: March 15, 2016, url: [<http://www.statcan.gc.ca/pub/95f0301x/4151237-eng.htm>]
- Benoy, G.A., Jenkinson, R.W., Robertson, D.M., Saad, D.A. (2016). Nutrient delivery to Lake Winnipeg from the Red—Assiniboine River Basin – A binational application of the SPARROW model. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, 41(3), 429-447. doi:10.1080/07011784.2016.1178601.
- Bourne, A., Armstrong, N., Jones, G. (2002). A Preliminary Estimate of Total Nitrogen and Total Phosphorus Loading to Streams in Manitoba, Canada. Manitoba Conservation Report No. 2002 – 04, Water Quality Management Section, pp. 49.
- Brakebill, J.W., Wolock, D.M., Terziotti, S.E. (2011). Digital Hydro logic Networks Supporting Applications Related to Spatially Referenced Regression Modeling. *Journal of the American Water Resources Association (JAWRA)*, 47(5), 916-932.
- Canadian Council on Geomatics. (2009). *National Hydro Network (NHN)*. accessed: March, 2014, url: [<http://www.geobase.ca/geobase/en/data/nhn/index.html>]
- Centre for Topographic Information. (2000). *Canadian Digital Elevation Data (CDED)*. Natural Resources Canada, accessed: December 2012, url: [<http://www.geobase.ca/geobase/en/find.do?produit=cded>]
- Centre for Topographic Information. (2009). Land cover, circa 2000-vector data product, edition 1.0., Earth Sciences Sector, Natural Resources Canada, pp. 21.
- Centre for Topographic Information. (2012). CanVec: Data Product Specifications - Edition 1.1. Natural Resources Canada, Sherbrooke, Quebec. accessed: March, 2014, url: [ftp://ftp2.cits.nrcan.gc.ca/pub/canvec/doc/CanVec_product_specifications_v1_1.pdf]
- Chambers, P.A., Guy, M., Roberts, E.S., Charlton, M.N., Kent, R., Gagnon, C., Gagnon, G., Grove, G., Foster, N. (2001). Nutrients and their impact on the Canadian environment. Agriculture and Agri-Food Canada, Environment Canada, Fisheries and Oceans Canada, Health Canada, Natural Resources Canada, pp. 241.
- Esri. (2010). ArcGIS Desktop: Release 10.2.1. . Redlands, CA: Environmental Systems Research Institute, Inc.
- Finocchiaro, R.G. (2016). *Agricultural Subsurface Drainage Tile Locations by Permits in North Dakota*. U.S. Geological Survey data release, url: [<http://dx.doi.org/10.5066/F7QF8QZW>]
- Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., Tyler, D. (2002). The National Elevation Dataset. *Journal of the American Society for Photogrammetry and Remote Sensing*, 68(1).
- Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J.N., Wickham, J. (2007). Completion of the 2001 National Land Cover Database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, 73(4), 337-341.
- Hong, B., Swaney, D.P., Howarth, R.W. (2011). A toolbox for calculating net anthropogenic nitrogen inputs (NANI). *Environmental Modelling and Software*, 26(5), 623-633. doi:10.1016/j.envsoft.2010.11.012.
- Jenkinson, R.W., Benoy, G.A. (2015). Red-Assiniboine Basin SPARROW Model Development Technical Document. National Research Council of Canada, Ottawa, ON, pp. 65. url: [<http://ijc.org/files/publications/IJC%20RA%20Model%20Development%20Report%202015%20FINAL.pdf>]
- Jobson, H.E. (1997). Predicting travel time and dispersion in rivers and streams. *Journal of Hydraulic Engineering*, 123(11), 971-978. doi:10.1061/(ASCE)0733-9429(1997)123:11(971).
- Johnston, C.M., Dewald, T.G., Bondelid, T.R., Worstell, B.B., McKay, L.D., Rea, A., Moore, R.B., Goodall, J.L. (2009). Evaluation of catchment delineation methods for the medium-resolution National Hydrography Dataset. Scientific Investigations Report 2009–5233, U.S. Geological Survey, pp. 88. url: [<http://pubs.usgs.gov/sir/2009/5233/>]
- Laitta, M. (2010). Canada-U.S. transboundary hydrographic data harmonization efforts gain momentum. International Joint Commission, pp. 6. accessed: September 1, 2015, url: [http://www.ijc.org/rel/boards/watershed/Canada-US_Hydro_Harmonization_e.pdf]
- Lake of the Woods Control Board. (2018). *Board Description*, accessed: November 7, 2018, url: [<https://lwcb.ca/BoardDesc/index.html>]

- Lakehead University. (2016). *The Nipigon River* Lakehead University, accessed: November 5, 2018, url: [\[https://www.lakeheadu.ca/sites/default/files/uploads/53/outlines/2015-16/GEOG4411/Nipigon%20River.pdf\]](https://www.lakeheadu.ca/sites/default/files/uploads/53/outlines/2015-16/GEOG4411/Nipigon%20River.pdf)
- Maupin, M.A., Ivahnenko, T. (2011). Nutrient Loadings to Streams of the Continental United States from Municipal and Industrial Effluent. *Journal of the American Water Resources Association*, 47(5), 950-964. doi:10.1111/j.1752-1688.2011.00576.x.
- McCabe, G.J., Wolock, D.M. (2011). Independent effects of temperature and precipitation on modeled runoff in the conterminous United States. *Water Resources Research*, 47(11), n/a-n/a. doi:10.1029/2011WR010630.
- McKay, L., Bondelid, T., Dewald, T., Johnston, J., Moore, R., Rea, A. (2012). NHDPlus Version 2: User Guide. pp. 180. accessed: February 16, 2018, url: [\[https://s3.amazonaws.com/nhdplus/NHDPlusV21/Documentation/NHDPlusV2_User_Guide.pdf\]](https://s3.amazonaws.com/nhdplus/NHDPlusV21/Documentation/NHDPlusV2_User_Guide.pdf)
- McKenney, D.W., Papadopol, P., Campbell, K.L., Lawrence, K.M., Hutchinson, M.F. (2006). Spatial models of Canada- and North America-wide 1971/2000 minimum and maximum temperature, total precipitation and derived bioclimatic variables. Frontline Technical Notes 106, Natural Resources Canada, Great Lakes Forestry Centre, Sault St. Marie, Ontario, pp. 9.
- Moore, R.B., Dewald, T.G. (2016). The Road to NHDPlus — Advancements in Digital Stream Networks and Associated Catchments. *JAWRA Journal of the American Water Resources Association*, 52(4), 890-900. doi:10.1111/1752-1688.12389.
- Mueller, D.K., Gronberg, J.M. (2013). County-level estimates of nitrogen and phosphorus from animal manure for the conterminous United States, 2002. U.S. Geological Survey Open-File Report Report 2013-1065.
- Nakagaki, N., Wieczorek, M.E., Qi, S.L. (2016). *Estimates of subsurface tile drainage extent for the conterminous United States, early 1990s*. U.S. Geological Survey data release, url: [\[http://dx.doi.org/10.5066/F7RB72QS\]](http://dx.doi.org/10.5066/F7RB72QS)
- National Agriculture Statistics Service. (2004). 2002 Census of Agriculture—State and County Data, Volume 1: Geographic Area Series. U.S. Department of Agriculture, Washington, D.C. accessed: February 21, 2018, url: [\[http://agcensus.mannlib.cornell.edu/AgCensus/censusParts.do?year=2002\]](http://agcensus.mannlib.cornell.edu/AgCensus/censusParts.do?year=2002)
- National High Impact Weather Laboratory. (2014). The Canadian Precipitation Analysis (CaPA). Environment and Climate Change Canada, pp. 5. accessed: January 19, 2018, url: [\[http://collaboration.cmc.ec.gc.ca/cmc/cmoi/product_guide/docs/lib/capa_information_leaflet_20141118_en.pdf\]](http://collaboration.cmc.ec.gc.ca/cmc/cmoi/product_guide/docs/lib/capa_information_leaflet_20141118_en.pdf)
- Natural Resources Canada. (2002). National Topographic Database. url: [\[http://open.canada.ca/data/en/dataset/1f5c05ff-311f-4271-8d21-4c96c725c2af\]](http://open.canada.ca/data/en/dataset/1f5c05ff-311f-4271-8d21-4c96c725c2af)
- Ontario Ministry of Agriculture Food and Rural Affairs. (2015). *Tile Drainage Area*. Land Information Ontario, accessed: November 28, 2016, url: [\[https://www.javacoeapp.lrc.gov.on.ca/geonetwork/srv/en/main.home\]](https://www.javacoeapp.lrc.gov.on.ca/geonetwork/srv/en/main.home)
- Ontario Ministry of Natural Resources and Forestry. (2012). *Ontario Integrated Hydrology Data, 1.0 Initial Data Release: November 2012*. Land Information Ontario,
- Peet, S.E., Day, J.C. (1980). The Long Lake Diversion: An Environmental Evaluation. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, 5(3), 34-38.
- Preston, S.D., Alexander, R.B., Woodside, M.D., Hamilton, P.A. (2009). SPARROW MODELING—Enhancing Understanding of the Nation's Water Quality. U.S. Geological Survey Fact Sheet 2009–3019, pp. 6.
- PRISM Climate Group. (2006). *30-Year Normals*. Oregon State University, url: [\[http://www.prismclimate.org\]](http://www.prismclimate.org)
- Python Software Foundation. (2018). *Python Language Reference, version 2.7*, accessed: November 9, 2018, url: [\[http://www.python.org\]](http://www.python.org)
- Robertson, D.M., Saad, D.A. (2011). Nutrient Inputs to the Laurentian Great Lakes by Source and Watershed Estimated Using SPARROW Watershed Models. *Journal of the American Water Resources Association*, 47(5), 1011-1033. doi:10.1111/j.1752-1688.2011.00574.x.
- Robertson, D.M., Saad, D.A. (2013). SPARROW models used to understand nutrient sources in the Mississippi/Atchafalaya River Basin. *Journal of Environmental Quality*, 42(5), 1422-1440. doi:10.2134/jeq2013.02.0066.
- Ruddy, B.C., Lorenz, D.L., Mueller, D.K. (2006). County-Level Estimates of Nutrient Inputs to the Land Surface of the Conterminous United States, 1982–2001. U.S. Geological Survey Scientific Investigations Report 2006-5012, Reston, VA, pp. 17.
- Saad, D.A., Benoy, G.A., Robertson, D.M. (2018). Estimates of long-term mean-annual nutrient loads considered for use in SPARROW models of the Midcontinental region of Canada and the United States, 2002 base year. U.S. Geological Survey, Scientific Investigation Report 2018-5051, pp. 14.
- Schwarz, G.E., Alexander, R.B. (1995). Soils data for the Conterminous United States Derived from the NRCS State Soil Geographic (STATSGO) Data Base. [Original title: State Soil Geographic (STATSGO) Data

- Base for the Conterminous United States.] Edition: 1.1. (95-449). U.S. Geological Survey, url: [\[https://water.usgs.gov/lookup/getspatial?ussoils\]](https://water.usgs.gov/lookup/getspatial?ussoils)
- Schwarz, G.E., Hoos, A.B., Alexander, R.B., Smith, R.A. (2006). The SPARROW Surface Water-Quality Model: Theory, Application, and User Documentation. Techniques and Methods 6-B3, U.S. Geological Survey, pp. 248.
- Schwede, D.B., Dennis, R.L., Bitz, M.A. (2009). The watershed deposition tool: A tool for incorporating atmospheric deposition in water-quality analyses. *Journal of the American Water Resources Association*, 45(4), 973-985. doi:10.1111/j.1752-1688.2009.00340.x.
- Sugarbaker, L.J., Eldridge, D.F., Jason, A.L., Lukas, V., Saghy, D.L., Stoker, J.M., Thunen, D.R. (2016). Status of the 3D Elevation Program: U.S. Geological Survey Open-File Report 2016-1196. pp. 13. url: [\[http://dx.doi.org/10.3133/ofr20161196\]](http://dx.doi.org/10.3133/ofr20161196)
- Sugg, Z. (2007). *Assessing U.S. Farm Drainage: Can GIS Lead to Better Estimates of Subsurface Drainage Extent?* World Resources Institute, Washington D.C., accessed: January, 2018, url: [\[http://pdf.wri.org/assessing_farm_drainage.pdf\]](http://pdf.wri.org/assessing_farm_drainage.pdf)
- U.S. Army Corp of Engineers. (2016). National Inventory of Dams, 2016 NID database: U.S. Army Corp of Engineers.
- U.S. Department of Agriculture. (1994). State Soil Geographic (STATSGO) Data Base, Data use information. . (U.S. Department of Agriculture Miscellaneous Publication Number 1492).
- U.S. Department of Agriculture, Natural Resources Conservation Service. (2004). *Watershed Boundary Dataset*. accessed: January, 2016, url: [\[https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/water/watersheds/dataset/\]](https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/water/watersheds/dataset/)
- U.S. Environmental Protection Agency. (2018). *CMAQ: The Community Multiscale Air Quality Modeling System*. accessed: January 2, 2018, url: [\[https://www.epa.gov/cmqa\]](https://www.epa.gov/cmqa)
- U.S. Geological Survey. (1999). The National Hydrography Dataset (NHD) - Fact Sheet 106-99. url: [\[http://nhd.usgs.gov\]](http://nhd.usgs.gov)
- U.S. Geological Survey, U.S. Environmental Protection Agency. (2006). NHDPlus User Guide. pp. 115.
- Wieczorek, M.E., Jackson, S.E., Schwarz, G.E. (2016). Attributes for NHDPlus Version 2.1 Reach Catchments and Modified Routed Upstream Watersheds for the Conterminous United States: Estimates of Subsurface Tile Drainage Extent for the Conterminous United States, Early 1990s. accessed: January 16, 2018, url: [\[https://www.sciencebase.gov/catalog/item/58d9dec6e4b0543bf7fda70e\]](https://www.sciencebase.gov/catalog/item/58d9dec6e4b0543bf7fda70e)
- Wieczorek, M.E., LaMotte, A.E. (2010). *Attributes for NHDPlus Catchments (Version 1.1) for the Conterminous United States-STATSGO Soil Characteristics*. U.S. Geological Survey, Digital Data Series, DS-490-26, url: [\[https://water.usgs.gov/GIS/metadata/usgswrd/XML/nhd_statsgo.xml\]](https://water.usgs.gov/GIS/metadata/usgswrd/XML/nhd_statsgo.xml)
- Wolock, D.M. (1997). *STATSGO soil characteristics for the conterminous United States*. U.S. Geological Survey, Open-File Report 97-656, url: [\[https://water.usgs.gov/GIS/metadata/usgswrd/XML/muid.xml\]](https://water.usgs.gov/GIS/metadata/usgswrd/XML/muid.xml)
- Wolock, D.M., McCabe, G.J. (1999). ESTIMATES OF RUNOFF USING WATER-BALANCE AND ATMOSPHERIC GENERAL CIRCULATION MODELS¹. *JAWRA Journal of the American Water Resources Association*, 35(6), 1341-1350. doi:10.1111/j.1752-1688.1999.tb04219.x.
- Wolock, D.M., McCabe, G.J. (2018). *Water Balance Model Inputs and Outputs for the Conterminous United States, 1900-2015*. U.S. Geological Survey data release, url: [\[https://doi.org/10.5066/F71V5CWN\]](https://doi.org/10.5066/F71V5CWN)

Appendix A – Coordinate System

The Albers Equal-Area Projection was used as a standard coordinate system for the entire study. Table A-1 below describes the details of the projection.

Table A-1: Albers Equal-Area Projection used for the SPARROW Midcontinent model

Central Meridian	-96°
Standard Parallel 1	29.5°
Standard Parallel 2	45.5°
Latitude of Origin	23°
Linear Unit	m
Ellipsoid	NAD83

Appendix B – Land cover

Table B-1: Final harmonized land-cover categories compared to original U.S. dataset

U.S. Land-Cover Dataset		Harmonized Dataset	
<i>Index</i>	<i>Description</i>	<i>Index</i>	<i>Description</i>
11	Open Water	1	Open Water
12	Perennial Snow/Ice	8	Perennial Snow/Ice
21	Developed	2	Urban
22	Developed	2	Urban
23	Developed	2	Urban
24	Developed	2	Urban
31	Barren Land	3	Barren
32	Unconsolidated Shor	3	Barren
41	Forest	4	Forest
42	Forest	4	Forest
43	Forest	4	Forest
51	Shrub	5	Grassland\Shrub
52	Shrub	5	Grassland\Shrub
71	Grassland/Herbaceous	5	Grassland\Shrub
72	Sedge/Herbaceous	5	Grassland\Shrub
73	Lichens	3	Barren
74	Moss	3	Barren
81	Pasture/Hay	9	Pasture
82	Cultivated Crops	6	Agricultural Land
90	Wetland	7	Wetland
91	Wetland	7	Wetland
92	Wetland	7	Wetland
93	Wetland	7	Wetland
94	Wetland	7	Wetland
95	Wetland	7	Wetland
96	Wetland	7	Wetland
97	Wetland	7	Wetland
98	Wetland	7	Wetland
99	Wetland	7	Wetland

Table B-2: Final harmonized land-cover categories compared to original Canadian dataset

Canadian Land-Cover Dataset		Harmonized Dataset	
<i>Index</i>	<i>Description</i>	<i>Index</i>	<i>Description</i>
0	No Data	0	No Data
10	Unclassified	0	No Data
11	Cloud	0	No Data
12	Shadow	0	No Data
20	Water	1	Open Water
30	Barren/Non-vegetated	3	Barren
31	Snow/Ice	8	Perennial Snow/Ice
32	Rock/Rubble	3	Barren
33	Exposed land	3	Barren
34	Developed	2	Urban
35	Sparsely vegetated bedrock	3	Barren
36	Sparsely vegetated till-colluvium	3	Barren
37	Bare soil with cryptogam crust - frost boils	3	Barren
40	Bryoids	3	Barren
50	Shrubland	5	Grassland\Shrub
51	Shrub tall	5	Grassland\Shrub
52	Shrub low	6	Grassland\Shrub
53	Prostrate dwarf shrub	7	Grassland\Shrub
80	Wetland	7	Wetland
81	Wetland - Treed	7	Wetland
82	Wetland - Shrub	7	Wetland
83	Wetland - Herb	7	Wetland
100	Herb	3	Barren
101	Tussock graminoid tundra	3	Barren
102	Wet sedge	7	Wetland
103	Moist to dry non-tussock graminoid/dwarf shrub tundra	5	Grassland\Shrub
104	Dry graminoid prostrate dwarf shrub tundra	3	Barren
110	Grassland	5	Grassland\Shrub
120	Cultivated Agricultural Land	6	Agricultural Land
121	Annual Cropland	6	Agricultural Land
122	Perennial Cropland and Pasture	9	Pasture
200	Forest/Tree classes	4	Forest
210	Coniferous Forest	4	Forest
211	Coniferous Dense	4	Forest
212	Coniferous Open	4	Forest
213	Coniferous Sparse	4	Forest
220	Deciduous Forest	4	Forest
221	Broadleaf Dense	4	Forest
222	Broadleaf Open	4	Forest
223	Broadleaf Sparse	4	Forest
230	Mixed Forest	4	Forest
231	Mixed wood Dense	4	Forest
232	Mixed wood Open	4	Forest
233	Mixed wood Sparse	4	Forest

Appendix C – Information to Estimate Fertilizer and Manure Inputs

Fertilizer

The following calculation was performed to calculate the fertilizer inputs:

$$\text{Fertilizer input} = \frac{\text{Fertilizer purchase by area}}{\text{Fertilizer purchase by province}} \times \text{Fertilizer expenditure by province} \times 1000$$

where *Fertilizer input* is calculated in units of kg (of TN and TP), *Fertilizer purchase* were provided in units of \$, and *Fertilizer expenditure* were provided in units of metric tonne.

Fertilizer purchase data were obtained from AAFC (Agriculture and Agri-Food Canada, 2001; 2002; 2007). Fertilizer expenditure data were obtained from Fertilizer Canada (personal communication, Amanda Giamberardino, September 15, 2015) and shown in Table C-1 below. TN and TP were calculated by multiplying values of Table C-1 by conversion factors provided in Table C-2. Calculated fertilizer loadings, by subdivided area, were attributed to *Agricultural Land* cover classes prior to processing into SPARROW catchments.

Table C-1: Retail sales statistics for fertilizer year ending June 30th, 2002

Fertilizer Material	Ontario	Manitoba Metric tonnes	Saskatchewan
Urea	163,648	209,769	492,202
Ammonium Nitrate	8,423	72,322	123,863
Ammonium Sulphate	18,502	44,047	52,793
Anhydrous Ammonia	35,050	167,781	176,713
Nitrogen Solutions	87,238	63,570	111,290
Monoammonium Phosphate	89,325	185,887	368,943
Ammonium Polyphosphate	1,656	25,406	36,270

Table C-2: Conversion factors for nitrogen and phosphorus using known amounts of fertilizer material

Fertilizer Material	Nitrogen Conversion Factor	Phosphorus Conversion Factor
Urea	0.46	0
Ammonium Nitrate	0.34	0
Ammonium Sulphate	0.2	0
Anhydrous Ammonia	0.82	0
Nitrogen Solutions	0.28	0
Monoammonium Phosphate	0.11	0.52
Ammonium Polyphosphate	0.1	0.34

Manure

The final harmonized livestock types for the Midcontinent SPARROW models are provided in Table C-3. Animal delivery coefficients based on Ruddy *et al.* (2006) were used to calculate manure inputs given in Table C-4 (under U.S. column). These coefficients are typically used by U.S. scientists to calculate manure inputs. In Canada, animal delivery coefficients typically used by Canadian scientists to calculate manure inputs were based on data published by Agriculture and Agri-Food Canada (2001). These Canadian delivery coefficients are provided in Table C-4 (under Canadian column) as a reference, but were not used in the Midcontinent SPARROW model study for base year 2002. U.S. delivery coefficients were used for all areas.

Table C-3: Final harmonized livestock types compared to original Canadian and U.S. datasets

Canadian Data	U.S. Data	Selection for SPARROW
Beef Cattle	Beef Cattle	Beef Cattle
Milk Cows	Milk Cows	Milk Cows
Heifer	Other Cattle*	Heifer
Steers	Other Cattle*	Steers
Broilers	Broilers	Broilers
Laying Hens	Chickens	Laying Hens
Pullets	Pullets	Pullets
Turkeys	Total Turkeys	Turkeys
Boars, Hogs, Sows, Nursing and Weaner Pigs	Hogs	Hogs and Pigs
Sheep and Lambs	Sheep and Lambs	Sheep
Horses and Ponies	Horses and Ponies	Horses

* An assumption was made that the U.S. Other Cattle category was made up of 50% heifers and 50% steers.

Table C-4: Animal coefficients used by U.S. and Canadian scientists to calculate manure loads

Animal Category	Nitrogen Multiplier		Phosphorus Multiplier	
	kg multiplier / animal / year			
	U.S.	Canadian	U.S.	Canadian
Beef Cattle	54.75	78.81	19.35	21.32
Milk Cows	74.46	121.97	11.68	26.8
Heifer	51.47	52.19	6.57	14.12
Steers	54.75	56.29	17.52	15.23
Broilers	0.37	0.36	0.11	0.1
Laying Hens	0.55	0.55	0.22	0.2
Pullets	0.37	0.36	0.11	0.1
Turkeys (toms)*	0.72	1.54	0.27	0.57
Turkeys (hens)*	0.38		0.15	
Hogs and Pigs	9.86	9.93	4.38	3.31
Sheep	8.40	6.95	1.46	1.44
Horses	46.36	49.28	8.03	11.66

* An assumption was made that half of the Canadian turkey numbers were toms and half were hens.

Canadian livestock data from each sub-sub-drainage area (SSDAs) were obtained from the *Interpolated Census of Agriculture to Soil Landscapes, Ecological Frameworks, and Drainage Areas of Canada* (Agriculture and Agri-Food Canada, 2001). In some regions head counts are suppressed⁷ and for this study were substituted by CD data (Agriculture and Agri-Food Canada, 2007). Table C-5 illustrates how data replacement was managed in this study and the number of suppressed livestock categories. It should be noted that CD data are occasionally suppressed in some cases. SSDAs are spatial regions delineated by sub-watershed drainages (i.e., physiographic divides) while CDs are spatial regions delineated by provincial, regional or municipal divisions (i.e., political divides). Therefore SSDA and CD boundaries overlap, in other words multiple SSDAs may be found in one CD and vice versa.

Table C-5: Canadian livestock data replaced by missing information (areas where CDs replaced SSDAs)

Missing SSDAs		Replacing CDs	
<i>Identifier</i>	<i>Suppressed livestock categories</i>	<i>Identifier</i>	<i>Suppressed livestock categories</i>
02AA	8	3558	2
02AC	3	3558	2
02CC	4	3557	0
02CD	6	3557	0
02CE	3	3552	2
02CE	3	3553	0
02CE	3	3557	0
02CG	1	3551	0
02CH	11 (all)	3551	0
02CH	11 (all)	3552	2
02CH	11 (all)	3553	0
02CH	11 (all)	3557	0
02DB	6	3552	2
02DB	6	3553	0
02DC	5	3548	0
02EA	2	3549	1
02EB	2	3549	1
050A	2	4605	0
05PB	11 (all)	3559	2

⁷ Data suppression is governed by the Statistics Act when released data may divulge potentially identifiable information (Agriculture and Agri-Food Canada, 2001). With regards to the Interpolated Census of Agriculture, the data are suppressed when a limited number of farms are reported in a geographic unit or when up to two farms dominate the given information. For more information, please refer to The Interpolated Census of Agriculture (Agriculture and Agri-Food Canada, 2001).