MUDAC 2019 Problem 1

Background

It is often said that the quality of the water in a river is a reflection of its watershed. Each watershed – a land area where all of the water drains to a particular point on a stream or river – consists of a complex mix of geology, soils, landforms, topography, land cover (e.g. forest, prairie), land uses (e.g. farming, urbanization), etc.

Beginning in 2007, the Minnesota Pollution Control Agency and Minnesota Department of Natural Resources began a systematic long-term water quality monitoring program of rivers throughout Minnesota (Figure 1). This monitoring network tests for a variety of water quality indicators related to ecological health of the rivers and suitability for recreation. The monitoring also quantifies how much pollution is carried downstream to lakes, larger rivers, and even the Gulf of Mexico. The data collected through this network provide a unique opportunity to develop models that are applicable statewide, and perhaps to adjacent states.

Problem

Your goal in problem 1 is to predict water quality based on watershed characteristics.

- Develop a model, using whatever tools and techniques you want, that best predicts the responses of the water quality variables (TSS and nitrate) in relation to the watershed variables.
- Analyze and comment on whether the responses in your model are related to each other, and if so, how.
- Analyze and comment on whether the predictors in your model are related to each other, and if so, how.
- Analyze and comment on whether particular watersheds seem to be "outliers" (see below).

Data

All data is contained in the spreadsheet MUDAC_data_Problem1.xlsx. The spreadsheet contains 15 watershed variables and 2 water quality variables for each of 50 watersheds (Figure 2). These variables are described in Table 1. The data is averaged over several years.

Table 1

COLUMN	VARIABLE NAME	DESCRIPTION
WATERSHED VARIABLES		
B	FOREST[1]♦	% of watershed with forest cover
C	GRASS AND HAY[1]♦	% of watershed with grass or hay cover
D	WETLANDS[1]♦	% of watershed in wetlands
E	CROPLAND[1]♦	% of watershed that is cropped
F	DEVELOPED/URBAN[1]♦	% of watershed covered with houses, buildings, roads, parking lots, etc.
G	ALTERED STREAMS[2]	% of all streams in watershed that have been altered (e.g. straightened)
H	TILE-DRAINED LAND[3]	% of watershed that has artificial subsurface drainage. The term “tile” reflects that when this drainage practice first began, clay pipe was utilized.
I	SHALLOW BEDROCK UNDER CROPLAND[4]	% of watershed with bedrock (e.g. limestone, granite) that is not covered with much or any soil
J	SAND[5]★	Average sand % in watershed soils
K	SILT[5]★	Average silt % in watershed soils
L	CLAY[5]★	Average clay % in watershed soils
M	ORGANIC MATTER[5]	Average organic matter % in watershed soils
N	LAND SLOPE[6]	Average topographic relief (hilly-ness) of watershed
0	LAKES[1]♦	% of watershed covered by lakes
P	LAKE INTERCEPTION[7]	% of watershed land that drains to a lake that is at least 25 acres in size
WATER QUALITY VARIABLES		
Q	TOTAL SUSPENDED SOLIDS (TSS)[8]	The amount of suspended solids in the river expressed as a multi-year average concentration. Units are millgrams/liter. Sometimes referred to as suspended sediment.
R	NITRATE[8]	The amount of nitrate in the river expressed as a multi-year average concentration. Units are millgrams/liter. Nitrate is a dissolved form of nitrogen.

Superscript numbers reference data sources

♦ these variables are mutually exclusive; however will not add up to 100% as not all variables are given

★ these variables are mutually exclusive; however organic matter in soil is distinct from sand, silt, and clay

Analysis Considerations

- Environmental data is messy. Water quality in a river is governed by a complex interrelated set of watershed conditions. There are likely many variables beyond those provided that influence water quality. Watershed conditions ~ Water quality
- The watershed variables that explain total suspended solids (TSS) levels may or may not be the same as those that explain nitrate. Analyze separately
- There is value in understanding if certain watershed variables explain both TSS and nitrate levels. Addressing these variables through proper land management activities might provide a "two-fer" in terms of water quality benefits. Find both significant variable
- Some of the watershed variables may be highly correlated with each other. High correlation problem
- Minnesota is a state with a very diverse landscape (Figure 3). Most but not all of the watershed variables have a relationship to that spatial diversity.
- There may be some "outlier" watersheds that don't fit overall patterns very well. There is not a problem with the data for those watersheds. Rather, there may be some unique conditions not fully accounted for in the variables provided. not forest & cropland?

Data Sources

1- National Land Cover Database (NLCD), Land Cover 2011. Multi-Resolution Land Characteristics Consortium. Available online at https://www.mrlc.gov/

2- Minnesota Pollution Control Agency led project described here: https://www.pca.state.mn.us/water/minnesota-statewide-altered-watercourse-project.

3- Derived from 1,5, and 6

4- Derived from Minnesota Geological Survey. https://mngs-umn.opendata.arcgis.com/pages/spatial-datasets, and 1.

5- Gridded Soil Survey Geographic (gSSURGO) Database & the U.S. General Soil Map (STATSGO2) for Minnesota. United States Department of Agriculture, Natural Resources Conservation Service. Available online at https://datagateway.nrcs.usda.gov/. 2015 official release.

6- Minnesota LiDAR Derived Digital Elevation Model (DEM). Various sources, data made available by Minnesota Geospatial Information Office. Available online at *https://gisdata.mn.gov/*. 2006-2012.

7- National Hydrography Dataset: https://www.usgs.gov/core-science-systems/ngp/national-hydrography

8- Minnesota Pollution Control Agency. https://www.pca.state.mn.us/wplmn/data-viewer.

Figure 1

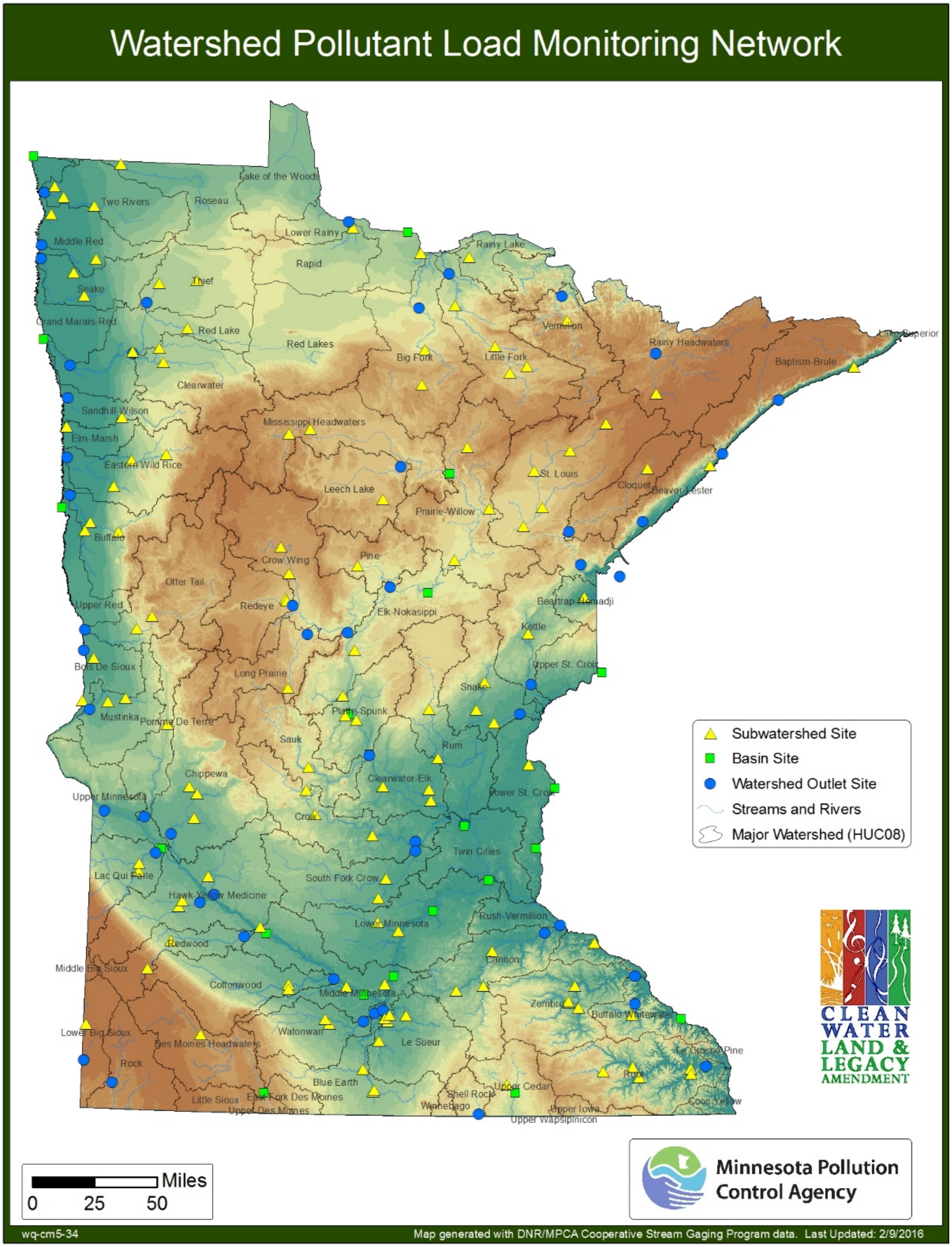

Watershed Pollutant Load Monitoring Network
Subwatershed Site
Basin Site
Watershed Outlet Site
Streams and Rivers
Major Watershed (HUC08)
Miles
0
25
50
CLEAN WATER LAND & LEGACY AMENDMENT
Minnesota Pollution Control Agency
wq-cm5-34
Map generated with DNR/MPCA Cooperative Stream Gaging Program data. Last Updated: 2/9/2016
Lake of the Woods
Two Rivers
Roseau
Lower Rainy
Rainy Lake
Middle Red
Rapid
Thief
Snake
Grand Marais-Red
Red Lake
Red Lakes
Big Fork
Little Fork
Vermilion
Rainy Headwaters
Baptism-Brule
Clearwater
Sandhill-Wilson
Mississippi Headwaters
Elm-Marsh
Eastern Wild Rice
St. Louis
Cloquet
Leech Lake
Prairie-Willow
Buffalo
Crow Wing
Pine
Otter Tail
Upper Red
Redeye
Elk-Nokasippi
Kettle
Bois De Sioux
Long Prairie
Upper St. Croix
Snake
Mustinka
Platte-Spunk
Sauk
Rum
Chippewa
Clearwater-Elk
Crow
Twin Cities
Lac Qui Parle
South Fork Crow
Hawk-Yellow Medicine
Lower Minnesota
Rush-Vermillion
Redwood
Cannon
Middle Big Sioux
Cottonwood
Middle Minnesota
Zumbro
Buffalo-Whitewater
Lower Big Sioux
Watonwan
Le Sueur
Des Moines Headwaters
Rock
Blue Earth
Upper Cedar
Shell Rock
Little Sioux
Winnebago
Upper Iowa
Upper Wapsipinicon

Figure 2

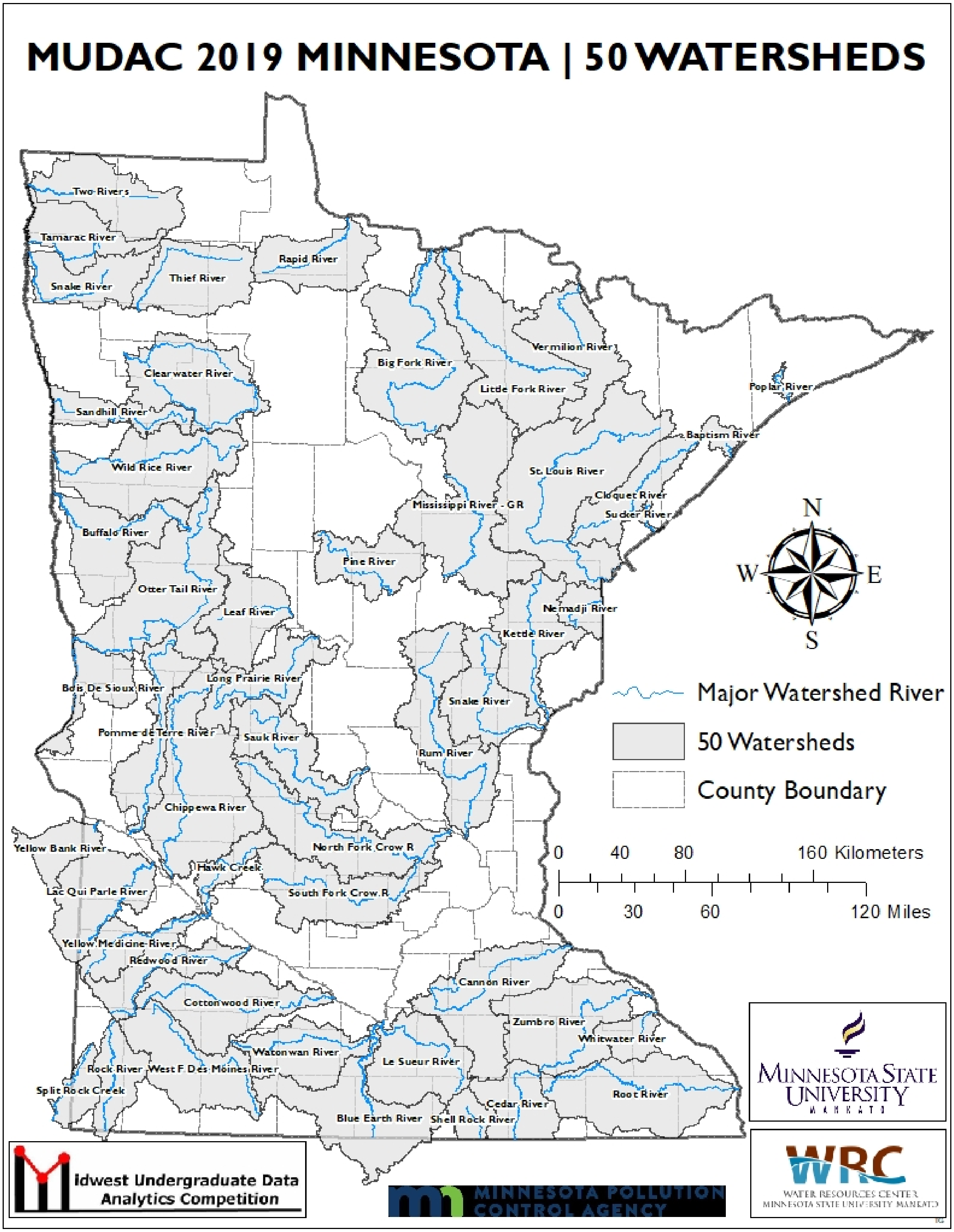
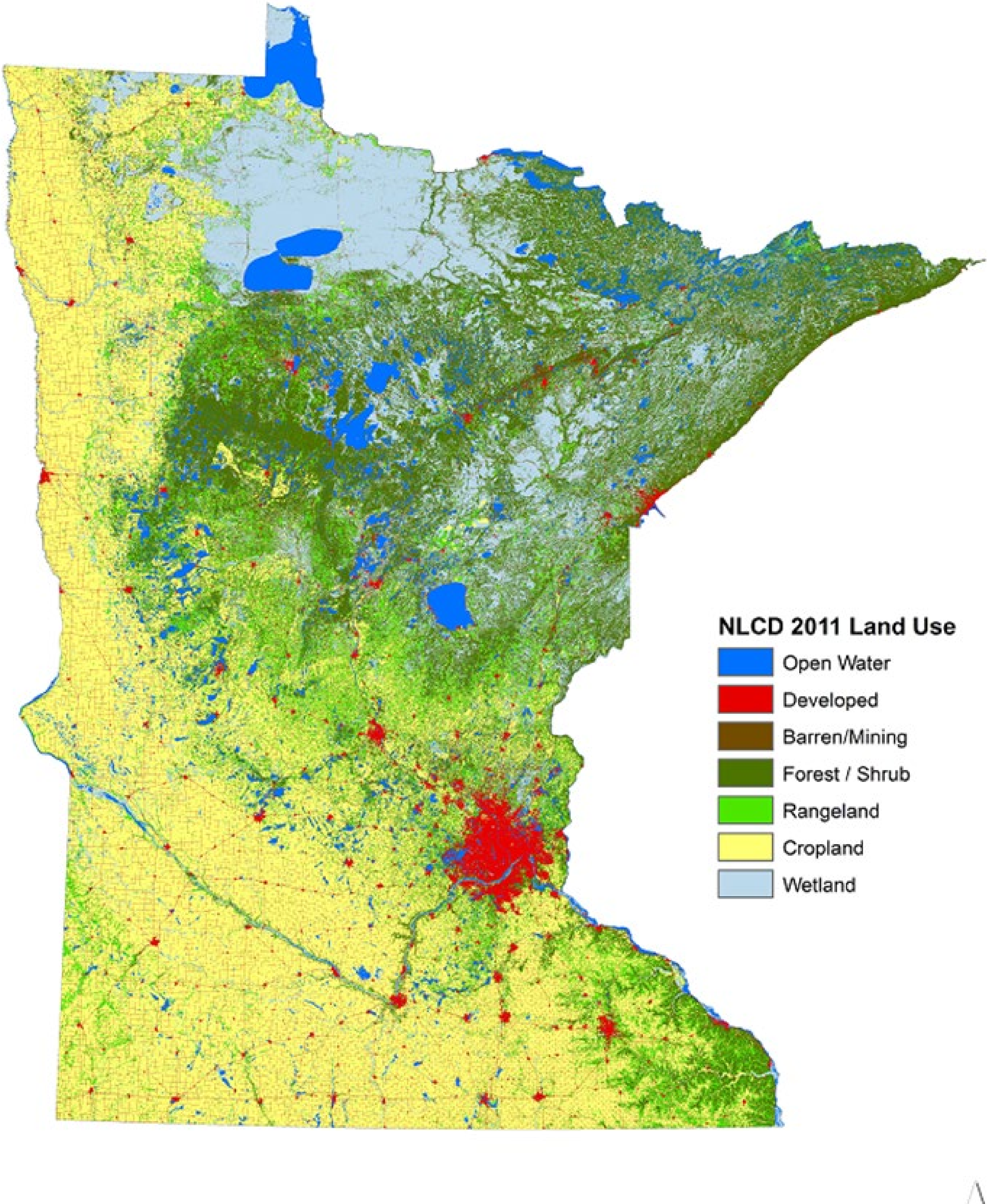

MUDAC 2019 MINNESOTA | 50 WATERSHEDS
Two Rivers
Tamarac River
Snake River
Thief River
Rapid River
Big Fork River
Vermilion River
Little Fork River
Poplar River
Baptism River
Clearwater River
Sandhill River
Wild Rice River
St. Louis River
Cloquet River
Sucker River
Mississippi River - GR
Buffalo River
Pine River
Otter Tail River
Leaf River
Nemadji River
Kettle River
Long Prairie River
Bois De Sioux River
Snake River
Pomme de Terre River
Sauk River
Rum River
Chippewa River
Yellow Bank River
Hawk Creek
North Fork Crow R
Lac Qui Parle River
South Fork Crow R
Yellow Medicine River
Redwood River
Cannon River
Cottonwood River
Zumbro River
Whitewater River
Watonwan River
Le Sueur River
Rock River
West F Des Moines River
Split Rock Creek
Root River
Cedar River
Blue Earth River
Shell Rock River
N
S
W
E
Major Watershed River
50 Watersheds
County Boundary
0 40 80 160 Kilometers
0 30 60 120 Miles
Minnesota State University Mankato
WRC
Water Resources Center
Minnesota State University Mankato
Midwest Undergraduate Data Analytics Competition
Minnesota Pollution Control Agency

Figure 3


NLCD 2011 Land Use
Open Water
Developed
Barren/Mining
Forest / Shrub
Rangeland
Cropland
Wetland