DEPARTMENT OF NATURAL RESOURCES

Transient NMLG Model Analysis and Revisions

August 2018

Contents

Tables

Figures

1. Introduction

As described in the model report (SSPA, 2017), the transient version of the NMLG model was developed to address the needs of the MNDNR that cannot be addressed using the steady-state NMLG model developed by the U.S. Geological Survey (Jones et al., 2017; Trost et al., 2017). Transient simulations require storage parameters and transient input data sets not used in steady-state simulations. In addition to the expanded input, SSPA made changes to model parameters and processes to make transient modeling tractable and to improve the fit of computed output to transient hydrological data. SSPA (2017) discuss considerations for the transition to a transient model in Sections 3.2.3, 3.3.4, and Section 4 of their report. Practical limits on computation time were an important consideration for the NMLG model, which is complex and can become numerically unstable.

The transient model, initial analyses, and report were developed on an accelerated schedule. The compressed schedule limited the amount of model review, exploration and conceptual-model testing that were undertaken. Following completion of the initial transient model, DNR staff conducted analyses to better understand the behavior of the model; to explore the sensitivity of results to selected modeling assumptions and choices; and to provide some initial analyses of the effects of model uncertainties and limitations on model predictions. These evaluations are intended to provide additional, useful information when applying the model to support management decisions. To incorporate new and updated data and information from the model analyses, DNR staff revised the transient NMLG model. The model is initially being used to support assessments of the effects of groundwater withdrawals on water levels in White Bear Lake, and, therefore, these evaluations focus on White Bear Lake.

This report is divided into three main sections. Section 2 reviews important model elements and input data against the conceptual model of the system and previous studies. Exploration of uncertainties in groundwater recharge and other water-balance components were areas of particular focus. Section 3 explores how selected input data and other features that were not adjusted during model parameter estimation affected model results. Section 4 describes development of and results from a revised version of the transient. A key change made for the revised model was re-parameterization of the Soil Water Balance (SWB) model used to compute groundwater recharge and surface runoff to lakes.

2. Review of Model Elements and Data

The spatial and temporal distribution of recharge is the foundation of the water balance in the groundwater-flow model. In their Discussion and Recommendations, SSPA (2017, Section 10) describe some model structure and data deficiencies that limit the ability of the SWB model to accurately compute recharge and runoff in the northeastern Metro area. The scale and complexity of the system are difficult challenges for any model, and DNR do not intend to replace the SWB model in the near term. Computed groundwater recharge directly affects computed groundwater and lake levels and discharges. In addition, hydraulic conductivities are highly correlated to groundwater recharge in the parameter estimation process. Section 2.1 assesses information on groundwater recharge and other water balance components relevant to understanding model results.

Some hydrogeological property values applied in the transient NMLG model are outside of the expected ranges for particular material types. Additionally, the relative values among some contrasting material types were inconsistent. Horizontal hydraulic conductivities of some unconsolidated, Quaternary material types and of the Prairie du Chien aquifer differ substantially from the values applied in Metro Model 3 (Metropolitan Council, 2014) and the USGS steady-state NMLG model (Jones et al., 2017). Uncertainties and limitations of the SWB model and limited available information on the system water balance were important drivers of the differences in estimated hydraulic conductivities. Section 2.2 evaluates hydraulic property information and uncertainties.

Section 2.3 discusses changes to the representation of some surface-water features using more accurate/updated data that could improve the model.

2.1. Recharge and Water Balance

The spatial and temporal distribution of recharge is the foundation of the water balance in the groundwater-flow model. Except at the point scale, recharge cannot be directly measured. Most recharge in the model derives from the potential recharge output from the SWB model. Many of the lakes and wetlands represented by the River (RIV) and Lake (LAK) packages also act as net sources of groundwater recharge. The SWB model was also used to calculate runoff to lakes in the NMLG model, which has also not been directly measured.

This report provides evaluation of available information on groundwater recharge and discharge and surface runoff. The evaluation focused on: streamflow and stream base flow, previous recharge estimates for the NMLG area, and on components of the water balance for White Bear Lake.

Long-term, net recharge in a watershed (recharge minus groundwater evapotranspiration and other consumptive groundwater uses) is often approximated by base flow estimated from streamflow records. Available data and previous analyses provide some constraints on groundwater recharge, but there remains significant uncertainty in both average and time-varying groundwater flow rates at the regional scale in the NMLG model area. The spatial and temporal distribution of recharge at local scales is also important, affecting spatial and temporal variations in groundwater heads and discharges. SSPA (2017) described inconsistencies in the satellite-based National Land Cover data sets and some of the limitations of the SWB model that affect the calculation of spatial and temporal variations in recharge. The effects of recharge representation in the transient NMLG model were considered both in terms of the regional water balance and local effects.

The SWB model was also used to calculate runoff to lakes in the NMLG model. Seasonal and annual variations in runoff can be important to lake water budgets including lakes with small contributing watersheds and limited surface inflow and outflow. Although runoff to the simulated lakes is relatively small compared to direct precipitation on the lakes, runoff may be more significant when compared to net precipitation minus evaporation.

To gain further insight into the hydrological behavior of White Bear Lake in particular, DNR staff evaluated components of the lake budget. Although there are uncertainties in direct precipitation, evaporation, and surface outflow volumes, these components can be more directly estimated from measured data. The effects of total groundwater inflows and outflows and surface runoff into the lake cannot be estimated directly from measured data, but the water-balance analysis provides some insights into these components.

2.1.1. Streamflow and Base Flow

In the northeastern Twin Cities region, available streamflow data provide limited information relevant to estimating large-scale groundwater discharge and corresponding recharge. There are always unavoidable uncertainties and assumptions that affect estimation of base flow and recharge from stream gaging records. A particular limitation on assessing the large-scale groundwater budget in the northeastern Twin Cities region is that most of the groundwater discharge cannot be estimated from streamflow data. Although groundwater discharge supplies base flow to tributary streams, much of the groundwater discharges to the Mississippi and St. Croix Rivers either directly or in seepage faces, springs, and wetlands in the river valleys.

Because the flow in these rivers upstream from the model domain is large, changes in stream flow due to groundwater inflows and outflows within the model domain are relatively small and typically within the range of stream-flow measurement errors. As a result, groundwater inflows within the model domain cannot be reliably estimated from streamflow records. For example, calculated net flow to the Mississippi River through the Twin Cities is sometimes negative, but it is unlikely that there is actually a net loss during those periods. The anomaly is most likely caused by measurement error. This limitation has been recognized in previous modeling studies in the region (Metropolitan Council, 2014).

Despite the limitations, comparing simulated discharge to measured streamflow provides a possible check on the simulated groundwater budget. Streamflow records for three streams were considered during development of the steady-state NMLG model (Jones et al., 2017): Rice Creek (a tributary to the Mississippi River) and Brown's Creek and Valley Creek (tributaries to the St. Croix River). These data sources were further evaluated to better understand the information they can provide for the regional groundwater model.

Different hydrograph separation methods produce a wide range of base flow estimates for most streams except for those with very limited quick flow or storm flow. In addition, there are characteristics specific to these streams that limit their use to assess overall groundwater recharge represented in the NMLG model. Rice Creek is not well suited to hydrograph separation techniques, but the data from Brown's Creek and Valley Creek do provide information on the local-scale performance of the model.

Groundwater-derived base flow cannot readily be separated from total streamflow in Rice Creek using hydrograph separation. Rice Creek flows through multiple lakes, and its tributaries receive discharges from lakes and wetlands. There is a concrete dam at the outlet of Peltier Lake on Rice Creek, and several tributary lakes have control structures at their outlets such as the dam at Bald Eagle Lake. Runoff discharging to these lakes is slowly released at the outlets, and hydrograph separation methods such as the local-minimum method used by Jones et al. (2017) do not generally distinguish between hydrograph shapes caused by lake storage and release versus groundwater discharge.

Some shallow groundwater flows into and out of the Rice Creek watershed, and a portion of groundwater recharged in the watershed exits the watershed via percolation to and flow through buried Quaternary and bedrock aquifers. For example, Jones et al. (2013) showed that White Bear Lake recharges groundwater that exits the Rice Creek watershed to the south and southwest of the lake. Groundwater evapotranspiration in wetlands and other areas with a shallow water table is expected to be a significant component of groundwater discharge in this watershed. There is also groundwater use within the watershed and use outside of the watershed that affects groundwater flow within the watershed.

Jones et al. (2017) applied a multiplier of 0.7, based on professional judgement, to local-minimum separation outputs from Rice Creek during parameter fitting of the SWB model. Considering that there are groundwater outflows not reflected in the Rice Creek streamflow, total groundwater recharge is greater than the groundwater component of streamflow for Rice Creek, but the groundwater component of streamflow is not well constrained (i.e. the assumptions for applying hydrograph separation methods are not met). For these reasons, base-flow estimates for Rice Creek were not considered further.

Almendinger (2003) estimated that base flow made up 87 percent and 92 percent of the total stream flow in Browns Creek and Valley Creek in 1998 and 1999 respectively. Only 4 and 3 percent of flows were identified as “quick flow” with the remainder described as “intermediate flow”. Some base flow and “intermediate flow” derives from lake outflows incorporating groundwater and surface-water inflows to tributary lakes, but groundwater discharge is the dominant component of streamflow. Computed groundwater discharges to these streams as computed using the River package in the groundwater-flow model can be compared to base-flow estimates.

DNR obtained the daily flow data for 2003 through 2013 that the USGS compiled from the source agencies (P. Jones, pers. comm., 2017). The average streamflow at Browns Creek and Valley Creek was 7.70 and 15.6 cfs respectively. Average groundwater discharge (base flow) is slightly less than average streamflow, but the difference may be within the range of measurement errors for these streams.

Jones et al. (2017) compare base flows estimated from gaging records to model-computed groundwater discharges in their Table 5, but the values of both the estimated and computed base flows in that table are in error. For example, the median “measured net base flow” at the Valley Creek gage is listed as 1.25 million cubic feet per second (Mft^3/s). The computed base flow is listed as being equal to the “measured net base flow” for Browns Creek and Valley Creek.

To check the simulated base flow in the steady-state NMLG model, DNR staff extracted discharge to River package reaches representing these streams using the USGS program Zone Budget version 3.01 (Harbaugh, 2008). The computed net flow to these River reaches is 3.2 cfs for Browns Creek and 3.5 cfs for Valley Creek.

Surface leakage computed by the UZF package is not included in these totals. Surface leakage represents a combination of rejected recharge (i.e. surface runoff) and groundwater discharge. Computed surface leakage near Browns Creek was 2.8 cfs (for a total computed discharge of 6 cfs), and computed surface leakage near Valley Creek was less than 0.1 cfs. Computed River plus surface-leakage discharge to Browns Creek was 77 percent of gaged streamflow, but computed discharge to Valley Creek was 23 percent of gaged streamflow.

For the transient, annual NMLG model, the computed River discharge for 2003 through 2013 averaged 3.0 cfs for Browns Creek and 2.2 cfs for Valley Creek. The UZF package was not used in the transient model, and, therefore, there was no surface leakage. For Valley Creek, to which surface leakage was minimal in the steady-state model, the computed discharge in the transient, annual model was 63 percent of the computed discharge in the original steady-state model. Computed discharge to Valley Creek was much lower than gaged flows for both models.

The computed base flows to these tributary streams does not necessarily reflect groundwater recharge in the larger model domain due to the limited geographic areas that influence base flow to these streams. Local-scale errors in hydrogeological properties could cause errors in the partitioning of groundwater discharge between these tributary streams and the St. Croix River Valley, even if the total discharge is approximately correct.

There are large horizontal and vertical groundwater gradients near these streams, especially in the lower reaches, which are not generally well represented in the regional-scale model. For example, Almendinger (2003) estimated that the north branch of Valley Creek supplied about 40% of the total base flow in 1999, despite finding that the north branch was perched at one piezometer site near the mid-section of the stream. Based on the limited piezometer evidence, Almendinger speculated that much of the base flow to the north branch may be supplied chiefly from Metcalf Marsh and Lake Edith at the headwaters. This could only be confirmed with more data, such as collecting a streamflow measurement at multiple points along the length of the stream. In the steady-state and transient NMLG models, Lake Edith is perched above the water table. This may be because spatially complex hydrogeological properties are not well represented locally, and relatively small head errors in the context of the large drop in head toward the river make the difference between simulating groundwater discharge versus perched conditions for this feature.

2.1.2. Recharge Estimates

Jones et al. (2017) noted that previous studies had estimated a wide range of average recharge rates in the Twin Cities area, but average recharge rates calculated in some earlier studies (Barr Engineering Co. and Washington Co., 2005; Metropolitan Council, 2014; Lorenz and Delin, 2007) were higher than the average recharge of 4.6 inches per year (in/yr) derived from the SWB model for the NMLG model. Jones et al. (2017) multiplied the SWB computed recharge by 1.75, a factor that was estimated as an adjustable groundwater-flow model parameter. The resulting infiltration recharge applied in the groundwater flow model was 6.45 in/yr spread over the active model domain. The 6.45 in/year total does not include recharge from surface-water bodies represented explicitly by the River and Lake packages and also does not exclude infiltration rejected by the UZF package where surface leakage was computed.

Infiltration recharge computed for a “typical” year (1979) for the southern Washington County model had an aerial average of 8.5 inches per year, including negative values where evapotranspiration (ET) from a specified water table exceeded infiltration (Barr Engineering Co. and Washington Co., 2005). The domain for this model overlapped with but did not include the entire NMLG model domain.

To compare data from other studies, recharge estimates had to be extracted for just the NMLG model area. DNR extracted average recharge rates for the Minnesota portion of the active NMLG model domain from several data sets. Within this subarea, the average recharge calculated by the Metro Model 3 SWB model for non-open water cells for 2003-2011 was 8.4 in/yr (this value would be smaller if averaged over all cells). A state-wide SWB model on a one kilometer grid (Smith and Westenbroek, 2015) calculated average recharge for the period 1996-2010 of 6.3 in/yr for non-open water cells in the area. The Regional Regression Recharge model (Lorenz and Delin, 2007) calculated an average recharge for the period 1971-2000 of 7.6 in/yr for the area, excluding large lakes (Forest, Big Marine, and White Bear). Small water bodes were lumped in with land areas in the latter model.

To compare with the steady-state NMLG model, the total recharge applied through the Recharge package to the transient NMLG model for the period 2003-2013 was extracted from the model output. This total averaged 4.78 in/yr spread over the entire active model domain. Rivers and streams represented by the River package were entered such that they could only serve as discharge features. Some lakes act primarily as groundwater discharge features and should not be considered when assessing total recharge. Other lakes act as recharge features or flow through features with net recharge. The leakage out of minus the leakage into these latter features is their net effective recharge added to the groundwater model. These different types of features are not split out in the total water budget calculated by MODFLOW. Parsing out just the net recharge from recharge and flow through lakes would require a substantial processing effort. The total inflow to groundwater from all River cells in the transient model was about 60 percent of the total in the original steady-state model.

The overall recharge applied in the transient model developed by SSPA (2017) was significantly less than recharge applied in other recent groundwater models for the area. The available base-flow estimates provide limited information about regional-scale recharge rates, but computed base-flow discharge to Browns Creek and Valley Creek in the transient model were much less than gaged streamflow. These comparisons suggest that overall recharge rates may be underestimated in the transient model.

Applying lower recharge rates requires lowering the system transmissivity in order to produce similar groundwater heads in the model. During parameter estimation, some horizontal hydraulic conductivity values, particularly for the Prairie du Chien aquifer, were substantially reduced from the values used in Metro Model 3 and the steady-state NMLG model. Uncertainties in recharge, therefore, directly affected the estimated hydraulic properties.

2.1.3. Precipitation

The climate inputs to the SWB model include daily grids of precipitation and maximum and minimum temperatures. Spatial variation in temperatures tends to be smoother than precipitation amounts. Precipitation is a key driver of temporal and spatial variability in computed recharge and is the largest inflow to lakes with limited surface inflows. Errors in precipitation inputs result from variable catch at precipitation gages, variable spatial coverage of precipitation gages, and interpolation/modeling of the distribution of precipitation from gage data. Although the temporal and spatial distribution of precipitation is better constrained than most other model inputs, precipitation is the ultimate hydrologic driver of the system. Exploring precipitation errors and uncertainties can shed light on model behavior and limitations.

Jones et al. (2017) selected the Daymet data set of climate data at daily intervals on a 1-km grid, which is available from 1980 through 2016 (Thornton et al., 1997; Thornton et al., 2017). Daymet data are available as 2 degree X 2 degree tile subsets in the efficient netCDF format that can be read directly by SWB. Therefore, it provided a readily available data set that could be used directly without further processing.

The data sources and methods used to generate the Daymet precipitation grids affect accuracy and how the data vary over time. The grids are generated from station data from the Global Historical Climatology Network (GHCN) using an elevation-adjusted, spatial weighting method (Thornton et al., 1997). In the NMLG model area, the GHCN network includes daily National Weather Service Cooperative (NWS Coop) data and hourly Automated Surface Observing System (ASOS) data from airports. These stations are relatively sparse, and few stations were in place at the same location for all or most of the 1980-2016 period of the transient model. Missing data thresholds are used to remove stations from the Daymet interpolation process in individual years.

DNR staff compared Daymet to other precipitation data at four NWS Coop station locations and at the center of White Bear Lake as a check on the Daymet data. The stations were: Minneapolis St Paul Intl AP (215435), Vadnais Lake (218477), Stillwater 1 SE (218037), and Forest Lake 5NE (212881). Data for these locations was extracted using the Single Pixel Extraction Tool on the Daymet website. Both monthly time series and the 1981-2010 averages (climate normal) were considered in the comparisons. The data sets used for comparison were: station data with filled missing values, monthly interpolations from the Minnesota HIDEN network on a 10-km grid (Minnesota State Climatology Office, 2017), and the PRISM product of monthly data interpolated on a 4-km grid (PRISM Climate Group; Daly et al., 2004).

The HIDEN data were derived from a “network of networks” including NWS Coop, Soil and Water Conservation District (SWCD), Metropolitan Mosquito Control District, and backyard volunteer network gages. Some of the stations used to produce the monthly HIDEN grids change more frequently and tend to have more missing data, but the HIDEN network has a much higher density of stations than the GHCN, especially in more densely populated parts of the Twin Cities metropolitan area. PRISM uses similar inputs to Daymet before 2002, but a different interpolation method. PRISM also incorporates Radar-based precipitation starting in 2002. Missing months in the NWS Coop station data were filled with the nearest available individual station from the Minnesota HIDEN network to generate complete, monthly time series at the station locations. This approach for the individual station data was selected to reflect the station data as closely as possible while providing a complete record.

The climate normal values provide a quick comparison of long-term average differences between the data sets. For the stations with complete or nearly complete data, the climate normal values from the National Centers for Environmental Information (NCEI) were also compared. These data are shown in Table 1. The Daymet normal values are larger than the other data sets at four of the five locations. Daymet is smaller than the other data sets at Stillwater 1 SE. Note that open exposure/wind results in under catch at the MSP airport, and precipitation measured at this station tends to be biased low (Peter Boulay, pers. comm.).

Table 1 – Climate normal (1981-2010 average) annual precipitation (inches) calculated from different data sets at three precipitation stations and at White Bear Lake

Station / Location	Daymet	Station (Filled)	Gridded HIDEN	NCEI	PRISM
Minneapolis St Paul Intl AP (215435)	32.79	30.57	32.11	30.57	31.97
Vadnais Lake (218477)	34.09	31.46	32.45	32.04	31.88
Stillwater 1SE (218037)	32.18	33.42	33.37	--	33.44
Forest Lake 5NE (212881)*	33.41	31.14	31.02	32.05	30.80
White Bear Lake	35.16	--	32.91	--	32.61

* This station was established November 1986; an earlier station with the same NWS ID but a different name was 14 km to south.

The State Climatology Office produced maps of climate normal precipitation using the gridded HIDEN data. Annual precipitation (1981-2010) was generally between 32 and 33 inches in the central and east-central parts of the Metro area, including four of the locations considered here. Annual precipitation was between 31 and 32 inches in the surrounding area that makes up most of the remainder of the NMLG model domain (including the Forest Lake 5NE station in Wyoming Township, Chisago County). The average Daymet precipitation for the entire SWB model domain for the 1981-2010 period was 33.98 inches per year, roughly 1.5 to 2 inches (5 to 6 percent) higher than expected.

Long-term model results can be affected by a relatively small but consistent precipitation bias. For example, increasing or decreasing the precipitation by 5 percent in the steady-state NMLG model changed computed lake levels by two to three feet (Jones et al, 2017). A consistent precipitation bias would have a lessor effect on assessing the effects of other changes to the system (e.g. groundwater pumping) because the precipitation bias would be common to model runs for such alternative scenarios.

Double mass plots were used to further evaluate the precipitation data at White Bear Lake. Double mass plots are a standard tool used to check data consistency at a gage by comparing cumulative time series at one station to a reference cumulative time series based on an average of multiple gages (e.g. Searcy and Hardison, 1960; Mizukami and Smith, 2012). Because the reference series is an average, inconsistencies in individual gages as well as the effects of local precipitation variability tend to smooth out. In addition to checking for possible persistent changes caused by station changes, the plots also are a visual indicator of shorter-term variability in the station data relative to the surrounding region.

Rather than generating a new reference series from gage data, the climate division data set for East-Central Minnesota (Climate Division 6) was used as the regional average data set for comparison (Vose et al., 2014a, 2014b). This data set is also based on GHCN data interpolated to a grid (Vose et al., 2014a). The individual “station” data compared against the division average were actually monthly values from the gridded datasets at the location of White Bear Lake: Daymet, HIDEN grid, and PRISM. The plots based on the period 1980 through 2016 are shown in Figure 1.

The plots all have a positive slope indicating that precipitation at White Bear Lake is generally higher than the average for the climate division. This is consistent with the general spatial trend of precipitation across the region as shown in normal precipitation maps (State Climatology Office). The plots generally fluctuate around a straight line with the highest slope for the Daymet data. Deviations from a straight line reflect real local variations in precipitation relative to the regional average and/or temporary biases caused by other factors such as periods of missing station data.

To better visualize these temporal variations, a straight line with intercept at zero was fitted to each double-mass curve. The differences between the double-mass curves and the corresponding straight lines are plotted versus time in Figure 2. These plots show multi-year tendencies in which the ratio between a “station” and regional precipitation appears to change. For example, the double-mass curves fall increasingly below the straight line from 1980 through about 1990 (increasingly negative values in Figure 2).

The variations over time appear to be similar for all three data sets at White Bear Lake, but there are notable differences. For example, the Daymet data has a much larger positive deviation in 2015 and 2016 than the gridded HIDEN data. These differences have not been tested for statistical significance, but the NMLG model is deterministic and responds to absolute inputs. The computed White Bear Lake stages have a lower range than measured stages--generally too little rise during periods of increasing stages and too little decline during periods with falling stages. In 2015 and 2016, however, the model-computed rise in stage was substantially larger than the observed rise. The high Daymet anomaly compared to HIDEN over that period suggests that a higher than average positive bias in Daymet precipitation at White Bear Lake in 2015 and 2016 could have been a significant factor in the divergent model behavior during that period.

2.1.4. Lake Evaporation

Evaporation is the largest outflow from closed-basin lakes or other lakes with minimal surface-water inflow and outflow, but lake evaporation is typically not well constrained by data. Lake evaporation is rarely measured because effective methods are expensive and difficult to employ. In most cases evaporation must be estimated using indirect and approximate models based on measured meteorological data and/or pan evaporation. These approaches are uncertain and can lead to a wide range of estimates. The eddy correlation (EC) method for determining evaporation is direct and is generally considered the most accurate method for measuring evaporation. The EC method uses expensive equipment deployed in the lake and requires complex data processing and analysis.

SSPA (2017) estimated lake evaporation as a function of pan evaporation measured at the University of Minnesota, St. Paul campus. This evaporation computation method was calibrated to preliminary evaporation estimates developed by University of Minnesota researchers using EC systems deployed in White Bear Lake during 2014 (late July through early November), 2015 (early May through October), and 2016 (late March through November). Applying a pan coefficient of 0.60 along with a storage coefficient to represent the seasonal heat storage in the lake resulted in a reasonably good fit to the preliminary EC data. These estimates did not include evaporation and sublimation during the winter months of December through March, which were assumed to be minimal.

Updated EC Observations and Lake Evaporation Modeling at White Bear Lake

Xiao et al. (2018) reported in late March 2018 on the final EC evaporation values and on simulations with a physically-based lake model (CLM4-LISSS) to extend the evaporation estimates to unmeasured periods during ice-free seasons. Xiao et al. (2018) also made EC measurements from late January through early February 2017 to gage the rate of sublimation from ice and snow. Latent heat flux, from which evaporation can be directly calculated, was measured over 30-minute intervals with occasional gaps (about 20%) filled by interpolation and empirical fitting functions. The DNR obtained the EC-derived evaporation and sensible heat flux, meteorological data, and CLM4-LISSS input data and computed evaporation data in May 2018 (Ke Xiao, pers. comm.)

Table 2 lists the annual total evaporation for 2014 through 2016 for three estimates: the transient NMLG model (SSPA); the preliminary EC data with empirical extrapolation to unmeasured periods (EC preliminary); and the total combining the final EC observations, CLM4-LISSS model extrapolation, and typical winter ice/snow sublimation rate based on EC measurements in January-February 2017 (EC + CLM4-LISSS). Annual pan evaporation is also listed for comparison. Xiao et al. (2018) reported four different annual evaporation totals for each year based on different methods for extrapolating to unmeasured periods. The values shown in Table 2 differ from these reported totals because they include estimated winter sublimation and used modeled evaporation to extrapolate only during observed ice-free periods (rather than modeled ice-free periods). One of the authors of the report concurred that this was an appropriate option for combining observed and modeled evaporation to extend to the full year (pers. comm., Timothy Griffis, April 2018).

Table 2 – Annual total evaporation from White Bear Lake from different methods and data sets in millimeters (inches).

Year	2014	2015	2016
SSPA	518 (20.4)	512 (20.2)	527 (20.7)
EC Preliminary	--	502 (19.8)	588 (23.1)
UMN pan (April 21 through Oct 10)	864 (34.0)	853 (33.6)	877 (34.5)
EC + CLM4-LISSS	576 (22.7)	751 (29.5)	771 (30.4)

The EC + CLM4-LISSS totals are higher than the estimates based on the preliminary EC data (EC preliminary and SSPA). This represents a 35 percent increase over the SSPA values for these three years. The difference in totals is caused by higher evaporation rates during the measured periods not just the model-extrapolated periods or the added winter sublimation (i.e. the interpretation of the raw data changed significantly). SSPA did not include an estimate for evaporation during the parts of March 2016 or Decembers 2015 and 2016 that were ice free. The EC preliminary totals include an estimate for March 2016 evaporation (17 mm). Ignoring sublimation in the SSPA and EC Preliminary contributes a small amount to the differences in the data sets. The estimated sublimation totals varied from 15 to 26 mm (0.6 to 1.0 inches) depending on the length of ice cover, 2.0 to 4.6 percent of the annual total EC + CLM4-LISSS evaporation.

The USGS used the Hargreaves-Samani method for estimating potential evapotranspiration (PET) from a vegetated surface to estimate lake evaporation during the ice-free season (Jones et al., 2017). It is commonly assumed that, over the course of the open-water season, total lake evaporation is close to shallow-water evaporation that is not significantly affected by heat storage. Shallow-water evaporation may also be similar to PET for the grass reference crop for with the Hargreaves-Samani was developed. Jones et al. (2017) calculated an average evaporation rate from White Bear Lake (2003 through 2013) of 719 mm/year (28.3 inches/year). The EC+CLM4-LISSS computed evaporation rate for the same time period is 751 mm/year (29.6 inches/year), just over 4 percent higher than the USGS estimate.

An important difference between the evaporation data sets is the year-to-year variation. The annual totals used by SSPA (2017) for 2014 through 2016 varied by only three percent reflecting very similar pan totals. This matched the preliminary EC estimates reasonably well. The EC + CLM4-LISSS value for 2016 is 34 percent higher than the value for 2014. The EC+CLM4-LISSS total for 2014 was less than 2015 and 2016 due to a lower daily average evaporation rate along with a shorter ice-free season. This suggests that, under some weather conditions, pan evaporation may not be a good indicator of relative variations in annual total evaporation from White Bear Lake.

Xiao et al. (2018) used the tuned CLM4-LISSS model for retrospective modeling of the period 1979 through 2016 as well as in a projected future climate scenario. The retrospective model results were used to generate revised evaporation inputs for the NMLG model.

The USGS-calculated average evaporation rate (2003 through 2013) corresponds to an annual pan coefficient of 0.79 using the UMN pan data compared to an annual pan coefficient of 0.82 for EC+CLM4-LISSS retrospective modeling. The *Evaporation Atlas for the Contiguous 48 United States* (Farnsworth et al., 1982) shows a calculated, May through October free-water surface (i.e. shallow water) evaporation of between 28 and 29 inches/year for the northeastern Twin Cities metro area and a pan coefficient between 0.79 and 0.80. The EC+CLM4-LISSS retrospective totals incorporate evaporation and sublimation outside of the pan-measurement season and are expected to be slightly higher than a May through October estimate. Therefore, the average EC+CLM4-LISSS evaporation rate is comparable to previous estimates for the region, whereas the preliminary EC data suggested a substantially lower average evaporation rate.

A key consideration for the retrospective modeling is the climate data used to force the model. During the observation period, Xiao et al. (2018) forced the CLM4-LISSS model with weather data collected at the EC system stations with missing data filled using a nearby Citizen Weather Observer Program weather station. A regional (North America) weather model reanalysis product (NLDAS-2 which is derived from the North American Regional

Reanalysis, NARR, See Xiao et al., 2018 for references) was used to fill any remaining data gaps and for downwelling longwave radiation. Reanalysis products are derived by running numerical weather models of the whole atmosphere with a consistent system for assimilating observations from satellites, weather balloon radiosondes, etc. at regular intervals. For the observed periods in 2014 through 2016, the CLM4-LISSS evaporation correlated closely with observations at the daily time scale (R^2 = 0.81) with the average computed evaporation 98 percent of the observed average.

Because on site meteorological data were not available, Xiao et al. (2018) used forcing inputs derived from the reanalysis product for the retrospective modeling. Although the reanalysis process assimilates a variety of data to update the weather model states, ground based measurements of some variables, such as wind speed, are not incorporated because they are affected by near-surface/localized features not represented in the model. Because the available assimilation data changed over time, the reanalysis product can have spurious fluctuations or trends.

Diurnal temperature variations and daily average humidity in the reanalysis data differed from on-site measurements primarily because of the local meteorological effects of the lake. Xiao et al. (2018) developed fitting functions to adjust the reanalysis temperature and humidity data for the entire retrospective modeling period. Other data such as wind speed and direction were not adjusted. The adjustments did not affect any trends in the reanalysis data, just the absolute values. To evaluate the CLM4-LISSS retrospective modeling results, DNR compared the retrospective model outputs to the available EC observations. At the daily time scale R^2 = 0.70, and the average computed evaporation was 107 percent of the observed average. The bias toward higher computed evaporation was largely driven by 2014 (late July through early November), which averaged 32 percent higher than observations.

The retrospective modeled evaporation showed an increasing trend (linear slope of 3.8 mm/yr) over 1979 to 2016 (Xiao et al., 2018). This general increasing trend was driven mainly by computed evaporation rates during the ice-free period. Xiao et al. (2018) attributed the increasing trend in evaporation rates to increased wind speed and increased temperature and lake-surface vapor pressure deficit during the ice-free period. The length of the modeled ice-free period also increased due to earlier modeled ice-out dates, although the trend in ice-out dates was less strong in the observed ice-out record. An increasing trend in air temperatures has been previously observed, but increasing trends in open-water evaporation and in wind speed deserve further exploration.

Historical Climatological Data and Lake Evaporation

Pan evaporation is generally considered to correspond closely to potential evaporation and has widely been used with pan coefficients to estimate open-water evaporation. Although heat storage in lakes affects the seasonality of evaporation, annual lake evaporation is commonly estimated with pan data as in SSPA (2017). The results in Xiao et al. (2018) indicate, however, that pan evaporation may not account for all factors affecting lake evaporation. It is nevertheless worth looking at trends in pan evaporation.

A number of researchers have demonstrated declining trends in pan evaporation in many parts of the world since the mid-20th century including much of North America (summarized in Roderick et al., 2009a). Roderick et al. (2009b) attribute the decline in pan evaporation at many locations to declines in radiation and/or wind speed.

DNR analyzed the University of Minnesota pan data and found a weak downward trend over the 1979 to 2016 period based on the Kendall's tau rank correlation (tau = -0.19, p =0.09) with a linear trend slope of -0.1 in/yr. Annual total, global shortwave radiation measured at the University of Minnesota showed no significant trend (Kendall's tau rank p = 0.58) over the 1979 to 2014 period (missing some data in 2015). The other factors affecting pan evaporation are vapor-pressure deficit (controlled by temperature and humidity) and wind speed.

Dadaser-Celik and Stefan (2008) calculated evaporation from shallow, open-water in equilibrium with the air temperature using several published mass-transfer type equations for weather stations around and immediately adjacent to Minnesota, including MSP International Airport. Input data were temperature, dew point temperature, and wind speed. They found a downward linear trend in calculated open-water evaporation at MSP for 1964 – 2005 (-0.99 in/yr, p = 0.07) with a weaker, statistically insignificant trend for 1986 – 2005 (-0.69 in/yr, p = 0.75). They did not describe any analysis or adjustment of the input climate data.

Changes in wind-speed measurement equipment and methods are important in assessing trends in wind speed, which can affect computation of evaporation. Like other climatological variables, a shift in measurement bias can throw off long-term modeling and analyses. Wind measurements at MSP changed from a height of 21 feet to 33 feet in October 1981 and relocated and changed from manual recording to an automated system with different equipment (ASOS) in June 1996 (Peter Boulay, State Climatology Office, pers. comm., May 2018). There was a later move of about 1,000 feet followed by a switch to a more accurate sonic anemometer. Before ASOS implementation, hourly wind speeds were recorded manually by visually averaging the anemometer dial.

Correcting for data inhomogeneity would be problematic and uncertain. For example, Klink (2002) isolated periods with homogenous data collection when assessing wind trends in Minnesota. She only evaluated MSP data from 1959 to 1980, eliminating measurement-height changes but also eliminating many years of data form the analysis. Errors in wind speed observations (instrumentation, human error, etc.) likely cannot be filtered out with certainty (Degaetano, 1998).

Lockhart (1999) compared synoptic hourly wind speed measurements in 1994-95 using both ASOS and manual measurements at 18 stations. Average differences in the measurements varied from -1 to +1 knots with standard deviations from 1 to 2.5 knots (1 knot = 1.15 mph = 0.51 m/s). The accuracy specification of the ASOS anemometers was +/- 2 knots. The bias varied with ASOS measured wind speed, typically with ASOS lower for lower speeds but higher for high speeds.

This points to a possible discontinuity in wind speed data with the transition to ASOS. The average November through April (typically ice-free months) wind speed at MSP from 1997 through 2016 was 9.0 mph. Therefore, a bias of 1 mph, although within the measurement error range for earlier ASOS instrumentation, represents an 11 percent shift in the average. To look further into this potential shift at MSP, the data post-height change and pre-ASOS (1982 through 1995) were compared to an equal length of ASOS data (1997 through 2010). The minimum April through November average for 1982 through 1995 (10.0 mph) was higher than the maximum April through November average for 1997 through 2010 (9.8 mph). A Wilcoxon rank-sum test (also known as

Mann-Whitney U test) of all the daily averaged data for these two data sets indicated a near zero probability that they are drawn from the same distribution or have the same median.

There is an apparent downward trend in wind speed data collected at MSP over the 1979 to 2016 period, but this is likely, at least in part, due to the change to ASOS in 1996. Other changes in the MSP equipment and measurement locations might also affect the homogeneity of the data. The MSP airport is the only long-term wind speed measurement location in the Twin Cities area. Unfortunately, the observation record cannot be used with confidence to verify or contradict the increasing trend in wind speed in the NLDAS-2 forcing data used by Xiao et al. (2018).

It is helpful to consider the NLDAS-2 data at the single grid point used in the White Bear Lake modeling in a larger context. Pryor et al. (2009) evaluated wind speed trends over the contiguous United States using two different observational data sets (mostly quality-controlled airport data), four reanalysis data sets (including the NARR used in NLDAS-2 for 1979-2006), and two regional climate models. They looked specifically at 50th and 90th percentile data from two specific hours of the day. The 1973-2005 observational data set indicated downward trends across most of the U.S., including MSP and other stations in Minnesota and surrounding states. The NARR reanalysis showed no trends in the Twin Cities area except for a weak downward trend in 90th percentile data for one of the two evaluated hours of the day in a small area south of the Twin Cities. More generally, the NARR data showed fewer and weaker downward trends in the central U.S. than the observational data.

Pryor et al. (2009) recognized the potential influence of the introduction of ASOS on their analysis. They looked at the years at each observation station that exhibited the largest discontinuity in wind speed data and compared the numbers of stations for each year to the years ASOS was introduced. There was not a correspondence. Although this test may not identify actual discontinuities caused by the transition to ASOS, the synoptic comparisons of Lockhart (1999) suggest that conversion to ASOS variably shifted average wind speeds upward or downward at different stations. Given this variability, the fact that downward wind speed trends were found for a large majority of U.S. stations by Pryor et al. (2009) is notable. Pryor et al. (2009) also note that differences in magnitude and direction of wind speed trends between observational data sets and reanalysis data sets were found in studies in Europe and Australia.

Additional Evaporation Analysis

DNR completed analysis of the available data to provide additional perspective on possible lake-evaporation trends. The analysis applied a semi-empirical approach to estimate lake-evaporation at the monthly time scale for comparison with the CLM4-LISSS retrospective modeling. Forcing the CLM4-LISSS with alternative, ground-based measurements could be explored in future work.

In addition to the use of pan coefficients, evaporation models also include combination-equation and mass-transfer methods. The combination method combines equations for surface-energy balance and turbulent transport of heat and water vapor (Penman, 1948). To apply to a lake, the combination-equation approach requires measurements or estimates of net radiation, lake-surface temperature, and heat storage in the lake along with standard meteorological measurements. Mass transfer (or aerodynamic) methods relate evaporation to the vapor pressure difference between the water surface and air and wind speed. Mass transfer equations include a constant that must be empirically fitted to a specific site or setting and also require measurements or estimates of the water temperature, dew point, and wind speed.

With the exception of lake heat storage, the other inputs required for the combination equation were measured at White Bear Lake as part of or along with the EC system measurements. Heat storage can be estimated as the dominant part of the energy-balance residual. In some cases researchers have fitted empirical functions for net radiation, surface temperature, and heat storage to measured data to extrapolate to periods without direct measurements of all of the components (e.g. Duan and Bastiaanssen, 2017). DNR explored both this approach and the mass-transfer approach but selected the mass-transfer approach based on simplicity and fit to EC observations.

Mass transfer equations take the general form:

$$E = Nf(u)(e_s - e_a)$$

Where N = an empirical coefficient
f(u) = a function of wind speed
es = vapor pressure of water surface (calculated from water temperature)
ea = vapor pressure of the air (calculated from dew point)

Mass transfer equations have been used successfully when empirically fitted to direct evaporation measurements such as from water-balance computations or EC system measurements (See Singh and Hu, 1997, for a review). Dadaser-Celik and Stefan (2008) applied several mass-transfer equations in their analysis of evaporation trends in Minnesota.

Vapor pressure of the air can be calculated directly from measurements of dew-point temperature. Wind speed measurements are also available, although inhomogeneity in the record can be a concern as discussed above. Although long-term water-surface temperature records are not available, water temperatures are highly correlated to air temperatures. Although the method is based on the vapor-pressure difference at the water-surface-air interface, temperatures at a shallow depth also work well (e.g. Meyer, 1944). Water temperatures at a shallow depth vary more smoothly than water temperatures at the surface.

Xiao et al. (2018) measured water-surface temperatures and temperatures at approximately 1-m depth at the EC stations. These data along with dew point and wind speed measured at the MSP airport were used to test the mass-transfer equation, first at the daily time scale in the form:

E [mm/day] = N u [m/s] $(e_s - e_a)$ [kPa]

Fitted with a coefficient (N) of 0.87, the daily values correlated closely with observations ($R^2 = 0.80$) with a bias of 1.9 percent. The fit was similar to the fit of the CLM4-LISSS model forced with on-site weather data. Wind speed was important to daily variations in evaporation. A model fitted without wind speed inputs had an $R^2 = 0.57$. Using monthly average data for all months with complete or nearly complete EC observations (17 months), an N of 0.81 was fitted with $R^2 = 0.90$ and a bias of 1.8 percent. The mass-transfer equation using monthly average data tended to slightly over-estimate evaporation in the spring and slightly under-estimate evaporation in the summer. Monthly relative errors varied from -18 to 27 percent.

To apply the method to the period before water-temperature records were available, the shallow water temperature had to be estimated. Monthly average water temperatures were estimated by regression to monthly average air temperatures for the current and previous months using the equation:

$T_w = a(T_{prev} – T_{curr}) + b\ T_{cur} + c$

Where
T_w is estimated monthly average water temperature [°C]
T_{prev} is average air temperature for the previous month [°C]
T_{curr} is average air temperature for the current month [°C]
a, b, and c are fitted coefficients

The estimated monthly shallow water temperatures fitted measurements with R^2 = 0.996 and a bias = 0.00 percent. An N = 0.81 also fit the data with estimated monthly temperatures including missing months filled with CLM4-LISSS modeled data (27 months, some partial due to ice cover); R^2 = 0.91 and bias = 0.96 percent.

Singh and Hu (1997) found that evaporation on the monthly time scale may not be correlated to wind speed and a mass-transfer equation with only vapor-pressure difference may work well for monthly data. Monthly average wind speeds at MSP during open water months varied from 3.3 to 5.3 m/s (7.4 to 12 mph) with annual averages of 4.2 to 4.4 m/s (9.4 to 9.9 mph). These variations may not be significant relative to the influence of vapor-pressure differences and overall model error.

A mass-transfer equations without wind speed was also fitted to all the monthly evaporation values (27 months). The results were: N = 3.21, R^2 = 0.92, and bias = -0.49 percent. These results were essentially the same as the monthly results that factored in wind speed. If extended to periods with significantly different wind speeds, particularly at the annual time scale, the wind speed would influence the evaporation calculated using a mass-transfer equation.

The two mass-transfer equations were applied with data from MSP airport from 1980 through 2016 and compared with the CLM4-LISSS retrospective modeling results. This includes periods before and after conversion of MSP to the ASOS system. For each model, the open-water season was specified using the best available data. White Bear Lake ice-out dates have been observed since 1978, but ice-in dates have been observed since 2012. Ice-in dates have been observed on Medicine Lake (located in central Hennepin County) since 1955. Observed ice-in and ice-out dates at the two lakes generally correspond closely. Observed Medicine Lake ice-in dates were used for White Bear Lake prior to 2012. During ice cover, the average sublimation rate of 0.17 mm/day measured in January-February 2017 by Xiao et al. (2018) was applied in all three models.

The computed annual total evaporation totals for all three models are shown in Figure 3 along with regression lines for each. The mass-transfer equation with wind speed is designated MTu, and the equation without wind speed is designated MTvpd. Higher measured wind speeds prior to 1996 cause MTu to be higher than MTvpd for those years. In the following years, the difference between MTu and MTvpd is variable and generally smaller.

The MTu totals have a generally downward trend due to higher values prior to 1996. The MTvpd values are closely correlated to the CLM4-LISSS values (R^2 = 0.85, relative difference = 0.39 percent).

The NLDAS-2 forcing data for the retrospective CLM4-LISSS model had a slightly increasing trend in wind speeds that affected the results to some degree. Nevertheless, the MTvpd model that ignores wind speed variations (i.e. it effectively assumes a constant wind effect incorporated into the empirical coefficient) corresponds reasonably closely to CLM4-LISSS at the annual time scale with very similar, generally upward trends over time. This suggests that any errors in the wind speed data applied in CLM4-LISSS may not have a substantially detrimental effect on computed evaporation at annual or longer time scales. Changes in vapor pressure difference may be the most important factor affecting variations in evaporation over longer time scales.

2.1.5. Groundwater Withdrawals

Groundwater-use volumes reported by permit holders are available in database form beginning in 1988. Earlier reports were previously available only as scanned paper records, and some of these older records are less complete. Metropolitan Council staff transcribed scanned water-use reports for municipal wells to spreadsheets for the period 1980 through 1987 and provided the data to DNR (John Clark, pers. comm., 2018). Wells are identified by their local names as indicated in the scanned permit reports (e.g. Well #1). These records are mostly complete for municipal permits, but some years were not reported for some permits. Permit reports for other types of water uses were not transcribed. These records would have to be connected to unique well identifiers and reformatted to transfer these data to the groundwater model. Some of these data were processed and used in the modified model described below in 4 Revised Model Development and Analysis.

2.1.6. White Bear Lake Water Budget Analyses

The Lake package calculates a dynamic water budget. Precipitation and evaporation rates and surface inflow/withdrawal volumes are model inputs. Cell-by-cell groundwater-lake exchange volumes, surface outflow volumes, and precipitation and evaporation volumes are computed for each time step using an iterative solver. The model grid cells assigned to the lake and computed water table control the surface over which groundwater-lake exchange occurs, but a stage-volume-area table defines total lake volume and surface area. An outlet rating curve is also specified. For the NMLG model, surface runoff into the lake was computed using the SWB model.

The characteristics of the model-computed components (groundwater exchange and surface runoff) were evaluated through examination of daily and monthly water budgets and an isolated, large precipitation event. Three years were initially chosen for the water-budget analysis: 1995, 2009, and 2016. These years encompass a range of climatic conditions and model errors that may shed light on actual system behavior. The DNR began collecting continuous stage and precipitation data in White Bear Lake in late 2013. These records were reviewed to find an isolated, large precipitation event that caused a sharp increase in lake stage. White Bear Lake received over three inches of rain on September 17, 2015. This event was examined to assess the potential contribution of storm-water runoff after large rainfall events.

Water Budget Analyses

The water budget analysis split out components that can be derived more directly from measurements (precipitation, evaporation, surface outflow, and change in volumetric storage) from the remaining balance made up of groundwater exchanges, surface runoff, and residual errors. Separating out the combined model-computed quantities (groundwater exchanges and surface runoff) yielded insights into sensitivities to input data errors and into the hydrologic behavior that the models should simulate.

This approach differed from a typical surface-water budget approach in which all surface components are estimated, and the remaining balance is assumed to be net groundwater exchange. In this case, the surface runoff component is also treated as an unknown quantity. Surface runoff computed by the SWB model is discussed further in Section 4. The budgets were calculated as cumulative volumes because changes in lake stage/volume reflect the cumulative effect of inflows and outflows.

The water budget equation is:

Closure = $\Delta V - (P - E - SO)$,

where

Closure is the cumulative volume (acre-feet) of combined net groundwater exchange, surface runoff, and data error (i.e. what is left after accounting for measurement-based volumes);

ΔV is the cumulative change in lake volume based on daily stage and the stage-volume curve (acre-feet);

P is the cumulative precipitation volume (acre-feet) calculated from daily precipitation multiplied by lake area at the daily stage;

E is the cumulative evaporation volume (acre-feet) calculated from the estimated evaporation rate multiplied by lake area at the daily stage;

SO is the cumulative surface-water outflow volume (acre-feet) calculated from the daily stage and the outlet rating curve.

Daily lake stages were estimated from measured lake stages using linear interpolation. The daily stages were converted to cumulative volume changes using the newly developed stage-volume-area table (See 2.3.2.below). Because the Daymet data appear to be biased high, a daily time series of the nearest available rain gage was extracted from HIDEN/MNGage data using the web based tool on the State Climatology Office web site. This time series will be referred to as Nr_MNGage. The gridded HIDEN data are monthly sums and could not be used for daily calculations. Water balances were calculated using both precipitation data sets.

The water-budgets were initially calculated prior to the final EC-based evaporation measurements and retrospective lake-evaporation modeling became available. Daily evaporation rates were applied as a uniform rate for each month using the monthly evaporation estimates calculated by SSPA (2017) or preliminary EC-based estimates when available. These evaporation values and the corresponding budgets were retained for comparison. Water budgets were also calculated using the evaporation modeling results (confined to the observed ice-free period) along with the average ice-period sublimation rate or revised EC observations, when available (Xiao et al., 2018). The preliminary evaporation rates originally used will be referred to as E_p, and the revised values will be referred to as E_r.

Initially, three years were selected for water-budget analysis based on characteristics of actual versus modeled lake-level changes: 1995, 2009, and 2016. Starting in 1991 (the wettest year on record at White Bear Lake) the lake level generally increased, rising above the lake outlet invert in 1995. The transient NMLG model computed less lake-level recovery during this period than actually occurred, with the largest error in annual lake change in 1995. The lake level fell from 2006 through 2010, with the largest annual drop in 2009. The lake level rose from 2013 through 2016, but the model-computed stage rose too much in 2015 and 2016. 2016 had the most precipitation since 1991. As explained below, 1993 was also evaluated for comparison after reviewing the 1995 and 2016 results.

Daily precipitation reporting with different recording times among different observers, errors in the daily interpolation of lake stages, and other input errors lead to some noise in the daily computations, even for 2016 in which stage was measured continuously for most of the year (Figure 4). Despite this noise, computing volumes at daily intervals provides good detail and accuracy for evaluating system behavior. Monthly values extracted from the daily computations are plotted in Figures 5 to 8 to remove the noise and to facilitate comparisons.

A flat water-budget closure (Daymet Close and Nr_MNGage Close in Figures 5 to 8) indicates that net groundwater exchange was close to zero or that low rates of net groundwater outflow were compensated by surface runoff. A downward slope in the water-budget closure indicates net groundwater outflow. An upward slope in the water-budget closure indicates net groundwater inflow and surface-runoff.

In most years, the net groundwater exchange for White Bear Lake is expected to be an outflow because precipitation plus surface runoff usually exceeds evaporation plus surface outflow. The net groundwater exchange is highly variable, and previous water balance analysis suggested that there may have been a net groundwater inflow in some years (DNR, 1998). The interpretation of net groundwater exchange, however, is affected by errors in precipitation, evaporation, and runoff.

Both 1995 and 2016 were high precipitation years, but there were significant differences between the two years in lake-level response, net groundwater exchange, and surface runoff. These differences in behavior are intriguing, particularly considering the contrasting model errors for these two years. The transient NMLG model predicted too little rise in stage in 1995 but too much rise in stage in 2016.

At the end of 1995, the cumulative volume needed to close the water budget (Daymet Close and Nr_MNGage Close in Figure 5) was positive, indicating that total groundwater inflows plus surface runoff inflows exceeded groundwater outflows. In contrast, the closure volume at the end of 2016 was negative (Figure 6), indicating that groundwater outflows exceeded groundwater inflows plus surface runoff.

Depending on the data sets used, precipitation minus evaporation was higher in either 1995 or 2016. Assuming that Nr_MNGage and E_r provide the best daily data, total precipitation was slightly higher in 2016 than in 1995, but evaporation was also higher in 2016. Net Nr_MNGage-E_r-SO was 16.4 inches in 1995 but was 12.0 inches in 2016. The annual total can also be calculated from the monthly, gridded HIDEN data, which cannot be used for daily computations. The gridded HIDEN data may be the most reliable monthly or annual precipitation data source since it incorporates multiple gages near the lake. The net HIDEN-E_r-SO was 13.3 inches for 1995 and 11.1 inches for 2016. These data sets indicate net lake-surface flux was either 4.4 or 2.2 inches higher in 1995 than in 2016.

Importantly for the transient NMLG model, the net Daymet-E_p-SO model input was 17.7 inches for 1995 and 25.4 inches for 2016, a net lake-surface flux 7.7 inches less in 1995 than in 2016. This represents a large contrast in both absolute and relative lake fluxes between 1995 and 2016 for the model inputs (Daymet and E_p) versus what is expected to be more accurate available data at the time of writing this report (gridded HIDEN or Nr_MNGage and E_r).

In 1995, the lake level rose steadily from March through August (peaking at an increase of 21 inches in mid-August). The closure volumes gradually declined from mid-April through June, indicating net groundwater outflow. There was then a large increase in closure volume in July and August, indicating that groundwater inflow plus surface runoff substantially exceeded groundwater outflow during that time. Sharp upward jumps in the daily closure volumes (not shown) indicate that several inches of runoff resulted from two large (3 to greater than 4 inches) precipitation events in July and August 1995. There was then rapid net groundwater outflow in September, gradual net outflow in October, and very gradual or near zero net groundwater outflow in November and December.

In contrast, the cumulative closure volume remains negative or close to zero throughout the year for both precipitation and evaporation data sets in 2016. The closure volume decreased slowly or held nearly steady in winter and autumn but decreased in May and July and also August in some data combinations. There might have been a small amount of runoff during and shortly after snowmelt in March and April, but direct precipitation on the lake can explain all or most of the rapid lake level response to precipitation. The lake closure volumes show very little immediate response after large (3 to greater than 4 inch) precipitation events in June, August, and September (Figure 6), very unlike the response to the two large events in 1995. Gradual increases in closure volume in June and August (for Nr_MNGage, Er) indicate net groundwater inflow. The net Closure volume for the year was negative (with a wide range in values depending on the precipitation and evaporation data sets used) indicating overall net groundwater outflow for the year.

Stormwater inflow by itself pushes the lake level above equilibrium with groundwater, inducing more groundwater outflow and less groundwater inflow. Given the large positive closure volume that developed in summer 1995, it appears that net groundwater exchange was positive (inflow) during this period. This indicates that there was a large increase in groundwater inflows and possibly a decrease in groundwater outflows despite significant stormwater inflows. This points to a large rise in shallow groundwater heads near White Bear Lake. Available observation-well data are consistent with this general interpretation.

DNR observation wells 62045 (shallow buried Quaternary) and 62044 (Priairie du Chien) are nested near the south shore of the lake. Water levels in 62045 are above water levels in 62044, indicating a downward component to the hydraulic gradient (Figure 9). With the exception of a brief period in June 2014, the lake level remained above the water level in 62045 throughout the period of record (1995-present). The differences between the lake and 62045 averaged about one foot in both years, but the differences appear to have varied over a larger range in 1995. Head differences were less in July and August 1995 compared to 2016. Interestingly, the vertical head difference between 62045 and 62044 was smaller throughout 2016 than in 1995, and water levels in 62044 rose higher in 2016 than in 1995 despite lower lake levels in 2016.

Therefore, differences in vertical gradients within the aquifer system at this location do not explain the different shallow groundwater responses. This observation well pair only represents one location, but the data suggest that groundwater outflows were comparable in 1995 and 2016 except in July and August. In 1995, the shallow groundwater head increased more than the lake level in July and August.

The model does not appear to capture the conditions that lead to substantial surface runoff and rapid water-table rise in 1995. The SWB model computed 168 acre-feet of surface runoff in 1995, only about 0.8 inches over the lake surface. Computed net groundwater outflow in the SSPA triannual model was -13.8 inches. Although errors in precipitation and evaporation inputs likely explain some of the model error, those errors do not explain the error in the modeled groundwater response. The differences in the real system that lead to the different hydrological responses may be complex and not well represented by the model structure.

The lake level at the beginning of 1995 was about 1.4 feet higher than at the beginning of 2016, and the water table near the lake was likely also correspondingly higher. Water levels in some Prairie du Chien and deep buried Quaternary observation wells to the north and east of White Bear Lake were higher in 1995 than in 2016: 82024 (QBAA, Stillwater Township), 82039 (OPDC, southeastern corner Hugo), and 82029 (OPDC, central Hugo). Water-table and runoff response may depend on the antecedent water-table elevation. The lake level at the beginning of 1993 (922.2 feet) was close to the lake level at the beginning of 2016 (922.3 feet) and the lake level remained below the outlet elevation. Therefore, the calculations were also carried out for 1993 (Figure 10).

In 1993, net Daymet-E_p was 20.2 inches, and net Nr_MNGage-E_r was 13.9 inches. These are variably lower (Daymet-E_p) or higher (Nr_MNGage-E_r) than the totals for 2016 (25.4 inches and 12.0 inches, respectively). The closure volumes indicate slow net groundwater outflow in winter and more rapid net groundwater outflow in September and October (2 to 3 inches over the two months). At other times, the net groundwater exchange was positive or close to zero. Lake-level measurements in 1993 were too infrequent to clearly identify runoff events, but there was likely a runoff contribution.

The net rise in lake level in 1993 was 1.5 feet (18 inches) compared to 0.9 feet (11 inches) in 2016. The closure volumes using either precipitation data set with the revised evaporation data (E_r) are positive for at least six months (May through October) in 1993. The net closure volume ends the year either close to zero (Daymet, E_r) or positive (Nr_MNGage, E_r). This contrasts with a negative closure volume for the Daymet, E_p combination applied in the SSPA NMLG model. The triannual model computed net groundwater exchange of -12.4 inches for 1993, but actual net groundwater exchange was close to zero or positive (depending on surface runoff and the precipitation data used).

Lake-level fluctuations and water-budget assessment for 2009 was less complex than for the wet years. There was below average precipitation with net Daymet-E_p of 7.7 inches and net Nr_MNGage-E_r of 0.4 inch. The lake level declined about 0.9 ft. The computed lake-stage decline in the triannual model was 0.85 ft. The limited number of lake stage measurements preclude detecting short-term stage fluctuations that may indicate runoff, but runoff appears to have been relatively minimal in 2009. Daymet and Nr_MNGage precipitation were similar, and there appears to have been net groundwater outflow for all months of the year. There may have been net groundwater inflow for brief periods following recharge events. The net closure volumes (Daymet, E_p Close and Nr_MNGage, E_r Close) ended the year at about -3,700 and -2,100 acre feet or -19 and -11 inches over the lake, respectively.

Although the pattern is qualitatively similar for all data-set combinations, the higher E_r values result in substantially lower net closure volume than the E_p used in the SSPA NMLG model. The net groundwater flow computed by the triannual model for 2009 was about -3,200 acre-feet (-18 inches over the computed lake area or about -16 inches over the actual lake area). Computed stages were 2 to 5 inches higher than observed stages, but the computed lake area was about 260 acres less than "observed" because of the stage-volume-area table used. The water-budget evaluations suggest that the model-computed, net groundwater outflow in 2009 was larger than actual, however, assuming that E_r is more accurate.

One cause of the contrasting model errors for 1993, 1995 and 2016 was errors in the direct precipitation to and evaporation from the lake surface. Nevertheless, it is not clear why net groundwater outflow was so much higher in 2016 than in 1993 or 1995. Net P-E-SO was most likely higher in 1993 and 1995 than in 2016, although only 1.9 inches higher in 1993 (Nr_MNGage-E_r). It appears that the biggest contrast between these periods was that groundwater inflow (which occurs mostly at shallow depths) and runoff were much higher in 1993 and 1995. Although computed pumping drawdown is imperfect, the model results and available observation-well data suggest that variations in groundwater pumping over time cannot explain the differences in observed behavior. Heads rose higher in 2016 than they were in 1995 in the Prairie du Chien aquifer at the only observation well with long-term data adjacent to White Bear Lake. Heads at this location were most likely even lower in 1993 than in 1995.

Contrasting shallow groundwater dynamics in 1993 and 1995 versus 2016 that are not closely matched in the model appear to have been important. Under-simulation of the water-table rises in 1993 and 1995 likely relates to a combination of limitations in the SWB model and its inputs (precipitation, land use, etc.) and errors in the distribution of hydrogeological properties near and beneath White Bear Lake. Under-representing surface runoff in the model in 1993 and 1995 also appears to have played a role in model errors.

The SSPA model computed rise in lake level in 2016 (and 2015) was too large, even with computed, net groundwater outflow of -7.7 inches in 2016 (-8.3 inches in 2015). The Daymet precipitation appears to be much too high along with E_p much too low at White Bear Lake in 2015 and 2016. Net precipitation minus evaporation errors appear to be an important factor in the 2015 and 2016 model errors.

Farther afield, groundwater conditions were variable. Some other observation wells in the area also had higher levels in 2016 versus 1993 and/or 1995, for example, 62038 (Prairie du Chien, White Bear Township), 62008/62055 (buried Quaternary, Maplewood), 62002 (St. Peter, Roseville), and 62030 (Jordan, Shoreview). The model also showed higher heads in 2016 at these wells. As mentioned above, water levels in observation wells to the north and east were higher in 1993 and/or 1995. Three of these wells (82024, 82031, and 82039) are distant from high capacity pumping and likely minimally influenced by groundwater use. The model computed heads at these latter locations were higher in 2016 than in 1993-1995. Modeled, multi-year head variations at these locations were generally too small, both rises and drops. The amount of head rise from 2014 through 2016, however, was much more closely matched.

These spatial and temporal patterns in model error may point to errors in the temporal and spatial variation of recharge. Groundwater is intimately connected with the high density of lakes and wetlands in the northeastern Twin Cities area. This complicates estimating groundwater recharge. Complex dynamics in the shallow, Quaternary system that are not always well matched by the model seem to be important to the hydrology of White Bear Lake, particularly during periods with a generally rising water-table. Changes in land use and stormwater practices may also be factors, but it is not clear how significant those changes may have been in the vicinity of White Bear Lake.

Two observation-well nests have been installed near White Bear Lake since 2013, one to the northwest and one to the southeast of the lake. The data collected from these wells will continue to gain in value as the length of the record increases. None of these wells is open to the shallow water-table, however. The shallowest well in the southeastern nest (82057) is 154-feet deep and screened immediately above the top of the Prairie du Chien. The screened aquifer at 82057 appears to be locally unconfined but is laterally connected to confined Quaternary materials and the St. Peter Sandstone. The shallowest well at the northwestern nest is open to the Prairie du Chien.

Event Analysis

The September 17, 2015 storms provide the opportunity to evaluate the rapid response of lake level to precipitation. Based on provisional, hourly data from the DNR tipping bucket gage adjacent to Manitou Island, a small amount of rain fell in the hour before 11 PM on September 16th. This was followed by two periods of more intensive rain in the morning and early afternoon of the 17th. Lake-level response was immediate.

There had been minimal rainfall for the previous 11 days. This represents a dry antecedent runoff condition in the curve number method employed in SWB. This event could indicate if there are impervious/low permeability areas that are hydrologically well connected with White Bear Lake, even under dry antecedent conditions. Rapid lake-level rise that could not be fully accounted for by direct precipitation on the lake would be an indicator of surface runoff. The key data limitation is that the interpretation can be sensitive to errors in the average precipitation over the lake surface.

The DNR tipping bucket gage appears to be biased low for this event and more generally for this time period. For example, the monthly total in the DNR gage (4.39 inches) was 1.06 to 1.45 inches less than the monthly totals for 6 HIDEN/MNGage stations located within about ½ mile of the lake. Monthly totals for the DNR gage were also lower for August and October 2015. Therefore, daily totals from the six HIDEN/MNGage stations were used for the event analysis. The reported observation times were used to determine which daily values to incorporate into the event total based on the times precipitation was recorded in the DNR tipping bucket.

Event totals at the six rain gages varied from 2.95 to 3.64 inches (0.25 to 0.30 ft). A spatially weighted average of 3.35 inches (0.28 feet) was calculated by grouping the gages by proximity to each of the lake's three basins/bays. This corresponds to a rainfall event with a recurrence interval of about 5 years (NOAA, Precipitation Frequency Data Server). The Daymet precipitation total for this event was 2.40 inches (0.20 ft) spread over two days. The

MNGage data indicate there was about 0.2 to 0.3 inch (0.017 to 0.025 ft) of precipitation the following afternoon of the 18th.

The lake level was declining at a rate of -0.0174 feet per day before the precipitation event. Projecting this rate to the afternoon of September 17th, the lake level would have been 921.23 feet (MSL, 1912). The recorded level rose to 921.53 feet and fluctuated within 0.01 feet of 921.52 on the afternoon of the 17th. Measurement noise appears to be +/- 0.01 feet both before and after the event. Therefore, an immediate lake-level rise of 0.29 feet (3.5 inches) corresponds to the rainfall event. The lake level appears to rise 0.02 feet (0.24 inches) in response to the rainfall on the 18th. The recorded lake level continued to fluctuate around 921.52 feet until beginning to decline again on September 21st.

It is possible that there was a small amount of runoff from this event, but the difference between direct precipitation and lake-level rise is within the measurement errors of precipitation and lake stage. The sustained lake level for several days after the event could indicate some delayed runoff. Rapid water-table response to rainfall immediately adjacent to the lake, particularly at locations where the shore slopes are very low, may also explain a small fraction of the rapid lake level rise and sustained lake levels several days after the event.

Lake levels were recorded at irregular intervals but typically at least once per week in 1995. As discussed above, the available data appear to indicate that there was significant runoff during/following large, summer precipitation events.

2.2. Hydrogeological Properties

DNR reviewed the hydrogeological property parameter values that were applied in the transient model and compared these values to previous models and other previous estimates.

2.2.1. Aquifer and Aquitard Hydraulic Conductivities

Horizontal and vertical hydraulic conductivity and conductance of quasi-three-dimensional confining units should be constrained to reasonable values based on literature values, analyses of aquifer tests, and consideration of the modeling scale.

The hydraulic conductivity parameterization of Jones et at. (2017) was preserved in the transient model, but the parameter values were changed (SSPA, 2017). One relatively minor modification was allowing a conductivity multiplier that applies to the top half of layer 1 (top 1.5 m or 4.9 ft) based on the hydrologic soil group (HSG). These multipliers were defined in the input processes for the steady-state model but fixed at 1.0.These multipliers affect a relatively thin part of the shallow groundwater-flow system and have no effect when the computed water table is below the top model layer. Most of the hydraulic conductivity parameters were modified in the parameter estimation process, but the largest and most systematic changes were concentrated in the Quaternary sediments parameters and the pilot-point multipliers for the Prairie du Chien and portions of the Jordan.

Quaternary Sediments and Decorah-Platteville-Glenwood

The parameterization of the Quaternary sediments in the NMLG models follows the grid-based mapping of Tipping (2011) that was applied to Metro Model 3 (Metropolitan Council, 2014). The sediments were divided into 11 classes based on texture and two depth categories (shallow and deep) for loamy textures. Jones et al. (2017) filled in gaps in the sediment-class grid left by Tipping where grid points were distant from any well/boring materials logs in the same elevation interval. The Decorah-Platteville-Glenwood (D-P-G) bedrock grouping was also included in this parameter group because it was lumped into layer 1 in Metro Model 3. Hydraulic conductivity values are assigned to each class, and effective cell hydraulic conductivities were calculated based on the intersection of the sediment-class grid with the groundwater-flow model grid. The division of these materials into four layers in the NMLG model allows for representation of some of the vertical differences in properties and corresponding vertical differences in computed heads.

The horizontal and vertical hydraulic conductivities of glacially associated sediments are highly variable, and the ranges of expected values of several of the sediment classes overlap. Nevertheless, when lumped into the 11 classes, there are reasonable expectations about the relative values among several of the classes. Sand and gravel is expected to have the highest value. Sandy loam is expected, on average, to have higher values than loam to clay loam. For the same texture, shallow classes have higher, or perhaps the same, value as the corresponding deeper class.

The horizontal and vertical hydraulic conductivities applied by Jones et al. (2017) in the steady-state model and by SSPA (2017) in the transient model are shown in Tables 1 and 2, respectively. These values are compared by rank (maximum to minimum) and against the minimum and maximum values allowed for each class listed in Table 3 of Jones et al. (2017). It should be noted that the Jones et al. (2017) report containing these allowable parameter ranges had not been released at the time the transient model was being developed. The parameter ranges of Jones et al. (2017) are reasonable, are generally consistent with previous estimates (e.g. Metropolitan Council, 2014), and provide a useful frame of reference. In some cases, parameter values close to but outside of these ranges may be reasonable.

Several of the horizontal hydraulic conductivity values in the transient (SSPA) model fall well outside of the expected range (e.g. quat_hk2). The relative relationships of several of the horizontal hydraulic conductivities also do not fit expectations. For example, sand is assigned a smaller value than several classes of loamy sediments generally representing glacial till deposits. These parameter values and relationships represent a departure from the conceptual model of Tipping (2011), Metropolitan Council (2014), and Jones et al. (2017).

Most of the vertical hydraulic conductivity values used in the transient (SSPA) model are within the expected ranges. The vertical hydraulic conductivity of deep loam to sandy loam (quat_vk9) is less than the minimum chosen by Jones et al. (2017), but, as expected, it is greater than the values for deep loam to clay loam (quat_vk8) and the D-P-G (quat_vk12).

The spatial distribution of classes was not modified during parameter estimation. The parameter values were not tightly constrained during parameter estimation because of the limitations and uncertainties imposed by the spatial distribution of parameters (Vivek Bedekar, SSPA, pers. comm.). Inaccuracies in the spatial distribution of materials assumed to have similar properties is a type of structural error in a model.

There are limitations to the sediment-class model and to the resulting spatial parameterization of hydraulic conductivity of Quaternary sediments. Nevertheless, it is likely that the sediment-class model captures many of the major geological features that affect the distribution of major sediment bodies having contrasting hydraulic conductivities. These may include features such as laterally extensive and/or thickly stacked till deposits and broad belts of outwash sand and gravel.

Applying parameter values outside of the expected ranges or that do not conform to expected relative values also risks parameter overfitting and its own structural errors. The resulting parameters could degrade the predictive capability of the model and may not reduce or effectively compensate for structural errors resulting from the spatial distribution of parameters and model layering. This was considered in development of the revised model.

Table 3 – Horizontal hydraulic conductivity parameter values (feet per day, ft/d) for Quaternary sediments and the Decorah-Platteville-Glenwood unit

USGS Parameter Name	Description	USGS Value	USGS Lower	USGS Upper	USGS Rank	SSPA Value	SSPA Rank
quat_hk1	loam to clay loam	2.3	0.51	13	9	61 [1]	2 [2]
quat_hk2	loam to sandy loam	36	14 [3]	41	3	129 [1]	1 [2]
quat_hk3	loam, silt rich; silt and clay	2.1	0.41	10	10	15	8
quat_hk4	loam to sandy clay loam	25 [3]	--	--	4	3.3	12 [2]
quat_hk5	sand and gravel	92	39	118 [3]	1	52	3 [2]
quat_hk6	fine sand	51	20	59	2	46	4 [2]
quat_hk7	sandy silt	7.9	3.0	12	6	3.4	11 [2]
quat_hk8	loam to clay loam (deeper than 60 feet below the land surface)	1.8	0.37	9.2	11	4.0	9-10
quat_hk9	loam to sandy loam (deeper than 60 feet below the land surface)	8.1	2.6 [3]	10	5	38 [1]	5
quat_hk10	loam, silt rich; silt and clay (deeper than 60 feet below the land surface)	3.9	2.1 [3]	8.5	8	26 [1]	6-7

USGS Parameter Name	Description	USGS Value	USGS Lower	USGS Upper	USGS Rank	SSPA Value	SSPA Rank
quat_hk11	loam to sandy clay loam (deeper than 60 feet below the land surface)	4.6 [3]	--	--	7	26 [1]	6-7
quat_hk12	Decorah Shale, Platteville Limestone, and Glenwood Formation confining unit	1.6	0.33	8.2	12	4.0	9-10

[1] Value more than 50 percent greater than expected maximum value

[2] Unexpected relative parameter rank

[3] Reasonable value may be outside of range defined by this maximum or minimum value.

Table 4 - Vertical hydraulic conductivity parameter values (feet per day, ft/d) for Quaternary sediments and the Decorah-Platteville-Glenwood unit

USGS Parameter Name	Description	USGS Value	USGS Lower	USGS Upper	USGS Rank	SSPA Value	SSPA Rank
quat_vk1	loam to clay loam	0.036	0.0036	0.36	8	0.011	8
quat_vk2	loam to sandy loam	1.9	0.18 [3]	4.6	2 [2]	1.1	2 [2]
quat_vk3	loam, silt rich; silt and clay	0.0010	4.9E-06	0.066 [3]	10	0.00082	10
quat_vk4	loam to sandy clay loam	0.62 [3]	--	--	5	0.33	5
quat_vk5	sand and gravel	23	16 [3]	47	1	0.66	3 [2]

USGS Parameter Name	Description	USGS Value	USGS Lower	USGS Upper	USGS Rank	SSPA Value	SSPA Rank
quat_vk6	fine sand	1.6	0.45	11	3 [2]	3.4	1 [2]
quat_vk7	sandy silt	0.66	0.0082	0.82	4	0.54	4
quat_vk8	loam to clay loam (deeper than 60 feet below the land surface)	0.00036	4.6E-06	0.033	11	0.00013	12
quat_vk9	loam to sandy loam (deeper than 60 feet below the land surface)	0.042	0.0049	0.49	6-7	0.0019 [1]	9 [2]
quat_vk10	loam, silt rich; silt and clay (deeper than 60 feet below the land surface)	0.023	0.0039	0.39	9	0.033	7
quat_vk11	loam to sandy clay loam (deeper than 60 feet below the land surface)	0.043 [3]	--	--	6-7	0.043	6
quat_vk12	Decorah Shale, Platteville Limestone, and Glenwood Formation confining unit	0.00013	3.3E-05	0.33	12	0.00024	11

[1] Value more than 50 percent greater than expected maximum value
[2] Unexpected relative parameter rank
[3] Reasonable value may be outside of range defined by this maximum or minimum value.

Prairie du Chien-Jordan

The Prairie du Chien Group is generally considered to consist of two main hydrostratigraphic units. The upper one-half to two thirds includes the Shakopee Fm. and sometimes the Upper Oneota Dolomite and acts primarily as an aquifer. The lower half to one third is generally an aquitard consisting of most of the Oneota Dolomite. The bulk hydraulic conductivity of the Prairie du Chien is highly variable and is controlled by flow through secondary porosity (i.e. connected, open and solution enhanced fractures).

Interconnected and solution-enhanced fractures are generally expected to be more abundant where the depth to the bedrock surface is shallow, defined as less than 200 feet by Runkel et al. (2003). The top of the Prairie du Chien is within 200 feet of the bedrock surface within most of the NMLG domain, except where there are extensive remnants of Decorah Shale, such as in and near the west side of St. Paul.

In general, enhanced permeability may be gradational toward the bedrock surface. At greater depths below the bedrock surface, hydraulically active secondary porosity appears to be concentrated in discrete intervals (Runkel et al., 2003). Overall transmissivity of carbonate bedrock units such as the Prairie du Chien is typically dominated by discrete intervals even in shallow settings, however. High permeability intervals are limited or absent in the lower aquitard portion of the Oneota Dolomite in deeper bedrock settings. In shallower settings, the upper aquifer is also expected to have higher transmissivity.

Variability in the hydraulically dominant fracture sets and dissolution features make local-scale bulk hydraulic conductivity unpredictable. The range of bulk hydraulic conductivity of the Prairie du Chien in relatively deeper bedrock settings overlaps with the range where it is the upper-most bedrock. This reflects the variability in the upper aquifer, but the lower part of the Oneota likely also affects variability in the bulk hydraulic conductivity of the Prairie du Chien Group. Where erosion has made the Oneota the upper-most bedrock unit and enhanced permeability, the bulk hydraulic conductivity likely differs from the bulk hydraulic conductivity of the Prairie du Chien in an adjacent area where the Shakopee is the top of bedrock.

Groundwater flow through the Jordan Sandstone is largely through inter-granular pores, but flow through fractures can also be important, particularly in shallow bedrock conditions (Runkel et al., 2003). The Jordan Sandstone includes fine-grained intervals that generally restrict vertical flow. There is commonly a fine grained interval at the base of the Jordan that is gradational with the underlying St. Lawrence Formation. Like the lower Tunnel City Group and other fine-grained bedrock units, bedding-plane fractures in fine-grained intervals in the Jordan may allow significant horizontal flow.

Information from Aquifer Tests

There are few estimates of the transmissivity/bulk hydraulic conductivity of the Prairie du Chien Group in or near the northeastern metro area. There are a greater number of aquifer test analyses for wells completed in the Jordan aquifer or open to both the Prairie du Chien and Jordan. Transmissivity estimates derived from aquifer tests were assembled during development of Metro Model 3 (Metropolitan Council 2014). This compilation represents most of the reasonably reliable aquifer-test analyses in the area (Table 5). The spatial distribution of these tests differs, with more Jordan aquifer tests closer to the perimeter of the area (e.g. southern Washington County) and the Prairie du Chien-Jordan tests in locations in the western half of the area. Information for additional aquifer tests and single-well tests are available from DNR and Minnesota Department of Health records, but these records would have to be individually reviewed to extract estimates comparable to the previous compilation.

Table 5 Summary of transmissivity estimates within the active NMLG model domain for the Jordan and Prairie du Chien-Jordan from aquifer-test analyses

Statistic	Jordan Transmissivity (ft^2/d)	Prairie du Chien-Jordan Transmissivity (ft^2/d)
Count	28	25
Minimum	2,000	690
Maximum	25,000	54,000
Median	4,300	9,800
Geometric Mean	4,700	9,800
25th percentile	2,700	8,000
75th percentile	8,500	15,000

The available transmissivity estimates provide useful guidance to constrain the transmissivity and corresponding bulk hydraulic conductivity of the Prairie du Chien-Jordan aquifer system, but there is a broad range in the estimates. There is inherent uncertainty in the estimates, but the broad range in transmissivity estimates for the area is likely substantially greater than the uncertainties in most of the individual estimates.

The Jordan is typically 85 to 105 feet thick, and most hydraulic conductivity estimates are in the range of 25 to 175 ft/d. The Prairie du Chien is 1.3 to 2 times the thickness of the Jordan, except where the Prairie du Chien is deeply eroded. Therefore, if the bulk hydraulic conductivity of the Prairie du Chien and Jordan are similar at a given location, the transmissivity of the combined units would be roughly double to three times the hydraulic conductivity of the Jordan alone. Given the ranges of transmissivity estimates for the Jordan and the combined units, it appears that the range of bulk hydraulic conductivities of the Prairie du Chien and Jordan are broadly overlapping. It is plausible that the two units have similar hydraulic conductivities at some locations, but there is substantial variability.

At a few locations there are relatively closely spaced estimates for single and combined aquifer pumping wells. Transmissivity estimates for St. Paul Regional Water Services wells open to both aquifers in Vadnais Heights and nearby Little Canada vary from 9,200 ft^2/d to 14,000 ft^2/d. A transmissivity estimate for a nearby Jordan well (Unique No. 208285) is 2,300 ft^2/d (hydraulic conductivity 23 ft/d). If the Jordan transmissivity estimate from Unique No. 208285 is applicable to the nearby Prairie du Chien-Jordan well sites, the transmissivity and hydraulic conductivity of the Prairie du Chien can be estimated by subtracting the assumed transmissivity of the Jordan from the bulk, Prairie du Chien-Jordan transmissivity.

The thickness of the Prairie du Chien varies from about 70 to 130 ft, and the thickness of the Jordan varies from about 85 to 100 feet at these wells. The overlying St. Peter bedrock is thin or absent in this area (i.e. shallow bedrock conditions for the Prairie du Chien). Given the range of thicknesses of each unit and the range in total transmissivity, the bulk hydraulic conductivity of the Prairie du Chien would be 3 to 6 times the bulk hydraulic conductivity of the Jordan in this area. This one estimate is tentative and should not be extrapolated. Nevertheless, it is notable because the Prairie du Chien is in relatively shallow conditions across most of the area and is closer to the bedrock surface than the Jordan. This may make the hydraulic conductivity of the Prairie du Chien typically higher than the Jordan in the NMLG area.

Model Comparisons

Model layering and lateral extents for the bedrock units (with the exception of the patchy Decorah-Platteville-Glenwood that was lumped with the Quaternary) were carried over directly from Metro Model 3. The Prairie du Chien Group is represented in Metro Model 3 and the NMLG models by one model layer with an underlying quasi-three-dimensional confining unit that adds to the vertical resistance between the Prairie du Chien aquifer (usually layer 6 in the NMLG models) and the Jordan aquifer (usually layer 7 in the NMLG models). The confining unit will be discussed separately under *Confining Units* below.

In Metro Model 3, there were two, overlapping processes that determined hydraulic conductivities of each model cell representing bedrock units: interpolation of base hydraulic conductivity from pilot points, and a multiplier that increases exponentially as the thickness of overlying bedrock units decreases. Within the NMLG active model domain west of the St. Croix River, four Prairie du Chien pilot points dominate the base hydraulic conductivity in Metro Model 3. These pilot points have values of 10.5, 20.0, 66.3, and 186 ft/d (Metropolitan Council, 2014, Figure 56). The depth-controlled multiplier varies from 1.0 at about 11 m (36 ft) to a maximum of 1.7 at zero depth. The pilot points vary by greater than a factor of 1.7, but the depth multiplier causes sharp, local variations in hydraulic conductivity where the St. Peter Sandstone is patchy.

The resulting hydraulic conductivity varies from 16 ft/d to 260 ft/d. The hydraulic conductivity generally varies from 30 to 150 ft/d across most of the focus area for the NMLG model. Bulk hydraulic conductivity of the Prairie du Chien in the model of southern Washington County falls within this range (50 to 60 ft/d). This model was later modified to include more hydraulic conductivity zones with hydraulic conductivity of the Prairie du Chien varying from 1 to 160 ft/d (Barr Engineering Co., 2010).

The spatially varying Metro Model 3 hydraulic conductivities were modified in the NMLG model via a different set of pilot points. Multiplier values rather than hydraulic conductivity were specified at these pilot points. These multipliers were interpolated to all cells representing the Prairie du Chien, and the resulting values were multiplied by the hydraulic conductivity assigned to each corresponding cell in Metro Model 3. Hence, conductivities specified through the two-component process in Metro Model 3 were modified by a smoothly varying multiplier. For the Prairie du Chien, there were 12, unevenly spaced multiplier pilot points.

For the steady-state NMLG model, 11 of the 12 pilot point values were assigned values close to one (1), resulting in limited changes to the spatial distribution of hydraulic conductivity. For the transient model, the pilot points were assigned values from 0.156 to 0.168 through automated parameter estimation. This relatively narrow range in multipliers did not greatly change the spatial distribution of hydraulic conductivity in a relative sense, but there was a significant reduction in hydraulic conductivity across the model domain. The hydraulic conductivity values at the 12 pilot point locations for the three models are shown in Table 6 as examples for comparison. Again, hydraulic conductivities vary cell-by-cell in all three models.

Table 6 – Horizontal hydraulic conductivity values for the Prairie du Chien at NMLG model pilot point locations

Pilot Point Name	Location Description	Metro Model 3 (ft/d)	Steady-State NMLG (ft/d)	Transient NMLG (ft/d)
opdc_hk1	Roseville	70	71	11
opdc_hk2	St. Anthony	114	115	18
opdc_hk3	southern Woodbury	21	16	3
opdc_hk4	eastern St. Paul	36	33	6
opdc_hk5	White Bear Lake	65	65	11
opdc_hk6	southern Hugo	112	110	19
opdc_hk7	West Lakeland Twp	85	81	14
opdc_hk8	northeastern Maplewood	56	57	9
opdc_hk9	southwestern St. Paul	32	30	5
opdc_hk10	Lake Elmo	90	94	14
opdc_hk11	northeastern Hugo	103	99	17
opdc_hk12	northern North Oaks	129	125	22

Four Jordan pilot points dominate the base hydraulic conductivity in Metro Model 3 within the NMLG active model domain west of the St. Croix River. These pilot points have values of from 10.5 to 45 ft/d (Metropolitan Council, 2014, Figure 57). The depth-controlled multiplier varies from 1.0 at about 70 ft to a maximum of 1.6 at zero depth. Jordan hydraulic conductivities vary more smoothly than in the Prairie du Chien, from a low near the center of the northern subcrop to higher values toward the Mississippi and St. Croix rivers. There are also higher values in wider sections of buried valleys where Jordan subcrop is resolved in the 500-meter Metro Model 3 grid. For comparison, hydraulic conductivity of the Jordan was assigned values from 3 to 160 ft/d in the models of southern Washington County (Barr Engineering Co. and Washington Co., 2005; Barr Engineering Co., 2010).

For the steady-state NMLG model, 10 of the 12 pilot point values were close to one (1) with two values at approximately 0.85, resulting in limited changes to the spatial distribution of hydraulic conductivity. For the transient model, 8 pilot points were assigned values of approximately 0.79. The four pilot points in northern part of the domain were assigned values of approximately 0.32. This resulted in an overall reduction in hydraulic conductivity but with more reduction in the area with already lower values in Metro Model 3 and the steady-state NMLG model. The hydraulic conductivity values resulting at the 12 pilot point locations for the three models are shown in Table 7 as examples for comparison.

Although there is uncertainty in the distribution of hydraulic conductivity at all spatial scales, the hydraulic conductivities assigned to the Prairie du Chien in the transient model are systematically lower than previous models and generally lower than expected from the available aquifer-test analyses. The hydraulic conductivities assigned to the Jordan appear to be reasonable across most of the domain, except the values in the northern part may be too low.

The transmissivities in Metro Model 3 at the aquifer-test locations are variably lower and higher than the corresponding estimates, but on average and at more locations the Metro Model 3 values are lower. There is significant variability between the differences, but this suggests that there is not a systematic high bias in Metro Model 3 hydraulic conductivities for the Prairie du Chien and Jordan aquifers.

Table 7 – Horizontal hydraulic conductivity values for the Jordan at NMLG model pilot point locations

Pilot Point Name	Location Description	Metro Model 3 (ft/d)	Steady-State NMLG (ft/d)	Transient NMLG (ft/d)
cjdn_hk1	western North Oaks	22	22	18
cjdn_hk2	central St. Paul	32	31	25
cjdn_hk3	southwestern Shoreview	28	27	22
cjdn_hk4	southeastern Lino Lakes	14	14	5
cjdn_hk5	White Bear Lake	22	22	17
cjdn_hk6	northwestern May Twp	17	17	6
cjdn_hk7	Big Marine Lake	29	29	10
cjdn_hk8	Lake Elmo	29	30	23
cjdn_hk9	southwestern White Bear Lake	24	25	19
cjdn_hk10	southeastern Hugo	17	17	5
cjdn_hk11	Newport	34	29	27
cjdn_hk12	southeastern Woodbury	37	31	29

Quasi-3D Confining Units

Quasi-3D confining units represent the effect of confining layers on the vertical resistance between model layers without explicitly representing them as separate layers. This capability was used to represent aquitards at the base of several bedrock hydrostratigraphic units: the St. Peter aquifer, the Prairie du Chien aquifer, and the Tunnel City aquifer. Like the bedrock-aquifer hydraulic conductivities, the NMLG model uses Metro Model 3 as a base or starting point for the hydraulic conductivities of quasi-3D confining units. In Metro Model 3, the thicknesses of the confining units were all set to 0.1 meters (Metropolitan Council, 2014). The hydraulic conductivities of the confining layers were assigned such that the resistance of the confining layer (thickness/hydraulic conductivity) represented the actual thickness combined with a reasonable vertical hydraulic conductivity.

According to Jones et al. (2017), the bedrock layer and confining unit elevations and thicknesses from Metro Model 3 were transferred to the NMLG model, and the confining units in the NMLG model also have a thickness of 0.1 meter. The thickness of the confining units in the NMLG model, however, is actually 0.2 meter. Presumably this resulted from an undetected error in the process used to translate the Metro Model 3 layer elevations to the NMLG grid. For the steady-state NMLG model, most of the pilot-point multipliers for the confining unit hydraulic conductivities were assigned values close to one (1). This resulted in an unintended doubling of the confining unit resistances except around two pilot points for the St. Peter aquitard (ostp_bk6 and ostp_bk7) that were assigned values of 3.25 and 1.42.

A typical resistance for the St. Peter aquitard in the steady-state NMLG model is about 220,000 days, which corresponds to a vertical hydraulic conductivity of 1.8×10^{-4} ft/d for an aquitard thickness of 40 ft. This is of the same order of magnitude as estimates for the vertical hydraulic conductivity of other fine clastic aquitards in the region (Runkel et al., 2003). For the transient model (SSPA, 2017), two pilot-point multipliers were applied to the St. Peter confining unit: 1.13 or 2.71. There are few estimates of the resistance of the St. Peter aquitard, such as from aquifer tests, but the St. Peter has traditionally been regarded as relatively leaky, at least in some areas. For example, the basal St. Peter is not explicitly defined as a confining layer in Minnesota Rules 4725.0100 (well code), but the Glenwood Shale, St. Lawrence Formation, and Eau Claire Formation are. The proportion of shale in the basal St. Peter is variable but is generally higher near the center of the Twin Cities basin such as much of Ramsey County (Runkel et al., 2003).

A typical resistance for the Prairie du Chien aquitard in the steady-state NMLG model is about 7,100 days, which corresponds to a vertical hydraulic conductivity of $5.6 \text{ X } 10^{-3}$ ft/d for an aquitard thickness of 40 ft. The resistance of the lower Prairie du Chien aquitard has been estimated from a few analyses of aquifer tests with values typically in the range of 200 to 5,000 days (e.g. Bonestroo, Rosene, Anderlik & Assoc., 2004; Champion, 2009, Blum, 2011, Blum, 2016a; Blum, 2016b). This suggests that the resistances used in Metro Model 3 are reasonable, but double those values (i.e. 7,100 days) may be too high. For the transient NMLG model (SSPA, 2017) two pilot-point multipliers were used for the lower Prairie du Chien aquitard: 0.746 or 2.24.

The deeper bedrock aquifers were not a focus of the NMLG model, and parameters applied to those models were not further evaluated.

2.2.2. Lakebed Leakance

Lakebed leakance is defined as the hydraulic conductivity of lakebed sediments divided by the thickness of the sediments. The horizontal and vertical hydraulic conductivities of adjacent groundwater-flow cells are used to calculate horizontal and vertical conductances between lake cells and adjacent groundwater-flow cells. Because leakance affects the conductance between the lake and groundwater, it represents a lumped parameter to adjust lake stage and bulk groundwater-lake exchange in practice.

Jones et al. (2017) divided the simulated lake beds into two or three leakance zones based on depth and/or information about the locations and thickness of low-permeability, organic deposits. Jones et al. (2017) defined a wide range of allowable lake-bed leakance values, but the low permeability zones were expected to have values one or more orders of magnitude less than the higher permeability zones. SSPA (2017) did not restrict leakance relationships for the lakebed zones, and the estimated leakance values were higher for the low permeability zone than the permeable zone for several lakes (White Bear, Big Marine, and Turtle).

During the parameter estimation process, adjustments to lake-bed leakance are affected by limitations in the spatial distribution of adjacent groundwater-domain hydraulic conductivities, model discretization, and errors in the spatial and temporal distribution of recharge. For example, thin deposits of low permeability materials beneath and adjacent to a lake may strongly influence lake-groundwater interaction. Where present, these materials may not be well represented in the distribution of horizontal and vertical hydraulic conductivities due to limitations in the Quaternary materials model and/or the vertical layering. Given the available information, however, it is likely that, bulk leakance is better represented with lower or equal values in deeper areas compared to shallow areas, as conceptualized by Jones et al. (2017).

2.3. Model Features and Processes

The transient NMLG model preserved most of the model features and processes in the steady-state NMLG model. Some features were necessarily added for transient simulation. A notable change was removal of the Unsaturated Zone Flow (UZF) package for the transient model. The effects of this on recharge timing were discussed in the transient NMLG model report (SSPA, 2017), but other effects are further reviewed here. Also, alternative representations of the stage-volume-area relationship and outflow rating curve for White Bear Lake are discussed.

2.3.1. Surface-Water Features

Wetlands are important hydrologic features in the northeastern Metro area, and wetland interactions with groundwater can be complex. Wetlands may serve as recharge and/or discharge features in relation to groundwater, and may vary their hydrologic function seasonally or over longer time periods. Groundwater may discharge to wetlands and drain over the surface via connected streams and ditches. Direct evapotranspiration (ET) from the water table also occurs in wetlands and other areas with a shallow water table. Groundwater evapotranspiration can be explicitly included in groundwater-flow models such as MODFLOW, but this transient process is dependent on an accurate representation of the depth to the water table and ET from any overlying unsaturated zone.

Larger lakes and open-water wetlands were represented in the NMLG model with the River package, which applies a boundary condition with specified head and a conductance between the feature and groundwater. These can act as either recharge or discharge boundaries. Some wetland complexes in the northern part of the area that were represented with the River package in Metro Model 3 were not included in the NMLG model. The steady-state NMLG model, however, employed the surface leakage capability of the UZF package, which removes water when the computed water table rises within a specified depth from the surface. This allows for groundwater discharge where surface-water features are not explicitly represented.

Similar to evapotranspiration, surface leakage is a transient process that depends on an accurate computed water table. For example, if the water table is modeled too high but within the range of typical model head errors, this can result in surface leakage being computed where there is not actually groundwater discharge. On the other hand, surface leakage provides a means to discharge water where groundwater evapotranspiration and surface leakage actually occur. One traditional, simplified approach to represent wetlands in regional MODFLOW models is to use the Drain package. This functions very similarly to surface leakage in UZF but is restricted to the model cells where the Drain package is specified.

To allow for tractable model run times and maintain numerical stability, the UZF package was not used in the transient NMLG model. Applying estimated recharge to the water table without routing through the unsaturated zone is reasonable across most of the model domain for the relatively long stress periods employed in the annual and triannual model versions.

Removing the UZF package, however, eliminates surface leakage. In the steady-state model, surface leakage was computed in a number of areas of the model, but it was most common near features explicitly represented with the River or Lake packages where discharge can occur without surface leakage. There were extensive areas of surface leakage in some other areas, however, such as along wetlands and streams/ditches tributary to Rice Creek.

Several small streams were represented with the River package in the NMLG model, but two significant tributary systems to Rice Creek were not included: Hardwood Creek and Clearwater Creek. Data collected by the Rice Creek Watershed district indicates that these stream systems have significant, perennial base flow (Emmons & Olivier Resources, Inc., 2009). The Clearwater Creek watershed includes White Bear Lake and Bald Eagle Lake. Bald Eagle Lake rises above its outlet and discharges to Clearwater Creek seasonally in most years. Computed heads in this area of the model domain were biased high in the transient NMLG model. This was likely caused by a combination of effects such as the distribution of hydraulic conductivity and recharge, but these missing surface-water features also contributed to the head bias.

The transient model may be improved by explicitly representing these stream systems with the River package similar to the way other streams are represented. Explicit representation of additional wetlands in discharge areas may also be warranted, such as with the Drain package, but appropriate features would have to be selected and created in the model.

2.3.2. White Bear Lake Geometry (Stage-Volume-Area Table)

Changes in lake stage are controlled by net volumetric inflows and outflows from the lake and the relationship between lake stage and storage volume. Direct precipitation and evaporation volumes are affected by the lake area, which also changes with stage. The stage-volume-area relationship also affects how groundwater-lake fluxes respond to hydrologic stresses such as changes in groundwater recharge and groundwater withdrawals.

The stage versus volume relationships for the NMLG model (Jones et al., 2017) were derived from digital bathymetric grids developed by the DNR (2014). These data were readily available and straightforward to apply, although there may be issues with tying these water-depth data to elevation. More detailed, LiDAR-derived elevation data can be combined with the bathymetry data to develop stage-volume-area relationships that are likely more accurate for stages within most of the historical range. Using the LiDAR data is not expected to have a large effect relative to other uncertain model elements, but the data provide an opportunity to improve the model. In the following paragraphs, potential sources of error in the NMLG stage-volume-area relationship are discussed, followed by development of a new relationship.

Jones et al. (2017) describe the process used to develop lake stage versus volume tables for lakes represented with the LAK package. In summary, digital bathymetry on a 5-m (16-ft) grid (DNR, 2015a) was aggregated to the 150-m (492-ft) model grid. The volume of the lake at 150 equal-depth intervals (number of depth intervals required by the Lake package) was then calculated from the derived, 150-m bathymetry grid. Elevations corresponding to zero depth were taken as the ordinary high water (OHW). Jones et al. (2017) do not document how lake areas for each stage were derived or how the volumes were extrapolated to stages above the zero-depth stage that was applied.

The lake stage-volume-area curve in the Lake package is independent of the model grid, and its purpose is to provide a more detailed and accurate representation of lake-geometry than the groundwater-flow model grid cells representing the lake. Aggregating the bathymetry to the coarser model-grid scale before calculating volumes removed some information. The method provided more detail in the vertical direction than the model grid, but not in the horizontal direction.

Another source of error in the derived stage-volume relationship is that the actual stage corresponding to zero depth in the bathymetry grid generally does not correspond to the OHW. According to the metadata for the bathymetry data (DNR, 2015a), the zero depths correspond to lake shorelines digitized from digital orthoquads (DOQ's) derived from 1991 aerial photography (USGS, 1992). Visual inspection of zero-depth outlines over the digitized photos shows that another source was used for the outline of some lakes, including White Bear Lake. This is likely because lake levels were below the relevant zero depth on April 17, 1991 when the photographs were taken. Bathymetry away from the shoreline was derived from water depths measured at the times of the bathymetric surveys. This means that there are discrepancies in the information used to delineate shallow depths (zero depth versus the shallowest data points used).

The date the bathymetry data were collected for each lake is included with the distributed data, as described in the metadata. For some lakes this date is missing or only provided as a year. If the date is provided and a stage measurement on or near that date is available, that measurement provides a more accurate value for the zero-depth stage that corresponds to the bathymetric survey data and most of the bathymetry grid. This was confirmed by comparing the bathymetry grid values to contours digitized from DNR bathymetric contour maps. The contours represent the depth below the lake level at the time of the bathymetric survey.

The NMLG model is initially being applied to evaluations for White Bear Lake, and the stage-volume-area relationship for that lake was specifically evaluated. The bathymetric survey of White Bear Lake was completed in May 1978. Lake stage measurements varied from 922.79 to 923.18 ft (MSL, 1912) during that month. Therefore, the zero elevation for the bathymetric survey was approximately 923 ft MSL, 1912 (923.5 ft NAVD88) compared to the OHW of 924.89 ft MSL, 1912 (925.35 ft NAVD88).

The lake area versus stage relationship in the NMLG model can be readily compared against topographic contours and aerial photography. During the development of the protection elevation for White Bear Lake, DNR staff computed the area encompassed by topographic contours derived from LiDAR data (DNR, 2014) at elevations of 920, 922, 924, and 926 ft (NAVD88). The LiDAR data is capable of producing 2-foot contours with 95% accuracy confidence and a root mean square error (RMSE) of point elevations less than 0.61 ft.

These areas are compared to the corresponding areas in the NMLG model in Table 8. The planimetered area of White Bear Lake in Bulletin No. 25 (DNR, 1968) was 2,410 acres based on a 1949 aerial photo. Water levels in White Bear Lake varied between approximately 924 and 925 ft (NAVD88) in 1949.

Table 8 – Area of White Bear Lake, LiDAR derived elevation contours versus NMLG model

Elevation (ft, NAVD88)	Contoured Area (acres)	NMLG Model Area (acres)
920	2,258	1,947
922	2,392	2,137
924	2,434	2,295
926	2,473	2,606

To further evaluate the LiDAR-based contours and DEM, DNR staff compared the apparent lake shoreline at several locations on digital aerial photography to lake stage measured on dates close to the dates the photos were taken: 2016 (Metropolitan Council, 2016, 2013 (Washington County, 2013), 2012 (USGS, 2012), 2010 (DNR and Metropolitan Council, 2010), 2006 (USGS, 2007), and 1997 (Metropolitan Council, 1999). The photo times correspond to lake stages from 920.5 to 925.1 ft (NAVD88). Each of the digital aerial photography data sets were generated from multiple flights over periods of less than one month.

In general, the orientation of the LiDAR contours followed the apparent shorelines on aerial photography. Contours generated from the bathymetric grid often did not closely follow the apparent orientation of the shoreline. The elevation of the apparent shoreline in the 1-meter DEM varied over a range up to two (2) ft but typically within about one (1) ft from location to location. The differences are due to a combination of uncertainty in identifying the position of the shoreline on the photography, errors in the DEM, and possibly different flight dates covering different parts of the lake. The DEM elevation at the interpreted shoreline was typically 0.2 to 1.5 ft lower than the measured stages over the periods of the flights. It is possible that the visually identified shorelines were biased toward actual non-zero depths or that the LiDAR-based DEM is biased slightly low at White Bear Lake.

A new stage-volume-area table was developed for stages above 918 ft MSL, 1912 (918.46 ft NAVD88) by combining the values from DNR (1998), with values at higher stages calculated from the LiDAR contours. The DNR (1998) values were based on five-foot contours of the bathymetry data and a zero elevation of 923 ft MSL, 1912. Volumes at each contour level were calculated using the conic method (e.g. HEC, 1998). To provide the necessary detail for computations in the Lake package, a cubic spline through the points at the contour levels was used to calculate volume and area values at 0.5 ft intervals. The resulting stage-volume and stage-area curves are shown in Figures 10 and 11.

An important feature of the lake stage-area relationship is that, as stages fall below approximately 921.5 ft MSL, 1912 (922 ft NAVD88), the lake area decreases steeply because the lake bed has a low slope in several areas within this elevation range. The decrease in lake area during periods of low lake stages is well documented and visible in aerial photographs. This appears to be better reflected in the new stage-area curve. The change in volume per change in lake stage, the slope of the stage volume curve, controls stage-volume interaction in the model. The slope of the new stage-volume curve is similar to slope of the USGS curve, but the new curve is slightly steeper except at stages exceeding about 926.5 ft, MSL 1912 (927 ft NAVD88).

2.3.3. White Bear Lake Outlet Rating Curve

The lake stage versus outflow rating curves were not used directly in the steady-state NMLG model. Outflows were calculated separately and entered as model input (Jones et al., 2017). For White Bear Lake, the actual and computed average stage during the period represented by the model (2003 through 2013) was below the outlet invert. Jones et al. (2017) used the outlet rating developed by the DNR (DNR, 1998) to estimate total surface-outflow volume from White Bear Lake during the steady-state model period.

The transient model required an outlet rating. SSPA (2017) applied a simplified rating that resembles the DNR (1998) rating and extended the rating to higher stages (Figure 12). The simplification of the rating curve has a relatively small effect on computed stages for White Bear Lake over the focal period for the transient models (1988 through 2016) because the computed lake stage was below the outlet for nearly this entire period. Also, the relatively long stress periods of either one year or four months smooth out shorter-term stage fluctuations that affect actual outflows. Nevertheless, it is straight forward to apply the DNR (1998) rating to White Bear Lake in the transient NMLG model. The DNR (1998) rating was entered into the alternate, transient NMLG model discussed later in this report.

3. Model Testing and Feature Modifications

After reviewing the transient NMLG model inputs and results in detail, DNR staff undertook two lines of model testing and analysis. Several input stress changes were explored through a series of tests based on the original model. The effects of selected model assumptions and estimated parameters were also further evaluated through development of a revised model configuration with modified parameter values and some feature modifications. This section describes the set of model tests. The revised model configuration is described in the Section 4.

One set of tests evaluated the effect of the stage-volume-area curve for White Bear Lake on computed lake stage variations and computed pumping impacts. Another set of tests explored the sensitivity of computed lake stages to input stresses applied during the initial, steady-state and warm-up (1981-1987) periods. The final test isolated the effect of applying an alternative data set for lake precipitation.

Evaluation of the test results focused on the impacts of the selected input changes to the computed hydrograph of White Bear Lake. Model output from these tests could also be reviewed at other locations at a future time. The tests are useful to explore the sensitivity of the model to several modeling choices, but they do not represent alternative or "re-calibrated" models. The estimated model parameters depend to varying degrees on the combination of all inputs. Some parameter estimates would have differed had the tested inputs been applied during parameter estimation.

3.1. Model Tests

Several tests were conducted to evaluate model inputs and features. The effects of these tests at other locations could also be examined using the generated model outputs.

3.1.1. Stage-Volume-Area Curves and Outlet Rating for White Bear Lake

For this test, the new stage-volume-area table and an outlet rating that more closely matches the DNR (1998) rating were used for White Bear Lake (See 2.3 Model Features and Processes above). All other inputs were maintained at their initial values.

These changes caused an overall increase in computed lake stage but also affected lake-stage variations over time. Computed lake levels were 0.13 to 0.81 ft higher than the original annual model (Figure 13) and 0.20 to 0.87 ft higher than the original triannual model (Figure 14). The differences generally increased over time from the smallest difference in 1987 to the largest difference in 2013. Except at the highest computed lake stages, the lake area is larger in the new table. Because precipitation is typically greater than evaporation, this increases the net amount of water entering the lake surface. The higher computed lake stages with the new stage-volume-area curve are consistent with increased input volume.

The differences between these computed hydrographs should be qualified by the caveat that the estimated model parameters were affected by the stage-volume-area table. Computed-lake stage is sensitive to lakebed leakance. Had the new stage-volume-area table and modified outlet rating been used for White Bear Lake in the original model, its estimated conductance values would have been affected. Higher conductance values generally decrease computed lake stage because computed groundwater heads are less than computed lake stage beneath most of the lake.

The lake stage-volume-area table affects the lake response to input stresses. Therefore, these data inputs affect computed impacts for the hypothetical pumping scenarios that were tested by SSPA (2017). To demonstrate this, the 25% shutoff scenario (75 percent of reported pumping from 1988 through 2016) was run with the modified annual model. The computed stage differences resulting from this pumping scenario for the original model and the modified model are shown in Figure 15. The computed stage differences are smaller in the modified model, with the maximum stage difference in 2013 at 2.2 ft in the modified model versus 2.7 ft in the original model. Again, a combination of the new table with a corresponding (i.e also modified) lakebed leakance would have different computed impacts.

3.1.2. Steady-State and Warm-Up Period Inputs

This set of tests explored the sensitivity of computed heads for the whole model period (1981-2016) to input stresses applied during the initial, steady-state (pre-1981) and warm-up (1981-1987) periods. The initial steady-state stress period provides a stable, internally consistent set of initial heads for the transient simulation, but steady-state conditions never actually occur in the real system. The analyses described in the model report (SSPA, 2017) suggested that changes in model stresses have residual impacts on computed heads for a long period of time.

The warm-up period was made as long as possible with the data sets used to develop the NMLG model. Pumping volumes prior to 1988 were estimated because earlier records had not been digitized. The Daymet precipitation and temperature data begin in 1980. The earliest National Land Cover data set is for 1992, and there are basic differences between the 1992 and later data sets in addition to the errors/differences between each subsequent data set.

Information on pumping and land-use for earlier periods is only available as original or scanned documents and maps. An alternative dataset of gridded, daily precipitation and temperature formatted for input to SWB could be developed with some effort, but changes to the station network and network density may become more significant issues for earlier periods. Finally, fewer groundwater-level data are available for earlier periods, limiting the ability to test the validity of input data sets and modeled conditions.

An earlier starting time for the warm-up period would be desirable, ideally before significant urban development occurred in/near the model focus area. Model errors during the earlier period would likely be large, however, relative to the amounts of groundwater and lake-level variations that are of interest in the upper aquifer systems.

Groundwater-flow simulations of long periods with limited or approximate data are more practical and meaningful for large, well-confined groundwater systems which have experienced large and persistent drawdown trends and that are less sensitive to variable near-surface processes. For example, Feinstein et al. (2005) simulated the southeastern Wisconsin groundwater-flow system over a 137-year period using pumping estimates but holding recharge constant. The focus was on deep, well-confined sandstone aquifers that have experienced hundreds of feet of drawdown. Even in that case, groundwater-level datasets dating to the 1940s were important.

Given the lack of readily available data and other limitations, the model simulation period will not be extended at this time. Therefore, it is important to gain an understanding of the sensitivity of model results during the focus period (1988-2016) to input stresses applied during the steady-state (pre-1981) and warm-up (1981-87) periods. Input stresses during these periods were modified, and the initial evaluation focused on the effects of these changes on the computed water levels of White Bear Lake. Pumping changes to selected wells based on improved data and estimates were made for the steady-state and 1981 through 1987 periods. Recharge, lake precipitation, lake evaporation, and runoff to lakes were also modified for the initial, steady-state period.

Groundwater-use volumes reported by permit holders are available in database form beginning in 1988. Earlier reports were previously available only as scanned paper records, and some of these older records are also less complete. For the transient NMLG model, pumping rates for the steady-state period were estimated as 50 percent of the reported 1990 value, and 1981 through 1987 rates were estimated as the reported 1990 value for most wells that existed at the time. The year 1990 was selected because 1988 and 1989 were drought years with unusually high water use.

SSPA (2017) assigned non-zero pumping rates in the 1980 through 1987 period to a small number of wells that had not yet been constructed because wells were lumped by permit in the process they used to determine when to add a well to the model. For example, if a well under a given permit was installed before 1980 (i.e. before the start of the modeled period), a different well that was installed in 1984 under the same permit may have been added to the model starting in the steady-state period instead of leaving it out until 1984. For this test and in the revised model, wells were not added to the model until the year they were constructed based on the construction date given in the CWI database.

Metropolitan Council staff transcribed scanned water-use reports for municipal wells to spreadsheets for the period 1980 through 1987 and provided the data to DNR (John Clark, pers. comm., 2018). Wells are identified by their local names as indicated in the scanned permit reports (e.g. Well #1). These records are mostly complete for municipal permits. Permit reports for other types of water uses were not transcribed.

Connecting data records to unique well identifiers and reformatting would be required to transfer these data to the groundwater model. As a first step, and to capture the data that might most affect analyses of White Bear Lake and vicinity, a subset of the data were processed for input into the annual model. This data set included best available reported/estimated annual use volumes for 1980 through 1987 for all wells under the 9 permits computed to have the largest impacts on White Bear Lake in the 1988-2016 scenario analyses. Eight of these were municipal, and the other (food processing) permit was estimated based on the permitted volume at the time. Model runs using the revised steady-state and 1981-87 pumping rates will be referred to as a) SS and 1981-87 Q.

For most wells, the pumping rate applied during the initial, steady-state stress period was set to the 1980 rate. The steady-state pumping rate was reduced by 50 percent for St. Paul Regional Water Services wells that first went into operation in 1978 or 1979. This was done because the effects of the brief operation of these wells were far from reaching steady state in 1980.

The model-wide, total pumping volumes with the partially amended data set were 74 to 85 percent of the totals applied by SSPA (2017) in these years. For most of the nine modified permits, the original estimates exceeded the revised volumes in 1981 through 1987, but not for all. For example, reported withdrawals from the City of White Bear Lake's three active, Prairie du Chien-Jordan wells varied from 70 to 107 percent of 1990 pumping volumes during this period. Because pumping from each well has a different impact on the system at any given time and location, the net effects of these pumping changes in the model can be complex.

The steady-state period was used to approximate heads and lake levels at the end of 1980, which were affected by climate and groundwater use over a number of previous years. Therefore, the most representative model stresses to apply to the steady-state period would not reflect just the year 1980. In the original transient model, recharge calculated for 1990 was applied to the steady-state, initial stress period because that was an approximately average recharge year and was also the year used to estimate pre-1988 pumping rates. Direct precipitation to and evaporation from the lakes for 1980 was applied in the LAK package, however. In general, the initial steady-state period is an imperfect initial condition for the transient simulation, which is an important reason the following 7 years were considered a warm-up period. The magnitude and length of the residual effects of the initial steady-state period were further evaluated by running test simulations with modified steady-state recharge and lake precipitation and evaporation.

Groundwater levels rose during the 1970s in areas of the eastern Metro distant from groundwater pumping centers (Schoenberg, 1984), but recharge in the years before 1980 may have still been less than the 1981-2010 average. The five-year moving average precipitation at White Bear Lake (HIDEN) at the end of 1980 was 90 percent of the 1981-2010 average. This may indicate that the groundwater recharge for the steady-state period should be somewhat less than the 30-year average. The computed recharge in 1990 was 107 percent of the 1981-2010 average. Note that the net, 1980 precipitation plus runoff minus evaporation applied to the lake surface in the steady-state stress period in the original model was less than average.

The averages of the 1981 through 1983 recharge, lake precipitation, lake evaporation, and runoff were used as alternative inputs for the steady-state period. These three years represent a relatively wide range of recharge amounts. The precipitation average was close to the 1980-2016 average, and the recharge was about 90 percent of the 1980-2016 average. Model runs with these modified climate inputs for the steady-state period are referred to as b) SS RCH, P, E, and RO. Model tests were also run for: c) just the modified steady-state lake inputs (SS P, E, and RO), d) with just the modified steady-state recharge (SS RCH), and with all of the modifications (All).

Computed hydrographs for White Bear Lake for the original annual model and the model with modified inputs are shown in Figures 16 and 17. To demonstrate the way the different input changes combine, the average of the outputs from pairs of runs that each applied a subset of the modified inputs are also plotted. These averages can be compared against the corresponding test in which the input modifications were combined. For example, the average of a) SS and 1981-87 Q and b) SS RCH, P, E, and RO can be compared with All, which made all the changes in a) and b).

For the test that applied all of the input modifications (Figure 16, All), the impact of the changes on computed lake stage is relatively small (maximum of 0.3 ft). The computed stage for the steady-state period (plotted as 1980) was nearly the same, reflecting a balance between increased precipitation minus evaporation, decreased recharge, and variably different pumping rates. The All test lowered the computed stages from 1981 through 1985, but they increased the computed stages from 1987 through 2002. The variable impact over time is notable; the tests of different combinations of the input changes provide stronger demonstrations of the transient response of the model.

As mentioned above, the modified pumping rates for the warm-up period were generally less than the 1990 reported rates. As expected, applying the modified steady-state and 1981-87 pumping rates (test a) raised the computed lake stage during this period (Figure 16). The difference in computed stage was relatively small while the computed lake stage remained above the outlet into 1987, but the difference then increased, peaking at 1.4 feet in 1994 (Figure 18). The difference was still 0.5 feet in 2015 despite unchanged inputs after 1987.

Simultaneously modifying all the steady-state climate inputs (test b) had a minimal effect on the steady-state lake stage, which was computed above the outlet elevation. The higher lake inputs offset the effect of lower recharge in the steady-state period. But, the changes during the steady-state period lowered the computed lake stage after the initial period (Figure 16). The computed stage for test b) was 1.14 feet lower by 1997 and remained 0.67 feet lower in 2015 (Figure 18). Changing only the steady-state lake inputs to a larger net input (test c) increased the computed steady-state lake stage, but the impact was minimal by 1983 (Figure 17). This is not surprising given the local nature of the input changes for test c).

Changing only the steady-state recharge (test d) lowered the computed lake stage throughout the simulation period (Figure 17). The impact on computed lake stage varied over time, peaking at 2.6 feet lower in 1982 but remaining 1.0 foot or more lower through 2011 (Figure 18). Note that the computed stages for tests b) and d) gradually converge (Figure 17) and differ by only 0.09 feet by 2016. The average of test c) and d) would match test b) if the mode behaved linearly, but the c and d average differs somewhat from b (Figure 17).

3.1.3. Lake Precipitation

Spatially averaged precipitation and precipitation at White Bear Lake appears to be biased high in Daymet relative to other interpolated precipitation data sets. This affects computed recharge and runoff and precipitation applied to lakes. The annual model was tested with precipitation to White Bear Lake from the gridded HIDEN data set (State Climatology Office) instead of Daymet. Isolating the lake precipitation data set in this way exaggerates the effect of precipitation bias by making lake precipitation incongruent with precipitation used in the SWB model to compute groundwater recharge and runoff. Nevertheless, this test provides an instructive demonstration of sensitivity to lake precipitation.

The results are shown in Figure 19. As expected the test simulation using gridded HIDEN for direct precipitation computed lake stages that trend lower over time relative to the original model. The effect is small during the steady-state period in which a high lake stage was computed in both the test and the original model, but the difference increases over time. The maximum difference between the two models (1.8 feet) occurs in 2011.

4. Revised Model Development and Analysis

DNR revised the transient NMLG model, beginning with a different approach to parameterizing the SWB model that resulted in higher average recharge rates than the rates computed for the initial transient model version. Additionally, changes to selected model features based on new or better information were tested and incorporated in the revised model. For example, the newly developed stage-volume-area table described previously replaced the original table in the revised model. The revised model parameters were adjusted to achieve a roughly comparable level-of-fit to observed groundwater heads and lake stages, while placing tighter constraints on the range of possible hydrogeological properties.

This work was initially completed prior to the final White Bear Lake evaporation estimates (Xiao et al., 2018) becoming available. The revised evaporation values were higher and reduced simulated lake levels but had a relatively small effect on the fit to groundwater heads. To incorporate the revised evaporation estimates, lakebed leakance values in the model were again adjusted.

The resulting revised model is more tightly constrained to match the system conceptual model and also incorporates new and/or improved data. Developing the revised model and comparing to the initial, SSPA model demonstrate how some model choices and assumptions and forcing input data/measurement errors influence parameter estimation, model behavior, and computed impacts.

The revised and original models do not represent the entire range of reasonable models. A complete, quantitative analysis of predictive error that accounts for all significant model uncertainties (forcing inputs/stresses, model structure, and parameters) is challenging for simpler models and is computationally infeasible for the transient NMLG model. The model tests and development of the revised model, nevertheless, provide useful insights into uncertainties in model predictions and provide additional information to support management decisions.

4.1. SWB Model, Recharge and Runoff

SSPA (2017) estimated/assigned a smaller number of distinct parameters in the SWB model than were used by Jones et al. (2017) or Metropolitan Council (2014) by lumping/grouping some parameters. As observed by SSPA (2017), the more detailed parametrization can introduce errors due to errors in the land-use and other data sets and structural deficiencies/limitations in the SWB model. One problem is that inconsistencies in the land-use data sets result in changing parameters over time for areas that actually remained unchanged. The parameter lumping approach of SSPA (2017) removed this time bias in some areas but not all locations. For example, some areas with single family homes are categorized variably as low intensity residential, forest, or pasture/hay in different NLCD data sets. Oversimplification likely introduces its own set of structural errors.

4.1.1. SWB Parameters

A larger set of parameters more similar to those applied by Metropolitan Council (2014) was used for the revised SWB model. Metropolitan Council (2014) used similar data to define soil types and available water capacity, but they used different land-use data sets. No changes were made to the soil and land-use data used in the NMLG model. The revised SWB parameterization is expected to have somewhat differing errors and biases from the SWB model of SSPA (2017). Similar to the parameter-estimation approach taken by SSPA (2017), SWB parameters were assigned based on a small number of model runs. The parameters of the SWB lookup table are listed in Table 9.

Potential evapotranspiration (PET) is a key driver of computed groundwater recharge. Jones et al. (2017) used the Hargreaves and Samani (1985) option for calculating PET from daily maximum and minimum temperatures. SSPA (2017) made no changes to the PET calculation parameters in SWB. Jones et al. (2017) applied a "slope" coefficient of 0.0027, higher than recommended values of 0.0022 to 0.0023 (Hargreaves and Samani, 1985; Hargreaves, 1994). Using the method of Allen (1995), a coefficient of 0.00226 was used for the revised model. This reduced the computed PET by 16 percent. The resulting decrease in computed actual evapotranspiration (AET) increased computed groundwater recharge.

For the revised model, runoff curve numbers (CN) for each land-use/soil combination (14 land uses and 5 soil types) were set based on USDA (1986) and held fixed. This contrasts with the reduced set of five different curve numbers specified only by land use applied by SSPA (2017).

SSPA (2017) applied a maximum daily recharge rate of 3.6 inches per day for all land uses and soil types to account for small closed depressions. Metropolitan Council (2014) applied even higher maximum recharge rates for type A and B soils (9 and 5.5 inches per day) in the SWB model for Metro Model 3 but applied lower values for type C and D soils (2.4 and 0.7 inches per day).

For the revised model, the maximum recharge rates varied by soil type, as in Metropolitan Council (2014), but they vary over a narrower range (1 to 6 inches/day for all but developed land uses). Developed areas have generally been graded to produce better drainage. Storm ponds and small wetlands within developed areas are typically large enough to register as at least one, separate 125-m cell in the land-use grid (i.e. are typically not assigned a developed land-use type). A high maximum recharge rate appeared to produce excessive recharge in grid cells with developed land uses, particularly in areas with higher density and/or heavier soils. Maximum recharge rates were reduced for low, medium, and high intensity developed areas.

Root zone depths and soil-water holding capacities control the amount of available water storage and the scaling of AET versus PET in the SWB model. Because moisture storage is a lumped quantity in the SWB model, the most appropriate root-zone depths may not correspond to the actual, maximum rooting depths. SSPA used root-zone depths from Metropolitan Council (2014) for most land use types, but they increased these values for the shrub/scrub, grassland/herbaceous, and pasture/hay types. For most land-cover types, effective root-zone depth is expected to be largest for type B soils and smallest for type D soils, although type A soils may have the deepest root zones for forest (Westenbroek et al., 2010; Smith and Westenbroek, 2015).

Table 9 – SWB model lookup-table parameters according to land-use class and hydrologic soil group

Land-use class	Description	Curve Number by Soil Group [1]					Maximum Recharge (inches/day)by Soil Group					Root Zone Depth (ft) by Soil Group				
		A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
21	Developed, Open Space	39	61	74	80	70	6	3.5	2.4	1	3	3.33	4.17	3.33	2.22	3.33
22	Developed, Low Intensity	57	72	81	86	77	3	1.5	1	0.4	0.4	1.67	2.08	1.33	1.33	1.33
23	Developed, Medium Intensity	77	85	90	92	85	1.5	1	0.5	0.2	0.2	1.33	1.39	1.33	1.33	1.33
24	Developed, High Intensity	89	92	94	95	92	1	0.75	0.4	0.1	0.1	1.33	1.33	1.33	1.33	1.33
31	Barren Land	77	86	91	94	90	6	3.5	2.4	1	3	0.5	0.5	0.5	0.5	0.5
41	Deciduous Forest	30	55	70	77	63	6	3.5	2.4	1	3	6.66	6.66	5.33	3.9	5.33
42	Evergreen Forest	30	55	70	77	63	6	3.5	2.4	1	3	6.66	6.66	5.33	3.9	5.33
43	Mixed Forest	30	55	70	77	63	6	3.5	2.4	1	3	6.66	6.66	5.33	3.9	5.33
52	Shrub/Scrub	30	48	65	73	68	6	3.5	2.4	1	3	3.33	4.17	3.33	2.22	2.22
71	Grassland/Herbaceous	30	58	71	78	75	6	3.5	2.4	1	3	3.33	4.17	3.33	2.22	2.22
81	Pasture/Hay	39	61	74	80	77	6	3.5	2.4	1	3	3.33	4.17	3.33	2.22	2.22
82	Cultivated Crops	64	75	82	85	83	6	3.5	2.4	1	3	1.67	2.08	1.67	1.33	0.83
90	Woody Wetlands	60	60	60	60	60	6	3.5	2.4	1	3	4.5	4.5	4.5	4.5	4.5
95	Emergent Herbaceous	60	60	60	60	60	6	3.5	2.4	1	3	4.5	4.5	4.5	4.5	4.5

[1] Base curve numbers for hydrologic soil groups for soil groups A, B, C, and D at antecedent runoff condition II (USDA, 1986). Soil group E is for wetland soils.

In the revised model, root-zone depths from Metropolitan Council (2014) were used for most land uses. The shallow root zone depths applied to developed land uses, especially for heavier type C and type D soils resulted in unexpectedly high recharge in those settings. SSPA (2017) applied an artificially high interception fraction in developed areas. For the revised model, root-zone depths were increased for medium and high intensity developed areas and also for type D soils in low intensity developed areas. Root zone depth was also increased for crops on type C and D soils. These changes increased AET and decreased recharge in these areas.

4.1.2. Precipitation Recharge and Runoff

Average precipitation recharge computed for the period 2003 through 2013 for the Jones et al. (2017), SSPA (2017), and revised models are listed in Table 10. These are the average recharge rates applied in the MODFLOW model scaled to the total active model area. Note that a recharge multiplier of 1.75 was applied to SWB output in the steady-state model, and recharge to cells computed to be saturated by MODFLOW was rejected as surface discharge in that model. The recharge rates were calculated using the total recharge in the MODFLOW volumetric budget output and the area of the active model domain. Mean recharge computed by the revised SWB model for the entire simulation period (1980 through 2016) is depicted in Figure 20 to illustrate the spatial distribution of recharge.

Table 10 – Spatially averaged, 2003 through 2013 precipitation recharge computed for the three versions of the model

Jones et al. (2017), inches/year	SSPA (2017), inches/year	DNR Revised, inches/year
6.45	4.78	6.32

Runoff rates computed by the SWB model were used to apply runoff to the six lakes simulated using the Lake Package in MODFLOW. Jones et al. (2017) applied runoff multipliers to adjust the SWB-computed recharge for each lake. These multipliers varied from 0.61 to 1.5.

The runoff extracted from SWB and applied to the lakes by Jones et al. (2017) includes runoff originating in cells representing the lakes. Jones et al. (2017) used separate SWB runs to calculate runoff versus recharge. For the runoff calculation, all cells containing water bodies were assigned to the land-use category representing open water, which was assigned a runoff curve number of 84.4. Direct precipitation on the lakes was applied directly to the lakes in the Lake package, however, and any runoff computed by SWB originating within the lakes should not have also been applied to the lakes.

For example, the contributing drainage area used for White Bear Lake in the model, excluding the lake itself, includes 879 cells (3,394 acres). The runoff applied to the lake in MODFLOW, after applying a multiplier of 0.61, was 9578 m^3/day, which is 10 inches/year over the drainage area and 15 inches per year over the computed lake area (2,211 acres). This amount of runoff is unreasonably high and greatly distorts the lake water budget. For example, using a larger watershed area of 4,995 acres and a relatively high base curve number of 82, DNR (1998) estimated annual runoff amounts to White Bear Lake in the 1980s from 2 to 4 inches over the watershed area (4 to 8 inches over the lake area).

SSPA (2017) excluded runoff originating within all lakes from the summation of watershed runoff applied in the Lake package. The same approach was continued for the revised model. This excludes runoff from lakes represented with the River package, but the surface hydrology of these lakes is not represented in the model. Some small lakes and wetlands may capture runoff under most conditions but release runoff under high water-table conditions or after large runoff events.

The drainage area used for White Bear Lake in the model (3,394 acres) excludes the drainage areas for Echo, Long, and Lost lakes and a wetland downstream from Long Lake. These lakes are represented by the River Package in MODFLOW, but any surface flow out of the lakes is excluded from the model. This contributing drainage area is smaller than the drainage area of 4,995 acres used in the earlier work by DNR (1998). DNR (1998) acknowledged that the total drainage area may include non-contributing areas. The water level in Long Lake is above the outlet elevation most of the time. Some of the outflowing water originates from groundwater, but, given its relatively large drainage area relative to the lake area of about 53 acres, a large portion of the outflow likely originates as runoff. Direct precipitation on the lake also induces outflow (i.e. surface runoff). The outflow from Long lake flows through several wetlands, but some of this outflow likely reaches White Bear Lake. These processes cannot be explicitly represented in the SWB model.

The standard curve number method was developed for flood analysis and may not work well for representing smaller events in continuous simulations (Woodward et al., 2003). The USGS (Trost et al., 2017)applied the “Hawkins” option in SWB that uses a smaller initial abstraction term combined with a corresponding adjusted maximum storage in the runoff calculation (Westenbroek, 2010). This results in more runoff for small events and low curve numbers but a rainfall-runoff relationship close to the original formulation for other conditions (Woodward et al, 2003). The continuous frozen ground index option to increase curve numbers on frozen soil was also applied. Both of these options were maintained in the SSPA (2017) and revised SWB models to address some of the limitations of the curve number method for use in continuous modeling.

Average runoff to White Bear Lake computed for the period 2003 through 2013 for the steady-state (Jones et al., 2017), transient (SSPA,2017), and revised models are listed in Table 11. Recall that runoff was computed differently for the steady-state model because it included runoff originating from precipitation falling directly on the lake and was then multiplied by 0.61. The runoff computed by the SWB models from the contributing watershed appear low, but there are few measurements of annual surface-runoff volumes from similar catchments lacking any well-defined stream. The water budget analyses for White Bear Lake suggest that the runoff response of the watershed was highly variable under different historical conditions.

Table 11 - Average 2003 through 2013 runoff to White Bear Lake computed for the three versions of the model

Jones et al. (2017), inches/year over lake area (cfs)	**SSPA (2017), inches/year over lake area (cfs)**	**DNR Revised, inches/year over lake area (cfs)**
15.4 (3.9)[1]	0.69 (0.17)	0.25 (0.07)

[1] The values for the USGS model include runoff originating within lakes followed by multiplication by 0.61. The values extracted from the SWB model in the same way as was done by SSPA (2017) are 0.59 inches/year and 0.23 cfs. This does not include surface leakage computed by the UZF package within the lakeshed, which was also added to the LAK package.

The daily runoff routing feature in SWB appears to route most initially computed runoff to cells where it infiltrates on the same day as the precipitation or snow melt. As explained above, the maximum recharge rates were set to allow relatively high amounts of infiltration representing both areas of concentrated infiltration and surface storage. The maximum recharge rates were set lower for developed land uses in the alternate model, but that change appears to have had only a small effect on computed runoff to White Bear Lake.

Standard runoff curve numbers (USDA, 1986) were derived for entire small catchments from runoff measurements for large storm events. The curve number technique was developed for analysis of flood risk and for stormwater design in un-gaged, small catchments and continues to be one of the most commonly used methods for those purposes. It may be that infiltration of runoff in downstream cells, as computed in SWB, reduces runoff too much using standard curve numbers. Using higher curve numbers to increase the ultimate computed runoff would reduce the computed recharge, however.

Recognizing the limitations of the curve number method as applied in SWB, Jones et al. (2017) attempted to simultaneously estimate groups of curve numbers and root zone depths by calibrating to estimated surface runoff volumes at stream gaging sites. SSPA (2017) discuss problems and inconsistencies in the resulting parameter set. This parameter set also did not produce sufficient recharge as reflected in the recharge multiplier of 1.75 used by Jones et al. (2017). Finally, the runoff in the White Bear Lake Watershed computed in the Jones et al. (2017) SWB model, excluding lake cells, was 0.59 inches per year (0.23 cfs), similar in magnitude to the runoff computed by SSPA (2017). These observations are made not to criticize previous studies but to demonstrate the difficulties and limitations that have been encountered in applying the SWB model to the study area's complex hydrology.

4.2. MODFLOW

The representation of several features and model parameters were modified for the revised model. First the feature modifications are described followed by the parameter estimation process and results.

4.2.1. Model Features

White Bear Lake

The new stage-volume-area table and more accurate outlet rating described in 2.3 Model Features and Processes were applied in the revised model. The most significant effect of this change was to increase the lake surface area to more accurate values.

The preliminary evaporation estimates developed by SSPA (2017) based on preliminary, eddy-covariance-based evaporation values developed by University of Minnesota researchers for parts of 2014 through 2016 were initially applied during development of the revised model. The revised, final eddy covariance-based observations and retrospective modeling (Xiao et al., 2018) were applied when they became available along with adjustments to the estimates lakebed leakance values.

Streams (Hardwood Creek and Clearwater Creek)

As noted previously under 2.3 Model Features and Processes, the Hardwood Creek and Clearwater Creek stream systems (both tributaries to Rice Creek) were not explicitly included in the NMLG model (Jones et al., 2017). Removing the UZF package from the transient model eliminated surface leakage. Surface leakage partially/approximately compensated for surface-water features that were not explicitly represented.

Computed heads are biased high across much of the Rice Creek watershed in the transient NMLG model. This bias is likely caused, in part, by missing discharge features including small streams/ditches, groundwater evapotranspiration in wetlands, and surface discharges in wetlands. The Hardwood Creek and Clearwater Creek stream systems were added to the revise model (Figure 21). The model may also benefit from representation of some wetland complexes in the Rice Creek watershed using the Drain (DRN) package or possibly one of the packages representing groundwater evapotranspiration, but the latter have not yet been developed for the NMLG model.

In the original NMLG model, GIS processing was used to generate a table listing the intersection of streams, rivers, and lakes with model cells. The table was used in processing scripts to generate input for the River package. Processing scripts were provided by the USGS to the DNR (P. Jones, pers. comm., 2017). DNR staff used geoprocessing in ArcMap software (ESRI) to intersect polylines representing likely perennial segments of the Hardwood and Clearwater creek systems with the NMLG model grid. The polylines were derived from DNR hydrography (DNR, 2015b). Minor modifications to the polylines were made manually based on review of digital aerial photography and to remove overlap with cells already representing ponds/lakes with the River package. The width of groups of stream segments were approximated through visual inspection of aerial photography. The model cells and corresponding stream areas in each cell were added to the table used in the generation of River package input.

River Package Conductance Zonation

In the steady-state NMLG model (Jones et al., 2017), the conductance of River package reaches was parameterized according to feature type (stream, major river, or lake) and the dominant, underlying geological material (7 material types) mapped to the model grid. SSPA (2017) modified the conductance parameterization system by over-riding this parameterization scheme at a majority of features. The overriding values assign higher conductance in major groundwater discharge areas but limit conductance to a maximum of 100 m^2/d (minimum resistance of 156 days) in areas where lakes were simulated as perched and/or recharged large volumes to groundwater. Two other special lake conductance zones were also applied in the River package.

For the revised model, the original data used to parameterize River package conductance were modified in a small number of locations to allow for fewer conductance value exceptions. For example, surficial materials around the Rice Creek chain of lakes are variably mapped as “peat” or sand (Meyer, 2007). These lakes were assigned to the “peat or modern lacustrine deposits” surficial geology category in the NMLG model (Jones et al., 2017). Because these lakes appear to be relatively well connected with the water table in a groundwater discharge area, the surficial geology category for these lakes was changed to “Des Moines Lobe sands” in the surficial geology raster used to create River package input, which was assigned a higher conductance.

In the transient NMLG model, the conductance of River cells representing the St. Croix River south of Marine on St. Croix was assigned a higher value than derived from the original parameterization by geological material type. This conductance appears reasonable, but the conductance assigned to the St. Croix River to the north was much too low. This resulted in computed groundwater heads tens of feet above the river valley bottom in this northern area. In the revised model, the conductance of these River cells was increased to compute reasonable heads in that area.

For the revised model, only one other special River conductance zone was maintained. This was for a cluster of lakes to the southeast of White Bear Lake where lake stages are 15 to more than 25 ft higher than nearby groundwater-head targets in the St. Peter and Prairie du Chien aquifers. Surficial geological materials are mapped as sand, but low-permeability near surface and lake-bed sediments may restrict the hydraulic connection of these features to shallow aquifers.

St. Peter Confining Layer in River Valleys

In Metro Model 3 (Metropolitan Council, 2014), the St. Peter Sandstone, including its basal aquitard, extends under the Mississippi River along a 3.3 mile reach near downtown St. Paul, but it is likely entirely eroded starting near the confluence with the Minnesota River (Mossler, 2013). Bedrock zonation was carried over from Metro Model 3 to the NMLG model. This disrupts discharge from the Prairie du Chien to the Mississippi River and causes computed heads in the Prairie du Chien and Jordan aquifers to be too high along this reach.

It appears that the St. Peter was included along this reach due to a combination of the relatively coarse grid resolution (500 meters) of Metro Model 3 and reliance on an earlier bedrock map (Mossler and Tipping, 2000). Water levels recorded in the CWI database and interpreted in Sanocki et al. (2009) indicate that heads in the Prairie du Chien and Jordan in this area are much closer to the river elevation. The effect is localized to the area near this reach of the Mississippi River, but this groundwater discharge area would be better represented by removing the St. Peter in this area of the valley.

For the revised model, the extent of the St. Peter in this area as mapped in Mossler (2013) was overlain on the model grid. Where the St. Peter is absent along this reach, the bedrock zone array was set to the value representing Quaternary sediments, and the hydraulic conductivity of the basal St. Peter confining layer was set to the value representing its absence.

The St. Peter is present beneath the Mississippi River gorge from just above St. Anthony Falls to the confluence with the Minnesota River (Mossler, 2013). Water levels recorded in the CWI database and interpreted in Sanocki et al. (2009) indicate that heads in the Prairie du Chien-Jordan are no more than 25 ft above the river elevation beneath and adjacent to the river along this reach. In the transient NMLG model, computed heads in the Prairie du Chien are 45 to more than 50 ft above the specific river stage in the reach below St. Anthony Falls. Computed heads along this reach are even higher in the steady-state NMLG model (Jones et al., 2017).

The high computed heads in the Prairie du Chien-Jordan aquifer system along the Mississippi River gorge are maintained by the relatively high resistance of the basal St. Peter confining layer in the model (See Section 2.2.1). Based on the available groundwater-head data, it appears that the basal St. Peter must be very leaky in this area. To try to improve the representation of the connection of the Prairie du Chien to the Mississippi River along the gorge, the hydraulic conductivity of the basal St. Peter confining layer was increased by a factor of 100 immediately beneath the Mississippi River in this reach in the revised model.

This decreased computed heads in the Prairie du Chien and Jordan by a relatively small amount (less than 3 feet). It may be that the effective hydraulic conductivity of the basal St. Peter is relatively high across a wider area than the width of the river. Because the Mississippi River valley is not the focus for the model, additional modifications to the basal St. Peter confining layer in this area were not tested.

4.2.2. Parameter Estimation

MODFLOW parameters were estimated in the revised model using a similar approach as was used for the original transient model. The most important change to the input that affected MODFLOW parameter estimation was the recharge calculated by the revised SWB model parameterization. Annual stress periods and the same groundwater head and lake level targets were used with the PEST (Doherty, 2016) parameter estimation software.

To ensure that relative values among different parameters maintained the expected relationships and to reduce the computational demands of parameter estimation, more parameters were tied together or held fixed. Tied parameters are forced to maintain a specified ratio. This approach reduced the number of estimated parameters and the corresponding flexibility of parameter estimation. The goal was to develop a revised model that could be compared against the original model using a reasonable set of parameters.

Parameter estimation was carried out over several steps to identify sensitive parameters. Parameters that were changed little during initial steps were removed from further evaluation. For example, bedrock hydraulic conductivity multipliers estimated at values close to 1.0 were eventually fixed at a value of 1.0. Parameter estimation focused on hydraulic conductivities of the Quaternary and Prairie du Chien based on issues previously identified with these parameters. River- and Lake-package leakance values were also relatively sensitive parameters that were estimated. Storage parameters were not re-estimated.

Improvement in model fit generally decreased with each estimation step in a PEST run. After the initial steps, parameter adjustments that pushed some parameters closer to the minimum or maximum made little improvement to the fit. These changes generally represent parameter over-fitting. This was considered in choosing the final parameter set. Final parameter values were rounded to two significant digits unless the results were highly sensitive to a parameter. It was confirmed that the final parameter set produced an objective function (sum of squared errors) within a few percent of the best PEST runs.

Lake levels were sensitive to lakebed leakance. Lakebed leakance values were refined as a last step of the parameter estimation process as was done for the original transient model. Computed water levels in Pine Tree and Snail lakes were very sensitive to leakance and could become unstable. Relatively low leakance values had to be applied to reasonably match lake levels in these two lakes, but relatively large head differences between the lakes and groundwater could cause instability in the lake-level solution.

The model-layer discretization and zonation of Quaternary material types did not allow the model to represent details of the shallow groundwater heads and gradients near these lakes that appear to be important to groundwater-lake interaction. Data in the CWI database indicate that the water table near the south shore of

Snail Lake was about 6 feet below the lake level. Water levels in shallow monitoring wells (~ 50 feet deep) near the lake were 15 to 30 feet below the lake level. This steep, downward gradient in shallow groundwater heads could not be represented in the model without addition of local hydraulic properties zonation and possibly increased vertical discretization. Head observations near Pine Tree Lake also indicate heads below the lake level (~ 14 feet) in buried Quaternary aquifers near the south end of the lake.

4.2.3. Estimated Revised Model Parameters

All the final parameters used in the revised model are listed alongside corresponding parameter values from Jones et al. (2017, separate Table 3) and SSPA (2017) in the Supplemental Table. Horizontal hydraulic conductivity of layers 1 through 8 is depicted in Figures 22 through 29. Layers 1 through 4 generally represent Quaternary materials and the Decorah-Platteville-Glenwood. Layers 5 through 8 generally represent the St. Peter, Prairie du Chien, Jordan, and St. Lawrence formations respectively. Where bedrock units have been eroded, Quaternary materials are present in these deeper layers. In the southeastern part of Washington County, bedrock faults offset the bedrock units. Bedrock zonation was derived from Metro Model 3, except for the modification to the St. Peter along the Mississippi River in St. Paul as described above.

The horizontal hydraulic conductivities of Quaternary deposits, the Prairie du Chien aquifer, and the Jordan aquifer for the steady-state and transient models were listed in Table 3, Table 6, and Table 7. These values are shown again below alongside the parameter values used in the Alternate model. Because the Quasi-3D confining layers were inadvertently made 0.2 meter thick instead of 0.1 meter thick in the NMLG model (Trost et al., 2017), the pilot point multipliers were adjusted accordingly. Pilot-point multipliers of 1.8 to 2.0 were used in the revised model.

Jones et al. (2017) divided each lake into two or three zones described as either permeable or low permeability. In the revised model, these relative permeability relationships between zones were enforced within a lake. Lakebed leakance values in the revised model are generally lower than but within the bounded range of values given in Jones et al. (2017). One reason lower leakance values were estimated for the revised model is that Jones et al. (2017) applied much higher runoff amounts to the lakes that included runoff generated from precipitation directly on the lakes (See 4.1.2 Precipitation Recharge and Runoff above). This double counted a large fraction of the direct lake precipitation and exaggerated total lake inflows. The steady-state model also included additional runoff computed as surface leakage in the UZF package within the lake watershed. Because the lakes act as recharge or groundwater flow-through features, high leakance was needed in the steady-state model to discharge the large input of surface inflow via the only possible outflow, groundwater.

Table 12 – Horizontal hydraulic conductivity values for Quaternary deposits (ft/day)

USGS Parameter Name	Description	USGS Lower	USGS Upper	USGS Value	SSPA Value	Revised Value
quat_hk1	loam to clay loam	0.51	13	2.3	61 [1,2]	3.6
quat_hk2	loam to sandy loam	14 [3]	41	36	129 [1]	18
quat_hk3	loam, silt rich; silt and clay	0.41	10	2.1	15	3.6
quat_hk4	loam to sandy clay loam	--	--	25 [3]	3.3 [2]	3.6
quat_hk5	sand and gravel	39	118 [3]	92	52 [2]	121
quat_hk6	fine sand	20	59 [3]	51	46	74
quat_hk7	sandy silt	3.0	12	7.9	3.4 [2]	12
quat_hk8	loam to clay loam (deeper than 60 feet below the land surface)	0.37	9.2	1.8	4.0	1.0
quat_hk9	loam to sandy loam (deeper than 60 feet below the land surface)	2.6 [3]	10	8.1	38 [1]	2.0
quat_hk10	loam, silt rich; silt and clay (deeper than 60 feet below the land surface)	2.1 [3]	8.5	3.9	26 [1]	1.6

USGS Parameter Name	Description	USGS Lower	USGS Upper	USGS Value	SSPA Value	Revised Value
quat_hk11	loam to sandy clay loam (deeper than 60 feet below the land surface)	--	--	4.6 [3]	26 [1]	1.6
quat_hk12	Decorah Shale, Platteville Limestone, and Glenwood Formation confining unit	0.33	8.2	1.6	4.0	3.3

[1] Value more than 50 percent greater than expected maximum value
[2] Unexpected relative parameter rank
[3] Reasonable value may be outside of range defined by this maximum or minimum value.

Table 13 - Horizontal hydraulic conductivity values for the Prairie du Chien at NMLG model pilot point locations

Pilot Point Name	Location Description	Metro Model 3 (ft/d)	Steady-State NMLG (ft/d)	SSPA (ft/d)	Revised (ft/d)
opdc_hk1	Roseville	70	71	11	53
opdc_hk2	St. Anthony	114	115	18	87
opdc_hk3	southern Woodbury	21	16	3	21
opdc_hk4	eastern St. Paul	36	33	6	36
opdc_hk5	White Bear Lake	65	65	11	65
opdc_hk6	southern Hugo	112	110	19	39
opdc_hk7	West Lakeland Twp	85	81	14	30
opdc_hk8	northeastern Maplewood	56	57	9	56
opdc_hk9	southwestern St. Paul	32	30	5	24
opdc_hk10	Lake Elmo	90	94	14	32
opdc_hk11	northeastern Hugo	103	99	17	36
opdc_hk12	northern North Oaks	129	125	22	142

Table 14 – Horizontal hydraulic conductivity values for the Jordan at NMLG model pilot point locations

Pilot Point Name	Location Description	Metro Model 3 (ft/d)	Steady-State NMLG (ft/d)	SSPA (ft/d)	Revised (ft/d)
cjdn_hk1	western North Oaks	22	22	18	18
cjdn_hk2	central St. Paul	32	31	25	25
cjdn_hk3	southwestern Shoreview	28	27	22	22
cjdn_hk4	southeastern Lino Lakes	14	14	5	10
cjdn_hk5	White Bear Lake	22	22	17	17
cjdn_hk6	northwestern May Twp	17	17	6	12
cjdn_hk7	Big Marine Lake	29	29	10	21
cjdn_hk8	Lake Elmo	29	30	23	21
cjdn_hk9	southwestern White Bear Lake	24	25	19	19
cjdn_hk10	southeastern Hugo	17	17	5	12
cjdn_hk11	Newport	34	29	27	27
cjdn_hk12	southeastern Woodbury	37	31	29	37

4.2.4. Revised Model Fit to Observations and Model Comparisons

Because the focus of the model is the upper parts of the groundwater system (water table through Jordan aquifer) and groundwater/surface-water interaction, only the fit to data from those upper units are discussed. Statistics of the fit of computed heads in the annual model to corresponding, observed annual average water levels are given in Table 15. The statistics are generally similar. The mean residual is slightly higher for the SSPA model, indicating more bias toward computed heads higher than observed. The ideal mean error is close to zero. The mean absolute error (MAE) is slightly lower for the revised model, but the root mean square error (RMSE) is slightly higher for the revised model. This suggests that, on average, head errors are slightly lower for the revised model but it has larger outliers that more heavily influence squared errors.

Table 15 – Statistics of fit of computed heads in the annual model to observed water levels in wells open to the QWTA through CSTL

Statistic	SSPA (2017) Model (feet)	DNR Revised Model (feet)
Mean residual (observed – computed)	-2.0	-0.77
Maximum residual	29.9	28.8
Minimum residual	-28.1	-45.7
Mean absolute error (MAE)	6.9	6.5
Root Mean Squared Error (RMSE)	8.7	9.2
Observed Range	256	256
RMSE/Range	0.034	0.036

Plots showing observed versus computed heads for each model version are shown in Figures 30 and 31. The plots generally show similar scatter around the one-to-one line. Heads computed at two St. Peter observation wells near Pig's Eye Lake in St. Paul (observed between 700 and 720 feet) were computed too high in both model versions, but the errors are larger in the revised model. These observation wells are near the bluffs along the river, and head gradients near the bluffs and in the river valley are not matched well. This appears to be a local problem near the river valley and bluffs.

To illustrate the spatial distribution of fit to heads, computed head residuals for the year with the most widely distributed target observations (2013) are shown in Figure 32. Residuals for all of the upper aquifers are shown together because the aquifer system is generally well connected at the regional scale. Residual patterns are similar for the different aquifers and better visualized together. The pattern is similar for both model versions within the resolution of the model residual categories shown. The bias toward high computed heads (negative residuals) to the north and northwest of White Bear Lake is larger in the SSPA (2017) version, but residuals also tend negative in this area in the revised model.

Again, computed residuals near the Mississippi River in St. Paul have larger magnitudes in the revised model. Computed heads in the Prairie du Chien-Jordan aquifer system are generally too high in all three model versions along the Mississippi River in Minneapolis and St. Paul based on data from CWI and data targets from other years.

Although not clearly evident in the data targets that were used, both transient model versions appear to have a bias toward computed heads that are too low in much of southern Washington County. The focus area for the model, however, is the lake-dense, central portion of the model domain.

Spatial and temporal head relationships can be more closely examined in hydrographs. To illustrate these relationships in a sub-region of the model, an area near the center of the model domain and White Bear Lake was selected. The locations of DNR observation wells completed in the upper aquifer systems for which there are data in the modeled period are shown in Figure 33.

Figures 34 and 35 show hydrographs for the observation wells completed in the Prairie du Chien aquifer (82039, 82029, 62044, and 82038) and for White Bear Lake. Also shown with these wells is 82057, which is screened immediately above the eroded top of the Prairie du Chien in a deep, locally unconfined sand aquifer. The computed temporal variation in heads at the annual time scale is similar for the two different models. The absolute heads are more closely matched by the revised model for 82029, 82057, and 62038. The fit to White Bear Lake is similar for the two models.

Vertical and horizontal head relationships are complex near the south shore of White Bear Lake reflecting complex geology, hydrogeological property distribution, and influences from well pumping. Figures 36 and 37 show hydrographs for observation wells completed in water table (QWTA) and buried Quaternary (QBAA) aquifers. The groundwater flow direction in the St. Peter and Prairie du Chien aquifers is generally to the southwest to south-southwest in this area (Jones et al., 2013), which is evident in the relationships shown in Figures 34 and 35. The groundwater-flow direction appears to be similar in the buried Quaternary system in this area even at relatively shallow depths. For example, the water level in 62045 (top of screen 25-feet deep) was consistently below the lake level (Figures 36 and 37) indicating a downward gradient from the surface and, combined with other data, horizontal flow away from the south shore of the lake.

There are water-table highs to the southwest of White Bear Lake where surficial till deposits and the Platteville-Glenwood confining bedrock units allow for large vertical head differences between the water table and the St. Peter. The water table appears to slope northeastward, toward the lake (Jones et al., 2013). There are few water-level measurements for Goose Lake, but its lake level appears to be typically above the level of White Bear Lake and about two feet above the water table at the adjacent observation-well nest (62039).

Therefore, horizontal groundwater flow directions are nearly in opposite directions near the water table versus at depths of a few tens of feet. At the observation-well nest adjacent to Goose Lake, the difference in water level in the QWTA (62039) and QBAA (62033) wells averaged from 14 to over 20 feet. This indicates the presence of an aquitard with relatively high resistance as well as differences in horizontal gradients with depth.

At this observation-well nest, both model versions computed Goose Lake to be perched with the water table elevation far below the observed water table (62039) and within a few feet of heads in the nearby QBAA well (62033). The low hydraulic conductivity and lateral extent of the shallow Quaternary aquitard are not captured in the model. The original Quaternary geology model (Tipping, 2011) has large gaps in this area at shallow depths due to spatial gaps in well materials logs. The shallow aquitard appears to be more extensive than in the Quaternary data filled in by Jones et al. (2017) in this area. In addition, hydraulic conductivity of till and lake deposits that form the aquitard can be highly variable, but only a single pair of horizontal and vertical hydraulic conductivities was assigned to each material class in the NMLG models. This local hydrogeological detail affects the computed water table and shallow buried heads, which affect groundwater-lake exchange in this area.

The magnitude of the errors in the computed water table at 62039 are larger for the revised model (average of 22 feet versus 13 feet for SSPA). This appears to result primarily from differences between the two model versions in horizontal hydraulic conductivity in the upper 6 model layers that affect the overall transmissivity and heads of the system in this area. At the observation well nest near Goose Lake, computed heads were 4 to 9 feet too high at the nest in the Prairie du Chien (62038, Figure 34) and 1 to 7 feet too high in the QBAA (62033, Figure 36) in the SSPA model. Computed heads in the revised model overlapped with observed values at 62038 and were 1 to 8 feet too low at 62033.

Hydrographs computed by the annual models for the six lakes modeled with the Lake Package are shown in Figure 38. Statistics describing the fit of annual model-computed lake levels versus observed annual-average lake levels for the 1988-2016 period are given in Table 16. The errors differ somewhat between the two model versions. Turtle Lake levels computed by the revised model vary over a larger range; Snail Lake levels are too low. The computed stage for Snail Lake was very sensitive to lakebed leakance, and slightly lower leakance values caused computed lake levels to be too high. Augmentation of Snail Lake has not yet been incorporated into any of the models. The revised model fits the range of observations more closely for Big Marine and Elmo.

For White Bear Lake, the mean residual for the SSPA model (-0.13 ft) indicates a slight bias toward high lake stages, but the mean residual for the revised model (0.13 ft) indicates a slight low bias. The mean absolute error and root mean square error are slightly lower for the revised model that for SSPA model (ideal of zero).

For the 1988 through 2016 comparison period, the revised, annual model computed a similar variation in White Bear Lake levels (3.3 ft) to the SSPA model (3.4 ft). Both models computed too little variation in lake levels (5.1 ft range in annual averages, 1988-2016). The revised triannual model computed a larger range in stages, but the under-simulation of lake stages during the period of high stages from 1995 through 2003 was similar for the triannual and annual models (Figure 39). As discussed in 2.1.6 White Bear Lake Water Budget Analyses, the models did not produce the large groundwater inflows that appear to have occurred during some years in the 1990s.

Table 16 – Statistics of the revised annual model fit to annual average lake levels, 1988-2016 (feet)

Lake	Turtle		Snail		Big Marine		Elmo		Pine Tree		White Bear	
Model Version	**SSPA**	**Rev.**	**SSPA**	**Rev.**	**SSPA**	**Rev.**	**SSPA**	**Rev.**	**SSPA**	**Rev.**	**SSPA**	**Rev.**
Number of Years	29	29	29	29	26	26	29	29	20	20	29	29
Mean Residual	-0.41	-0.11	0.79	4.35	-0.47	-0.08	1.23	0.45	0.84	0.64	-0.13	0.13
Min Residual	-1.49	-1.59	-1.55	0.94	-1.42	-0.71	-0.37	-0.63	-0.73	-1.24	-1.90	-1.47
Max Residual	0.53	1.96	2.95	8.26	0.10	0.70	4.47	1.65	4.33	3.33	1.71	1.71
Mean Absolute Error	0.53	0.66	1.11	4.35	0.48	0.30	1.26	0.58	1.33	1.18	0.91	0.81
Root Mean Square Error	0.66	0.81	1.31	4.71	0.59	0.36	1.75	0.74	1.85	1.50	1.01	0.93
Observation Range	3.33	3.33	5.22	5.22	1.85	1.85	2.51	2.51	2.38	2.38	5.13	5.13
RMSE / Obs. Range	0.20	0.24	0.25	0.90	0.32	0.19	0.70	0.29	0.78	0.63	0.20	0.18

The tests described in this report showed that the computed lake levels were affected throughout a simulation by the recharge and pumping inputs during the initial steady-state and warm-up (1981-1987) periods. The revised model computed lake levels closer to observations and about one foot lower that the SSPA model in the early 1980s. During initial work on the revised model, this difference in computed stage gradually dissipated until about 2007. Although the changes in steady-state recharge were carried through, updating to the revised lake-evaporation estimates and corresponding lakebed-leakance adjustments eventually had a greater effect, bringing the revised model stage above the SSPA model by 1989.

Similar to the large rise in lake stage from 1991 through 1995, the revised model computed too little rise in lake stage for 1983 through 1986. Because the SSPA model starts out too high and above the outlet, it computed little stage increase through 1986.

Replacing the stage-volume-area table with the improved version in the SSPA model caused computed lake levels to be higher by an amount that generally increased over time (See 3.1 Model Tests above). The stage-volume-area table likely had a similar relative effect on stage variations in the revised model. The revised lake-evaporation appears to have had a larger effect, however. Computed White Bear Lake stages were higher in the revised model than in the SSPA model from 1989 through 1997 but dropped lower for the remainder of the simulation period.

Replacing the preliminary lake evaporation estimates used in the SSPA model with the revised evaporation estimates affected the lake budgets by reducing the net precipitation minus evaporation in most years. The generally increasing trend in the revised evaporation led to decreasing trends in computed lake levels relative to model versions using the preliminary evaporation data set.

Computed groundwater discharges are compared against 2003-2013 average total streamflow for Brown's Creek and Valley Creek in Table 16. Average base flow in these streams is expected to be approximately 85 to more than 90 percent of average total streamflow.

Table 17 – Computed average discharge to Brown's Creek and Valley Creek for 2003-2013

Stream	Average Streamflow	SSPA	DNR Revised	Steady State [1]
Brown's Creek	7.7	3.0	8.1	6.0
Valley Creek	15.6	2.2	7.7	3.6

[1] Computed discharge includes surface leakage computed by the UZF Package for the steady-state model

4.3. Predictive Scenarios Comparisons

SSPA (2017) used the transient NMLG model to evaluate the effect of several hypothetical scenarios with modified historical groundwater pumping on White Bear Lake levels. In the first set of model runs (Scenario 1) wells permitted under each of 45 active permits within 5 miles of White Bear Lake were shut off, one permit at a time, in 1988. In the Scenario 2 runs, reported pumping from the same permits was multiplied by 0%, 25%, 50%, or 75% beginning in 1988. In Scenario 3, the fraction of municipal pumping (among the same group of permits) estimated as outdoor use (primarily irrigation) was removed beginning in 2007. The impacts of reported use relative to each of these scenarios was evaluated by subtracting the unmodified, computed lake stage or volume from the lake stage or volume computed in each hypothetical scenario run. See SSPA (2017) for details about these scenarios.

To compare the annual revised model to the original, SSPA model, a subset of the scenarios were selected. Subsets of Scenario 1 and Scenario 2 were selected that showed a range of relatively larger impacts on lake stage. From Scenario 1, the three permits with the largest computed impacts were selected. From Scenario 2, the end members were selected: 25% shutoff and 100% shutoff (no pumping).

Plots showing the differences in computed lake stage between the hypothetical scenarios and the unmodified models are shown in Figures 40 and 41. The revised model (designated as Revised Er in Figures 40 and 41) computed less stage impact for most of the scenarios. Notably, the SSPA and revised annual models computed similar impacts after 2004 for the no pumping scenario. The revised model computed a larger relative contrast between the no pumping and 25% shutoff scenarios.

The stage differences developing from pumping changes applied beginning in 1988 generally developed more gradually in the revised model. Differences in hydraulic properties (hydraulic conductivity and lakebed leakance) between the two models likely influenced the timing as well as the magnitude of computed impacts.

For comparison, the computed stage differences for the initial model revision that used the preliminary lake evaporation estimates (Ep) are also shown in Figures 40 and 41. The lakebed leakance values were somewhat higher in the initial revised model, allowing more net groundwater outflow. The decreased net precipitation minus evaporation with the revised evaporation estimates (Er) required decreasing the lakebed leakance, which reduced net groundwater outflow to more closely rebalance the lake water budget. The revised evaporation also changed the pattern of lake-level variation over time in the base scenario.

This demonstrates how input stresses that affected the lake water budget both directly (due to the non-linear system behavior) and indirectly (by influencing the estimated parameters) affected the computed pumping impacts. The shortcomings in representation of shallow hydrogeological properties and shallow groundwater-flow dynamics identified above (See Section 2.1.6 White Bear Lake Water Budget Analysis) also likely affected the computed stage differences in White Bear Lake caused by pumping from deeper aquifers (i.e. Prairie du Chien and Jordan) for all model versions. This type of predictive uncertainty should be considered when applying the model to assess impacts or evaluate hypothetical scenarios.

SSPA (2017) also evaluated two of the permits (1969-0174 and 1977-6229) and the no pumping scenario using the triannual model. The triannual model computed impacts were larger in most years than the impacts computed by the annual model. The revised, triannual model was also run for these scenarios. The triannual results were very similar to the annual model results for the revised model (Figure 42). For the no pumping scenario, the computed stage differences were 7 to 25 percent less (0.1 to 1 foot) than in the triannual, SSPA model from 2004 through 2015, but the revised model stage differences were 0 to 0.3 foot greater in 2016. Like the annual model, the contrasts were greater between the SSPA and revised model stage differences for the individual permits.

5. Summary and Recommendations

This report describes an effort by DNR staff to better understand the behavior of the transient NMLG model; to explore the sensitivity of results to selected modeling assumptions and choices; to provide insight into the effects of model uncertainties and limitations on model predictions; and to revise the model using updated data and improved information.

Evaluation of model input data focused on precipitation, groundwater recharge, and lake evaporation. Uncertainties in groundwater recharge and other water-balance components and limitations of the SWB model affected estimation of hydrogeological properties during model history matching (i.e. "calibration"). The implications of these factors were explored through a series of model tests. In addition, new and additional data provided opportunities to refine the representation of some model features.

In conjunction with and following the model review and testing, the model was revised to incorporate new data and more tightly constrain model parameters. The revised model used a different approach to parameterizing the SWB model that resulted in higher average recharge rates than the rates computed for the USGS or SSPA models. In addition, the revised model incorporated changes to some model features based on new or better information, such as a new stage-volume-area table for White Bear Lake. Recently developed, revised White Bear Lake evaporation estimates were also applied in the revised model, a substantial change to lake water budgets. The revised model fit observed heads and lake levels similarly to the original transient model, but it used parameters more tightly constrained to be consistent with the conceptual model of the system.

The resulting revised model does not represent a unique solution, but it incorporates improvements and new data. Computed pumping impacts on White Bear Lake were qualitatively similar but smaller for the revised model than for the original transient model.

The two models (SSPA and revised) do not represent end members of the potential range in possible impacts, but the differences in the computed impacts are a demonstration of how model uncertainties and data limitations can affect computed impacts. Despite their differences, both models indicate that White Bear Lake levels might have been higher with reduced historical pumping under certain permits. The models also generally agree on the relative ranking of which scenarios produced more or less impacts than others. Taking a particular management action may have more or less effect on lake stages than computed by the model.

As recommended by SSPA (2017), the triannual model should be used for evaluating specific management scenarios. The revised model could be further modified to use shorter stress periods. Shorter stress periods would be most practical if applied to limited time periods, such as monthly stress periods during the summer for a limited number of years with triannual stress periods for the remainder of the simulation. The similarities between the annual and triannual model-computed stage differences for pumping scenarios in the revised model indicates, however, that longer-term effects may be more important than short-term variations in pumping rates. Additionally, the computed stage differences should be evaluated in the context of predictive uncertainty.

Using monthly stress periods could distort the timing of recharge where the water table is deep. Deep water tables are not common in the focus area for the model, however. To compute percolation through the vadose zone accurately using the UZF package, the model would have to be run on shorter stress periods (e.g. 1 day). If needed, a simple method to lag and attenuate recharge could be developed to provide an adequate approximation at the monthly time scale without having to use the Unsaturated Zone Flow (UZF) package.

6. References

Allen, R.G, 1995, Evaluation of procedures for estimating mean monthly solar radiation from air temperature: Report submitted to the United Nations Food and Agricultural Organization (FAO), Rome Italy.

Almendinger, J. E., 2003, Watershed hydrology of Valley Creek and Browns Creek - Trout streams influenced by agriculture and urbanization in eastern Washington County, Minnesota, 1998-99: St. Croix Watershed Research Station, Science Museum of Minnesota, 80 pp., ill. maps, diagr., tables.

Barr Engineering Company, 2010, Southern Washington County Groundwater Model Refinements: Technical Memorandum to Washington Co. Dept. of Public Health and Environment, 58 p.

Barr Engineering Company and Washington County, 2005, Intercommunity Groundwater Protection – Sustaining Growth and Natural Resources in the Woodbury/Afton Area, Report on Development of a Groundwater Flow Model of Southern Washington County, Minnesota: Barr Engineering Co., 67 p., appendices, figures. https://www.leg.state.mn.us/edocs/edocs?oclcnumber=63690522.

Blum, J., 2011, Evaluation of Pumping Test – Hugo Well 6 (773400) on October 28, 2010, Jordan Sandstone Aquifer: Minnesota Department of Health memo to file, PWSID 1820007, 25 p.

Blum, J., 2016b, Analysis of the MDNR Test Well (767123) - Pumping Test, July 21, 2015, Jordan Aquifer – Vermillion Wildlife Management Area: Minnesota Department of Health memo, 31 p.

Blum, J., 2016b, Analysis of the Lakeville 3 (207727) - Pumping Test, February 25, 2015, Jordan Sandstone - MDNR: Minnesota Department of Health memo to file, PWSID 1190015, 28 p.

Bonestroo Rosene Anderlik and Assoc., 2004, City of Woodbury Well 15 aquifer test report, prepared for the City of Woodbury, Proj. No. 31-02-175, 27 p.

Champion, G.S., 2009, Part 1 Wellhead Protection Plan, City of White Bear Lake: report prepared for the City of White Bear Lake.

Dadaser-Celik, F. and H.G. Stafan, 2008, Lake evaporation response to climate in Minnesota, Project Report No. 506: Univ. of Minn. St. Anthony Falls Lab. report prepared for the Legislative Citizens Committee on Minnesota Resources, 63 p.

Daly, C., Gibson, W.P., M. Doggett, J. Smith, and G. Taylor, 2004, Up-to-date monthly climate maps for the conterminous United States. 14th AMS Conf. on Applied Climatology, 84th AMS Annual Meeting Combined Preprints, Seattle, WA, January 13-16, Paper P5.1.

DeGaetano, A. T., 1998, Identification and implications of biases in US surface wind observation, archival, and summarization methods: Theor. Appl. Climatol., 60, 151–162, doi:10.1007/s007040050040.

DNR, 1968, An Inventory of Minnesota Lakes: Minnesota Conservation Department, Division of Waters, Soils, and Minerals Bulletin No. 25.
http://www.dnr.state.mn.us/waters/watermgmt_section/pwi/bulletin25.html

DNR, 2014, LiDAR Elevation, Twin Cities Metro Region, Minnesota, 2011: Minnesota Department of Natural Resources digital data, http://www.mngeo.state.mn.us/chouse/metadata/lidar_metro2011.html#Distribution_Information.

DNR, 2015a, Lake Bathymetric Outlines, Contours, Vegetation, and DEM: Minnesota Department of Natural Resources digital data, https://gisdata.mn.gov/dataset/water-lake-bathymetry.

DNR, 2015b, DNR Hydrography - Stream Routes with Stream Types: Minnesota DNR Division of Fish & Wildlife, digital data. https://gisdata.mn.gov/dataset/water-dnr-hydrography

DNR and Metropolitan Council, 2010, Digital Orthoimagery, Twin Cities, Spring 2010, 1-ft Resolution: Digital data accessed http://www.mngeo.state.mn.us/chouse/wms/geo_image_server.html.

Doherty, J., 2016, PEST—Model-independent parameter estimation user manual (6th ed.): Brisbane, Australia, Watermark Numerical Computing, http://pesthomepage.org/.

Duan, Z and W G M Bastiaanssen, 2017, Evaluation of three energy balance-based evaporation models for estimating monthly evaporation for five lakes using derived heat storage changes from a hysteresis model: Env. Res. Lett. 12, 024005. doi:10.1088/1748-9326/aa568e

Emmons & Olivier Resources, Inc., 2009, Groundwater Management & Planning Program Final Report: prepared for Rice Creek Watershed District, 117 p.

Farnsworth, R.K., Thompson, E.S., and Peck, E.L., Evaporation Atlas for the Contiguous 48 United States, NOAA Technical Report NWS 33: National Weather Service, Washington, D.C.

Feinstein, D.T., Eaton, T.T., Hart, D.J., Krohelski, J.T., and Bradbury, K.R., 2005, Regional Aquifer Model for Southeastern Wisconsin Report 1 – Data Collection, Conceptual Model Development, Numerical Model Construction, and Model Calibration: Report to the Southeastern Wisconsin Regional Planning Commission, 63 p.

Hargreaves, G.H.,1994, Simplified coefficients for estimating monthly solar radiation in North America and Europe: Departmental Paper, Dept. of Biol. And Irrig. Engrg., Utah State University, Logan, Utah.

Hargreaves GH and Samani ZA., 1985, Reference crop evapotranspiration from temperature: Applied Engineering in Agriculture, 1(2):96–99.

Hofstra, N., New, M. & McSweeney, C., 2010, The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data: Climate Dynamics 35(5), 841–858. https://doi.org/10.1007/s00382-009-0698-1

Hydrologic Engineering Center (HEC), 1998, HEC-1 Flood Hydrograph Package Users Manual: U.S. Army Corps of Engineers Hydrologic Engineering Center, Davis, CA.

Jones, P.M., Roth, J.L., Trost, J.J., Christenson, C.A., Diekoff, A.L., and Erickson, M.L., 2017, Simulation and assessment of groundwater flow and groundwater and surface-water exchanges in lakes in the northeast Twin Cities Metropolitan Area, Minnesota, 2003 through 2013: U.S. Geological Survey Scientific Investigations Report 2016–5139–B, 88 p., https://doi.org/10.3133/sir20165139B.

Jones, P.M., Trost, J.J., Rosenberry, D.O., Jackson, P.R., Bode, J.A., and O'Grady, R.M., 2013, Groundwater and surface-water interactions near White Bear Lake, Minnesota, through 2011: U.S. Geological Survey Scientific Investigations Report 2013–5044, 73 p.

Klink, K. ,2002, Trends and interannual variability of wind speed distributions in Minnesota: J. Clim., 15, 3311 – 3317, doi:10.1175/1520-0442(2002)015<3311:TAIVOW>2.0.CO;2.

Levy, M.C., Cohn, A., Lopes, A.V., and Thompson, S.E., 2017, Addressing rainfall data selection uncertainty using connections between rainfall and streamflow: Scientific Reports, 7, Article number: 219. doi:10.1038/s41598-017-00128-5

Lockhart, T., 1999, Climate continuity of wind speed with ASOS: AMS 11th Conference on Applied Climatology, Dallas, TX, 11 Jan.

McKee, T.B., N.J. Doesken, C.A. Davey, and R.A. Pielke, Sr., 2000, Climate data continuity with ASOS - Report for period April 1996 through June 2000: Climo Report 00-3, November, 77 pp.

Metropolitan Council, 1999, 1997 Digital Orthophotos, Twin Cities Metropolitan Area: Digital data https://gisdata.mn.gov/dataset/us-mn-state-metc-base-metro-orthos-1997.

Metropolitan Council, 2014, Twin Cities Metropolitan Area Regional Groundwater Flow Model, Version 3.0, Prepared by Barr Engineering: Metropolitan Council, Saint Paul, MN.

Metropolitan Council, 2016, Digital Orthoimagery, Twin Cities, Minnesota, Spring 2016, 1-ft Resolution: http://www.mngeo.state.mn.us/chouse/wms/geo_image_server.html

Minnesota State Climatology Office, 2017, climate data. http://climateapps.dnr.state.mn.us/index.htm

Mizukami, N. and Smith, M.B., 2012, Analysis of inconsistencies in multi-year gridded quantitative precipitation estimate over complex terrain and its impact on hydrologic modeling: Journal of Hydrology 428-429, 129-141. https://doi.org/10.1016/j.jhydrol.2012.01.030

Mossler, J.H., 2013, Bedrock Geology and Structure of the Seven-County Twin Cities Metropolitan Area, Minnesota: Minnesota Geological Survey Miscellaneous Map Series Map M-104, 1:125,000 scale.

Mossler, J.H. and Tipping, R.G., 2010, Bedrock Geology of the Twin Cities Ten-County Metropolitan Area, Minnesota: Minnesota Geological Survey Miscellaneous Map Series Map M-194, 1:125,000 scale.

NOAA, Precipitation Frequency Data Server (PFDS): National Weather Service Hydrometeorological Design Studies Center, https://hdsc.nws.noaa.gov/hdsc/pfds/index.html, accessed March 2018.

Penman, H. L., 1948, Natural evaporation from open water, bare soil and grass: Proc. R. Soc. London, Ser. A,193, 120–145.

PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, accessed February 2018.

Pryor, S. C., R. J. Barthelmie, D. T. Young, E. S. Takle, R. W. Arritt, D. Flory, W. J. Gutowski Jr., A. Nunes, and J. Roads, 2010, Wind speed trends over the contiguous United States: J. Geophys. Res., 114, 14105.

Reeder, H.O., Wood, W.W., Ehrlich, G.G., and Sun, R.J., 1976, Artificial Recharge through a Well in Fissured Carbonate Rock, West St. Paul, Minnesota: U.S. Geological Survey Water-Supply Paper 2004, 80 p.

Roderick, M.L., Hobbins, M.T., and Farquhar, G.D., 2009a, Pan evaporation trends and the terrestrial water balance - I. principles and observations: Geography Compass, 2(2), 746-760. doi.org/10.1111/j.1749-8198.2008.00213.x

Roderick, M.L., Hobbins, M.T., and Farquhar, G.D., 2009b, Pan evaporation trends and the terrestrial water balance - II. Energy balance and interpretation: Geography Compass, 2(2), 761-780. doi.org/10.1111/j.1749-8198.2008.00214.x

Runkel, A.C., Tipping, R.G., Alexander, E.C., Jr., Green, J.A., Mossler, J.H., and Alexander, S.C., 2003, Hydrogeology of the Paleozoic bedrock in southeastern Minnesota: Minnesota Geological Survey Report of Investigations 61, 105 p., 2 pls.

Samani, Z., 2000, Estimating solar radiation and evapotranspiration using minimum climatological data: Journal of Irrigation and Drainage Engineering, Vol. 126, No. 4., pp. 265-267, doi:10.1061/(asce)0733-9437(2000)126:4(265).

Sanocki, C.A., Langer, S.K., and Menard, J.C. 2009, Potentiometric surfaces and changes in groundwater levels in selected bedrock aquifers in the Twin Cities Metropolitan Area, March–August 2008 and 1988–2008: U.S. Geological Survey Scientific Investigations Report 2009–5226, 67 p. with appendices.

Schoenberg, M.E., 1984, Water levels and water-levels changes in the Prairie du Chien-Jordan and Mount Simon-Hinckley aquifers, Twin Cities Metropolitan Area, Minnesota, 1971-80: U.S. Geological Survey Water-Resources Investigations Report 83-4237, 23 p. and two plates.

Searcy, J.K. and C.H. Hardison, 1960, Double-Mass Curves, Manual of Hydrology – Part 1, General Surface-Water Techniques: U. S. Geological Survey Water-Supply Paper 1541-B, 66 p.

Singh, V.P. and C.-Y. Xu, Evaluation and generalization of 13 mass-transfer equations for determining free water evaporation: Hydr. Proc., 11, 311-323.

Smith, E.A., and Westenbroek, S.M., 2015, Potential groundwater recharge for the State of Minnesota using the Soil-Water-Balance model, 1996–2010: U.S. Geological Survey Scientific Investigations Report 2015–5038, 85 p., http://dx.doi.org/10.3133/sir20155038.

Soil Conservation Service (SCS) Engineering Division, 1986, Urban hydrology for small watersheds: U.S. Department of Agriculture, Technical Release 55.

S. S. Papadopulos & Associates, Inc. (SSPA), 2017, Development of a Transient Version of the Northeast Metro Lakes Groundwater-flow (NMLG) Model - with Simulation and Calibration from 1980 through 2016:

prepared for the Minnesota Department of Natural Resources, October 2017. Available at https://www.dnr.state.mn.us/gwmp/area-ne-model.html

Thornton, P.E., Running, S.W., White, M.A, 1997, Generating surfaces of daily meteorological variables over large regions of complex terrain: Journal of Hydrology 190: 214 - 251. http://dx.doi.org/10.1016/S0022-1694(96)03128-9

Thornton; P.E.; M.M. Thornton; B.W. Mayer; Y. Wei; R. Devarakonda; R.S. Vose; and R.B. Cook, 2017, Daymet: Daily Surface Weather Data on a 1-km Grid for North America; Version 3: ORNL DAAC; Oak Ridge; Tennessee; USA. http://dx.doi.org/10.3334/ORNLDAAC/1328

Trost, J.J., Roth, J.L., Jones, P.M., and Christenson, C.A., 2017, MODFLOW-NWT model used to simulate and assess groundwater flow and surface-water exchanges in lakes of the Northeast Twin Cities Metropolitan Area, Minnesota, 2003 through 2013 (ver. 1.0): U.S. Geological Survey data release, https://dx.doi.org/10.5066/F7445JRM.

U.S. Geological Survey (USGS), 1992, Digital Orthophoto Quadrangles (DOQs), Minnesota, 1991-92.

U.S. Geological Survey (USGS), 2007, USGS High Resolution Orthoimagery, Ramsey County, Minnesota, Spring 2006: Digital data accessed http://www.mngeo.state.mn.us/chouse/wms/geo_image_server.html.

U.S. Geological Survey (USGS), 2012, USGS High Resolution Orthoimagery, Minneapolis-St. Paul, MN, Spring, 2012: Digital data http://www.mngeo.state.mn.us/chouse/wms/geo_image_server.html.

Vinayak, S.S., King, K.W., Brown, L.C., Fausey, N.R., Heckel M., and Harmel, R.D., 2009, Measurement errors in tipping bucket rain gauges under Different Rainfall Intensities and their under different rainfall intensities and their implication to hydrologic models, ASABE Paper No. 097368: ASABE, St. Joseph, Michigan.

Vose, R.S., S. Applequist, M. Squires, I. Durre, M.J. Menne, C.N. Williams, C. Fenimore, K. Gleason, and D. Arndt, 2014a, Improved historical temperature and precipitation time series for U.S. Climate Divisions: J. Appl. Meteor. Climatol., 53, 1232–1251, https://doi.org/10.1175/JAMC-D-13-0248.1

Vose, Russell S.; Applequist, Scott; Squires, Mike; Durre, Imke; Menne, Matthew J.; Williams, Claude N., Jr.; Fenimore, Chris; Gleason, Karin; Arndt, Derek, 2014b, NOAA's Gridded Climate Divisional Dataset (CLIMDIV), nClimDiv division climate indices dataset: NOAA National Climatic Data Center. doi:10.7289/V5M32STR accessed February 2018.

Washington County, 2013, High Resolution Orthoimagery, Washington County, MN, Spring, 2013: Digital data https://gisdata.mn.gov/dataset/base-washington-2013-airphotos.

Woodward, D.E., Hawkins, R.H., Jiang, R., Hjelmfelt, A.T., Van Mullem, J.A., and Quan, Q.D., 2003, Runoff curve number method – Examination of the initial abstraction ratio, in World Water and Environmental Resources Congress 2003, Conference Proceedings Papers, June 24-26, Philadelphia: American Society of Civil Engineers, 12 p.

Xiao, K., Griffis, T.J., Baker, J.M., Bolstad, P.V., Erickson, M.D., Lee, X., Wood, J.D., Hu, C., and Nieber, J.L., 2018, Evaporation from a temperate closed-basin lake and its impact on present, past, and future water level: Journal of Hydrology 561, 59-75.