USGS
science for a changing world

Water Availability and Use Science Program

Initial Estimates of Net Infiltration and Irrigation from a Soil-Water-Balance Model of the Mississippi Embayment Regional Aquifer Study Area

Open-File Report 2021–1008

U.S. Department of the Interior
U.S. Geological Survey

Front and back cover image. Two combined Landsat 8 photographs (Path 23, rows 26 and 27) show the Mississippi River and its confluence with the Arkansas River and the White River. The images were taken on August 20, 2020. Landsat photographs are courtesy of the U.S. Geological Survey.

Initial Estimates of Net Infiltration and Irrigation from a Soil-Water-Balance Model of the Mississippi Embayment Regional Aquifer Study Area

By Stephen M. Westenbroek, Martha G. Nielsen, and David E. Ladd

Water Availability and Use Science Program

Open-File Report 2021–1008

U.S. Department of the Interior
U.S. Geological Survey

U.S. Geological Survey, Reston, Virginia: 2021

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit https://www.usgs.gov or call 1–888–ASK–USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit https://store.usgs.gov/.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this information product, for the most part, is in the public domain, it also may contain copyrighted materials as noted in the text. Permission to reproduce copyrighted items must be secured from the copyright owner.

Suggested citation:
Westenbroek, S.M., Nielsen, M.G., and Ladd, D.E., 2021, Initial estimates of net infiltration and irrigation from a soil-water-balance model of the Mississippi Embayment Regional Aquifer Study Area: U.S. Geological Survey Open-File Report 1008, 29 p., https://doi.org/10.3133/ofr20211008.

Associated data for this publication:
Westenbroek, S.M., Nielsen, M.G., and Ladd, D.E., 2021a, OFR 2021–1008 MODEL ARCHIVE—Soil-Water-Balance model developed to simulate net infiltration and irrigation water use for the Mississippi Embayment Regional Aquifer System, 1915 to 2018: U.S. Geological Survey data release, https://doi.org/10.5066/P98PBR8O.

Westenbroek, S.M., Nielsen, M.G., and Ladd, D.E., 2021b, OFR 2021–1008 MODEL OUTPUT—Soil-Water-Balance net infiltration and irrigation water use output datasets for the Mississippi Embayment Regional Aquifer System, 1915 to 2018: U.S. Geological Survey data release, https://doi.org/10.5066/P9U484X5.

ISSN 2331-1258 (online)

Contents

Figures

Tables

Conversion Factors

U.S. customary units to International System of Units

Multiply	By	To obtain
	Length	
inch (in.)	2.54	centimeter (cm)
mile (mi)	1.609	kilometer (km)
	Area	
square mile (mi^2)	2.590	square kilometer (km^2)
	Flow rate	
inch per year (in/yr)	25.4	millimeter per year (mm/yr)

International System of Units to U.S. customary units

Multiply	By	To obtain
	Length	
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)

Abbreviations

CDL	Cropland Data Layer
ET	evapotranspiration
HSG	hydrologic soil group
MAP	Mississippi Alluvial Plain
MERAS	Mississippi Embayment Regional Aquifer Study
netCDF	Network Common Data Format
NRCS	Natural Resources Conservation Service
PEST	Parameter ESTimation
SWB	Soil-Water-Balance
USDA	U.S. Department of Agriculture
USGS	U.S. Geological Survey

Initial Estimates of Net Infiltration and Irrigation from a Soil-Water-Balance Model of the Mississippi Embayment Regional Aquifer Study Area

By Stephen M. Westenbroek, Martha G. Nielsen, and David E. Ladd

Abstract

The Mississippi embayment encompasses about 100,000 square miles and covers parts of eight States. In 2016, the U.S. Geological Survey began updating previous work for a part of the embayment known as the Mississippi Alluvial Plain to support informed water use and agricultural policy in the region. Groundwater, water use, economic, and other related models are being combined with field surveys and observations to create a quantitative framework for evaluating regional groundwater withdrawals and their effects on long-term water availability in the Mississippi Alluvial Plain.

As part of this effort, the U.S. Geological Survey's Soil-Water-Balance code (version 2.0) is being used to model potential groundwater recharge and irrigation water use, as necessary inputs to the long-term groundwater modeling efforts. The Soil-Water-Balance code is designed to estimate the distribution and timing of net infiltration leaving the root zone. Soil-Water-Balance makes use of gridded datasets of elevation, soils, land use (including specific crop types), and daily weather datasets to calculate other components of the root-zone water balance, including soil moisture, reference, actual evapotranspiration, snowfall, snowmelt, and canopy interception. Parameters on plant height and growing-season water needs are used to estimate crop-water demand and potential irrigation water use.

This report documents the initial construction, calibration, and application of a Soil-Water-Balance model of the Mississippi Embayment Regional Aquifer Study area for simulations running from 1915 to 2017. Further refinements of the model calibration for an expanded model area are planned.

Introduction

The U.S. Geological Survey (USGS) is involved in a multiyear, multidiscipline regional water availability study for the Mississippi River Alluvial Plain (MAP). More information about the study is available online at https://www2.usgs.gov/water/lowermississippigulf/map/index.html. For the MAP study, the USGS; in cooperation with more than 10 Federal, State, and local stakeholders; is comprehensively studying and modeling the Mississippi embayment aquifer system and associated hydrogeologic units (Clark and others, 2013) underlying the MAP. The USGS is currently (2020) developing groundwater model simulations for the MAP and the Mississippi embayment aquifer system that are intended to quantify groundwater conditions for several areas in which groundwater pumping has resulted in large cones of depression (Alhassan and others, 2019a, 2019b; Clark and others, 2013; Painter and Westerman, 2018).

In order to simulate the potential effect of changes in pumping behavior into the future, and to support informed water use and agricultural policy in the region, the groundwater model simulation is planned to cover the time over which the cones of depression have developed. Inputs needed for the groundwater models include recharge and information on groundwater pumping. The modeling approach documented in this report was used to provide these inputs to the long-term groundwater-flow model; a modeling approach is necessary because groundwater recharge is not directly observable and long-term pumping records often contain missing data or only qualitative information. The USGS Soil-Water-Balance (SWB) model (Westenbroek and others, 2018) was chosen to simulate net infiltration (or potential recharge to groundwater) and irrigation water demand (met by pumping from groundwater) for the long-term Mississippi Embayment Regional Aquifer Study (MERAS) groundwater modeling effort.

The USGS SWB model has been applied throughout North America and in study areas in Europe and Asia (Pirellas, 2019; Mair and others, 2013; Smith and Westenbroek, 2015; Trost and others, 2018). The SWB method of estimating net infiltration uses inputs of precipitation, temperature, land cover, and soil information, and estimates of potential and actual evapotranspiration (ET). The calculations done using this method traditionally have been used in the agricultural sector to estimate crop water demands, but they also can be used to provide estimates of excess soil moisture, which is the source of recharge to groundwater (Westenbroek and others, 2018). In this report, the term "net infiltration" refers to the calculation of water that passes through the rooting zone of the soil and is available for potential recharge to the groundwater system.

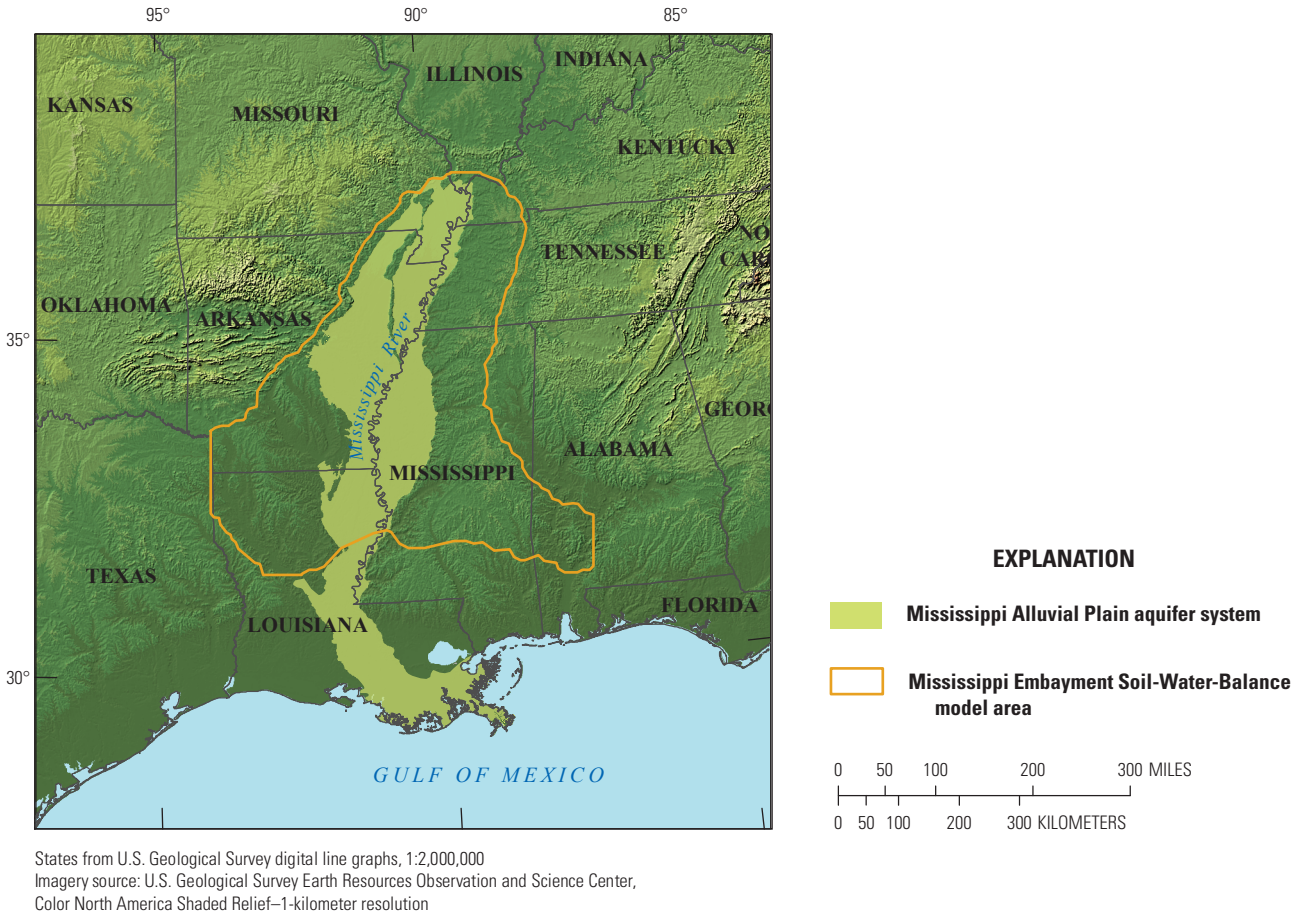


Figure 1. Location of the Mississippi embayment Soil-Water-Balance model area and the Mississippi Alluvial Plain aquifer system.

The SWB model was used to estimate net infiltration (or groundwater recharge) for two time periods for use in a long-term groundwater-flow model of the MAP and the broader Mississippi embayment from northern Louisiana to southern Missouri (fig. 1). Previous modeling of this study area was done for MERAS (Clark and others, 2013; Clark and Hart, 2009). The Mississippi embayment Soil-Water-Balance model area (fig. 1) encompasses about 100,000 square miles and covers parts of eight States. In 2016, the USGS began updating all aspects of previous water-budget work for the MERAS area and more of the MAP aquifer system to support informed water use and agricultural policy in the region (Alhassan and others, 2019a, 2019b; Clark and others, 2013; Painter and Westerman, 2018). Groundwater, water use, economic, and other related models are being combined with field surveys and observations to create a quantitative framework for evaluating regional groundwater withdrawals and their effects on long-term water availability in the MERAS/MAP study area.

The modeling described in this report spans the period from 1915 to 2017. Two separate SWB models were developed because of changes in data availability during this period. The model covering the earlier period (1915–2011) relies on spatial estimates of land use derived from aerial photographs and county-level data reports; for the later period (2000–17), land-use data are derived from satellite observations.

Purpose and Scope

The purpose of this report is to document SWB model construction and parameter estimation for the Mississippi embayment model study area and the estimated net infiltration and irrigation water use grids produced by the model for the following two time periods: 1915–2011 and 2000–17. This report describes the input datasets, observation datasets, SWB model output processing, and general SWB model calibration.

Data Sources and Preparation

The USGS SWB version 2.0 model code (Westenbroek and others, 2018) uses gridded input data for soil properties, land use, and daily weather to run. Because of differences in the availability of land use and weather data over time, two SWB models were developed for the Mississippi embayment area, each one using a different set of source data. The two models together span the years from 1915 through 2017.

The first SWB model for the MAP area (the "long-term historical model") spans the years from 1915 through 2011. Spatial datasets describing agricultural practices and land use are almost nonexistent for the earliest decades of the 20th century. Application of SWB to the earliest period requires the adoption of many assumptions and generalizations about agricultural practices and irrigation during years when no data were collected. The data used are described in the "Long-Term Historical Time Period" section in this report.

The second SWB model for the MAP area (the "short-term historical model") spans the years from 2000 through 2017. This period roughly corresponds to a period in which remotely sensed datasets became more widely available. There is an 11-year overlap between the two models.

Spatial Extent and Model Resolution

The work described in this report was primarily aimed at covering the spatial extent of the existing MERAS model (Clark and Hart, 2009; fig. 1). The updated MODFLOW model and the SWB model of the MERAS area have a spatial resolution of 1,609 meters (m), or 1 mile. All input data grids were reprojected to match the spatial bounds and projection used in the underlying MERAS MODFLOW model and then resampled to the 1,609-m grid-cell size from the native resolution to the model resolution using software from the Geospatial Data Abstraction Library (GDAL) (Warmerdam, 2016). Additional information about the spatial data extraction is given in appendix 1.

Common Spatial Datasets

Some of the input datasets are independent of the time period and are used for both of the SWB models. These input datasets include the hydrologic soil group (soil groups A, A/D, B, B/D, C, C/D, and D), available water capacity (AWC), and irrigated crop areas. Hydrologic soil group and AWC data are derived from work by Wieczorek (2014), who calculated area-weighted values of certain soil properties from the U.S. Department of Agriculture Natural Resources Conservation Service (NRCS) Soil Survey gSSURGO and STATSGO datasets (Soil Survey Staff, 2014a, 2014b). The hydrologic soil group data include four primary hydrologic soil groups, A through D, that range from low to high runoff potential (high to low infiltration capacity) and comprise less than 10 percent clay to more than 40 percent clay, respectively (U.S. Department of Agriculture, 2007). Most of the soil types in the MAP model area are moderately to poorly conductive, in soil hydrologic groups B, B/D, C, C/D, and D (fig. 2), which is reflective of the surficial geology and alluvial silt and fine sand of the Mississippi Alluvial Plain.

Most of the cropland in the model area has soils in the D hydrologic soil group (table 1). The forested land cover types are primarily in the C and B hydrologic soil groups, and the woody wetlands primarily are in the D and C hydrologic soil groups. The hay-pasture-grassland-shrubland category is fairly evenly divided between the C, B, and D soil hydrologic soil groups (table 1).

The irrigated acres dataset is taken from a grid representing the extent of irrigation as estimated for 2012 (Brown and Pervez, 2014), which has a spatial resolution of 250 m. The 2012 irrigation dataset was used for the entire simulation period. This obviously overestimates the irrigated acres present earlier in the 20th century, as irrigation began slowly in the Mississippi Alluvial Plain region. However, in the SWB application, the irrigated area grid is only used as a mask; irrigation is only supplied to grid cells that were cultivated during each modeled time period.

Long-Term Historical Time Period

The existing MERAS MODFLOW model (Clark and Hart, 2009) has 69 stress periods that extend from the late 1800s through 2007 and was later updated (Clark and others, 2013) to include 78 stress periods with the final stress period ending in 2011. The long-term historical SWB model covers the years 1915 through 2011 and overlaps all but the earliest years of the MERAS MODFLOW model time steps. Sources of gridded land use and daily weather data for the long-term historical SWB model are given in table 2.

The gridded daily weather data were taken from Livneh and others (2013). These datasets were generated from the National Weather Service Cooperative Monitoring Network Stations. The network station data were interpolated by Livneh and others (2013) using the SYMAP algorithm (Shepard, 1968). The resulting gridded values were adjusted to ensure that their long-term statistical properties matched those of the Parameter-elevation Regressions on Independent Slopes Model (PRISM) gridded weather data product (Daly and others, 1994), which was used for the short-term historical time period and are described in a later section of this report.

Reconstructed land-use grids created by Sohl and others (2016) represent annual estimates of 14 general land-use types for the years 1938 through 2011. As the earliest year of land-use data available, the 1938 land-use distribution was applied to the preceding years of the historical long-term period for the years 1915–37. A separate grid of estimated land use was supplied to SWB to represent each year from 1939 through 2011. Figure 3 illustrates the general land-use distribution across the model area for 1938, the earliest year for which land-use

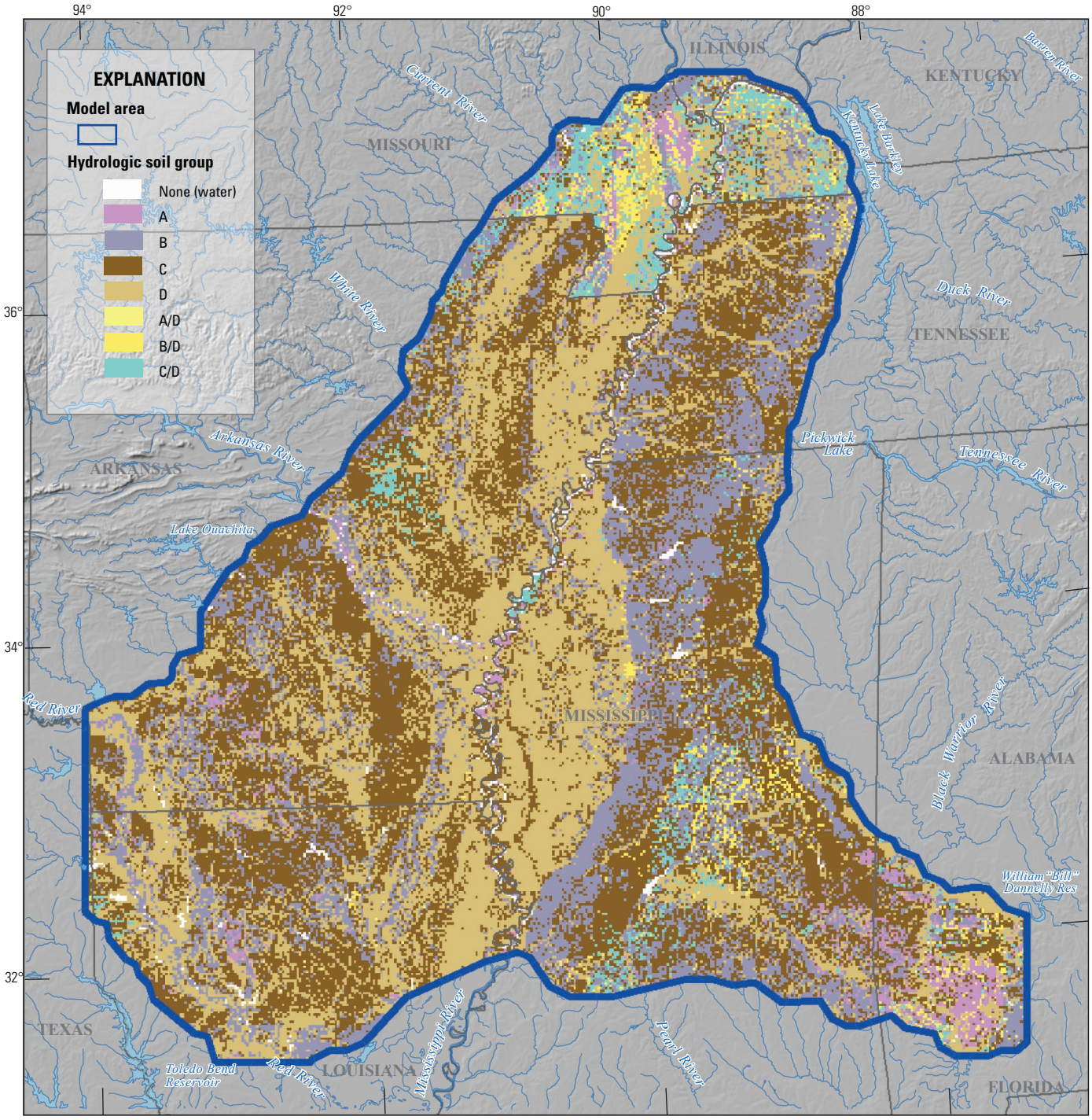


Figure 2. Distribution of hydrologic soil groups within the Mississippi embayment Soil-Water-Balance model area.

Table 1. Distribution of land use and soil types by area within the model domain.

[Area in square miles. --, no data]

Generalized land use, 2016	Hydrologic soil group							
	A	A/D	B	B/D	C	C/D	D	Soil data absent[1]
Water	85	--	330	14	544	50	825	546
Developed	116	5	1,340	118	2,421	216	1,439	28
Barren	9	--	13	1	16	2	12	13
Forest	975	16	9,130	675	13,298	831	5,983	46
Hay-pasture-grassland-shrubland	359	6	4,068	297	6,143	488	3,748	31
Cultivated crops	408	16	3,083	789	6,877	1,167	11,812	59
Herbaceous wetland	4	1	24	7	49	18	97	11
Woody wetland	303	11	1,413	288	4,019	340	6,282	169

[1]Natural Resources Conservation Service soil group codes are absent primarily over water and open-water wetlands.

Table 2. Data sources for the long-term historical Soil-Water-Balance application.

[°, degree]

Dataset type	Dataset name	Date range	Spatial resolution	Reference
Daily weather	Livneh and others	1915–2011	1/16°, approximately 7,000 meters	Livneh and others (2013)
Land use	Sohl and others	1938–92	250 meters	Sohl and others (2016)

data were available during the long-term historical period. The total area belonging to each of the land-use categories is shown in figure 4; the proportion of the cultivated crops that is assumed to be irrigated is shown overlain on the cultivated crops category.

The amount of irrigated land is likely overestimated because we used an irrigation "mask" representing 2007 conditions to define areas of irrigated and nonirrigated land (Brown and Pervez, 2014) for all time steps in the model. Plans for the MAP project are to develop better estimates of irrigated and nonirrigated land use with time, but that output was not available for use in this iteration of model development.

Short-Term Historical Time Period

The short-term historical SWB model for the MERAS area has a time span that roughly coincides with the increased availability of remotely sensed satellite datasets, spanning the years 2000 through 2017. The Wieczorek (2014) gridded soil data and the Brown and Pervez (2014) irrigated areas data sources were used for the soil inputs for this model along with the weather and land-use data sources given in table 3.

The DayMet (version 3) daily weather data from the Oak Ridge National Laboratory (Oak Ridge, Tennessee; Thornton and others, 2018) were the source of daily minimum and maximum temperature data and daily precipitation data used in the short-term historical period SWB model for the MERAS area. These data, available in the Network Common Data Form (netCDF) format, have a spatial resolution of 1 kilometer. DayMet data are derived from an algorithm that models daily temperature and precipitation based on ground-surface observations, elevation, and incident solar radiation (Thornton and others, 2018).

The land-use data for the short-term historical time period come from the U.S. Department of Agriculture National Agricultural Statistical Service (NASS) Cropland Data Layer (CDL) for 2016 (Boryan and others, 2011; U.S. Department of Agriculture, National Agricultural Statistics Service, 2016). The 2016 data were used for the entire short-term historical model (2000–17). Figure 4 shows the area of the model area associated with each of the major land-use categories for the 2016 data. To allow comparisons to be made with the earlier time period data, the specific land-use and crop descriptions for the 2016 CDL were condensed into categories that matched the historical backcast data. The proportion of cropland that is assumed irrigated is shown as an overlay in figure 4.

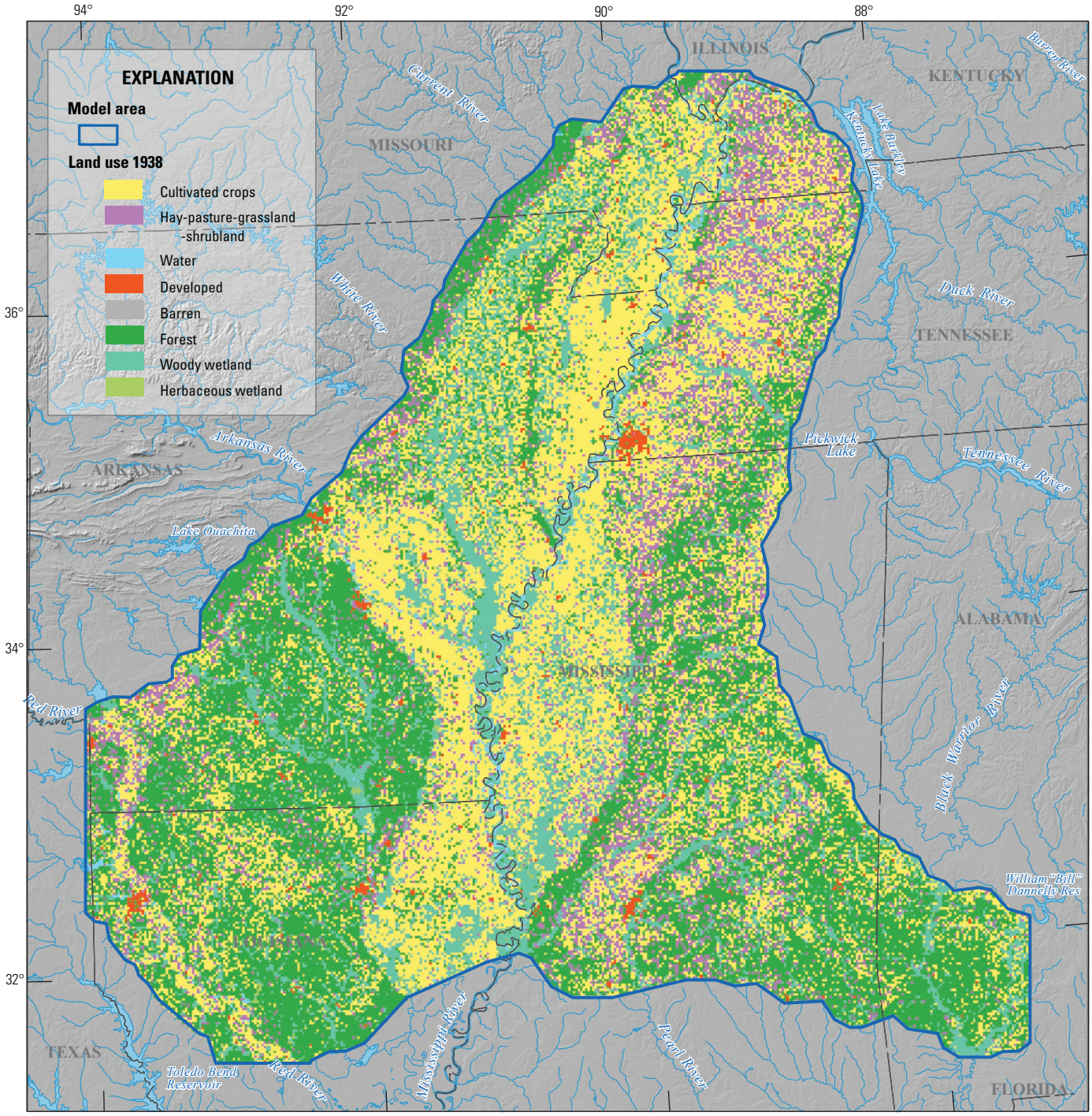


Figure 3. The 1938 distribution of land-use within the Mississippi embayment Soil-Water-Balance model area.

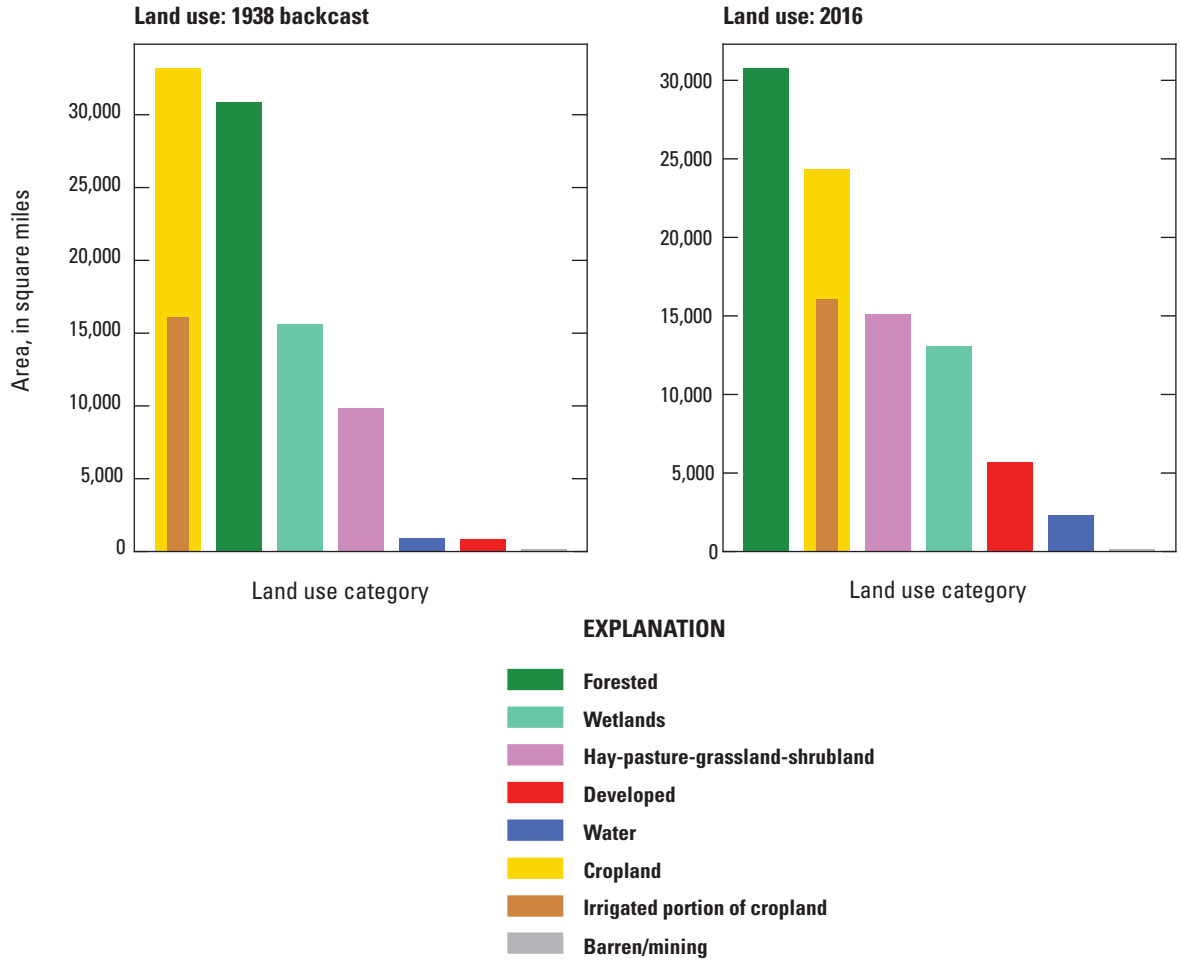


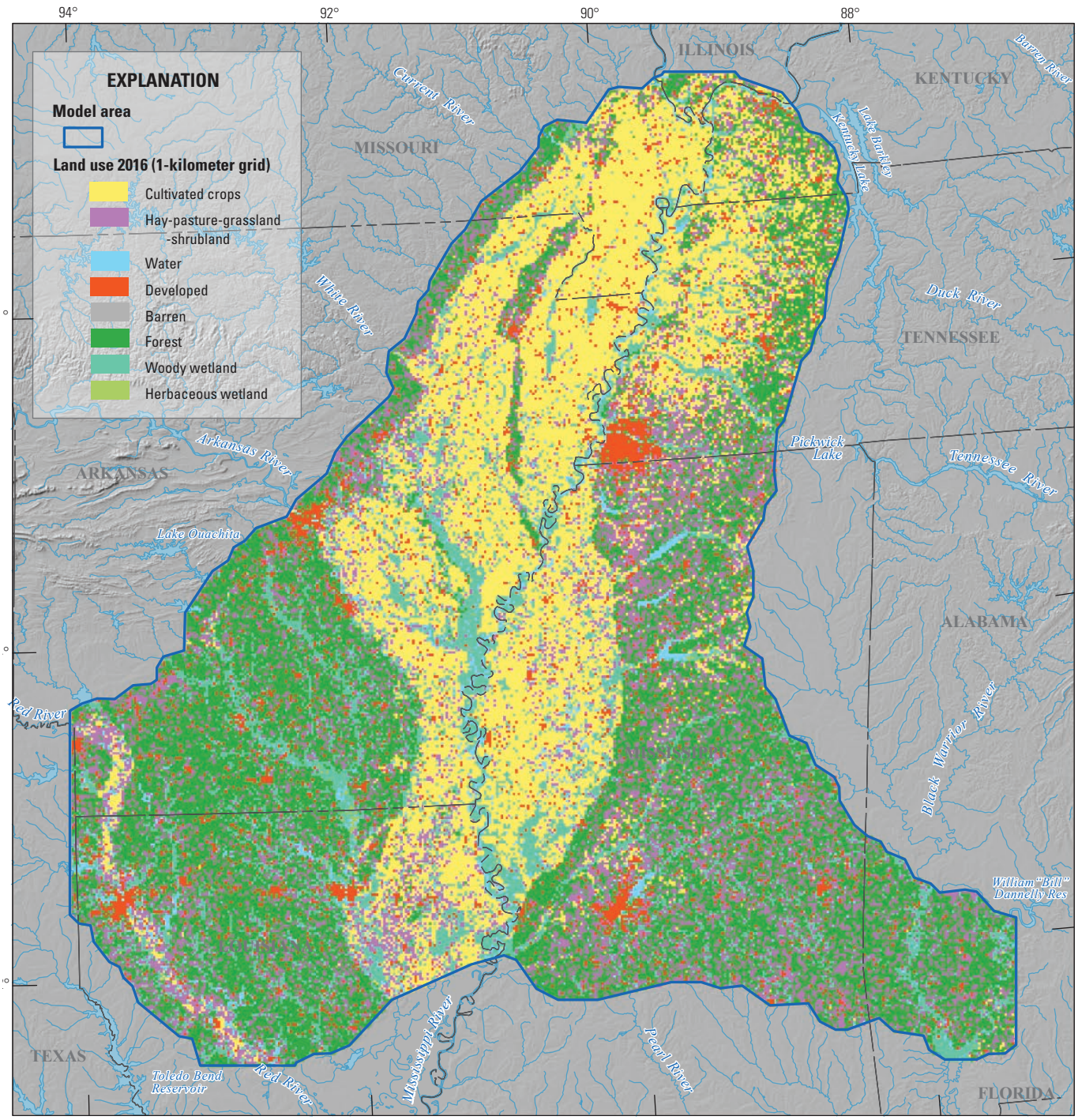
Figure 4. Abundance of each land-use category for 1938 and 2016 within the model area.

Table 3. Data sources for the short-term historical Soil-Water-Balance application.

[USDA, U.S. Department of Agriculture]

Dataset type	Dataset name	Date range	Spatial resolution	Reference
Daily weather	DayMet	2000–17	1,000 meters	Thornton and others (2018)
Land use	Cropland Data Layer	2016	30 meters	Boryan and others (2011), USDA, National Agricultural Statistical Service, 2016

The spatial distribution of the land-use categories for 2016 is shown in figure 5. Compared to the distribution of cropland in 1938 (fig. 3), the total amount of cultivated cropland in the Mississippi Alluvial Plain increased, whereas cropland in the rural uplands outside the alluvial plain decreased, having converted to forest, developed areas, grassland, hay, pasture, and shrubs.



States from U.S. Geological Survey digital line graphs, 1:2,000,000, USA Contiguous Albers Equal Area Conic coordinate system, U.S. Geological Survey version, North American Datum of 1983
Landform imagry from U.S. Geological Survey National Map 3D Elevation Program (3DEP)
Hydrography from U.S. Geological Survey National Hydrography Dataset

Land use data for 2016 generalized from U.S. Department of Agriculture Cropland Data Layer, 2016, 1-kilometer grid

0 25 50 100 150 MILES
0 25 50 100 150 KILOMETERS

Figure 5. The 2016 distribution of land-use within the Mississippi embayment Soil-Water-Balance model area.

Parameter Estimation and Observation Data

Calibration of the SWB model parameters was done using parameter estimation (Parameter ESTimation, or "PEST"). The PEST++ (Welter and others, 2015) software was used to calibrate the short-term historical SWB model described above using observation data spanning the years from 2000 to 2017. Observations of stream base flows and watershed evapotranspiration (ET) were used as targets in the parameter estimation of model parameters. The best-fit calibrated parameters were used for the final runs of the long-term historical SWB model and the short-term historical SWB model.

Observation Data

Observation data for the PEST parameter estimation consisted of base-flow estimates from processed streamflow records at USGS streamflow gaging stations and a series of monthly gridded actual ET estimates generated from satellite data and ground-based flux towers (Reitz and others, 2017). Both types of observations were processed at the watersheds scale for this study. The base-flow estimates were calculated at an annual time scale, and the actual ET estimates were calculated monthly.

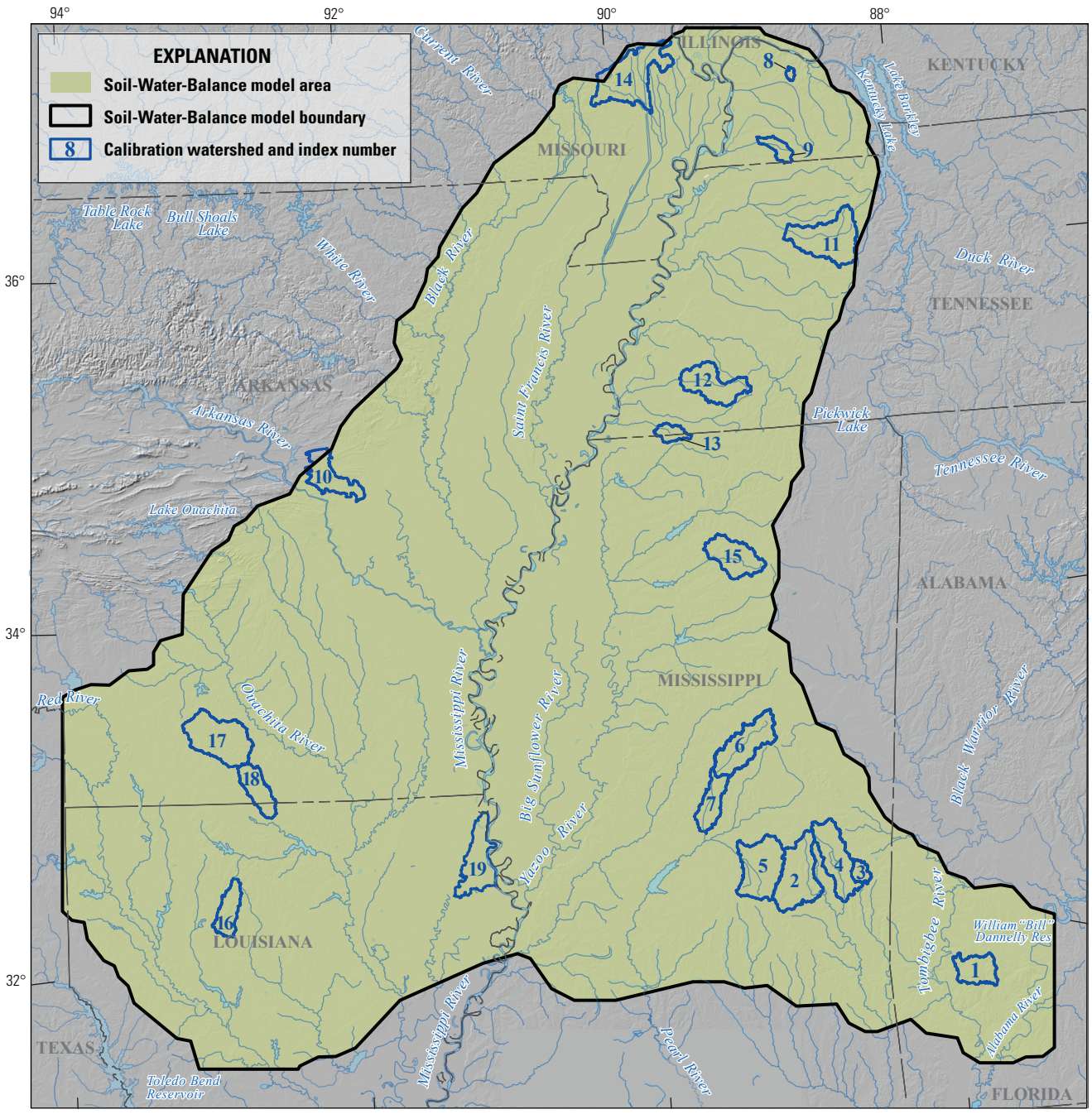
Base Flow Separation Analysis

Indirect estimates of groundwater recharge may be made through the application of base-flow separation techniques to daily surface-water streamflow records (Healy, 2010). Base flow is often assumed to represent long-term discharge from groundwater, which should equal groundwater recharge if ET from the water table is negligible and the groundwater and surface-water watersheds are coincident (Healy, 2010). Nineteen USGS streamflow gaging stations (U.S. Geological Survey, 2020) in the GAGES II dataset (Falcone, 2011), with continuous streamflow records that coincided with the calibration period (2000–17), were selected as calibration targets using the criteria that the gages have a low amount of surface-water diversion or other hydrological disturbances present within their drainage areas. The watersheds associated with the gaging stations that were used to estimate base flow are listed in table 4 and shown in figure 6. Appendix 2 contains a summary of the steps used to include base flow separation data values in the PEST calibration process. The base flow separation observations were assigned to an observation group called "ni," for "net infiltration" in the parameter estimation process.

Table 4. U.S. Geological Survey streamflow gaging stations used for base flow separation.

[Streamflow data can be accessed at U.S. Geological Survey (2020) using the site number. mi², square mile; AL, Alabama; NR, near; MS, Mississippi; KY, Kentucky; TN, Tennessee; No., number; MO, Missouri; AR, Arkansas; LA, Louisiana]

Map index number (fig. 6)	U.S. Geological Survey site number	Station name	Watershed area (mi²)
1	02469800	SATILPA CREEK NEAR COFFEEVILLE AL	163.2
2	02475500	CHUNKY RIVER NR CHUNKY, MS	369.7
3	02476500	SOWASHEE CREEK AT MERIDIAN, MS	51.8
4	02476600	OKATIBBEE CREEK AT ARUNDEL, MS	351.0
5	02483000	TUSCOLAMETA CREEK AT WALNUT GROVE, MS	355.2
6	02484000	YOCKANOOKANY RIVER NR KOSCIUSKO, MS	303.3
7	02484500	YOCKANOOKANY RIVER NR OFAHOMA, MS	467.7
8	03611260	MASSAC CREEK NEAR PADUCAH, KY	13.3
9	07024000	BAYOU DE CHIEN NEAR CLINTON, KY	68.7
10	07024500	SOUTH FORK OBION RIVER NEAR GREENFIELD, TN	206.6
11	07030240	LOOSAHATCHIE RIVER NEAR ARLINGTON, TN	382.6
12	07032200	NONCONNAH CREEK NEAR GERMANTOWN, TN	261.7
13	07043500	Little River Ditch No. 1 near Morehouse, MO	67.9
14	07264000	Bayou Meto near Lonoke, AR	436.3
15	07274000	YOCONA RIVER NR OXFORD, MS	241.4
16	07352000	Saline Bayou near Lucky, LA	148.4
17	07362100	Smackover Creek near Smackover, AR	384.6
18	07366200	Little Corney Bayou near Lillie, LA	169.1
19	07369500	Tensas River at Tendal, LA	278.4



States from U.S. Geological Survey digital line graphs, 1:2,000,000, USA Contiguous Albers Equal Area Conic coordinate system, U.S. Geological Survey version, North American Datum of 1983
Hydrography from U.S. Geological Survey National Hydrography Dataset
Landform imagry from U.S. Geological Survey 3D Elevation Program (3DEP)

Land use data for 2016 generalized from U.S. Department of Agriculture Cropland Data Layer, 2016, 1-kilometer grid

0 25 50 100 150 MILES
0 25 50 100 150 KILOMETERS

Figure 6. Watersheds associated with streamflow gaging stations selected for observations of base flow and evapotranspiration, with index numbers from table 6.

Actual Evapotranspiration

Actual ET values generated by SWB from the short-term historical SWB model were compared to actual ET estimated from a statistical relation between landscape variables and remotely sensed datasets of land surface temperatures (Reitz and others, 2017). Adding actual ET data in conjunction with the base flow separation data helps to constrain the parameter optimization process and is expected to result in more realistic calibration parameters (Nielsen and Westenbroek, 2019; Hunt and others, 2006). The ET observations were calculated for the same watersheds used for the base-flow separation observations. Appendix 2 contains a simplified workflow for incorporating remotely sensed actual ET into the PEST control file. The ET observations were assigned to an observation group named "et" in the parameter estimation process.

Parameter Estimation

The tabular SWB input has dozens of tabular input values that can affect the projected water-budget component values and can be treated as model parameters. A complete listing and description of all potential SWB version 2.0 parameters are listed in appendix 3 of USGS Techniques and Methods Report 6A–59 (Westenbroek and others, 2018). The Mississippi embayment Soil-Water-Balance model described in this report used the optional Food and Agriculture Organization of the United Nations Drainage and Irrigation Paper 56 (FAO56) approach (Allen and others, 1998; Westenbroek and others, 2018) for simulating evapotranspiration (table 5 summarizes the types of processes simulated in SWB; these processes are controlled by tabular inputs that could be treated as parameters that the authors felt were most important to include in the parameter estimation effort for this project, based on how sensitive the SWB output would be to those parameters.

During the calibration process using PEST, parameters were allowed to vary within set upper and lower limits until the best fit was obtained between the model-generated values for the set of observations and the observation values themselves. Some tabular input values were treated as a single parameter across all soil hydrologic groups in a given land-use class, and some tabular input values were treated as a single parameter for all land-use classes in a given soil hydrologic group. Several parameters that varied across soil hydrologic groups were made to preserve a relative ranking within the tabular data category across the soil hydrologic groups. For example, the PEST procedure was only allowed to modify the curve number for A soils; the curve numbers for all remaining soil groups were determined from a set of "curve number aligner" equations (Hawkins and others, 2009), which keeps a relative ranking between the curves, eqs. 1 through 4. Use of the curve number aligner equations ensures that the optimized curve number values for each land use will reflect the relative differences in soil characteristics between the soil groups.

$$CN_A = CN_A \tag{1}$$

$$CN_B = 37.8 + 0.622 CN_A \tag{2}$$

$$CN_C = 58.9 + 0.411 CN_A \tag{3}$$

$$CN_D = 67.2 + 0.328 CN_A \tag{4}$$

where:

CN is the curve number with subscript corresponding to the hydrologic soil group.

Table 5. Soil-Water-Balance parameter descriptions.

[SWB, Soil-Water-Balance]

SWB process	Tabular input description	SWB parameter name
Runoff	Curve number, by soil number	cn_1, cn_2, cn_3, cn_4
Evapotranspiration and crop growth calculations	Basal crop development coefficients	kcb_ini, kcb_min, kcb_mid, kcb_end
Evapotranspiration and crop growth calculations	Crop coefficient time values	L_ini, L_dev, L_mid, L_late
Infiltration through the rooting zone	Rooting depth in inches, by soil number	rz_1, rz_2, rz_3, rz_4
Infiltration through the rooting zone	Maximum infiltration rate	Mxinfil1, mxinfil2, and others
Interception	Plant interception	growing_season_interception, nongrowing_season_interception

Rooting depths were allowed to change for each land use and soil combination. Rooting depths for the various soil types was enforced, similar to the relations asserted by Thornthwaite and Mather (1957), who wrote that:

> "One factor which complicates the relation between depth of rooting of a plant and the type of vegetation is that the same plants will send roots to different depths in different types of soil. Thus in a sandy soil plants tend to be more deeply rooted than in silts and clays. The rooting habit of plants in different types of soils, in this way tends to compensate somewhat for the different water holding capacities of soils." (p. 185).

In the parameter estimation scheme, the rooting depth for a given land use was allowed to vary for hydrologic soil group A only; rooting depths for hydrologic soil groups B, C, and D were assumed to take on values of 0.98, 0.96, and 0.94 times the value assigned to hydrologic soil group A. The original Thornthwaite and Mather (1957) table values indicate more dramatic differences between soil groups, although little evidence is presented for this assertion. The approach taken for this calibration is more conservative with respect to differences between hydrologic soil groups. The factors used in the PEST parameter optimization were chosen so that the resulting rooting depths would take on values proportional to their general particle size—greater rooting depths for sandier (group A) soils, with decreasing rooting depths for loamy (groups B and C) and clayey (group D) soils.

The basal crop development coefficient parameter values (Kc_min, Kc_ini, Kc_mid, Kc_end) and planting season length values (L_ini, L_dev, L_mid, and L_late) are typically some of the most sensitive parameters when applying the FAO56 methodology built into SWB (fig. 7). The planting season length values control the elapsed time since planting allocated to each plant growth stage. The basal crop development values define the amount of water required for plant growth for each growth stage. Another study, which used PEST for the calibration of the crop coefficient parameters for SWB (Nielsen and Westenbroek, 2019), determined that the crop growth coefficient values were more sensitive than most other parameters tested in the model.

For the SWB model parameter optimization, the general shape of the crop coefficient curve associated with each land use or crop type was left intact; meta parameters representing offsets (Loffset in fig. 7) and scale factors (Kmult and

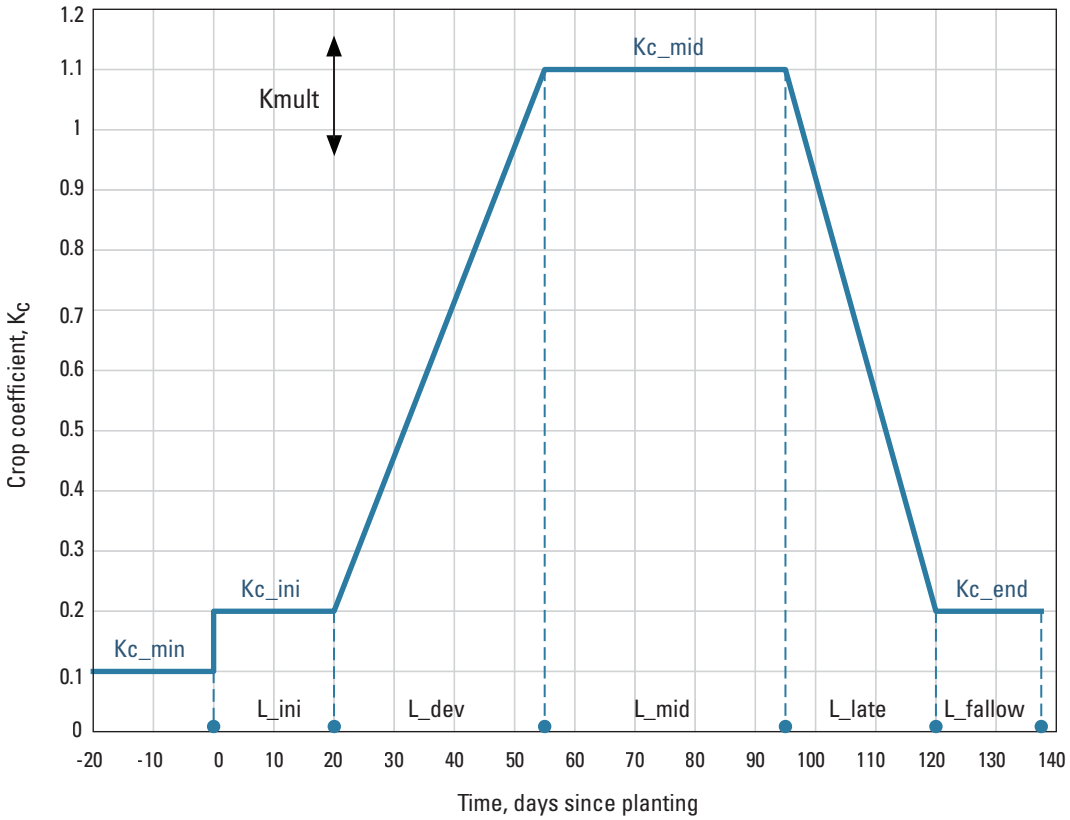


Figure 7. Generalized crop coefficient curve, with parameters identified to adjust the height and width of each curve section.

Lmult in fig. 7) were created to adjust the width and height of the crop coefficient curve of several land-use types during the parameter estimation process. The PEST++ software was allowed to optimize the meta parameters (scale factors and offset amounts); these meta parameters were used to shift and expand or contract the crop coefficient curves. Values of Kmult were multiplied by the initial Kcb values (K_ini, K_mid, K_late) to yield final model Kcb values; likewise, values of Loffset and Lmult were added and multiplied, respectively, to the time values of the Kcb curve (L_ini, L_dev, L_mid, L_late). Table 6 summarizes the way in which relations between parameter values were enforced for curve number, rooting depth, and crop coefficient parameter classes.

Calibration Results

Using the best-fit parameter values from the calibration, the model fit (comparison of simulated values with the observation values for watershed net infiltration and ET) is plotted in figure 8. (The final calibrated values of the tabular input datasets for SWB are given in the model archive for this report (Westenbroek and others, 2021a, 2021b). Figure 8*A* and 8*B* show the observation values versus the simulated equivalents, and figure 8*C* and 8*D* show the residuals, or the difference between the observed and simulated values plotted against the observation values. The current SWB simulation slightly overpredicts net infiltration (fig. 8*D*) for cases where the observed values are low (between 0 and about 7 inches [in.]); conversely, the SWB simulated values are underpredicted to varying degrees for cases where the observed values are high (10 in. and more). The residuals for the actual ET observations (fig. 8*C*) are mostly quite a bit lower than the residuals for the net infiltration observations, likely because each observation was weighted the same in the calibration. Because there were so many more actual ET observations, the parameter fitting gave more weight overall to the actual ET values.

For models to exhibit bias in residual plots is not uncommon; the bias shown in this report indicates that either the model input or the model structure is inadequate to allow simulation of higher net infiltration events. This is a phenomenon that could be investigated in future work.

Table 6. Parameters for which an assumed relation between terms was enforced.

[SWB, Soil-Water-Balance]

SWB model calculation	Parameter group	Notes
Runoff	Curve number	Only curve numbers for hydrologic soil group A soils could change; other values were determined by means of a set of "curve number aligner" equations.
Soil-moisture reservoir	Rooting depth	Only rooting depths for hydrologic soil group A soils were allowed to change; other values were determined by means of approximate relations as included in Thornthwaite and Mather (1957).
Evapotranspiration	Crop coefficient	Select crop coefficients could vary by applying a multiplier, the shape of the crop coefficient curve was allowed to expand and contract, and the initial planting date was allowed to vary.

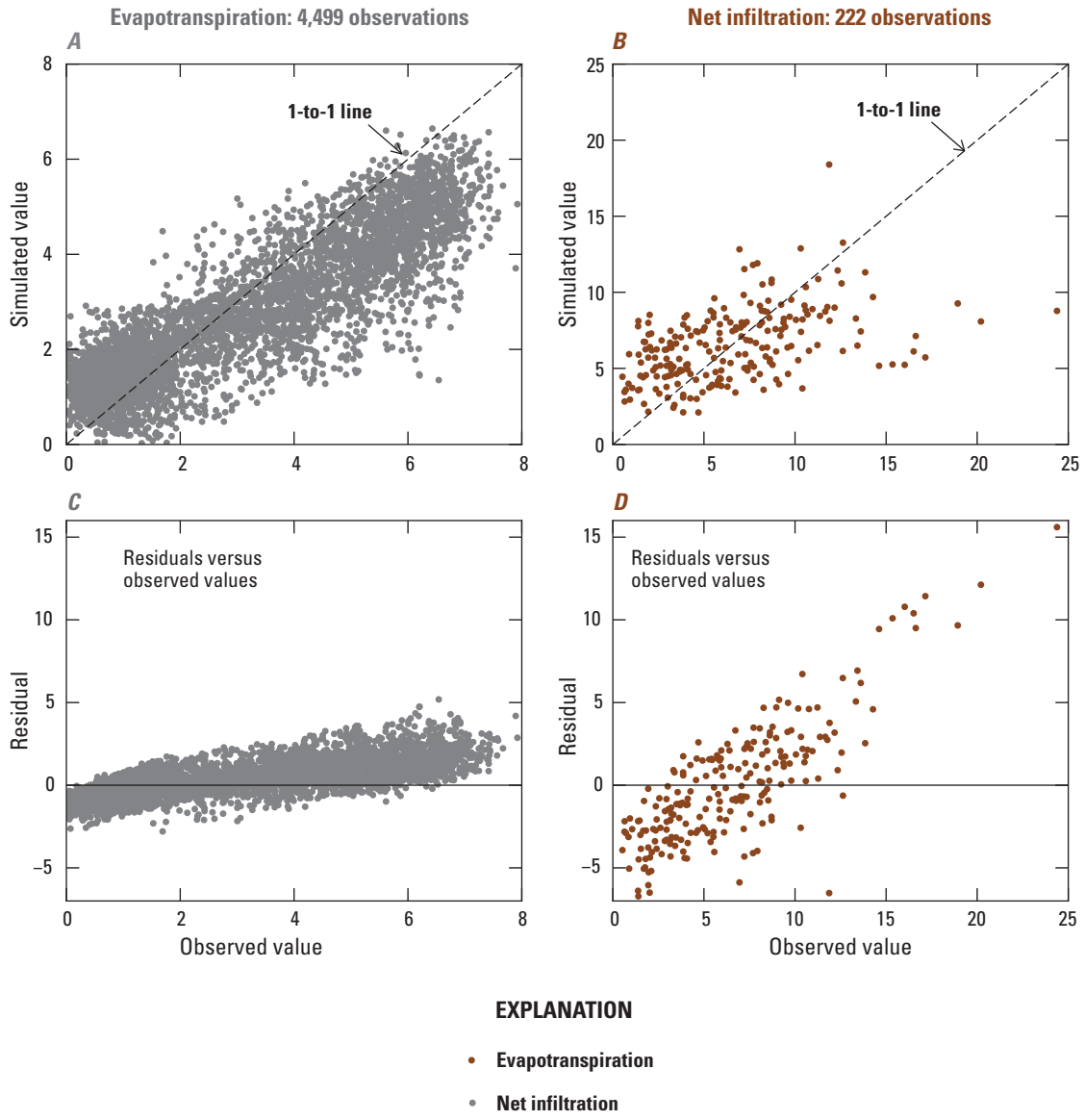


Figure 8. Relations between observed and simulated values for observations of evapotranspiration and net infiltration from 2000 to 2017 and plots of residuals compared to observed values. Observed versus simulated values of *A*, evapotranspiration observations and *B*, net infiltration observations. Residuals versus observed values of *C*, evapotranspiration observations and *D*, net infiltration observations.

Simulations of Net Infiltration and Irrigation, 1915–2017

Simulations of the net infiltration and irrigation from the two SWB models (the long-term historical period from 1915 to 2011 and the short-term historical period from 2000 to 2017), using the best-fit parameter values from the parameter estimation, were run for use in the groundwater simulations of the MERAS model area. Although the SWB model time step is daily, the groundwater-flow model uses inputs that are aggregated over larger time periods, for time steps that range from 5 years, for the early part of the 20th century (starting in 1920), to seasonal (6 months), from 1986 through 2017. Net infiltration varies considerably from year to year and also by land-use type.

The backcast land-use data from Sohl and others (2016) and daily weather data from Livneh and others (2013) were used for the SWB simulations from 1915 through 1999. DayMet daily weather data (Thornton and others, 2018) and the 2016 CDL (Boryan and others, 2011; U.S. Department of Agriculture, National Agricultural Statistics Service, 2016) were used for the SWB simulations for 2000–2017. When the two land-use data sources are combined, the land-use history through time (fig. 9) shows a generally gradual decline in the number of acres of cultivated crops and wetlands, and an increase in forested, hay-pasture-grassland-shrubland, and developed land uses. There are abrupt changes in the distribution of most land-use classes because the model input source changed in 2000 from the backcast data of Sohl and others (2016) to the 2016 CDL. The potential implications of these changes are discussed below.

Figure 10 shows a comparison of the annual DayMet and Livneh total precipitation amounts for 1986 through 2015 and includes the 11-year period when they overlap. On an annual total basis, the DayMet precipitation data are somewhat higher than the Livneh data for the period of overlap (fig. 10).

The time steps used to summarize the SWB-modeled net infiltration as input for the groundwater-flow models ranged from 1 to 5 years before 1986. Starting in 1986, the time steps in the groundwater model changed to 6-month seasonal simulations (fig. 10). The seasonal simulations have a distinct pattern of low recharge in the growing season (April through September) and higher recharge in the nongrowing season

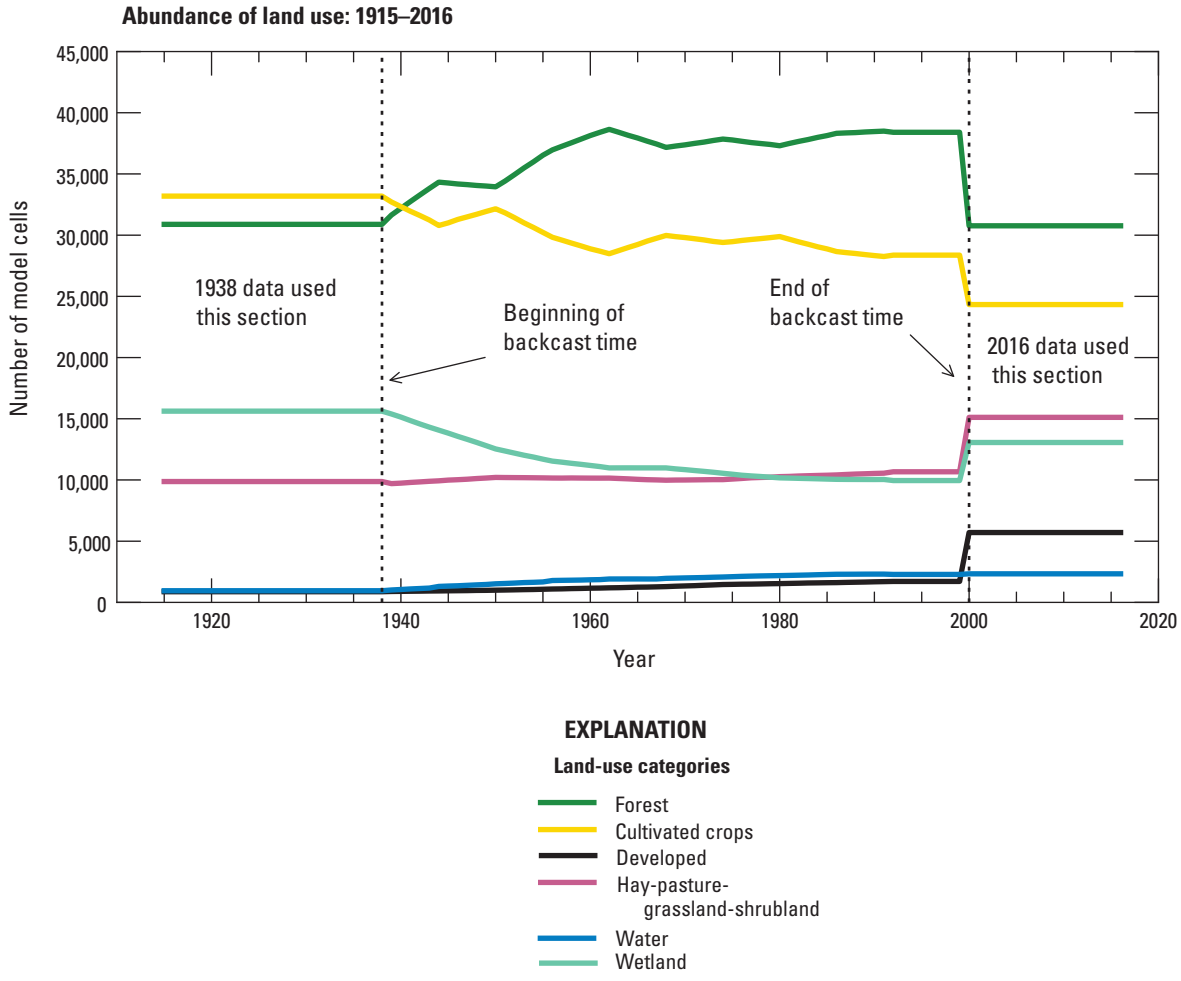


Figure 9. Land-use categories in the Mississippi embayment Soil-Water-Balance model area, 1915–2017.

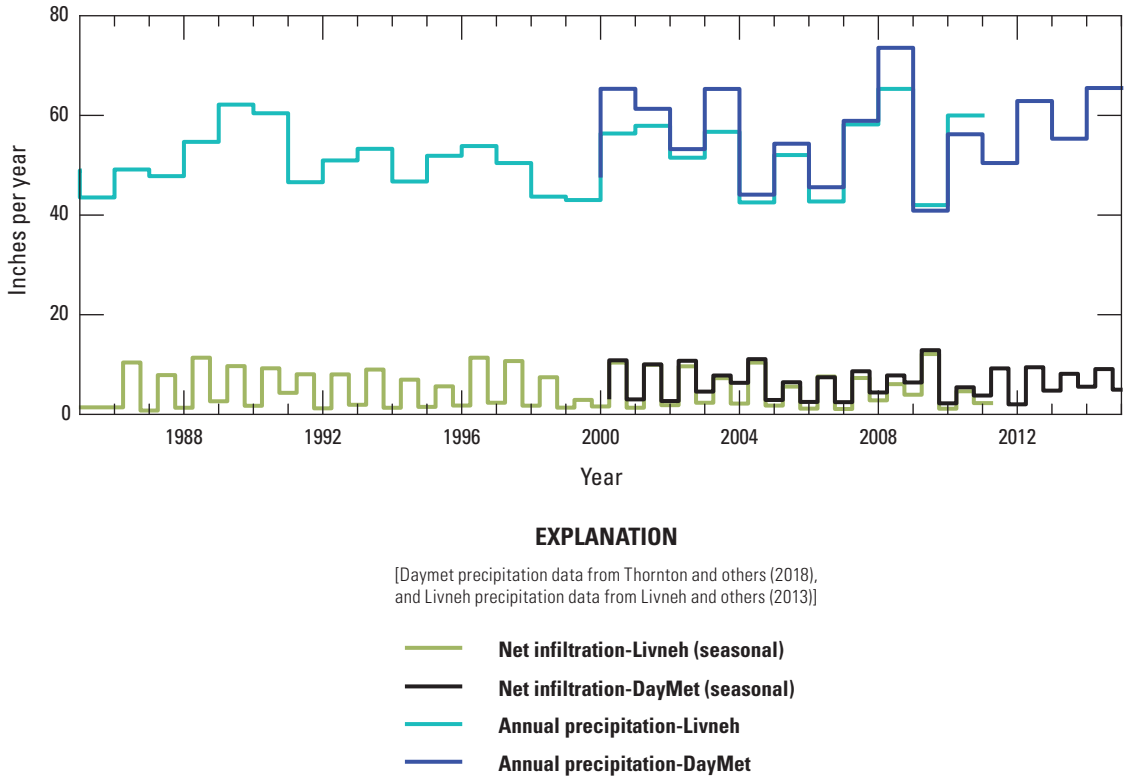


Figure 10. Annual precipitation inputs and seasonal net infiltration simulations for 1986—2015. Daymet precipitation data from Thornton and others (2018), and Livneh precipitation data from Livneh and others (2013).

(October through March). The final dataset used in the groundwater simulations switches from the Livneh-data subset to the DayMet-data subset in 2000 (where land-use data change as well).

Although the land-use categories vary with time, the SWB simulations of average monthly net infiltration and irrigation are not strikingly different during the two simulation periods (fig. 11). Furthermore, the absolute difference in net infiltration rates between forested and crop land uses is small. For both time periods, the average fall-winter net infiltration is a bit less than 1 inch per month for both land uses, with the summer net infiltration being much lower.

In the irrigation simulations, a fixed amount of the cropland was assumed irrigated, which was a smaller fraction of the total cropland in 1938 than in 2016. Actual historical irrigation practices were not simulated, because data on areas where irrigation was applied historically are not readily available. The amount of simulated irrigation applied to irrigated croplands in the 1940s is somewhat higher than for the 2005–2014 time period (assuming a constant amount of irrigated acreage with time). Estimates of simulated irrigation amounts are likely better for the short-term historical model (2000–2017) than for the long-term model (1915–2011).

The applied irrigation appears to boost the summertime net infiltration over cropland to be more than the forested net infiltration, but by a small fraction of the total simulated, applied irrigation amounts. Agricultural water use in the Mississippi embayment varies widely depending on the year, location within the embayment, and the crop (Powers, 2007); the average simulated irrigation amount for 2014–17 is about 10 in., with a minimum value of about 1.5 in. and a maximum value of about 40 in. The sum of monthly values shown in figure 11 is about 14 in., which is within the range of reported crop water use but is higher than the reported average value (about 10.8 in. for row crops, corn, and soybeans, 16.8 in. for all crops, for 1991–2006). Because the differences in recharge rate between the two primary land-use types in the area are small, the effect of the changes in the land-use distribution shown in figure 9 are likely to be small as well.

Previous estimates of net infiltration (recharge) rates in the MERAS model area range from nearly zero to about 5.7 inches per year (Sumner and Wasson, 1990; Arthur, 2001; Clark and Hart, 2009; Clark and others, 2013). SWB simulations of net infiltration for short-term model period (2000–17) are higher than the MERAS model values. The overall average net infiltration rate for the entire groundwater model simulation period (1920–2017) and the period after the simulations switched to seasonal simulations (1986–2017; table 7) was 5.4 inches per year, ranging from 2.1 to 9.6 inches per year, substantially higher than the older MERAS modeled rates. The seasonal rates for the growing season averaged 1.5 in., and the rates for the nongrowing season averaged 4.4 in. The recharge rates for forested land uses was somewhat lower than the overall average, and the recharge rates for cropland and

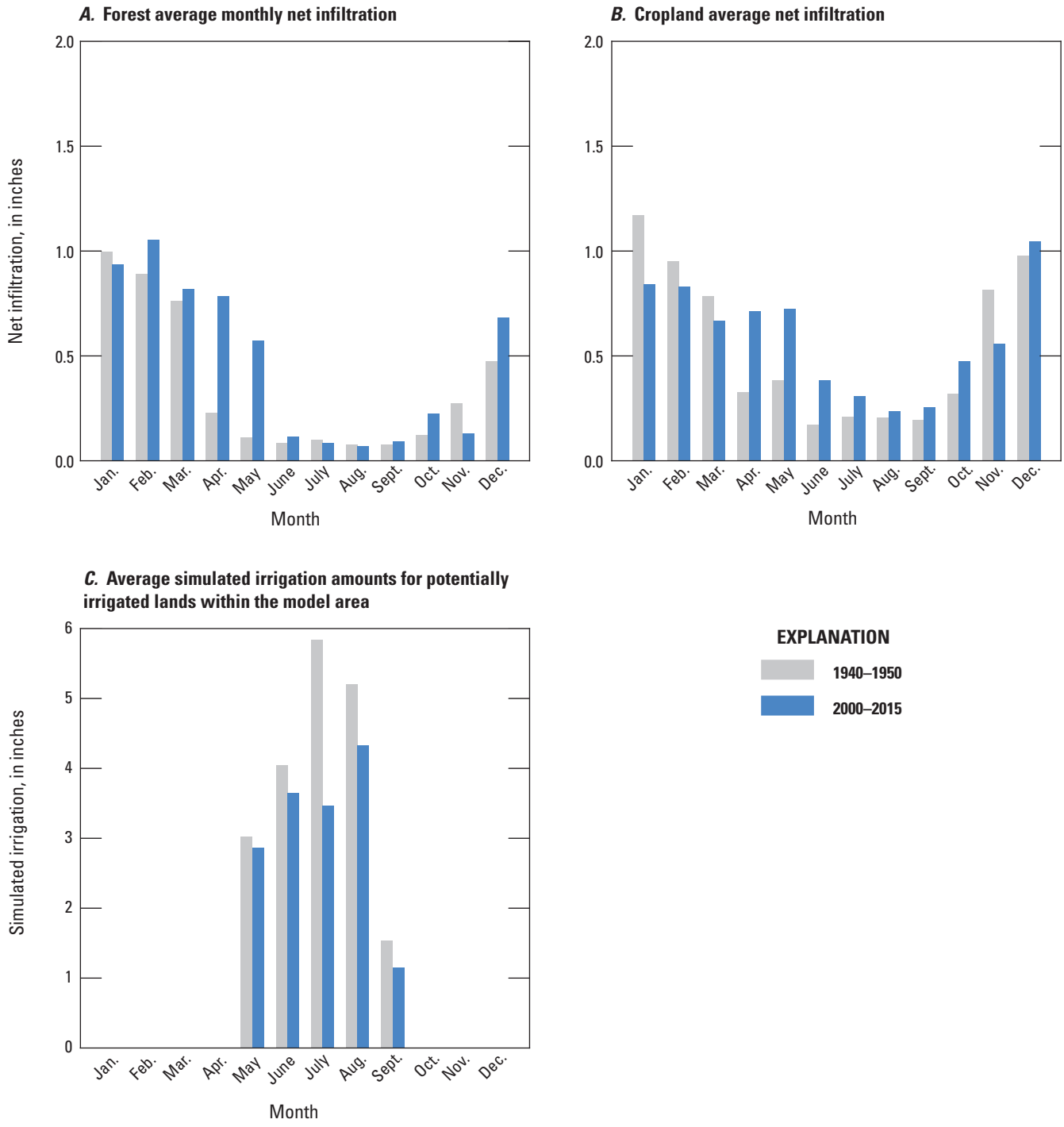


Figure 11. Graphs showing, *A*, forest average monthly net infiltration; *B*, cropland average net infiltration; and *C*, average simulated irrigation amounts for potentially irrigated lands within the model area.

hay-pasture-grassland-shrubland areas was higher than the overall average. Most of the recharge to the model area occurs during the nongrowing season for all land-use types (table 7).

Much of the difference between the current SWB model results and previous studies may be the result of structural issues in the model. SWB is known to perform less well when there are barriers to flow in the unsaturated zone beneath the root zone and above the water table (Nielsen and Westenbroek, 2019), which may be the case in some parts of this study area. Ongoing improvements to the model are aimed at using input datasets that will take areas where clay layers below the rooting zone may be impeding net infiltration into account.

Table 7. Statistical summary of net infiltration rates, annually and seasonally, for the Mississippi embayment Soil-Water-Balance model area.

Geographic area	1920–2017 annual rates (inches per year)			1986–2017 seasonal rates (inches per season)	
				Growing season	Nongrowing season
	Mean	Minimum	Maximum	Mean	Mean
Mississippi embayment model area (entire)	5.4	2.1	9.6	1.5	4.4
Cultivated crops	6.3	3.7	9.6	2.1	4.5
Forest	4.4	0.5	9.2	1.2	4.0
Hay-pasture-grassland-shrubland	8.2	4.7	12.2	1.9	6.5
Wetlands	3.7	1.2	8.2	1.1	3.1
Developed and other	4.4	1.9	8.1	1.1	4.0

Possible Improvements for Future Work

The results presented here represent the best estimates possible at the time this report was published. Researchers extending this work may consider adding data from numerous additional sources either as direct model input or as a set of observed data in the calibration process. These additional or alternate data sources could include the following:

Evapotranspiration estimates from eddy covariance towers.— Tower data may be difficult to use as an observation type, because SWB simulates ET for a grid cell, whereas the eddy covariance tower provides a time series of ET estimates for a much smaller footprint, but an attempt at using these data could be done.

Dynamic crop data layers (2010–2017).—This initial version of the model used a single estimate of land-use coverage (2016) for the entire short-term historical simulation (2000–2017). Although this estimate is adequate as a first cut at the water-budget components, substituting a more dynamic set of land-use grids into SWB will result in a more accurate representation of the crops that are grown and should improve irrigation water estimates.

Remotely sensed estimates of irrigated land.—A better representation of which fields receive irrigation water, updated on an annual basis, would improve water use estimates generated by SWB.

Independent estimates of total irrigation.—Reported irrigation volumes are available as monthly grids for much of the study area (the Aquaculture and Irrigation Water Use Model) (Wilson, 2021; Wilson and others, 2019), which could be aggregated into county-level data and used as calibra-tion targets.

PRISM daily weather data.—Replacing DayMet daily weather data with PRISM data from Oregon State University (PRISM Climate Group, Oregon State University, 2012) daily weather data would remove one factor that currently makes it more difficult to compare SWB output to other water-budget tools such as the Empirical Water Budget (Reitz and others, 2017) and the Aquaculture and Irrigation Water Use Model (Wilson, 2021).

gNATSGO soils data.—In 2020, the NRCS released a product that fills in gaps in the gSSURGO gridded soils data with STATSGO in a similar process to what was done for this study using 2014 data. The 2020 data from NRCS would be an improved input dataset to work with.

Extension of the model area beyond the Mississippi Embayment Regional Aquifer Study model boundary to parts of coastal Louisiana and Mississippi within the Mississippi Alluvial Plain could also be considered to provide similar model simulation results to a wider geographic area.

Summary and Conclusions

The Mississippi embayment encompasses about 100,000 square miles and covers parts of eight States. In 2016, the U.S. Geological Survey began updating previous work for a part of the embayment known as the Mississippi Alluvial Plain to support informed water use and agricultural policy in the region. Groundwater, water use, economic, and other related models are being combined with field surveys and observations to create a quantitative framework for evaluating regional groundwater withdrawals and their effects on long-term water availability in the Mississippi Alluvial Plain.

The Mississippi embayment Soil-Water-Balance model constructed for this project was calibrated using a combination of annual base-flow (recharge) data from 19 watersheds and monthly evapotranspiration data for the 2000–2017 period. Using the calibrated model parameters, estimates of net infiltration of water past the root zone and crop irrigation water amounts were simulated for 1915–2017 using a set of

two overlapping input data sources. Although the model time step is daily, the summary statistics are reported for time steps that range from 5 years, for the early part of the 20th century, to seasonal (6 months), from 1986 through 2017. Net infiltration varies considerably from year to year and also by land-use type.

The most abundant land uses simulated were forested and cultivated crops; for the earliest time periods (until about 1940), cropland was slightly more abundant than forested land, but during the rest of the 20th century, cropland decreased and forested land increased. In the simulations, a fixed amount of the cropland was assumed irrigated, which was a smaller fraction of the total cropland in 1938 than in 2016. Actual historical irrigation practices were not simulated because data on areas where irrigation was applied historically are not readily available.

The average net infiltration rates for the entire time period was 5.4 inches per year, which is substantially higher than what has been modeled for this area in the past. The range in annual simulated net infiltration rates for this study ranged from 2.1 to 9.6 inches per year. Cultivated crops and hay-pasture-grassland-shrubland areas had higher simulated infiltration rates than forested and other land-use categories. The Soil-Water-Balance model is sensitive to the rooting depths used in the simulations, and shallow-rooted land uses allow more of the initial infiltration to pass beyond the rooting zone than deeper-rooted plants (forests).

Net infiltration of water beyond the rooting zone is much greater during the nongrowing season when evapotranspiration pressures are lower than during the growing season. The 1986–2017 average of total net infiltration during the growing season was 1.5 inches for all land-use types, and the amount during the nongrowing season was 4.4 inches The cultivated crops had relatively more net infiltration during the growing season, which could be because of the shallow rooting zones and (or) the applied irrigation, or a combination of the two.

Improvements for future model development could include incorporation of a wider variety of observation datasets for use in model calibration, updated soil input data, and extension of the model area beyond the Mississippi Embayment Regional Aquifer Study model boundary to encompass parts of coastal Louisiana and Mississippi within the Mississippi Alluvial Plain.

References Cited

Alhassan, M., Lawrence, C., Richardson, S., and Pindilli, E.J., 2019a, The Mississippi Alluvial Plain aquifers—An engine for economic activity: U.S. Geological Survey Fact Sheet 2019–3003, 4 p., accessed May 11, 2020, at https://doi.org/10.3133/fs20193003.

Alhassan, M., Lawrence, C., Richardson, S., and Pindilli, E.J., 2019b, The Mississippi Alluvial Plain aquifers—An engine for economic activity: U.S. Geological Survey data release, accessed May 11, 2020, at https://doi.org/10.5066/P9RW8Y2A.

Allen, R.G., Pereira, L.S., Raes, D., and Smith, M., 1998, Crop evapotranspiration—Guidelines for computing crop water requirements: Rome, Food and Agriculture Organization of the United Nations, Irrigation and Drainage Paper 56, 174 p. [Also available at http://www.fao.org/3/X0490E/x0490e00.htm#Contents.]

Arthur, J.K., 2001, Hydrogeology, model description, and flow analysis of the Mississippi River alluvial aquifer in northwestern Mississippi—Pearl, Miss: U.S. Geological Survey Water-Resources Investigations Report 01–4035, 47 p.

Boryan, C., Yang, Z., Mueller, R., and Craig, M., 2011, Monitoring US agriculture—The US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program: Geocarto International, v. 26, no. 5, p. 341–358.

Brown, J.F., and Pervez, M.S., 2014, Merging remote sensing data and national agricultural statistics to model change in irrigated agriculture: Agricultural Systems, v. 127, p. 28–40.

Clark, B.R., and Hart, R.M., 2009, The Mississippi Embayment Regional Aquifer Study (MERAS)—Documentation of a groundwater-flow model constructed to assess water availability in the Mississippi embayment: U. S. Geological Survey Scientific Investigations Report 2009–5172, 61 p.

Clark, B.R., Westerman, D.A., and Fugitt, D.T., 2013, Enhancements to the Mississippi Embayment Regional Aquifer Study (MERAS) groundwater-flow model and simulations of sustainable water-level scenarios: U.S. Geological Survey Scientific Investigations Report 2013–5161, 29 p., accessed December 14, 2020, at https://pubs.usgs.gov/sir/2013/5161/.

Daly, C., Neilson, R.P., and Phillips, D.L., 1994, A statistical-topographic model for mapping climatological precipitation over mountainous terrain: Journal of Applied Meteorology, v. 33, no. 2, p. 140–158.

Falcone, J., 2011, GAGES-II—Geospatial attributes of gages for evaluating streamflow: U.S. Geological Survey database, accessed June 10, 2019, at https://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml.

Hawkins, R.H., Ward, T.J., Woodward, D.E., and Van Mullem, J.A., 2009, Curve number hydrology: American Society of Civil Engineers, 106 p.

Healy, R.W., 2010, Estimating groundwater recharge: Cambridge University Press, 245 p. [Also available at https://doi.org/10.1017/CBO9780511780745.]

Hunt, R.J., Feinstein, D.T., Pint, C.D., and Anderson, M.P., 2006, The importance of diverse data types to calibrate a watershed model of the Trout Lake Basin, Northern Wisconsin, USA: Journal of Hydrology (Amsterdam), v. 321, no. 1–4, p. 286–296.

Livneh, B., Rosenberg, E.A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K.M., Maurer, E.P., and Lettenmaier, D.P., 2013, A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States—Update and extensions: Journal of Climate, v. 26, no. 23, p. 9384–9392.

Mair, A., Hagedorn, B., Tillery, S., El-Kadi, A.I., Westenbroek, S., Ha, K., and Koh, G.W., 2013, Temporal and spatial variability of groundwater recharge on Jeju Island, Korea: Amsterdam, Journal of Hydrology, v. 501, p. 213–226.

Nielsen, M.G., and Westenbroek, S.M., 2019, Groundwater recharge estimates for Maine using a Soil-Water-Balance model—25-year average, range, and uncertainty, 1991 to 2015: U.S. Geological Survey Scientific Investigations Report 2019–5125, 58 p., accessed June 15, 2020, at https://doi.org/10.3133/sir20195125.

Painter, J.A., and Westerman, D.A., 2018, Mississippi Alluvial Plain extent, November 2017: U.S. Geological Survey data release, accessed June 15, 2020, at https://doi.org/10.5066/F70R9NMJ.

Pirellas, A., 2019, Approccio metodologico comparato per la determinazione della capacità di infiltrazione dei suoli dell'area di Muravera: Sardinia, Italy, University of Cagliari, 84 p.

Powers, S., 2007, Agricultural water use in the Mississippi Delta: Jackson, Miss,37th Annual Mississippi Water Resources Conference.

PRISM Climate Group, Oregon State University, 2012, United States average annual precipitation: Corvallis, Oreg., Oregon State University, Northwest Alliance for Computational Science & Engineering, digital data, accessed February 10, 2021, at https://prism.oregonstate.edu/recent/.

Reitz, M., Sanford, W.E., Senay, G.B., and Cazenas, J., 2017, Annual estimates of recharge, quick-flow runoff, and evapotranspiration for the contiguous U.S. using empirical regression equations: Journal of the American Water Resources Association, v. 53, no. 4, p. 961–983.

Shepard, D., 1968, A two-dimensional interpolation function for irregularly-spaced data, *in* ACM national conference, 23rd,1968, Proceedings: Association for Computing Machinery, p. 517–524.

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., accessed June 15, 2020, at https://doi.org/10.3133/sir20155038.

Sohl, T., Reker, R., Bouchard, M., Sayler, K., Dornbierer, J., Wika, S., Quenzer, R., and Friesz, A., 2016, Modeled historical land use and land cover for the conterminous United States: Journal of Land Use Science, v. 11, no. 4, p. 476–499.

Soil Survey Staff, 2014a, Gridded Soil Survey Geographic (gSSURGO) database for the conterminous United States: United States Department of Agriculture database, accessed May 5, 2015, at https://gdg.sc.egov.usda.gov.

Soil Survey Staff, 2014b, Web soil survey: U.S. Department of Agriculture web page, accessed March 13, 2014, at https://websoilsurvey.nrcs.usda.gov/.

Sumner, D.M., and Wasson, B.E., 1990, Geohydrology and simulated effects of large ground-water withdrawals on the Mississippi River alluvial aquifer in northwestern Mississippi: U.S. Geological Survey Water-Supply Paper 2292, 60 p.

Thornthwaite, C.W., and Mather, J.R., 1957, Instructions and tables for computing potential evapotranspiration and the water balance: Publications in Climatology, v. 10, no. 3, p. 1–104.

Thornton, P.E., Thornton, M.M., Mayer, B.W., Wei, Y., Devarakonda, R., Vose, R.S., and Cook, R.B., 2018, Daymet—Daily surface weather data on a 1-km grid for North America, version 3: National Aeronautics and Space Administration web page, accessed August 16, 2016, at https://dx.doi.org/10.3334/ORNLDAAC/1328.

Trost, J.J., Roth, J.L., Westenbroek, S.M., and Reeves, H.W., 2018, Simulation of potential groundwater recharge for the glacial aquifer system east of the Rocky Mountains, 1980–2011, using the Soil-Water-Balance Model: U.S. Geological Survey Scientific Investigations Report 2018–5080, 64 p., accessed June 15, 2020, at https://doi.org/10.3133/sir20185080.

U.S. Department of Agriculture, 2007, Hydrologic soil groups, *in* Part 630 Hydrology—National Engineering Handbook: U.S. Department of Agriculture, p. 14.

U.S. Department of Agriculture, 2016, National Agricultural Statistics Service—CropScape—Cropland Data Layer: U.S. Department of Agriculture database, accessed June 15, 2020, at https://nassgeodata.gmu.edu/CropScape/.

U.S. Geological Survey, 2020, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed March 20, 2020, at https://doi.org/10.5066/F7P55KJN.

Warmerdam, F., 2016, GDAL—Geospatial data abstraction library, version 2.1: Open Source Geospatial Foundation, accessed June 15, 2020, at http://www.gdal.org/.

Welter, D.E., White, J.T., Hunt, R.J., and Doherty, J.E., 2015, Approaches in highly parameterized inversion—PEST++ Version 3, a Parameter ESTimation and uncertainty analysis software suite optimized for large environmental models: U.S. Geological Survey Techniques and Methods, book 7, chap. C12, 54 p., accessed August 22, 2017, at https://doi.org/10.3133/tm7C12.

Westenbroek, S.M., Engott, J.A., Kelson, V.A., and Hunt, R.J., 2018, SWB version 2.0—A Soil-Water-Balance code for estimating net infiltration and other water-budget components: U.S. Geological Survey Techniques and Methods, book 6, chap. A59, 118 p., accessed October 2018, at https://doi.org/10.3133/tm6A59.

Westenbroek, S.M., Nielsen, M.G., and Ladd, D.A., 2021a, OFR 2021-1008 MODEL ARCHIVE—Soil-Water-Balance model developed to simulate net infiltration and irrigation water use for the Mississippi Embayment Regional Aquifer System, 1915 to 2018: U.S. Geological Survey data release, https://doi.org/10.5066/P98PBR80.

Westenbroek, S.M., Nielsen, M.G., and Ladd, D.E., 2021b, OFR 2021-1008 MODEL OUTPUT—Soil-Water-Balance net infiltration and irrigation water use output datasets for the Mississippi Embayment Regional Aquifer System, 1915 to 2018: U.S. Geological Survey data release, https://doi.org/10.5066/P9U484X5.

Wieczorek, M.E., 2014, Area- and depth- weighted averages of selected SSURGO variables for the conterminous United States and District of Columbia: U.S. Geological Survey Data Series 866, accessed February 10, 2016, at https://doi.org/10.3133/ds866.

Wilson, J.L., 2021. Aquaculture and Irrigation Water-Use Model (AIWUM) version 1.0—An agricultural water-use model developed for the Mississippi Alluvial Plain, 1999–2017: U.S. Geological Survey Scientific Investigations Report 2021–5011, 36p., https://doi.org/10.3133/sir20215011.

Appendix 1. Spatial Subset Creation

Input data grids were clipped and reprojected so they correspond to the bounds, resolution, and alignment of the underlying MODFLOW model grid. For each of the input data grids, the gdalwarp and gdal_translate command-line applications (Warmerdam, 2016) were used to create input data grids clipped and reprojected to conform to the underlying MODFLOW grid.

Mississippi Embayment Regional Aquifer Study Area

The boundaries used in all clipping operations are derived directly from the MODFLOW model for the Mississippi Embayment Regional Aquifer Study (MERAS) model area. The bounds for the Soil-Water-Balance model applied to the MERAS model area are given in figure 1.1.

Shellmound Inset

A smaller, higher-resolution subset of input data grids were prepared to support a MODFLOW model for an area near Shellmound, Mississippi. The bounds for this model are given in figure 1.2.

References Cited

Warmerdam, F., 2016, GDAL—Geospatial data abstraction library, version 2.1: accessed June 15, 2020, at http://www.gdal.org/.

```
# PROJ4 string used in all data reprojections – USGS Albers Equal Area
PROJ4=”+proj=aea +lat_1=29.5 +lat_2=45.5 +lat_0=23 +lon_0=-96 +x_0=0 +y_0=0
+datum=NAD83 +units=m +no_defs”

LLX=178388.99987163        # lower left-hand x-coordinate
LLY=938511.58387163        # lower left-hand y-coordinate
URX=812470.51617163        # upper right-hand x-coordinate
URY=1604779.9791716        # upper right-hand y-coordinate
RES=1609.34395000          # gridcell resolution in meters (1-mile cells)
```

Figure 1.1. PROJ4 string and bounding box coordinates used to subset datasets to the Mississippi Embayment Regional Aquifer Study model area.

```
PROJ4="+proj=aea +lat_1=29.5 +lat_2=45.5 +lat_0=23 +lon_0=-96
+x_0=0 +y_0=0 +datum=NAD83 +units=m +no_defs"
MAP_LLX=501405.              # lower left-hand x-coordinate
MAP_LLY=1175735.             # lower left-hand y-coordinate
MAP_URX=536505.              # upper right-hand x-coordinate
MAP_URY=1205835.             # upper right-hand y-coordinate
MAP_RES=100.                 # gridcell resolution in meters
```

Figure 1.2. PROJ4 string and bounding box coordinates used to subset datasets to the Shellmound inset of the Mississippi Embayment Regional Aquifer Study model area.

Appendix 2. Incorporating Observations into PEST++ Workflow

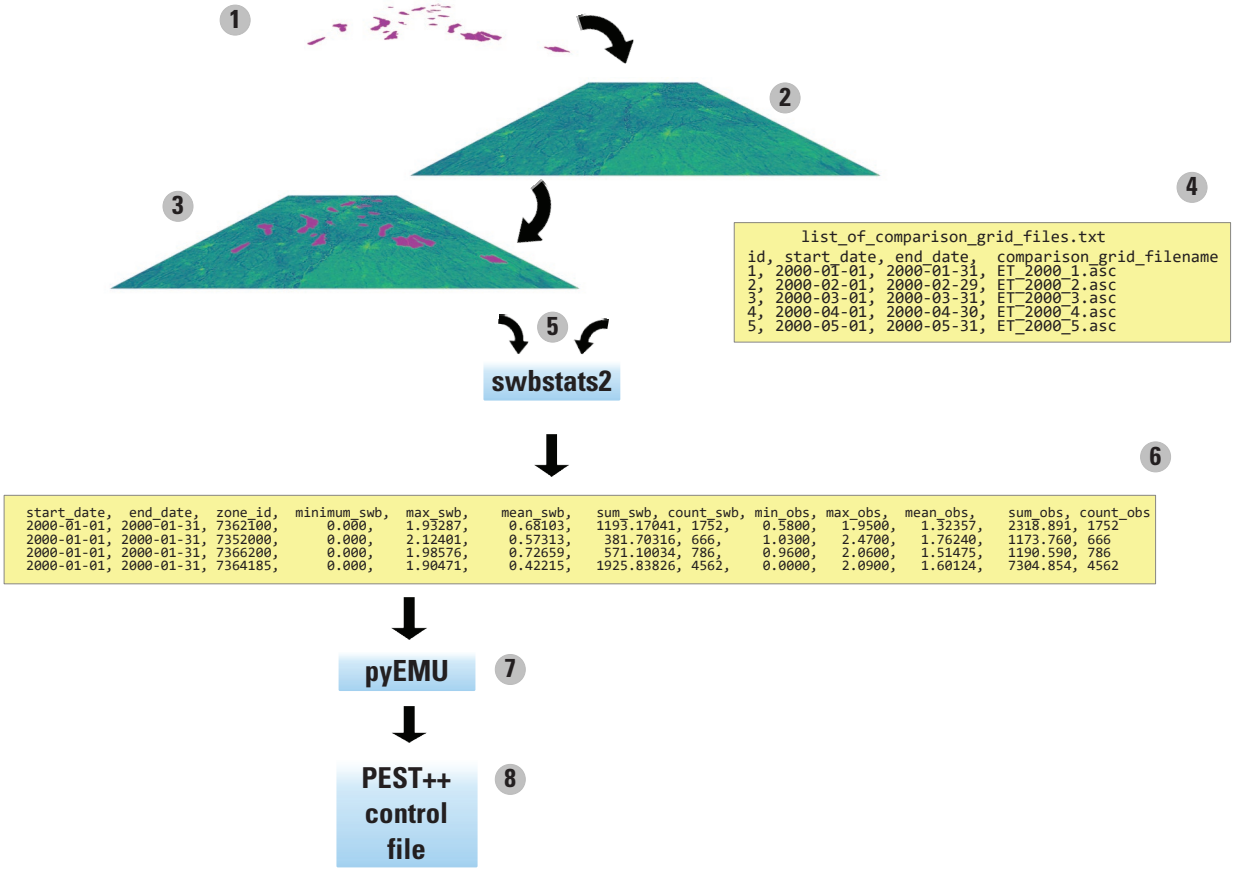
Comparing environmental model output to observed data always requires a degree of data carpentry in order to ensure that the comparisons are being made at appropriate time and spatial scales. The processes included here will likely change somewhat in planned future model calibration exercises as newer datasets are swapped into the process. All data processing scripts for future work are planned to be included in a model archive rather than in a report appendix.

The data reduction steps are usually very similar even if the specific data types are dissimilar. For observations that are already in gridded form, one straightforward way to match simulated and observation data is to calculate zonal statistics (sum, mean, range) for an identical set of polygon areas and time periods. For example, one might calculate the mean of zonal sums for identical watersheds for all monthly time steps in the simulation. For observations that are representative of point data, such as streamflow data, one can assume that the total base flow or streamflow can be represented by the sum of Soil-Water-Balance (SWB) output values for model cells within the drainage area upstream from the streamflow gaging station. In this case, only the model SWB output is subject to zonal statistics calculation.

This section focuses on the specific steps taken to incorporate (a) gridded evapotranspiration and (b) stream base-flow estimates into the mix of observations supplied to PEST++ (Welter and others, 2015) as part of parameter estimation.

Actual Evapotranspiration

Actual evapotranspiration based in part on satellite observations was used in the PEST++ parameter optimization process. For this work, a companion program to the SWB code (version 2 (swb2) was used. The swbstats2 code (Westenbroek and others, 2018) is designed to read the output Network Common Data Form files from swb2, calculate zonal statistics on the basis of a user-supplied set of zones, and optionally, calculate zonal statistics simultaneously on a set of observation grids (fig. 2.1)



EXPLANATION

1. Shapefile defining drainage basins for which statistics are desired
2. American Standard Code for Information Interchange (ASCII) grid of actual evapotranspiration (ET) values
3. Rasterized ASCII drainage basins overlayed on ET grid
4. Text file listing calculation start and end dates and name of ASCII ET grid
5. Text file, ASCII drainage basin outlines fed into 'swbstats2'
6. 'swbstats2' calculates mean value of ET for each of the drainage basins
7. Table of calculated actual ET values read into Python and processed with pyEMU (White and others, 2016)
8. pyEMU writes out new PEST++ (Welter and others, 2015) control file, now including the actual ET values as observations

Figure 2.1. Schematic showing steps taken to include remotely sensed actual evapotranspiration data in a PEST++ control file.

Several text files are required for use with swbstats2. A snippet of the first text file is shown in figure 2.2; this file lists comparison start and end dates and the corresponding observation grid to compare swb2output against during each observation time period. The command syntax for swbstats2 is shown in figure 2.3, and the resulting output is shown in figure 2.4.

The swbstats2 program is run from the command line (fig. 2.3), supplying the name of the comparison_grid file (fig. 2.2) and the name of the SWB output file.

The output lists the statistics for SWB (minimum_swb, maximum_swb, and so forth) and for the observation grids (minimum_obs, maximum_obs, and so forth). The preferred workflow at the time this document was written (August 2019) is to write a short Python script making use of the pyEMU package (White and others, 2016) to add the swbstats2 output to the collection of observations, then use the script to write out a new PEST++ control file. The script is shown in fig. 2.5 to give a sense of how the processing was done.

```
id,start_date,end_date,comparison_grid_filename
1,2000-01-01,2000-01-31, model/external_files/grid_actual_et/ET_2000_1.asc
2,2000-02-01,2000-02-29, model/external_files/grid_actual_et/ET_2000_2.asc
3,2000-03-01,2000-03-31, model/external_files/grid_actual_et/ET_2000_3.asc
4,2000-04-01,2000-04-30, model/external_files/grid_actual_et/ET_2000_4.asc
5,2000-05-01,2000-05-31, model/external_files/grid_actual_et/ET_2000_5.asc
```

Figure 2.2. File ('comparison_grid_filenames__actual_ET.csv') supplied to swbstats2 defining statistics calculation time periods and raw observed actual evapotranspiration grid filenames.

```
> ./swbstats2 --monthly_statistics --comparison_grid=comparison_grid_
filenames__actual_ET.csv  \
actual_et__2000-01-01_to_2017-12-31__888_by_845 .nc
```

Figure 2.3. Command syntax for swbstats2.

```
start_date  end_date      zone_id   minimum_swb maximum_swb mean_swb    sum_swb
count_swb minimum_obs maximum_obs  mean_obs     sum_obs     count_obs

1/1/2000     2000-01-31  7362100    0      1.93287      0.68103   1193.17041  1752
0.58  1.95     1.32357  2318.8916   1752

1/1/2000     2000-01-31  7352000    0      2.12401      0.57313   381.70316   666
1.03  2.47     1.7624   1173.76062  666

1/1/2000     2000-01-31  7366200    0      1.98576      0.72659   571.10034   786
0.96  2.06     1.51475  1190.59058  786

1/1/2000     2000-01-31  7364185    0      1.90471      0.42215   1925.83826  4562  0
2.09     1.60124  7304.85498  4562
```

Figure 2.4. Output of swbstats2 showing actual evapotranspiration statistics.

```
    import pandas as pd
    import numpy as np
    import pyemu

    # need to start with an existing PEST control file
    # could also use pyEMU to create a basic PEST control file
    my_pst_file           = 'map_swb_pest_calibration.pst'
    my_act_et_obs_file    = 'zonal_stats_actual_et.csv'

    # create a Pst object
    my_pst = pyemu.Pst(my_pst_file)
    my_act_et_obs = pd.read_csv(my_act_et_obs_file)

    my_model_obs = my_pst.observation_data

    # data cleanup: create a 'month' and 'year' column from the start date
    df = my_act_et_obs
    df=df.assign(month=pd.to_datetime(df.start_date).dt.month)
    df=df.assign(year=pd.to_datetime(df.start_date).dt.year)

    # now create a unique PEST observation by concatenating:
    # 1) zone id
    # 2) observation type ('et')
    # 3) year
    # 4) month
    df=df.assign(pest_obs_name=df.zone_id)
    for index in df.index:
        rec=df.loc[index,:]
        df.loc[index,'pest_obs_name']=str(rec.zone_id) + 'et' + str(rec.year) +
'{0:0{width}}'.format(rec.month, width=2)
    my_act_et_obs=df

    # merge new observations with existing observations
    df2 = pd.merge(left=my_model_obs,right=my_act_et_obs,left_on='obsnme',right_
on='pest_obs_name',how='left')

    # copy 'mean_obs' calculated by swbstats2 into pyEMU observations data structure
    for index in df2.index:
        if not np.isnan(df2.loc[index,'obsval']):
            df2.loc[index,'obsval']=df2.loc[index,'mean_obs']

    #
    for index in df2.index:
        if 'et' in df2.loc[index,'obsnme']:
            df3.loc[index,'obgnme'] = 'actual_et'
        if my_model_obs.loc[index,'obsnme']==df2.loc[index,'obsnme']:
            my_model_obs.loc[index,'obsval']=df2.loc[index,'obsval']
            my_model_obs.loc[index,'obsgnme']=df2.loc[index,'obgnme']

    # last, write out the new PEST++ control file
    my_pst.write(my_pst_file)
```

Figure 2.5. Example Python script showing incorporation of SWBSTATS2 output into pyEMU pst object for use in generating a new PEST++ control file.

Last, a small Python script (fig. 2.6) was written to create the PEST++ instruction file needed to allow PEST++ to process SWB-generated model output.

```
pif *
l1
l1 w w w w w !7362100et200001!
l1 w w w w w !7352000et200001!
l1 w w w w w !7366200et200001!
l1 w w w w w !7364185et200001!
l1 w w w w w !7073500et200001!
```

Figure 2.6. Snippet of PEST++ instruction file specifying that 'mean_swb' statistic be read and used as the observed actual et value

Net Infiltration (Potential Recharge)

Net infiltration amounts from the SWB model also were processed using swbstats2 software to yield a series of annual mean sums for each of the calibration basins, which also were used as observations in the parameter estimation process. These processed observations were supplied to PEST++ software. swbstats2 sums the values for each grid cell within the specified observation watershed for each year of the SWB simulation. The mean sum of net infiltration for a given year and basin is then calculated by dividing by the surface area of the watershed and applying a unit conversion factor. Figure 2.7 shows a simplified workflow for incorporating stream base flow observations into a PEST++ parameter optimization.

Base flow separations were calculated in the R statistical language (R Development Core Team, 2018), using the smwrStats package; this package includes an implementation of the PART and HYSEP base flow separation methods. The Python scripts for the last two steps given in fig. 2.7 are not shown; they are very similar to those discussed in the previous section for incorporating the actual evapotranspiration observations into the PEST++ parameter optimization workflow.

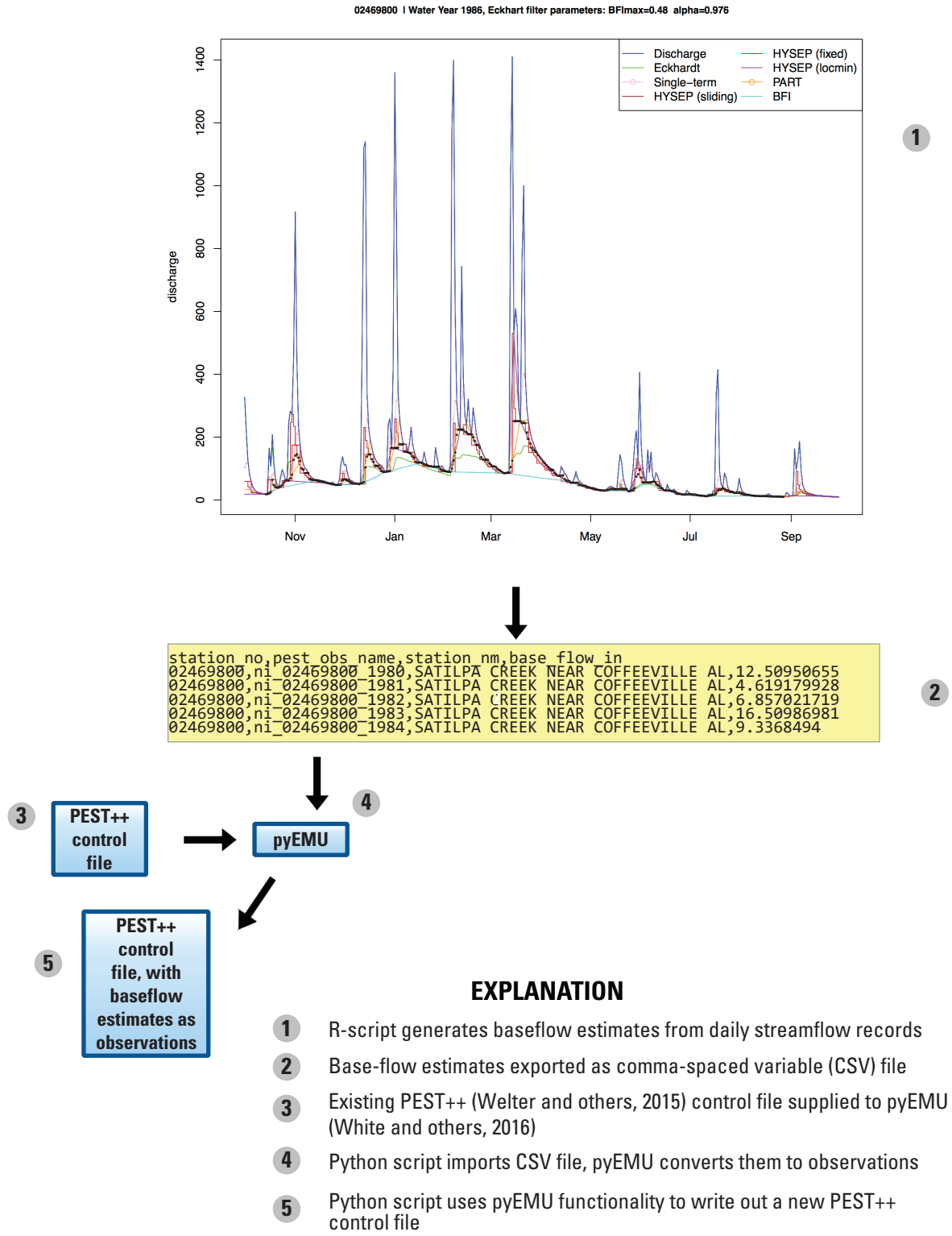


Figure 2.7. Schematic showing steps taken to include base-flow separation data in a PEST++ control file.

References Cited

R Development Core Team, 2018, R: A language and environment for statistical computing.: Vienna, Austria, R Foundation for Statistical Computing web page, accessed March 15, 2021, https://www.R-project.org.

White, J.T., Fienen, M.N., and Doherty, J.E., 2016, A python framework for environmental model uncertainty analysis: Environmental Modelling & Software, v. 85, p. 217–228.

Welter, D.E., White, J.T., Hunt, R.J., and Doherty, J.E., 2015, Approaches in highly parameterized inversion—PEST++ Version 3, a Parameter ESTimation and uncertainty analysis software suite optimized for large environmental models: U.S. Geological Survey Techniques and Methods, book 7, chap. C12, 54 p., accessed August 22, 2017, at https://pubs.er.usgs.gov/publication/tm7C12.

For more information, contact
Director, Upper Midwest Water Science Center
U.S. Geological Survey
8505 Research Way
Middleton, WI 53562

Publishing support provided by the
U.S. Geological Survey, Science Publishing Network,
Indianapolis Publishing Service Center

ISSN 2331-1258 (online)
https://doi.org/10.3133/ofr20211008