≡USGS
science for a changing world

INFORMATICS OF GROUNDWATER-LEVEL MEASUREMENTS, ANOMALOUS DATA DETECTION, AND EXPLORATORY LONG-TERM MONITORING NETWORK IDENTIFICATION FOR THE MISSISSIPPI RIVER VALLEY ALLUVIAL AQUIFER, USA
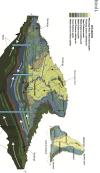
William H. Asquith, Ph.D., P.G., *Research Hydrologist*, U.S. Geological Survey, Lubbock, Texas; Ronald C. Seanor, *Hydrologic Technician*, U.S. Geological Survey, Ruston, Louisiana; Wade H. Kress, *Hydrologist*, U.S. Geological Survey, Nashville, Tennessee

THE MISSISSIPPI RIVER VALLEY ALLUVIAL AQUIFER

A START OF A REGIONAL WATER AVAILABILITY STUDY, the U.S. Geological Survey (USGS) is comprehensively studying and modeling the Mississippi River Valley alluvial aquifer (MRVA; fig. 1; https://www2.usgs.gov/water/lowermississippigulf/map/mrva.html)...

Figure 1. Cross-sectional view of associated hydrogeologic units to the Mississippi River Valley alluvial aquifer (MRVA) with post-development flow (2006) conceptual water-movement (Clark and others, 2011, fig. 10).

USGS Groundwater Database and Furnished Records

W ATER-LEVEL DATA currently are (*Great* December 2017) available for the MRVA and outliers compared to other measurements made at that well (*time-series outliers*) or the measurements for the well in relation to the measurements from neighboring wells (*spatial outliers*).

Figure 2. Mississippi Alluvial Plain Level III Ecoregion 8.224 kilometer buffer and wells color coded by unique years of water-level measurements across various agencies within the Mississippi River Valley alluvial aquifer through April 2017.

Figure 3. Hydrograph of well USGS:331745090260401 showing water-level measurements at neighboring wells, metadata, and outlier detection (residual > ±20 feet [SVM] trend lines — Well location is shown in figure 4.

GROUNDWATER-LEVEL INFORMATICS

I NFORMATICS involves the science and practice of *information processing* (https://www.informatics.jax.org/silverbook/)...
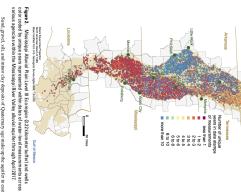
Anomalous Data Identification

G ROUNDWATER INFORMATICS FOR THIS STUDY are based on the R language (R Development Core Team, 2017) and specialized add-on packages of `kern1a.b` and `ngcv` (Karatzoglou and others, 2004; Wood, 2017)...

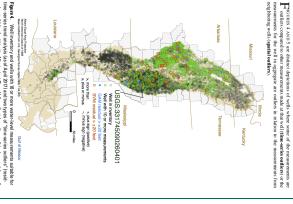**Figure 4.** Well inventory and wells with 10 or more water-level measurements suitable for time-series trend analysis (as of April 2017) and two types of "time-series outliers" (residual > 20 feet) within the Mississippi River Valley alluvial aquifer—A hydrograph for well USGS:331745090260401 is shown in figure 3.

OUTLYING MEASUREMENTS AND OUTLYING WELLS

F IGURES 4 AND 5 are distinct depictions of wells where some of the measurements are outliers compared to other measurements made at that well (*time-series outliers*)...
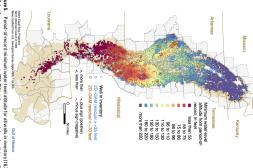

Figure 5. Period-of-record minimum water-level altitude for all wells in inventory (18,756 wells) with one or more measurements (429,028 total measurements) and outlying wells (as of August 2017) identified by the 2D-GAM within the Mississippi River Valley alluvial aquifer.

TRENDS — DECADAL ESTIMATION OF WATER-LEVEL ALTITUDE

F IGURES 6 AND 7 depict spatial patterns in water-level altitudes using arithmetic mean water-level altitude for a period-of-record from January 1960 to August 2017...
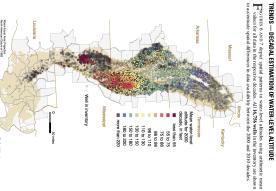

Figure 6. Mean decadal water-level altitude for the 2000 decade (January 2000–December 2009) based on 2,533 wells and 32,753 measurements for the Mississippi River Valley alluvial aquifer—Note the disparity in data availability between states and in areas of rapid change in water levels.
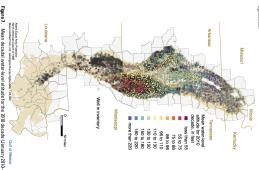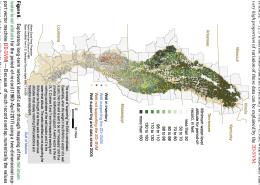

Figure 7. Mean decadal water-level altitude for the 2010 decade (January 2010–April 2017) based on 1,370 wells and 22,298 measurements for the Mississippi River Valley alluvial aquifer—Note reduction in data availability compared to the 2000 decade shown in figure 6.

LONG-TERM MONITORING NETWORK—EXPLORATORY EFFORTS

A TWO-DIMENSIONAL SVM (2D-SVM) was made to estimate minimum water-level altitude for a period-of-record from January 1960–April 2017 (13,218 wells, 149,156 measurements)...

Figure 8. Exploratory long-term network identification through mapping of the minimum water-level altitude for the period-of-record (1960–April 2017) using a two-dimensional support vector machine (2D-SVM)—Because of much record overlap, numerically the values in green shades are similar to those in figure 4.

FINAL REMARKS

C ONSIDERABLE PROGRESS has been made by the authors towards the construction of a comprehensive informatics system for MRVA water-level database review...

REFERENCES

CLARK, B.R., HART, R.M., and GURDAK, J.J., 2011, Groundwater availability of the Mississippi embayment: U.S. Geological Survey Professional Paper 1785, 62 p. [Also available at https://pubs.usgs.gov/pp/1785.]

KARATZOGLOU, O., Alexandros, SMOLA, Alex, HORNIK, Kurt, and ZEILEIS, Achim, 2004, kern1a—An S4 package for kernel methods in R: Journal of Statistical Software, v. 11, no. 9, p. 1-20, https://www.jstatsoft.org/v11/i09/ and R package version 0.9-25 accessed on April 14, 2017, at https://cran.r-project.org/package=kernlab.

R DEVELOPMENT CORE TEAM, 2017, R—A language and environment for statistical computing: R Foundation for Statistical Computing, Vienna, Austria, version 3.3.3 (2017-03-06), ISBN 3-900051-07-0, accessed on April 14, 2017, at https://www.r-project.org.

STEINWART, Ingo, and CHRISTMANN, Andreas, 2008, Support vector machines: New York, Springer, 601 p., ISBN 978-0-387-77241-7.

U.S. ENVIRONMENTAL PROTECTION AGENCY, 2013, Level III ecoregions of the continental United States: National Health and Environmental Effects Research Laboratory, accessed on April 14, 2017, at https://www.epa.gov/eco-research/level-iii-and-iv-ecoregions-continental-united-states.

U.S. Geological Survey, 2017, National Water Information System: accessed December 13, 2017, at https://nwis.waterdata.usgs.gov/nwis/gwlevels.

WOOD, S.N., 2006, Generalized additive models—An introduction with R: Boca Raton, Chapman & Hall/CRC, 391 p., ISBN 978-1-58488-474-3.

WOOD, S.N., 2017, mgcv—Mixed GAM computation vehicle with GCV/AIC/REML smoothness estimation: R package version 1.8-17, accessed on April 14, 2017, at https://cran.r-project.org/package=mgcv.