Application of Stage IV Precipitation Data to Estimate Spatially Variable Recharge for a Groundwater Flow Model

HEATHER M. MOSER

Iowa State University, Ames, Iowa

ABSTRACT

Groundwater recharge is the volume of infiltrated water that reaches the water table and becomes part of the groundwater flow system. Although a required parameter in groundwater models, recharge is difficult to measure directly, so it must be estimated indirectly. Indirect methods of recharge estimation include empirical methods based on stream discharge and baseflow or assuming that areal recharge is spatially and temporally homogeneous and can be represented by a mean value that is a percentage of total precipitation. However, the distribution of precipitation is inherently heterogeneous, and recharge resulting from it should likewise vary areally. The goals of this study are: 1) to evaluate if NEXRAD rainfall data is useful for estimating spatially variable recharge for groundwater models, and 2) to compare the results to those obtained using a common empirical method of recharge estimation (Rorabaugh method). The model output showed that Rorabaugh recharge and uniform recharge input best approximated modeled hydraulic head to observed values. The Rorabaugh method produced better results than the NEXRAD Stage IV data because recharge from Stage IV precipitation fields do not account for geologic effects on recharge and hydraulic head distributions while they did have an impact on the Rorabaugh input. Spatial variability had only a small impact on the distribution of hydraulic head levels, and did not improve model accuracy.

1. Introduction

Groundwater recharge (R) is the volume of infiltrated water that reaches the water table and becomes part of the groundwater flow system (Anderson and Woessner, 1992). The parameter, R, is an important and necessary part of a groundwater model. Groundwater may be recharged by infiltration from three sources: indirectly from storm runoff that flows into rivers and then percolates through the streambed; localized infiltration from runoff that collects in small depressions and flows downward through joints and fractures; and direct infiltration to the water table wherever precipitation occurs (de Vries and Simmers, 2002; Figure 1). This diagram was created for a semi-arid or arid environment; humid regions rely primarily on vertical infiltration.

In central Iowa, recharge is difficult to estimate. First, the aquifer and aquitard, consisting of Wisconsinan-age glacial deposits of the Des Moines Lobe contains till, sand and gravel, and buried alluvial valleys, which produce a heterogeneous material with a wide range of hydraulic conductivities (K). Recharge is related to K because it is generally thought that sandy soils recharge more water than clayey soils. Second, agricultural tile drains can intercept infiltrated water prior to reaching the water table. Tile drainage is a particularly difficult problem for understanding recharge because the locations of tile drains are poorly documented and the existence

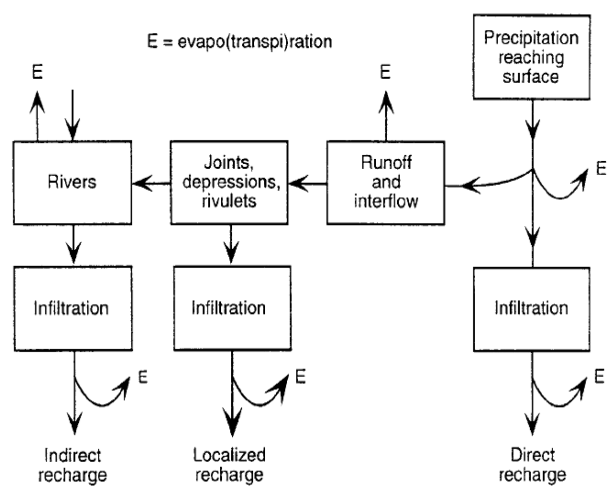

FIG 1. Flow diagram of recharge sources for an arid climate, including losses due to evapotranspiration (de Vries and Simmers, 2002).

of many are completely unknown.

Many methods have been proposed in the literature on how to estimate recharge. De Vries and Simmers (2002) suggest water-balance calculations, tracers, application of Darcy's Law, empirically-derived equations, stream baseflow and discharge calculations, isotope and chemical mass-balance methods, and the use of remotely-sensed variables (e.g., soil moisture) coupled with ground data. Differing scales of investigation and climates may require a variety of methods; thus, a standard method has not been developed (Scanlon et al., 2002). For this reason, recharge is often estimated as a percentage of precipitation that is varied in a groundwater model until a reasonable match between calculated and observed hydraulic heads (water table elevation above sea level for unconfined aquifers) is achieved in groundwater flow models. For the humid Midwest region, recharge has been shown to range from 10 to 20% of total precipitation (Simpkins, 2005). The percentage can range from less than 1% in arid regions where evapotranspiration dominates to as high as 50% where caves and sinkholes provide highly efficient channels for precipitation and runoff to reach groundwater (de Vries and Simmers, 2002).

In more recent years, modelers have examined radar-derived rainfall totals as a means for estimating not only recharge, but also for modeling runoff, streamflow, and flooding potential (Carpenter et al., 2001; Hardigree et al., 2003; Di Luzio and Arnold, 2004; Smith et al., 2004; Gourley and Vieux, 2005; Yilmaz et al., 2005). Rainfall totals from radar have been examined for hydrological use for several decades, but were previously considered to be inadequate due to errors in rainfall detection (Seo, 1998). Rainfall totals derived from radar data alone (from the Z-R power relationship) may be erroneous because of factors affecting reflectivity values. These include bright band enhancement in stratiform regions, beam widening/lengthening with range, underestimation beyond the melting layer, calibration problems, overshooting of low-topped features at long ranges, false precipitation (e.g., convective outflow, ground clutter, etc.), anomalous propagation of the beam, variability of the Z-R relationship, and underestimation resulting from the data only being sampled every five minutes rather than continuously (Zawadski, 1975; Fabry et al., 1992; Kitchen and Jackson, 1993; Hunter, 1996; Smith et al., 1996; Hardegree et al., 2003). However, advancements have been made over the past 20 years to improve the accuracy of radar rainfall estimates. One such advancement was to use rain gages to adjust the radar-derived totals toward more accurate values by scaling them up or down to the gage data nearby. This "multisensor precipitation estimator" (MPE) created by the National Weather Service is known as Stage III and is a 4 km resolution, gridded dataset generated for each radar site in the WSR-88D network by the local River Forecast Centers. By 2002, the centers had begun to mosaic all the MPE data together into a national composite known as Stage IV. Stage IV utilized various averaging protocols to yield the best data available for locations at the edge of the radar range where errors were highest in the original Stage III output, improving overall accuracy of the data. Further quality control checks were also put in place, including adjustments to account for many of the sources of error mentioned previously. Some argue that radar-derived rainfall totals are still not accurate enough for hydrological prediction and modeling (Hardegree et al., 2003), but its use is nevertheless becoming more widespread within the literature.

This study acts as to further test the potential of using NEXRAD in hydrological modeling, specifically groundwater modeling. The goals of this study are: 1) to evaluate if NEXRAD rainfall data is useful for estimating spatially varying recharge for groundwater models, and 2) to compare those results with ones from an empirical model of recharge estimation (the Rorabaugh method).

2. Models

a. GFLOW

GFLOW is a steady-state, single-layer, analytic element groundwater flow model that can be used to simulate flow in confined or unconfined aquifers. The model solves flow equations across the model domain, so that unique solutions for hydraulic head can be obtained at any point. It does this by using stream segments as line sinks or line sources of water and then using the Principle of Superposition (of linear differential equations) to solve for flow in the aquifer. This is in contrast to grid models that use finite difference methods, for example, to solve for hydraulic head at the center of a grid cell. The model incorporates the Dupuit-Forchheimer assumptions where hydraulic head gradients are constant with depth and only vary horizontally, resulting in primarily horizontal flow. Figure 2 illustrates the simplifications of Dupuit-Forchheimer flow.

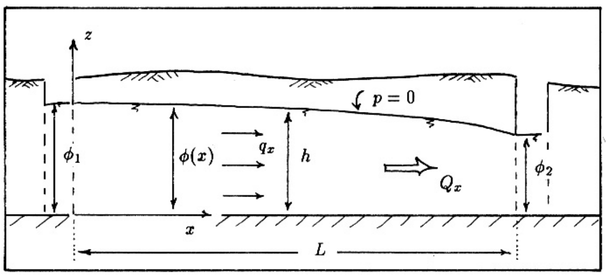

FIG 2. Dupuit-Forchheimer flow for an unconfined aquifer where q is discharge, h is aquifer thickness (depth of water from bottom of aquifer), and Φ is hydraulic head (Haitjema, 1995).

Dupuit-Forchheimer assumptions reduce groundwater flow to a 2-dimensional plane that does not consider any component of vertical flow. Recharge is by definition a vertical flow into the aquifer, so in order to incorporate it into a Dupuit-Forchheimer model, it must be included in the continuity equation as a constant (Equation 1 and Figure 3) and combined with Darcy's Law to develop a form of Poisson's equation (Equation 2):

$$\frac{\partial Q_x}{\partial x} + \frac{\partial Q_y}{\partial y} = R \tag{1}$$

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = -\frac{R}{T} \tag{2}$$

where Q is discharge in units of volume per time, R is recharge in depth per time, T is transmissivity (the product of hydraulic conductivity K and aquifer thickness), and h is hydraulic head (Haitjema, 1995).

Recharge is entered into the GFLOW model generally as a constant value in L/T units (ft/day) and applied over an area. However, it allows for spatial variability in the form of "added recharge

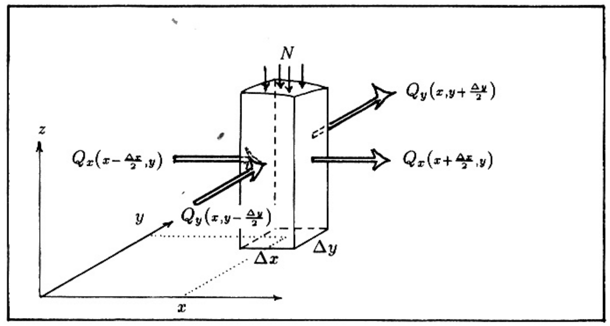

FIG 3. Conceptualized diagram of the continuity equation including recharge (N) (Haitjema, 1995).

rates," which are zones of extra recharge that are then added to the default value (the user specifies where these areas are within the domain by drawing polygons where higher recharge would occur). Haitjema (1995) advised that the uniform default recharge should be used without any inhomogeneity for simplicity, which is equivalent to rainfall occurring at the same rate everywhere in the domain. The ability to introduce inhomogeneities was not originally put in place to reflect spatially varying rainfall, however. The overview of the theory behind the model by Haitjema (1995) showed that inhomogeneities were only placed where aquifer properties changed (e.g., placing more recharge in sandy alluvial valleys where infiltration occurs more readily and in higher amounts).

Using the K and R values, the groundwater model solves a series of simultaneous equations to simulate the hydraulic head observed from field data. Because any ratio of R/T can provides a solution to the equation, R has traditionally been used as a tuning factor to calibrate observed vs. modeled hydraulic head values.

b. RORA

RORA is a USGS FORTRAN program that calculates groundwater recharge in a basin based upon USGS gaging station streamflow data. The program, named for the Rorabaugh method (Rorabaugh, 1964), looks within these data for periods of continuous streamflow recession on hydrographs and calculates an average baseflow level for the stream at that location (Rutledge, 1998). To minimize any error resulting from tile drainage, RORA uses wintertime data, which is considered to be the baseflow value. RORA then uses an empirical equation developed by Rorabaugh (1964) that estimates recharge to be about one-half the amount of discharge (any volume of water in the streamflow exceeding the calculated baseflow) and averages that volume across the area of the drainage basin to arrive at a depth of water that infiltrates to groundwater, as in the equation

$$R = \frac{2(Q_2 - Q_1)K}{2.3026} \tag{3}$$

where R is recharge in L^3 (later divided by watershed area), Q_1 is groundwater discharge before a rainfall event (L^3/T), Q_2 is groundwater discharge after a rainfall event (L^3/T), and K is a recession index constant determined from the hydrograph's master recession curve (MRC).

The Rorabaugh method makes several assumptions in its calculation of recharge. Because all analysis is based on stream discharge at gaging stations, those stations must be located such that they intercept all water that goes into the watershed; thus, the model may perform poorly in low-relief regions. The model may not accurately represent processes in drainage basins of less than 1 km^2 or greater than 2000 km^2. Finally, the model was designed to calculate average recharge based on normal streamflow, so several years of continuous streamflow data are needed to approximate mean conditions (Rutledge, 1998).

3. Methods and Analysis

a. The GFLOW Domain

The watershed selected for this study was the South Fork Iowa River watershed (HUC-08 05451210) in north-central Iowa. GFLOW requires the specification of a near-field and far-field within a fixed domain, so a domain was chosen with South Fork in the center and surrounding watersheds acting as the far-field or boundary conditions (Figure 4).

b. NEXRAD Recharge Input

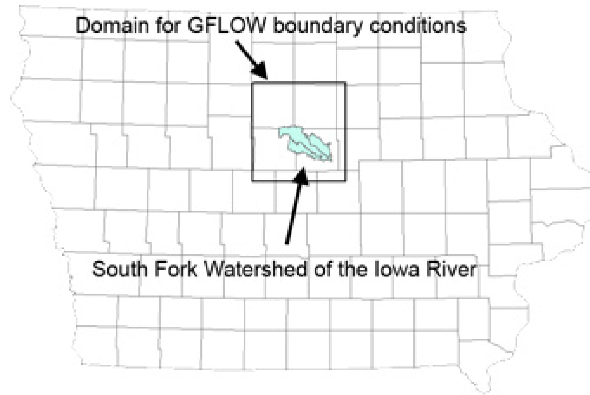

FIG 4. Illustration of the South Fork Iowa River Watershed within the designated GFLOW domain.

Stage IV precipitation data were available as GIS shapefiles of daily rainfall totals from the Iowa Environmental Mesonet. GFLOW runs on an annual time step, so the rainfall data needed to be compiled into yearly totals. The observed hydraulic head measurements that the model was tested against had only been available since August 2005, so it was necessary to select years of rainfall that were climatologically similar to what 2005 has been up to the time of this study. Instead of relying on just one year, three separate years of data (2002, 2003, and 2004) were used to reflect what impact slight variances in rainfall distribution had on the GFLOW output.

The NEXRAD data for the three years were then checked against 15-20 available rain gages from three networks (ASOS, NWS COOP, and USDA ARS) within the GFLOW domain. Gages were thrown out from the sample if they were missing more than an inch worth of rainfall observations during the year. The missing amount was determined by averaging all other available data for the month and totaling those averages for every missing month (the idea being that missing data in January or December will not significantly affect the yearly total while missing data in June or July likely will). In order to compare gages to radar rainfall, the gages had to be combined into spatial averages to put the two datasets on comparable scales. Rain gages and NEXRAD rainfall are sampled both at different time intervals and spatial scales, so they cannot be adequately compared unless the totals are integrated over time and averaged out to similar scales (Zawadzki, 1975). Spatial averaging of the rain gage data was accomplished through the Theissen Polygon method (described by Ward and Trimble, 2004). The mean rainfall of rain gage data was compared to the mean Stage IV rainfall over the domain to test the difference (Table 1).

TABLE 1. Comparison of Stage IV and rain gage annual rainfall totals for 2002, 2003, and 2004.

Year	Type	Rainfall	Stage IV - Gages
2002	Stage IV: Gages:	28.51 in 30.04 in	-1.53 in
2003	Stage IV: Gages:	26.49 in 28.32 in	-1.83 in
2004	Stage IV: Gages:	28.55 in 34.47 in	-5.92 in

As mentioned before, a good starting point for estimating groundwater recharge in Iowa is that it is between 10% and 20% of total precipitation. In order to encompass this entire range in the GFLOW model, two GIS shapefiles were created showing 10% and 20% of the annual Stage IV totals for each of the three years considering that a reasonable output of hydraulic head in GFLOW would lie somewhere between those two extremes. Figure 5 shows the distribution of rainfall amounts for all Stage IV gridpoints contained within the domain for the three years of study (with 20% recharge calculated), and Figure 6 shows a spatial representation within the domain.

Since the user is required to manually input

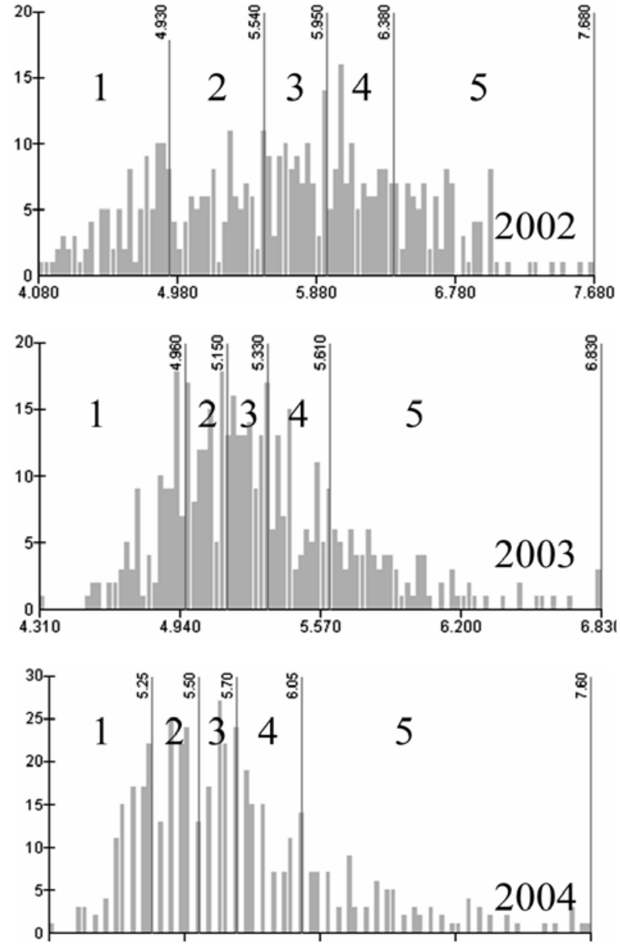

FIG 5. Distributions of Stage IV 20% recharge, separated into five segments: (1) Min to 20th Quantile; (2) 20th to 40th Quantile; (3) 40th to 60th Quantile; (4) 60th to 80th Quantile; and (5) 80th Quantile to Max. Quantile ranges were averaged to determine GFLOW input.

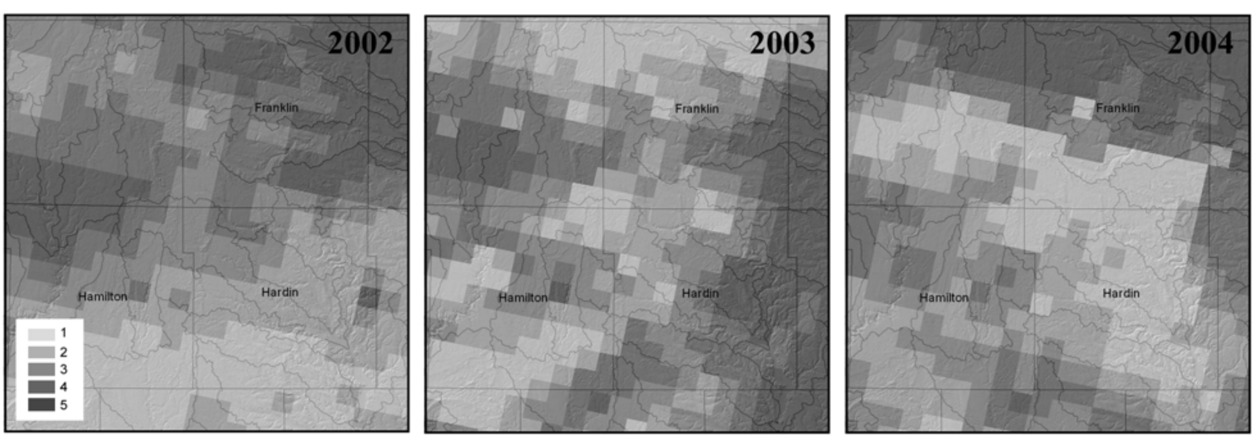

FIG 6. Spatial distributions by quantile separation of recharge within the GFLOW domain based on Stage IV precipitation. Lighter regions correspond to lower amounts of recharge and darker regions correspond to higher amounts.

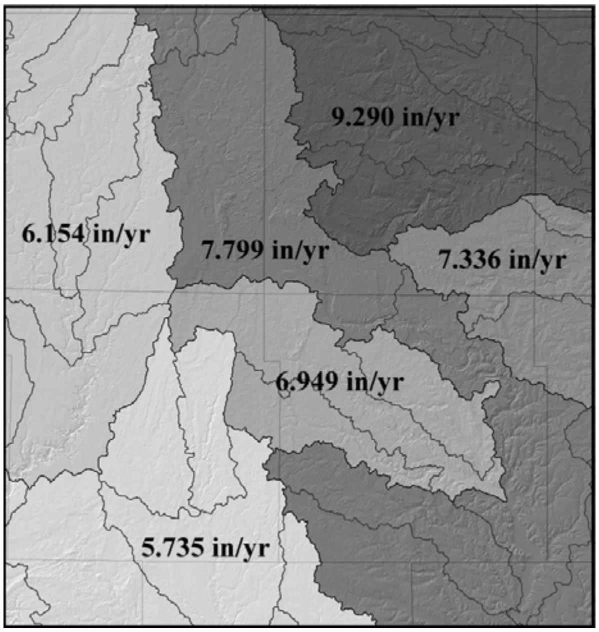

FIG 7. RORA recharge input for the GFLOW model, separated into six distinct areas corresponding to USGS gaging stations.

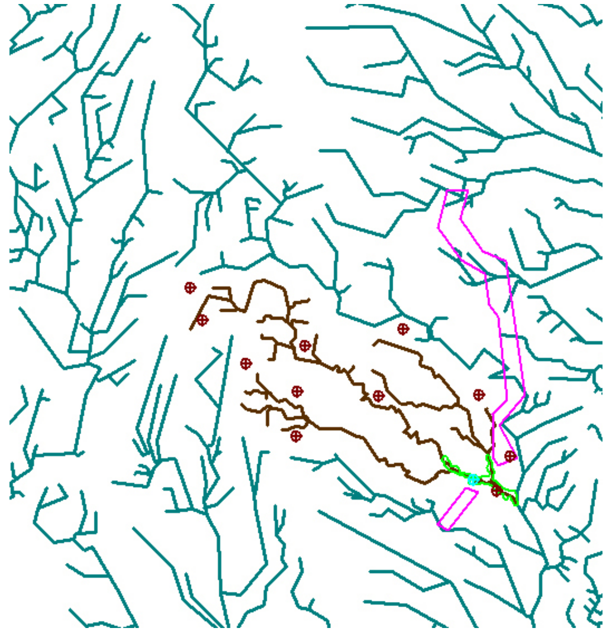

FIG 8. GFLOW analytic elements (stream segments), test points (wells for observed head measurements), and geological inhomogeneities.

the recharge inhomogeneities into the GFLOW model, the rainfall distribution was filtered by statistical quantiles to avoid the tedious task of designating hundreds of gridpoint values separately. Means of each quantile were used as model input. Quantiles were used rather than standard deviation due to two years having skewed, non-normal distributions in the data.

c. RORA Recharge Input

The RORA model was run for six separate USGS streamflow gaging stations that were fed by watersheds within the domain. All six stations had continuous streamflow data from 1996 to 2004. RORA averages a recharge value over entire drainage basins; thus, six unique values were generated within the domain for the six drainage basins covered.

RORA was intended to be run over long periods of time to arrive at average conditions within the watersheds (Rutledge, 1998). However, it outputs the recharge in yearly totals for every year that was included in the record, so the nine years from 1996 to 2004 were averaged together to create a single, average recharge map for the GFLOW domain (Figure 7). Each of the six watershed areas were entered individually as inhomogeneities into GFLOW for spatial variability.

d. Uniform Recharge Input

The mean RORA recharge data was used to generate the uniform recharge input. The single number was calculated by averaging all RORA drainage basins together for all nine years worth of data to best approximate a single, mean condition within the domain. This value was 7.09 inches per year (21.4% of mean annual precipitation for the region).

4. Results

Because the uniform recharge was used as the control, it was run in GFLOW first to calibrate the model's other parameters to levels that would best match the model output head to observed head in the near-field watershed.

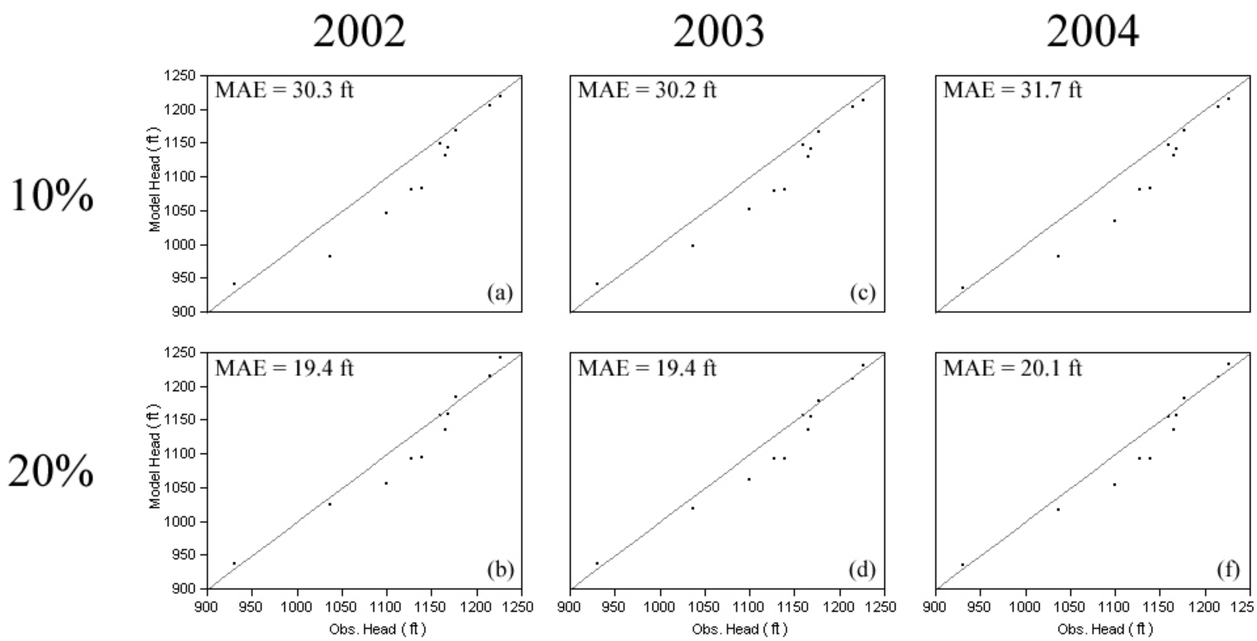

FIG 9. Correlation plots of observed hydraulic head vs. modeled hydraulic head for 10% and 20% Stage IV recharge scenarios. Mean absolute error (MAE) is listed for each dataset.

Inhomogeneities were put in place to account for isolated areas within the domain where the differences between the model output and observed head were unusually large relative to surrounding test points (Figure 8). The additional inhomogeneities were justified based on existing geological features that would have an impact on groundwater levels and movement, such as alluvial material and stratified soil layers of varying hydraulic conductivity. Since GFLOW is a single-layer model, the multiple layers of soil with different K values were averaged into a single K that extended through the entire depth of the aquifer.

Once the seven spatially variable recharge datasets were run, plots were produced for each scenario correlating modeled head to observed head with mean absolute errors calculated (Figures 9 and 10). Mean absolute error is determined by averaging the absolute difference between modeled and observed head for all data points (Anderson and Woessner, 1992).

The RORA and uniform recharge inputs resulted in the lowest or best mean absolute errors, respectively, followed by the three 20% recharge distributions from NEXRAD. The 10% distributions led to greatly underestimated head levels throughout the domain compared to the 20%

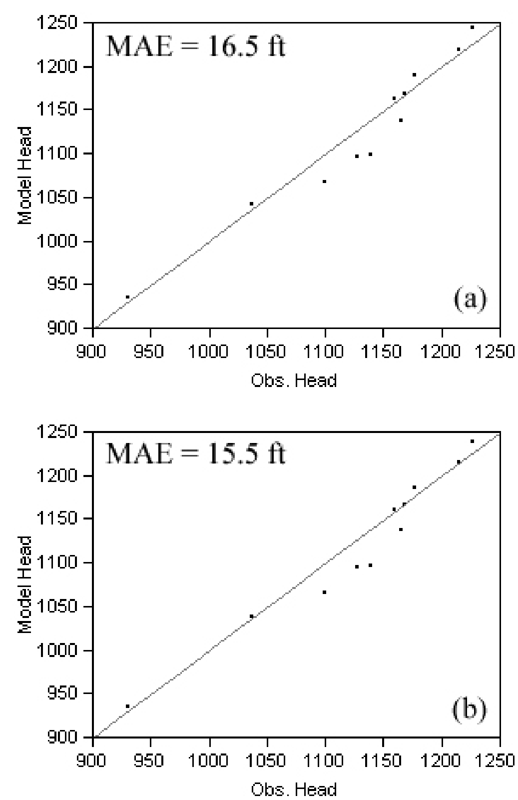

FIG 10. Correlation plots of observed hydraulic head vs. modeled hydraulic head for uniform (a) and RORA (b) scenarios. Mean absolute error (MAE) is listed for each dataset.

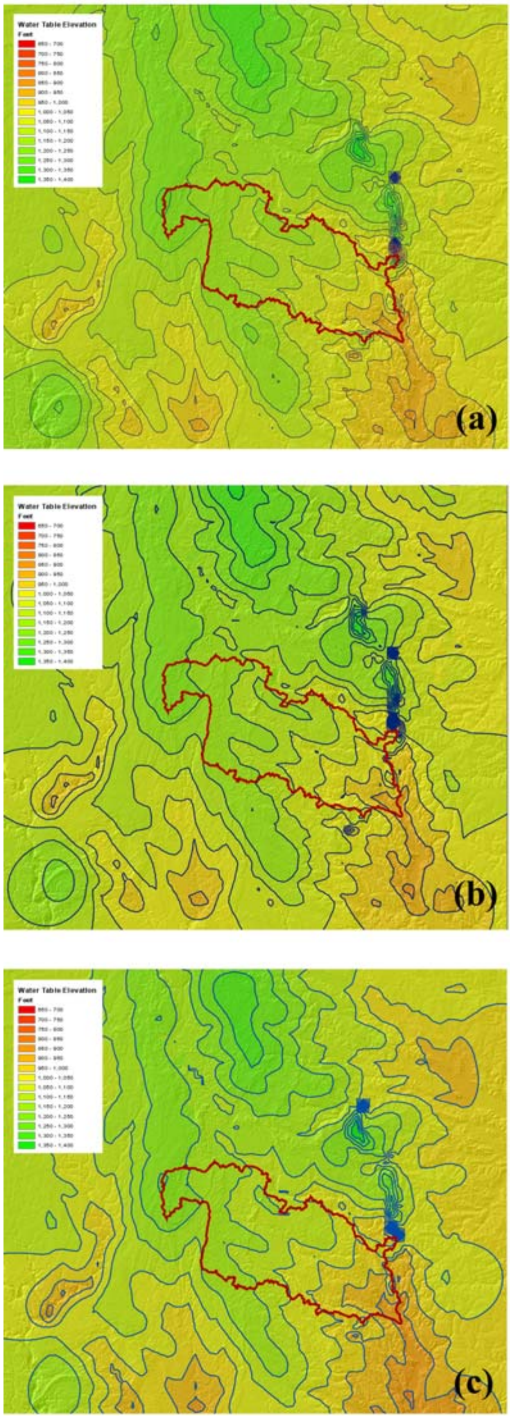

FIG 11. GFLOW modeled hydraulic head contour plot for (a) uniform recharge, (b) RORA recharge, and (c) 2002 Stage IV 20% recharge. The South Fork watershed is outlined in red.

recharge (Figure 9). This is reflected in the correlation plots since nearly all data points fall below the fit line, indicating a tendency for observed head to be greater than modeled head (Figure 9).

5. Conclusions

Spatially variable recharge only improves the modeling of hydraulic head in GFLOW if a) the recharge itself is estimated accurately, and b) geologic factors are considered. The Stage IV recharge performed poorly because it failed both of those conditions. If the rainfall data itself is underestimating precipitation, then the recharge will also be too low. Stage IV rainfall estimates were calculated to be from 5% (2002) and 6% (2003) to 20% (2004) lower than local rain gage measurements, and the mean absolute errors generated for hydraulic head in GFLOW seemed to also reflect that degree of error from year to year (2002 and 2003 performed similarly while 2004 was worst). Also, precipitation data alone estimates recharge without any consideration for geologic heterogeneities and features that have a large impact on where recharge occurs. Therefore, precipitation cannot be used to estimate recharge without considering varying hydraulic conductivities within the aquifer. The approximate residence time of a water volume from infiltration to stream discharge should also be taken into account when using precipitation data for recharge estimation. If the residence time is a year, for example, spatial rainfall distribution from 2002 will have little or no relevance to head measurements taken in 2005.

The RORA and uniform recharge scenarios performed best in GFLOW for a few reasons. First, they likely contributed more accurate recharge amounts because they were calculated with some consideration for geologic factors. Baseflow in a stream comes from groundwater, so the recharge amounts are based on water that had directly interacted with the geologic variability in the aquifer. Thus, although RORA did not directly consider stratified soil units of differing hydraulic conductivity, it was working with data that was directly affected by it. Also, there is less of a short-term time dependence in the RORA recharge because it approximates mean conditions rather than changes in recharge on a year-to-year basis. There was some concern that since RORA calculations are based on streamflow they would be affected by tile drain contributions to the streams. This may be occurring on some level, but the degree of contribution is beyond the scope of this study.

Even if the earlier Stage IV data did not adequately reflect current water table levels, they did have the benefit of demonstrating how much spatial variability affects modeled hydraulic head, however, when compared from year to year. Recharge from 2002 and 2003 were useful for this purpose because the magnitudes were comparable while the distributions were very different. Comparing the two years did not indicate that spatial variability improved the modeled head, however. The mean absolute errors were equal for the correlations, and the plotted data points showed very little relative change. Even comparing all eight spatial recharge scenarios show only minor adjustments in the differences between modeled and observed head for each test point, but variation does exist regarding distribution of head over the domain (Figure 11). Therefore, spatial variability of recharge does have an affect on modeled hydraulic head, but it does not seem to significantly improve the overall accuracy of it. The magnitude of recharge tends to have a greater influence in matching model output to observed values.

Rainfall estimates from weather radar have improved over the last 20 or 30 years, but further improvements in accuracy still need to be made to make it a useful hydrologic tool. A possible next step toward exploring the use of the Stage IV precipitation data in groundwater modeling could be creating composite recharge maps that incorporate both rainfall data and area soil properties into single distributions. This would account for not only the spatial variability of rainfall, but also the geologic variability.

6. Acknowledgements

I would like to thank Dr. William Simpkins for his guidance and suggestions and Daryl Herzmann for his assistance with the NEXRAD data. I would especially like to thank Lucie Macalister for efforts with GFLOW, GIS, and for all of the time she invested to help this project succeed.

7. References

Anderson, M. P., and W. W. Woessner, 1992: *Applied Groundwater Modeling: Simulation of Flow and Advective Transport*. Academic Press, Inc., 381 pp.

Carpenter, T. M., K. P. Georgakakos, and J. A. Sperfslegea, 2001: On the parametric and NEXRAD-radar sensitivities of a distributed hydrologic model suitable for operational use. *J. Hydrol.*, **253**, 169-193.

De Vries, J. J., and I. Simmers, 2002: Groundwater recharge: an overview of processes and challenges. *Hydrogeology J.*, **10**, 5-17.

Di Luzio, M., and J. G. Arnold, 2004: Formulation of a hybrid calibration approach for a physically based distributed model with NEXRAD data input. *J. Hydrol.*, **298**, 136-154.

Fabry, F., G. L. Austin, and D. Tees, 1992: The accuracy of rainfall estimates by radar as a function of range. *Quart. J. Roy. Meteor. Soc.*, **118**, 435-453.

Gourley, J. J., and B. E. Vieux,, 2005: A method for evaluating the accuracy of quantitative precipitation estimates from a hydrologic modeling perspective. *J. Hydrometeor.*, **6**, 115-133.

Haitjema, H. M., 1995: *Analytic Element Modeling of Groundwater Flow*. Academic Press, Inc., 394 pp.

Hardegree, S., and Coauthors, 2003: Multi-watershed evaluation of WSR-88D (NEXRAD) radar-precipitation products. *Proc., First Interagency Conf. on Research in the Watersheds*, Benson, Arizona, U.S. Department of Agriculture, 133-138.

Hunter, S. M., 1996: WSR-88D radar rainfall estimation: capabilities, limitations and potential improvements. *Natl. Wea. Dig.*, **20**, 26-38.

Kitchen, M., and P. M. Jackson, 1993: Weather radar performance at long range – simulated and observed. *J. Appl. Meteor.*, **32**, 975-985.

Prior, J. C., J. L. Boekhoff, M. R. Howes, R. D. Libra, and P. E. VanDorpe, 2003: *Iowa's Groundwater Basics*. Iowa Department of Natural Resources, 83 pp.

Rorabaugh, M. I., 1964: Estimating changes in bank storage and ground-water contribution to streamflow. *International Association of Scientific Hydrology*, **63**, 432-441.

Rutledge, 1998: Computer programs for describing the recession of ground-water discharge and for estimating mean ground-water recharge and discharge from streamflow records—update. USGS Water-Resources Investigations Report 98-4148, 43 pp.

Scanlon, B. R., R. W. Healy, and P. G. Cook, 2002: Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeology J.*, **10**, 18-39.

Simpkins, W. W., 2005: A multiscale investigation of ground water flow at Clear Lake, Iowa. *Ground Water*, (in press).

Smith, J. A., D. -J. Seo, M. L. Baeck, and M. D. Hudlow, 1996: An intercomparison study of NEXRAD precipitation estimates. *Water Resour. Res.*, **32**, 2035-2045.

Smith, M. B., V. I. Koren, Z. Zhang, S. M. Reed, J. –J. Pan, and F. Moreda, 2004: Runoff response to spatial variability in precipitation: an analysis of observed data. *J. Hydrol.*, **298**, 267-286.

Ward, A. D., and S. W. Trimble, 2004: *Environmental Hydrology*. CRC Press, 475 pp.

Yilmaz, K. K., T. S. Hogue, K. -L. Hsu, S. Sorooshian, and T. Wagener, 2005: Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting. *J. Hydrometeor.*, **6**, 497-517.

Zawadzki, I. I., 1975: On radar-raingage comparison. *J. Appl. Meteor.*, **14**, 1430-1436.