

We thank the reviewer, Marios Sophocleous, for his kind and careful assessment of our work.

He raises a number of questions that are interesting not just for our work but for any attempt at quantifying flow of regional, macro-scale aquifers located in areas where data are scarce. This is the case, for instance, of other transboundary aquifers such as the Nubian sandstone aquifer, which has even less data available than the Guaraní.

In these cases, it is not possible to build a model in the traditional sense, where enough data is available for defining model parameters and hydraulic head measurements are sufficient to perform a steady state calibration preceding the transient calibration.

Instead, in our case the model is used to test the consistency of assumptions in an attempt to point the strengths and weaknesses of the currently available conceptual model.

Having said all this, we address each of the questions pointed out by the reviewer.

1) How the K (hydraulic conductivity) values of the different spatial zonations, especially the ones for the increasing number of zones, such as Z3 to Z5, were determined? How many K determinations are available and what is their distribution? Given the scarcity of data, it seems unnecessary to correct the K values for temperature (eq. 1). By the way, how much difference did the K correction make?

For all zonations, K values resulted from the automatic calibration algorithm available in TRANSIN, minimizing an objective function written in terms of heads, parameters and concentrations (if a transport problem is solved). In TRANSIN, the minimizing algorithm uses the Marquardt Method, an iterative algorithm to solve non-linear problems for parameter estimation by the least square method. At all times, we sought calibrated K values to be coherent with the expected K values for sandstones sediments. The K values used for comparison with calibrated ones were obtained from the literature, and were mainly located in exploited areas of the aquifer (these locations will become clear once we add a new figure with the distribution of pumping wells (see response to comment 5)

In spite of K data scarcity, we consider the correction adequate because of the wide range of temperatures found in the aquifer, from 25 up to 65°C. For example, considering a K value for sandstones of 1 m/d at 20°C (Freeze and Cherry, 1979), applying equation (1) considering an aquifer temperature of 35°C (typical of extensive areas of the aquifer), would result in a correction factor of 1,375, increasing then the reference K value.

2) How certain are the boundaries where the GAS is not outcropping, such as in the Argentinean side, north in Brazil, etc? Was some sensitivity analysis run on the boundaries to check the impact of boundary conditions on the results?

The aquifer limits were defined by consensus among geologists and hydrogeologists participating in the PSAG. The delineation of the Northern boundary is well supported by the geology of the area. That is probably one of the best known areas of the aquifer as it has been the subject of numerous studies due to its location within the Sao Paulo State, one of the most densely populated in Brazil. See for instance Hirata et al. (2011).

On the contrary, the boundary in the Argentinean side is still a matter of debate and has recently been revisited by Rosello and Veroslasky (2011). Opposite to what happens along the

other borders of the aquifer, these authors state that the western boundary in the Gran Chaco of Argentina, has no strong geo-hydrogeologic expressions that can be recognized in both, surface and subsurface. Therefore, the definition of such a boundary offers different interpretations depending on the factors and interests taken into account. Consequently, some technical experts propose alternative locations that include or exclude extensive areas of the Gran Chaco Argentino which, in turn, pose administrative and economic controversies for the aquifer exploitation and management. Nonetheless, in our work we have used the boundary defined at the time of the PSAG execution. Being the model a dynamic tool, it is expected to incorporate new features and to be used to test the consistency of new hypothesis as new data and re-interpretation of existing one postulate alternative conceptual models.

- Hirata, R., Geisicki, A.; Sracek, O.; Bertolo, R.; Giannini, P.C. y Aravena, R., 2011. Relation between sedimentary framework and hydrogeology in the Guarani Aquifer System in São Paulo state, Brazil. *Journal of South American Earth Sciences*, 31(4): 444-456.
- Rossello, E. A. y Veroslavsky, G. 2012. Definición del límite occidental del Sistema Acuífero Guaraní (Gran Chaco, Argentina): ¿técnico o convencional? *Boletín Geológico y Minero*, 123 (3): 297-310 ISSN: 0366-0176

3) How was the steady-state (initial) condition was determined and how was it checked to ensure that it was indeed at equilibrium?

As stated in the introduction, we faced the problem of insufficient head data to perform a conventional modeling approach, i.e. model calibration for steady state with field data representing pre-development conditions as close as possible, and then use the steady state simulated aquifer head as the initial condition for the transient simulation.

Instead, we resorted to a different calibration strategy. We simulated a steady state condition for the first time step of a time series, after which the transient simulation was run. Instead of attempting to temporally distribute the few observed head measurements following some criteria (year of reading was not available for many of them), we assumed that all observed heads could be assigned to the last simulated year. See also comment 7).

4) How were the various leakance coefficients determined? The statement that “they were previously estimated based upon riverbed characteristics” seems vague and needs specification.

In order to estimate leakance coefficients α for the application of the following boundary condition or stream/aquifer interaction representation, we considered a K values for streambed sediments two orders of magnitude less than K values for sandstones (1 m/d according to Freeze

$$Q = \frac{AK}{L}(h - H_{ext}) = \alpha(h - H_{ext})$$

and Cherry (1979), a thickness of streambed sediments L equal to 1 m, and an area A ranging from 100 to 1000 m² for all streams except for the Uruguay and Paraná river, for which higher values were adopted. Similar to recharge, on the early stages of the model development (a first version of a steady state model, Vives et al. 2008) we performed sensitivity runs to this parameters and found no significant differences for the range of α tested.

Vives, L. S., Rodríguez, L. B., Gómez, A. A., and Cota, S. D. S.: Modelación numérica regional del Sistema Acuífero Guaraní (Regional numerical modeling of the Guarani Aquifer System),

Project for Environmental Protection and Sustainable Development of the Guaraní Aquifer System, Global Environment Facility (GEF), Technical Report, Montevideo, Uruguay, 144 pp., 2008.

5) Showing the spatial distribution of pumping would be instructive to the readers.

We uploaded a draft figure with this information that will be incorporated in the final version of the manuscript after the discussion period ends.

6) Showing the distribution of the “317 observed piezometric levels” (also indicating at least the decade they were measured) would be useful to the readers.

Please, see Figure 7 and Page 9901- Lines 16-17. The text reads “The geographic distribution of errors, with their corresponding sign and magnitude, not only highlights the location and density of calibration data ...” .Anyhow, we will consider re-working some of the figures to clarify this point.

7) The authors state (in p. 9900, L. 8-9) that “In the absence of transient head data for calibration, all available observations were assigned to the last period.” (The authors also stated that they used “yearly periods.”) However, earlier they stated that “those [piezometric] levels span a 30-40 yr time window starting in the 70’s: :” which means that they have some idea of at least the decade in which the measurements were performed. Why then all those measurements were assigned in the “last period,” i. e., the last year of the simulation?

Please, see response to comment 3) above. It is worth remembering that head measurements were taken at the time of well drilling and that about 70-80% of the wells were constructed during the last decade, when there was a significant increment in groundwater exploitation (needless to say that this increment is not evenly distributed across the entire aquifer area, but is concentrated in densely populated regions). The rest of the measurements are spread in time since the seventies. Those are the reasons that support our strategy of assigning all data to the last stress period of our 39 year simulation time (from 1970 until 2008)

Nonetheless, the reviewer is correct; we could have attempted an alternative approach to handle aquifer head data.

8) Recharge was determined as a fixed percentage of precipitation without much justification. Are there any independent studies to verify or support the adopted percentage of precipitation?

In actuality, during the preliminary stages of the model development (Vives et al., 2008), we performed a sensitivity analysis of model results to recharge rates, using values between 1 and 10 % of mean annual precipitation. Those extreme values were not arbitrarily chosen. In a previous work we had compiled recharge rates from several authors and aquifer locations that helped us to bracket recharge rates (see Gómez et al., 2010). That sensitivity analysis was also performed in order to reduce the number of calibration parameters during the inverse simulation runs and, with that, help convergence and diminish simulation times.

Recharge studies at regional level are lacking. Currently, detail studies are underway at micro-scale basins located in Sao Paulo State and Porto Alegre, Brazil.

Gómez, A. A., Rodríguez, L., and Vives, L.: The Guaraní Aquifer System: estimation of recharge along the Uruguay–Brazil border, *Hydrogeol. J.*, 18, 1667–1684, 2010.

9) With regard to model structure identification, it would be helpful to readers to summarize the meaning of each criterion used in Table 3.

We will prepare a summary with their meaning to accompany Table 3.

10) It would also be helpful if an explanation of the different colors in the left figure of Fig. 4 were provided. By the way, I missed noting an explanation/justification of the southern boundary conditions, which should be pointed out.

The colours have the only purpose of delineating different zones which were distinguished as being exclusively-recharge zones and recharge/discharge zones. We agree that an appropriate legend would have been informative. We will rework figure 4 accordingly.

With regard to the southern boundary condition (and all boundary conditions), they are represented in Figure 4. The southern boundary is represented as a stream-like boundary condition pictured in black. Anyway, we will clarify the text to better identify the different boundary sectors over with the water budget analysis is performed.

11) English language mistakes/misprints that need to be corrected

We apologize for so many mistakes in our English style, we will incorporate all the modifications as well as all the technical responses to the issues raised by the reviewer in the revised version of the paper, once the discussion period has finished.