A methodology for the preliminary characterisation of the river boundary condition in finite difference groundwater flow numerical models

Una metodologia per la caratterizzazione preliminare della condizione al contorno “river” in modelli numerici di flusso delle acque sotterranee alle differenze finite

Manuel M. Oliveira, Tiago N. Martins

Introduction

The definition of the river boundary condition in finite difference numerical models such as MODFLOW (McDonald and Harbaugh 1988) or MODFLOW-2005 (Harbaugh 2005) requires the specification of three parameters for each cell of the model, namely the stage of the river (HRIV), the elevation of the riverbed bottom (RBOT), and the hydraulic conductance of the stream-aquifer interconnection (CRIV). This last parameter is a combination of other four parameters: the width of the river (W), the length of the stream in the cell (L), the hydraulic conductivity of the riverbed material (K) and the thickness of the riverbed (M) - Figure 1:

\[
CRIV = \frac{K \times L \times W}{M} \quad (1)
\]

Flow rate between aquifer and river (QRIV) is computed in MODFLOW using HRIV, CRIV and RBOT, by (McDonald and Harbaugh 1988):

\[
QRIV = CRIV \times (HRIV - h) \quad \text{if } h > RBOT \quad (2)
\]

\[
QRIV = CRIV \times (HRIV - RBOT) \quad \text{if } h \leq RBOT \quad (3)
\]

where \( h \) represents the simulated hydraulic head in the cell. If the hydraulic head of a cell is higher than the stage of the river \( h > HRIV \), equation (2) applies (situation represented...
by h₂ in Figure 1), the river will drain the aquifer and hence QRIV will be negative. On the contrary, if the stage of the river is higher than the hydraulic head (HRIV > h, situation represented by h₁ in Figure 1), the river will recharge the aquifer, and two situations may occur: if h is higher than the elevation of the riverbed bottom (RBOT), equation (2) applies and flow is proportional to (HRIV – h); or, if h is lower than RBOT, equation (3) applies and flow will be constant no matter the value of h.

The purpose of this paper is to define a methodology for an initial characterization of all the terms present in the river boundary condition, using available information and applying GIS tools, at a regional scale and in an expeditious way. This methodology had its first presentation in Oliveira and Martins 2017 and has been applied and further updated in two projects developed in Brazil (Leitão et al. 2017; Costa et al. 2019). The most recent updates are presented in this paper.

**Materials and Methods**

Some comments are presented together with the methodologies concerning the evaluation of each of the terms of the river boundary condition (represented in red in Figure 1). This section is mainly descriptive while the Application of the proposed methodology section illustrates its application.

**Stage of the river (HRIV)**

When the information required for the characterization of the stage of the river is not available or may be difficult to obtain, it is proposed to use an approach in GIS that joins the water course lines, the digital elevation model (DEM), and the model grid.

One of the aspects in numerical modelling is that each model cell must be assigned a river stage value. In each model cell the stage of the river may vary and this variation will increase as the cell size increases. Also, sometimes, inside a cell there may be a confluence of one or more streams. These situations raise a first challenge to define which river stage value should be assigned to a cell.

The proposed methodology calculates the river stage value for each cell of the model. It goes through the following steps to be performed in a GIS environment:

a. Preparing the polygon shapefile of the model grid. In this shapefile each cell polygon is a square or a rectangle and must be identified by a unique ID integer value (e.g. field PKUID);

b. Preparing the DEM raster file;

c. Preparing the streams’ line shapefile that the modeler wants to include in the numerical model;

d. Intersecting the streams’ line shapefile with the model grid polygon shapefile: the result is a line shapefile which is sectioned by the limits of the model grid shapefile. The PKUID of the model grid shapefile must be preserved in the line output;

e. Dividing each line section in segments using a predetermined interval (variable LEN_SEG – a value of 25 m is suggested if the DEM cell side is larger than 25 m, otherwise a value equal to the DEM cell size is suggested). Despite this is not necessary for the computation of the river stage, the length of each segment must be computed (field LENGTH - this field will be used for the calculation of the length of the stream in the cell - term L);

f. Creating a point shapefile with the central point of each line segment. The PKUID value must be preserved in this shapefile;

**Fig. 1 - Rappresentazione 3D di una cella del modello numerico (contorno verde) e rappresentazione dei vari termini necessari per la condizione al contorno “river” (termini in rosso: W indica l’ampiezza del fiume, L la lunghezza del segmento di fiume nella cella, K la conducibilità idraulica dei sedimenti del letto fluviale, M lo spessore del letto fluviale e RBOT l’elevazione del fondo del letto fluviale), h₁ e h₂ rappresentano il carico idraulico calcolato dal modello nel caso in cui questo si attesti ad una quota maggiore o minore del livello idrometrico del fiume, QRIV1 e QRIV2 la portata di scambio tra il fiume e l’acquifero.**
g. Overlapping this point shapefile with the DEM and get the elevation at each central point;

h. Calculating, for each model cell represented by its PKUID, minimum and mean elevations (fields Z_MIN and Z_MEAN respectively) of the segment central points, using a statistical tool. The same tool is used to compute the total length of the segments in each cell (field SUM_LENGTH), that will be used later for the calculation of the length of the stream in the cell.

As a first approach one could use Z_MEAN, the average surface elevation, to represent the river stage (HRIV). However, note that each DEM cell value used to assign the elevation of the central point of the stream segment is the average elevation of the area that the DEM cell represents. Considering that the river flows through the lowest elevation line inside the DEM cell, it must have an average elevation which is lower than the value of the DEM cell.

For that reason, the Z_MIN and the Z_MEAN values are used to assign HRIV. A representative value for the river stage could be situated between Z_MEAN and Z_MIN, and it is proposed the middle value of these two:

$$HRIV = \frac{(Z_{MIN} + Z_{MEAN})}{2} \quad (4)$$

Even a value lower than Z_MIN might be possible, as this value still represents, for the lowest elevation cell of the DEM, an average elevation value. But, taking into consideration all the DEM cells inside the model cell, HRIV calculated with equation (4) may be a reasonable approach.

In order that the computation of Z_MIN, Z_MEAN and HRIV have more statistical meaning, a minimum of 9 DEM cells should exist per model cell. It is more likely that the average river stage of a model cell approaches the value calculated with equation (4) as the number of DEM cells per model cell increases. For lower number of DEM cells per model cell (less than 9) the average river stage is expected to be closer to Z_MIN or even lower than Z_MIN.

A situation may occur in which the elevation of the top of the cell may be lower than HRIV. In this situation, a correction must be made, and either the cell top is increased or HRIV must be lowered. Also, if HRIV is lower that the elevation of the bottom of the cell, again a correction must be carried out, either increasing HRIV or decreasing the cell bottom.

**Riverbed bottom elevation (RBOT)**

In many cases there is no information on the riverbed bottom elevation (RBOT). Also, as mentioned before, the stream bed may not even exist. Hence an expedient approach is proposed. This should be based on knowledge of the river depth and the streambed thickness at some locations, and then by defining a relation between these and the river stage in the study area. If this knowledge is not available, an arbitrary relation like the suggested one depicted in Table 1 could be used.

RBOT would be calculated subtracting to the HRIV value computed in the last step, the value of the depth of the riverbed bottom given in Table 1 or any other value defined by the modeler.

Control must be carried out if by the application of this procedure RBOT is below the elevation of the model cell bottom. If this happens either the cell bottom must be lowered or the RBOT must be increased to be situated between HRIV and the cell bottom.

**Length of the river (L)**

The length of the river (L) in each cell is calculated taking advantage of the steps used to characterize the river stage (HRIV). As mentioned in the explanation of the presented methodology, for each cell, L is given directly by the field SUM_LENGTH calculated in step h of the Stage of the river (HRIV) section. The advantage of using this methodology is that a real stream length is used. This value is calculated whether there is only a small portion of the stream intersected by a model cell, a model cell comprises one or more streams, or whatever the directions of the streams: the computed value always represents the total length of the streams existing in the model cell.

**Width of the river (W)**

Riverbed width may be characterized using cartography, satellite images, or aerial photos. An expeditious way of using this information observed in several places of the model region would be to correlate it with the stage of the river, in a similar way to the one used for the characterization of the elevation of the riverbed bottom. A table could be produced with this relation and this would be applied to the model area. Table 2 provides an example of such a possible relation.

**Hydraulic conductivity of the riverbed material (K)**

Concerning the vertical hydraulic conductivity of the riverbed, if the riverbed exists, and if there is no measured information, a published value from the bibliography could be used if there is any knowledge about the riverbed sediments. In many cases

Tab. 1 - Esempio e relazione suggerita tra la profondità del fondo del letto fluviale e il livello idrometrico del fiume.

<table>
<thead>
<tr>
<th>River stage (m)</th>
<th>0 - 2</th>
<th>2 - 10</th>
<th>10 - 20</th>
<th>&gt; 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth of the riverbed bottom below the river stage (m)</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Tab. 2 - Esempio e relazione suggerita tra l’ampiezza del fiume e il livello idrometrico del fiume.

<table>
<thead>
<tr>
<th>River stage (m)</th>
<th>0 - 1</th>
<th>1 - 2</th>
<th>2 - 10</th>
<th>10 - 50</th>
<th>&gt; 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth of the riverbed bottom below the river stage (m)</td>
<td>100</td>
<td>50</td>
<td>10</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>
maybe it is not possible to define a riverbed with geological features distinct from the geological aquifer material; in this case it is suggested to use the value of the vertical hydraulic conductivity of the model cell where the stream is located.

**Thickness of the riverbed (M)**

This is the most awkward parameter to estimate, as it is very difficult to define the depth of the riverbed, and, if it doesn’t exist, what is its meaning. So, a unitary value may be arbitrarily chosen for M.

As the values of M, K, L and W are considered in the model as lumped parameters (CRIV given by equation (1)), assigning a unitary value for M will not be a major problem. The final CRIV value will later be obtained during the calibration phase.

**Application of the proposed methodology**

The application of the methodology presented in the previous section is illustrated with data from an anonymous real situation.

The original DEM was prepared using the Shuttle Radar Topography Mission (SRTM) 1 arc-second data (EROS 2017). Each area of the 1 arc-second SRTM data corresponds to a square of around 30 m side in the areas near the equator. As one moves away north or south, the area becomes a rectangle, keeping the same around 30 m cell side along the longitude (vertical side) but shortening the cell side as latitude (horizontal side) increases. For instance, at 40° latitude, the horizontal cell side reduces to less than 24 m.

In order to keep a square side grid, the original SRTM is projected from its original geographical coordinates to an orthogonal coordinate system. At the same time the resolution of the grid is changed. In the following example it was changed to a 50 m cell size.

All procedures will be demonstrated using QGIS version 2.18.21, a free and open source GIS software (https://qgis.org) available from http://download.osgeo.org/qgis/windows/. It is not intended to make a thorough step by step guide of the procedures but just to point out the main used tools and the results obtained.

**Preparing the model grid polygon shapefile**

The model grid polygon shapefile must be produced in order to make all the subsequent procedures. This will be produced taking advantage of the FREEWAT platform (http://www.freewat.eu/) which is a QGIS plugin. FREEWAT is an open source and public domain GIS integrated modelling environment for the simulation of water quantity and quality in surface water and groundwater with an integrated water management and planning module (Rossetto et al. 2018). One of its main advantages is the ability to conciliate the QGIS environment with the MODFLOW code. Using the FREEWAT plugin, a rotated grid of the model area was generated. This model grid is composed of 200 m side square cells (Figure 2).

**Preparing the DEM and stream network**

Using the 1 arc-second SRTM, the projected 50 m cell size DEM is obtained in QGIS using the capability available in QGIS to export a raster file. This DEM is sink filled for depressions with the aim of producing the river stream network that will be used in the numerical model. This is a choice of the modeler that guarantees total coherence between the DEM and the stream network. Otherwise there might be problems with the stream network following areas of higher elevations instead of the most depressed areas. In order to fill sinks and draw the stream network, three SAGA tools available in QGIS toolbox are used in the following sequence: “Fill sinks (wang & liu)”, “Flow accumulation (qm or esp)” and “Channel network”. Figure 2 partially illustrates the obtained sink filled DEM and the stream network, already superimposed by the rotated model grid.

**Sectioning the stream network by using the model grid cells**

This task is accomplished intersecting the streams’ line shapefile with the model grid polygon shapefile. This action, that uses the QGIS “Intersection” command, breaks the stream network in the model cells where it occurs while keeping the PKUID identifier of the model cell. The result can be observed in Figure 3.

**Dividing each cell stream section in segments and creating the corresponding central points**

Each cell stream section is divided into segments. This task is carried out in QGIS using the “Densify geometries given an interval” tool followed by the “Explode lines” tool, both available in the Processing Toolbox. The first tool generates new vertices along the stream section distant at maximum 25 m from each other (25 m was the value assigned to the
variable LEN_SEG as presented in the Stage of the river (HRIV) section). These vertices are used in the second tool to effectively break the line section into line segments. The length of each segment is calculated in the LENGTH field. The used tools may have produced several segments with zero length that must be deleted from the attribute table. The obtained result is shown with larger detail in Figure 4a for cell 8437 where there is a confluence of streams. In this cell 17 segments were created.

For each segment the central point is created using the QGIS analysis tool “Mean coordinate(s)”. Figure 4b shows the results obtained after this process.

Getting the elevation of each segment central point from the DEM

Having defined the location of each segment center, next step is to assign their elevation values. This is accomplished with the “Add raster values to points” SAGA command in QGIS, using, as suggestion, the “Nearest Neighbor” resampling option. Figure 5 shows the overlap of the stream segments with the DEM, and Table 3 the values obtained for each segment after assignment of the elevation values, only for cell 8437 as an example. Table 3 also shows the values of the segment lengths.
Tab. 3 - Values of elevation \( (Z, \text{ in meters}) \) and length \( (m) \) of each segment centroid of the river network (with unique identifier ID) located in the model cell with unique identifier PKUID = 8437.

Tab. 3 - Valori di quota \( (Z, \text{ in metri}) \) e lunghezza \( (m) \) di ciascun centroide del segmento del reticolo fluviale (con identificativo ID univoco) all’interno della cella del modello con identificativo univoco PKUID = 8437.

<table>
<thead>
<tr>
<th>PKUID</th>
<th>ID</th>
<th>LENGTH</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>8437</td>
<td>390</td>
<td>25.0000</td>
<td>523.0000</td>
</tr>
<tr>
<td>8437</td>
<td>764</td>
<td>25.0000</td>
<td>523.0000</td>
</tr>
<tr>
<td>8437</td>
<td>765</td>
<td>8.1898</td>
<td>523.0000</td>
</tr>
<tr>
<td>8437</td>
<td>766</td>
<td>25.0000</td>
<td>523.0000</td>
</tr>
<tr>
<td>8437</td>
<td>389</td>
<td>16.8101</td>
<td>523.0000</td>
</tr>
<tr>
<td>8437</td>
<td>391</td>
<td>25.0000</td>
<td>521.0000</td>
</tr>
<tr>
<td>8437</td>
<td>392</td>
<td>20.7107</td>
<td>521.0000</td>
</tr>
<tr>
<td>8437</td>
<td>393</td>
<td>25.0000</td>
<td>521.0000</td>
</tr>
<tr>
<td>8437</td>
<td>394</td>
<td>25.0000</td>
<td>519.0000</td>
</tr>
<tr>
<td>8437</td>
<td>396</td>
<td>25.0000</td>
<td>519.0000</td>
</tr>
<tr>
<td>8437</td>
<td>767</td>
<td>25.0000</td>
<td>519.0000</td>
</tr>
<tr>
<td>8437</td>
<td>769</td>
<td>25.0000</td>
<td>519.0000</td>
</tr>
<tr>
<td>8437</td>
<td>397</td>
<td>25.0000</td>
<td>518.0349</td>
</tr>
<tr>
<td>8437</td>
<td>770</td>
<td>25.0000</td>
<td>518.0349</td>
</tr>
<tr>
<td>8437</td>
<td>771</td>
<td>20.7107</td>
<td>518.0349</td>
</tr>
<tr>
<td>8437</td>
<td>89</td>
<td>25.0000</td>
<td>518.0349</td>
</tr>
<tr>
<td>8437</td>
<td>90</td>
<td>5.7158</td>
<td>518.0262</td>
</tr>
</tbody>
</table>

Calculating the required statistics for each numerical model cell

Last step in this phase is to compute the statistics \( Z_{\text{MIN}} \), \( Z_{\text{MEAN}} \) and \( \text{SUM}_\text{LENGTH} \) for each model cell. This is accomplished in QGIS using the “Dissolve with stats” plugin tool, using as dissolve field the unique PKUID cell model identifier and the statistics “Sum” for the Field LENGTH, and “Min” and “Mean” for the field Z. Table 4 shows the results obtained using this methodology, including the statistics obtained for the model cell with PKUID = 8437.

<table>
<thead>
<tr>
<th>PKUID</th>
<th>LENGTH</th>
<th>MIN_Z</th>
<th>MEAN_Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>8321</td>
<td>119.8592</td>
<td>516.0000</td>
<td>517.3362</td>
</tr>
<tr>
<td>8322</td>
<td>100.5289</td>
<td>518.0087</td>
<td>518.0192</td>
</tr>
<tr>
<td>8323</td>
<td>56.6815</td>
<td>523.0087</td>
<td>523.0145</td>
</tr>
<tr>
<td>8435</td>
<td>231.6169</td>
<td>536.0000</td>
<td>538.0080</td>
</tr>
<tr>
<td>8436</td>
<td>88.0528</td>
<td>531.0000</td>
<td>533.4000</td>
</tr>
<tr>
<td>8437</td>
<td>372.1371</td>
<td>518.0262</td>
<td>520.2450</td>
</tr>
<tr>
<td>8438</td>
<td>263.9904</td>
<td>523.0087</td>
<td>524.5440</td>
</tr>
<tr>
<td>8439</td>
<td>65.0934</td>
<td>526.0123</td>
<td>527.3416</td>
</tr>
<tr>
<td>8551</td>
<td>175.0350</td>
<td>527.0000</td>
<td>530.8000</td>
</tr>
<tr>
<td>8552</td>
<td>253.0660</td>
<td>523.0000</td>
<td>527.4615</td>
</tr>
</tbody>
</table>

Discussion and Conclusions

Based on the stream network and the DEM it is possible to define some of the parameters required for the cells where the river boundary condition is to be defined, namely the stage of the river (HRIV) and the length of the river (L). Concerning the remaining parameters some suggestions are presented in section Materials and Methods, including, depending on the HRIV values, the estimation of the riverbed bottom elevation (RBOT) and the width of the river (W).

One of the main advantages of these procedures is that they allow, using base information usually available, to make a first characterization of the boundary condition, in a consistent way, providing a starting point for the model calibration process.

Despite base information on river network may be available, in case the methodology is being used in areas without detailed information, it may be advisable, for the sake of consistency, to produce the river network using the available DEM. Otherwise HRIV may assume values higher than the acceptable. The decision on the size of the river network to be used in the model depends on the way the model reacts to its extension.

The methodology can be applied to DEM of different sources and sizes. A DEM widely available is the 1 arc-second SRTM but this methodology can be applied to any DEM as far as its cell size is lower than the model cell size. So that the computed values of river stage may be more representative, at least a minimum of 9 DEM cells per model cell would be advisable.

The automation of these processes might also be of interest because, in case there is a change in the model geometry, it would be very easy to compute the new boundary condition values and assign them to the corresponding model cells.

The values hereby generated may be used to produce the river boundary condition input file of the MODFLOW program (file with extension RIV). They can also be incorporated in specific files of the several modelling softwares that preprocess the information used to run MODFLOW. For instance, if the FREEWAT platform is being used, it is very easy to assign the calculated values to the river model data object that is used to later produce the MODFLOW RIV file.

Finally, the same methodology applies if the rivers in a modeling area are not permanent and they are modeled with the DRAIN boundary condition, meaning that the aquifer may discharge to the river, but the river never feeds the aquifer. In this situation all the methodology applies, except the definition of the riverbed bottom elevation which is not required, with drain elevation being given by HRIV.

Acknowledgment: This paper was developed in the framework of LNEC’s 2013-2020 Research and Innovation Plan, line Risk Management and Safety in Hydraulics and Environment (Process Nr. 0605/112/20383). Tiago N. Martins thanks the Fundação para a Ciência e Tecnologia (FCT), Portugal for the Ph.D. Grant PD/BD/135590/2018.
REFERENCES


