Use of remote sensing and long-term in-situ time-series data in an integrated hydrological model of the Central Kalahari Basin, Southern Africa

Moiteela Lekula1,2 & Maciek W. Lubczynski 1

Received: 5 June 2018 / Accepted: 17 December 2018
© The Author(s) 2019

Abstract
Distributed numerical models, considered as optimal tools for groundwater resources management, have always been constrained by availability of spatio-temporal input data. This problem is particularly distinct in arid and semi-arid developing countries, characterized by large spatio-temporal variability of water fluxes but scarce ground-based monitoring networks. That problem can be mitigated by remote sensing (RS) methods, which nowadays are applicable for modelling not only surface-water but also groundwater resources, through rapidly increasing applications of integrated hydrological models (IHMs). This study shows implementation of various RS products in the IHM of the Central Kalahari Basin (~200 Mm²) multi-layered aquifer system, characterized by semi-arid climate and thick unsaturated zone, both enhancing evapotranspiration. The MODFLOW-NWT model with UZF1 package, accounting for variably saturated flow, was set up and calibrated in transient conditions throughout 13.5 years using borehole hydraulic heads as state variables and RS-based daily rainfall and potential evapotranspiration as driving forces. Other RS input data included: digital-elevation-model, land-use/land-cover and soils datasets. The model characterized spatio-temporal water flux dynamics, providing 13-year (2002–2014) daily and annual water balances, thereby evaluating groundwater-resource dynamics and replenishment. The balances showed the dominant role of evapotranspiration in restricting gross recharge to only a few mm yr⁻¹ and typically negative net recharge (median, −1.5 mm yr⁻¹), varying from −3.6 (2013) to +3.0 (2006) mm yr⁻¹ (rainfall of 287 and 664 mm yr⁻¹ respectively) and implying systematic water-table decline. The rainfall, surface morphology, unsaturated zone thickness and vegetation type/density were primary determinants of the spatio-temporal net recharge distribution.

Keywords  Groundwater/surface-water interaction · Remote sensing · Numerical modelling · Water balance · Namibia · Botswana

Introduction
Groundwater is often the only, but vulnerable, source of potable water in arid and semi-arid areas, hence it must be well evaluated and managed. Nowadays, distributed integrated hydrological models (IHMs), coupling surface with groundwater processes, are considered optimal tools for groundwater resources management, but their reliability is largely constrained by the availability and quality of input data (Meijerink et al. 2007). The largest data problem is found in arid and semi-arid areas, because in those areas, ground-based monitoring networks are scarce (Brunner et al. 2007; Leblanc et al. 2007). That problem can be mitigated by remote sensing (RS) methods.

In recent years, RS has played an increasingly important role in providing spatio-temporal information for water resources evaluation and management (Coelho et al. 2017). Its applications in surface hydrology, including surface-water modelling, are already well known and typically include digital elevation derivatives, land use/cover, and spatio-temporal rainfall and evapotranspiration evaluations (Schmugge et al. 2002). However, the RS contribution to groundwater hydrology and groundwater resources evaluation is less distinct, so less well known.
Standard published RS applications in groundwater hydrology have involved: assessment of groundwater recharge (e.g., Awan et al. 2013; Brunner et al. 2004; Coelho et al. 2017; Jasrotia et al. 2007; Khalaf and Donoghue 2012), surface-water/groundwater interaction (e.g., Bauer et al. 2006; Hassan et al. 2014; Leblanc et al. 2007; Sarma and Xu 2017), and groundwater storage (resources) evaluation and change (e.g., Henry et al. 2011; Rodell et al. 2007; Rodell and Famiglietti 2002; Taniguchi et al. 2011; Yeh et al. 2006). With recent advancement of IHMs, the RS contribution to such models is rapidly increasing, mainly because of continuously increasing amounts of downloadable RS-based hydrological products, for example, rainfall or potential evapotranspiration data.

The most complex models are those coupling atmospheric-land-energy processes (so-called land surface models, LSMs, or atmospheric models, AMs) with physically based models, integrating surface with subsurface processes (further referred to as the integrated hydrological model, IHM). Examples of such complex couplings include: (1) coupling of hydrological model Parallel Flow (ParFlow) IHM (Ashby and Falgout 1996; Kollet and Maxwell 2006) with a LSM called Common Land Model (Dai et al. 2003) referred as CLM.PF (Maxwell and Miller 2005; Rihani et al. 2010); (2) coupling of CATchment HYdrology (CATHY; Paniconi et al. 2003) IHM with a LSM called NoahMP (Niu et al. 2011) referred as CATHY/NoahMP (Niu et al. 2014) and (3) coupling of HydroGeoSphere IHM (Brunner and Simmons 2012; Therrien 1992; Therrien et al. 2006) with the Weather Research and Forecasting (WRF) AM (described in Davison et al. 2018), referred to as HGS/WRF (Davison et al. 2018).

Among the IHMs, there are complex models based on a three-dimensional (3D) solution of Richards’ equation and models simplifying Richards’ equation to simulate surface-water/groundwater interactions. Examples of the complex IHMs include the three models already mentioned (ParFlow, CATHY and HydroGeoSphere), as well as HYDRUS-3D (Šimůnek et al. 2012), MODHMS (Panday and Huyakorn 2004), and WASH123D (Yeh et al. 2003). All these IHMs are computationally demanding and require fine spatial and temporal discretization due to large nonlinearity of Richards’ equation (Downer and Ogden 2004; Sheikh et al. 2009). The IHMs simplifying Richards’ equation are more robust. For example, used worldwide in lots of surface-water/groundwater interaction studies, MIKE SHE (Danish Hydraulic Institute 1998), simplifies Richards’ equation to one-dimension (1D), although it is still a very complex code, requiring a variety of skills such as hydrogeology, soil science, agronomy, computational hydraulics (Refsøgaard 2010) and many kinds of data for spatial heterogeneity description (Ma et al. 2016). Similar to MIKE SHE is the Gridded Surface/Subsurface Hydrological Analysis (GSSHA) model (Downer and Ogden 2004), which also simplifies Richards’ equation to 1D.

Relatively simpler and computationally more efficient are models simplifying Richards’ equation, not only to 1D but also simplifying vertical, variably saturated flow as driven only by the gravity potential gradient, i.e. ignoring negative potential gradients (Harter and Hopmans 2004; Niswonger and Prudic 2004; Smith and Hebert 1983). Such a solution, applying the kinematic wave (KW) approximation of Richards’ equation, solved by the method of characteristics, is, for example, proposed within the widely used (also in this study) Unsaturated-Zone Flow (UZF1) Package (Niswonger et al. 2006) under MODFLOW-NWT (Niswonger et al. 2011). In the regional-scale modelling, such as in this study, that simplification can even be advantageous, because the errors introduced by averaging or upscaling soil hydraulic parameters, makes the KW approximation of the Richards’ equation and its solution comparable in accuracy, while the KW equation requires less input data and much less computational power (Bailey et al. 2013; Hassan et al. 2014; Morway et al. 2013). Besides, the MODFLOW related codes are public domain.

MODFLOW-NWT, with its surface-water/groundwater interaction packages, including UZF1 package, has already been applied worldwide, either directly or within the GSFLOW (Markstrom et al. 2008) IHM. Most of these applications focused on simulation of hydrological processes of surface-water/groundwater interactions (El Zehairy et al. 2018; Gong et al. 2012; Hassan et al. 2014; Huntington and Niswonger 2012) and climate change impact on groundwater resources (e.g., Gong et al. 2012; Hay et al. 2010; Huntington and Niswonger 2012; Surfleet and Tullos 2013; Surfleet et al. 2012). However, none of such applications has ever been dedicated to IHM simulation of a regional, multi-layered aquifer system with a very thick unsaturated zone such as the Central Kalahari Basin (CKB), integrating RS data and long-term in-situ hydro-meteorological time-series data.

The main objectives of this study were: (1) to present the use of various RS products coupled with long-term in-situ monitoring data, as input of a regional-scale distributed numerical IHM of the CKB; (2) to characterize spatio-temporal water flux dynamics of a semi-arid, multi-layered aquifer system characterized with very thick unsaturated zone; and (3) to provide a long-term quantitative water-balance estimate of such a system, evaluating its groundwater resources.

The Central Kalahari Basin (CKB) was chosen as the study area because it not only complies with the aforementioned characteristics, but also because it hosts the most productive, important and exploited transboundary groundwater resources of the Karoo System Aquifer (SMEC and EHES 2006), the focus of interest of Botswana and potentially also of Namibia.
Data and methods

Study area and conceptual model

The Central Kalahari Basin (CKB) study area (Fig. 1), occupies central Botswana (~181,000 km²) and a small part (~14,000 km²) of Eastern Namibia. It is a large-scale hydrogeological basin, which formerly was a catchment of the fossil Okwa-Mmone River system (de Vries 1984). It is nearly flat due to surficial accumulation of eolian sand, known as Kalahari Sand. About 90% of the CKB is occupied by Kalahari Desert, characterized by semi-arid to arid climate because of its position under the descending limb of the Hadley cell circulation (Batisani and Yamal 2010).

The rainfall in the CKB is highly spatially and temporally variable (Lekula et al. 2018b; Obakeng et al. 2007), with localized showers (Bhalotra 1987; Lekula et al. 2018b). Almost all the rainfall occurs from September to April, i.e. mainly during summer rainy season. The mean annual rainfall in the CKB ranges from 380 mm yr⁻¹ in the south-western, to 530 mm yr⁻¹ in the north-eastern side of the study area (Lekula et al. 2018b). Annual potential evapotranspiration in the CKB exceeds annual rainfall, ranging from 1,350 to 1,450 mm (Choudhury 1997). The majority of the study area is covered by savannah grassland, sparse shrubs and acacia trees, which increase their density towards the east. The CKB is sparsely inhabited by people, mainly at the fringes, with the interior part occupied by Central Kalahari Game Reserve.

The whole CKB study area is covered by a mantle of Kalahari Sand of variable thickness, ranging from a few meters in the west to >60 m in the central and eastern part. In approximately two-thirds of the CKB area, the sand is underlain by rocks of Karoo Supergroup Formation, while the remaining third part is underlain by Pre-Karoo rocks (Lekula et al. 2018a). The principal aquifers are Lebung, Ecca, and Ghanzi Aquifers (SMEC and EHES 2006). It is remarkable that despite the deep occurrence of groundwater (typically >60 m b.g.s.), in the majority of the CKB, the main regional groundwater flow follows topography, i.e. it is directed from the higher elevated areas along the water divides of the CKB in the west, south and east, towards the lowest depression area in the central part and towards the north-east of Makgadikgadi Pans (de Vries et al. 2000; Fig. 1). There are no permanent surface-water bodies in the CKB study area; thus de Vries et al.
characterized it as a closed surface-water basin with an internal groundwater drainage system outflowing towards a natural discharge area of Makgadikgadi Pans (Fig. 1).

The hydrogeological framework of the CKB conceptual model presented by Lekula et al. (2018a) consists of six hydrostratigraphic units (HU; Fig. 2), including: (1) Kalahari Sand Unit (KSU); (2) Stormberg Basalt Aquitard (SBA); (3) Lebung Aquifer (LA); (4) Inter-Karoo Aquitard (IKA); (5) Ecca Aquifer (EA); and (6) Ghanzi Aquifer (GA).

The top Kalahari Sand Unit (KSU) is composed of unconsolidated to semi-consolidated deposits with thickness ranging from 6 m in the western part to more than 100 m in the central and northern parts of the CKB (de Vries et al. 2000; Lekula et al. 2018a). The major part of the KSU thickness comprises the unsaturated zone, while the remaining, thin part at the KSU bottom, is saturated and spatially discontinuous. In a large part of the CKB, the KSU is underlain by the second HU, i.e. Stormberg Basalt Aquitard (SBA). In the majority of that area, the bottom of the KSU is saturated. The locally fractured SBA is composed of spatially nonuniform tholeiitic flood basalts (Smith 1984), characterized by abruptly changing thickness from 0 even up to ~200 m, as a result of intrusion or block faulting around major fault zones. Where the KSU is underlain by LA, EA or GA aquifer, it is hydraulically connected with that aquifer, forming one unconfined unit, typically with the water table in the aquifer underlying the KSU. The third HU, i.e. Lebung Aquifer (LA) consists of well-sorted, reddish to white, massive, but fractured sandstone, also nonuniformly distributed. Its spatially variable thickness ranges from 0 m in the north-western part of the CKB where it wedges out and in the southern part of the Zoetfontein Fault where significant uplifting has resulted in LA erosion, to ~230 m in the northeastern and south-western parts of the CKB (Lekula et al. 2018a; SMEC and EHES 2006). The fourth HU, i.e. Inter-Karoo Aquitard (IKA), is composed of sedimentary, interchanging and nonuniformly distributed, mudstones and siltstones. Its thickness ranges from 0 m in the north-western and southern part of the CKB, to ~250 m in the central part. The fifth HU, i.e. Ecca Aquifer (EA), consists of nonuniformly distributed sandstone, inter-layered with siltstone and carbonaceous mudstone (Smith 1984). Its thickness ranges from 0 m in the north-western CKB where it wedges out towards the Ghanzi Aquifer, to ~290 m in the southern part of the Zoetfontein Fault (Fig. 1), where significant uplifting is followed by erosion of the SBA, LA and even IKA, so that EA is directly overlain by KSU (Lekula et al. 2018a; Smith 1984). The sixth HU, i.e. Ghanzi Aquifer (GA), consists of weakly metamorphosed, purple-red arkosic sandstone, siltstone, mudstone and rhythmite (Smith 1984), present only in the north-western part of the CKB. Its thickness ranges from 0 m in the centre of CKB, to ~230 m towards the north-western CKB. The basement, not included in the model, is represented by impermeable rocks underlying the deepest aquifer at given locations of the CKB flow system (Fig. 2).

<table>
<thead>
<tr>
<th>Hydrostratigraphic Unit (HU)</th>
<th>Sub-Basin</th>
<th>Karoo Division</th>
<th>Group</th>
<th>Period</th>
</tr>
</thead>
</table>
| Kalahari Sand Unit (KSU; HU 1) | Kalahari Group | Post Karoo | Kalahari | CENOZOIC
| Stormberg Basalt Aquitard (SBA; Unit 2) | Lebung Aquifer (LA; HU 3) | Upper Karoo | Stormberg Basalt | MESOZOIC
| Inter-Karoo Aquitard (IKA; HU 4) | Nkon De Sandstone Formation | Beaufort | Lebung | Triassic
| Ecca Aquifer (EA; HU 5) | Tshwane Fm. | Ecca | Mnemonic Fm. | LOWER PERMIAN
| Ghanzi Aquifer (GA; HU 6) | Dwyka Fm. | Middlepits Fm. | Mnemonic Fm. | PROTEROZOIC
| Basement | Waterberg, Transvaal, Gabarone Granite, Kanye Formation, Otka Complex | Pre Karoo | | ARCHAEOAN

Fig. 2 Stratigraphy and hydrostratigraphy of Kalahari Group, Karoo Super-Group and Pre-Karoo in the CKB, after Lekula et al. (2018a); the colours correspond to hydrostratigraphic units and the dash-line defines a regional unconformity. Permission for reuse granted by Physics and Chemistry of the Earth, License No. 4470031150263
Numerical model

The MODFLOW-NWT model with active UZF1 Package, further referred as MOD-UZF, was chosen as the IHM to be used in this study because: (1) it is a relatively simple but still integrated modelling solution, allowing one to compute groundwater fluxes (gross recharge, groundwater evapotranspiration and groundwater exfiltration) internally, based on unsaturated-zone parameterization and external input driving forces such as rainfall reduced by interception and potential evapotranspiration (Fig. 3), rather than assigning them arbitrarily, as is the case in a standard standalone groundwater model (Hassan et al. 2014); (2) the study area is pretty flat with a poor drainage network, active only shortly after long heavy rains, which justifies the use of MOD-UZF rather than more sophisticated IHM solution with complex surface modelling domain; (3) it is a computationally efficient IHM solution, optimal for large areas such as the CKB (~200,000 km²); (4) it is public domain software with extensive web materials.

For pre- and post-processing of the MOD-UZF, the ModelMuse graphical user interface (Winston 2009) was used because: (1) it is public domain software; (2) it is easy and straightforward software with good technical support. Post-processing of cell by cell water budgets for each HU or of a specific part of the model, was evaluated with ZONEBUDGET (Harbaugh 1990).

Model setup

A six-layer 3D regional-numerical model was built over the CKB area, following the hydrogeological conceptual model (Fig. 3) of Lekula et al. (2018a), in which six model layers directly corresponded to the six HUs as per Figs. 2 and 4. The top model boundary represented by topographic surface, was assigned using 90-m spatial resolution digital elevation model data obtained from the Shuttle Radar Topography Mission (SRTM; Jarvis et al. 2008). Each subsequent layer boundary was defined by subtraction of the HU thicknesses, interpolated using borehole log data within the 3D geological model (Lekula et al. 2018a) developed in Rockworks 17 software (RockWare 2017), further referred to as Rockworks. Where HUs pinched out (Fig. 4), layers were extended throughout the model domain, applying a fictitious 1-m-thick layer, with hydraulic properties representative of the overlying layer, in order to have continuous hydraulic connections in all the six layers, as per the solution proposed by Anderson et al. (2015) and for example implemented by (Masterson et al. 2016).

In the first, KSU top layer, 1D-vertical, variably saturated flow between land surface and water table, was simulated by the UZF1 package. All the six model layers, including the upper KSU with partially unsaturated zone, were set as “convertible” to be able to simulate spatio-temporally varying groundwater flow in either confined or unconfined conditions, depending on the head position. A quadratic 5 × 5-km² grid, consistent with the WGS84 ARC coordinate system in which the CKB falls entirely in one zone 10 was used, instead of the commonly used WGS 1984 UTM coordinate system, where the CKB falls in the two zones, 34 and 35. Such grid size was found to be a trade-off between model computational time and model accuracy.

Model input

The input data for the CKB consisted of driving forces, parameters and state variables. Rainfall reduced by interception, further referred to as effective precipitation ($P_e$), is the main driving force of the MOD-UZF model. In this study, the spatio-temporally variable, daily satellite rainfall of Famine Early Warning System Network Rainfall Estimate v.2 (Herman et al. 1997), further referred to as RFE, was downloaded from the United States Geological Survey (USGS) Famine Early Warning
System Network (FEWSNET) data portal (USGS 2007) for the period from 1 June 2001 to 31 December 2014. The choice of RFE was because of its superior daily rainfall detection capability in the CKB (Lekula et al. 2018b). The RFE of 0.1° (~11 km) spatial resolution was resampled in ArcGIS to 5-km spatial resolution to match the model grid size, then converted to ASCII format and finally imported into ModelMuse.

The interception losses were assigned as spatially variable based on 1 × 1-km spatial resolution Land Use Land Cover (LULC) map (Loveland et al. 2000). Five land cover types were defined using information from Le Maitre et al. (1999), Miralles et al. (2010) and Werger and van Bruggen (1978), each with attributed interception losses defined in percentages of rainfall: bare soil and water bodies (0%); grasslands (2%); shrubs (4%); savannah (a mixture of grassland, shrubs and forest) (6%); forest (12%). That map was reclassified according to the corresponding interception loss classes and resampled to 5 × 5-km spatial resolution, all done in ArcGIS. Finally, for each day of simulation, the interception loss map was subtracted from the rainfall map to obtain spatio-temporally variable infiltration rate, i.e. the effective precipitation ($P_e$), applied as the model driving force for each simulation day.

The second important driving force of the MOD-UZF is potential evapotranspiration (PET). The spatio-temporally variable daily PET data at 1° (~110 km) spatial resolution was downloaded from the same USGS FEWSNET data portal as the rainfall and for the same period as the rainfall. The FEWSNET PET is calculated by USGS, using the Penman-Monteith equation formulation of Shuttleworth (1993) for reference crop evaporation, with external input parameters such as air temperature, atmospheric pressure, wind speed, relative humidity, and solar radiation, obtained from the Global Data Assimilation System (GDAS), generated every 6 h by the National Oceanic and Atmospheric Administration (NOAA), standardized in accordance with FAO Publ. 56 (Allen et al. 1998) and finally aggregated into daily totals. The FEWSNET PET was chosen because it was the only RS PET product available at daily time step for the CKB. The FEWSNET PET, further referred to as PET, was resampled in ArcGIS to 5-km spatial resolution to match the model grid. The resampled daily PET, was finally converted to ASCII format and imported into ModelMuse.

The third driving force of the model was well abstractions, sourced from the two Debswana Diamond Mining Company Wellfields (Orapa and Jwaneng) and Water Utilities Corporation...
Wellfields (Greater Ghanzi and Gaathobogwe areas). The abstraction rates were arranged according to the daily simulation time step.

The parameterization of the unsaturated and saturated zones is presented in Table 1. The relation between unsaturated hydraulic conductivity and the unsaturated-zone water content was defined by the Brooks and Corey function (Brooks and Corey 1966; Eq. 1):

\[ K(\theta) = K_s \left( \frac{\theta - \theta_r}{\theta_s - \theta_r} \right)^\varepsilon \]

where \( K(\theta) \) is unsaturated hydraulic conductivity, \( K_s \) is vertical saturated hydraulic conductivity, \( \theta \) is current volumetric water content, \( \theta_r \) is soil residual water content; \( \theta_s \) is soil saturated water content; and \( \varepsilon \) is the Brooks and Corey exponent.

In the UZF1 Package, continuity between the unsaturated zone and saturated zone in the top unconfined aquifer is maintained through \( S_y \) estimated as \( \theta_s - \theta_r \), where \( \theta_r \) approximates specific retention capacity (Niswonger et al. 2006). The \( \theta_r \) and \( \theta_s \) (Table 1) were defined as spatially variable input, with the help of the 1 x 1-km resolution Africa soil map data (Jones et al. 2013); for each soil class, \( \theta_r \) and \( \theta_s \) were assigned following studies by Carsel and Parrish (1988) and Joshua (1991) and then spatially aggregated into 5 x 5-km grid. The evapotranspiration extinction water content (EXTWC) was assigned as \( \theta_r + 0.01 \) and initial water content (\( \theta_i \)) as \( \theta_r \). The UZF1 vertical hydraulic conductivity (\( K_v \)) was assigned as ten times lower than \( K_h \) of the first layer (Domenico and Schwartz 1998; SMEC and EHEES 2006) and was adjusted in the calibration process. The Brooks and Corey exponent (\( \varepsilon \)) was kept as default (Table 1).

The spatial distribution of EXTDP was defined as per the different vegetation types, following the 1 x 1-km spatial resolution of LULC classification (Loveland et al. 2000). The values of the EXTDP classes were assigned based on the areal contributions of plants with different rooting depths, deduced from Obakeng et al. (2007), Kleidon (2004) and Canadell et al. (1996) as: bare soil and water bodies (0 m); grasslands

### Table 1

<table>
<thead>
<tr>
<th>ZONE</th>
<th>Parameter</th>
<th>Min value</th>
<th>Max value</th>
<th>Unit</th>
<th>Model</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsaturated</td>
<td>( \theta_s )</td>
<td>0.37</td>
<td>0.43</td>
<td>m³ m⁻³</td>
<td>MOD-UZF (UZF1)</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>( \theta_r )</td>
<td>0.05</td>
<td>0.15</td>
<td>m³ m⁻³</td>
<td>MOD-UZF (UZF1)</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>( \theta_i )</td>
<td>0.05</td>
<td>0.15</td>
<td>m³ m⁻³</td>
<td>MOD-UZF (UZF1)</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>EXTWC</td>
<td>0.06</td>
<td>0.16</td>
<td>m³ m⁻³</td>
<td>MOD-UZF (UZF1)</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>EXTDP</td>
<td>1</td>
<td>25</td>
<td>m</td>
<td>MOD-UZF (UZF1)</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon )</td>
<td>3.5</td>
<td>3.5</td>
<td>–</td>
<td>MOD-UZF (UZF1)</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>( K_v )</td>
<td>1.1</td>
<td>1.96</td>
<td>m d⁻¹</td>
<td>MOD-UZF (UZF1)</td>
<td>C</td>
</tr>
<tr>
<td>Saturated</td>
<td>( K_h ) (layer 1)</td>
<td>11</td>
<td>19.6</td>
<td>m d⁻¹</td>
<td>MOD-UZF</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>( K_h ) (layer 2)</td>
<td>3.1E-05</td>
<td>0.009</td>
<td>m d⁻¹</td>
<td>MOD-UZF</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>( K_h ) (layer 3)</td>
<td>0.42</td>
<td>0.95</td>
<td>m d⁻¹</td>
<td>MOD-UZF</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>( K_h ) (layer 4)</td>
<td>1.02E-07</td>
<td>0.004</td>
<td>m d⁻¹</td>
<td>MOD-UZF</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>( K_h ) (layer 5)</td>
<td>0.132</td>
<td>0.67</td>
<td>m d⁻¹</td>
<td>MOD-UZF</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>( K_h ) (layer 6)</td>
<td>0.585</td>
<td>0.96</td>
<td>m d⁻¹</td>
<td>MOD-UZF</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>( S_y ) (layer 1)</td>
<td>0.22</td>
<td>0.33</td>
<td>–</td>
<td>MOD-UZF</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>( S_y ) (layer 2)</td>
<td>1.0E-05</td>
<td>1.0E-05</td>
<td>–</td>
<td>MOD-UZF</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>( S_y ) (layer 3)</td>
<td>0.02</td>
<td>0.08</td>
<td>–</td>
<td>MOD-UZF</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>( S_y ) (layer 4)</td>
<td>1.0E-06</td>
<td>1.0E-06</td>
<td>–</td>
<td>MOD-UZF</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>( S_y ) (layer 5)</td>
<td>0.03</td>
<td>0.04</td>
<td>–</td>
<td>MOD-UZF</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>( S_y ) (layer 6)</td>
<td>0.01</td>
<td>0.06</td>
<td>–</td>
<td>MOD-UZF</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>( S_s ) (layer 2)</td>
<td>1.0E-09</td>
<td>1.0E-09</td>
<td>m⁻¹</td>
<td>MOD-UZF</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>( S_s ) (layer 3)</td>
<td>2.0E-06</td>
<td>8.0E-06</td>
<td>m⁻¹</td>
<td>MOD-UZF</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>( S_s ) (layer 4)</td>
<td>1.0E-09</td>
<td>1.0E-09</td>
<td>m⁻¹</td>
<td>MOD-UZF</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>( S_s ) (layer 5)</td>
<td>2.0E-06</td>
<td>4.8E-06</td>
<td>m⁻¹</td>
<td>MOD-UZF</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>( S_s ) (layer 6)</td>
<td>1.2E-06</td>
<td>6.5E-06</td>
<td>m⁻¹</td>
<td>MOD-UZF</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>Cond</td>
<td>0.5</td>
<td>67</td>
<td>m² d⁻¹</td>
<td>MOD-UZF (GHB)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>Cond</td>
<td>2</td>
<td>307</td>
<td>m² d⁻¹</td>
<td>MOD-UZF (DRN)</td>
<td>C</td>
</tr>
</tbody>
</table>

Wellfields (Greater Ghanzi and Gaathobogwe areas). The abstraction rates were arranged according to the daily simulation time step.

The parameterization of the unsaturated and saturated zones is presented in Table 1. The relation between unsaturated hydraulic conductivity and the unsaturated-zone water content was defined by the Brooks and Corey function (Brooks and Corey 1966; Eq. 1):

\[ K(\theta) = K_s \left( \frac{\theta - \theta_r}{\theta_s - \theta_r} \right)^\varepsilon \]

where \( K(\theta) \) is unsaturated hydraulic conductivity, \( K_s \) is vertical saturated hydraulic conductivity, \( \theta \) is current volumetric water content, \( \theta_r \) is soil residual water content; \( \theta_s \) is soil saturated water content; and \( \varepsilon \) is the Brooks and Corey exponent.

In the UZF1 Package, continuity between the unsaturated zone and saturated zone in the top unconfined aquifer is maintained through \( S_y \) estimated as \( \theta_s - \theta_r \), where \( \theta_r \) approximates specific retention capacity (Niswonger et al. 2006). The \( \theta_r \) and \( \theta_s \) (Table 1) were defined as spatially variable input, with the help of the 1 x 1-km resolution Africa soil map data (Jones et al. 2013); for each soil class, \( \theta_r \) and \( \theta_s \) were assigned following studies by Carsel and Parrish (1988) and Joshua (1991) and then spatially aggregated into 5 x 5-km grid. The evapotranspiration extinction water content (EXTWC) was assigned as \( \theta_r + 0.01 \) and initial water content (\( \theta_i \)) as \( \theta_r \). The UZF1 vertical hydraulic conductivity (\( K_v \)) was assigned as ten times lower than \( K_h \) of the first layer (Domenico and Schwartz 1998; SMEC and EHEES 2006) and was adjusted in the calibration process. The Brooks and Corey exponent (\( \varepsilon \)) was kept as default (Table 1).

The spatial distribution of EXTDP was defined as per the different vegetation types, following the 1 x 1-km spatial resolution of LULC classification (Loveland et al. 2000). The values of the EXTDP classes were assigned based on the areal contributions of plants with different rooting depths, deduced from Obakeng et al. (2007), Kleidon (2004) and Canadell et al. (1996) as: bare soil and water bodies (0 m); grasslands
yield (SBA and IKA aquitards were assigned with spatially uniform conductivities (Kv) were derived from the aquifer transmissivity data, extracted from pumping tests of projects executed in the CBK and aquifer thicknesses deduced using Rockworks (Lekula et al. 2018a). The vertical hydraulic conductivities (Kh) of all the layers were assigned as ten times lower than Kh (Domenico and Schwartz 1998). The K-values were adjusted in the calibration process as per Table 1. These were the basis for demarcating internally homogeneous and isotropic K-zones for the aquifers, i.e. for layer 1–26 zones, for layer 3–26 zones, for layer 5–27 zones and for layer 6–11 zones. The SBA and IKA aquitards were assigned with spatially uniform K-values. The zones of aquifer storage parameters, i.e. specific yield (Sw) and specific storage (Sw), were delineated the same way as K-zones, while their values were assigned following borehole lithology and various literature sources elaborated in Lekula et al. (2018a). The Sw and Sw were further calibrated in the transient simulation, mainly following expansion of the cones of depressions around the wellfields.

Based on the conceptual model of Lekula et al. (2018a), the following external boundary conditions (Fig. 5) were assigned for the CBK numerical model: (1) no-flow boundaries all around the CBK, matching the Okwa-Mmone River Catchment boundary for the first KSU, while for the subsequent units, either at the contact with an impermeable unit or along groundwater flowlines (Lekula et al. 2018a); (2) the head-dependent inflow/outflow boundary assigned using MODFLOW General Head Boundary (GHB) Package (McDonald and Harbaugh 1988), to simulate lateral groundwater inflow or outflow (Fig. 5); and (3) the head dependent outflow boundary assigned using MODFLOW Drain (DRN) Package (McDonald and Harbaugh 1988), to simulate lateral groundwater outflow to the Makgadikgadi Pans in the northern part of the study area (Fig. 5). There is also an internal model boundary along the regional structural feature of the fifth hydrostratigraphic unit (EA) called Zoetfontein Fault (Figs. 1 and 5), which is simulated using the Horizontal Flow Barrier (HFB) Package (Hsieh and Freckleton 1993), applying thickness of 3 m and hydraulic conductivity, adjusted during the calibration process.

Groundwater levels converted to hydraulic heads were sourced from Department of Water Affairs (Botswana), Water Utilities Corporation (Botswana), Debswana Diamond Mining Company (Botswana) and Directorate of Water Resources Management (Namibia). The hydraulic heads were used as state variables in the model calibration. In the steady-state calibration, all the water levels in investigation boreholes and saturated zone, i.e. groundwater exfiltration (EXFgw), groundwater evapotranspiration (ETg), gross recharge (Rg) and net recharge (Ra), to changes in selected parameters.

Model calibration and sensitivity analysis

First, a steady-state model was developed and calibrated using 13.5-year means of daily driving forces and state variables (from 1 June 2001 to 31 December 2014) to initialize the transient model. However, the steady-state simulation did not provide satisfactory initial hydraulic heads and initial water contents (θ) to start transient simulation, resulting in unrealistically large gross recharge. Similar problem was also observed by Niswonger et al. (2006). To fix it, a 6-month warm-up (spin-up) period, lasting from 1 June 2001 until 31 December 2001, was applied; hence, the final transient model was calibrated with data from 1 January 2002 till 31 December 2014 applying daily stress periods and daily time steps.

For model calibration, the Newtonian solver (MODFLOW-NWT) was used, applying the code option of “calculating groundwater heads even if below cell bottom” to prevent drying cells. The head tolerance was adjusted to 0.5 m, the flux tolerance to 5,000 m3 d−1 and the model complexity, to “complex”. All the remaining solver criteria were left as default settings. The model was calibrated manually because of its complexity; using optimization codes such as PEST (Doherty and Hunt 2010) or UCODE (Hill and Tiedeman 2006) turned to be computationally and time-wise too demanding. Besides, the manual, trial-and-error calibration allows users to better understand the model behaviour (Hassan et al. 2014).

The steady-state and transient model calibrations aimed at minimising the mean absolute error (MAE) and the root mean square error (RMSE) of the differences between the simulated and measured groundwater heads and also water balance discrepancies at each time step as per Eqs. (2) and (3). The calibration process was done in all six simulated layers, adjusting the initially assigned zones of hydraulic conductivity (Kh), mainly those with scarcity of pumping test data. In the same way, also zones of specific yield (Sw) and specific storage (Sw) were adjusted (Table 1). The complete information explaining which parameter was calibrated and which not, can be found in Table 1.

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |H_{\text{obs}} - H_{\text{sim}}|
\]

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (H_{\text{obs}} - H_{\text{sim}})^2}
\]

The sensitivity analysis mainly focused on testing sensitivity of fluxes representing water exchange between unsaturated and saturated zone, i.e. groundwater exfiltration (EXFgw), groundwater evapotranspiration (ETg), gross recharge (Rg) and net recharge (Ra), to changes in selected parameters.
Various parameters were tested in that respect, to find those influencing the most surface-groundwater exchange; finally, three parameters were selected to be tested, i.e. vertical hydraulic conductivity ($K_v$) of the KSU, evapotranspiration extinction depth (EXTDP) and soil saturated water content ($\theta_s$).

Water balances

Water balancing of the multi-layered aquifer system, particularly when simulated with variably saturated models, can be a complex issue because of many interacting unsaturated and saturated zone components (Fig. 3). The water balance of the whole CKB model domain can be expressed as follows:

$$ P + q_{GHB} = I + ET_{ss} + q_{ABS} + q_{DRN} \pm \Delta S $$

where $P$ is precipitation, $q_{GHB}$ is lateral groundwater inflow into the modelled area across the GHB boundary, $q_{DRN}$ is lateral groundwater outflow out of the modelled area across the DRN boundary, $I$ is canopy interception loss, $ET_{ss}$ is subsurface evapotranspiration, $q_{ABS}$ is groundwater abstraction, and $\Delta S$ is total change in storage.

The $ET_{ss}$ and $\Delta S$ can be expressed as follows:

$$ ET_{ss} = ET_{uz} + ET_g $$

$$ \Delta S = \Delta S_{uz} + \Delta S_g $$

where $ET_{uz}$ is unsaturated zone evapotranspiration; $ET_g$ is groundwater evapotranspiration; $\Delta S_{uz}$ is storage change in unsaturated zone; and $\Delta S_g$ is storage change in the saturated zone.

The unsaturated zone water balance is expressed as:

$$ P_e = P_e + EXF_{gw} = R_g + ET_{uz} \pm \Delta S_{uz} $$

where: $P_e$ is effective precipitation ($P_e = P - I$), $EXF_{gw}$ is groundwater exfiltration; $R_g$ is gross recharge; $P_a$ is actual infiltration (El-Zehairy et al. 2018).

The saturated zone water balance for all the simulated four layers can be expressed as follows:
\[ R_g + q_{\text{GHB}} = q_{\text{ABS}} + ET_g + q_{\text{DRN}} + EXF_{gw} + \Delta S_g \]  \hfill (8)

The net recharge \( (R_n) \) is expressed as follows (Hassan et al. 2014):

\[ R_n = R_g - EXF_{gw} - ET_g \]  \hfill (9)

**Results and discussion**

**Model calibration**

The estimated and calibrated MOD-UZF hydraulic parameters are presented in Table 1. Figure 6 shows the comparison between the simulated and the measured heads for the 13-year calibration period for the selected representative boreholes as in Fig. 1. In general, there is a good match of the simulated with the measured temporal head patterns. The MAE values for the control points ranged from 0.02 to 2.70 m and the RMSE from 0.02 to 3.13 m (Eqs. (2 and 3)). The likely explanations for discrepancies between the simulated and the measured heads include: (1) averaging of the simulated heads within the 25-km² model cell; (2) potential errors in the abstraction data of piezometers/boreholes affected by wellfield groundwater abstraction (TP34J, W14 J, WF6 OB18O, WF5 OB100, W47 J, WF2 OB10 and WF2 OB3O); (3) unrepresented heterogeneity within the 25-km² model cell; (4) uncertainty in the measured water levels; (5) eventual errors in model parameterization.

It can be seen in Fig. 6 that there are wide ranges of slopes of the head declines. The particularly steep declines of groundwater heads are observed in boreholes W13J, W43J and TP34J located in the wellfield operated by Debswana Diamond Mining Company (DDMC) in Jwaneng and in boreholes EB17O, WF2 OB10 and WF6 OB18O, also operated by DDMC, but in Orapa. The heads in areas outside the wellfields’ influences (BH4743, BH7763, BH7764, BH9294, BH9297 and BH10224) also decline but with substantially gentler slopes. These declines are because the Kalahari area is affected by: (1) relatively low rainfall within the 13-year simulation period; (2) substantial ET_{uz} due to large PET and a thick unsaturated zone restricting rainfall infiltration and recharge; and (3) considerable ET_{g} due to groundwater uptake by deep rooted trees (Alaghmand et al. 2014; Obakeng et al. 2007), and possibly also due to direct groundwater evaporation from the water table (Balugani et al. 2016), both reducing net recharge and as such declining the water table and the groundwater resources.

The general head decline throughout the 13-year simulation period shows that the relatively low \( R_g \) was not able to compensate groundwater discharge occurring mainly by \( ET_g \) (Lubczynski 2000; Lubczynski 2009) and by lateral groundwater outflow while only marginally by abstractions for livestock watering and by \( EXF_{gw} \). In the areas affected by mine groundwater abstractions, that disproportion was much more distinct.

In the Kalahari, substantial replenishment of groundwater resources, on average, occurs only once per decade, in response to exceptionally high rainfall years (Lubczynski 2011, 2009; Obakeng et al. 2007; Wanke et al. 2008), while within this study period, there was no such rainfall year. The last exceptionally high rainfall year and aquifer replenishment was in the wet season of 1999/2000 characterized by rainfall of 970 mm yr\(^{-1}\) (Obakeng et al. 2007), i.e. before this study simulation period. Since then, the CKB heads have declining trend as can be seen in Fig. 6.
Water balances

The yearly means of water balance components of the whole model domain per each of the 13 simulated hydrological years are presented in Table 2, while the 13-year means per each HU, presenting quantitative groundwater exchange between the six layers, are shown in the schematic block-diagram in Fig. 7. Note that in the CKB, the hydrological year starts from 1 September of the previous year and ends 31 August of the analysed year. The 13-hydrological year starts from 1 September of the previous year and ends 31 August of the analysed year. The 13-hydrological-year annual water balance of the Central Kalahari Basin as per Eqs. (4), (7) and (8). All values are in mm yr$^{-1}$

The input of the saturated zone water balance (Eqn (8)) consists of $R_g$ (1.87 mm yr$^{-1}$) and $q_{GHB}$ (0.30 mm yr$^{-1}$), while the output is dominated by $ET_g$ (3.12 mm yr$^{-1}$), followed by $q_{DRN}$ (0.94 mm yr$^{-1}$) reflecting lateral groundwater outflow and $q_{ABS}$ (0.22 mm yr$^{-1}$). In the 13 investigated years, in the CKB, there was dominance of groundwater output as compared to input, which is reflected by the negative mean $\Delta S_g$ (−2.11 mm yr$^{-1}$), as only in one hydrological year (2006) with the largest rainfall of 664.45 mm yr$^{-1}$, was the $\Delta S_g$ positive. This also explains the declining water table within the 13 simulated years.

The $R_e$ was estimated as $R_e = ET_g$ (Eq. 9) because EXFgw ~0. The positive $R_e$ indicates $R_g > ET_g$ and the negative $R_e$ indicates $R_g < ET_g$. Throughout the 13 hydrological years of the model simulation (Table 2), the $R_e$ was typically negative, except for the 2 years with rainfall distinctly above-average, i.e. 2006 when $P = 664.45$ mm yr$^{-1}$ and $R_e = 3.42$ mm yr$^{-1}$ and 2014 when $P = 605.90$ mm yr$^{-1}$ and $R_e = 0.98$ mm yr$^{-1}$. However, these two, relatively wet years could not compensate the remaining 11 years with negative $R_e$, so the 13-year mean $R_e = −1.25$ mm yr$^{-1}$. It is interesting that the largest yearly $R_e$ (2006) coincided, as expected, with the largest $P$ and $R_g$, but, unexpectedly, the lowest $R_e$ (2002) coincided with the highest $ET_g$, not with the lowest $P$ and $R_g$. The simulated $ET_g$ was the highest in 2002, because at the beginning of the simulation period, the water table was still pretty high after the replenishment in the extremely wet season of 2000 (Obakeng et al. 2007). Unfortunately, there were no sufficient data available in this study to start the model simulation from that year 2000 or earlier.

<table>
<thead>
<tr>
<th>Hydrological year</th>
<th>$P_e$</th>
<th>$P_g$</th>
<th>$I$</th>
<th>ET$_{au}$</th>
<th>ET$_{uz}$</th>
<th>ET$_g$</th>
<th>$R_g$</th>
<th>$R_n$</th>
<th>$q_{GHB}$</th>
<th>$q_{DRN}$</th>
<th>$q_{ABS}$</th>
<th>$\Delta S$</th>
<th>$\Delta S_{au}$</th>
<th>$\Delta S_g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>445.40</td>
<td>363.49</td>
<td>8.91</td>
<td>413.34</td>
<td>408.00</td>
<td>5.34</td>
<td>1.88</td>
<td>−3.46</td>
<td>0.30</td>
<td>0.94</td>
<td>0.22</td>
<td>22.29</td>
<td>26.61</td>
<td>−4.32</td>
</tr>
<tr>
<td>2003</td>
<td>357.29</td>
<td>350.14</td>
<td>7.15</td>
<td>364.99</td>
<td>361.10</td>
<td>3.88</td>
<td>0.66</td>
<td>−3.22</td>
<td>0.30</td>
<td>0.94</td>
<td>0.22</td>
<td>−15.70</td>
<td>−11.63</td>
<td>−4.08</td>
</tr>
<tr>
<td>2004</td>
<td>439.49</td>
<td>430.70</td>
<td>8.79</td>
<td>419.82</td>
<td>416.91</td>
<td>2.91</td>
<td>1.42</td>
<td>−1.49</td>
<td>0.30</td>
<td>0.95</td>
<td>0.22</td>
<td>10.02</td>
<td>12.37</td>
<td>−2.35</td>
</tr>
<tr>
<td>2005</td>
<td>386.60</td>
<td>378.87</td>
<td>7.73</td>
<td>383.38</td>
<td>380.57</td>
<td>2.82</td>
<td>0.97</td>
<td>−1.85</td>
<td>0.30</td>
<td>0.94</td>
<td>0.22</td>
<td>−5.38</td>
<td>−2.67</td>
<td>−2.71</td>
</tr>
<tr>
<td>2006</td>
<td>664.45</td>
<td>651.16</td>
<td>13.29</td>
<td>533.78</td>
<td>530.98</td>
<td>2.81</td>
<td>6.23</td>
<td>3.42</td>
<td>0.30</td>
<td>0.94</td>
<td>0.22</td>
<td>96.51</td>
<td>93.96</td>
<td>2.55</td>
</tr>
<tr>
<td>2007</td>
<td>330.48</td>
<td>323.87</td>
<td>6.61</td>
<td>329.45</td>
<td>324.67</td>
<td>4.78</td>
<td>1.52</td>
<td>−3.26</td>
<td>0.30</td>
<td>0.94</td>
<td>0.22</td>
<td>−6.43</td>
<td>−2.32</td>
<td>−4.12</td>
</tr>
<tr>
<td>2008</td>
<td>483.41</td>
<td>473.74</td>
<td>9.67</td>
<td>473.55</td>
<td>470.23</td>
<td>3.31</td>
<td>1.33</td>
<td>−1.98</td>
<td>0.30</td>
<td>0.94</td>
<td>0.22</td>
<td>−0.67</td>
<td>2.17</td>
<td>−2.84</td>
</tr>
<tr>
<td>2009</td>
<td>510.24</td>
<td>500.04</td>
<td>10.20</td>
<td>489.85</td>
<td>487.13</td>
<td>2.72</td>
<td>1.40</td>
<td>−1.32</td>
<td>0.30</td>
<td>0.94</td>
<td>0.22</td>
<td>9.33</td>
<td>11.51</td>
<td>−2.18</td>
</tr>
<tr>
<td>2010</td>
<td>556.79</td>
<td>545.65</td>
<td>11.14</td>
<td>543.24</td>
<td>540.92</td>
<td>2.32</td>
<td>1.55</td>
<td>−0.78</td>
<td>0.30</td>
<td>0.94</td>
<td>0.23</td>
<td>1.54</td>
<td>3.19</td>
<td>−1.65</td>
</tr>
<tr>
<td>2011</td>
<td>590.72</td>
<td>579.91</td>
<td>11.81</td>
<td>567.16</td>
<td>564.79</td>
<td>2.37</td>
<td>2.02</td>
<td>−0.35</td>
<td>0.30</td>
<td>0.94</td>
<td>0.22</td>
<td>10.89</td>
<td>12.11</td>
<td>−1.21</td>
</tr>
<tr>
<td>2012</td>
<td>307.21</td>
<td>301.07</td>
<td>6.14</td>
<td>304.30</td>
<td>301.38</td>
<td>2.92</td>
<td>1.29</td>
<td>−1.63</td>
<td>0.30</td>
<td>0.94</td>
<td>0.21</td>
<td>−4.07</td>
<td>−1.60</td>
<td>−2.48</td>
</tr>
<tr>
<td>2013</td>
<td>287.84</td>
<td>282.08</td>
<td>5.76</td>
<td>284.00</td>
<td>281.62</td>
<td>2.38</td>
<td>1.09</td>
<td>−1.29</td>
<td>0.30</td>
<td>0.94</td>
<td>0.21</td>
<td>−2.76</td>
<td>−0.63</td>
<td>−2.14</td>
</tr>
<tr>
<td>2014</td>
<td>605.90</td>
<td>593.78</td>
<td>12.12</td>
<td>545.77</td>
<td>543.76</td>
<td>2.01</td>
<td>2.99</td>
<td>0.98</td>
<td>0.30</td>
<td>0.94</td>
<td>0.22</td>
<td>47.15</td>
<td>47.03</td>
<td>0.12</td>
</tr>
</tbody>
</table>

The hydrological year starts from 1 September of the previous year and ends 31 August of the analysed year.
The lateral and vertical water flux exchanges through the six layers of the CKB are presented in Fig. 7 as 13-year means. In that diagram, each layer receives a number of input and output water fluxes, and the difference between them per layer represents its storage change. The presented water balance is pretty complex because of the very thick top KSU layer, redistributing rainfall water into various underlying layers and partitioning that water between $R_{g}$, $E_T$, and $ET_g$. That complexity is also because of the complex structural geology and hydrogeology of the simulated area, with step-wise pinching out layers underlying KSU, which implies complex water flux exchanges between layers. For example, the $R_{g}$, consists of five $R_{g}$-components (Fig. 7), each addressing a different layer and each constrained by the presence of the shallowest, unconfined water table, located in one of the following layers; these $R_{g}$-components are: 1.49 mm yr$^{-1}$ to saturated KSU, 0.23 mm yr$^{-1}$ to SBA, 0.04 mm yr$^{-1}$ to LA, 0.03 mm yr$^{-1}$ to EA and 0.08 mm yr$^{-1}$ to GA, all five summing up to total of 1.87 mm yr$^{-1}$ (Table 2).

The total $R_{g}$ was deduced directly from the MOD-UZF water budget output; however, the $R_{g}$ components were defined indirectly by delineation of water budget zones with ZONEBUDGET postprocessor in layers underlying fully unsaturated KSU and by calculating downward water fluxes in those zones. Such an additional, indirect calculation protocol was applied, because the current UZF1 package, estimates $R_{g}$ only within the water-table extent of the layer to which the UZF1 Package is assigned—in this study case, the KSU. Considering $ET_g$, there was no analogic water budgeting problem, because the EXTDP was everywhere less than the KSU thickness, while the EXFgw was negligible.

Considering external groundwater exchange with CKB (Fig. 7), the lateral groundwater inflow enters LA from the east (0.18 mm yr$^{-1}$) and the EA from the two sides in the south (0.03 + 0.09 mm yr$^{-1}$) as shown in Fig. 5. The presence of these lateral inflows (defined in the model throughout the GHB boundary condition) rejected the hydrogeological conceptual model hypothesis that the CKB is a fully isolated basin (Lekula et al. 2018a), although the simulated inflows were pretty low, and possibly triggered by the wellfields’ abstractions. The north-eastern lateral groundwater outflows towards Makgadikgadi Pans were (Fig. 5): (1) in KSU, negligible; (2) in LA, 0.29 mm yr$^{-1}$, so more than lateral input to LA; (3) in EA, 0.28 mm yr$^{-1}$, so much more than lateral input to EA; and (4) in GA, which did not have any lateral input, the groundwater outflow was the largest (0.39 mm yr$^{-1}$) mainly because of the largest, positive balance of the interlayer, water flux exchange with the saturated KSU (Fig. 7), where the input from the saturated KSU was 0.62 mm yr$^{-1}$, while the output, only 0.28 mm yr$^{-1}$. The large downward water input was because of the peripheral GA position with respect to the CKB and relatively large area of saturated KSU being in direct hydraulic contact with GA (Lekula et al. 2018a). In contrast, the lowest interlayer water
flux exchange difference between different layers was observed between KSU and EA where the same 0.09 mm yr\(^{-1}\) went in and out between the layers. Surprising is relatively large groundwater exchange across the SBA (Fig. 7). This basaltic layer has low storage but pretty high vertical permeability due to locally occurring vertical fracture zones. The 13-year mean groundwater abstractions (\(q_{ABs}\)) in all the three aquifers look pretty small as compared to other water fluxes, because they are referenced to the whole CKB model domain (~200 Mm\(^2\)), while at the local scales, these abstractions are significant.

**Spatial variability of groundwater fluxes**

Figure 8a presents the spatial variability of \(R_g\), \(ET_g\) and \(R_n\) in the wettest simulated hydrological year (2006) and Fig. 8b in the driest year, 2013 (Table 2). It can be seen that in both years, \(R_g\), \(ET_g\) and \(R_n\) were highly localized, being limited to small areas such as fossil river channels and depressions (Fig. 8a). This is in agreement with de Vries et al. (2000), who also observed only locally enhanced recharge of up to 50 mm yr\(^{-1}\) in pans and fossil valleys in the southern part of the CKB. The spatial restriction of groundwater fluxes to relief depressions and fossil channels is mainly due to periodic rainfall water storing in these locations, to local increase of soil moisture and to shallowing of water table, all creating favourable conditions for \(R_g\), although also enhancing \(ET_{uz}\) and \(ET_g\).

The \(R_g\) (6.23 mm yr\(^{-1}\)) in the 2006 hydrological year, was larger and covered a much bigger area than the \(R_g\) (1.09 mm yr\(^{-1}\)) in 2013 (Fig. 8b), while the \(ET_g\) in 2006 and 2013, were comparable (2.81 and 2.38 mm yr\(^{-1}\) respectively). As such, the \(R_n\) in 2006 was positive (3.42 mm yr\(^{-1}\)), having spatial extent similar to \(R_g\), while the \(R_n\) in 2013 was negative (~1.29 mm yr\(^{-1}\)) and restricted to similar locations as the \(ET_g\). It can be concluded that the spatio-temporal CKB patterns of \(R_g\) depend mainly on spatio-temporal variability of rainfall, surface morphology, thickness of unsaturated zone, and vegetation type and density.

**Long-term temporal variability of water fluxes**

Large temporal variability of surface and subsurface water fluxes, both on a daily (Fig. 9) and yearly basis (Table 2), is observed. The CKB is characterized by erratic high-rainfall days (restricted to wet season), relatively low interception (Table 2), and erratic actual infiltration events, some of them even >20 mm d\(^{-1}\) (Fig. 9a). However, the majority of that infiltration is removed from the unsaturated zone by generally large \(ET_{uz}\), ranging from nearly zero in dry season when soil moisture is low or negligible, to even 5 mm d\(^{-1}\) during the wet season, so only a small portion of the infiltrated water arrives at the water table. This is because of: (1) the extremely large Kalahari PET, the largest in the hottest wet season; (2) the very thick unsaturated zone (with ‘thirsty’ Kalahari plants), which enhances water loss and restricts \(R_g\) to erratic daily episodes that in the 13 years of this study period, ranged from 0 up to only 0.13 mm d\(^{-1}\) in 2006 (Fig. 9b). The \(ET_g\) is less temporally variable, varying from 0 to 0.02 mm d\(^{-1}\) in the similar manner as the water table, i.e. its peak is delayed several months with respect to the peak of the wet season rains. As such, the \(ET_g\) peaks are also offset with respect to \(ET_{uz}\) peaks, as the latter are mainly dependent on climatic factors, so rather matching the PET peaks.

The daily variability of \(R_n\) is presented in Fig. 9b. As defined by the difference between highly temporally variable \(R_g\) and moderately variable \(ET_g\), the resultant \(R_n\)-pattern follows the \(R_g\)-pattern, being also highly temporally variable, ranging from ~0.02 to 0.13 mm d\(^{-1}\). It is remarkable that on a yearly basis, there are only relatively short periods with \(R_n\) > 0, occurring not even every year. The exceptions are years 2006 and 2014 with above average annual precipitation (Table 2), when relatively large \(R_g\) was observed during many days, resulting in a positive annual \(R_n\). In the other 11 years, the \(ET_g\) was dominant, so annual \(R_n\) < 0. The annual variability of \(R_n\) as well as of other water fluxes, is presented in Table 2. The nature of \(R_g\) and \(R_n\) dependence on precipitation is illustrated in Fig. 10, where it can be seen that below the ~600-mm yr\(^{-1}\) annual rainfall threshold, there is nearly no change of \(R_g\) and \(R_n\). The substantial \(R_g\) and \(R_n\) increments take place only when annual rains exceed 600 mm. Assuming a linear trend of rainfall-recharge in the years with the largest rainfall, as in Fig. 10, the backward-estimated annual \(R_g\) and \(R_n\) for the exceptionally wet year 1999/2000, with 970-mm rainfall (year not simulated), were ~23 and 16 mm respectively. The preceding assumption of linear trend is still quite modest, so most likely the recharge input was even larger.

The episodic nature of recharge events in the CKB is mainly attributed to erratic rainfall, thick unsaturated zone and very high PET as well as large \(ET_{uz}\). Significant recharge events occur only in response to cumulated-in-time rainfall, consisting of a number of sequential above-average events. A similar observation was made also by Wanke et al. (2008) in a comparable environment.

**Sensitivity analysis**

The replenishment and therefore sustainability of groundwater resources largely depends on \(R_n\) (Lubczynski 2011, 2006); therefore, the sensitivity analysis in this study focused on testing \(R_g\), \(ET_g\) and \(EXF_{gw}\) as well as the resultant \(R_n\), all characterizing water exchange between the unsaturated and saturated zone. However, after preliminary tests, the \(EXF_{gw}\) was excluded from further sensitivity analysis, as the \(EXF_{gw}\) was negligible in all analysed years and in all tests, due to the generally deep water-table depth; thus, its sensitivity was also not relevant for the \(R_n\) estimate. Therefore, hereafter, the sensitivity analysis is shown only for \(R_g\) and \(ET_g\) and for the resultant \(R_n\) (Fig. 11), all in the wettest year within the study period (hydrological year 2006).
As the analysed year 2006 was relatively wet and the \( R_g \) was substantially larger than \( ET_g \), therefore it had generally much larger effect upon the \( R_n \) than in other years while varying unsaturated zone parameters \( \theta_s \), \( K_v \) and EXTDP. In contrast, in dry years, the \( R_n \) was totally dependent on \( ET_g \), as \( R_g \) was negligible. The \( ET_g \) sensitivity to changes of \( \theta_s \) was generally low (Fig. 11), regardless of the \( ET_g \) seasonal variability, driven mainly by the water-table fluctuation characterized by peaks delayed with respect to the occurrence of the wet seasons. The little peak of the \( ET_g \) in 0.6 \( \theta_s \) simulation at the end of April 2006, was likely attributed to the \(~1.5\)–\(~2.0\)-month delayed-water-table rise, in response to the \( R_g \) peak occurring around 1 March (Fig. 11). The \( K_v \) changes had very similar impact upon \( ET_g \), having a similar little peak of \( ET_g \) in April 2006 for 10-\( K_v \) simulation, likely due to the same reason as in the 0.6-\( \theta_s \) simulation. The sensitivity of the \( ET_g \) to EXTDP

---

**Fig. 8** Spatial variability of gross recharge (\( R_g \)), groundwater evapotranspiration (\( ET_g \)), and net recharge (\( R_n = R_g – ET_g \) as \( EXF_{gw} = 0 \)) for: a 2006; b 2013 hydrological years

---
was different compared to that of the other two parameters analysed. In general, the larger the \( \theta_s \) itself, the larger the differences were between different EXTDP simulations. It is remarkable that the \( \theta_s \) maxima and the largest differences between the three EXTDP simulations, were just before the wet season started, i.e. in October (Fig. 11c), with the largest \( \theta_s \) for 1.5-EXTDP simulation. After that peak, the differences between the three simulations gradually declined to be negligible already in February.

In the selected wet year 2006 (Fig. 11c), the \( R_g \) was sensitive to changes of all the three tested parameters (\( \theta_s \), \( K_v \) and EXTDP). The 0.6-\( \theta_s \) and 10-\( K_v \) simulations, as well as the 1.4-\( \theta_s \) and 0.1-\( K_v \) simulations, had very similar effects upon \( R_g \), as both parameters, i.e. \( \theta_s \) and \( K_v \), similarly influence \( K(\theta) \) in Eq. (1). What is remarkable in the presented \( R_g \) sensitivity patterns, is the peak around 1 March in response to wet-season accumulation of rain, with a clear sequence of peak occurrences, the fastest for the lowest 0.6\( \theta_s \) and for the largest 10\( K_v \), both, due to the largest \( K(\theta) \). It is also remarkable that only the two simulations, i.e. 0.6 \( \theta_s \) and 10 \( K_v \), resulted in the delayed substantial \( R_g \) extending throughout the dry season, while in all the other \( \theta_s \) and \( K_v \) simulations, \( R_g \) converged to

Fig. 9 Daily variability of different water balance components over the 13-year CKB simulation period: a actual infiltration (\( P_a \)), unsaturated zone evapotranspiration (\( \text{ET}_{uz} \)), gross recharge (\( R_g \)) and groundwater evapotranspiration (\( \text{ET}_g \)); b net recharge (\( R_n \))

Fig. 10 Cross-dependencies of yearly means of rainfall (\( P \)) versus gross recharge (\( R_g \)) and net recharge (\( R_n \))
zero shortly after the wet season. In contrast, the nonzero $R_g$ ‘tail’ extending throughout the dry season, was not present in any of the EXTDP simulations. Considering the sequential $R_g$ peaks, they had similar timing and pattern as the $\theta_s$ and $K_v$ simulations. The largest $R_g$ was attributed to the lowest EXTDP and the opposite, as the increment of EXTDP reduces the amount of water potentially available for $R_g$. The $R_g$ sensitivity, presented in Fig. 11, was very similar to the $R_g$ because of the small impact of $\text{ET}_g$ in the wet year 2006 and generally negligible impact of EXF$_{gw}$.

Experience of using remote sensing in data-scarce central Kalahari Basin

The recent introduction of integrated hydrological models (IHMs) creates promising ‘avenues’ for novel remote sensing (RS) applications, not only in surface water but also in groundwater studies. This is because, in contrast to the standard standalone groundwater models, where driving forces, i.e. $R_g$ and $\text{ET}_g$ were not quantifiable by RS, in the IHMs, the driving forces, i.e. rainfall and PET (Hassan et al. 2014), are well quantifiable by RS, while the $R_g$ and $\text{ET}_g$ are estimated internally by IHMs based on land surface and unsaturated zone parameterization.

One of the main challenges of integrated hydrological modelling, particularly in arid and semi-arid areas characterized by large spatio-temporal variability of water-related fluxes, has been insufficient availability and quality of surface and subsurface input data. This study shows that RS can contribute to regional-scale IHMs, providing various types of readily available (downloadable) RS products, most importantly, spatio-temporally variable driving forces such as rainfall and PET.

With advancement in RS techniques, various satellite rainfall products at different spatial and temporal resolutions are now readily available and their spatial and temporal resolution increases. However, these rainfall products still need to be validated against in situ data to select the optimal product and eventually to remove the bias (Lekula et al. 2018b; Rahmawati and Lubczynski 2017). Also satellite-derived PET data are available as an RS product, although not as widely as rainfall and at much coarser spatial and temporal resolution. However, even with that limitation, the RS-based PET estimates are still useful because the PET is much less spatio-temporally variable than rainfall. Besides, if necessary (e.g. in local-scale assessments), with some effort, PET can be also defined at much better spatio-temporal resolution from raw multispectral RS data, as for example by Kim and Hogue (2008).

The remotely sensed earth observation from space cannot contribute to subsurface hydrostratigraphy of a model, except the upper model boundary, i.e. the topographic surface. The topographic surface, is nowadays derived by RS techniques, applying for example interferometry (e.g. Noferini et al. 2007; Wegmüller et al. 2009), LiDAR (e.g. Liu et al. 2005; Ma 2005) or analysis of stereoscopic images (e.g. Haala and Rothermel 2012; Xu, et al. 2010). These methods provide digital elevation models (DEMs) already at fine spatial resolution—in this study the SRTM 90-m DEM was downloaded from the CGIAR-CSI database (Jarvis et al. 2008).
The UZF1 package of MODFLOW, which links surface input with groundwater, requires soil physical parameters and evapotranspiration extinction depth. The soil physical parameters can be estimated based on field sampling and literature sources but the challenge is how to spatially distribute them in the IHMs. For such a task as that, remotely sensed soil maps with quite detailed spatial distribution of different soil types is a solution. In this study, the spatial distribution of the soil physical parameters was defined using the “Soil Atlas of Africa” (Jones et al. 2013), largely based on remote sensing soil assessment, while parametric values were extracted from literature sources. The spatial distribution of the evapotranspiration extinction depth can be defined based on RS-based vegetation maps, provided the rooting depths of individual species can be realistically estimated from other sources. In this study, spatial distribution of evapotranspiration extinction depth was defined based on the RS-based, land use/land cover (LULC) maps (Loveland et al. 2000), while the rooting depth of plant species was defined based on Obakeng et al. (2007), Kleidon (2004) and Canadell et al. (1996). Spatial and temporal resolution of the RS products can be an issue limiting their applicability as IHM input. If temporal resolution of most of the currently available RS products is already sufficient for typical daily input data requirements of most of the IHMs, the spatial resolution, (especially of RS rainfall) is still a limiting factor in their applicability to local-scale IHMs, except maybe for some DEM products available at <100 m resolution. Because of that limitation, the RS products are still mainly used in the regional-scale models, particularly those over areas with lack of or scarce monitoring networks such as the CKB. However, as the RS products continuously improve their spatial (and temporal) resolution, it is expected that shortly, they will also be more frequently applied in IHMs at the local-scale applications.

Considering the scarcity of fine spatial-resolution RS products, in the local scale IHMs, only tailor-made quantitative RS applications based on moderate, high- or very high-spatial resolution images can be utilized such as for example: (1) evapotranspiration mapping using Moderate Resolution Imaging Spectroradiometer (MODIS) with spatial resolution ranging from 0.5 to 1 km (Bastiaanssen et al. 1998a, b; Su 2002) or at high-resolution using Landsat 8 imagery with spatial resolution of 100 m (Senay et al. 2016); (2) tree transpiration mapping using very high-resolution QuickBird or WorldView at 60–40 cm per pixel (Reyes-Acosta and Lubczynski 2013); (3) tree interception using QuickBird and WorldView at 60–40 cm per pixel (Hassan et al. 2017); the aforementioned tree transpiration and tree interception mapping methods require tree-scaling functions, which for nine dominant Kalahari tree species are presented by Lubczynski et al. (2017).

In RS applications at the local scale, unmanned aerial systems (drones) that can carry on-board multispectral cameras, can provide the required very high spatial resolution and are also cost effective (Colomina and Molina 2014). They are very convenient and efficient, so very promising in various environmental applications. However, drone data processing is still cumbersome and requires specialized knowledge; besides, in many countries use of drones requires a specific license, which can be difficult to obtain.

The reliability of models depends not only on driving forces and parameters, but also on state variables that represent model calibration reference. The RS techniques of earth observation from satellites, cannot detect the most commonly calibrated state variable, i.e. the water table, even at shallow water-table conditions. In that respect, the Gravity Recovery and Climate Experiment (GRACE) satellite, a joint mission launched by National Aeronautics Space Administration (NASA) and German Aerospace Center (DLR), was promising; this system detects changes in subsurface water storage, and thus also changes of aquifer water levels, through the analysis of gravity change (e.g. Leblanc et al. 2009; Rodell and Famiglietti 2001; Seoane et al. 2013; Yeh et al. 2006). However, GRACE assessment is done at a very coarse resolution of ~400 × 400 km (Sutanudjaja et al. 2014; Tapley et al. 2004), restricting its applicability to regional, or even continental-scale basins.

Another popular state variable applied in IHMs is river discharge. In contrast to groundwater levels, river discharges in ungauged catchments, can be approximated applying RS techniques (e.g Brakenridge et al. 2007; De Groeve 2010; Hirpa et al. 2013), although in this study such an assessment was not needed because in the CKB there are no flowing rivers.

With advancement in IHMs and improvement of accuracy of RS solutions of spatio-temporal soil moisture and actual evapotranspiration, these two variables can also be used as state variables in model calibration (e.g. Li et al. 2009; Lopez et al. 2017), although the currently available RS solutions of soil moisture and actual evapotranspiration still involve substantial error, particularly in dryland applications, where the error of RS quantification can be comparable or larger than the size of recorded water fluxes. Besides, the web-based products of soil moisture and actual evapotranspiration are still available only at a coarse resolution, which restricts their use in IHMs to large-scale regional assessments only, otherwise forcing researchers to process raw, higher-resolution RS data, which, considering the typical, daily IHM data input requirement, is not only specialized, but also cumbersome and time consuming.

**Conclusions**

The increasingly used integrated hydrological models (IHMs) need spatio-temporally distributed surface-input data such as P, I and PET (driving forces), unsaturated zone parameters, i.e. soil physical properties and evapotranspiration extinction depths, hydrogeological parameters and also state variables...
such as hydraulic heads or river discharges. Typically, such data on the ground are scarce or unavailable, particularly in remote areas of developing countries such as the Central Kalahari Basin (CKB). The alternative source is Remote Sensing (RS), although it does not provide all data types that are required by IHM input. The main objective of this study was to present the use of RS data in the setting of the IHM of a multi-layered aquifer system characterized by a very thick unsaturated zone such as the CKB and to provide its quantitative assessment. The key findings of this study are listed below:

1. The presented multi-layered CKB hydrogeological system has three aquifers—Lebung, Ecca and Ghanzi—whereby each aquifer receives diffuse recharge from rainfall through the top unconfined Kalahari Sand layer (either partially saturated or entirely unsaturated), while the Lebung and Ecca aquifers receive additional groundwater input from other overlying layers and small lateral inflows from outside the CKB. The flow patterns in all the aquifers are similar, with piezometric surfaces radially converging towards the central part of the basin from where all three aquifers discharge groundwater towards Makgadikgadi Pans. A considerable amount of groundwater is also discharged by groundwater evapotranspiration.

2. In semi-arid aquifer systems with a thick unsaturated zone such as the CKB, subsurface evapotranspiration is the dominant discharging water flux comparable with rainfall, while groundwater exfiltration is negligible because of the deep water table. As a consequence of the latter, gross recharge and groundwater evapotranspiration are comparable, so their yearly balance represented by the net recharge, is typically low or close to zero. Whether that balance is positive or negative, depends primarily on the annual rainfall amount and distribution, and secondarily on the water-table depth.

3. The groundwater resources in semi-arid aquifer systems with a thick unsaturated zone such as the CKB, are sustained only through recharge events of exceptionally wet years; however, within the 13-year simulation period, from 2002 to 2014, there was no such wet year. The wettest year was 2006 with annual rainfall 664 mm, gross recharge 6.2 mm and net recharge 3.0 mm; in the majority of other simulated years with lower rainfall, gross recharge was less than groundwater evapotranspiration, resulting in typically negative net recharge; the lowest, $-3.5$ mm, was in 2002. The generally negative net recharge within the simulation period is the main reason for the observed declining trend of the water table and groundwater storage in the CKB.

4. Amount and temporal distribution of rainfall, surface morphology, thickness of the unsaturated zone and vegetation type/density, are primary determinants of the $R_n$ spatial distribution in the CKB. The porosity, specific yield and vertical hydraulic conductivity of the unsaturated zone material are also important.

5. Advancement in integrated hydrological models (IHM$s$), coupling surface processes and groundwater flows, created new opportunities for using remote sensing (RS) techniques in hydrogeology. The RS can nowadays provide input data for IHMs at reasonable spatial and quite good temporal resolutions. Lots of such data is readily and freely-available as web-based products, which is particularly important in data-scarce areas with insufficient density of monitoring networks and/or inaccessible areas such as in the CKB. In this CKB-IHM study, the main RS contributions addressed rainfall, potential evapotranspiration, land cover and land use types, and terrain elevation, all acquired as web-based products.

6. As a follow up to this study it is recommended in the CKB to: (1) densify monitoring of water-table, to better control the IHM calibration process; (2) densify monitoring network of rainfall to improve bias correction of remotely sensed rainfall estimates; (3) investigate rooting depth of Kalahari plants, their water uptake and interception; (4) attempt to use the RS-based soil moisture, actual evapotranspiration and eventually GRACE satellite storage change as state variables of the IHM.

Acknowledgements Special thanks go to the Department of Water Affairs in Botswana, Botswana Geoscience Institution, Debswana Diamond Mining Company and Directorate of Water Resources Management in Namibia, for providing the geological and hydrogeological data that made this study possible. Dr. Richard G Niswonger and Dr. Richard B Winston from USGS are greatly acknowledged for their interactive discussion about the UZF1 package and two anonymous reviewers for their constructive critiques that substantially contributed to the improvement of this paper. Free online databases for geological data for Namibia used in this study are also acknowledged.

Funding information This study is part of the first author’s PhD research, which was supported financially by the Botswana Government through Botswana International University of Science and Technology.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


Bhalotra YPR (1987) Climate of Botswana Part II: elements of climate. Department of Meteorological Services, Gaborone, Botswana


Bhalotra YPR (1987) Climate of Botswana Part II: elements of climate. Department of Meteorological Services, Gaborone, Botswana


Su Z (2002) The surface energy balance system (SEBS) for estimation of ground-based methods in multi-scale hydrology. IAHS Series of


