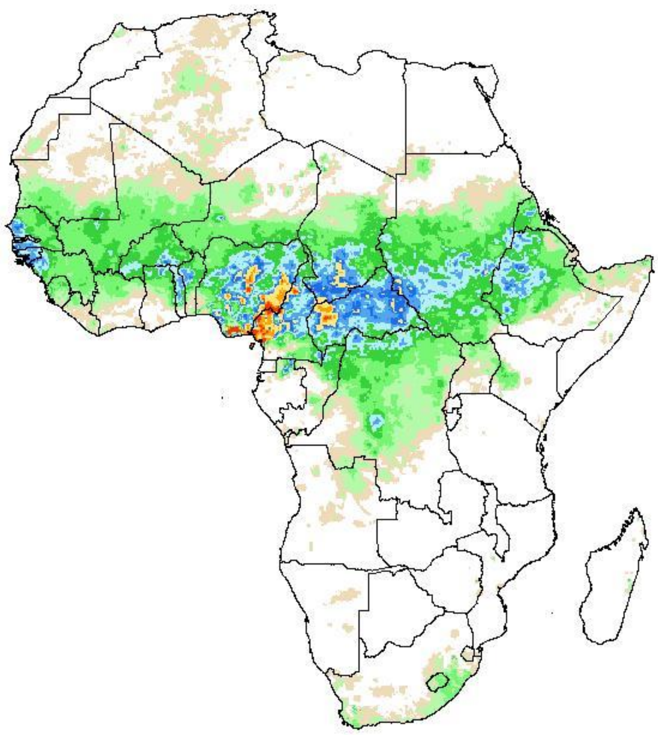
British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

A review of continent scale hydrological datasets available for Africa

Groundwater Science Programme

Internal Report IR/11/016



BRITISH GEOLOGICAL SURVEY

GROUNDWATER SCIENCE PROGRAMME

INTERNAL REPORT IR/11/016

A review of continent scale hydrological datasets available for Africa

Bonsor HC

The National Grid and other Ordnance Survey data are used with the permission of the Controller of Her Majesty's Stationery Office.
Licence No: 100017897/2010.

Keywords

Africa; aquifer properties; systematic data review; hydrogeology data

Bibliographical reference

BONSOR HC . 2010. A review of continent scale hydrological datasets available for Africa. *British Geological Survey, Internal Report*, IR/11/016. 28pp.

Copyright in materials derived from the British Geological Survey's work is owned by the Natural Environment Research Council (NERC) and/or the authority that commissioned the work. You may not copy or adapt this publication without first obtaining permission. Contact the BGS Intellectual Property Rights Section, British Geological Survey, Keyworth, e-mail ipr@bgs.ac.uk. You may quote extracts of a reasonable length without prior permission, provided a full acknowledgement is given of the source of the extract.

Maps and diagrams in this book use topography based on Ordnance Survey mapping.

© NERC 2010. All rights reserved

Keyworth, Nottingham British Geological Survey 2010

BRITISH GEOLOGICAL SURVEY

The full range of our publications is available from BGS shops at Nottingham, Edinburgh, London and Cardiff (Welsh publications only) see contact details below or shop online at www.geologyshop.com

The London Information Office also maintains a reference collection of BGS publications, including maps, for consultation.

We publish an annual catalogue of our maps and other publications; this catalogue is available online or from any of the BGS shops.

The British Geological Survey carries out the geological survey of Great Britain and Northern Ireland (the latter as an agency service for the government of Northern Ireland), and of the surrounding continental shelf, as well as basic research projects. It also undertakes programmes of technical aid in geology in developing countries.

The British Geological Survey is a component body of the Natural Environment Research Council.

British Geological Survey offices

BGS Central Enquiries Desk
Tel 0115 936 3143 Fax 0115 936 3276
email enquiries@bgs.ac.uk

Kingsley Dunham Centre, Keyworth, Nottingham NG12 5GG
Tel 0115 936 3241 Fax 0115 936 3488
email sales@bgs.ac.uk

Murchison House, West Mains Road, Edinburgh EH9 3LA
Tel 0131 667 1000 Fax 0131 668 2683
email scotsales@bgs.ac.uk

Natural History Museum, Cromwell Road, London SW7 5BD
Tel 020 7589 4090 Fax 020 7584 8270
Tel 020 7942 5344/45 email bgslondon@bgs.ac.uk

Columbus House, Greenmeadow Springs, Tongwynlais, Cardiff CF15 7NE
Tel 029 2052 1962 Fax 029 2052 1963

Maclean Building, Crowmarsh Gifford, Wallingford OX10 8BB
Tel 01491 838800 Fax 01491 692345

Geological Survey of Northern Ireland, Colby House, Stranmillis Court, Belfast BT9 5BF
Tel 028 9038 8462 Fax 028 9038 8461

www.bgs.ac.uk/gsni/

Parent Body

Natural Environment Research Council, Polaris House, North Star Avenue, Swindon SN2 1EU
Tel 01793 411500 Fax 01793 411501
www.nerc.ac.uk

Website www.bgs.ac.uk
Shop online at www.geologyshop.com

Foreword

As rainfall becomes less reliable with predicted climate change the ability to assess the spatial and seasonal variations in groundwater availability on a large-scale (catchment and continent) is becoming increasingly important (Bates, et al. 2007; MacDonald et al. 2009). The scarcity of observed hydrological data, or difficulty in obtaining such data, within Africa means remotely sensed (RS) datasets must often be used to drive large-scale hydrological models. The different applicability of RS data in different regions of the world means that the correct datasets must be selected for Africa for accurate hydrological models to be developed.

This report provides a review of the main hydrogeological datasets available for Africa to develop and validate continental-scale hydrological models. The review assesses the suitability of each data set specifically for use in Africa. This report feeds into the 'Groundwater and Environmental Change' science budget project initiated in 2010 to develop a continental-scale recharge model of Africa, however, the review also forms a general resource for BGS projects in Africa.

Acknowledgements

As with all projects there are a number of BGS and external colleagues to thank for their help and assistance:

Chris Taylor (CEH) – for his invaluable practical experience and advice in using different precipitation datasets and contact to African Monsoon Multidisciplinary Analyses (AMMA).

Tanya Waneraars (CEH) – WATCH project manager – for help in general information to WATCH WFD project and contact to the researchers.

Graham Weedin (CEH) – WATCH WFD data processing researcher – for insight into development of the WATCH WFD dataset.

Luke Bateson (BGS) – for general assistance with the background of remotely sensed datasets, particularly with issues surrounding different digital elevation models.

Alan Nicol (NBI) – for assistance with investigating other non-web sources of river discharge gauge data in northeast Africa.

Declan Conway (UEA) – for assistance with investigating other non-web sources of atmospheric and river discharge gauge data in northeast and west Africa.

Contents

TABLES

1 Introduction

This report provides a review of the main hydrogeological datasets available for Africa to develop and validate continental-scale hydrological models. The work was undertaken as part of the BGS Groundwater Science research programme, under the 'Groundwater and Environmental Change' project, however, the review forms a general resource to any BGS project using remotely sensed (RS) or gauged, hydrological data within Africa.

One of the main tasks set within the 'Groundwater and Environmental Change' project is to develop a continental-scale recharge model of Africa, that builds on the recharge model of the Nile Basin using the BGS ZOODRM model with RS data in 2008-9 (Bonsor et al. 2010; Mansour and Hughes, 2004). To develop a continental-scale recharge model for Africa requires:

1. **an understanding of the different recharge mechanisms** operating in different parts of the continent
2. **the development of the BGS ZOODRM code** to include the main recharge processes, and
3. **knowledge of the most appropriate input data** to drive and validate a recharge model at a large-scale.

The scarcity of observed hydrological data, or difficulty in obtaining such data, within Africa means largely RS datasets must be used to drive and validate large-scale (catchment or continent) hydrological models in Africa. The different applicability of RS datasets in different regions of the world, however, means that the datasets must be selected carefully to develop accurate hydrological models.

This report provides a review of the main, readily available, hydrogeological datasets. The review assesses the suitability of each dataset specifically with respect for use in Africa. Available, continental-scale gauged datasets are also reviewed.

This report is an internal report for BGS projects based on current available data. As technology progresses and new datasets are issued, the review will require updating.

2 Hydrological data requirements of the ZOODRM recharge model

The main task within the BGS Groundwater and Environmental Change project is to develop a continental-scale recharge model for Africa using the BGS ZOODRM recharge model. ZOODRM is designed to calculate spatial and temporal variations in groundwater recharge, and has been applied successfully before in semi-arid areas (e.g. Palestine - Hughes et al., 2008; and, Inner Mongolia – Ó Dochartaigh et al., 2010) and in wet temperate areas (e.g. Europe – Jackson et al., 2005). Recharge is calculated over a daily time step within ZOODRM by applying the soil moisture deficit (SMD) recharge method (Penman, 1948; Grindley,1967) or the Wetting Threshold (WT) method (Lange et al. 2003). These methods can be used in different areas of the same recharge model (e.g. Hughes, et al. 2008) depending on climatic variability. The ZOODRM model outputs monthly estimates of rainfall, evapotranspiration (ET), surface run-off, river discharge, groundwater recharge and change in soil moisture.

The data demands of the ZOODRM model are relatively low compared to other more sophisticated land surface hydrological models (e.g. JULES, SWAT). The hydrological data requirements of the ZOODRM model are, spatially distributed:

- Daily rainfall data
- Daily potential evapotranspiration data (PET)
- Land cover data (affects ET)
- Digital Elevation Model (DEM) data (affects distribution of run-off)
- Geology data
- Monthly river discharge data

3 Review of continent scale hydrological datasets available for Africa

3.1 RAINFALL DATA

3.1.1 Background to RS rainfall datasets

Rainfall is the primary input to any hydrological model, and the accuracy of the rainfall dataset is significant to the accuracy of the final modelled outputs – e.g. groundwater recharge. RS rainfall data are modelled datasets, not actual measurements of total rainfall in a day at a point on the Earth's surface. To produce RS precipitation datasets, remote sensor measurements (e.g. average daily cloud cover and atmospheric moisture) are converted to rainfall rates, using different algorithms. The results are validated and/or corrected to the gauge data (as available) to produce the final RS dataset.

Accuracy of data for Africa

The accuracy of the rainfall datasets modelled from the satellite sensors for a point on the Earth's surface depends on:

a) The frequency of the orbit of the satellite above the point on the Earth's surface.
b) The assumptions made to derive a rainfall rate from the RS measurements for different types of rainfall.
c) The amount of observed rainfall data used/available to validate the RS dataset.

a) Frequency and density of RS data

Depending on the orbit of the satellite, sensors take precise measurements relating to rainfall at a point on the Earth's surface at different frequencies. In Africa, seasonal rainfall is predominantly incurred by movement of the Inter-Tropical Convergence Zone and, on a daily time-scale, by intense convection. Rainfall in Africa is largely seasonal and falls within very intense rainfall events, which usually last only a few hours. In contrast, rainfall within Europe occurs throughout the year and is driven by the convergence of pressure systems, which often gives several days of rainfall. A higher frequency of RS measurements, or use of a greater number of sensors of different frequencies, is therefore required to model rainfall accurately in Africa.

b) Assumptions used to model RS dataset

Depending on the type of rainfall typical in a region, different algorithms are required to generate an accurate daily rainfall rate from the RS measurements. Often RS datasets are global datasets, so rainfall in relatively high latitudes (e.g. in Europe) might be better modelled than rainfall in the tropics (e.g. Africa), or vice versa, depending on the algorithm used.

For example, the TRMM dataset which is focused to modelling rainfall in the tropics, is accurate to within ±14% of observed data in the tropics, but is highly inaccurate at higher latitudes (errors of 100-200% common). Selecting the best RS 'global' dataset for a given area is therefore important.

c) Level of validation

Due to variability in observed/gauged rainfall data worldwide, global RS datasets are often validated and corrected to a much greater accuracy in higher latitudes (e.g. the U.S.A. and

Europe). The stated accuracy of the RS dataset might not, therefore, reflect the accuracy of the dataset in Africa.

Ultimately, all RS and forced RS datasets have advantages and disadvantages, and no one dataset is perfect for use in a continent-scale recharge model of Africa.

3.1.2 RS rainfall datasets available for Africa

There are three main RS datasets of daily rainfall data which cover Africa (Table 1):

- NOAA – National Oceanographic and Atmospheric Administration
- TRMM – Tropical Rainfall Measurement Mission
- EUMETSAT – European Organisation for the Exploitation of Meteorological Satellites

(GPCC and CRU data are discussed in section 3.1.3 as part of the WATCH WFD dataset)

The RS datasets are derived from four main sensors in orbit on European and American satellites: the Special Sensor Microwave/Imager (SSM/I); the Advanced Microwave Sounding Unit (AMSU); the Precipitation Radar (PR); and the Spinning Enhanced Visible and Infrared Imager (SEVIRI).

Many other RS global rainfall datasets appear available from a web search, but nearly all are actually derived from the three listed above. The NOAA dataset is one of the most widely used within higher latitudes, and the TRMM is widely published from work in the tropics (Taylor, *pers. comm.*).

The **TRMM** dataset is aimed at modelling rainfall accurately in tropical latitudes. Within this dataset, daily rainfall values are derived for a point of the Earth's surface, from one very accurate PR measurement taken once in every 24 hours, and then several 6-hourly and-12 hourly satellite-based measurements for the remaining 23 hours in the day.

The dataset provides a very accurate (±14%) rainfall estimate for the tropics, with increasing error in the data away from the tropics. Error in the dataset becomes significant >50^{o} N/S of equator (Nicolson, et al. 2003). Due to most rainfall in Africa occurring within the tropics, this dataset is an appropriate dataset to use for Africa, and it is widely used by work in the tropics (Taylor, *pers. comm.*).

BGS has the internal capacity to generate ASCII files as required by the ZOODRM recharge model, from the TRMM data files available for download. Fortran files are also provided from the data source for standard data transformations. Multiple algorithms (i.e. datasets) of the TRMM data are available, but the most appropriate for daily hydrological modelling is the **TRMM 3B 42** dataset (Taylor, *pers. comm.*).

The **EUMETSAT** and **NOAA** datasets (e.g. Symeonakis, et al. 2009) are also found to model rainfall in the tropics accurately, and are appropriate for Africa. The NOAA rainfall data are the most widely used RS data for Africa within published work, and the data are available from various portals.

3.1.3 Forced RS rainfall data for Africa

WATCH WFD (CRU/GPCC derived) – WATCH Forcing Data

The WATCH WFD dataset – is a forced dataset which has been generated as a dataset ready to use for hydrological modelling, alongside temperature, soil moisture and wind data (Weedon, et al. 2010). The rainfall dataset is available in two forms – one derived from precipitation data from the Climate Research Unit (CRU) and the other derived from the Global Precipitation Climatology Centre (GPCC) precipitation dataset. These gridded RS precipitation datasets are forced to observed/gauged rainfall data – thereby increasing the accuracy of the data – to generate the WATCH WFD dataset. Both the GPCC and CRU datasets significantly overestimate rainfall in the ITCZ regions (i.e. majority of rainfall in Africa) (Weedin, *pers. comm.*). The accuracy of the WATCH WFD in ITCZ regions in Africa, therefore, depends upon the number of gauging stations available to force (i.e. correct) the modelled rainfall data – limited in some parts of Africa.

In comparison to the TRMM or NOAA datasets, which are available from approximately 2000 to the present day, the WATCH WFD data are available for the 1980s and 1990s. The WATCH WFD data may, therefore, be more appropriate than the TRMM/NOAA data depending on the time period being modelled, and the time period for which validation data (i.e. gauged river discharge data) are held.

Daily and 3-hourly WATCH WFD data are freely downloadable to BGS (Net CDF files), once a letter of agreement is obtained from the project administrator – Tanya Warnaars (CEH) at: twarnaars@ceh.ac.uk.

3.1.4 Observed rainfall data for Africa

There is an extensive rainfall gauge network in Africa which relays data to the World Meterological Organisation (WMO) from around 760 telecommunication stations (GTS) in Africa (Symeonakis, et al. 2009). This is the only readily available gauged rainfall data for Africa at a continental-scale.

In some regions national, or regional data (e.g. the Somalia Dekadal Rainfall Update project) is available for calibration of hydrological models in specific parts of Africa.

3.1.5 Summary

The largest choice in hydrological data for large-scale hydrological modelling in Africa lies in precipitation data. There is no one ‘perfect’ dataset; all have some disadvantages and advantages. The TRMM dataset is considered the most appropriate for Africa if modelling the last decade (2000-present day). For pre-2000 modelling work, the best available precipitation datasets for Africa are the NOAA data or, WATCH WFD.

Data type	Source	Applicability of data	
		Advantages	Disadvantages
Rainfall data – remotely sensed			
NOAA rainfall data Daily 0.5° spatial resolution Units (mm/day)	NOAA – disseminated from FEWS NET Africa Data dissemination network (US Aid) http://earlywarning.usgs.gov/fews/africa/index.php http://earlywarning.usgs.gov/fews/africa/web/dwndailyrfe.php and also available from NOAA Climate Prediction Centre (CPC) web portal: http://www.cpc.ncep.noaa.gov/products/fews/data.html	• Complete daily rainfall data set for 2003-2005. Data also now available for 2006-2011. • Until TRMM data – this is the most commonly used dataset used in published work for Africa. • Africa rainfall estimate is a merged product derived from 6 hourly rainfall estimates from Special Sensor Microwave/Imager (SSM/I) on defence satellites and 12 hourly rainfall estimates from Advanced Microwave Sounding Unit (AMSU) on NOAA satellite. • Used successfully in groundwater modelling in BGS before, and is used successfully elsewhere (e.g. IRI). • Available as a continent tile, or country tile (http://www.fao.org/giews/english/ierf/index.htm) • Accurate estimation of rainfall in tropics	• Incomplete data - significant number of days missing in any one month in years beyond 2006. • Downloaded as sequential binary .BIL files that contain no header (coordinate) information. ACSII files have to generated using spatial reference of data supplied by FEWS NET source. • Projection of data ALBERS Conical Equal-area projection – unusual but easily re-projected.
TRMM rainfall rate 3 hourly and daily 0.25° spatial resolution Units (mm/day and mm/hour)	TRMM – NASA-JAXA data http://trmm.gsfc.nasa.gov/ 3B42 algorithm data http://trmm.gsfc.nasa.gov/3b31.html ftp://trmmopen.gsfc.nasa.gov/pub/merged/mergeIRMicro/	• The daily rainfall estimate is derived from satellite-based precipitation radar (PR) measurement which accurately measures rainfall once every 24 hours, and other less accurate satellite datasets for the remaining 23 hours in a day. • High spatial resolution (0.25°) • Very accurate satellite-based estimate of tropical rainfall (matches observed data in Africa to within 10-14%). • Downloaded as binary .BIN files. These can be easily converted to .ASCII files for modelling, using fortran programs provided ftp://trmmopen.gsfc.nasa.gov/pub/merged/software/. • Accurate estimation of rainfall in tropics –	• Very inaccurate (100-200% error) at higher latitudes – but little rainfall in Africa away from tropics.

		appropriate for Africa.	
SEVIRI 0.25° spatial resolution Units (mm/hr)	EUMETSAT data – SEVIRI sensor aboard Meteosat-8 and -9 satellites	• Rainfall rate measured by SEVIRI imager • Able to detect seasonal and synoptic variations in rainfall in West Africa accurately from daytime observations (Wolters, et al. 2010) http://www.hydrol-earth-syst-sci-discuss.net/7/6351/2010/hessd-7-6351-2010.pdf	
Forced RS rainfall data			
WATCH – WFD CRU rainfall rate. 3 hourly 0.5° spatial resolution Units ($kg/m^2/s$)	WATCH ftp site: ftp://ftp.iiasa.ac.at Derived from Climate Research Unit (CRU) gridded station observations (multiple variables) www.cru.uea.ac.uk/~timm/grid/CRU_TS_2_1.html http://www.cru.uea.ac.uk/cru/data/	• A 'Water Forcing Dataset' (WFD) designed and collated specifically for use in hydrological modelling of 0.5° resolution. • Data available for 1980-1990s – i.e. when most river discharge data is available to calibrate the rainfall data and any hydrological modelling.	• *Significant error in ITCZ region*, where extremely high rainfall is modelled (based on comparison or GPCC data to ECMWF ERA-40 global analysis of atmosphere from 1957-2002). • *Data is corrected – i.e. forced – in 3 steps:* a. Correction of wet days – days of lowest rainfall in a month in ITCZ region, are designated 'dry days'. In ITCZ regions, a log-normal distribution of rainfall is clipped to locally-based threshold levels. b. Correction for rainfall totals – using GPCC version 4 product and GPCC gridded gauge data (both 0.5° resolution). c. Correction of 'catch ratios' – correction of some underestimation of average monthly rainfall measured at gauges from some catchments, when forcing WATCH data. • Availability of gauge data, and therefore accuracy of correction and the dataset, variable in ITCZ regions of Africa.
WATCH – WFD GPCC rainfall rate 3 hourly	WATCH ftp site: ftp://ftp.iiasa.ac.at	• Same as above.	• Same as above.

0.5° spatial resolution Units (kg/m²/s)	Derived from Global Precipitation Climatology Centre (GPCC) gridded station observations (multiple variables) ftp://ftp-anon.dwd.de/pub/data/gpcc/html/gpcc_normals_download.html		• The GPCC monthly gridded precipitation totals from rain-gauge observations are used to correct WFD modelled dataset, instead of the CRU gridded observation rainfall data as above. • Little difference in accuracy of GPCC or CRU based WATCH rainfall datasets for Africa (Weedin, *pers comm.*).
Rainfall data – gauged datasets			
WMO records – GTS/NCDC stations Daily observed data 760 stations Units (mm/day)	WMO Global Telecommunication System meteorological data from NCDC stations http://www.wmo.int/pages/prog/www/TEM/GTS/index_en.html and http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCDC/.DAILY/ Also distributed by GCOS http://www.wmo.int/pages/prog/gcos/index.php?name=ObservingSystemsandData	• Only readily available observed daily rainfall data • Approximately 760 GTS stations in Africa; 580 of these are within sub-Saharan Africa • Near real time data	• Total number of rain-gauge reports varies daily – on average on 550 data stations report data each month. • Generally <45% stations in sub-Saharan are active – data bias therefore to western and southern Africa.
Somalia Dekadal Rainfall Update (SWALIM) Manual and automatic gauged rainfall data	SWALIM – FAO, UNICEF project http://www.faoswalim.org/ Data download: http://www.faoswalim.org/subsites/climate/index.php		• Data for Somali only and not Africa wide

Table 1 – Types and applicability of rainfall data sources in Africa.

3.2 EVAPOTRANSPIRATION DATA

There are four main RS datasets of daily actual evapotranspiration (ET) and potential evapotranspiration (PET) data which cover Africa (Table 2):

- NOAA – PET data
- FAO AQUASTAT – ET data
- MODIS (MOD16) – ET and PET data
- NASA global gridded land surface PET (now superseded by MOD16 dataset)

The PET/ET datasets available are all global datasets, in which ET or PET values are modelled from satellite sensor measurements of humidity, atmospheric moisture, atmospheric pressure, temperate, using the Penman-Monteith equation.

Limited validation of the datasets is possible, as ET cannot be monitored or 'gauged' in the same way as rainfall. Instead, the datasets are often validated to average values of ET/PET according to land cover and climate within regions, rather than by detailed point calibration. It is difficult to assess which is the most accurate and precise dataset for Africa.

Some of the datasets are, however, better for hydrological modelling than others.

- The **NASA MODIS** (**MOD16) PET and ET dataset** – is the highest resolution ET/PET data available. The dataset has a spatial resolution of 0.05^o and a temporal resolution of mm/day (based on 8 daily-data). Validation of the data by Mu et al. (2007) found the dataset to compare well to observed data available globally.
- The **NOAA PET** dataset – provides daily PET data from 2003-present day, based on 6 hourly satellite measurements of humidity, temperature, solar radiation, wind speed, etc. The spatial resolution of the data (0.1^o) is, however, coarse relative to the MOD16 data and available rainfall data. The dataset has been used successfully before by BGS in large-scale hydrological modelling work in Africa – e.g. Nile recharge model.
- The **NASA global gridded land surface PET** data – is of a relatively coarse resolution, both spatially (0.5^o) and temporally (mm/month) (Mintz and Walker, 1993). The data are still available for download, but have now been superseded by the NASA MODIS dataset, which is of a much higher spatial and temporal resolution.
- The **FAO AQUASTAT** data – is the only dataset to provide long-term mean monthly values of ET (based on 1961-1990). The dataset is, therefore, less suitable for hydrological models analysing beyond 1990, as significant land-use changes (e.g. along the Nile river) and rainfall changes having occurred in some regions of Africa from the 1961-1990 period.

All of the datasets provide PET or ET data for Africa; however, the most recent datasets – MODIS and NOAA – are most suitable for future hydrological modelling work, based on the accuracy, and temporal and spatial resolution, of the datasets.

Data type	Source	Applicability of data	
		Advantages	Disadvantages
Potential Evapotranspiration (PET) data – remotely sensed			
NOAA PET data Daily/Monthly/Yearly 1° spatial resolution Precision 0.01mm/day Units (mm/day)	NOAA – disseminated from FEWS NET Africa Data dissemination network (US Aid) http://earlywarning.usgs.gov/fews/africa/index.php Data download: http://earlywarning.usgs.gov/fews/global/web/dwnglobalpet.php	• Complete daily PET data set. Data available for 2003-2011. • Daily PET estimate is derived from 6 hourly air temperature, atmospheric pressure, wind speed, relative humidity and solar radiation fields measured from the Global Data Assimilation System (GDAS). Daily PET is calculated on spatial basis using Penman-Monteith equation. • Used successfully in groundwater modelling in BGS before, and is used successfully elsewhere (e.g. IRI). • Precision of PET 0.01 mm/day	• Downloaded as sequential binary .BIL files that contain no header (coordinate) information. ACSII files have to generated using spatial reference of data supplied by source. • Spatial resolution is coarse, relative to available rainfall data.
FAO ET data Long-term mean monthly data 0.5° spatial resolution Units (mm/month)	FAO AQUASTAT ET data based on CRU CLv1.0 Global Climate Dataset http://www.cru.uea.ac.uk/~timm/grid/CRU_CL_1_0.html and http://www.cru.uea.ac.uk/cru/data/	• Spatial resolution of data same as that of available rainfall data. • Based on long-term average data. • Used within FAO global agro-ecological assessment studies • ET data calculated according to Penman-Monteith equation	• Data available only for 1961-1990. • Long-term mean monthly values only (based on period 1961-1990). •
MODIS (MOD16) ET and PET 8 day/yearly 0.05° spatial resolution Units (mm/day)	NASA data ET calculated from MODIS land products (vegetation induces and surface temperature) using algorithms. Global grids. http://secure.ntsg.umt.edu/projects/index.php/ID/9a3dae27/fuseaction/projects.detail.htm	• High spatial and temporal resolution. • Used widely in other published work • Compares well to observed data - domparison of MODIS data to observed ET data by Mu et al. 2007, found slight (but not significant) overestimate of ET in 38% of MODIS data	• Not available from website – but freely available for non-commercial research purposes. http://modis.gsfc.nasa.gov/data/dataprod/index.php
NASA global land surface PET 5° spatial resolution Units (mm/month)	NASA global gridded data based on observed air surface temperatures using Thornthwaite regression equations (ET is derived from empirical regression against	• Dataset based on 8565 station values of land surface air temperature globally. • Correction adjustments to dataset away from high latitudes is small (In very large river catchments	• Coarse spatial and temporal data resolution. • Density of air surface temperate monitoring in points in Africa unknown – possibly low (Walker, *pers. comm.*.).

	the observed daily mean temperature) http://gcmd.nasa.gov/records/GCMD_MINTZ_WALKER_SOIL_AND_EVAPO.html	the annual mean ET calculated is adjusted so it agrees with the annual integrated precipitation over the catchment, minus the annual river discharge from the catchment.).	• PET Thornthwaite regression equation is based on PE measured in North America.
ET data – gauged datasets			
No observed or gauged data readily available for Africa.			

Table 2 – Types and applicability of evaporation and evapotranspiration datasets in Africa.

3.3 LAND COVER DATA

There are only two readily available datasets of land cover data for Africa (Table 3):

- USGS Global Land Cover (GLCC) data
- NASA MODIS (MOD11) land cover and land cover change data

Both land cover datasets are derived from thematic maps of land use, ecosystems, surface temperature, etc, and both are validated to ground data. The datasets each use 17 land-use categories, following the International Geosphere Biosphere Programme (IGBP) classification, and are equally appropriate for use in hydrological modelling work in Africa at a continental-scale.

- The **USGS Global Land Cover (GLCC)** data – maps 17 land categories across Africa at 0.5^{o} spatial resolution, based on the data from the AVHRR satellite sensor, as well as land surface thematic maps. The dataset has been used successfully before by BGS in large-scale hydrological modelling work in Africa – e.g. Nile recharge model.

- The **NASA MODIS** (**MOD11) land cover and land cover change** dataset – maps 17 land categories across Africa at 0.5^{o} spatial resolution, based on the land surface data from MODIS dataset. Also maps yearly changes – although update frequency/lag-time between observations and updates, is unknown.

Most other land use datasets available from a web search, are sourced from the two listed above – e.g. the Global Mapping Project (GLCNMO data product) http://www.iscgm.org/, and the Global Land cover facility – UMD land cover classification (http://www.landcover.org/data/landcover/data.shtml) are based on the GLCC dataset.

<table>
<tr><td rowspan="2">Data type</td><td rowspan="2">Source</td><td colspan="2">Applicability of data</td></tr>
<tr><td>Advantages</td><td>Disadvantages</td></tr>
<tr><td colspan="4">Potential Evapotranspiration (PET) data – remotely sensed</td></tr>
<tr><td>USGS Global Land Cover Data
0.5° spatial resolution
Global/Continental tile
17 land use categories</td><td>USGS GLCC dataset
http://edc2.usgs.gov/glcc/af_int.php

Global land cover facility
http://www.landcover.org/data/landcover/data.shtml</td><td>• Mapped land cover category are based on 0.5o AVHRR data from April 1992 to March 1993, as well as derived thematic maps for each continent (e.g. ecosystems map, IGBP cover classification, biosphere models)
• Used successfully in groundwater modelling in BGS before, and is used successfully elsewhere (e.g. IRI).</td><td>• Data based on AVHRR observations from 1992-1993: 18 years old. Appropriate for general trends, rather than specific small-scale data.</td></tr>
<tr><td>MODIS (MOD11) land cover and land cover change data
Global coverage – yearly changes
0.5° spatial resolution
17 land use catagories</td><td>NASA data
ET calculated from MODIS land products (vegetation induces and surface temperature) using algorithms. Global grids.
http://modis.gsfc.nasa.gov/data/dataprod/index.php
and
https://lpdaac.usgs.gov/lpdaac/products/modis_products_table/land_cover/yearly_l3_global_1km2/mod12q1</td><td>• High spatial and temporal resolution.
• Good validation to ground truth data globally and changes.
• Land use categories as defined by International Geosphere Biosphere Programme (IGBP)</td><td>• Update frequency unknown</td></tr>
<tr><td colspan="4">Observed datasets</td></tr>
<tr><td colspan="4">No observed continental-scale data readily available for Africa land cover.</td></tr>
</table>

Table 3 – Types and applicability of land cover data for Africa.

3.4 DIGITAL ELEVATION DATA

Digital elevation data are required in land surface hydrological modelling to route water through catchments accurately according to topography.

There are four main Digital Elevation Models (DEM) which cover Africa (Table 2):

- ACE2 (European Space Agency)
- SRTM (NASA)
- GTOPO30 (USGS)
- ESRI global grid (derived from merging the USGS GTOPO30 and ETOPO2 datasets)

DEMs provide global modelled digital elevation data, typically to a high spatial resolution (0.001-0.5^{o}) and are accurate to within 20 m. The DEMs are all well validated to spot topographic data.

All four DEMs are widely used; however, some are more accurate than others.

- The **ACE2** DEM – has been produced by merging the latest SRTM DEM (the most accurate and highest resolution DEM) with accurate satellite altimetry data. This has enabled void areas within the SRTM data to be filled by interpolation and using the altimetry data. The ACE2 data is, therefore, regarded as the most accurate DEM dataset currently available (typically accurate to ±6 m) and it is most accurate within latitudes ±60^{o}N/S of the equator – i.e. within Africa.

- The **SRTM** DEM – provides the most accurate DEM without correction to altimetry data. The SRTM has a high spatial resolution (0.001-0.01^{o}), and elevation data is accurate typically to ±8 m. The DEM is most accurate within latitudes ±60^{o}N/S of the equator – i.e. within Africa.

- The **GTOPO30** DEM –is of a coarser spatial resolution than the ACE2 or SRTM DEMs. It was completed in 1996, and so it is also one of the older DEMs now available. The GTOPO30 DEM was derived from several sources of topographic data (e.g. Digital Terrain Elevation data). Within Africa, particularly in sub-Saharan Africa, the main source of data is the Digital Chart of the World, which provides only basic elevation and geographic data. The GTOPO30 is, therefore, less accurate than the ACE2 or SRTM DEMs. The GTOPO30 DEM is also split across 4 separate tiles for Africa making it awkward to use. The four tiles have been merged to produce a single DEM tile for Africa – the **USGS Streamless dataset –** but the merged tile is of a much coarser spatial resolution.
- The **ERSI** DEM grid – was produced by merging the GTOPO30 (elevation data) and ETOPO2 (geographic and topographic data) datasets. It is, therefore, less accurate than the ACE2 and SRTM datasets, and of a coarser spatial resolution.

The ACE2 and SRTM undoubtedly provide the most accurate digital elevation data. The resolution of both the DEMs is likely to be too high for the needs of most continental-scale hydrological modelling, however, if the DEMs are re-sampled to produce a coarser resolution DEM – the re-sampled grid will be generated from the most accurate data available if these DEMs are used.

Corruption of DEM data within ArcGIS

Within previous large-scale modelling work in BGS, all readily downloaded DEMs have been found to corrupt when manipulated within Imaging Software or ArcGIS (e.g. if the DEM is re-projected or re-sampled). Issues with all the DEMs have been experienced by users across BGS in ArcGIS, except the ERSI gridded data, which are in a format directly compatible with ArcMap.

Corruption occurs when ArcMap or Imaging Software can not differentiate between the pixel values and the Z (elevation) values of the DEM files, giving a total grid error. Within ArcGIS, the pixel values often relate to the colour value of the pixel and not the elevation data. To avoid the corruption, USGS advise that DEMs must be manipulated within ArcInfo, rather than ArcMap or Arc Catalogue.

Derived elevation datasets

Elevation data from Digital Elevation Models (DEMs) is often re-sampled to produce aspect maps for hydrological models. This reduces the complexity and resolution of DEM data to a manageable level for hydrological models.

- ***GTOPO30 Hydro 1K*** *- Aspect data*

  The Hydro 1K dataset is produced by the USGS specifically for hydrological modelling work and has been derived from the GTOPO30 DEM. The Hydro 1K dataset includes gridded: elevation, aspect, flow accumulation, drainage basins, flow direction, compound topographic index, slope, and stream data for Africa, at continental scale. The grids are all produced at 0.5^{o} spatial resolution for Africa and as a single tile.

  Use of the Hydro 1K grids avoids problems of corruption of the DEM data, and the Hydro 1K dataset provides a useful resource for hydrological modelling work. However, depending on the accuracy and spatial resolution of the DEM data required in the hydrological model, it might still be better to generate the aspect grids from the ACE2 or SRTM DEMs.

Data type	Source	Applicability of data	
		Advantages	Disadvantages
Potential Evapotranspiration (PET) data – remotely sensed			
ACE2 Accuracy ±16m (typically ±6m) 0.001° spatial resolution	ESA-EAPRSE data product http://tethys.eaprs.cse.dmu.ac.uk/ACE2/shared/overview Data already held by BGS on disc from De Montfort University	• Dataset produced by merging SRTM 3 data, with satellite rader altimeter data. SRTM void areas are filled by interpolation and/or altimeter data. Dataset better validated than SRTM DEM. • Most accurate within ±60oN/S of equator – therefore appropriate data for Africa. • Accuracy typically ±6 m. • Very high spatial resolution – also available in lower resolutions (up to 0.01°). • Presented as being one of the most accurate DEM available at ESA conference 2009.	• Large data files due to high spatial resolution of data.
SRTM DEM Accuracy ±20m (typically ±8m) 0.001-0.01° spatial resolution	NASA data Version 2 – most recent available. http://www2.jpl.nasa.gov/srtm/ Africa data is downloadable from: http://dds.cr.usgs.gov/srtm/version2_1/SRTM3/Africa/ and http://dds.cr.usgs.gov/srtm/version2_1/SRTM30/	• High spatial resolution. • SRTM version 2, exhibits well-defined water bodies and coastlines well, and individual pixel errors have been smoothed out. • Most accurate within ±60oN/S of equator – therefore appropriate data for Africa. • Dataset completed in 2000.	• Void areas (missing data) still present. • Validation less rigorous than ACE2 DEM. • Single continent data tile not available for Africa – multiple data tiles therefore have to downloaded indivdidually to construct the DEM for Africa. • Large data files due to high spatial resolution of data. • DEM data can corrupt within ArcGIS (total grid error).
GTOPO30 DEM 0.01° spatial resolution	USGS EROS dataset http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30_info	• DEM derived from several raster and vector sources of topographic information. • Resolution sufficient for continental scale modelling	• In Africa the topographic data sources are: Digital Terrain Elevation data, and for much of sub-Saharan Africa, the Digital Chart of the World (basic geographical information). The DEM is therefore based on less accurate data for Africa than more recent SRTM and ACE2 DEMs. • GTOPO30 DEM completed in 1996 – now among the older DEM datasets available. • Single continent data tile not available for Africa –DEM spilt over 4 data-tiles for

			Africa.
Hydro 1K 0.5° spatial resolution	USGS EROS dataset, derived from GTOPO30 DEM http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30_info	• Dataset derived from GTOPO30 DEM, and includes gridded: elevation, aspect, flow accumulation, drainage basins, flow direction, compound topographic index, slope and stream data for Africa at continental scale. • All available at 0.5° spatial resolution for Africa - appropriate scale for any continental-scale hydrological modelling.	• Based on GTOPO30 DEM which is have been superseded by other DEMs now in terms of accuracy.
USGS Streamless DEM 0.02° spatial resolution	A resampled dataset of GTOPO30 data, to produce a single global DEM tile. Available from USGS streamless server: http://seamless.usgs.gov/	• DEM data available in single tile for Africa. • Reduced file size.	• Half the spatial resolution of GTOPO30 DEM. • Same disadvantages as listed from GTOPO30
ESRI DEM 0.01 and 0.05° spatial resolution	Already held by BGS: S:\GIStore\ESRIdata\ArcImageData\world\elevation	• Adequate resolution for continental-scale hydrological modelling • Both GTOPO30 and ETOPO2 data are provided in native ESRI grid format – therefore no issues with DEM corruption when used in ArcGIS.	• Not based on most accurate DEM datasets • Relatively coarse resolution DEM

Table 4 – Types and applicability of digital elevation data, and derived products for Africa.

3.5 RIVER DISCHARGE DATA

There are three main datasets of gauged river discharge datasets available for Africa at a continent-scale (Table 5):

- GRDC – Global Run-off Data Centre
- WHYCOS – World Hydrological Cycle Observing System
- FRIEND – Flow Regimes from International Experimental and Network Data

In general there is a scarcity of river discharge data in Africa, or considerable difficulties in obtaining modern discharge data, particularly within the large transboundary basins (e.g. the Nile) where allocation of water resources is contentious and river discharge data sensitive. The scarcity of readily available modern discharge for many large rivers within Africa makes validation of groundwater flow and recharge models difficult.

The most comprehensive gauged river discharge dataset for Africa, at a continent scale, is that of the GRDC. It is most easily downloaded from the SAGE portal (http://www.sage.wisc.edu/riverdata/).

- The **GRDC** dataset – is the most comprehensive river discharge dataset available for Africa at continent-scale. Mean monthly discharge data are available for 44 African countries over 931 gauging stations. The dataset includes historic and modern data for all major rivers and tributaries in Africa. The dataset is updated annually – although modern discharge data are not available for many rivers in Africa from the mid-1990s. In cases where modern gauge data are unavailable, GRDC provide long-term average monthly discharge data. The WATCH project was able to obtain privileged access to daily gauged data for some regions in Africa from the GRDC – contact the project administrator – Tanya Warnaars (CEH) at: twarnaars@ceh.ac.uk.
- The **SAGE** dataset – provides the same RivDis (version 2) data as provided by the GRDC, but the data can be viewed and downloaded from SAGE in a much clearer and user-friendly format. Annual hydrographs and tabular data are available online for all gauging stations in Africa from a satellite map. The SAGE portal has been successfully used to download the GRDC RivDis data for other large-scale BGS modelling work in Africa, and is widely used by CEH.
- **WHYCOS** data – provides modern river discharge data for West Africa and SADC regions in Africa. Monitoring of river discharge in Niger and Volta basins is due to begin under WHYCOS in the next 2-3 years. WHYCOS is a WMO initiative. Limited data is available to download from the WHYCOS website.
- **FRIEND** – the FRIEND initiative encompasses all of Africa with exception of the Congo. The FRIEND initiative is aimed at improved river basin and catchment management, and not focused solely to collating and disseminating river discharge data. FRIEND does not, therefore, provide as comprehensive a river discharge dataset as GRDC or SAGE, and the data are not available in as systematic format. However, FRIEND may provide data which fills voids in the GRDC dataset.

Data type	Source	Applicability of data	
		Advantages	Disadvantages
River discharge data, Africa – gauged data			
GRDC gauge data (RIVDIS v2) Mean monthly river discharges Units (m^3/s)	GRDC dataset http://www.bafg.de/cln_007/nn_294172/GRDC/EN/Home/homepage__node.html?__nnn=true GRDC RivDis data also disseminated through wider **GCOS (**Global Climate Observing System) portal http://www.wmo.int/pages/prog/gcos/index.php?name=ObservingSystemsandData	• Main source of readily available, current and historic, discharge data for Africa. • Monthly discharge data from 44 African countries, over 931 gauging stations. • Most recent discharge data updated online annually.	• Often only long-term average monthly river discharge data available from many Africa rivers –absence of monthly discharge data since mid-1990s. • Incomplete discharge records (months and years of data often missing) • Absence of recent discharge data (2000-present) in some major rivers (e.g. Nile).
SAGE Mean monthly river discharges Units (m^3/s)	SAGE dataset http://www.sage.wisc.edu/riverdata/ Based on RivDis2.0 discharge data (as GRDC data)	• Easier download and viewing facility than GRDC – but same data.	• Same as GRDC data.
WHYCOS (WMO initiative)	WMO/IRD monitoring initiative http://www.whycos.org/rubrique.php3?id_rubrique=12 SADC-HYCOS data: http://sadchycos.dwaf.gov.za/ Western Africa ACO-HYCOS data: http://aochycos.ird.ne	• Data available for western Africa and SADC region of southern Africa. • Modern river discharge data • Monitoring in Niger and Volta basins due to start within 1-2 years.	• Data only available for western and southern Africa • Limited data available from web server from current monitoring projects.
FRIEND	UNESCO initiative http://typo38.unesco.org/fr/about-ihp/ihp-partners/friend.html	• Operates initiative across most of Africa – excluding DRC.	• Limited discharge data available to download from the web

Table 5 – Types of gauged discharge data available for Africa.

4 Summary

The scarcity of observed hydrological data, or difficulty in obtaining such data, within Africa means largely remotely sensed (RS) datasets must be used to drive and validate large-scale (catchment or continent) hydrological models. The different applicability of RS data in different regions of the world, however, means that the most appropiate datasets must be selected for Africa for accurate hydrological models to be developed.

Accuracy of RS datasets for Africa depends on: 1) the frequency of the orbit of the satellite holding the relevant sensors above Africa and the number and accuracy of sensors used to derive the data; 2) the validity of the processing assumptions made to derive a rainfall rate, or ET rate from the RS measurements, relative to the main hydrological processes in Africa; and 3) the amount of observed data within the region of interest, to validate RS datasets.

The most applicable continental-scale hydrological datasets identified for Africa are:

- Rainfall – TRMM 3B 42 data (for modelling of the last decade). If modelling a time period pre-2000, the NOAA or the WATCH forced datasets are the most appropriate.
- ET and PET – NOAA and MODIS (MOD16)
- Land cover – GLCC and MODIS (MOD11)
- Digital elevation data – ACE2 and SRTM
- Gauged river discharge data – GRDC (via the SAGE portal)

The largest choice in hydrological input data for Africa lies within RS precipitation datasets. The TRMM dataset is considered the most appropriate dataset for Africa if modelling the last decade (2000-present day). For pre-2000 modelling work, the best available precipitation datasets for Africa are the NOAA data or, WATCH WFD.

Gauged discharge data are readily available for most of the major rivers in Africa up until the mid-1990s from GRDC. However, there are significantly less gauged data available from the last decade, making validation of hydrological models using modern RS data, difficult. Novel methods, such as using satellite altimetry data to estimate river discharges, could enable a more robust validation of hydrological models, and this is an area which researchers are already trying to develop (e.g. Phillipa Berry, De Monfort University).

There are no 'perfect' datasets within the identified datasets for large-scale hydrological modelling in Africa. It would be useful, to conduct multiple model runs of a hydrological model, using the different datasets to test the sensitivity of the model to different input datasets.

References

British Geological Survey holds most of the references listed below, and copies may be obtained via the library service subject to copyright legislation (contact libuser@bgs.ac.uk for details). The library catalogue is available at: http://geolib.bgs.ac.uk.

BATES, B., KUNDZEWICZ, K., WU, S. and PALUTIKOF, J (eds) IPCC: Climate change and water, IPCC Working Group II, Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva, 2007, 210 pp.

BONSOR, H. C., MANSOUR, M. M., MACDONALD, A. M., HUGHES, A. G., HIPKIN, R. G. and BEDADA, T. (2010) Interpretation of GRACE data of the Nile Basin using a groundwater recharge model, *Hydrology Earth System Science Discussions, HESSD*, 7; 4501-4533.

BONSOR, H. C., MANSOUR, M. M., MACDONALD, A. M., HUGHES, A. G., HIPKIN, R. G. and BEDADA, T. (2009) Developing a preliminary recharge model of the Nile Basin to help interpret GRACE data, *British Geological Survey,* Open Report, OR/09/018, pp18.

GRINDLEY, J. (1967) The estimation of soil moisture deficits, *Meterological Magazine,* 96; 1137, 97-108.

HUGHES, A. G., MANSOUR, M. M. and ROBINS, N. (2008) Evaluation of distributed recharge in an upland semi-arid karst system: the West Bank Mountain Aquifer, Middle East, *Hydrogeology Journal,* 16; 5, pp845-854.

JACKSON, C. R., HUGHES, A. G., O DOCHARTAIGH, B. E., ROBINS, N. S. and PEACH, D. (2005) Numerical testing of conceptual models of groundwater flow: a case study using the Dumfries Basin Aquifer, *Scottish Journal of Geology,* 41; 1, 51-60.

LANGE, J., GREENBAUM, N., HUSARY, S., GHANEM, M. and SCHICK, A. P. (2003) Runoff generation from successive simulated rainfalls on a rocky, semi-arid Mediterranean hillslope, *Hydrological Processes,* **17;** 279-296.

MACDONALD, A. M., CALOW, R. C., MACDONALD, D. M. J., DARLING, W. G. and Ó DOCHARTAIGH, B. É. (2009) What impact will climate change have on rural water supplies in Africa? *Hydrological Sciences Journal,* 54; 690-703.

MANSOUR, M. M. and HUGHES, A. G. (2004) User's manual for the distributed recharge model ZOODRM, *British Geological Survey,* Internal Report, IR/04/150.

MINTZ, Y. And WALKER, G. K. (1993) Global Fields of Soil Moisture and Land Surface Evapotranspiration Derived from Observed Precipitation and Surface Air Temperature, *Journal of Applied Meteorology,* 32; 1305-1334

MU, Q., HENSCH, F. A., ZHAO, M. and RUNNING, S. W. (2007) Development of a global evapotranspiration algorithm on MODIS and global meteorology data, *Remote Sensing of Environment,* 111; 519-536.

O DOCHARTAIGH, B. E., MACDONALD, A. M., DARLING, W. G., HUGHES, A. G., LI, J. and SHI, L. (2010) Determining groundwater degradation from irrigation in desert-marginal northern China, *Hydrogeology Journal,* 18; 8, 1939-1992.

NICOLSON, S. E., SOME, B., McCULLUM, J., NELKIN, E., KLOTTER, D., BERTE, Y., DIALLO, B. M., GAYE, I., KPABEBA, G., NDIAYE, O., NOUKPOZOUNKOU, J. N., TANA, M. M., THIAM, A., TOURE, A. A. And TRAORE, A. K. (2003) Validation of TRMM and Other Rainfall Estimates with a High-Density Gauge Daatset for West Africa. Part II: Validation of TRMM Rainfall Products, *Journal of Applied Meteorology,* 42; 1356-1368.

PENMAN, H. L. (1943) Natural evaporation from open water, bare soil and grass, *Proceedings of Royal Society London,* Series A, 193; 120-145.

SYMEONAKIS, E., BONIFACIO, R. and DRAKE, N. (2009) A comparison of rainfall estimation techniques for sub-Saharan Africa, *International Journal of Applied Earth Observation and Geoinformation,* 11; 15-26.

WEEDEN, G. P., GOMES, S., VITERBO, P., ÖSTERLE, H., ADAM, J. C., BELLOUIN, N., BOUCHER, O. and BEST, M. (2010) The WATCH Forcing Data 1958-2001: A Meteorological Forcing Dataset for land surface- and hydrological-models, WATCH Technical Report No. 22, pp 41.