Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Hydrology 279 (2003) 224–243
Journal of Hydrology
www.elsevier.com/locate/jhydrol

Hydrologic process simulation of a semiarid, endoreic catchment in Sahelian West Niger. 1. Model-aided data analysis and screening

Christophe Peugeot[a], Bernard Cappelaere[a,*], Baxter E. Vieux[b], Luc Séguis[a], Ana Maia[a]

[a]*UMR HydroSciences, IRD—BP 64501, 34394 Montpellier Cedex 5, France*
[b]*University of Oklahoma, 202 West Boyd Street—CEC334, Norman, OK 73019, USA*

Received 20 August 2002; accepted 6 May 2003

Abstract

The drought in the African Sahel, persisting since 1968, emphasizes the vulnerability of local water resources to climatic variations and land use changes in this semiarid environment. Following the Hapex–Sahel experiment (1992–1994), research aims at understanding and modeling the impact on the hydrological cycle of these changing climatic and environmental conditions over a 1500-km^2 area east of Niamey, Niger. The hydrological landscape consists of a myriad of small endoreic catchments feeding temporary or permanent pools from which water percolates to an unconfined aquifer. As the first step in a local-to-regional scale approach to the water resource recharge in the area, a physically based rainfall–runoff model is developed for a typical, 1.9-km^2 endoreic system named Wankama, which is monitored continuously since 1992. This first of two companion papers describes the key elements of the hydrological landscape, the experimental setup with which the Wankama catchment and pool are monitored, the *r.water.fea* distributed hydrological model and its application to the construction of the Wankama catchment model. Because of the difficulties of long-term field experimentation in the region, not all observed events can be considered equally well recorded; in particular, detection of exogenous water inflow occurrences that may alter Wankama runoff estimates is needed. The paper describes the careful, model-aided data analysis performed in order to select the observation set with which the model can most reliably be tuned and operated. In this rainfall–runoff data analysis phase, the model is used prior to any calibration. Model calibration and verification, including output uncertainty, are the subject of the second paper [J. Hydrol. (2003)]. All steps of analysis and modeling were performed on the 1992–1998 data only, before the 1999–2000 data became available, and are tested a posteriori against the latter.

Keywords: Sahel; Rainfall–runoff; Distributed hydrologic modeling; Outliers; Uncalibrated model; Sample construction; Hortonian runoff

1. Introduction

Local effects of climatic changes impact the hydrological cycle and resulting water resources. Prospects of reduced water supply, crop, and food

* Corresponding author. Fax: +33-467-144-774.
E-mail address: bernard.cappelaere@mpl.ird.fr (B. Cappelaere).

0022-1694/03/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0022-1694(03)00181-1

Nomenclature

C.Pap.	companion paper (Cappelaere et al., 2003)
DEM	digital elevation model
EXO(*)	exogenous inflow to Wankama pool, from pool outside Wankama catchment (proven)
GIN(*)	general inundation of pool–kori system (proven)
GIS	geographical information system
IRD	Institut de Recherche pour le Développement (formerly ORSTOM), France
LOW(*)	event with very low observed-to-simulated runoff ratio
pEXO(*)	probable EXO, rejected from reference data set after statistical and model-aided analyses
pVAL(*)	event accepted in reference data set after statistical and model-aided analyses
UMR	Unité mixte de recherche (CNRS—IRD—Universités de Montpellier 1&2, France)
VAL(*)	validated, doubt-free runoff event (free from any exchange between pools)
(*)	designates categories in the event classification of Sections 4.3–4.5

production have prompted the consideration of the vulnerability of the local hydrological balance to global climatic changes. The current drought in the Sahel, persisting since 1968 and causing recurrent famine, emphasizes the vulnerability of water resources to climatic variations, as well as to environmental and land use changes, in this semiarid region. The hydrological response to climatic changes such as reduced rainfall is complicated by the nature of the terrestrial system, which transforms rainfall into runoff, infiltration, and evapotranspiration. This system itself undergoes concomitant anthropogenic and natural alterations that make it difficult to quantitatively explain the phenomena observed as resulting from any specific causal factors. Besides the complicated nature of the hydrological processes, incomplete or error-prone data hinder our understanding and simulation of these mechanisms. Any better knowledge we may gain about the process of rainfall–runoff transformation in this region is key to predicting the hydrological response to climatic and environmental modifications and to assessing future water resources. The Hapex–Sahel experiment that took place in Niger (Goutorbe et al., 1994) provided an opportunity to study the links between rainfall fluctuations and the overland water cycle in a typical Sahelian region. The hydrologic data collected were used as the basis for several studies published in a special issue of the Journal of Hydrology (vol. 188–189, 1997), which highlighted some specifics of the hydrological cycle in this region of the world. As described by Desconnets et al. (1997), the hydrologic landscape of the Niamey area is a myriad of small watersheds with inward drainage into temporary pools. These endoreic watersheds are the basic hydrological units of the region, and it is thus of prime importance to be able to model their response to the rainfall forcing in a realistic way if one is to assess the impact of deficient rain and accompanying environmental changes on the region's water resources. It is the more so since the pools are the only surface water resource available to people and cattle during the wet season and since they control the recharge of the aquifer (Desconnets et al., 1997; Favreau et al., 2002a) from which drinking water is obtained all year long.

The work presented in these two companion papers (the present paper and Cappelaere et al. (2003), hereafter referred to as *C.Pap.*, cf. Nomenclature) is one step on the way to building a hydrological model over a 1500-km^2 region of the Hapex–Sahel area, to serve as a tool for estimating medium-to-long-term impacts of climatic fluctuations and environmental modifications on water resources. This first step consists in modeling runoff processes in the Wankama watershed (1.9 km^2), typical of this area (Desconnets et al., 1997) and where a significant number of rainfall–runoff events have been recorded since 1992. This first paper reports on the model construction, on the analysis and classification of the available field data, and on the model-aided selection of the most appropriate data set for model tuning. The need for

screening the raw data arises from some heterogeneities in the record due to experimental conditions. The difficulties involved in obtaining good quality data records of ephemeral and episodic floods in semiarid and arid catchments have been underlined by others (Lange et al., 1999), and the location of the study area further complicates long-term experimentation. Careful critical review of data is required if one is to extrapolate the observational information content through catchment modeling. Halfway between the two opposite methodological extremes of either uncalibrated, process-based (Lange et al., 1999) or purely data-driven, 'black-box' models, the approach followed is to seek for a 'best' combination of the two available sources of information about the system, namely the understanding of dominant relevant processes with a faithful description of the physical system on one hand, and the observation (albeit incomplete) of system responses to natural stimuli, on the other hand. This means building a physically based, process-oriented model that is rather tightly constrained by prior physiographic information, i.e. with only limited calibration from carefully validated rainfall–runoff data. Model tuning and verification, including output uncertainty analysis, are the topic of *C.Pap*. The present paper describes the input/output data analysis-and-screening work, which makes use both of statistical techniques and of the as yet-uncalibrated hydrological model. After reviewing some key elements of Sahelian hydrology relevant to this study (Section 2), material found in the literature on similar attempts is briefly discussed and the bases of the *r.water.fea* model and its implementation for the Wankama catchment are presented (Section 3). The model-aided analysis and screening of data are detailed in Section 4. As the data discrimination method was developed and performed on the 1992–1998 data series before the 1999–2000 data became available, the method is then tested against the latter series (in Section 4.5).

2. Study area

2.1. General physiographic description

The 1500-km^2 study area (Fig. 1) is located in southwestern Niger (West Africa), close to the center of Hapex–Sahel's square-degree (2–3°E, 13–14°N). It coincides with the topographic basin of a fossil river valley called the Dantiandou Kori, where a major part of Hapex–Sahel (Goutorbe et al., 1994) and of later experiments (Leduc et al., 2001) have been concentrated. This region lies along

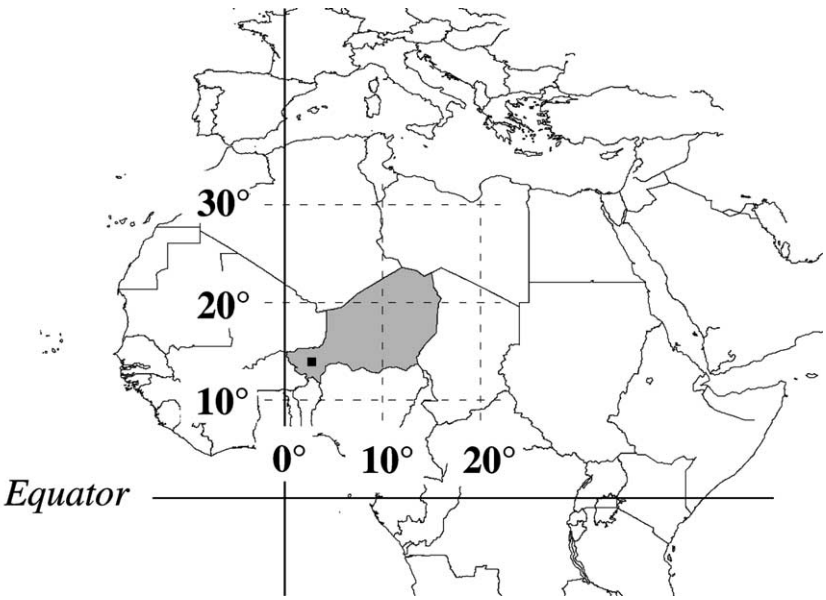

Fig. 1. General location map of study area (■) in Niger (gray).

the southwestern edge of the large Iullemmeden sedimentary basin (Lang et al., 1990). Outcropping sediments, mainly silty sandstones, belong to the Continental Terminal series, from Eocene to Pliocene. The associated unconfined aquifer, 15–50 m below the soil surface, is the only permanent though barely tapped water resource in the area. The present landscape is left over from previous wetter, valley-carving and drier, aeolian-erg Quaternary periods. Moving from high to low elevations (Fig. 2), the following dominant geomorphologic features are encountered: flat, dissected, armored plateaus (relics of the primitive, nearly horizontal laterite plateau), gently sloping (1–3%) gully carved sand plains, and fossil riverbeds named koris in the valley bottoms. The topographic relief between plateaus and valley bottoms is generally around 50 m with elevations ranging from 200 to 270 m amsl. Plateaus, characterized by minute slopes and banded 'tiger-bush' vegetation, may only export water from a narrow outer bare-soil strip (Galle et al., 1999; Valentin and D'Herbès, 1999). Unlike the many gullies (ravines) running down from steep plateau edges, the remnant regional drainage network no longer conveys runoff because of Pleistocene aeolian sand deposits: hillslope runoff is trapped in sand-blocked pools along the koris. Hence, surface runoff is confined to small-scale endoreic catchments in the order of a few square kilometers or less. Halfway along the rather smooth slope between a plateau foot and a kori, a ravine may vanish in a sand-clogged spreading zone and reappear downstream, eventually reaching the collecting pool. Pool water evaporates or percolates to the aquifer, which is the main source of groundwater recharge (Desconnets et al., 1997; Leduc et al., 2001). Besides kori blockages and mid-slope spreading zones, smaller-scale network degradation may occur at other locations in the watersheds, and this, combined with the weakly marked topography, makes catchment delineation a difficult task. Aside from occasional spillover from one pool to its neighbor, there is no runoff at any larger scale. A typical toposequence and the associated hydrological processes are shown in Fig. 2, together with the estimated annual water budget for a small catchment in the area (Peugeot, 1995): around 8% of annual rainfall are transferred to pools through surface runoff and subsequently reach the aquifer, the remaining 92% being evapotranspired back to the atmosphere. This

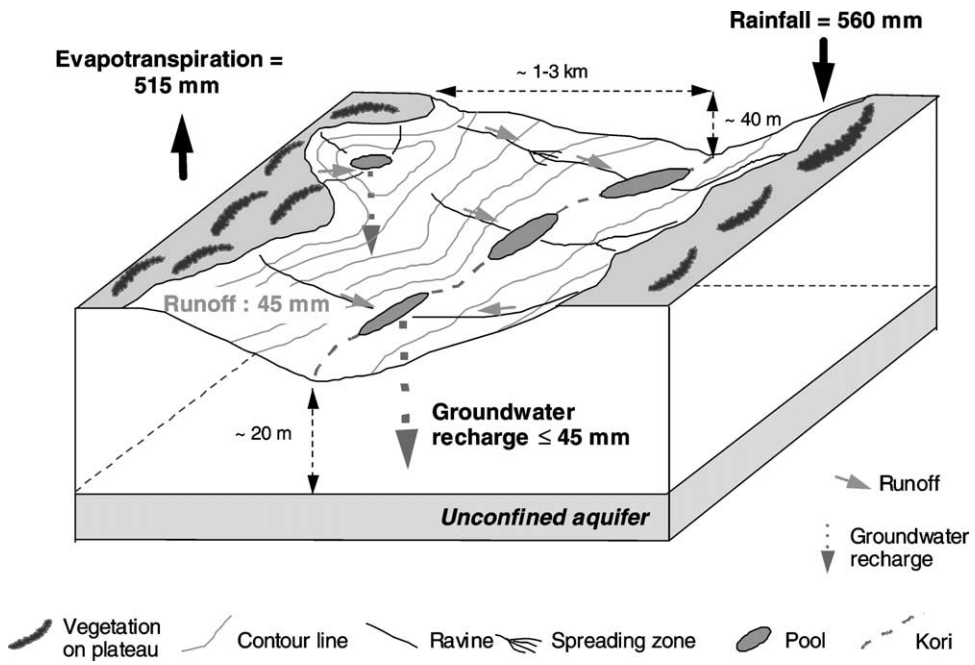

Fig. 2. Typical toposequence and major hydrological processes (after Peugeot, 1995).

estimate of the current annual aquifer recharge rate is consistent with figures published by Leduc et al. (2001) and Favreau et al. (2002a).

Hillslope soils are mainly sandy, with 10–15% of fine particle content (clay and silt) and less than 0.5% of organic carbon (Nagumo, 1992). They are classified as 'weakly leached tropical ferruginous soils' (CPCS, 1967), or 'Sandy Psammentic Haplustalf' (Soil Survey Staff, 1975). They are weakly structured, and highly prone to crusting and erosion (Hoogmoed and Stroosnijder, 1984; Valentin and Bresson, 1992). Extensive rain-fed, hand-cultivated food crops, such as pearl millet, sorghum or cowpea, alternate with fallow areas covered by annual grass and scattered bush savanna. Two mechanisms appear to be of key importance for the transformation of rainfall into runoff in this area. First, the amount of runoff production at the field scale strongly depends on soil surface crusting, as already acknowledged in various arid and semiarid zones (Casenave and Valentin, 1992; Cantón et al., 2001) and confirmed by Peugeot et al. (1997) for this particular area. Since crusting often shows large space variability, a fine-scale description of surface conditions is considered important for hydrological modeling of these small watersheds. The second important mechanism involved is water loss by infiltration along the conveying network, and especially in the intermediate spreading zones. Field measurements by Estèves (1995) showed that infiltration capacity in the ravines is as high as 45 cm/h.

Despite sharply reduced rainfall observed since 1970 (see Section 2.2), the groundwater table has undergone a general and continuous rise, as shown by piezometric records available since 1950 (Leduc et al., 2001). This behavior may be attributed to the hydrological impact of dramatically increased land-clearing over the latest 50 years: Leduc and Loireau (1997) have estimated from aerial photographs that the cropland fraction has grown from 20% in 1950 up to 80% in 1995 in the study area. Reduced evapotranspiration (Bromley et al., 2002) as well as increased runoff production and concentration are equally likely hydrological consequences, but here it is the latter effect that is believed to be primarily responsible for the rise in the water table (Leduc et al., 2001; Favreau et al., 2002b). Such opposite variations in rainfall and runoff have also been reported at other Sahelian locations (Albergel, 1987; Mahé et al., 2000).

2.2. Climate

Rainfall is confined to a single wet season from May–June to September–October. The long-term (1905–1989) average yearly rainfall is 562 mm in Niamey (Le Barbé and Lebel, 1997). The mean annual temperature is 29 °C, with a daily maximum around 45 °C in May. Potential evapotranspiration is about 2500 mm per year and rainfall dominates the water budget only in July and August. The Sahel has been subject to drought for 30 years now. For the Niamey area, Lebel et al. (1997) found that the average yearly rainfall of 645 mm observed over the period 1940–1967 had decreased to an average of 495 mm over the period 1968–1990. Comparing the 1950–1970 and 1970–1990 periods, Le Barbé and Lebel (1997) found the reduction in mean annual rainfall to be in the order of 180–200 mm over the whole Sahel region. More than 90% of the annual rainfall is associated with mesoscale convective systems (D'Amato and Lebel, 1998). At any given location, the annual rainfall is essentially produced by as few as 10–15 events, each a few hours long. The sparse structure of the systems, described by Lebel and Le Barbé (1997), is responsible for intermittency in space and strong spatial gradients: per-event and per-season point-rainfall depths largely depend on the chances that westbound storms travel over a particular location. In the Hapex–Sahel experiment, gradients of seasonal rainfall as high as 150 mm over 6 km (year 1990) or 275 mm over 9 km (year 1992) have been reported (Lebel et al., 1997). Amani and Lebel (1996) identify a typical hyetograph as having a distinctive form: an initial burst followed by an extended period of low intensity rainfall, for a few hours.

2.3. The Wankama catchment and experimental setup

As runoff to pools and subsequent percolation to the unconfined aquifer are the most important components of the hydrological cycle from the standpoint of water resources, the Wankama catchment (1.9 km^2) is being monitored since 1992 to investigate the complex interactions between rainfall, surface flow and infiltration characteristics that

control pool inflows. It is located 70 km east of Niamey, in the 1500-km^2 study area of the Dantiandou Kori basin and in the East-Central Supersite of the Hapex–Sahel Experiment (Goutorbe et al., 1994), and is considered as representative of the endoreic catchments of the area. West to east, the watershed extends from the plateau down to the 5-ha pool located in the kori near Wankama village (Fig. 3). It includes only a very small area east of the pool, and is limited to the north by a tarred road (straight boundary in Fig. 3). The watershed is now intensively cultivated, with 54% fields, 26% fallow and 20% uncultivated, 'natural' land. As commonly observed in the area, the main catchment ravine, originating near the plateau, gets lost at mid-slope into a spreading zone and reappears 350 m further down to eventually reach the pool. This ravine is 1.3–2.3 m wide and 1–1.5 m deep. Visual inspection of the junction area between the spreading zone and the receiving ravine reach indicates that through-flow occurs only very occasionally at this location, with only limited discharge. The lengthwise topographic profile of the kori (Fig. 4) shows that the catchment pool is surrounded by two other pools, the so-called 'North pool' upstream, and the small 'South pool' downstream (the expression 'the pool' with no qualifier will hereafter designate the catchment pool only). All but the South pool can potentially spill over to the neighboring pool(s).

Rainfall is measured since 1991 by a tipping-bucket recording raingauge (SERPE-IESM, Guinel, France; resolution: 0.5 mm) named Mare, located near the pool. At a distance of 2.6 km in the upper part of the catchment, a second, similarly equipped station, named Ouest, was operated in 1993 and 1994. Since storm systems predominantly travel east to west, the upper, Ouest station generally receives rainfall later than the lower, Mare station, with an average lag of 3–4 min. Catchment runoff is measured by recording pool level variations. For practical reasons, it has not been possible to install and maintain a long-term, direct discharge measurement system in the catchment ravine. Besides, a large catchment fraction drains directly to the pool. The water level in the Wankama pool is continuously measured (SERPE-IESM recorder) to the nearest 1-cm at 5 min intervals. Stage is converted into pool water volume through a stage-area relationship derived from topographic surveys performed in 1992 and 1999 (483 points), which only show minor change. Differentiation in time gives the hydrograph of net water inflow to the pool. It follows that event volumes are known with greater precision than discharge, with the former also being of much greater concern in this resource-oriented study. Until 1998, the North pool was only equipped with a staff gauge, for irregularly spaced, manual stage observations, resulting in an incomplete dataset (overflowing of the North pool was initially thought to be a rare occurrence); since 1999 incl. it is also equipped for continuous monitoring. Pool water level below filling depends on intermittent water inflow (essentially runoff from the Wankama catchment, but also, at times, from a larger

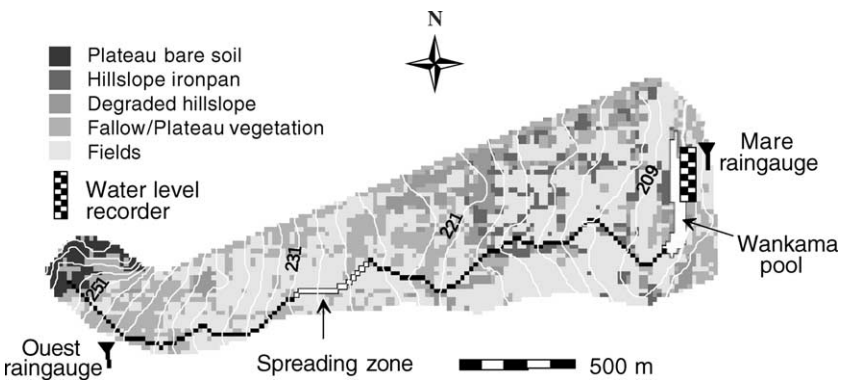

Fig. 3. Wankama catchment: soil surface features, drainage network, 2 m-spaced elevation contours (meters above sea level), and experimental setup.

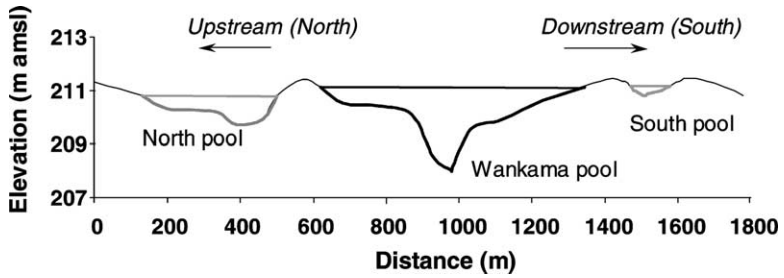

Fig. 4. Lengthwise topographic profile of kori showing Wankama pool, and North and South pools.

catchment due to overflow from the North pool), and on continuous water losses by evaporation and seepage towards the aquifer through the pool bottom. Direct evaporation is negligible during rainfall events and, compared to inflow, seepage may become significant only in the later part of certain events, when the pool level gets sufficiently high. Inflow estimates at the end of each event are therefore corrected for a seepage rate drawn from the pool recession curves. This procedure is used to derive, for each event, the total runoff volume (V, hereafter expressed as a depth in mm after division by the catchment's surface area) and related discharge hydrograph.

3. Model description

3.1. Background

Rainfall–runoff in these small endoreic catchments (in the order of a few km^2) calls for modeling at the event scale, even though our purpose is not to model each individual event as closely as possible but to do so in a more statistically meaningful sense, in line with our water resources perspective. Hereafter are discussed a few of the key issues reported in the literature on building rainfall–runoff models in a semiarid environment. Model structure has important consequences on the ability of a model to reproduce observed behavior. Michaud and Sorooshian (1994) compared three models for 24 severe storms in a semiarid watershed in the western US. A lumped model, a complex distributed model, and a simple distributed model were evaluated as to performance in predicting basin response. The lumped model performed the poorest, the simple and the complex distributed models gave about equal performance. Without calibration, the complex distributed model performed the best of the three. Lumping of rainfall and model parameters for partial area Hortonian runoff resulted in relatively poor performance. Unsuitable model structure or poor representation of soil properties, microtopography, and rainfall variability contribute to bad agreement between model results and field measurements. The spatial variability patterns of infiltration and overland flow parameters, especially hydraulic conductivity and surface roughness, must be accounted for. The effects of their spatial structure on Hortonian catchment runoff were studied by Saghafian et al. (1995) and Séguis et al. (2002), showing a greater impact for small storms. Woolhiser et al. (1996) point out that these effects are all the more important in semiarid regions where partial equilibrium runoff generated from Hortonian mechanisms dominates. The suitability of various simplified formulations of overland flow equations, including the kinematic wave approximation, for a single slope plane under stationary or moving storm conditions, is described by Richardson and Julien (1994). A number of studies have shown the necessity to account for channel processes in hydrological studies of arid or semiarid watersheds (Walters, 1990; Ye et al., 1997; Martin-Vide et al., 1999), with particular importance to be given to a proper representation of channel losses in hydrological models. Beven (1989) observes that distributed deterministic models may be over-parameterized unless prior knowledge or assumptions can be made that constrain the possible choices of parameters. This background favors the use of a physically based, spatially distributed,

process-oriented hydrological model for Wankama runoff simulation, with only limited calibration (i.e. few, constrained degrees of freedom). As shown in Section 3.2, the *r.water.fea* model meets these requirements and is chosen for this study.

3.2. Model principles

The *r.water.fea* model developed by Vieux and Gaur (1994) is a physically based distributed Hortonian runoff model integrated with the *Grass* geographical information system (GIS) software (USACE, 1993). The production and routing functions are fully coupled, allowing simulation of runoff from and run-on to each cell of a raster grid within the catchment. The kinematic-wave, Manning and Green-Ampt equations are solved concurrently using finite-elements in space and an explicit finite-difference scheme in time (Vieux, 2001). The GIS is used to manage the maps of parameters controlling infiltration and routing in overland and channel areas, as well as rainfall which may be input as a spatially lumped time series or as a time series of spatially distributed rainfall rate maps. The GIS also provides the model with two raster maps, derived from a digital elevation model (DEM) using *Grass' r.watershed* command (USACE, 1993), to describe the drainage structure of the discretized catchment, i.e. (i) the flow directions and junctions defining cell connectivity, and (ii) the location of the channeled drainage network, respectively. The connectivity of the n_c grid cells is used to assemble an $n_c \times n_c$ system of equations, which is solved with a time step limited by the Courant–Friedrich–Levy condition for the storm duration and catchment response. Hillslope cells contribute runoff to channel flow cells where it can still infiltrate or be routed downstream by the same, coupled equations. Channel flow differs from overland flow by specific geometrical characteristics representing a spatially variable channel cross-section. Thus, the model is capable of representing coupled runoff production and transfer over complex terrain, land use, and soil conditions. A major new version of the model has been developed in the context of this study to match the specifics of a semiarid, endoreic catchment, and to ease operation in parameter-exploration mode (see *C. Pap.*).

3.3. Model setup

A GIS of physiographic and remote sensing data covering the project area was developed for the specific purpose of hydrological modeling, both at the regional and at the catchment (local) scales, in the *Grass* software environment (Desconnets et al., 1996). The data described hereafter for the Wankama catchment is part of this GIS. Topography, which controls both slope and drainage structure, was obtained from a 1:5000 field survey (580 points) of the Wankama catchment, with elevation in cm to resolve small-scale differences. A *Grass* thin-plate spline interpolation module, *s.surf.tps* (Mitasova and Mitas, 1993), was used to fit a DEM surface to the surveyed point data. A single, 20 m resolution is used for the DEM and all other raster maps involved, for a total of $n_c = 4688$ grid cells over the 1.9-km^2 catchment. This choice arose from the 20 m resolution of the pre-existing surface condition map derived from SPOT imagery (see below), as well as from computational considerations. Resulting elevation contours are shown in Fig. 3 with 2 m spacing. As mentioned above, *Grass*'s *r.watershed* function was used to derive the drainage raster map that defines all cell-to-cell connections along which water is routed downwards, based on principal land surface slope direction. The stream network was satisfactorily delineated using a drained area threshold of 0.1 km^2. Minor differences with field and aerial photography information were manually corrected in the drainage structure. Based on thorough field observations, the channel network was subdivided into six reaches of uniform rectangular cross-sections, with respective widths of 1.3–2.3 m and a Manning roughness coefficient of 0.03, plus a 225 m-wide reach to represent the flat mid-slope spreading-zone. Slope is taken locally at each cell of the channel network from the overall slope raster map of the Wankama watershed.

All topography and geometry related information layers for the channel network and hillslopes have been set once and for all in the model. On the other hand, hydraulic roughness and infiltration data layers are subject to sensitivity analysis and tuning (*C.Pap.*), around prior spatially distributed reference values that will now be described. Hydraulic conductivity and roughness maps are based on

D'Herbès and Valentin's land surface condition map (1997), where surfaces are classified according to crust, vegetation, and land use, as well as to other conditions such as erosion, termites, or presence of rock or gravel armoring, using the scheme of Casenave and Valentin (1992) applied to the 20 m resolution SPOT imagery XS 9-25-99. Fig. 3 shows an extract of the land surface map for the Wankama catchment (surface types have been simplified in the figure from 16 to 5 classes for legibility). The map of saturated hydraulic conductivity (K_s) is obtained by reclassifying the land surface map using the values associated with each class by Casenave and Valentin (1992), displayed in Table 1. Values as low as 3 mm/h are frequent for the Wankama hillslopes, indicating that large amounts of runoff can be generated on a local or point basis due to crusted soils. Channel infiltration is accounted for by mapping appropriate values of hydraulic conductivity within the confines of the stream network. In situ conductivity data are not available for the Wankama ravine, but ring infiltrometer tests in a neighboring, physiographically similar catchment yielded a K_s of 450 mm/h for sandy ravine bottoms (Estèves, 1995). This is therefore used as the prior value for channel conductivity in the sandy reaches of the Wankama ravine (Desconnets et al., 1996). The other Green-Ampt infiltration parameters, namely the effective capillary drive H_f and the porosity, were set uniform over the whole basin to values of 20 cm (representative of the watershed's sandy soils) and 0.27 m^3/m^3 (field measurements), respectively. H_f normally depends on soil moisture, but sensitivity analysis (Séguis et al., 2002) showed that it can reasonably be held constant over the range of moisture contents encountered in this study.

Desconnets et al. (1996) described how estimates for Manning's hydraulic roughness parameter were derived from the land surface classification map. Since only a few of the classified land surface conditions have published hydraulic roughness coefficients, it was necessary to supplement the remaining categories. Four hydrologists who are familiar with the site were asked to rank all land surface classes in terms of hydraulic roughness. Using the published Manning values while maintaining relative rankings, each remaining classification was then assigned a proportionate Manning roughness coefficient. Table 1 summarizes the reference values for both K_s and roughness coefficient, assigned to each land surface type that occurs in the Wankama catchment. These values are used as prior estimates for the model calibration procedure, as described in *C.Pap*. Because of the limitations experienced with the field hydrological observations (Section 2.3), the choice was made to keep the physically based model tightly

Table 1
Hydraulic roughness and conductivity (K_s) values for land surface classifications. (sources: (♣) D'Herbès and Valentin, 1997; (◆) Casenave and Valentin, 1992; (†) Desconnets et al., 1996; (‡) Estèves, 1995; (♠) Peugeot, 1995; (♥) Chow, 1959)

Land surface condition (♣)	Main soil crust (♣)	Herbaceous cover (%) (♣)	K_s (mm/h) (◆)	Manning's hydraulic roughness (†)
Plateau: bare soil	Gravel	0	2	0.02 (♠)
Plateau: dense vegetation	No crust	<10	40	0.25
Hillslope iron pan	Erosion	<15	3	0.05
Degraded hillslope	Erosion	<5	4	0.02 (♠)
Hillslope low density field	Structural	15–25	13	0.15
Hillslope high density field	Structural	25–35	13	0.10
Valley bottom low density field	Erosion	15–25	9	0.17
Valley bottom high density field	Structural	25–50	13	0.12
Old dense shrub fallow	Drying	25–50	13	0.12
Old mid-dense shrub fallow	Drying	50–75	20	0.17
Old sparse shrub fallow	Drying	>75	20	0.20
Mid-old high grass fallow	Drying	50–75	20	0.17
Mid-old low grass fallow	Drying	>50	20	0.15
Recent fallow	Drying	50–75	20	0.12
Channel	No crust	0	450 (‡)	0.03 (♥)

constrained by the prior parameter values, with only limited tuning through a small number of degrees of freedom in the parameterization scheme. Rainfall inputs are managed as time series of spatially uniform rainfall fields, derived from the Mare raingauge (Section 4.1). A final input needed to run the *r.water.fea* model is the initial soil moisture distribution in the watershed at the beginning of a simulated storm event. The saturation index SI defined in Section 4.2 is used to describe the moisture status of the basin, also considered as uniform.

4. Event data analysis and screening

To calibrate and validate the Wankama model (*C.Pap.*), a set of paired rainfall and runoff observations that are representative of catchment behavior is needed. Because of occasional pool overflow and/or concomitant inflow from the North pool, not all recorded rainfall–runoff events may be included in the reference data set. This section is devoted to analysis of available event data, to event selection, and to preparation of the set of rainfall, initial soil moisture, and runoff information necessary for model operation. A special event selection method had to be developed to make up for the largely incomplete North pool level record (pool exchange instances may be missed) over the 1992–1998 period. It is based on a thorough study of the Wankama rainfall–runoff relationship for that period, consisting both in a data-driven statistical analysis (Sections 4.1 and 4.3) and in a process-driven, physically based simulation approach using the uncalibrated model (Section 4.4). The events of 1999–2000, which were obtained subsequent to this development and were fully monitored (continuous log of North pool level), are used to verify the event selection method a posteriori (Section 4.5).

4.1. Rainfall

The Mare gauge recorded 163 storms over the 1992–1998 period and 49 in the 1999–2000 period, while 50 storms were observed at the Ouest station (1993–1994). For events recorded concurrently at the two stations 2.6 km apart, there are some differences in the measured rain depths, as expected from the aforementioned spatial rainfall variability observed in the region (Section 2.2): the largest differences (Mare − Ouest) per event are +21 mm (October 10, 1994) then +10.5 mm on one hand, and −14 mm on the other hand. However, event rain depth statistics are essentially the same for the two stations over the concomitant recording period: average depth of 21 mm and mode of 10–15 mm; the eventwise regression equation is Ouest = 0.9725 × Mare, with a R^2 coefficient of 0.86. The overall largest rain depth observed at Mare is 88.5 mm (65.5 mm at Ouest). The longer Mare rainfall dataset will be used for the modeling study, but possible effects of rainfall variability across the catchment will be investigated and discussed (Section 4.4 hereafter, and Section 3.6 in *C.Pap.*).

The main characteristics of rainfall events with respect to runoff generation were extracted from the 163 Mare 5-min hyetographs (1992–1998), for principal-component analysis (PCA). These storm variables are total rain depth P (mm), maximum 5-min intensity I_{max} (mm/h), plus four variables derived from a plot of intensity i versus the cumulated rainfall depth $p(i)$ produced by all the higher-than-i 5-min intensities. These four variables are: P_{20} (mm) = p(20 mm/h) (i.e.: cumulated above-20 mm/h intensities, hereafter named 'the thresholded rain depth'), and $IP_{0.2}$, $IP_{0.5}$ and $IP_{0.8}$ (mm/h), the 5-min intensities such that $p(IP_{0.2})$, $p(IP_{0.5})$, and $p(IP_{0.8})$ are 20, 50 and 80% of the total event depth P, respectively. Peugeot et al. (1997) showed that P_{20} provides an estimate of the quantity of rainfall that can actually produce runoff in the area, and can be used to efficiently describe the rainfall–runoff relationship. The three variables $IP_{0.2}$, $IP_{0.5}$, and $IP_{0.8}$ are synthetic indicators of the distribution of intensities within the rainstorm. This set of six variables is considered to adequately represent the shape and magnitude of any given individual in this quite homogeneous population of essentially convective rainfall events, particularly with respect to potential runoff generation. Correlation coefficients between these six variables for the 163-storms sample are high, ranging from 0.50 (P versus $IP_{0.8}$) to 0.96 (P versus P_{20}); all other values are in the range 0.61–0.83, except for $IP_{0.2}$ versus $IP_{0.5}$ which is 0.95. The PCA analysis performed on the correlation matrix shows that 79% of the total variability is described by the first axis,

90% by the first plane. Correlations with this first axis are 0.83, 0.92, 0.85, 0.94, 0.94, and 0.83 for P, P_{20}, I_{max}, $IP_{0.2}$, $IP_{0.5}$, and $IP_{0.8}$, respectively. This axis, denoted P_{synth}, is a quite even combination of the six standardized variables, with weights in the range 0.38–0.43. No particular clustering or pattern, nor any outlying individual is apparent on the first principal plane. These results confirm the expected homogeneity of rainfall events in this area, and single out P_{synth} as a highly representative variable for the population of storm hyetographs in the Wankama basin. Although P_{synth} and P_{20} are partly redundant (P_{20} is included in P_{synth}) and correlated variables, both will be used successively for rainfall–runoff analysis in Section 4.3, since on one hand P_{synth} is the single most comprehensive rainfall descriptor, integrating both intensities and depths, while on the other hand P_{20} is simpler and has been used successfully in a similar context (Peugeot et al., 1997).

4.2. Soil moisture

Initial soil moisture at the beginning of a rainstorm event is a factor in the rainfall–runoff transformation and an input required for model operation. Soil moisture was monitored in 1993 using two neutron-probe access tubes located near the pool (Desconnets et al., 1997). However, the representativity of these data is too limited in time and space to fully characterize the moisture status of the catchment, over the study period. As an alternative, a global initial soil saturation ratio (SI) is computed, based on the Antecedent Precipitation Index (API) formulation of Kohler and Linsey (1951):

$$\mathrm{API}_\alpha(i+1) = [\mathrm{API}_\alpha(i) + P(i)]\mathrm{e}^{-\alpha\Delta t(i)} \quad (1)$$

where $\mathrm{API}_\alpha(i)$ is the index value at the beginning of storm i (mm), $P(i)$ is the rain depth of storm i (mm), α is the decay coefficient (day^{-1}), and $\Delta t(i)$ is the time span between the end of storm i and the beginning of storm $i+1$ (day). This index has been used successfully by others in the Sahelian context (Casenave and Valentin, 1992). To translate the API into soil moisture content, its relationship with the water storage of the top-40 cm soil horizon was established on the basis of the soil moisture data from the two neutron-probes in the Wankama watershed and from eight probes distributed in a neighboring catchment (Cuenca et al., 1997; Peugeot et al., 1997). The best fit is obtained with the formulation:

$$\frac{\mathrm{WS}(i) - \mathrm{WS}_0}{\mathrm{WS}_{sat} - \mathrm{WS}_0} = 1 - a^{\mathrm{API}_\alpha(i)} \quad (2)$$

where $\mathrm{WS}(i)$ is the water storage prior to storm i (mm), WS_0 and WS_{sat} are water storage at residual water content and saturation, respectively (mm), and $\mathrm{API}_\alpha(i)$ is defined above (mm). The left-hand side of Eq. (2) represents the fraction of soil moisture saturation at initiation of storm i, denoted $\mathrm{SI}(i)$. The WS_0 and WS_{sat} estimates from in situ top-40 cm moisture data are 12 and 120 mm, respectively. Parameter calibration produces $a = 0.995$ and $\alpha = 0.093\ \mathrm{day}^{-1}$.

4.3. Runoff

Among the 163 storm events observed at the Mare raingauge for the 1992–1998 period, 111 of them produced properly recorded stage graphs of the Wankama pool. Pool overflow periods can be detected by comparing the continuous pool stage records with the elevations of the upstream and downstream sills that separate the pool from the North and South pools (Fig. 4). Manual stage observations in the North pool, when available, inform about possible exogenous inflow to the catchment pool. Nine very large rainfall events ($P = 48$ to 88.5 mm), hereafter referred to as 'GIN'(cf. Nomenclature), were identified as having produced the general inundation of the valley bottom, resulting in the total submersion of the three pools. Spillover from the North pool into the Wankama pool is evidenced for 13 storms, denoted 'EXO' (for exogenous inflow). These 22 floods are considered inadequate for Wankama catchment modeling. For all remaining events, the records testify that the Wankama pool did not overflow. The occasional manual readings of the North pool staff gauge allow to rule out any overspilling from this pool for 5 of them (named 'VAL', for validated), whereas no absolute conclusion can be drawn for the remaining 84 events.

In order to identify other possible instances in which one could suspect North pool overflow into the Wankama pool (meaning potential overestimation of runoff volume), an analysis of pool recharge volumes (hereafter called observed or measured runoff, for

simplicity) against recorded rainfall parameters was undertaken for all 111 events. Fig. 5(a) is the plot of the synthetic storm-event variable P_{synth} versus the 'apparent' runoff coefficient C_r, which is the ratio V/P of measured runoff to rainfall depths. In Fig. 5(b), the thresholded rainfall depth P_{20} is plotted against C_{r20}, defined as V/P_{20}. In both figures, polygonal symbols are used to mark events for which a firm diagnosis has already been delivered above about the occurrence or not of overflow for any of the two pools: 9 filled triangles for the GIN events (general pool submersion), 13 empty triangles for the EXOs (spillover from North pool to Wankama pool), and filled squares for the 5 VAL

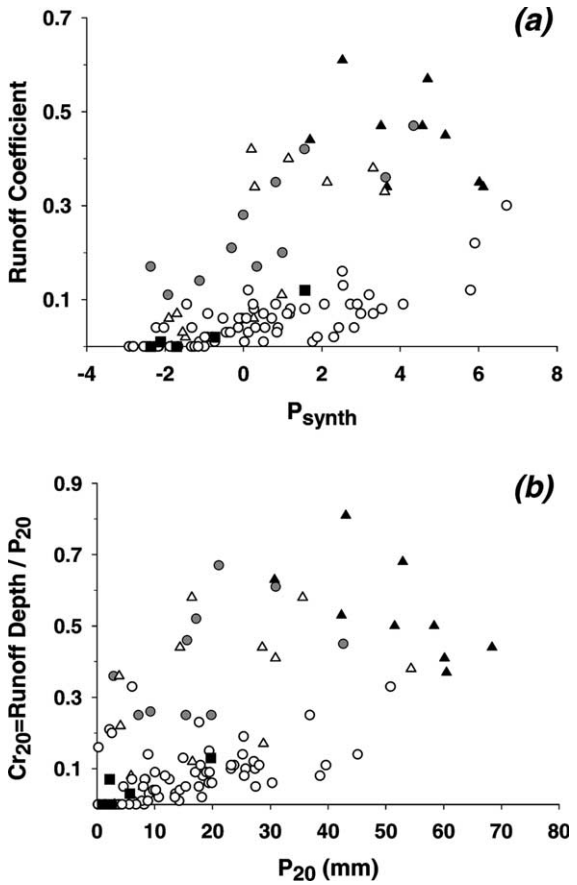

Fig. 5. Rainfall–runoff event data analysis, 1992–1998: (a) runoff coefficient C_r versus synthetic rainfall variable P_{synth}; (b) 'thresholded' runoff coefficient C_{r20} versus 'thresholded' rain depth P_{20}. Symbols: ■, no pool-transfer (VAL); ▲, general pool submersion (GIN); △, inflow to Wankama pool from North pool (EXO); ○, ●, unknown North pool stage.

Table 2
Correspondence between figure symbols and abbreviated names (see Nomenclature) for event classification categories

Squares	VAL, i.e. proven valid (filled: ■, 1992–1998; dotted: ⊡, 1999–2000)
Triangles	GIN and EXO, i.e. proven non-valid (filled: ▲, GIN 1992–1998; empty: △, EXO 1999–2000; dotted: ▽, GIN and EXO 1999–2000)
Circles	pEXO, pVAL and LOW, i.e. events of 1992–1998 classified only after statistical and model-aided analyses, due to unknown north pool stage (filled: ●, pEXO; empty: ○, pVAL and LOW)

events (established as pool-transfer free). Circles are used for all 84 other, as yet-unclassified events; the distinction between circle types (either filled or empty) is ignored until Section 4.4. Table 2 summarizes the correspondence between all figure and text symbols used for the classification; category definitions are recalled in Nomenclature. The various events are organized in very similar patterns in the two graphs. Inclusion of the second rainfall PCA axis (Section 4.1) does not change the distribution patterns significantly. Apparent runoff coefficients C_r (Fig. 5(a)) are high for known sill-overflowing events (GIN and EXO triangles). Among the as yet-unclassified events too (circles), some of the C_r values must be legitimately questioned as being excessively high (>30%) for the type of catchment, in this area, making the circles very dispersed compared with other known watersheds. If the apparent C_r values for all circle-marked events were actual runoff coefficients, the variance for any given range of rainfall (expressed as P or as P_{20}) would be considerably larger than for the similar, though smaller, close-by catchments studied by Peugeot et al. (1997). In Fig. 5(b), it can be seen that a number of circles have C_{r20} values as large as those of the GIN and of the bulk of the EXO events (triangles), i.e. above 30%, markedly higher than the mass of unclassified events. This may indicate exogenous inflow to the pool for those particular circle events, all the more so as these large C_{r20} coefficients occur for generally lower P_{20} levels compared with GIN and EXO events (triangles). In this respect, even the unclassified events (circles) with C_{r20} in the range 20–30% and low P_{20} values (say below 20 mm) may be questioned. However, this questioning must be taken further through

an analysis that takes into account not only each event's rainfall characteristics as done in this section, but also the other general variable that controls runoff, namely the catchment state when the storm occurs. The best way to do this analysis in quantitative terms is through the use of a hydrological model that is able to combine rainfall and catchment-state properties in a physically realistic manner. If the model is soundly built, and the prior model parameters are not too far off, then, without calibration, the distribution of events in a simulated-versus-observed runoff scatterplot should give clear indications of the possible bias of some data (exogenous inflow), since a reasonably good linear or curvilinear correlation is to be expected for unbiased observations. So far, the two steps of rainfall–runoff analysis (Fig. 5(a) and (b)) have consisted in performing simple transformation and/or combination of the raw rainfall variables before graphical comparison with observed runoff; the use of the uncalibrated hydrological model in Section 4.4 amounts to taking this analysis one step further by performing a 'hydrologically' more judicious combination/transformation of raw variables, likely to improve graphical correlation with runoff in the case of homogeneous, spill-free events, and therefore to increase confidence in the discrimination of doubtful events. Specifically, as compared to the above purely rain-based statistical analysis, the added value expected from the hydrological model is 3-fold: (i) rainfall information most relevant to runoff production is selected implicitly in the model, rather than explicitly and somewhat arbitrarily as in this section; (ii) non-linearity of rainfall–runoff transformation is accounted for, unlike in statistical analysis with raw rainfall variables; (iii) the catchment moisture-state variable is incorporated into the analysis.

4.4. 'Hands-off' simulation of recorded rainfall–runoff events

The *r.water.fea* model built for the Wankama catchment is used to help discriminate the population of 1992–1998 recorded events. The prior parameter values derived from available remote sensing and field data (Section 3.3) are applied without calibration. Fig. 6 shows the scatterplot of observed (apparent) versus simulated runoff coefficients C_r. The dotted line materializes the main diagonal, i.e. identity between observations and simulations (it falls well inside the overall cloud of points, lending credit to the prior set of values used for model parameters). Again, from a data analysis viewpoint, the simulated runoff coefficient may be considered as the best available a priori predictor variable to describe actual runoff coefficient variations, as it hydrologically integrates the effects of the various runoff-controlling factors (relevant rainfall and catchment-state characteristics). Fig. 6 confirms and emphasizes the overly scattered layout of points seen in Fig. 5. Grouping of points into distinct, well-apart subsets is even more pronounced, with a remarkable absence of points in the middle of the cloud. Unsurprisingly, known overflowing events (filled and, even more so, empty triangles, i.e. GIN and EXO events, respectively) generally show comparatively large to very large ratios of observed over simulated runoff coefficients. Subtraction of these known 'outliers' still leaves a heterogeneous observed-versus-simulated behavior: among the 84 as yet-unclassified events (circles), a fraction of them (11 filled circles) follow, or even

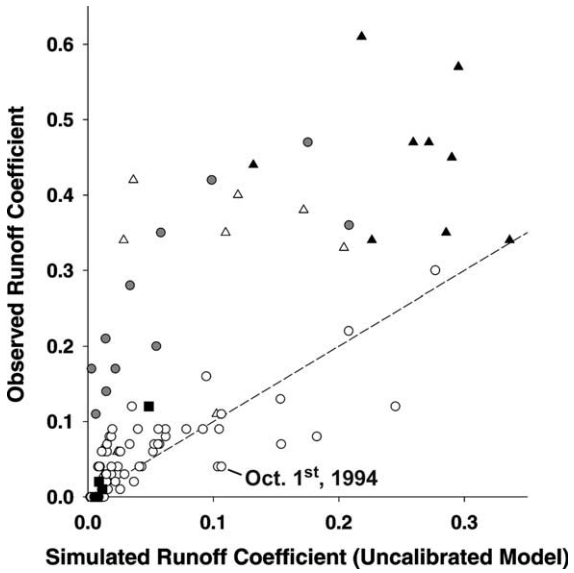

Fig. 6. Observed versus 'hands-off' simulated runoff coefficients, using untuned physically based, distributed hydrological model of Wankama watershed, 1992–1998. Symbols: same as Fig. 5, with circle differentiation: ●, suspected exogenous inflow (pEXO); ○, no exogenous inflow suspicion; dotted line = first bisector.

surpass, the triangles' trend of large observed-to-simulated ratios (GIN and EXO events). This trend differs significantly from the one followed by the other fraction of circles (empty symbols) and by the fully reliable VAL individuals (squares). This first group of 11 events may then be considered as highly questionable, and is also shown as filled circles in Fig. 5. Remarkably enough, the discrimination that can be made using the hydrological model (Fig. 6) is consistent with the one produced in Section 4.3 based solely on rainfall characteristics (i.e. similar trend for the GIN or EXO triangles and the questionable, filled circles, as opposed to the empty circles), indicating that the effect of soil moisture at initiation of each storm event cannot account for the large observed runoff coefficients displayed by the questionable events (filled circles). Comparing the two groups of circles in Fig. 6, it can be seen that for a given simulated runoff coefficient, observed coefficients for the questionable events are several times (from 3 to 10 times) larger than those of the other, unquestioned events (empty circles). The only convincing physical interpretation is that the excessively higher runoff observed for the questionable events is due to extra water supply to the Wankama pool, from North pool overflow; it is therefore reasonable to hold these runoff events, hereafter referred to as 'pEXO', as probably corrupted by exogenous inflow. Admittedly, for a small number of rather intermediate events, there is a certain subjectivity in their assignment into one group or the other (empty or filled circles), however, tests have shown that this is of only minor importance for the subsequent stages of the study (*C.Pap.*).

When one considers the empty circles only, five events (hereafter denoted 'LOW') appear in Fig. 6 to be on a markedly lower trend line, i.e. a comparatively lower observed-to-simulated runoff ratio, than the remaining points (in Fig. 5(a) and (b), these five points delineate the cloud's lower border as well, although they are not as clearly separated from the other points). One of the five is October 10, 1994 (Fig. 6), singled out in Section 4.1 as it shows by far the largest measured rainfall difference between the two Mare and Ouest gauges: using the Mare hyetograph uniformly over the whole catchment leads to overestimating the actual rain depth distribution for that event and hence its simulated runoff. Given this fact and the small number (four) of similarly behaving events, compared to the much greater number of individuals showing markedly higher observed-to-simulated runoff ratios, it is legitimate to hypothesize that, for these four events too, some unaccounted-for factor(s) may have induced overestimation of modeled runoff. Besides rainfall heterogeneity (which cannot be assessed for all events, due to short Ouest raingauge operation), the most likely such factor is the ephemeral increase in soil infiltration capacity due to farming activities. These mainly consist in soil crust breaking and weeding through hand-hoe tillage. For a nearby small plot, Peugeot et al. (1997) showed that crust-destructing soil tillage temporarily reduces runoff by a factor of 4. In a similar study in this area, Rockström and Valentin (1997) found that runoff can be reduced to nil after tillage. Such results were also obtained at other Sahelian locations by Hoogmoed and Stroosnijder (1984), Lal (1991) and Lamachère (1991). Tillage operations are usually performed prior to or at the very beginning of the rainy season, and may be repeated a second time in the middle of the growing period, around August. Infiltrability enhancement due to such practices is short-lived because of the fast crust restoration process on these sandy soils under Sahelian rain (Hoogmoed and Stroosnijder, 1984; Peugeot et al., 1997; Rockström and Valentin, 1997). Other recognized crust-destruction factors are cattle crossing in dry season (Hiernaux et al., 1999), and termite activity fed by cattle dejections and plant material (Mando and Stroosnijder, 1999; Lēonard and Rajot, 2001). All these factors contribute to the possibility of an ephemeral increase in soil infiltrability at the onset of and/or, less predictably, in the course of the growing season. Three of the four above 'singular' events (June 13, 1993, July 8, 1995, and June 9, 1996) are the first large rainstorms ($P > 30$ mm) in the respective seasons; comparatively lower observed runoff may then be imputed to season-beginning increased infiltrability. For the last of these four events, recorded on August 13, 1993, extensive mid-season weeding in the catchment is a possible explanation. However, these hypotheses cannot be validated in the field for lack of any precise record of actual farming operations. In order to verify their relevance, they are tested in *C.Pap.*, using the physically based, distributed model.

Finally, here are a few additional comments on this hands-off simulation with the uncalibrated hydrological model:

- the excellent overall performance of the prior parameter set is unintentional and fortuitous; events were not eliminated based on their individual proximity to the main diagonal in the simulated-versus-observed runoff scatterplot, but through an overall classification of the event sample into homogeneous groups based on their relative positions with respect to one another;
- the above picture of the relative event locations and groupings drawn from Fig. 6 is not changed in any noticeable way by model calibration or use of other realistic parameter values, at least under the limited number of degrees of freedom used for the present study (see *C.Pap.*);
- the simulated contribution to runoff from the upper part of the catchment is very small at the lower end of the spreading zone, a major feature of the catchment behavior.

4.5. Event classification

The above analysis suggests that observed events should be discriminated as to their further use in the catchment runoff study. The discrimination produced by the model-aided approach of Section 4.4 is in agreement with the results obtained from the statistical data-analysis approach of Section 4.3. Before a final classification is accepted, an analysis is performed on observed hydrograph shapes. Peak discharges are plotted against runoff depths in Fig. 7, again showing two rather separate trends. The five VAL events (squares, known to be non-overflowing) and the unquestioned, empty circles line up along a very linear trend, whereas the questioned, pEXO (filled circles) and the EXO events (empty triangles) follow a distinct, upper trend (the very large GIN events show a less clear picture, presumably due to more variable pool-exchange patterns). This upper trend means a flatter hydrograph with a higher volume for a given peak flow, as illustrated by Fig. 8 which shows the hydrograph and hyetograph for one of the EXO events. The graph exhibits an initial sharp discharge peak followed by a secondary delayed smoother

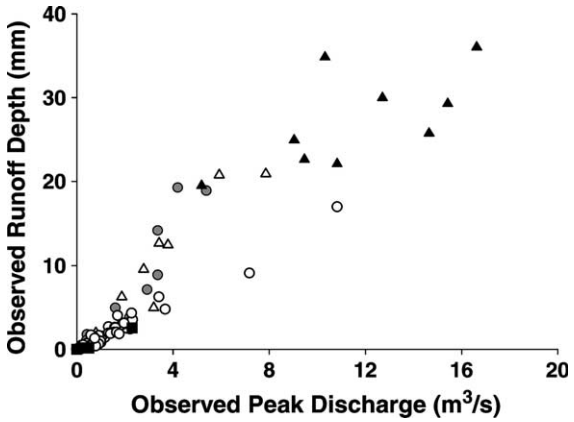

Fig. 7. Runoff depth versus peak discharge for observed flood events, 1992–1998; same symbols as Fig. 6.

peak, while the hyetograph only shows the typical single intense and short body with a thin tail usually observed for Sahelian storms. The secondary discharge peak is interpreted as the impact of delayed water supply from the North pool. This two-peak pattern is never observed for the VAL or unquestioned events (considered as non-overflowing), but neither is it systematically evidenced for EXO and pEXO events, as two contribution sources to the Wankama pool may not always be as clearly distinguished over time as in Fig. 8. The patterns in Fig. 7 strengthen the separation of the initially unclassified events (circles) into the two 'questioned, pEXO' (filled circles) and 'unquestioned' (empty circles) categories, as well as

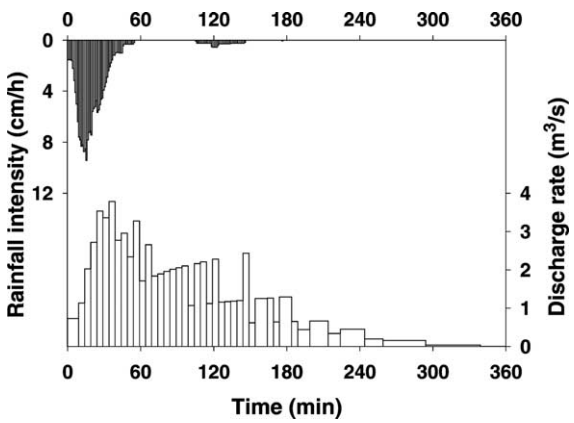

Fig. 8. Hyetograph and hydrograph for EXO event (north pool spillover) of August 28, 1998 ($P_{synth} = 2.13$; $C_r = 0.35$).

the interpretation that water supply from the North pool is the reason for these non-homogeneous behaviors.

Finally, these converging analyses led to adopting the following event classification and corollary reference sample selection, for the 1992–1998 period:

- 73 events, namely the five VAL (squares) plus the unquestioned empty circles other than the five LOWs (i.e. 68 events hereafter denoted 'pVAL'), are considered as reliably representing the dominant infiltration/runoff processes in the catchment; they are therefore suitable for tuning and testing of the hydrological model, and make up the so-called 'reference set';
- the five LOW events will be used to analyze the effects on catchment runoff of occasional, short-lived increase in soil infiltrability by farming practices or of recorded spatial rainfall heterogeneity;
- the 22 EXO and GIN events (empty and filled triangles) are affected by proven spillover or submersion;
- the 11 pEXO events (filled circles) are strongly suspected of inadequate catchment runoff measurement (inflow from North pool, like the EXO events); they, like the 22 EXOs and GINs, will consequently not be used for model calibration.

The recently acquired data for the last two 1999 and 2000 rainy seasons can be used to validate this classification a posteriori, as well as to enlarge the reference data set. For these 2 years, continuous recording of the North pool level provides a means for doubt-free classification of the 43 observed events (6 out of 49 observed rain events were not recorded at the pool) into 24 VAL (many of small magnitude), 4 GIN and 15 EXO. Re-drawing of Figs. 5(a) and (b) and 6 with superimposition of these new points would show excellent overlay of the 1999–2000 VAL area with the 1992–1998 reference-set area (VAL and pVAL) on the one hand, and of the 1999–2000 GIN and EXO area with the 1992–1998 region of GIN, EXO and pEXO events on the other hand. For instance, Fig. 9 is a replication of Fig. 6 (simulated versus observed runoff coefficients) with the addition of dotted squares and dotted triangles for the VAL and the GIN-and-EXO 1999–2000 events, respectively. Good agreement between the two periods can be seen, both for the rejected individuals including the 1992–1998 pEXO filled circles on one side, and for those accepted as the reference set including the 1992–1998 pVAL empty circles on the other. This concordance contributes to validating the method used and results obtained for the 1992–1998 event discrimination, particularly between the incompletely documented pEXOs and pVALs: had the North pool level not been known for any of the 43 1999–2000 events, they would still have been correctly classified by our method. No events in the small category of LOW events were identifiable in 1999–2000. Finally, the reference sample is made of a total of 97 rainfall–runoff events, the first 73 of which (1992–1998 period) being dedicated to model calibration and the last 24 (1999–2000 period) to its testing (*C.Pap.*).

Distribution of the reference data is sketched in Fig. 10, separately for the 1992–1998 (black) and the 1999–2000 (gray) periods. Fig. 10(a) shows rainfall against runoff depths for each individual

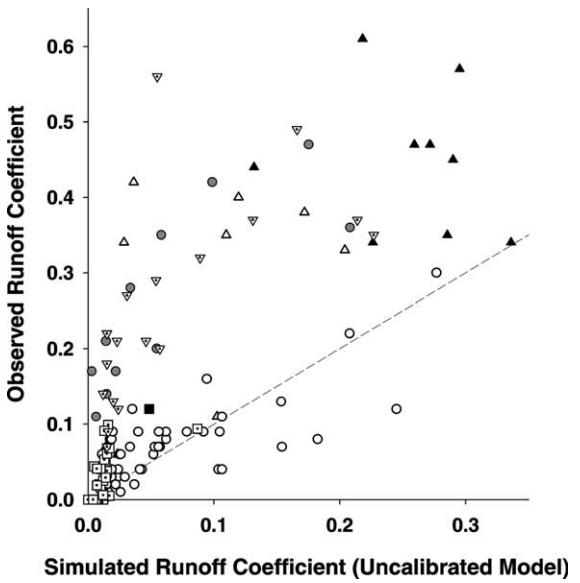

Fig. 9. Observed versus 'hands-off' simulated runoff coefficients, 1992–1998 (same as Fig. 6), plus superimposed 1999–2000 events: ⊡, no pool-transfer (VAL); ▽, inflow to Wankama pool from North pool (EXO) or general inundation (GIN).

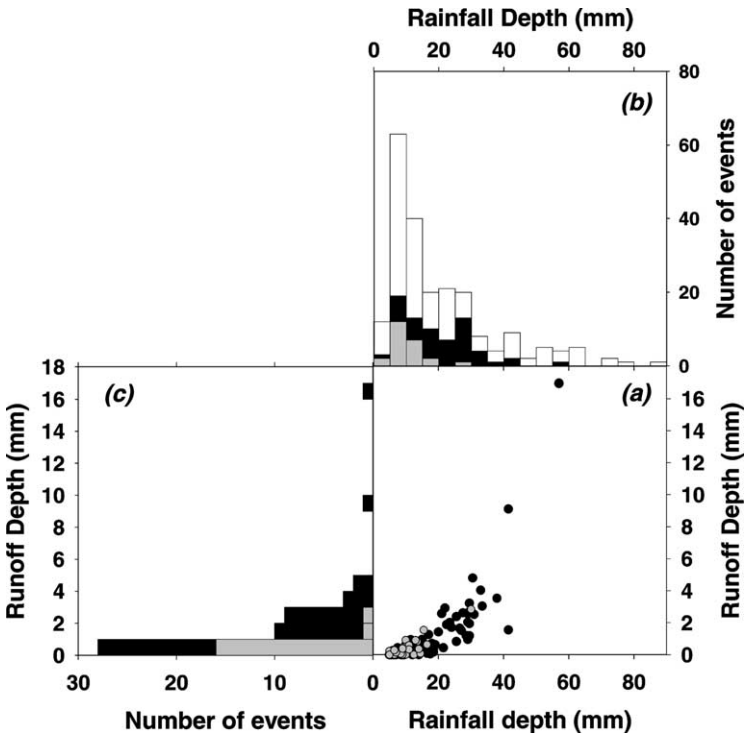

Fig. 10. Data distributions for rainfall and runoff depths: (a) rainfall versus runoff scatterplot, (b) rainfall frequency histogram, (c) runoff frequency histogram. Color codes: *black* = 1992–1998 reference events; *gray* = 1999–2000 reference events; *white* = all observations, in (b) only.

event, Fig. 10(c) is the frequency histogram of runoff depths in 1 mm-wide classes, and Fig. 10(b) is the frequency histogram of rainfall depths in 5 mm-wide classes, including the raw, observed population of 212 storms over the 1992–2000 period (white). The latter plot highlights the rather poor representation, by the reference sample, of the larger events in the original population, for which pool overflow is frequent. With regard to model calibration, this imbalance in the reference sample is partly counterbalanced by the higher responsiveness, pointed out by Séguis et al. (2002), of hydrologic parameter calibration to smaller storms, and is tentatively accounted for in the modeling strategy (*C.Pap.*).

5. Conclusion

With the aim of studying the hydrological response to climate changes over a 1500-km^2 Sahelian region of small endoreic catchments, a distributed, physically based model is being developed initially for the 1.9-km^2 Wankama watershed, using the *r.water.fea* model. Model construction and assignment of prior values to model parameters are based on detailed information obtained from field survey and remote sensing images. Heterogeneities in the raw sample of observed storm events, arising from occasional but poorly documented exogenous inflow to the Wankama pool, led us to perform a classification of the 163 events recorded in the 1992–1998 period. It is based first on statistical analysis of rainfall/runoff data, using rainfall variables that carry the largest information content with respect to runoff production, second on the relationship between runoff observations and results of simulation with the hydrological model, prior to any parameter calibration, and finally on the analysis of hydrograph shape, represented by the observed runoff volume-to-peak discharge relation. All three independent investigation pathways

produce consistent event behavior patterns, leading to a classification that is validated by tests against the more thoroughly monitored 1999–2000 period. A subset made of 73 events from the 1992–1998 period and of 24 events from the 1999–2000 period is selected as our reference sample for further catchment modeling, the latter period being dedicated to model testing. Five other storms in the first period appear to have produced comparatively low runoff, a behavior that may be explained by specific conditions (farming practices, non-uniformity of rain), which are evaluated in *C.Pap.*

The fact that the different approaches produce converging outputs lends credit not only to the derived classification, but also to the modeling principles on which they are based. Indeed, the suitability of the P_{20} variable (cumulated above-20 mm/h intensities) is strengthened by our analysis. The first axis (P_{synth}) of the PCA on rainfall characteristics, mixing cumulated depths and intensities at the storm scale, also appears to be a judicious synthetic variable for statistical runoff prediction in the area. The realistic results provided by hands-off operation of the uncalibrated hydrological model, when confronted with observations, are an encouraging indication with regard to model validity. Finally, the similitude in event distribution patterns found when observed runoff is plotted either against rainfall variables only or against simulated runoff suggests that initial soil moisture at storm onset is of secondary importance for the rainfall–runoff process in the area, at least in the range of values generally encountered in this study. Similar conclusions have been drawn for other semiarid areas (Cantón et al., 2001). The methodology used and results obtained for model calibration and verification, including parameter output uncertainty characterization, are the topic of the second, companion paper.

Acknowledgements

This study was funded by IRD (formerly ORSTOM) and by the French Programme National de Recherche en Hydrologie. Special thanks to the many people, from France and Niger, who contributed to field data collection at Wankama over the past decade.

References

Albergel, J., 1987. Sécheresse, désertification et ressources en eau de surface—Application aux petits bassins du Burkina-Faso. Proceedings of International Symposium on the Influence of Climate Change and Climatic Variability on the Hydrological Regime and Water Resources, Vancouver, Canada, IAHS Publication 168, pp. 355–365.

Amani, A., Lebel, T., 1996. Typology of rainfall fields to improve rainfall estimation in the Sahel by the area threshold method. Water Resour. Res. 32 (8), 2473–2487.

Beven, K., 1989. Changing ideas in hydrology—the case of physically based models. J. Hydrol. 105, 157–172.

Bromley, J., Taylor, C.M., Gash, G.H.C., et al., 2002. Comment on 'Long-term rise in a Sahelian water-table: the Continental Terminal in South-West Niger' by Leduc et al. (2001). J. Hydrol. 255, 260–262.

Cantón, Y., Domingo, A., Solé-Benet, J., Puigdefábregas, J., 2001. Hydrological and erosion response of a badlands system in semiarid SE Spain. J. Hydrol. 252, 65–84.

Cappelaere, B., Vieux, B., Peugeot, C., Maia, A., Séguis, L., 2003. Hydrologic process simulation of a semiarid, endoreic catchment in Sahelian West Niger, Africa. 2. Model calibration and uncertainty characterization. J. Hydrol.

Casenave, A., Valentin, C., 1992. A runoff capability classification system based on surface features criteria in semiarid areas of West Africa. J. Hydrol. 130, 231–249.

Chow, V.T., 1959. Open Channel Hydraulics, McGraw-Hill, New York.

CPCS (Commission Pédologique de Classification des sols), 1967. Classification des Sols, Ecole Nationale Supérieure d'Agronomie, Grignon, France, p. 87.

Cuenca, R.H., Brouwer, J., Chanzy, A., Droogers, P., Galle, S., Gaze, S.R., Sicot, M., Stricker, H., Angulo-Jaramillo, R., Boyle, S.A., Bromley, J., Chebhouni, A.G., Cooper, J.D., Dixon, A.J., Fies, J.-C., Gandah, M., Gaudu, J.-C., Laguerre, L., Lecocq, J., Soet, M., Steward, H.J., Vandevaere, J.P., Vauclin, M., 1997. Soil measurements during Hapex–Sahel intensive observation period. J. Hydrol. 188/189, 224–266.

D'Amato, N., Lebel, T., 1998. On the characteristics of the rainfall events in the Sahel with a view to the analysis of climatic variability. Int. J. Climatol. 18 (9), 955–974.

Desconnets, J.C., Vieux, B.E., Cappelaere, B., Delclaux, F., 1996. A GIS for hydrological modeling in the semiarid, Hapex–Sahel experiment area of Niger, Africa. Trans. GIS 1 (2), 82–94.

Desconnets, J.C., Taupin, J.D., Lebel, T., Leduc, C., 1997. Hydrology of Hapex–Sahel central super site: surface drainage and aquifer recharge through the pool systems. J. Hydrol. 188/189, 155–178.

D'Herbès, J.M., Valentin, C., 1997. Land surface conditions of the Niamey region: ecological and hydrological implications. J. Hydrol. 188/189, 18–42.

Estèves, M., 1995. Rapport de campagne hydrologique, saison 1994 [Report of field collection of hydrological data, 1994 season]. Orstom, Niamey (Niger).

Favreau, G., Leduc, C., Marlin, C., Dray, M., Taupin, J.D., Massault, M., Le Gal La Salle, C., Babic, M., 2002a.

Estimate of recharge of a rising water table in semiarid Niger from ^{3}H and ^{14}C modeling. Ground Water 40 (2), 144–151.

Favreau, G., Leduc, C., Schroeter, P., et al., 2002b. Reply on 'Long-term rise in a Sahelian water-table: the Continental Terminal in South-West Niger' by Leduc et al. (2001). J. Hydrol. 255, 263–265.

Galle, S., Ehrmann, M., Peugeot, C., 1999. Water balance in a banded vegetation pattern. A case study of tiger bush in western Niger. Catena 37, 197–216.

Goutorbe, J.P., Lebel, T., Tinga, A., Bessemoulin, P., Brouwer, J., Dolman, A.J., Engman, E.T., Gash, J.H.C., Hoepffner, M., Kabat, P., Kerr, Y.H., Monteny, B., Prince, S., Said, F., Sellers, P., Wallace, J.S., 1994. Hapex–Sahel: a large scale study of land-atmosphere interactions in the semiarid tropics. Ann. Geophys. 12, 53–64.

Hiernaux, P., Bielders, C.L., Valentin, C., Bationo, A., Fernandez-Rivera, F., 1999. Effects of livestock grazing on physical and chemical properties of sandy soils in Sahelian rangelands. J. Arid Environ. 41, 231–245.

Hoogmoed, W.B., Stroosnijder, L., 1984. Crust formation on sandy soils in the Sahel. I. Rainfall and infiltration. Soil Tillage Res. 4, 5–23.

Kohler, M.A., Linsey, R.K., 1951. Predicting the runoff from storm rainfall. Weather Bureau, US Department of Commerce. Research Paper 34, Washington, USA

Lal, R., 1991. Current research on crop water balance and implications for the future. In: Sivakumar, M.V.K., Wallace, J.S., Renard, C., Giroux, C. (Eds.), Soil Water Balance in the Sudano-Sahelian Zone, Proceedings of International Workshop, February 1991, Niamey, Niger. IAHS Publication 199, pp. 31–44.

Lamachère, J.M., 1991. Aptitude au ruissellement et à l'infiltration d'un sol sableux fin après sarclage. In: Sivakumar, M.V.K., Wallace, J.S., Renard, C., Giroux, C. (Eds.), Soil Water Balance in the Sudano-Sahelian Zone, Proceedings of International Workshop, February 1991, Niamey, Niger. IAHS Publication 199, pp. 109–119.

Lang, J., Kogbe, C., Alidou, S., Alzouma, K.A., Bellion, G., Dubois, D., Durand, A., Guiraud, R., Houessou, A., De Klasz, I., Romann, E., Salard-Cheboldaeff, M., Trichet, J., 1990. The Continental-Terminal in West Africa. J. Afr. Earth Sci. 10, 79–99.

Lange, J., Leibundgut, C., Greenbaum, N., Schick, A.P., 1999. A noncalibrated rainfall–runoff model for large, arid catchments. Water Resour. Res. 35 (7), 2161–2172.

Le Barbé, L., Lebel, T., 1997. Rainfall climatology of the Hapex–Sahel region during the years 1950–1990. J. Hydrol. 188/189, 43–73.

Lebel, T., Le Barbé, L., 1997. Rainfall monitoring during Hapex–Sahel. 2. Point and areal estimation at the event and seasonal scales. J. Hydrol. 188/189, 97–122.

Lebel, T., Taupin, J.D., D'Amato, N., 1997. Rainfall monitoring during Hapex–Sahel. 1 General rainfall conditions and climatology. J. Hydrol. 188/189, 74–96.

Leduc, C., Loireau, M., 1997. Evolution du couvert végétal et de la recharge de la nappe phréatique en zone sahélienne (Niamey Niger). Sustainability of Water Resources under Increasing Uncertainty, IAHS Publication 240, pp. 193–200.

Leduc, C., Favreau, G., Schroeter, P., 2001. Long-term rise in a Sahelian water-table: the Continental Terminal in South-West Niger. J. Hydrol. 243, 43–54.

Lēonard, J., Rajot, J.L., 2001. Influence of termites on runoff and infiltration: quantification and analysis. Geoderma 104 (1/2), 17–40.

Mahé, G., Olivry, J.C., Dessouassi, R., Orange, D., Bamba, F., Servat, E., 2000. Relations eaux de surface-eaux souterraines d'une rivière tropicale au Mali. C. R. Acad. Sci. (Sér. 2a) 330, 689–692.

Mando, A., Stroosnijder, L., 1999. The biological and physical role of mulch in the rehabilitation of crusted soil in the Sahel. Soil Use Manage. 15 (2), 123–127.

Martin-Vide, J.P., Ninerola, D., Bateman, A., Navarro, A., Velasco, E., 1999. Runoff and sediment transport in a torrential ephemeral stream of the Mediterranean coast. J. Hydrol. 225 (3/4), 118–129.

Michaud, J., Sorooshian, S., 1994. Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed. Water Resour. Res. 30 (3), 593–605.

Mitasova, H., Mitas, L., 1993. Interpolation by regularized spline with tension. I. Theory and implementation. Math. Geol. 25, 641–655.

Nagumo, F., 1992. Pedological environment and agro-ecological system of the Sudano-Sahelian zone, in Niger, West Africa. Master Course Thesis, Graduate School of Environmental Science, Hokkaido University, Japan.

Peugeot, C., 1995. Influence de l'encroûtement superficiel du sol sur le fonctionnement hydrologique d'un versant sahélien (Niger). Expérimentations in situ et modélisation [Influence of superficial soil crusting on the hydrological behaviour of a Sahelian hillslope (Niger)—Field experiments and modeling. PhD Thesis]. Thèse, Université Joseph Fourier, Grenoble, France.

Peugeot, C., Estèves, M., Galle, S., Rajot, J.L., Vandervaere, J.P., 1997. Runoff generation processes: results and analysis of field data collected at the East Central SuperSite of the Hapex–Sahel experiment. J. Hydrol. 188/189, 179–202.

Richardson, J.R., Julien, P.Y., 1994. Suitability of simplified overland flow equations. Water Resour. Res. 30 (3), 665–671.

Rockström, J., Valentin, C., 1997. Hillsope dynamics of on-farm generation of surface water flows: the case of rain-fed cultivation of pearl millet on sandy soil in the Sahel. Agric. Water Manage. 33, 183–210.

Saghafian, B., Julien, P.Y., Ogden, F.L., 1995. Similarity in catchment response. 1. Stationary rainstorms. Water Resour. Res. 31 (6), 1533–1541.

Séguis, L., Cappelaere, B., Peugeot, C., Vieux, B., 2002. Impact on Sahelian runoff of stochastic and elevation-induced spatial distributions of soil parameters. Hydrol. Process. 16, 313–332.

Soil Survey Staff, 1975. Soil Taxonomy. A Basic System on Soil Classification for Making and Interpreting Soil Surveys, 436, Soil Conservation Service, US Department of Agriculture, Washington, DC, USA, p. 754.

USACE, 1993. *Grass 4.1* User's Reference Manual, US Army Corps of Engineers Construction Engineering Research Laboratories, Champaign, IL, USA, p. 556.

Valentin, C., Bresson, L.M., 1992. Morphology, genesis and classification of soil crusts in loamy and sandy soils. Geoderma 55, 225–245.

Valentin, C., D'Herbès, J.M., 1999. Niger tiger bush as a natural water harvesting system. Catena 37, 231–256.

Vieux, B.E., 2001. Distributed Hydrologic Modeling Using GIS, Water Science and Technology Library, vol. 38. Kluwer Academic Publishers, Dordrecht.

Vieux, B.E., Gaur, N., 1994. Finite-element modeling of storm water runoff using *Grass* GIS. Microcomput. Civil Engng 9 (4), 263–270.

Ye, W., Bates, B.C., Viney, N.R., Sivapalan, M., Jakeman, A.J., 1997. Performance of conceptual rainfall–runoff models in low-yielding ephemeral catchments. Water Resour. Res. 33 (1), 153–166.

Walters, M.O., 1990. Transmission losses in arid region. J. Hydraul. Engng 116 (1), 129–138.

Woolhiser, D.A., Smith, R.E., Giraldez, J.V., 1996. Effects of spatial variability of saturated hydraulic conductivity on Hortonian overland flow. Water Resour. Res. 32 (3), 671–678.