
Suitability of a lumped rainfall–runoff model for flashy tropical watersheds in New Caledonia

Desclaux T rence ^{1,*}, Lemonnier Hugues ¹, Genthon Pierre ², Soulard Benoit ¹, Le Gendre Romain ¹

¹ IFREMER, Unit  de Recherche Lagons, Ecosyst mes et Aquaculture Durable, Noum a, New Caledonia

² IRD, Universit  Montpellier, CNRS, HydroSciences Montpellier, Noumea, New Caledonia

* Corresponding author : T rence Desclaux, email address : terence.desclaux@ifremer.fr

Abstract :

The GR4H lumped hourly rainfall–runoff model was assessed for its integration in a ridge-to-reef modelling framework. Particular attention was paid to rainfall representation, robustness of parameter estimates and ability to reproduce main runoff features. The study was conducted in four tropical mountainous watersheds in New Caledonia, exposed to intense rainfall events, large annual climatic variations triggered by El Ni o oscillation, and wildfires. The inverse distance and elevation weighting algorithm outperformed other classical rainfall interpolation methods under data-limited conditions. The time-span of data needed for robust calibration was site-specific and varied from 6–7 years to 10 years, which may be linked to El Ni o events and to wildfires. With sufficient data, simulation quality was equivalent during calibration and validation periods. The GR4H model was generally able to simulate both flash floods and large annual variations. The model was more reliable when simulating wet years and watersheds not subject to land-cover changes.

Keywords : hydrological model, ridge-to-reef, GR4H, rainfall–runoff model, integrated modelling, New Caledonia, lagoon, wildfire

Funding: This work was undertaken within the PRESENCE (PRESsures on coral Ecosystems of New CalEdonia) scientific project, sponsored by New Caledonia institutions (Government, South Province, North Province).

1 Introduction

Rivers constitute a major pathway for the impact of human footprint on marine ecosystems (Doney, 2010). As coral reefs host high biodiversity and provide a wide range of ecosystem services (Burke *et al.*, 2011), multidisciplinary management of transfers of freshwater, suspended and dissolved matter toward these marine ecosystems appears to be necessary (Richmond *et al.*, 2007, David *et al.*, 2010). This has generally been investigated using in-situ measurements and modelling of the transport of terrestrial tracers (Devlin and Brodie, 2005, Fernandez *et al.*, 2006) with remote sensing (Ouillon *et al.*, 2008, Petus *et al.*, 2016). However, in tropical mountainous islands, where many of the coral ecosystems are found, rivers and lagoons are subject to sudden events and processes, at typical time steps ranging from hours to days (Supplementary material 1, Arfi *et al.*, 1993, Smith *et al.*, 1998, Torr ton *et al.*, 2010) and present feedbacks effects between sea and land. They are therefore best analysed with fully coupled models working at appropriate time steps. Such integrated models can be used to efficiently manage land-based pollution and accurately target catchments to be managed (Brodie *et al.*, 2012). Moreover, they can address the sensitivity of coastal areas to extreme events, including both oceanic and hydrological effects (Dresback *et al.*, 2013).

Ridge-to-reef models should be able to account for land-surface changes and their spatio-temporal influence on the coastal area (Inoue *et al.*, 2008, Rude *et al.*, 2016). They should also be easily extended in space, e.g. regionalised to deal with ungauged or poorly gauged locations (Hrachowitz *et al.*, 2013) or in time, e.g. to investigate the impact of climate change (Manabe *et al.*, 2004, Holt *et al.*, 2010). However, in integrated modelling studies, classical atmospheric and marine models are generally used, but rainfall-runoff processes are often simplified (e.g. Manabe *et al.*, 2004, Rude *et al.*, 2016). This makes the extension of hydrological results difficult, especially in flash flood prone basins, where hydrological processes may be complex (Harpurachchi *et al.*, 2011). Therefore, the first aim of this paper is to propose an original approach to represent more realistically the hydrological component of a ridge-to-reef model, in particular in tropical flash-flood prone basins. In particular, in order to facilitate the temporal extension of the results, robust estimates of parameters are provided, and their spatial extensibility (e.g. for regionalisation) is discussed. To allow comparison with other hydrological studies, the model quality is presented during calibration and validation periods (Klemeš *et al.*, 1986, Dawson *et al.*, 2007). Also, as results are rarely reported to be equivalent during calibration and validation period in hydrological modelling studies (e.g. Coron *et al.*, 2012) – though this is common in ocean studies (e.g. Hollingsworth, 1994) –, the possibility of merging these periods is analysed. Lastly, the possible utility of the model for accounting for sudden land-cover changes is discussed.

The present study is conducted in the mainland of New Caledonia, an archipelago surrounded by one of the ten coral reef biodiversity hotspots and the second largest lagoon in the world (Roberts *et al.*, 2002). This region has a tropical climate, characterised by high inter-annual variability of rainfall amounts and recurrent

atmospheric extreme events. High and steep watersheds induce very rapid water transfers to the lagoons and sudden flash floods (Terry *et al.*, 2008), while large wildfires can severely change landscapes (Gomez *et al.*, 2015). Even though hydrological processes controlling such flash flood prone basins are known to be complex (Hapuarachchi *et al.*, 2011), there is no detailed knowledge of them in this region (Terry and Wotling, 2011). It is therefore important to provide a first estimate of the variability and stability of runoff generation in these watersheds. However, the small size of the basins, generally below 100 km², and the high spatial variability of rainfall (Terry and Wotling, 2011) compromises the use of soil moisture and rainfall satellite products (Maggioni and Massari, 2018), as well as extensive in-situ measurement of soil characteristics. Each basin is sampled by only one or two rain gauges, which provide relatively little information on the spatial distribution of rainfall on the watersheds. Therefore, modelling flash flood with a distributed model is considered not to be feasible at present (Hapuarachchi *et al.*, 2011, Merheb *et al.*, 2016) and a lumped rainfall-runoff model is adopted. Also, particular attention is paid to spatial interpolation of rainfall, accurate reproduction of flash flood and annual volumes.

As far as we know, this study provides the first rainfall-runoff modelling application on the watersheds of New Caledonia. This paper begins with a description of the four basins investigated. Methods, and the lumped hydrological model are then presented. Results are presented regarding i) spatial interpolation of rainfall, ii) model calibration and iii) simulation of the streamflow during flash floods and annual volumes. Lastly, the proposed approach is discussed, including model calibration, model performance, rainfall representation, limitations and further improvements needed for building a fully coupled model for interaction of runoffs with the New Caledonian lagoons.

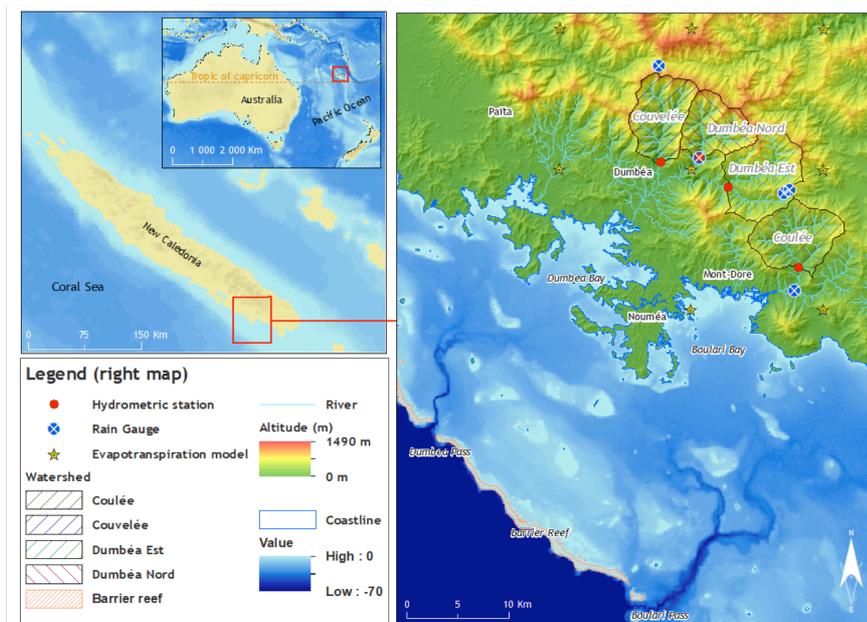


Fig. 1: Study area: location of instruments, topography, bathymetry, surface hydrological network, delimitation of the watersheds, toponyms and sampling points. The Couvelée rain gauge is located at ~ 920 m of altitude ~ 600 m north of the crest of the basin.

2 Study site

New Caledonia is located in the Coral Sea, 1,200 km east of Australia, around the twentieth parallel south (Fig. 1). It has been shown to be of exceptional biodiversity both on land and in its coral lagoons (Myers *et al.*, 2000, Roberts *et al.*, 2002). Approximately 400 km long and 40 to 70 km wide, the mainland, Grand Terre, is needle-shaped with a chain of mountains running along its entire length (Fig. 1). It is underlain by the peridotites nappe that sculpts the relief into high and steep watersheds (Trescases, 1975). Its rainfall is controlled by the position of the South Pacific Convergence Zone (SPCZ) and modulated by the El Niño Southern Oscillation (ENSO). Droughts are common during El Niño years and higher rainfalls characterize La Niña years (Morlière & Rebert, 1986, Vincent *et al.*, 2011). Throughout the hot wet season (December to May), passing tropical depressions can mature into tropical cyclones, triggering intense rainfall and suddenly increasing the runoff of rivers (Terry *et al.*, 2008). These results were confirmed by a multi-scale analysis of the runoff time series (presented in the Supplementary material – S1). Four watersheds were selected, namely Couvelée, Dumbéa Nord, Dumbéa Est and Coulée, which are situated behind the Nouméa peninsula in the southwest part of the mainland (Fig. 1). Their main characteristics are listed in Table 1 (sources: Terry & Wotling, 2011, OEIL, 2012 and our own calculations). These catchments are subject to relatively high rainfall, with annual amounts exceeding 2100 mm. They are delineated on their northern side by a mountain ridge with several summits higher than 1100 m (Fig. 1). All four watersheds are mountainous and their mean elevation ranges from 330 m for the Coulée watershed to 571 m for Dumbéa Nord, while mean slopes range from 17.3° at Coulée to 25° at Dumbéa Nord. The Coulée catchment was used for large-scale exploitation between 1966 and 1980, until mining ceased in 1981 (Fernandez *et al.*, 2006). Since the end of mining operations, the vegetation has slowly recovered. During the study period, the land cover of the watersheds underwent several abrupt changes (Fig. 2), all caused by uncontrolled wildfires, the extent of which was assessed from the Global Tree Change results (Hansen *et al.*, 2013). A relatively small wildfire burned 5% of the Dumbéa Nord area in late 2004. The Coulée and the Dumbéa Est catchments experienced a large wildfire event in late 2005 to early 2006, which burned more than 59% and 15% of their respective areas. As a result, in 2010, forest covered more than 75% of the Couvelée watershed, whereas it represented only 26.3% of the Coulée area. In turn, bushes and savannah covered more than 54% of Coulée, against just 22% of Couvelée. Together, former mining sites and bare land accounted for 18.8% of the Coulée area.

Table 1: Physical, hydrological, and land-cover features of the four watersheds. (a) Adapted from (Terry & Wotling, 2011), (b) computed from data used in this study, (c) from (OEIL, 2012).

	Coulée	Dumbéa Est	Dumbéa Nord	Couvelée
Area (km ²) (a)	44	51	32	40
Mean elevation (m) (a)	331	513	571	488
Mean hillslope (°) (a)	17.3	18.4	25.0	24.3
In 2010, part of the watershed covered by: (%) (c)				
Forest	26.3	48.3	67.3	76.2
Bushes / savannah	54.3	49.5	32.1	21.9
Mining sites (disused and used)	9.0	0	0	0.9
Bare land	9.8	1.4	0	0
Others	0.6	0.8	0.6	1.0
Annual rainfall (mm/year) (b)	2387	2470	2117	2301
Annual runoff (mm/year) (b)	1447	1809	1489	929
Mean runoff coefficient (%) (b)	60	73	70	40
Accuracy of high flow rating curve (a)	Good	Very good	Very Good	Poor
Annual PE (mm) (b)	1012	893	790	789

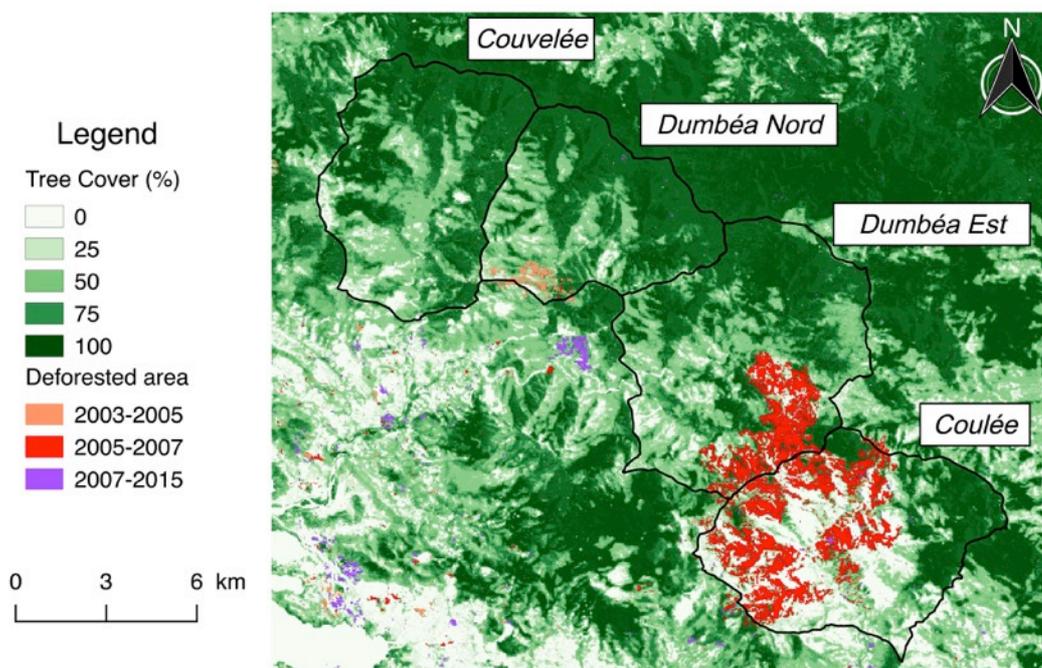


Fig. 2: Tree cover and deforestation on the watersheds studied, adapted from (Hansen *et al.*, 2013).

3 Material and methods

3.1 Data sources and analyses

3.1.1 Data sources

Fig. 1 summarizes the locations of the instruments and grid points used in this paper. Precipitation was measured at hourly time steps with five bucket rain gauges (Data source: DAVAR, Direction des Affaires Vétérinaires, Alimentaires et Rurales, and Météo France). Streamflows were measured at hourly time steps using four limnimeters (Data source: DAVAR). Hourly potential evapotranspiration (PE) is taken from computation at daily time steps on a 0.125° grid (Data source: Meteo-France according to Prats, 2009). These values were bilinearly interpolated on a $100\text{ m} \times 100\text{ m}$ grid, then spatially averaged on each watershed. Hourly PE values were then obtained by keeping PE constant throughout the day (equal to $1/24$ of the daily value), which was simple yet consistent with previous results (e.g. Oudin et al., 2004).". Elevation data was used at the resolution of 50 m (Data source: Government of New Caledonia). Land cover of 2010 was established by OEIL (2012) (Data source: Observatoire de l'Environnement en Nouvelle Calédonie). Forest changes were determined following Hansen *et al.*, 2013 (Data source: Global Forest Change project). Because of the New Caledonia seasonality of precipitation, hydrological years were deemed to last from the 1 October to 30 September. Lastly, this study focuses on a twelve-year period, from 01/10/2003 to 31/09/2015.

3.1.2 Rainfall spatial interpolation

Because it has been shown that rainfall-runoff modelling is particularly sensitive to rainfall data quality (Oudin *et al.*, 2006), three rainfall spatial interpolation methods were evaluated. The first, here termed Nearest, involved taking rainfall data from the station nearest to the centroid of each catchment. When data was missing, the next closest station was chosen to fill the gaps. The second method was the classical Thiessen polygon method (Nalder & Wein, 1998). Lastly, the Inverse Distance and Elevation Weighting algorithm (IDEW) (Masih *et al.*, 2011) was used to complete our third rainfall dataset. For this last method, rain gauge is weighted by scalars whose value represents the relative importance of each station and decreases as distance and elevation differences increase, following:

$$p_k = W \frac{\sum_i \frac{p_i}{d_i^\alpha}}{D} + (1 - W) \frac{\sum_i \frac{p_i}{\Delta z_i^\beta}}{Z}$$

Where p_k is the interpolated precipitation of a grid cell k , W is the relative importance factor for distance vs. elevation, p_i is the precipitation value in mm/day of the i th gauge station, d is the distance between the current grid and the gauge station, Δz is the absolute elevation difference between the current grid cell and the gauge station, α and β are the exponent factors for distance and elevation weighting, and D and Z are the normalization quantities given by the sum of individual weighting factors $1/(d_i)^\alpha$, and $1/(\Delta z_i)^\beta$ for all the gauges used in the interpolation. The weights are normalised so that

the algorithm automatically adjusts whenever a station goes in or out the monitoring pool and Δz is bounded between Δz_{\min} and Δz_{\max} .

For IDEW algorithm, only 3 parameters need to be calibrated: W , α and β as Δz_{\min} and Δz_{\max} were set at the default value of 100 m and 1,500 m respectively (Daly *et al.*, 2002, Masih *et al.*, 2011). To calibrate the parameters, the entire possible domain was covered for each of the watersheds studied: W was set between 0 and 1 with steps of 0.1, α between 0.5 and 2.5 with steps of 0.5, and β from 0.5 to 2 with steps of 0.5. For each of the 220 possible IDEW parameter sets, and for each of the watersheds, 4 additional parameters were calibrated for the rainfall-runoff model using the procedure described in the next section. Simulation quality was assessed using Kling Gupta Efficiency (Gupta *et al.*, 2009) on the 12 years of available data. The best IDEW calibration was then determined as the closest to the utopia point (Marler *et al.*, 2004).

3.1.3 Flood description and analysis

The flood analysis procedure consisted first in selecting the maximum discharge peaks in the whole hourly time series (Ficchi *et al.*, 2016). The flood beginning (and end) was specified as the time when the flow becomes higher (lower) than 10% of the flood maximum. If precipitation was not initially zero, the beginning of the event was shifted to the first precipitation time steps at which it was zero. Doing so enabled us to calculate the median height of precipitation and runoff during flood and the 24h flow shape coefficient, which represents the mean of the ratio between the maximum hourly flow (Ficchi *et al.*, 2016).

3.2 Model implementation

3.2.1 Rainfall-runoff model

The GR4H (Hourly Rural Engineering model with 4 parameters) model is the hourly variant of the lumped rainfall-runoff GR4 suite of models (Perrin *et al.*, 2003), that has been applied in a wide range of countries and climates, including tropical climates (e.g. Coron *et al.*, 2012). The model can be calibrated with four parameters and two time series inputs (rainfall and evapotranspiration). The GR4 structure (Fig. 3) has been extensively described (Perrin *et al.*, 2003, Bennett *et al.*, 2014) and is composed of a soil moisture accounting function followed by routing and underground water exchange functions. The soil moisture accounting (SMA) store receives the net rainfall and satisfies the net potential evapotranspiration. The routing function consists of two unit hydrographs and a non-linear routing store. The routing store also includes exchanges with groundwater. The four parameters are: X1 (mm) the maximum capacity of the SMA store; X2 (-), the water exchange coefficient, which is positive in the case of water gain; X3 (mm), the maximum capacity of the routing store; and X4 (h), the time base of the unit hydrographs.

the R packages “airGR” and “sensitivity” (Coron *et al.*, 2017, Pujol *et al.*, 2017). The Sobol method decomposes the model’s output variance into relative contributions from individual parameters and parameter interactions. Each parameter is then allocated a total sensitivity index (TSI), equal to its partial contribution, including all its interactions with other parameters, to the output variance. A Monte Carlo estimate of the sensitivity indices was used following the asymptotically efficient formulas described by Monod *et al.*, (2006), using initial an sample size of 5,000. Lastly, the confidence intervals were estimated by means of bootstrapping.

3.3.2 Length of the calibration period

As already mentioned, hydrologists often use cross-validation in order to assess the robustness of their models. Therefore, it is important to i) separate the calibration and validation periods and ii) furnish sufficient guidelines toward the identification of stable, robust parameter sets.

K-fold partitioning was applied (Kohavi, 1995, Bennett *et al.*, 2013) to determine the stability of the parameters and model quality for each calibration length. The aim of this methodology is to split the data into k sets and to use one for calibration and rest for validation. The process is then repeated k times, allowing all results to be averaged and analysed. Here, we defined $k=12$ sets of one year. This method was repeated with increasing length of calibration period (L), from 1 to 11 years. The data was subdivided in a deterministic manner, according to the following algorithm:

- (i) The first set starts in 2003 and ends in 2003+ L
- (ii) For the next set, the beginning and end is shifted to one year later.
- (iii) If the end of the set is later than last available date (30 September 2015) it is wrapped at the beginning of the available period. For example, with $L=2$, the last set is 2014-2015 and 2003-2004.
- (iv) Repeat (ii) and (iii) 11 times.

The k sets of L hydrological years were therefore not mutually exclusive when $L>1$, which is slightly different from what was initially done in k -fold partitioning (Kohavi, 1995). However, to ensure the consistency of the results, all hydrological years were present for the same number of times in the k sets. Changes of parameters were followed by calculating the deviation from their respective benchmark values (Li *et al.*, 2015), by $D_i = |(P_i - P_w)/P_w| \times 100$, where i is the index of the sub-period, P_i is the calibrated parameter value on the i_{th} sub-period and P_w is the calibrated parameter value on the whole period.

To simplify interpretation of the results, the calibration and validation results were presented for the hydrological year 2014-2015. When this year was part of the calibration set, the model results were flagged as “calibration” results. When it was part of the validation set, the results were flagged as “validation” results. In this way the performance of the model during calibration and validation was kept distinct. Moreover,

doing this kept the results independent of the period used to estimate the KGE, and was similar to the method used by Perrin *et al.* (2007), in that they kept a fixed validation set while changing the calibration data.

3.3.3 Simulation quality estimation

In order to make the results of this study accessible to a broad readership, the simulation quality is presented using a number of common, previously defined metrics (Bennett *et al.*, 2013, Ficchi *et al.*, 2016): Relative Volumetric Error (RVE), the square of the Pearson Product moment correlation (Rsqr), Index of Agreement (IOA), Nash-Sutcliffe Efficiency (NSE), Persistence Index (PI), Volumetric Efficiency (VE), and KGE. In addition, as already mentioned, a multi-scale analysis of the runoff time series was conducted (see Supplementary material S1 for additional details), and the model results were investigated at the two most important time scales of runoff variation. In particular, simulated and observed flash floods were compared by means of various metrics, such as time-to-peak error and median peak flow error. Relative volumetric error (RVE) was computed for three periods: the whole flood period, the rising limb, and the falling limb. The measured and simulated runoffs were visually compared on the 5 largest flood events that impacted the watersheds studied. The annual runoff was also computed for each watershed from measurements and simulations. This value was defined on a daily increment as the total amount of runoff between 6 months before and 6 months after the day concerned.

4 Results

4.1 Model implementation

4.1.1 Rainfall spatial interpolation

After calibration of the model parameters over the 12 years of available data, different validation metrics, described in section 3, were computed on the whole dataset (Table 2). The best IDEW calibration was obtained with $W=0.8$, $\alpha=1.5$ and $\beta=0.5$. This calibration was used for all watersheds (Masih *et al.*, 2011). The model quality was better with the IDEW method than with the others, except for Couvelée (Table 2). For this watershed, the Thiessen method gave slightly better results (PI = -1.8 versus -2.1 with IDEW). However, the difference was small and, even using IDEW-generated rainfall, Couvelée still remained the best-simulated watershed. Therefore, the IDEW method was chosen for the next parts of the study. With this method, whichever river was considered, RVE was always lower than 1.6% (in absolute value), Rsqr was higher than 0.76, IOA higher to 0.93, PI higher than -2.6, VE higher than 0.55, KGE higher than 0.87 and NSE higher than 0.75. Model quality was generally better for Dumbéa Est, Dumbéa Nord and Couvelée, where KGE was higher than 0.94 compared to Coulée, where KGE was 0.87.

Table 2: Evaluation metrics computed for the different rainfall spatial interpolation methods on the watersheds studied. Abbreviations: Near: nearest, Thie: Thiessen. Best values are highlighted in green.

	Coulée			Dumbéa Est			Dumbéa Nord			Couvelée		
	Near	Thie	IDEW	Near	Thie	IDEW	Near	Thie	IDEW	Near	Thie	IDEW
RVE (%)	-1.6	-0.1	-1.6	0.1	0.04	0.1	-0.1	-0.07	0.07	-0.01	0.03	-0.03
Rsqr	0.71	0.73	0.76	0.88	0.90	0.92	0.82	0.83	0.88	0.90	0.93	0.92
IOA	0.91	0.92	0.93	0.97	0.97	0.98	0.95	0.95	0.97	0.97	0.98	0.98
NSE	0.69	0.72	0.75	0.88	0.89	0.92	0.81	0.82	0.87	0.90	0.93	0.92
PI	-3.4	-3.0	-2.6	-1.2	-0.82	-0.4	-3.5	-3.2	-2.0	-3.0	-1.8	-2.1
VE	0.54	0.55	0.55	0.67	0.68	0.69	0.60	0.62	0.66	0.69	0.71	0.71
KGE	0.84	0.86	0.87	0.94	0.95	0.96	0.90	0.91	0.94	0.95	0.96	0.96

4.1.2 Parameter sensitivity

Results of the sensitivity analysis are plotted in Fig. 4. The sensitivity of GR4H to X4, the reference time of the unit hydrograph, was the lowest with TSI always under 0.01. The sensitivity of X1, the size of the production store, was also relatively small, except for Dumbéa Est, where it was 0.36. X2, the underground exchange coefficient, and X3, the size of the routing store, were found to be the most sensitive parameters, with TSI lying between 0.33 and 0.77. X3 and X4 sensitivity did not exhibit differences between catchments, whereas X2 showed a wide spread in its values. TSI of X2 was highest for Coulée (0.77) and lowest for Dumbéa Est (0.33).

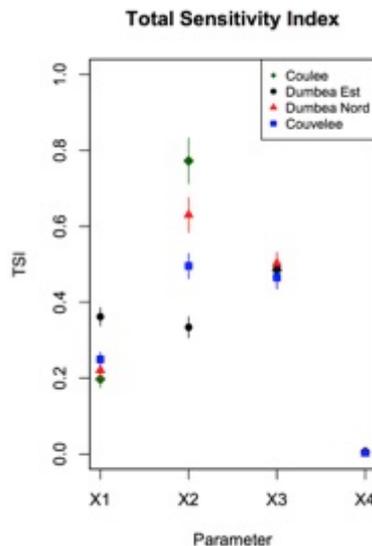


Fig. 4: Total sensitivity index computed for each river: Coulée (green), Dumbéa Est (black), Dumbéa Nord (red), Couvelée (blue). Means (dots) are plotted with confidence intervals (segments).

4.1.3 Length of the calibration period and optimal parameters set

Results of the experiments introduced in section 3.3.2 are presented on Fig. 5 and Fig. 6. For each calibration length, and for each watershed, 12 parameter sets were calibrated. Their mean was plotted together with the deviations from the benchmark values on Fig. 5. Changes in the mean of the parameters provided information about calibration stability and its evolution with increasing length of dataset used for

calibration. From an overall perspective, the parameters stabilised when the calibration period length increased and, with 11 years of data, whatever the parameter and watershed, the deviation from the benchmark value was less than 23%.

The stability of the calibration was found to be site specific. The parameter values generally stabilised within 6-7 years, with deviation from the benchmark value below 50%, except for Coulée. For this watershed, X2 underwent large variations as calibration length increased and 10 years were needed before reaching less than 50% of deviation. Also, X3 experienced large variation, with the mean value around 280 mm with 1 year of calibration, compared to a final value of 172 mm. Dumbéa Est and Dumbéa Nord showed relatively large variations of X1, before reaching stability at 6-7 years of calibration.

To assess the significance of these variations on the model quality, the KGE was computed during calibration and validation, as detailed in section 3.3.2. For each watersheds, the KGE at calibration and at validation converged toward the same value. For Dumbéa Est, Dumbéa Nord and Couvelée, KGE values were almost identical at calibration and validation after 6-7 years. After 7 years, the KGE remained almost constant. However, for Coulée, 10 years were needed for the model quality to converge at calibration and validation toward 0.7. In any case, with 11 years of data, the quality was almost identical in calibration and validation. Lastly, likely because of over-calibration (Andréassian *et al.*, 2012), for all watersheds, the KGE was always the highest at calibration with 1 year of data.

The parameter set obtained with 12 years of calibration is presented in Table 3. The underground loss rate (-X2) was lowest on Coulée and Dumbéa Est, where it was lower than 0.8, while it was higher than 1.6 on the other watersheds. The production store was almost non-existent on Dumbéa Nord and Dumbéa Est, where its size (X1) was lower than 14 mm, while it reached almost 96 mm on Coulée and 536 mm on Couvelée. The size of the routing store (X3) was lowest on Coulée, with approximately 172 mm, while it was higher than 225 mm on the other watersheds. It was highest on Dumbéa Nord, where it almost reached 350 mm. Eventually, X4 was close to 3 for Coulée, Dumbéa Nord and Couvelée and equal to 2 for Dumbéa Est.

Table 3: Optimal parameters set with rainfall computed on the basis of the IDEW algorithm and 12 years of calibration. The absolute deviation from the benchmark value at 11 years is shown in parentheses.

	Coulée	Dumbéa Est	Dumbéa Nord	Couvelée
X1 (mm)	99.6 (2.7)	13.7 (0.7)	5.5 (0.8)	587.7 (2.1)
X2 (-)	-0.196 (0.05)	-0.78 (0.002)	-1.67 (0.05)	-6.18 (0.01)
X3 (mm)	171.6 (6.2)	263.3 (0.7)	344.3 (0.5)	227.87 (0.9)
X4 (h)	2.99 (0.08)	2 (0.008)	2.98 (0.008)	3.1 (0.018)

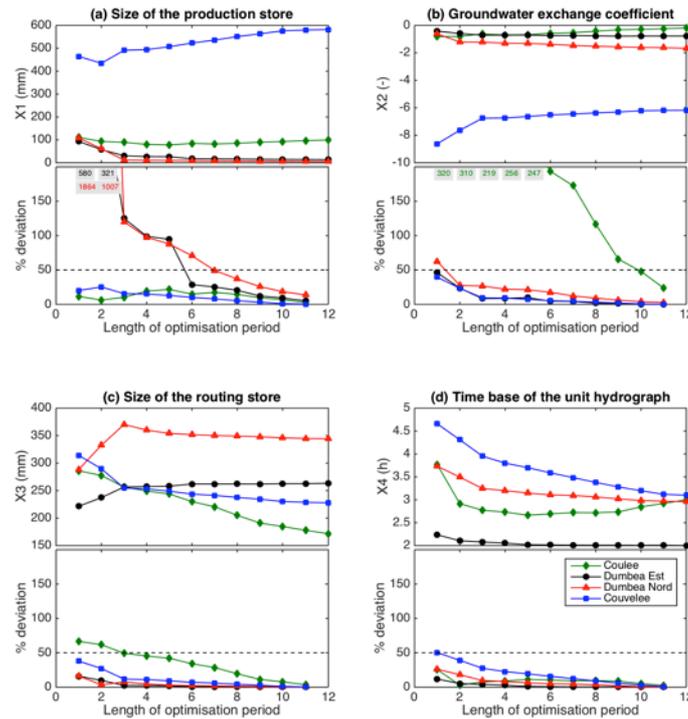


Fig. 5: Value of the optimal parameters for increasing length of calibration period and percentage of deviation from their benchmark values.

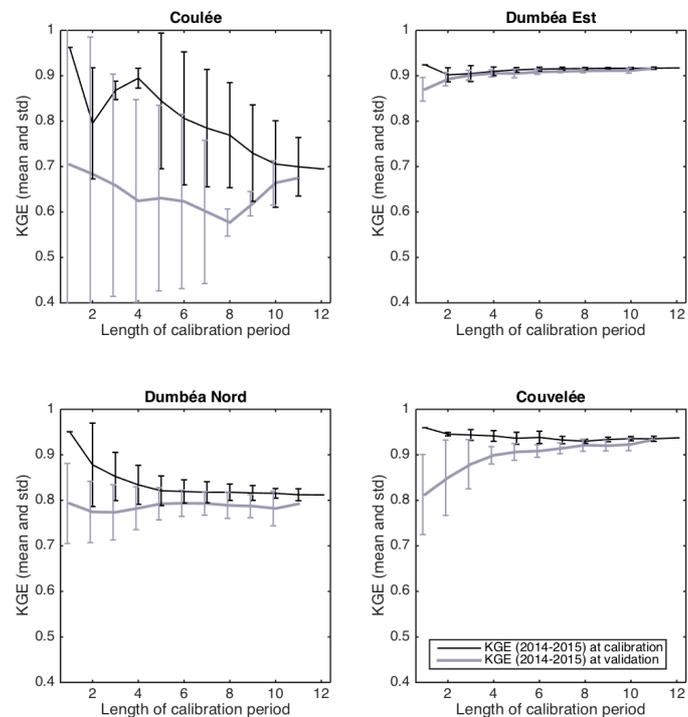


Fig. 6: Value of the KGE computed in 2014-2015 with different lengths of calibration period. The results obtained when 2014-2015 was within the calibration period (black) were plotted separately from results obtained when 2014-2015 was within the validation period (grey).

4.2 Analysis of the model's results

4.2.1 Measured flash flood description

The selection of an appropriate rainfall interpolation method, presented in section 4.1.1, allows the description of the most significant flood peaks. To remain consistent with the next results, the described flood peaks are analysed after 1 October 2003. The period covered is 12 years, and so the 12 largest floods were selected. The resulting metrics, based on measurements, are listed in Table 4. Median amount of rainfall during flood was highest for Couvelée, with 304 mm of precipitation, and lowest for Dumbéa Nord with 209 mm. The mean amount of runoff was, however, highest for Dumbéa Est, with more than 210 mm, and lowest for Couvelée, with only 140 mm. Whatever the watershed considered, the mean volume produced during one flood was between 10 % and 15 % of the annual runoff. The mean 24h flow shape coefficient, representing the flashiness of the floods, was higher for Dumbéa Est and Coulee (more than 2.5) than for Dumbéa Nord and Couvelée (less than 1.9).

Table 4: Median flash flood description on the watersheds studied.

	Coulee	Dumbéa Est	Dumbéa Nord	Couvelée
Amount of precipitation (mm)	233	245	209	304
Amount of runoff during flood (mm)	169	214	196	140
24h flow shape coefficient (-)	2.52	2.95	1.64	1.85

4.2.2 Flash flood simulation

Direct comparison of measured and simulated runoffs during the 5 largest floods of the study period for each of the studied watersheds is shown in Fig. 7. Except in the case of Coulee in January 2011, the flood peaks were generally well represented, both in terms of amplitude and timing of the flood peak.

In order to get a broader and more precise overview of flood simulation quality, the metrics defined in section 3.1.3 were computed on the 12 largest floods and are presented in Table 5 and Fig. 8. The mean peak flow error was highest for Coulee and Dumbéa Est, where, on average, they were overestimated by 17% and underestimated by -13%, respectively. On the other basins, they were simulated with less than 8% error. The flood volume was overestimated by 16% on average for Coulee, while it was underestimated on the other basins with less than 8% error. The falling limb was systematically underestimated, with errors between -3% and -12.4%. The rising limb was generally better represented, except for Coulee, where the flood volume was overestimated by around 20%.

The best flood peak timing was achieved on Dumbéa Est where 92% of the selected floods peaks were simulated with a time-to-peak error lower than or equal to 2h, while it was worst on Couvelée, where only 66% of the floods were simulated with less than 2h of delay. Lastly, only 2 flood peaks were simulated with positive time-to-peak error (i.e. advance).

Table 5: Evaluation metrics of the model's ability to reproduce flash floods.

	Coulée	Dumbéa Est	Dumbéa Nord	Couvelée
Mean peak flow error (%)	17.2	-13.2	2.9	-7.9
Mean RVE (%)	16.2	-4.0	-7.2	-4.2
Mean RVE during rising limb (%)	19.2	-0.4	5.3	7.0
Mean RVE during falling limb (%)	-3.0	-3.6	-12.4	-11.2

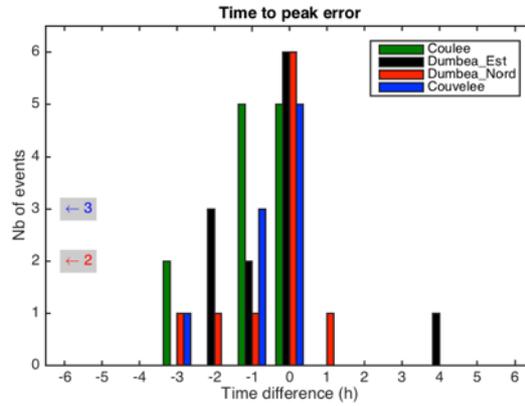


Fig. 7: Histogram of the time-to-peak error distribution for Coulée (a), Dumbéa Est (b), Dumbéa Nord (c), Couvelée (d).

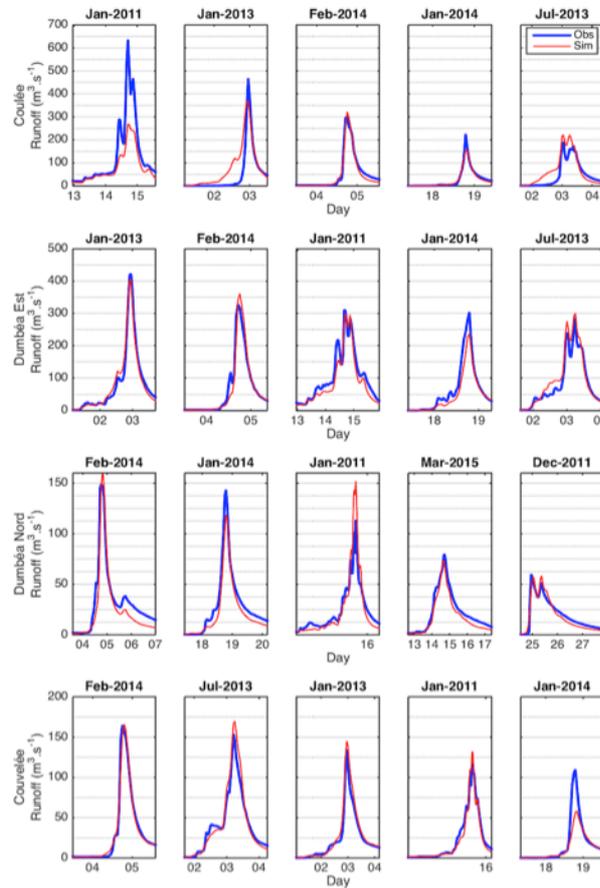


Fig. 8: Comparison of observed (blue) and simulated (red) runoffs during the 5 largest floods of the study period for each of the watersheds studied.

4.2.3 Annual runoff simulation

The annual runoff was calculated and simulations compared to measurements in Fig. 9. The measured and simulated annual runoffs covaried, with variations in line with the climatic context described in section 2. The difference between dry years and wet years was considerable. For example, the annual runoff was more than 3.5 times larger for Couvelée between 2010 and 2011. The sinusoidal form of the annual streamflow, and its relation with ENSO phases, was more visible from 2004 to 2011, as the number and amplitude of floods remained low compared to the 2011-2015 period, when floods represented a significant proportion of the annual runoff. In terms of volume, the largest flood occurred on all watersheds in January 2011, after the passage of the Vania tropical cyclone.

Simulated runoffs generally reproduced these variations. Inter-basin differences could be noted, especially between burned, i.e. Coulée and Dumbéa Est, and unburned watersheds, i.e. Couvelée and Dumbéa Nord. For Couvelée and Dumbéa Nord, GR4H reproduced the variation of the annual runoff, but the magnitude of the variation was underestimated, which led to general underestimation during wet years and overestimation during dry years. In particular, the runoff was overestimated from 2004 to 2008 and from mid-2012 to mid-2013 and underestimated in 2009, 2011 and in 2014. For these watersheds, the error was generally contained with +/- 15% of annual runoff, except for particularly dry years, like in 2010 for Couvelée, where annual runoff was only around 350 mm while the simulation was 500 mm.

In the case of Coulée and, to a lesser extent, Dumbéa Est, the annual runoff was overestimated from 2004 to mid-2006 by around 20% and 15%, respectively, and between 2012 and mid-2013 by around 15% and 10%, respectively. On these watersheds, it was also underestimated from mid-2006 to 2009, in 2011 and in 2014 by around -20% and -10%. This pattern was significantly different from what occurred on Dumbéa Nord and Couvelée, especially from mid-2006 to 2009, when runoff from unburned basins was overestimated, whereas it was slightly underestimated on burned basins. The difference was highest from mid-2006 to early 2007, when underestimation amounted to -20% and -8% for Coulée and Dumbéa Est, respectively, while for Dumbéa Nord and Couvelée the runoff was overestimated by more than 9%. The inter-basin difference was less clear after 2010, as the annual runoff was underestimated for all basins in 2011.

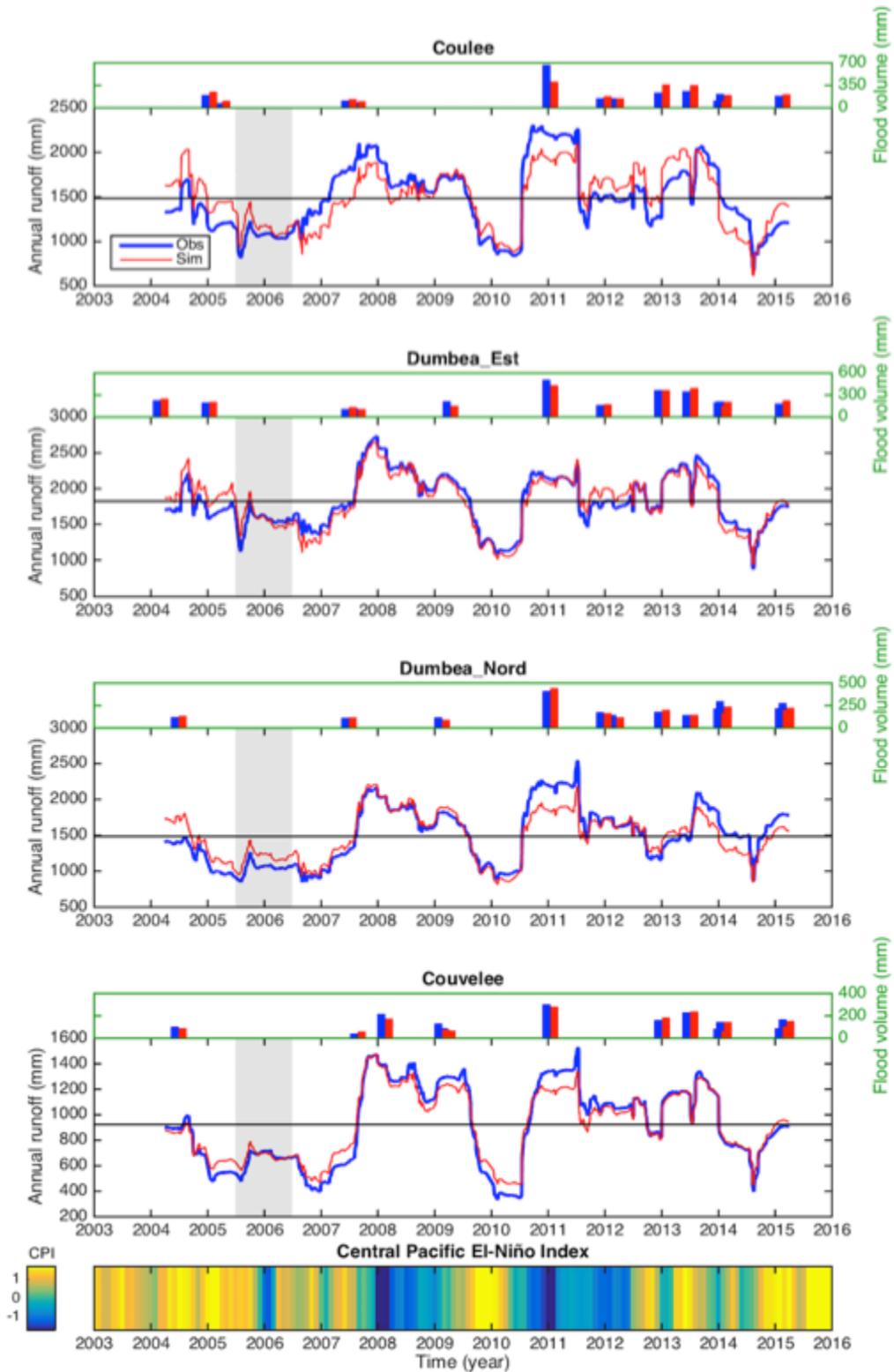


Fig. 9: Annual runoff expressed in millimetres computed from observations (blue) and simulations (red). Average over the study period is represented for each watershed by a black line. The grey strip refers to the period ranging from 6 months before to 6 months after the wildfire in late 2005 that burned 59% of Coulee and 15% of Dumbéa Est.

5 Discussion

5.1 Model Calibration

To represent more realistically the pathways of freshwater in integrated models of tropical flash-flood prone basins, this study proposed an original approach that allowed to, i) provide an estimate of the record length needed to obtain robust calibration of model parameters, ii) present the model quality during calibration and validation periods and assess the suitability of merging these periods.

The method relies on several k-fold partitioning tests, with increasing calibration length. It may be seen as a generalisation of split-sample tests (e.g. Klemeš *et al.*, 1986). While few general guidelines are found in the literature on the record length needed for calibration of a hydrologic model (Perrin *et al.*, 2007), the proposed method involves calibration lengths ranging from 1 year to the maximum length of the hydrological record. Site-by-site sensitivity analysis allowed us to distinguish parameters with high TSI values, which need to be estimated as precisely as possible, and insensitive parameters with TSI under 0.2, whose estimation can be coarse. In this study, X4 was deemed insensitive, which is in line with the literature (Shin *et al.*, 2013). In addition, X1 was almost insensitive, except for Dumbéa Est, and X2 was a critical parameter on Coulée. The minimum length of calibration period was estimated on the basis of these results. Overall, the calibration tended to stabilise within 6-7 years of input data. After 6 years of calibration, Coulée still underwent relatively large variations for X2 and X3, which were sensitive parameters, and 10 years were needed before they reached relative stability. It is therefore very likely that the hydrological processes of watersheds not exposed to wildfire are well described with ~ 6 years of data, with potential influence of El Niño oscillation, which follows an irregular cycle of 2 to 7 years and is one of the strongest climate driver of the region (see S1). As wildfires are known to severely change soil properties and deplete infiltration (Moody *et al.*, 2013), they have the potential to alter hydrological processes of a watershed for at least 3 years, as suggested by the Coulée results (section 4.2.4). That would explain why on this watershed, more data was needed to account for additional variability of runoff processes.

In the second step, the model quality during calibration period was made distinct from validation period and the equivalence of model results during these periods was examined. Results showed that, with short calibration periods, model quality during validation was significantly lower than during calibration, which is a classical result (Andréassian *et al.*, 2012). However, with sufficient data, the parameters and results are independent of the calibration period chosen, whatever the watershed studied. This suggests that the approach chosen provide robust results for long-term simulations in the region. In addition, the approach presented here determined both robust parameter sets and a methodology that might be generalizable, after testing on wide samples of watersheds.

5.2 Performance of GR4H

The model's ability to reproduce the measured runoffs was assessed using several efficiency metrics, which is a recommended way to estimate model quality (e.g. Bennett *et al.*, 2013, Hrachowitz *et al.*, 2013). However, translation of numerical values into subjective indicators of model quality is not straightforward and direct comparison with the literature can produce puzzling results. For example, during calibration, according to the hydrological classification of Dawson *et al.* (2007), the Coulée simulation should be rated as “poor” on NSE and PI metrics (0.75 and -2.6 resp.), “satisfactory” on Rsqr (0.76) and “good” on the IOA value (0.93). However, the overall model quality was deemed satisfactory when compared to the literature (Mathevet, 2005, Lerat *et al.*, 2012, van Esse *et al.*, 2013) and for future developments of our work. This result was not unexpected. Indeed, the GR models are lumped rainfall-runoff models that have been applied to several hundred catchments over a wide range of countries and climates (e.g. Mathevet, 2005, Harlan *et al.*, 2010, Coron *et al.*, 2012, Boumenni *et al.*, 2017, Gaba *et al.*, 2017) and GR4H has already been used for simulating flash floods (Bennett *et al.*, 2014, Ficchi *et al.*, 2016). Moreover, the most cited limitation of GR4H is its poor performance during low flows (Pushpalatha, *et al.*, 2011). Due to humid climate, and to general purposes of the modelling exercise, this had little effects in this study.

The overall quality of the models was further confirmed by the analysis of flash floods and annual runoff simulations. On average, all peak flow errors were less than 18.3% and the total volume error was less than 16.8%, which is satisfactory compared to the uncertainty associated with observations in torrential flood conditions (e.g. Harmel *et al.*, 2006, Baldassarre and Montanari, 2009). Furthermore, the time-to-peak error is generally less than 2 hours. This is one order of magnitude lower than what had already been reported for GR models, where it can exceed one day (e.g. Ficchi *et al.*, 2016 on France catchments). This finding might be due to the high speed of flash floods. The 24 h flow shape coefficient was particularly high, compared to values obtained in temperate countries (Ficchi *et al.*, 2016) and the delay between rainfall and runoff, given by the X4 parameter, was around 2-3 hours, which is almost 2 times lower than the common definition of flash floods (Hapuarachchi *et al.*, 2011). High inter-annual variations (> 350%) of the measured runoff volume were reproduced, generally with less than 15% of error. For unburned watersheds, the annual runoff was generally best reproduced in 2008, 2012 and 2013, which were wet years. This was also noted by (Hapuarachchi *et al.*, 2011) who suggest that hydrological processes are easier to account for in humid areas than in dry areas.

While the annual runoff was overestimated by 10% on watershed unexposed to wildfire, it was underestimated by -20% in early 2007 in Coulée, just after the massive wildfire that burned around 60% of its area. It is therefore likely that, even though the calibration procedure provided robust estimates of the parameters, the simulation exercise had difficulty in accounting for large land-cover changes, which is a known issue of lumped models (e.g. Devi *et al.*, 2015). Indeed, in the literature, the impact of land-cover change on hydrology is usually grasped through variations of parameters

after calibration on sub-periods (Andréassian *et al.*, 2003, Thirel *et al.*, 2015), that can be relatively long for lumped models. Therefore, such methods may encounter difficulties with rapid land-cover changes, such as those triggered by wildfires (Folton *et al.*, 2015). Instead, in this study, the impact of wildfires was based on parameters averaged over the whole study period. In fact, sudden land-cover changes were studied as a source of uncertainty. Doing so yielded valuable results and showed that the impact of wildfires on the annual runoff was visible and quantifiable for at least 3 years, which is consistent with temporal scales of resilience induced by such disturbances (Moody *et al.*, 2013).

5.3 Rainfall interpolation

The use of the IDEW algorithm, integrating topography, offered the best performance for the spatial interpolation of rainfall, thereby confirming that it is well adapted to mountainous terrain (Ly *et al.*, 2013), where orography effects significantly impact rainfall budgets (Terry and Wotling, 2011). In particular, this algorithm outperformed the classical Thiessen method, and the annual amount of rainfall falling on each watershed (Table 1) was generally in line with previous studies (Terry and Wotling, 2011). The algorithm worked with a limited dataset, which is an asset in tropical regions and developing countries, where data is often scarce (Sampson *et al.*, 2015, Rude *et al.*, 2016). In particular, this limitation prevented the use of kriging or co-kriging methods, which are often used for interpolating rainfall in mountainous areas (Ly *et al.*, 2013). However, rainfall in the region is known to be highly variable, spatially and temporally, and some punctual events might have not been very accurately represented. This may have been the case in early 2011, after the passage of the Vania cyclone on Coulée. Further analysis revealed that, during this event, errors in observations may have compromised the modelling exercise, as the total amount of water flowing through the limnimeter was higher than the amount of measured rainfall (whatever the instrument analysed). Future studies will clearly benefit from extensive measurements of rainfall fields. Ongoing local development of precise atmosphere models is also a promising perspective for producing accurate estimation of rainfall spatial patterns, which might bring valuable information for expanding punctual time series into gridded products (e.g. Teufel *et al.*, 2017).

5.4 Limitations and perspectives

All parameters values presented large variations between contiguous watersheds, except for X4, which was relatively constant. This finding suggests that, in New Caledonia, future efforts in terms of modelling or of prediction in ungauged basins should be cautious regarding a proximity-based approach and should consider the history and composition of land-cover (Hrachowitz *et al.*, 2013). However, as only 4 watersheds were considered here, parameters could not easily be linked to observed characteristics of the catchments. One way to overcome this limitation would be to rely on a large sample of basins in order to draw statistical relationships between parameters and descriptors of the basins (Hrachowitz *et al.*, 2013). In New Caledonia, a good study area for this purpose would be the eastern part of the Grande Terre, where hydro-climatic conditions and topography features are relatively homogenous (Terry and

Wotling, 2011), and where extensive mining operations have greatly altered the landscape.

Because it has been established that most of the rain falls on mountains (e.g. Terry and Wotling, 2011), it is believed that the most of the runoff that eventually reaches the lagoon was captured and simulated. However, it is estimated worldwide that around 10% of the freshwater received by the ocean comes as underground water (Burnett *et al.*, 2006), and peridotites, forming the bedrock of the study area, are known to present high hydraulic conductivity aquifer layers at shallow depth (Join *et al.*, 2005). Hence, a significant proportion of the rainfall may have exited the watersheds through subsurface flow. This might explain why the Couvelée runoff was particularly low, compared to the rainfall entering the watershed. Also, the runoff from small, low watersheds was also neglected. Future integrated modelling of hydrology with lagoon hydrodynamics compared to precise measurements of the lagoon salinity may allow to reduce some of these uncertainties.

6 Conclusions

To our knowledge, this study contains the first published results of rainfall-runoff modelling in New Caledonia. The approach and parameters adopted here provided reliable simulations and seem to be promising for coupling with hydrodynamics for long hindcasts of the lagoon dynamics. When sufficient data was available, the model's results were almost identical during the calibration and validation periods. Except for the Coulée basin, which was extensively burned during the study period, NSE criteria were around 0.9 and KGE metrics around 0.95. The largest floods were generally simulated with error magnitude in the range of measurement uncertainty; however, the flood triggered by cyclone Vania could not be represented accurately, likely because of rainfall underestimation and/or runoff overestimation during measurement. The flood timing error was generally lower than 2h and the general shape of runoff correctly simulated. The GR4H model demonstrated its ability to work on watersheds with contrasting runoff coefficients (41% for Couvelée and 81% for Dumbéa Est) and with annual rainfall changes of more than 300% (between El Niño dry years and La Niña wet years). While lessons have been learned about the influence of land-cover shifts and climatic variability, the GR4H model and the calibration procedure proposed here have produced accurate results. Coupling with the lagoon hydrodynamics will be the next step toward an integrated ridge-to-reef model.

7 Acknowledgements

The authors thank the SDE (Service De l'Eau) and especially Geoffroy Wotling and Nicolas Romieux from DAVAR (Direction des Affaires Vétérinaires, Alimentaires et Rurales) for the provision of flow and rainfall datasets and for valuable discussions about hydrology and rainfall-runoff modelling. They also thank Meteo-France New Caledonia for supplementary precipitation data and Nicolas Le Moine from UPMC (Université Pierre et Marie Curie) for providing the GR4H code.

8 References

- Andréassian, V., Parent, E. & Michel, C., 2003. A distribution-free test to detect gradual changes in watershed behavior. *Water Resources Research*, **39**(9).
- Andréassian, V., *et al.*, 2012. All that glitters is not gold: the case of calibrating hydrological models: Invited Commentary. *Hydrological Processes*, **26**(14), 2206–2210.
- Arfi, R., Guiral, D. & Bouvy, M., 1993. Wind Induced Resuspension in a Shallow Tropical Lagoon. *Estuarine, Coastal and Shelf Science*, **36**(6), 587–604.
- Baldassarre, G. D. & Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. *Hydrology and Earth System Sciences*, **13**(6), 913–921.
- Bennett, N. D. *et al.*, 2013. Characterising performance of environmental models. *Environmental Modelling & Software*, **40**, 1–20.
- Bennett, J. C., *et al.* 2014. A System for Continuous Hydrological Ensemble Forecasting (SCHEF) to lead times of 9 days. *Journal of Hydrology*, **519**, 2832–2846.
- Boumenni, H., Bachnou, A. & Alaa, N. E., 2017. The rainfall-runoff model GR4J optimization of parameter by genetic algorithms and Gauss-Newton method: application for the watershed Ourika (High Atlas, Morocco). *Arabian Journal of Geosciences*, **10**(15), 343.
- Brodie, J. E., *et al.*, 2012. Terrestrial pollutant runoff to the Great Barrier Reef: An update of issues, priorities and management responses. *Marine Pollution Bulletin*, **65**(4–9), 81–100.
- Burke, L., Reyntar, K., Spalding, M. & Perry, A., 2011. *Reefs at risk revisited*. Washington, DC: World Resources Institute.
- Burnett WC, *et al.*, 2006. Quantifying Submarine Groundwater Discharge in the Coastal Zone via Multiple Methods. *Science of the Total Environment*, **367**(1–2), 498–543
- Coron, L., Andréassian, V. & Perrin, C., 2012. Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments. *Water Resources Research*, **48**(5).
- Coron, L. *et al.*, 2017. The suite of lumped GR hydrological models in an R package. *Environmental Modelling & Software*, **94**, 166–171.
- Daly, C. *et al.*, 2002. A knowledge-based approach to the statistical mapping of climate. *Climate Research*, **22**, 99–113.
- David, G. *et al.*, 2010. Integrated coastal zone management perspectives to ensure the sustainability of coral reefs in New Caledonia. *Marine Pollution Bulletin*, **61**(7–12), 323–334.
- Dawson, C. W., Abrahart, R. J. & See, L. M., 2007. HydroTest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environmental Modelling & Software*, **22**(7), 1034–1052.
- Devlin, M. J. & Brodie, J., 2005. Terrestrial discharge into the Great Barrier Reef Lagoon: nutrient behavior in coastal waters. *Marine Pollution Bulletin*, **51**(1–4), 9–22.
- Devi, G. K., Ganasri, B. P. & Dwarakish, G. S., 2015. A Review on Hydrological Models. *Aquatic Procedia*, **4**, 1001–1007.
- Doney, S. C., 2010. The Growing Human Footprint on Coastal and Open-Ocean Biogeochemistry. *Science*, **328**(5985), 1512–1516.
- Dresback, K. M. *et al.*, 2013. Skill assessment of a real-time forecast system utilizing a coupled hydrologic and coastal hydrodynamic model during Hurricane Irene (2011). *Continental Shelf Research*, **71**, 78–94.
- Fernandez, J.-M. *et al.*, 2006. A combined modelling and geochemical study of the fate of terrigenous inputs from mixed natural and mining sources in a coral reef lagoon (New Caledonia). *Marine Pollution Bulletin*. **52**(3), 320–331.
- Ficchi, A., Perrin, C. & Andréassian, V., 2016. Impact of temporal resolution of inputs on hydrological model performance: An analysis based on 2400 flood events. *Journal of Hydrology*, **538**, 454–470.
- Folton, N., Andréassian, V. & Duperray, R., 2015. Hydrological impact of forest-fire from paired-catchment and rainfall-runoff modelling perspectives. *Hydrological Sciences Journal*, **60**(7–8), 1213–1224.
- Gaba, C., Alamou, E., Afouda, A. & Diekkrüger, B., 2017. Improvement and comparative assessment of a hydrological modelling approach on 20 catchments of various sizes under different climate conditions. *Hydrological Sciences Journal*, **62**(9), 1499–1516.
- Gomez, C., *et al.*, 2015. Wildfire risk for main vegetation units in a biodiversity hotspot: modeling approach in New Caledonia, South Pacific. *Ecology and Evolution* **5**(2), 377–390.
- Gupta, H. V. *et al.*, 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, **377**(1–2), 80–91.
- Hansen, M. C. *et al.*, 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, **342**(6160), 850–853.

- Harmel, R. D. *et al.*, 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Transactions of the ASABE*, **49**(3), 689–701.
- Hapuarachchi, H. A. P., Wang, Q. J. & Pagano, T. C., 2011. A review of advances in flash flood forecasting. *Hydrological Processes*, **25**(18), 2771–2784.
- Harlan, D., Wangsadipura, M. & Munajat, C., 2010. Rainfall-Runoff Modeling of Citarum Hulu River Basin by Using GR4J. *Proceedings of the World Congress on Engineering*, **2**.
- Hollingsworth, A., 1994. Validation and diagnosis of atmospheric models. *Dynamics of Atmospheres and Oceans*, **20**(3), 227–246.
- Holt, J., Wakelin, S., Lowe, J. & Tinker, J., 2010. The potential impacts of climate change on the hydrography of the northwest European continental shelf. *Progress in Oceanography*, **86**(3–4), 361–379.
- Hrachowitz, M. *et al.*, 2013. A decade of Predictions in Ungauged Basins (PUB)—a review. *Hydrological Sciences Journal*, **58**(6), 1198–1255.
- Inoue, M. *et al.*, 2008. A high-resolution integrated hydrology-hydrodynamic model of the Barataria Basin system. *Environmental Modelling & Software*, **23**(9), 1122–1132.
- Join, J.-L., Robineau, B., Ambrosi, J.-P., Costis, C., Colin, F., 2005. Système hydrogéologique d'un massif ultrabasique de Nouvelle-Calédonie. *Comptes Rendus Geoscience*, **337**(16), 1500–1508.
- Klemeš, V., 1986. Operational testing of hydrological simulation models. *Hydrological Sciences Journal* **31**(1), 13–24.
- Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence*, **14**, 1137–1145.
- Lerat, J. *et al.*, 2012. Do internal flow measurements improve the calibration of rainfall runoff models? *Water Resources Research*, **48**(2).
- Li, H., Beldring, S. & Xu, C.-Y., 2015. Stability of model performance and parameter values on two catchments facing changes in climatic conditions. *Hydrological Sciences Journal*, **60**(7–8), 1317–1330.
- Luersen, M. A. & Le Riche, R., 2004. Globalized Nelder–Mead method for engineering optimization. *Computers and Structures* **82**(23–26), 2251–2260.
- Ly, S., Charles, C. & Degré, A., 2013. Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. A review. *Biotechnology, Agronomy, Society and Environment* **17**(2), 392.
- Maggioni, V. & Massari, C., 2018. On the performance of satellite precipitation products in riverine flood modeling: A review. *Journal of Hydrology*, **558**, 214–224.
- Manabe, S., Milly, P. C. D. & Wetherald, R., 2004. Simulated long-term changes in river discharge and soil moisture due to global warming / Simulations à long terme de changements d'écoulement fluvial et d'humidité du sol causés par le réchauffement global. *Hydrological Sciences Journal* **49**(4), null-642.
- Marler, R. T. & Arora, J. S., 2004. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization* **26**(6), 369–395.
- Masih, I. *et al.*, 2011. Assessing the Impact of Areal Precipitation Input on Streamflow Simulations Using the SWAT Model. *Journal of the American Water Resources Association*, **47**(1), 179–195.
- Mathevet, T., 2005. *Quels modèles pluie-débit globaux au pas de temps horaire? Développement empiriques et comparaison de modèles sur un large échantillon de bassins versants*. Thesis (PhD). Ecole Nationale du Génie Rural, des Eaux et des Forêts.
- Merheb, M., *et al.*, 2016. Hydrological response characteristics of Mediterranean catchments at different time scales: a meta-analysis. *Hydrological Sciences Journal*, **61**(14), 2520–2539.
- Monod, H., 2006. Uncertainty and sensitivity analysis for crop models. In: Naud, C. & Makowski, D. eds. *Working with dynamic crop models: evaluation, analysis, parameterization, and applications*. Amsterdam: Elsevier, 55–100.
- Moody, J., Shakesby, R. & Robichaud, P., 2013. Current research issues related to post-wildfire runoff and erosion processes. *Earth-Science Reviews* **122** 10-37.
- Morlière, A. & Rebert, J.-P., 1986. Rainfall shortage and El Nino–Southern Oscillation in New Caledonia, southwestern Pacific. *Monthly Weather Review*, **114**(6), 1131–1137.
- Myers, N. *et al.*, 2000. Biodiversity hotspots for conservation priorities. *Nature*, **403**(6772), 853–858.
- Nalder, I. & Wein, R., 1998. Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. *Agricultural and Forest Meteorology*, **92**, 211–225.
- Nelder, J. A. & Mead, R., 1965. A Simplex Method for Function Minimization. *The Computer Journal*, **7**(4), 308–313.
- OEIL, 2012. *Suivi de l'évolution des paysages de 1998 à 2010 en Province Sud*. Nouméa: Observatoire de l'environnement.
- Oudin, L. *et al.*, 2006. Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models. *Journal of Hydrology* **320**(1–2), 62–83.

- Ouillon, S. *et al.*, 2008. Optical Algorithms at Satellite Wavelengths for Total Suspended Matter in Tropical Coastal Waters. *Sensors*, **8**(7), 4165–4185.
- Perrin, C., Michel, C. & Andréassian, V., 2003. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, **279**(1–4), 275–289.
- Perrin, C., Oudin, L. & Andréassian, V., 2007. Impact of limited streamflow data on the efficiency and the parameters of rainfall—runoff models. *Hydrological Sciences Journal*, **52**(1), 131–151.
- Petus, C. *et al.*, 2016. Estimating the Exposure of Coral Reefs and Seagrass Meadows to Land-Sourced Contaminants in River Flood Plumes of the Great Barrier Reef: Validating a Simple Satellite Risk Framework with Environmental Data. *Remote Sensing*, **8**(3), 210.
- Prats, S., 2009. *Table GRILLE_ETP_Q_NC Structure et fonctionnement*. Nouméa: Météo-France.
- Pujol, G. *et al.*, 2017. *Global Sensitivity Analysis of Model Outputs*. [R package]. Available from <https://CRAN.R-project.org/package=sensitivity>
- Pushpalatha, R., Perrin, C., Le Moine, N., Mathevet, T. & Andréassian, V. (2011) A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. *Journal of Hydrology*, **411**(1–2), 66–76.
- Richmond, R. H. *et al.*, 2007. Watersheds and coral reefs: conservation science, policy, and implementation. *BioScience*, **57**(7), 598–607.
- Roberts, C. M. *et al.*, 2002. Marine Biodiversity Hotspots and Conservation Priorities for Tropical Reefs. *Science*, **295**(5558), 1280–1284.
- Rude, J. *et al.*, 2016. Ridge to reef modelling for use within land-sea planning under data-limited conditions: Ridge to reef modelling under data-limited conditions. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **26**(2), 251–264.
- Sampson, C. C., *et al.*, 2015. A high-resolution global flood hazard model. *Water Resources Research*, **51**(9), 7358–7381.
- Shin, M.-J. *et al.*, 2013. Addressing ten questions about conceptual rainfall—runoff models with global sensitivity analyses in R. *Journal of Hydrology*, **503**, 135–152.
- Smith, L. C., Turcotte, D. & Isacks, B., 1998. Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrological Processes*, **12**, 233–249.
- Sobol, I., 1993. Sensitivity estimates for nonlinear mathematical models. *Mathematical Model Computing Experiment*, **1**(4), 407–414.
- Terry, J., Kostaschuk, R. A. & Wotling, G., 2008. Features of tropical cyclone-induced flood peaks on Grande Terre, New Caledonia. *Water and Environment Journal*, **22**(3), 177–183.
- Terry, J. & Wotling, G., 2011. Rain-shadow hydrology: Influences on river flows and flood magnitudes across the central massif divide of La Grande Terre Island, New Caledonia. *Journal of Hydrology*, **404**(1–2), 77–86.
- Teufel, B., *et al.*, 2017. Investigation of the 2013 Alberta flood from weather and climate perspectives. *Climate Dynamics*, **48**(9–10), 2881–2899.
- Thirel, G. *et al.*, 2015. Hydrology under change: an evaluation protocol to investigate how hydrological models deal with changing catchments. *Hydrological Sciences Journal*, **60**(7–8), 1184–1199.
- Torréon, J.-P. *et al.*, 2010. Variability of primary and bacterial production in a coral reef lagoon (New Caledonia). *Marine Pollution Bulletin*, **61**(7–12), 335–348.
- Trescases J.-J. 1975. *L'évolution géologique supérogène des roches ultrabasiqes en zone tropicale. Formation des gisements nickélicifères de Nouvelle-Calédonie*. Thesis (PhD). Université Louis Pasteur de Strasbourg.
- van Esse, W., *et al.*, 2013. The influence of conceptual model structure on model performance: a comparative study for 237 French catchments. *Hydrology and Earth System Sciences*, **17**(10), 4227–4239.
- Vincent, E., *et al.*, 2011. Interannual variability of the South Pacific Convergence Zone and implications for tropical cyclone genesis. *Climate Dynamics*, **36**, 1881–1896.