Journal of Hydrology 519 (2014) 3436–3447

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Rainfall-runoff modeling in a flashy tropical watershed using the distributed HL-RDHM model

Ali Fares [a,*], Ripendra Awal [a], Jene Michaud [b], Pao-Shin Chu [c], Samira Fares [d], Kevin Kodama [e], Matt Rosener [f]

[a] College of Agriculture and Human Sciences, Prairie View A&M University, Prairie View, TX, USA
[b] Geology Department, University of Hawai'i-Hilo, HI, USA
[c] Department of Meteorology, University of Hawai'i-Manoa, HI, USA
[d] University of Hawai'i-Manoa, HI, USA
[e] Hawaii Weather Forecast Office, National Weather Service, HI, USA
[f] North Shore Hydro Consulting, Kauai, HI, USA

ARTICLE INFO

Article history:
Received 19 November 2013
Received in revised form 20 June 2014
Accepted 17 September 2014
Available online 7 October 2014
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Purna Chandra Nayak, Associate Editor

Keywords:
HL-RDHM
SAC-SMA
Rainfall-runoff
Floods
Modeling

SUMMARY

Many watersheds in Hawai'i are flash flood prone due to their small contributing areas and frequent intense rainfall. Motivated by the possibility of developing an operational flood forecasting system, this study evaluated the performance of the National Weather Service (NWS) model, the Hydrology Laboratory Research Distributed Hydrologic Model (HL-RDHM) in simulating the hydrology of the flood-prone Hanalei watershed in Kaua'i, Hawai'i. This rural watershed is very wet and has strong spatial rainfall gradients. Application of HL-RDHM to Hanalei watershed required (i) modifying the Hydrologic Rainfall Analysis Project (HRAP) coordinate system; (ii) generating precipitation grids from rain gauge data, and (iii) generating parameters for Sacramento Soil Moisture Accounting Model (SAC-SMA) and routing parameter grids for the modified HRAP coordinate system. Results were obtained for several spatial resolutions. Hourly basin-average rainfall calculated from one HRAP resolution grid (4 km × 4 km) was too low and inaccurate. More realistic rainfall and more accurate streamflow predictions were obtained with the ½ and ¼ HRAP grids. For a one year period with the best precipitation data, the performance of HL-RDHM was satisfactory even without calibration for basin-averaged and distributed *a priori* parameter grids. Calibration and validation of HL-RDHM were conducted using four-year data set each. The model reasonably matched the observed peak discharges and time to peak during calibration and validation periods. The performance of model was assessed using the following three statistical measures: Root Mean Square Error (RMSE), Nash–Sutcliffe efficiency (NSE) and Percent bias (PBIAS). Overall, HL-RDHM's performance was "very good (NSE > 0.75, PBIAS < ±10)" for the finer resolution grids (½ HRAP or ¼ HRAP). The quality of flood forecasting capability of the model was accessed using four accuracy measures (probability of false detection, false alarm ratio, critical success index, and probability of detection) for three return periods 1.005, 1.05 and 1.25. HL-RDHM is more accurate at the lower threshold discharge (65 $m^3 s^{-1}$) than at the higher ones (140 and 248 $m^3 s^{-1}$) as indicated by lower values of probability of detection and critical success index, and a higher value of the false alarm ratio. These results suggest that HL-RDHM may be suitable for flood forecasting applications in watersheds with steep terrain and strong spatio-temporal variability of precipitation, e.g. US Pacific North-West, and tropical islands.

1. Introduction

Small island mountainous watersheds such as those in Hawai'i are subject to strong spatial variability of their hydrologic cycle components, e.g. rainfall, evapotranspiration, and groundwater recharge. This variability is a challenge to the accuracy of both hydrologic measurements and predictive models. Spatial variability coupled with non-linear processes demands that hydrologic predictions utilize distributed models and finer resolution input data. Flash flood forecasting is one arena where this is particularly true, with the resolution of precipitation data being a limiting factor (Michaud and Sorooshian, 1994a,b). Most Hawaiian watersheds

* Corresponding author at: P.O. Box 519, MS 2008 Prairie View, TX 77446, USA. Tel.: +1 (936) 261 5019; fax: +1 (936) 261 2548.
E-mail address: alfares@pvamu.edu (A. Fares).

http://dx.doi.org/10.1016/j.jhydrol.2014.09.042
0022-1694/

are flash flood prone due to their steep topography and frequent, intense precipitation. Nevertheless, there is no operational flash flood forecasting in Hawai'i. A first step in remedying this situation is to evaluate rainfall-runoff models that could be used in a forecasting system such as the U.S. National Weather Service (NWS) Hydrology Laboratory Research Distributed Hydrologic Model (HL-RDHM) which is nationally used by several River Forecast Centers (RFCs) and Weather Forecast Offices (WFOs) for flood prediction.

HL-RDHM is a complete and flexible tool for distributed hydrologic modeling research and development; it serves as a prototype to validate techniques before being operational in NWS offices (HL-RDHM user manual v. 3.0.0, 2009). It has been used in several studies and operational applications (Moreda et al., 2006; Reed et al., 2007). HL-RDHM has also been used in a number of studies related to model identification, evaluation, and parameter estimation (Yilmaz et al., 2008; Tang et al., 2007; van Werkhoven et al., 2008; Wagener et al., 2009; Pokhrel et al., 2008). Reed et al. (2004) summarized the results of the Distributed Modeling Intercomparison Project (DMIP). They reported that HL-RDHM was one of the best among the tested distributed models; it reasonably matched the observed streamflow including peak flow. There are also well-defined procedures to derive *a priori* parameters for HL-RDHM from soil and land use data (Koren et al., 2000, 2004) which represent major advantages of this model compared to other distributed models.

The main objective of this study is to evaluate the performance of a model, developed for temperate continental settings, in a steep small tropical watershed. This watershed, Hanalei watershed in Kaua'i, Hawai'i, is flood-prone and has one of the highest rainfall on the globe. This is the first time that such a study has been conducted in Hawai'i. From a logistical viewpoint, HL-RDM was developed for the continental U.S.; hence, there were several challenges to its application in Hawai'i. These include: (i) lack of *a priori* parameter grids, (ii) absence of hourly and multi-sensor WSR-88D (Weather Surveillance Radar – 1988 Doppler) based precipitation grids, and (iii) an internal coordinate system that is not valid outside the continental US.

2. Materials and methods

2.1. Study basin

The study focuses on Hanalei watershed, which extends from Mt. Wai'ale'ale (1570 m AMSL) to Hanalei Bay. Hanalei watershed is located on the northern part of Kaua'i Island, Hawai'i. The simulation work focused on the upper portion of the watershed located above the USGS 16103000 stream gauge (latitude 22°10′46.5″N, longitude 159°27′59.0″W) with an area of 48 km^2 (Fig. 1). Below the gage is a highway that is frequently inundated by flooding, isolating rural residents and tourists from the rest of the island. The watershed has steep mountain slopes and deep fluvial valleys through which flows the upper 20 km of Hanalei River and its tributaries. The lower portion of the river flows through a floodplain.

Hawai'i's precipitation is highly spatially and temporally variable; its average over the ocean around the Islands is about 650 mm yr^{-1}. However, some locations, especially on the higher windward slopes (e.g. Mt. Wai'ale'ale), receive rainfall ten times greater than the open ocean surrounding the islands. This is mainly caused by orographic precipitation, which forms within the moist trade winds air as they move in the north-east direction over the steep and high terrain of the Hawaiian Islands. The average rainfall during 2001–2010 was 8893 mm yr^{-1} and 2992 mm yr^{-1} on Mt. Wai'ale'ale and at the watershed outlet, respectively. The streamflow in Hanalei River is extremely variable; its average daily discharge is 5.6 $m^3 s^{-1}$, its daily minimum discharge is 1.5 $m^3 s^{-1}$, and its daily maximum discharge is 149 $m^3 s^{-1}$ for the 2001–2010 period. The mean monthly temperature near sea level varies between 21.5 °C in January and 24.5 °C in July.

The entire Hanalei River basin supports an array of land uses, including recreational and conservation areas, and traditional wet taro farming in the lower floodplains. It was also designated as an American Heritage River by President Bill Clinton among 14 rivers or river systems in 1998. It is a multi-jurisdictional area, comprising the Hanalei National Wildlife Refuge, Halele'a State Forest Reserve, Kaua'i County beach parks, and private land holdings.

2.2. Hydrology laboratory research distributed hydrologic model

HL-RDHM can be run in lumped, semi-distributed, or fully distributed modes. The model structure is based on regular rectangular grids represented in the Hydrologic Rainfall Analysis Project (HRAP) projection. Use of the HRAP grids facilitates operational forecasting. The HRAP grid is defined by a 4 km × 4 km resolution that corresponds directly to the NWS WSR-88D precipitation products. The model can also use finer resolution grids, i.e., 2 km × 2 km (1/2 HRAP), and 1 km × 1 km (¼ HRAP).

The model predicts soil moisture, actual evapotranspiration at each grid cell by tracking its water balance using the Sacramento Soil Moisture Accounting Model (SAC-SMA) (Burnash et al., 1973), and includes hillslope and a channel routing components for each cell. Fast model responses such as overland flow and direct runoff on impervious area are routed from the hillslope and drained into a conceptual channel. The slow model responses such as interflow and baseflow go straight into the conceptual channel without going through hillslope routing. The inter-cell channel routing is conducted using a connectivity file which reflects the surface flow directions. Parameters of the HL-RDHM include water balance and routing parameters, which are assumed to be constant within a cell but vary among cells. *A priori* SAC-SMA parameters are used to account for spatial variability over a basin. The three main parameters of hillslope routing are drainage density, hillslope slope, and roughness. There are two channel routing parameters relating the channel discharge to its cross-section.

Calibration of distributed models is more complicated than that of lumped models. Typically, the value of scale factors that are multiplied across each grid is adjusted during calibration. HL-RDHM uses the Stepwise Line Search (SLS) technique and the multi-scale objective function (MSOF) to auto-optimize the scalar multipliers of all the gridded parameters (SAC-SMA and routing). More details on calibration techniques used for HL-RDHM can be found in HL-RDHM User Manual (HL-RDHM user manual v. 3.0.0, 2009) and more details on HL-RDHM can be found in Koren et al. (2004).

Some of the major challenges of using HL-RDHM in Hawai'i are: (i) lack of *a priori* SAC-SMA parameter grids, (ii) Hawai'i lies outside the HRAP coordinate system of the continental US (CONUS), and (iii) absence of hourly and multi-sensor WSR-88D based precipitation grids.

Coordinates of Hawai'i in the HRAP coordinate system are negative. Thus, the programs which automatically derive flow direction, channel slopes, and overland slope grids were re-coded. The value of HRAPXOR (=401 in original HRAP coordinate system) was changed to 2357.41553 (=2 × 978.207765 + 401). This modification was incorporated in all programs and subroutines used to run the model, e.g. the two subroutines used (i) to determine the cell-to-cell connectivity, and (ii) to generate morphological parameters as a function of cell drainage areas.

The following sections explain the process followed to derive routing parameter grids, *a priori* SAC-SMA parameter grids, and

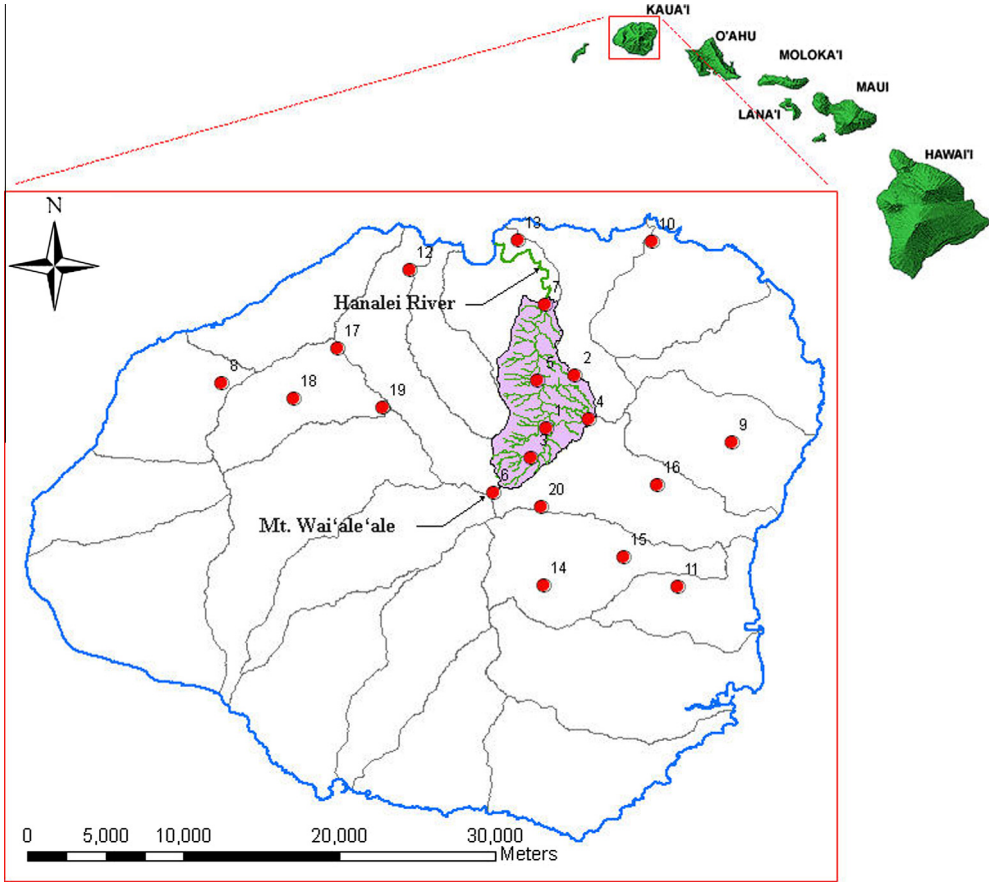

Fig. 1. Hanalei watershed and rain gauges located in the northern part of Kaua'i Island, Hawai'i (Rain gauges within and close to Hanalei watershed: 1. Keanaawi Ridge, 2. Middle Powerline, 3. Upper Hanalei, 4. Upper Powerline, 5. Middle Hanalei, 6. Mt. Wai'ale'ale and 7. Hanalei).

different forcing data (rainfall grids and potential evapotranspiration) for HL-RDHM model to simulate the hydrology of Hanalei watershed.

2.3. Generation of different resolution routing parameter grids

Details of HL-RDHM's mathematical formulation for hillslope and channel routing and their numerical solutions are found in Koren et al. (2004). They also describe algorithms used to derive distributed hillslope and channel routing parameters. The model uses the following relationship describing the relationship between the average hillslope water depth (h) in each cell and the flow discharge per unit area of hillslope (q_h):

$$q_h = 2k_q D \frac{\sqrt{S_h}}{n_h} h^{5/3} \quad (1)$$

where k_q is the unit transformation coefficient, D (km^{-1}) is the stream channel density, S_h is the hillslope slope, and n_h is the hillslope roughness coefficient.

The channel discharge, Q_C, for each cell is a function of the cross sectional area (A); it is calculated as follows:

$$Q_c = q_o A^{q_m} \quad (2)$$

where q_o is a specific channel discharge per unit channel cross-section area, and q_m is the power value of Q–A relationship.

Shape and top width parameters are defined based on the following relationship between channel top width (B) and depth (H):

$$B = \alpha H^{\beta} \quad (3)$$

where α and β are the channel-shape based routing parameters.

Routing parameter grids were generated according to HL-RDHM's user manual. First, routing parameters were determined based on observed stream flow at the watershed gauging station. The values of the coefficients of Eqs. (2) and (3) were determined based on observed stream flow data (Table 1). Default values of drainage density (D = 2.5) and Hillslope roughness (n_h = 0.15) were used to generate routing parameter grids. Based on all these parameter grids, any of the routing techniques (rutpix7: channel shape based or rutpix9: rating curve based) can be used. In this study, we used rutpix9 routing technique. Flow direction grids for different resolution grids are shown in Fig. 2.

2.4. Generation of SAC-SMA parameter grids at different resolutions

The SAC-SMA parameters were derived using the Soil Survey Geographic Database (SSURGO) and algorithms developed by Koren et al. (2000, 2003). SAC-SMA parameter grids for the ½ HRAP and 1 HRAP resolutions were derived based on the modified ¼ HRAP SAC-SMA grids. The range of SAC-SMA parameter ratios were derived based on the minimum, maximum and mean values of *a priori* SAC-SMA parameter grids (Table 2) at the ¼ HRAP resolution

Table 1
Coefficients for discharge–cross-section relationship and channel top width–depth relationship.

q_o	q_m	α	β	n_c
0.048	1.82	13.78	1.01	0.07

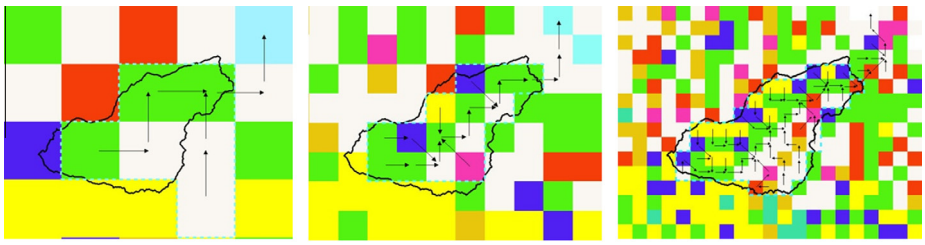

Fig. 2. 1 HRAP, ½ HRAP and ¼ HRAP resolution flow direction grids within Hanalei watershed (Solid polygon – actual watershed, dashed polygon – watershed delineated in HL-RDHM).

Table 2
Range of parameter ratios based on *a priori* SAC-SMA parameter of Hanalei watershed (based on ¼ HRAP resolution grids).

SAC-SMA parameter	*A priori* values			Multiplier boundaries	
	Min	Max	Mean	Lowest	Highest
UZTWM (mm)	1.495	116.667	43.891	0.03	2.66
UZFWM (mm)	0.766	50.916	21.133	0.04	2.41
UZK (day^{-1})	0.008	0.282	0.183	0.04	1.54
ZPERC	2.655	116.796	65.052	0.04	1.80
REXP	0.073	2.628	1.693	0.04	1.55
LZTWM (mm)	1.611	69.205	35.167	0.05	1.97
LZFSM (mm)	0.285	11.489	5.937	0.05	1.94
LZFPM (mm)	0.681	28.349	14.357	0.05	1.97
LZSK (day^{-1})	0.003	0.107	0.0712	0.04	1.50
LZPK (day^{-1})	0.000	0.004	0.00095	0.00	4.20
PFREE	0.009	0.404	0.223	0.04	1.81

grids. The lowest and highest parameter ratios were determined by dividing the minimum and maximum values by the mean, respectively.

2.5. Generation of different resolution precipitation grids

In the U.S., flood forecasting is often conducted using hourly multi-sensor WSR-88D-based precipitation grids. Because these are not generated for Hawai'i, we instead used hourly precipitation estimated from hourly rain gage data (Fig. 1) for the 2001–2010 period.

Hourly rain grids for the three resolutions (1, ½ and ¼ HRAP) were developed, following the ordinary kriging interpolation method as recommended by Mair and Fares (2011), using rain data from 20 rain gauges across north shore Kaua'i. Five of these 20 rain gauges (numbered from one through five) are located within the study watershed (Fig. 1). These rain gauges were established during 2006. Based on the results of Mair and Fares (2010), the normal ratio method was adopted to estimate missing hourly rainfall data for these stations. Ordinary kriging method requires the definition of a variogram characterizing the spatial variability of the precipitation field. The variograms for each year were developed using annual average of hourly precipitation. The optimum model parameters of the semi-variogram (sill and range) were derived for each year (2001–2010) based on the annual average of the hourly rainfall using the gstat R package (Pebesma and Wesseling, 1998). Gstat is a computer program for variogram modeling, and geostatistical prediction and simulation (Pebesma and Wesseling, 1998). Gstat models variograms using a two-step procedure of calculating the sample variogram and then fitting a model to this sample variogram (by estimating model parameters). According to Haberlandt (2007) the impact of the semivariogram model on interpolation performance is not very high, thus, we used spherical semivariogram (isotropic).

We tested the performance of the annual averaged variogram of the year 2008 by temporarily removing one observation at a time from the data set and re-estimating the removed value from the remaining stations using the interpolation method. We followed this procedure for the five stations within the watershed numbered from one to five in Fig. 1. The summary of the results of variogram parameters derived from annual average of hourly precipitation is shown in Table 3 and scatter plots of observed and predicted hourly rainfall for different stations within Hanalei watershed are shown in Fig. 3. The r^2 for the five stations varies from 0.48 to 0.79. Watershed-average precipitation (annual average of hourly) increased with increase in the grid resolution (Fig. 4).

The estimated annual average potential evapotranspiration (ET_o) for Hanalei watershed based on annual pan evaporation data is 758 mm yr^{-1}. Thus, the daily average ET_o was 2.08 mm.

2.6. Model performance measures

The performance of HL-RDHM was determined using three commonly used statistical performance measures: Root Mean Square Error (RMSE), Percent bias (PBIAS), and Nash–Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970). These performance measures have been used with data based on daily and/or monthly time steps. The root-mean square error computes the standard deviation of the model prediction error which is the difference between measured (Q_{obs}) and simulated values (Q_{sim}). The smaller the RMSE ($m^3\ s^{-1}$) value, the better the model performance is. It is computed as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\left(\sum_{t=1}^{n}\left((Q_{sim})_t - (Q_{obs})_t\right)^2\right)} \quad (4)$$

Percent bias measures the tendency of the simulated flows to be larger or smaller than their observed counterparts; its optimal value is 0.0, positive values indicate a tendency to overestimation, and negative values indicate a tendency to underestimation. It is estimated as follows:

$$\mathrm{PBIAS} = \frac{\sum_{t=1}^{n}\left((Q_{sim})_t - (Q_{obs})_t\right)}{\sum_{t=1}^{n}(Q_{obs})_t} \times 100\% \quad (5)$$

Nash–Sutcliffe efficiency measures the fraction of the variance of the observed flows explained by the model in terms of the relative magnitude of residual variance to the variance of the flows; the optimal value is 1.0 and values should be larger than 0.0 to indicate 'minimally acceptable' performance. It is computed as follows:

$$\mathrm{NSE} = 1 - \frac{\sum_{t=1}^{n}\left((Q_{sim})_t - (Q_{obs})_t\right)^2}{\sum_{t=1}^{n}\left((Q_{obs})_t - (\overline{Q}_{obs})\right)^2} \quad (6)$$

The ranges of models' performance ratings, RMSE, NSE and PBIAS reported by Moriasi et al. (2007) were based on monthly time step data. Models' performances are poorer for shorter time steps than

Table 3
Semivariogram parameters.

Parameter	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Sill (m^2)	0.090	0.079	0.085	0.133	0.100	0.167	0.103	0.097	0.099	0.059
Range (m)	12284	19862	17383	13942	15408	19364	17416	16666	19371	17609

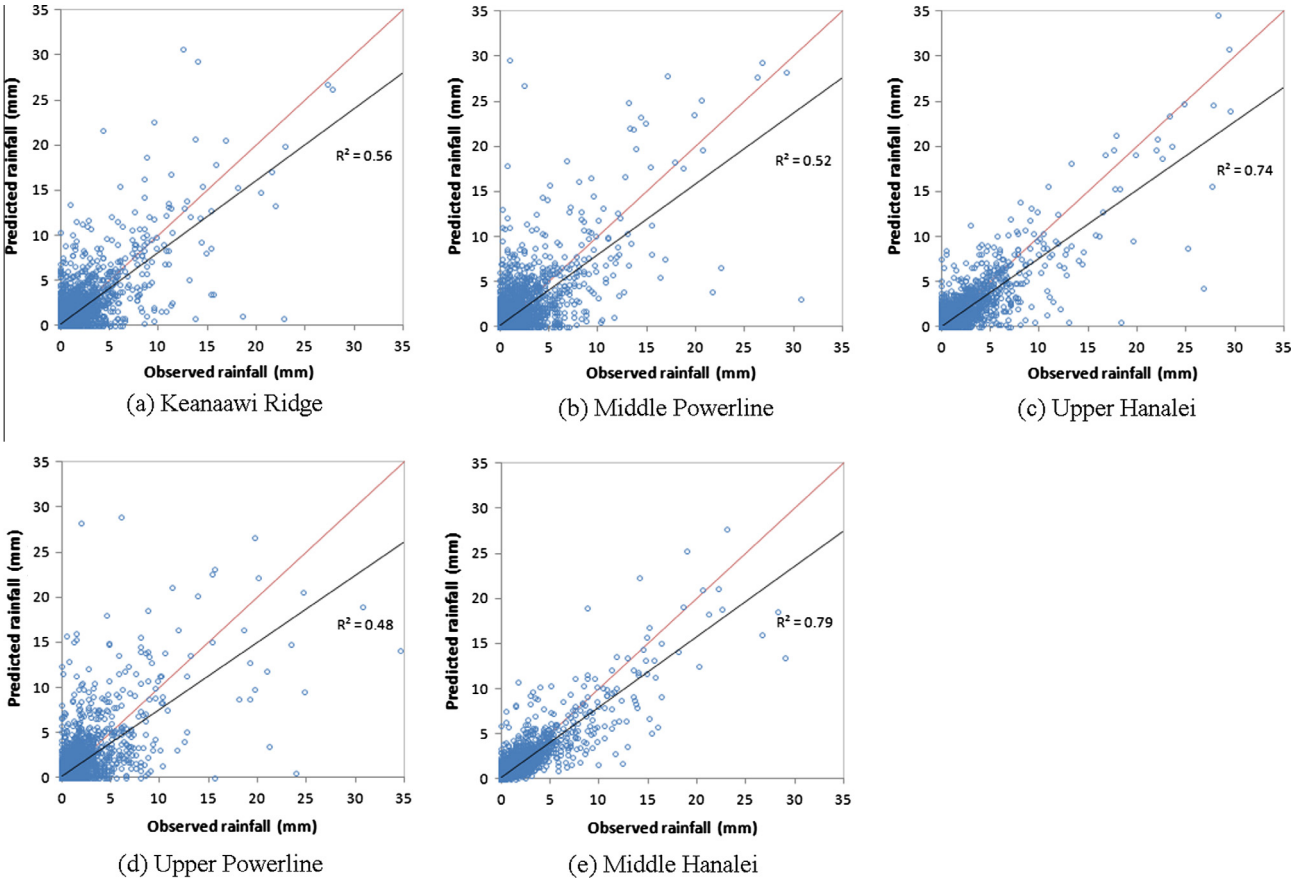

Fig. 3. Scatter plots of observed and predicted rainfall for different stations within Hanalei watershed (Variogram derived from annual average of hourly precipitation).

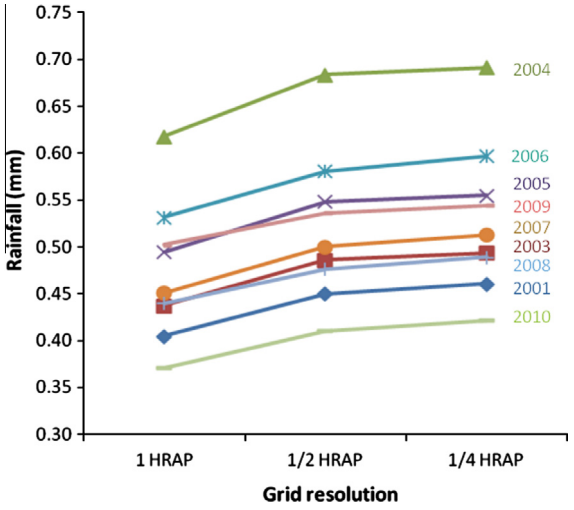

Fig. 4. Watershed wide annual average of hourly precipitation for different resolutions.

for longer time steps (Engel et al., 2007). Ajami et al. (2004) reported value of RMSE, PBIAS and NS of calibration and validation of semi distributed model for an hourly data (Calibration: RMSE = 21.88, PBIAS = 6.4, NS = 0.62 and Validation: RMSE = 27.8, PBIAS = 1.42, NS = 0.76). There were no optimum ranges available in the literature for these model's performance parameters based on daily or sub-monthly time steps. Thus, it was decided to use Moriasi's performance ratings for this work (Table 4). It is expected that our model ratings, using hourly time steps, are strict due to the use of hourly statistics along with rating categories based on monthly statistics.

Table 4
General Performance ratings for a monthly time step (Moriasi et al., 2007).

Performance rating	NSE	PBIAS (%)
Very good	$0.75 < NSE \leqslant 1.00$	PBIAS < ±10
Good	$0.65 < NSE \leqslant 0.75$	$\pm 10 \leqslant PBIAS < \pm 15$
Satisfactory	$0.50 < NSE \leqslant 0.65$	$\pm 15 \leqslant PBIAS < \pm 25$
Unsatisfactory	$NSE \leqslant 0.50$	PBIAS > ±25

2.7. Accuracy measures

The hourly accuracy of HL-RDHM was accessed in its ability to exceed certain flood thresholds (65, 140 and 248 $m^3\ s^{-1}$) for three return periods (1.005, 1.05, and 1.25). The number of flood events

exceeding the 2-year return period flood is very small. Four flood forecasting accuracy measures were calculated for peak floods with three return periods, i.e. 1.005, 1.05 and 1.25. These accuracy measures are defined as follows: probability of detection (POD)

$$\mathrm{POD} = \frac{hits}{hits + misses} \tag{7}$$

probability of false detection (POFD),

$$\mathrm{POFD} = \frac{false\ alarms}{correct\ negatives + false\ alarms} \tag{8}$$

false alarm ratio (FAR)

$$\mathrm{FAR} = \frac{false\ alarms}{hits + false\ alarms} \tag{9}$$

critical success index (CSI)

$$\mathrm{CSI} = \frac{hits}{hits + misses + false\ alarms} \tag{10}$$

The values of these accuracy measures range between 0 and 1. A perfect forecast would have a POD of 1, a POFD of 0, a FAR of 0 and a CSI of 1.

3. Simulation results

In the beginning, we tested performance of the model using different SAC-SMA parameter grids (watershed wide averaged, distributed without calibration and distributed with calibration) of different resolutions (1 HARAP, ½ HARAP and ¼ HARAP) with 2008 data to select the appropriate spatial resolution of the model calibration. We also derived applicable range of SAC-SMA parameter ratios for use during calibration. Then, the model was calibrated and validated using four-year data sets (2003–2006) and (2007–2010), respectively.

3.1. HL-RDHM performance for different SAC-SMA parameter grids and resolutions

Average watershed grids of *a priori* SAC-SMA parameters for all resolutions were produced and HL-RDHM was tested in a lumped mode using average watershed precipitation grids for 2008 data. The model was also tested using distributed *a priori* SAC-SMA parameters with and without calibration. The model's performance in the lumped mode was the poorest among averaged, distributed, and distributed with scaled *a priori* SAC-SMA parameters (Fig. 5). However, the model's performance was very good based on NSE (NSE > 0.75) for all three SAC-SMA parameter grids of different resolutions. This is a practical indicator of the potential applicability of *a priori* SAC-SMA parameter grids for rainfall-runoff simulation of tropical watersheds. Based on these results, we used grids of all three resolutions (1 HARAP, ½ HARAP and ¼ HARAP) in model's calibration and validation.

3.2. Applicable range of SAC-SMA parameter ratios

HL-RDHM's performance was tested for different ranges of SAC-SMA parameter ratios to determine the practical ranges of their values for small island tropical watersheds. Initially, SAC-SMA parameter ratios were optimized using values of the original SAC-SMA parameter ratios (Lower ratio = minimum value/average value, Upper ratio = maximum value/average value). SAC-SMA parameter ratios were then optimized based on expanded ranges of *a priori* grids. Four new ranges of *a priori* parameters were established by expanding lower and upper boundaries of the original ranges by 20%, 40%, 60% and 80%. SAC-SMA parameters were

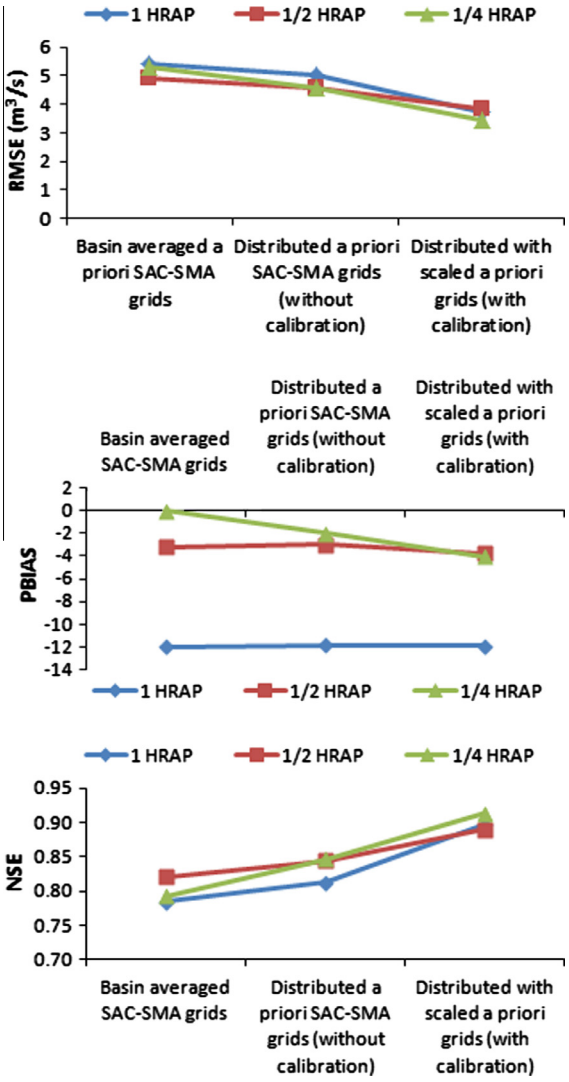

Fig. 5. Performance ratings for different SAC-SMA parameter grids and resolutions for input data of 2008 (Basin averaged, distributed and distributed with scaled).

optimized using these four new ranges with four-year (2003–2006), two-year (2005–2006) and one-year (2008) data sets. In other words, a ±20% expansion of the original range of a given *a priori* SAC-SMA parameter means, the lower range of this parameter was decreased by 20% and the value of the upper range of the same parameter was increased by 20%. The model's performance improved as the range of the values of *a priori* parameters expanded for all resolution grids (Fig. 6). Based on PBIAS, the model's performance was less sensitive to the expansion of the values of SAC-SMA parameter ratios. The magnitudes of improvement in the performance ratings were not significant if the range of SAC-SMA parameter was expanded beyond ±60%. The performance of the ¼ HRAP resolution was clearly better than that of the other two resolutions (1 and ½ HRAP) based on RMSE and NSE (Fig. 6). The same results are even clearer based on PBIAS. The poor performance of 1 HRAP resolution is mainly due to the poor delineation of the watershed at that resolution (see Fig. 2a). Based on these results, ±80% expanded SAC-SMA parameter ratios were used in model's calibration exercise.

3.3. Model calibration

Four years (2002, 2007, 2009 and 2010) had incomplete streamflow data. Therefore, HL-RDHM was calibrated with four-year

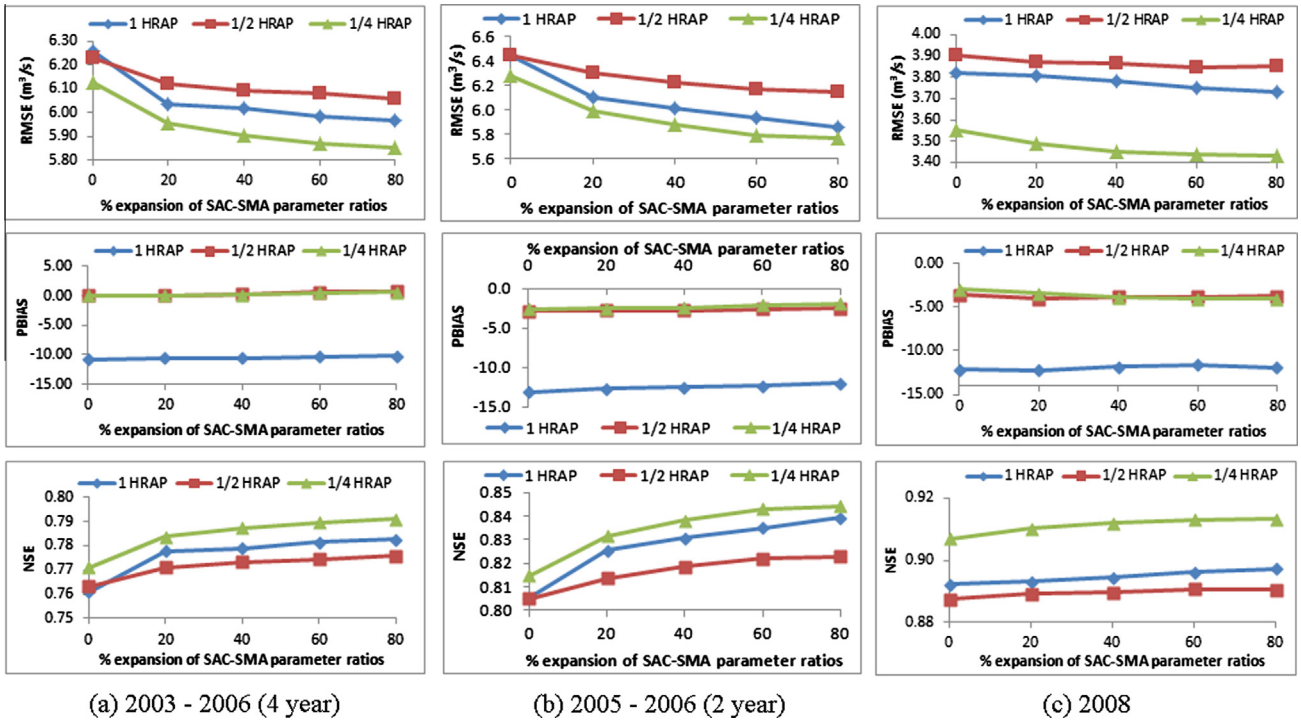

Fig. 6. Performance ratings for expanded range of parameter ratios.

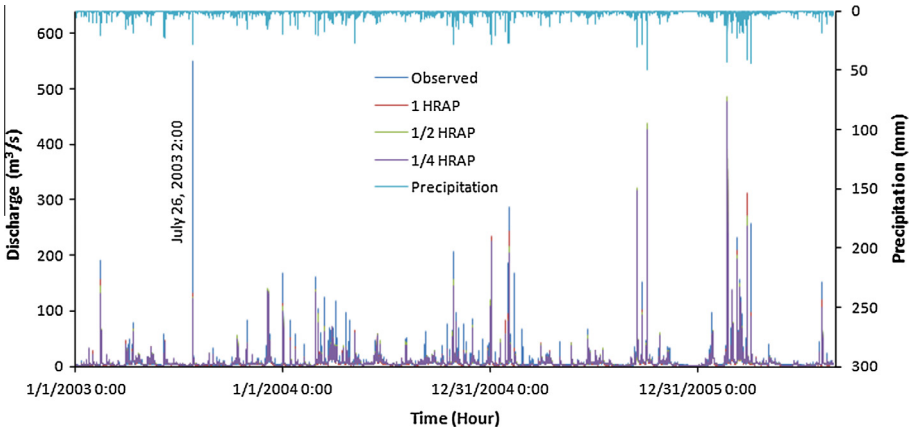

Fig. 7. Comparison of observed and simulated outflow hydrograph at the USGS Hanalei river stream gauge (Range of parameter ratios = ±80% of *a priori* grid based range of Hanalei watershed for period 2003–2006).

streamflow data set (2003–2006). The model closely simulated the observed streamflow during the entire simulation period (Fig. 7) for all the resolutions as reflected by the high correlation coefficient between simulated and observed data (Fig. 8). However, the model tended to overestimate streamflow in all resolutions for observed streamflow higher than 300 $m^3 s^{-1}$ (Fig. 8) except for one event. Based on NSE, the model performance was very good for all resolutions (Table 5); whereas, based on PBIAS, the model's performance was very good for the ½ and ¼ HRAP resolution and even surprisingly good for the 1 HRAP resolutions. HL-RDHM underestimates the observed streamflow for the 1 HRAP resolution as indicated by the negative value of PBIAS.

The observed peak flow during 2003 was 1.4–2.6 times higher than that during the other years (Table 6). The model's simulation of July 26 2003 flood event was very poor irrespective of resolution grids (Figs. 7 and 8).

Model performance was also evaluated by examining whether or not the model could predict flow above a given flood threshold. Seventy to 74% of the observed streamflow at or above a threshold of 65 $m^3 s^{-1}$ (recurrence interval of 1.005 years) are accurately simulated using the three resolution grids (Table 7). However, the value of the detection probability of 140 and 248 $m^3 s^{-1}$ threshold discharges are lower than that of 65 $m^3 s^{-1}$. The probability of false detection is less than 0.002 for all threshold discharges. This indicates that only 0.2% of the simulated discharges were less than the observed discharges. One-quarter of all simulated discharges equal or higher than 65 $m^3 s^{-1}$ were false as indicated by a value of FAR equals to 0.25 (Table 7). In other words, 25% of the simulated peak discharges of at least 65 $m^3 s^{-1}$ were falsely simulated. The level of FAR increases with increase of the threshold discharges. The critical success index is higher than 0.5 for the 65 and 140 $m^3 s^{-1}$ threshold discharges; however it is lower than

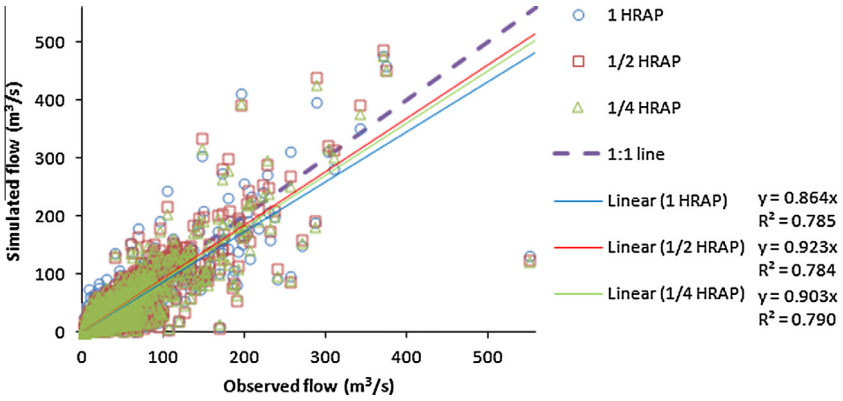

Fig. 8. Scatter plot of observed and simulated flow at the USGS Hanalei river stream gauge (Range of parameter ratios = ±80% of *a priori* grid based range of Hanalei watershed for period 2003–2006).

Table 5
Performance ratings for calibration of four year data (2003–2006).

Performance measures	Input grid resolution			Performance rating (Moriasi et al., 2007)
	1 HRAP	½ HRAP	¼ HRAP	
RMSE (m^3/s)	5.969	6.060	5.851	
PBIAS	−10.201	0.800	0.664	Good (1 HRAP) & very good (1/2 and 1/4 HRAP)
NSE	0.783	0.776	0.791	Very good

Table 6
Maximum discharge for different years (hourly averaged).

Year	2001	2003	2004	2005	2006	2008
Maximum discharge – hourly (m^3/s)	250.67	550.76	206.36	310.35	373.78	277.79

Table 7
Accuracy measures of simulated flow based on different threshold peak discharge (Calibration, 2003–2006).

Recurrence interval (*n*-year flood)	1.005			1.05			1.25		
Threshold peak discharge values (m^3/s)	65			140			248		
Number of floods higher than threshold peak discharge	245			62			13		
Grid resolution	1 HRAP	1/2 HRAP	1/4 HRAP	1 HRAP	1/2 HRAP	1/4 HRAP	1 HRAP	1/2 HRAP	1/4 HRAP
POD	0.718	0.739	0.702	0.613	0.645	0.613	0.538	0.538	0.538
POFD	0.001	0.002	0.001	0.000	0.001	0.000	0.000	0.000	0.000
FAR	0.174	0.249	0.204	0.224	0.322	0.208	0.417	0.500	0.417
CSI	0.624	0.593	0.595	0.521	0.494	0.528	0.389	0.350	0.389

Note: POD = probability of detection, POFD = probability of false detection, FAR = false alarm ratio, and CSI = critical success index.

0.5 for threshold discharge of 248 $m^3 s^{-1}$. This indicates that slightly more than half of 65 and 140 $m^3 s^{-1}$ discharges were correctly simulated; however, more than half of the 248 $m^3 s^{-1}$ threshold discharges were not accurately simulated indicating a decline in the accuracy of the model with an increase in the threshold discharge level.

3.4. Model validation

The model was validated with four-year stream flow data (2007–2010) using optimized parameters from the 2003–2006 period. There is no continuous streamflow data for 2007, 2009 and 2010; the total number of missing data was 3077 h for 2007–2010 period. The maximum observed peak flow was less than 290 $m^3 s^{-1}$ even though 2009 was a wet year compared to some other years, i.e. 2001, 2003, 2007, 2008, and 2010 (Fig. 4). This suggests that there might be missing peaks in this four-year data.

We also tested the performance of the model using optimized parameter ratios based on *a priori* SAC-SMA grids. Expanding the range of the *a priori* SAC-SMA grid did not improve the performance of HL-RDHM during the validation period (2007–2010). In fact, the model's performance was better with the original range than with that of the expanded range (Table 8). The four-year 2003–2006 period was wetter than the 2007–2010 period.

The simulated hydrograph seems to reasonably match the observed data (Fig. 9). There is a strong correlation between measured and simulated data irrespective of the resolution of the grids (Fig. 10); although the model seems to slightly under estimate the observed data especially during low flows.

Model performance was also evaluated by examining whether the model could predict the flow above a given flood threshold. Fifty-eight to 62% of the observed flows with a magnitude of at least 65 or 140 $m^3 s^{-1}$ were correctly simulated with different resolution grids, respectively (Table 9). However, for threshold discharge higher or equal to 248 $m^3 s^{-1}$, POD reached 0.5 at best.

Table 8
Performance of HL-RDHM during validation period (2007–2010) using optimized parameter ratios for period 2003–2006.

S. no.	Simulation option	Performance measures	Input grid resolution			Performance rating (Moriasi et al., 2007)
			1 HRAP	½ HRAP	¼ HRAP	
1	Optimized parameter ratios based on *a priori* grids of Hanalei	RMSE (m^3/s)	3.884	3.993	3.759	
		PBIAS	−7.037	2.959	3.737	Very good
		NSE	0.845	0.836	0.855	Very Good
2	Optimized parameter ratios based on *a priori* grids of Hanalei ± 80%	RMSE (m^3/s)	3.945	4.156	3.929	
		PBIAS	−6.196	3.769	4.288	Very good
		NSE	0.840	0.823	0.842	Very Good

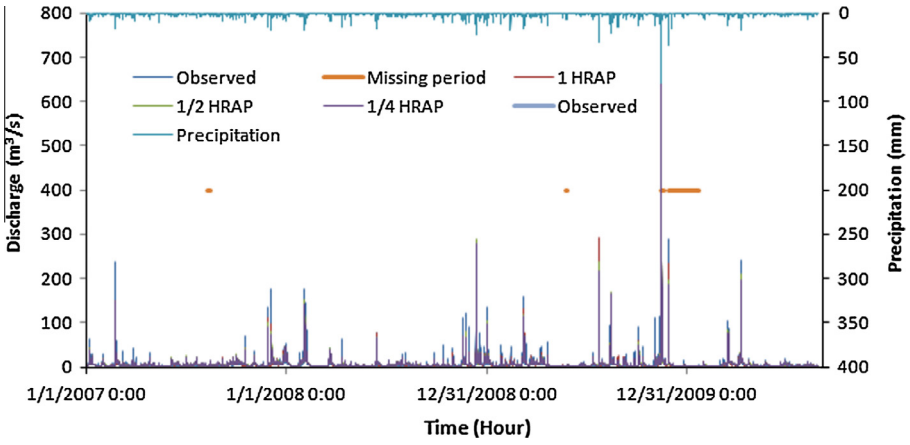

Fig. 9. Outflow hydrograph (Validation for period 2007–2010 using optimized SAC-SMA parameter ratios for period 2003–2006 derived by ±80% expansion of *a priori* grid based range of Hanalei watershed).

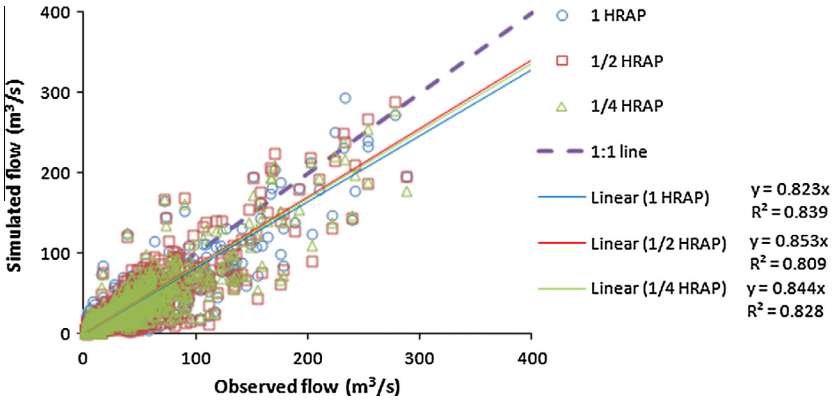

Fig. 10. Scatter plot of observed and simulated flow (Validation for period 2007–2010 using optimized SAC-SMA parameter ratios for period 2003–2006 derived by ±80% expansion of *a priori* grid based range of Hanalei watershed).

There were fewer flood events at three threshold peak discharge (65, 140 and 248 $m^3 s^{-1}$) during the validation period (2007–2010) than during the calibration period (2003–2006) (Tables 7 and 9). This might be due to the missing streamflow data during validation period. It is difficult to access the accuracy of the model for events equal or higher than 248 $m^3 s^{-1}$ because there were only 4 events that met that criterion during the validation period. The probability of false detection is equal to 0.1% indicating a reasonable detection of threshold discharges by the model. One quarter of the simulated peak flow events above 65 $m^3 s^{-1}$ were false alarms (FAR < 0.25). False alarm ratio for peak flow events of 140 was less than that of 65 $m^3 s^{-1}$. False alarm ratio varied as a function of simulation resolution for the 248 $m^3 s^{-1}$ threshold discharge; there is no false alarm at ¼ HRAP resolution; however, FAR reached 0.66 for the 1 HRAP resolution. The model was more than 50% of the time successful (CSI > 0.5) in simulating 65 and 140 $m^3 s^{-1}$ threshold discharges; however, this success decreased (CSI < 0.5) for the 248 $m^3 s^{-1}$ threshold discharge. These results indicate that the accuracy of simulation decreases as threshold discharge increases; thus, the model is more accurate in simulating lower discharges than higher discharges.

4. Results and discussions

4.1. Precipitation grids of different resolutions

The watershed averaged precipitation based on ¼ HRAP grid is higher than those of the other grid resolutions. The difference between ¼ HRAP and ½ HRAP is about 1.1–2.7% whereas the dif-

Table 9
Accuracy measures of simulated flow based on different threshold peak discharge (Validation, 2007–2010).

Recurrence interval (*n*-year flood)	1.005			1.05			1.25		
Threshold peak discharge values (m^3/s)	65			140			248		
Number of floods higher than threshold peak discharge	147			34			4		
Grid resolution	1 HRAP	1/2 HRAP	1/4 HRAP	1 HRAP	1/2 HRAP	1/4 HRAP	1 HRAP	1/2 HRAP	1/4 HRAP
POD	0.585	0.599	0.578	0.588	0.618	0.588	0.250	0.500	0.500
POFD	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
FAR	0.173	0.241	0.183	0.130	0.160	0.091	0.667	0.333	0.000
CSI	0.521	0.503	0.512	0.541	0.553	0.556	0.167	0.400	0.500

Note: POD = probability of detection, POFD = probability of false detection, FAR = false alarm ratio, and CSI = critical success index.

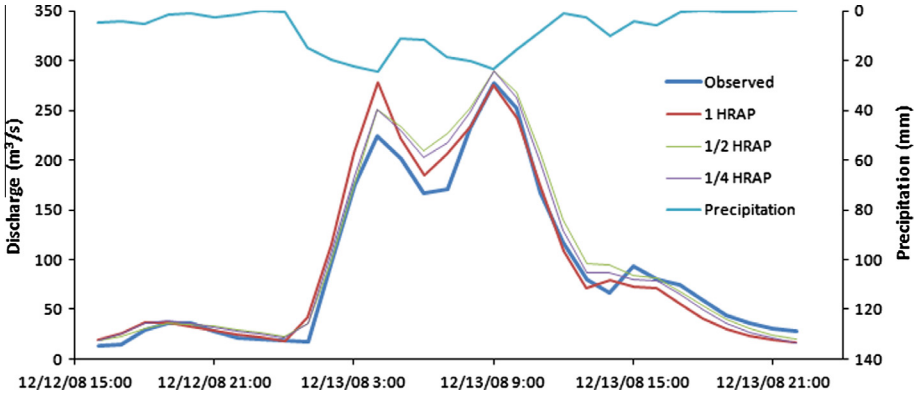

Fig. 11. Part of simulated and observed hydrograph for different resolution grids (From 12/12/2008 16:00 PM to 12/13/2008 10:00 AM).

ference between ¼ HRAP and 1 HRAP is about 7.8–12.1% based on the yearly comparison of watershed averaged precipitation from 2001 to 2010. The main reason for the higher rainfall difference between that of 1 HRAP and ¼ HRAP resolution grids is the inaccurate watershed delineation based on the grid resolution of the former. The delineated watershed based on the 1 HRAP covers six grids of 4 km × 4 km. The resulting area of the watershed is 96 km^2 which is double its actual size. As a result, an area of the lower part of the watershed, 16 km^2 (4 × 4 km), with relatively low rainfall is part of the watershed; this does not represent the actual watershed. However, a similar area in the top part of the watershed with a rainfall much higher than that at the lower part was not included. This would explain the consistently lower average rainfall based on 1 HRAP as compared to ½ and ¼ HRAP resolutions. The 4 km × 4 km grid resolution does not accurately represent the watershed and consequently will result in substantial errors if used with a distributed hydrological model to simulate the hydrology of this watershed.

4.2. Applicable range of SAC-SMA parameter ratio

The performance of HL-RDHM using the expanded range of *a priori* parameters was better than that with the original, non-expanded range; the model's performance improved by up to 4.6% and 2.8% based on NSE and RMSE, respectively, for the four-year data (2003–2006) (Fig. 6a). However, during the validation period (2007–2010), the model's performance worsened when the extended *a priori* range based optimized parameter ratios were used instead of the original range based optimized parameter ratios (Table 8).

4.3. Applicability of a priori SAC-SMA parameter grids

A priori SAC-SMA parameter grids were derived using SSURGO soils data. The performance rating is "very good" for both watershed averaged *a priori* grids and distributed *a priori* grids of Hanalei watershed for input data of 2008 (Fig. 5). Input data of 2008 was used because it had the least missing precipitation data from rain gauges within the watershed. This shows the possibility of using HL-RDHM without calibration. However, the result may be different with data from other years, so we need to test it with data for longer periods. We can improve the performance of the model by fine tuning some of the values of its parameters during calibration.

4.4. Hydrograph comparison

The model reasonably simulated the observed magnitudes of peaks and time to peak during calibration (2003–2006) and validation period (2007–2010). However, the model was unable to reproduce the maximum peak of 2003 during calibration (Figs. 7 and 8). The main reason for the poor performance is due to the uncertainty in the rainfall data as a result of the missing data from a number of rain gauge stations. Before 2006, there were only two operating rain gauges within the watershed, one at Mt. Wai'ale'ale (top of the watershed) and the second one at the USGS streamflow gauge. The other gauges were established during 2006.

The simulated flow based on the 1 HRAP grids was lower than those based on ½ HRAP and ¼ HRAP grids (Fig. 11, peak flow at 12/13/2008 9:00 AM) when the rainfall was higher in the upper part of the watershed (Fig. 12b). However, the simulated flow was higher (Fig. 11, peak flow at 12/13/2008 4:00 AM) when the precipitation in the lower basin was higher (Fig. 12a). The watershed represented by 1 HRAP resolution grid is not reasonable. The delineated basin covers a larger area at the lower part of the watershed. The rainfall intensity at the upper part of the watershed is higher than that of the lower part in most of the flood generating rainfall events; thus, the simulated result using 1 HRAP is not so poor compared to those of the ½ HRAP and ¼ HRAP grid resolution even though the watershed delineated in the 1 HRAP grid covers a larger area in the lower part of the watershed.

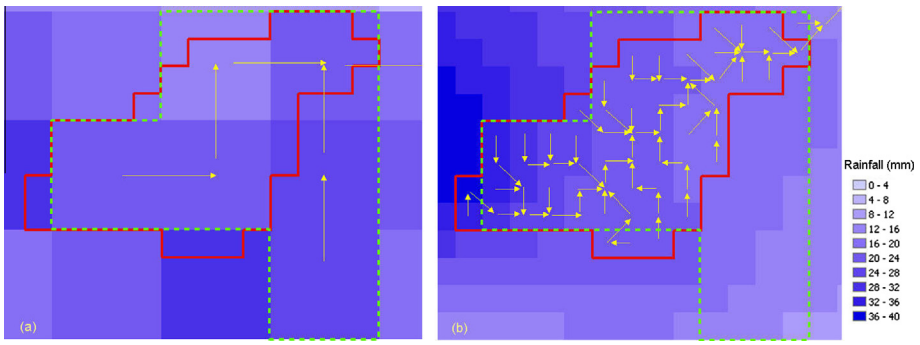

Fig. 12. (a) Precipitation grid (1 HRAP) – 1 h before first peak (12/13/2008 3:00 AM), (b) Precipitation grid (1/4 HRAP) – 1 h before second peak (12/13/2008 8:00 AM). (In this figure, solid polygon is the watershed delineated in ¼ HRAP resolution grid and dashed polygon is the watershed delineated in 1 HRAP resolution grid).

4.5. Performance measures

The modeling performance measures (RMSE, PBIAS and NSE) were used to evaluate the performance of HL-RDHM during calibration and validation periods. During calibration and based on PBIAS, the performance ratings of HL-RDHM are lower for 1 HRAP resolution grids resulting in a relatively higher underestimation of stream flow. However, based on PBIAS, the model's performance with the ½ HRAP and ¼ HRAP grid resolutions is "very good". The performance of HL-RDHM, based on NSE, is "very good" for all resolution grids. During model validation, the performance of the model is "very good" for all resolution grids based on PBIAS and NSE. Overall, the model's performance is "good" to "very good" if we use input data of finer grids including precipitation grids derived from high resolution rain gauges data within the watershed.

5. Conclusions and recommendations

In this study, we assessed the performance of HL-RDHM in a Hawaiian watershed. We modified the HRAP coordinate system, generated input data of precipitation grids for different resolutions using data from rain gauges, and generated SAC-SMA parameter grids and routing parameter grids for the modified HRAP coordinate system. Overall, the comparison of simulated and observed hydrographs shows reasonably accurate simulation of peak flows and time to peak during calibration and validation periods. The exception was for results based on 1 HRAP resolution grids. Simulations using 1 HRAP precipitation grids were not accurate. The resolution was too coarse in comparison to the spatial scale of precipitation variability; 1 HRAP precipitation values were lower than those from ½ HRAP and ¼ HRAP grids.

The performance of HL-RDHM for basin-averaged *a priori* grids and distributed *a priori* grids was evaluated without calibration for one year (2008). Relatively, favorable results for the distributed case indicate the possibility of using HL-RDHM without calibration. This would be a definite advantage for flash flood forecasting applications since most stream basins and flood-prone areas do not have operational gauges for data collection.

With calibration, the performance of HL-RDHM was "very good" during calibration and validation for finer resolution grids (½ HRAP or ¼ HRAP) and precipitation grids derived from interpolation with sufficient number of rain gauges.

The above results show the potential of using HL-RDHM in small islands environment such as Hawai'i. It should be noted, however, that Hanalei watershed is exceptional in its rugged topography and extreme rainfall. Hence, further studies should focus on other Hawaiian watersheds representing other land use conditions, topography, and degrees of orographic forcing of precipitation. Tentatively, however, results are promising for both flash flood forecasting applications and other uses such as flood studies. In either case discharges from HL-RDHM could be used as input data for hydraulic modeling. Hydraulic modeling would be necessary to predict details of flood extent, flood depth and duration of flooding which will help in disaster anticipation, preparation, and mitigation. Streamflow response to rainfall events in small islands tropical watersheds is short. The time of concentration for the watershed of this study is only 1.5 h based on empirical equation (Kirpich, 1940). Extending the lead-time for flood forecasts for these watersheds requires accurate rainfall prediction. Thus, further efforts are needed to evaluate the performance of HL-RDHM using the Quantitative Precipitation Forecasts (QPFs) grids. Development of operational flood forecasting systems would require, in addition to a rainfall-runoff model, investments in staffing and operational precipitation products. There is a particular need for Hawaiian HRAP quantitative precipitation estimates (QPEs) grids. The size and coordinate system of quantitative precipitation forecasts (QPFs) grids for Hawai'i are different from those for the CONUS; hence, further work should focus on generating QPE grids and incorporating QPF grids into the HL-RDHM model.

Acknowledgements

The project was supported by a grant from the Collaborative Science, Technology, and Applied Research (CSTAR) Program of NOAA-National Weather Service Grant number NA08NWS4680036. Partial funding was also provided by the National Institute of Food and Agriculture of the United States Department of Agriculture Evans-Allen funds. The authors wish to thank Safeeq Mohamad, Sanjit Deb and Farhat Abbas for their assistance during different stages of the project implementation. Special thanks to Dave Streubel, NOAA NWS, APRFC for his cooperation by providing SAC-SMA parameter grids. The authors also wish to thank people from NOAA including Michael Smith, Victor Koren, and Zhengtao Cui for their suggestions to resolve problems running the HL-RDHM model.

References

Ajami, N.K., Gupta, H., Wagener, T., Sorooshian, S., 2004. Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. J. Hydrol. 298, 112–135.

Burnash, R.J.C., Ferral, R.L., McGuire, R.A., 1973. A Generalized Streamflow Simulation System – Conceptual Modeling for Digital Computers. U.S. Dept. of Commerce National Weather Service and State of California Department of Water Resources, p. 204.

Engel, B., Storm, D., White, M., Arnold, J.G., Arabi, M., 2007. A hydrologic/water quality model application protocol. J. Am. Water Resour. Assoc. 43 (5), 1223–1236. http://dx.doi.org/10.1111/j.1752-1688.2007.00105.x.

Haberlandt, U., 2007. Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. J. Hydrol. 332, 144–157.

HL-RDHM user manual v. 3.0.0, 2009. <http://www.cbrfc.noaa.gov/present/rdhm/RDHM_3_0_0_User_Manual.pdf>.

Kirpich, Z.I.P., 1940. Time of concentration of small agricultural watersheds. Civ. Eng. 10 (6), 362.
Koren, V.I., Smith, M., Wang, D., Zhang, Z., 2000. Use of soil property data in the derivation of conceptual rainfall-runoff model parameters. In: Proceedings of the 15th Conference on Hydrology. AMS, Long Beach, CA, pp. 103–106.
Koren, V., Smith, M., Duan, Q., 2003. Use of a priori parameter estimates in the derivation of spatially consistent parameter sets of rainfall-runoff models. In: Duan, Q., Sorooshian, S., Gupta, H., Rosseau, H., Turcotte, H. (Eds.), Calibration of Watershed Models Water Science and Applications, vol. 6. AGU, pp. 239–254.
Koren, V.I., Reed, S., Smith, M., Zhang, Z., Seo, D.J., 2004. Hydrology laboratory research modeling system (HL-RMS) of the US national weather service. J. Hydrol. 291, 297–318.
Mair, A., Fares, A., 2010. Assessing rainfall data homogeneity and estimating missing records in Makaha Valley, Oahu, Hawaii. J. Hydrol. Eng. 15 (1), 61–66.
Mair, A., Fares, A., 2011. Comparison of rainfall interpolation methods in a mountainous region of a tropical island. J. Hydrol. Eng. 16 (4), 371–383.
Michaud, J.D., Sorooshian, S., 1994a. Effect of rainfall sampling errors on simulations of desert flash floods. Water Resour. Res. 30 (10), 2765–2775.
Michaud, J.D., Sorooshian, S., 1994b. Comparison of simple versus complex distributed runoff models on a mid-sized, semi-arid watershed. Water Resour. Res. 30 (3), 593–605.
Moreda, F., Koren, V., Zhang, Z.Y., Reed, S., Smith, M., 2006. Parameterization of distributed hydrological models: learning from the experiences of lumped modeling. J. Hydrol. 320 (1–2), 218–237 (Special Issue SI).
Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Binger, R.L., Harmel, R.D., Veith, T., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. T. ASABE 50 (3), 885–900.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models: Part 1. A discussion of principles. J. Hydrol. 10 (3), 282–290.
Pebesma, E.J., Wesseling, C.G., 1998. Gstat: a program for geostatistical modelling, prediction and simulation. Comput. Geosci. 24 (1), 17–31.
Pokhrel, P., Gupta, H.V., Wagener, T., 2008. A spatial regularization approach to parameter estimation for a distributed watershed model. Water Resour. Res. 44, W12419. http://dx.doi.org/10.1029/2007WR006615.
Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.J., 2004. Overall distributed model intercomparison project results. J. Hydrol. 298 (1–4), 27–60.
Reed, S., Schaake, J., Zhang, Z.Y., 2007. A distributed hydrologic model and threshold frequency-based method for flash flood forecasting at ungauged locations. J. Hydrol. 337 (3–4), 402–420.
Tang, Y., Reed, P., van Werkhoven, K., Wagener, T., 2007. Advancing the identification and evaluation of distributed rainfall-runoff models using global sensitivity analysis. Water Resour. Res. 43, W06415. http://dx.doi.org/10.1029/2006WR005813.
van Werkhoven, K., Wagener, T., Reed, P., Tang, Y., 2008. Rainfall characteristics define the value of streamflow observations for distributed watershed model identification. Geophys. Res. Lett. 35, L11403. http://dx.doi.org/10.1029/2008GL034162.
Wagener, T., van Werkhoven, K., Reed, P., Tang, Y., 2009a. Multiobjective sensitivity analysis of the information content in streamflow observations for distributed watershed modeling. Water Resour. Res. 45, W02501. http://dx.doi.org/10.1029/2008WR007347.
Yilmaz, K.K., Gupta, H.V., Wagener, T., 2008. A process-based diagnostic approach to model evaluation: application to the NWS distributed hydrologic model. Water Resour. Res. 44, W09417. http://dx.doi.org/10.1029/2007WR006716.