

Original Research Article

Stochastic Disaggregation of Daily Rainfall Using Barlett Lewis Rectangular Pulse Model (BLRPM): A Case study of middle Gujarat

ABSTRACT

Having accurate and ample data on rains is the sole golden input for deciding ultimate success of any progressive efforts towards natural resource management. Ultimate conquest of any pertinent schemes on developing and managing watersheds, canals, commands, irrigation net-works, soil-erosion, soil-conservation, drylands, forests, pastures, livestock, land use changes and many ecology-based errands; is entirely governs by the precision, relevancy and quality of rainfall data. Even the ending success of present days smart hydrologic models, modelling entirely remains regulated by the precision & relevance of rainfall data used therein. Most commonly available rain data happens to be daily rain values. However, for precise planning at microscale, we need to have its finer sub-daily temporal distribution. Rainfall disaggregation is a newly emerging applied option where utilities of advanced stochastic architecture is utilized across the globe to offer desired location specific and even rainy day specific best possible temporal disaggregated outcomes. Present paper offers some of the crisped outcomes from a detailed study performed in Gujarat. The predictive ability of one of the most popular BLRP model in this regard is shared by incorporating its basic architecture followed by its predictive performances on randomised sample rainy days covering 6 explicit locations in middle Gujarat region of western India. Preliminary findings reported herein will serve as a food for thought for smarter ways of managing water, land, watersheds and ecology.

Comment [RT1]: Put more emphasis on the need and potential benefits from model

Key Words: Rainfall Disaggregation, BLRMP, Stochastic Disaggregation, hydro-meteorological

INTRODUCTION

Precipitation is one of the most vital climatological information which is widely analysed on varied span of space and time for multiple utilities by hydrologists, water resource experts/planners/policy makers, watershed managers, irrigation planners, floods & drought regulatory experts and multiple agencies (government & non-government) dealing with plethora of agrarian as well as non-agrarian interventions. Analysis of rainfall data has high standing as it facilitates vital policy decisions by all above experts and agencies for fetching better levels of productions, protections, and end benefits for multiple stakeholders and societies. It remains more sensitive for farming issues like, cropping pattern, sowing date, constructions, soil & water conservation planning, domestic & irrigation-based water.

Comment [RT2]: Every statement should be supported by suitable references

Since rainfall happens to be weather-related key data which has its leading importance and utilities in kind of R&D under agricultural as well as non-agricultural domains. Accurate data on rains is the sole golden input for deciding ultimate success of any progressive efforts towards integrated management of lands, water, vegetation and ecological resources for delivering optimum benefits for farming interventions and societies involved therein. Success of any NRM based schemes on watersheds, canal commands, irrigation, soil conservation, Soil erosion controls, rainfed & irrigated agriculture, forests, pastures, livestock, land use and land cover etc.; is entirely governs by the precision, relevancy and quality of rainfall data. For entire domain and land and water management it remains main input information. The success of

present days smart hydrologic models, modelling entirely remains regulated by the precision and relevance of rainfall data used therein, which is still a weakest link for tropical countries like India. However, some of the variation in daily, monthly and annual rainfall for Indian situations; as caused by factors like, terrain elevation, slope and aspect are fairly explored by few researchers (). It provided some generic understanding on this crucial gap of information.

Variability of sub-daily rainfall as caused by above factors is even not amply explored even at global scale; and now a days emerging as one of the most prioritised R&D issues (Gaur, 2022). Further, the end possessions of multiple factors on properties of rain events such as duration, depth, and intensity, and length of time period between two consecutive events, (inter-event time) are not adequately explored. In other words, the rainfall studies are commonly restricted by non-availability of time series observations, where mere non-recording rain gauges serve as the main source of rainfall data. Further existence of data remains at discrete points, where rain gauges often located sparsely with uneven distribution in space. There remain many applied reasons/constraints in this regard. Say for example the rain gauges (even in developing countries) are commonly installed in towns or at locations that are easily approachable/connected with main roads and thus facilitate easy accessibility. As a result, relatively inaccessible areas such as sloppy lands, non-arable lands, forests, pastures, real outlets or rain-contributing zones of watersheds/sub-watershed and remotely located landscapes always remain uncovered by such observation network for rains. Thus, the regional/local water resource planners or watershed managers/policy makers, entirely compromise with whatever the level, type, accuracy and relevancy of rainfall record exists in their operational domain. Necessity with realized capabilities is the mother of invention. Under current smart era of technology and science, attempting realistic and proven synthetic transformations or renovations (e.g. desegregations) on existing daily rainfall data is appearing as one of the useful options for NRM based planning and operational executions. It remains equally vital for quantitative as well as qualitative management of water on different land parcels, having sizeably relevance towards river water qualities too (Paramar and Gaur, 2022).

Rainfall Disaggregation

As part of an effort to overcome rainfall data unavailability, various models/modelling approaches are projected in literature to synthetically generate rainfall data along with its dispersals. Examples include models that are based on dimensionless event hyetographs (Huff, 1967) and models that are based on the scaling properties of rainfall (El- Sayed ; 2018). Markov process based stochastic model are another segmented R&D effort in this direction, which had a bright significance toward developing effectual synthetic hyetographs for urban design for early or delayed peaking storms. A general mathematical technique for disaggregating rainfall for time scales finer than monthly was proposed by Koutsoyiannis (1992), which performed well encompassing an arbitrary and varying number of sub periods k , to facilitate the use of the model for the intermittent rainfall process.

A temporal rainfall disaggregation model to convert daily time series into an hourly resolution was proposed by Guntner *et al.* (2001). They also found that the model reproduces a range of hourly rainfall characteristics with a high accuracy in both climates. However, the overall model performance was better for the semi-arid tropical rainfall. The overall high accuracy of disaggregated data supports the potential usefulness of the model in hydrological applications. Many such works have reported that optimal designs of stormwater systems remained vastly rely on the rainfall IDF curves for data sparse sites/regions, moreover recently, Muller *et al.* (2018) analysed the influence of disaggregated rainfall products with different

degrees of spatial consistence on rainfall–runoff modelling. For the disaggregation of daily rainfall time series into hourly values, a multiplicative random cascade model was applied. Their findings suggested that simple model structures might compensate for deficiencies in spatial representativeness through parameterization and (ii) highly resolved hydrological models benefit from improved spatial modeling of rainfall. Muller and Sikorska (2019) compared a simple method (M1, one parameter), focusing on the effective precipitation duration for flood simulations, with a multiplicative cascade model (M2, 32/36 parameters). The results indicated that differences identified in precipitation characteristics of disaggregated time series vanish when introduced into the lumped hydrological model. Moreover, flood peaks were more sensitive than flood volumes to the choice of disaggregation method. Likewise, a few other researches (Miao *et al.*, 2016) too have reported their findings in this regard, where the purpose remained confined to either the derivation of threshold values or the generalized associated uncertainties for rainfall to meet the sole purpose of flood warning and other weather-related meteorological projections.

MATERIAL AND METHOD

Comment [RT3]: Elaborate how you approached and completed the study

Stochastic Disaggregation of Rainfall

The ability of the Barlett Lewis Rectangular Pulse Model (BLRPM) to reproduce important features of rainfall data ranging from an hourly to daily scale and higher has been well discussed by Rodriguez-Iturbe *et al.* (1987) reflecting vital features for representing rainfall in continuous time series including multiple calibrations & applications to several climates across the globe. There remains 5 general inherent assumptions while opting and applying a BLRP rainfall disaggregation model: (1) Storm origins t_i occur following a Poisson process with a rate λ ; (2) Each storm is characterized by a random number of cells, C ($C \geq 1$), and each storm origin is followed by a Poisson arrival at rate β of the cell's origin; (3) The intervals between successive cells are independent and identically distributed random variables, starting at each t_i , the cell arrival process terminates after an exponentially distributed time v_i with parameter γ ; (4) Each cell has an exponentially distributed duration of parameter η ; and (5) a uniform intensity X_{ij} with a specified distribution and mean μ_x (Fig. 1).

Model typically assumed to be exponential (parameter $1/\mu_x$), or alternatively a two-parameter gamma with mean μ_x and expected mean square of cell intensity μ_x^2 . The number of cells per storm, C , thus has a geometric distribution of mean $\mu_c = 1 + k/\phi$ where $k = \beta/\eta$ and $\phi = \gamma/\eta$ are dimensionless parameters. Therefore, for this BLRP model there are five parameters governing the process: λ , k , μ_x , ϕ and η , and all parameters are assumed to be constant. In order to overcome generation of fewer wet periods than required, and underestimates the proportion of dry periods added one extra parameter the mean cell duration, $1/g$, which is allowed to vary randomly from storm to storm. The parameter η now follows two-parameter gamma distributions with a shape parameter α and scale parameter ν . Subsequently, parameters β and γ also vary such that the ratios $k = \beta/\eta$ and $\phi = \gamma/\eta$ become constant. Each cell depth is a random constant that is exponentially distributed with mean $E[x]$. This results in a six-parameter model (λ , k , μ_x , ϕ , α and ν).

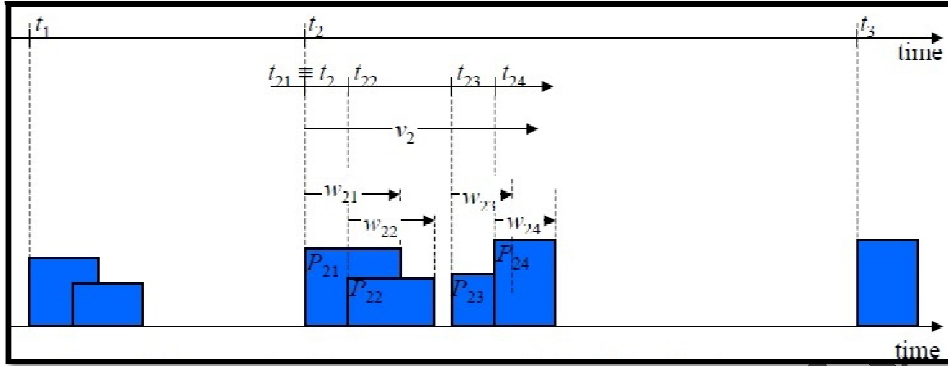


Fig 1. General conceptual configuration for understanding BLRP process

BLRPM Parameter Estimations

Couple of leading researchers (Christoph, 2017; Rodriguez-Iturbe *et al.* 1987) have recommended equalizing the characteristic features computed from the observed data with the corresponding model values. Four statistical properties are required to estimate the parameters of the BLRPM, *i.e.*, the mean, variance, lag-1 auto-covariance coefficient, and proportion of dry periods at 1, 24 and 48 h levels of aggregation. Letting k be the number of parameters to be fitted, p the statistics to be chosen from the historical data to fit the parameters, and these are denoted by the set T , where $T = (t_1, t_2, \dots, t_p)$, which can include the mean, variance, etc. of various timescales. Then the functions to calculate the various statistics from the parameter values in the BLRPM are using the following equations:

$$\text{Mean} = \lambda\mu \times \mu c \frac{v}{\alpha-1} T \quad (\text{Eq.1})$$

$$\text{Variance} = \frac{2v^{2-\alpha}T}{\alpha-2} \left(k - \frac{k_2}{\varphi} \right) - \frac{2v^{3-\alpha}}{(\alpha-2)(\alpha-3)} \left(k_1 - \frac{k_2}{\varphi^2} \right) + \frac{2}{(\alpha-2)(\alpha-3)} \left[k_1(T+v)^{3-\alpha} - \frac{k_2}{\varphi} (\varphi T + v)^{3-\alpha} \right] \quad (\text{Eq.2})$$

Where,

- i. $k_1 = \left(2\lambda\mu c E^2[x] + \frac{\lambda\mu c^k \varphi E^2[x]}{\varphi^2-1} \right) \left(\frac{v^\alpha}{\alpha-1} \right)$
- ii. $K_2 = \left(\frac{\lambda\mu c^k E^2[x]}{\varphi^2-1} \right) \left(\frac{v^\alpha}{\alpha-1} \right)$

$$\text{Prob}[\text{zerorainfall}] = \exp \left\{ -\lambda T - \left[\frac{\lambda v}{\varphi(\alpha-1)} \left(1 + \varphi \left(k + \frac{\varphi}{2} \right) - \frac{1}{4} \varphi (k + \varphi) (k + 2\varphi) \right) (k + 4\varphi) + \frac{\varphi(k+\varphi)(4k^2+27k\varphi+36)}{72} \right] + \frac{\lambda v}{(\alpha-1)(k+\varphi)} \left(1 - k - \varphi + \frac{3}{2} k\varphi + \varphi^2 + \frac{k^2}{2} \right) + \frac{\lambda v}{(\alpha-1)(k+\varphi)} \left(\frac{v}{v+(k+\varphi)T} \right)^{\alpha-1} \frac{k}{\varphi} \left(1 - k - \varphi + \frac{3}{2} k\varphi + \varphi^2 + \frac{k^2}{2} \right) \right\} \quad (\text{Eq.3})$$

$$\begin{aligned} \text{Autocovariance}(lags) = & \frac{k_1}{(\alpha-2)(\alpha-3)} \{ [T(s-1) + v]^{3-\alpha} + [T(s+1) + v]^{3-\alpha} - \\ & 2(Ts + v)^{3-\alpha} \} + \frac{k_2}{\varphi^2(\alpha-2)(\alpha-3)} \{ 2(\varphi Ts + v)^{3-\alpha} - [\varphi T(s-1) + v]^{3-\alpha} - [\varphi T(s+1) + \\ & v]^{3-\alpha} \} \end{aligned} \quad (\text{Eq. 4})$$

Since the components have different orders of magnitude, they will be first normalized to bring the adjusted values close to the historical rainfall. This will be achieved by calculating the error-residual term, Z using below equation.

$$Z = \min \left[\sum_{i=1}^N W_i \left(\frac{F_i(x)}{F_i'} - 1 \right)^2 \right] \quad (\text{Eq. 5})$$

RESULTS AND DISCUSSION

A vast magnitude of rainfall data domain was dealt in this study, where 20 years long daily rainfall records for 6 important rain gauge stations in study region were critically recognised, screened, analysed, and interpreted on advanced R-platform. Using BLRPM. After multiple runs, the best possible and most optimistic sets of location specific parametrization was attained. The blended sets of on-ground as well as satellite based observed rainfall record were suitably utilized for developing large number of relevant sets of BLRPM parameters.

Majority of results revealed adequate relevancy of such stochastic rainfall desegregation approach for the area and location adopted in present study. Plethora of outcomes were attained, however owing to paucity of space, only a few sample results (random based) are made part of present segment of manuscript. Location specific standard sets of synthetic matrices of parameters for stochastic predictability of rainfall values in the shape of sub-daily temporal configuration was attained with ample magnitudes of time series precipitations data; as utilized and demonstrated herein. Upon simultaneous comparisons of end products from stochastic disaggregation, couple of outcomes, which are of foremost practical utilities for area under study are produced below. They could enhance the quality of key input domain (basically the intensity, duration, excess rainfall, design hyetographs etc) that is commonly utilized in majority of rainfall-runoff models as well as modelling approaches adhered therein.

Comment [RT4]: In discussion part simulate your each study outcome with similar conclusions drawn in other studies with references

Synthetized Matrix of BLRPM Parameters for Study Locations

Preliminary results obtained by adopting BLRPM approach via its standard sets of procedural protocols; remained highly encouraging. was used to disaggregate daily rainfall values in to sub daily (hourly) rain depth. First and foremost, step in deploying this model remained estimation of location specific matrix of model parameters using the historical rainfall depths at higher temporal resolution. Such model parameters for all the 6 station of study area were estimated using the 21 years rainfall records. The estimated parameter values are furnished in table 1.0. Using these values, it is quite conceivable to get disaggregated rainfall values (from daily to sub daily) by mere deploying BLRPM.

The earnest effort by using R package transformation of BLRPM, made it feasible to generate storms by espousing number of location specific parameters as generated herein for 6 explicit

locations (RStudio, 2020). The magnitudes, ranges, and the trends of variabilities in such parameters remains highly diverse and uncertain. A relative contrast of such variabilities is illustrated in Table 1, where seven key parameters of BLRPM model are reflected for four key monsoon months, including all six locations where equivalent ground-based precipitation data too existed. Out of total annual records, these 4 months data (June to September) reflected wider variability domain on the scale of time and location. The information projected in Table 1, is self-explanatory to portray the explicit sphere of magnitudes and trends of such variations.

Table 1 Characteristics of appraised BLRPM parameters for six study sites

Monsoon Months	Generated BLRPM Parameters [#] for Different Locations				
	lamda (λ)	Kapa ($\kappa = \beta/\eta$)	phi ($\phi = \gamma/\eta$)	ni (ν)	mi_X (μ_X)
Rain Station : Dakor Gujarat					
June	0.004	0.080	0.001	12.977	99.000
July	0.012	0.069	0.001	11.647	83.841
August	0.193	0.016	0.005	20.878	70.325
Sept	0.104	0.060	0.040	33.540	77.336
Rain Station : Kathlal Gujarat					
June	0.004	0.061	0.001	7.710	99.000
July	0.004	0.061	0.000	11.648	83.841
August	0.040	0.041	0.002	16.485	68.296
Sept	0.082	0.001	0.000	32.627	76.643
Rain Station : Gudel Gujarat					
June	0.002	0.083	0.001	12.972	99.000
July	0.012	0.040	0.000	11.548	83.823
August	0.193	0.001	0.000	20.814	70.304
Sept	0.003	0.061	0.001	32.015	74.052
Rain Station : Sojitra Gujarat					
June	0.002	0.084	0.001	12.975	99.000
July	0.004	0.057	0.000	8.913	83.561
August	0.011	0.030	0.000	17.423	68.457
Sept	0.001	0.104	0.000	13.557	77.281
Rain Station : Pilol Gujarat					
June	0.004	0.080	0.001	12.994	99.000
July	0.006	0.067	0.000	11.647	83.841
August	0.197	0.011	0.004	20.876	70.324
Sept	0.011	0.064	0.002	15.453	95.206
Rain Station : Waghodiya Gujarat					
June	0.003	0.087	0.001	12.971	99.000
July	0.007	0.052	0.000	11.021	83.752
August	0.204	0.002	0.001	20.687	70.262
Sept	0.227	0.000	0.000	33.321	77.230

[#]Note: Among other two parameters, the values in respect of Alpha (α) in majority of cases remained 99, however the sets of numerical values for parameters namely 'sigma_X (σ_X)' remained exactly identical as of 'mi_X (μ_X)', hence excluded in this Table.

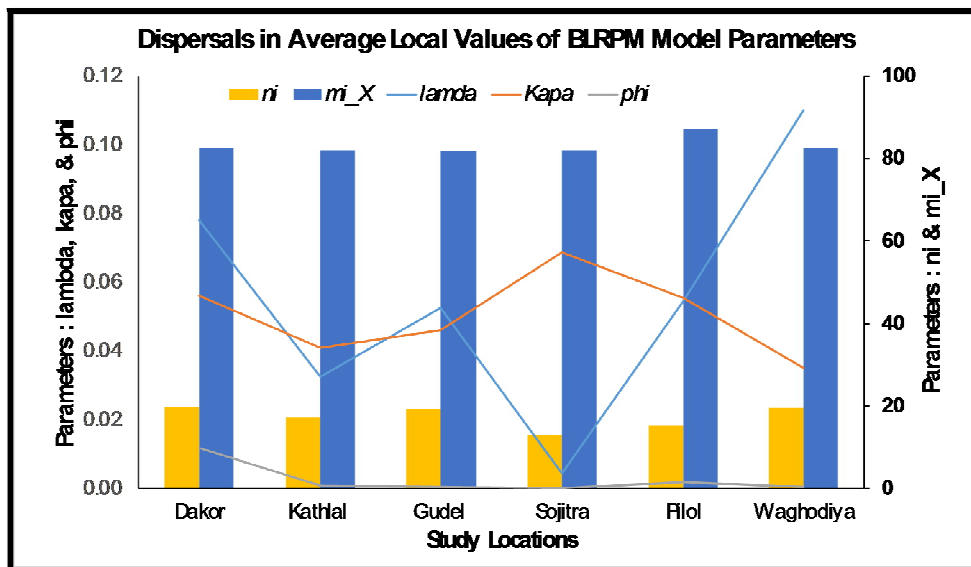


Fig. 2 Comparison of numerical dispersals in location specific key parameters of BLRPM disaggregation model for study region

Experimental Validation for Daily Rainfall Values

The resulted simulated storms events are pictorially described in figure 2.0, which showed highly cheering trends and perfections of applied utilities for totally ungauged situations in the region. Moreover, being unreal and stochastic in nature the overall end deliverables via BLRPM remained merely at coarser resolutions of time step; hence were not sufficiently capable to mimic the real storm events what actually happens and modelled for real ground scenarios for 6 different stations in middle Gujarat region of India.

Beside logical comparisons of above narrated BLRPM parameters, an effort was also made to visualize and realize the reliability and utilities of these region-specific values for predicting r simulating daily and thereafter sub-daily precipitations. In the first instant, the ultimate comparison of observed and predicted disaggregated rains was judged for a sample study station (Dakor). Six randomly selected rainy days across 21 years massive archives, are used in devising a sample pictorial view to compare their respective observed and predicted values (Fig. 3). It very well illustrates the decisive contrasts, showing the valid trends in predictions of disaggregated rain entities for a particular location. Similar exercises were done for all other stations and records too, where no homogeneous and well acceptable trends of relevancies were noticeable. Owing to paucity of space all those results are excluded herein, moreover the biggest inference from this particular exercise remains a recommendation that such stochastic approaches remains highly uncertain, unrealistic and unreliable for modelling natural entities like rains; even when location specific parameters are generated. One must remain more reliable on truly observed precipitation and its temporal distributions; followed by location specific standard or synthetic type curves that are extracted from real observed fine scales (space & time) time series of precipitations data; as utilized and demonstrated in this study.

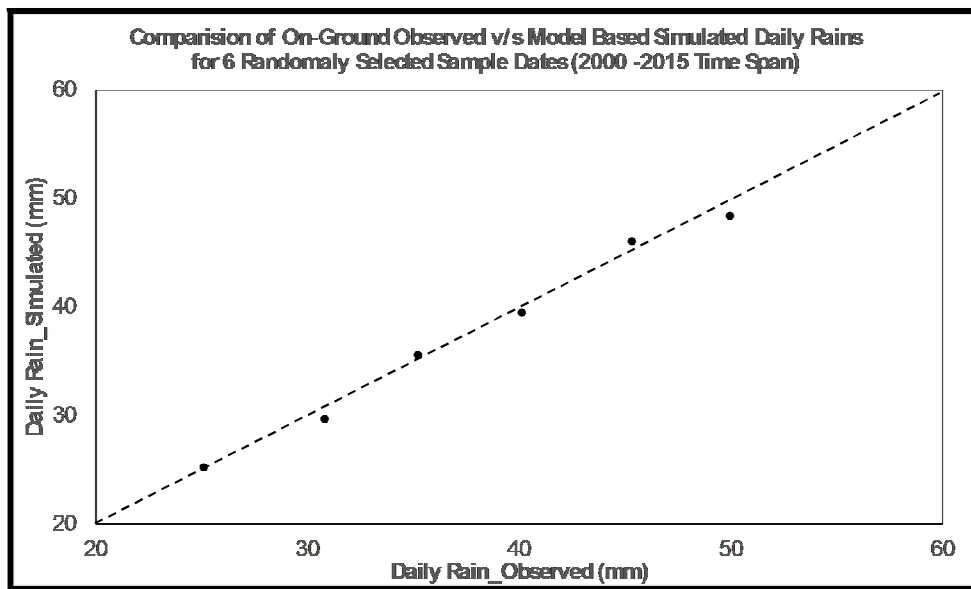


Fig. 3 Comparison of observed and simulated daily rains for randomly selected few sample matching rainy days at six different location in study region

Model Ratification on Disaggregated Sub-Daily Rainfall Values

Variety of trial runs were performed to generate numerous sub-daily time distributions from a single value of daily rainfall (i.e. of 24 hours duration 8 am to 8 am). The ultimate equivalences across paired sets of ground observed daily precipitation depth and satellite based daily rains were found highly matched (predictive efficiencies > 95 % in majority of cases). Moreover, as an advanced effort the individual daily rains were transformed into sub-daily values adopting enormous time steps (1, 3, 6, 12 hour). These synthetically generated i.e. disaggregated values were compared in hundreds of permutation and combinations for very high number of days cutting across 20 years time span over 6 different locations in study region.

Though results for coarser time steps were significantly better, still here a sample set of end results in this regard is presented for a condition where the adopted sub-daily rain time step is minimum i.e. 1 hour only. These detections and inclusive simulation results as reflected in dozens of sub-plots in Fig. 4, are self-explanatory to offer a clear-cut contrast of paired sets of observed and simulated sub-daily rains over a 24-hour duration for varying locations and randomly identified rain day i.e. dates. Using so arrived disaggregated rains from daily single value of rain, will certainly remain a valued asset for attaining realistic sets of rainfall hyetographs for these areas, which in turn will add values to plethora of rainfall runoff models and modelling approaches for working out rainfall-runoff relationships on variety of natural watersheds under diverse soil-vegetation-cover complexes and land use patterns.

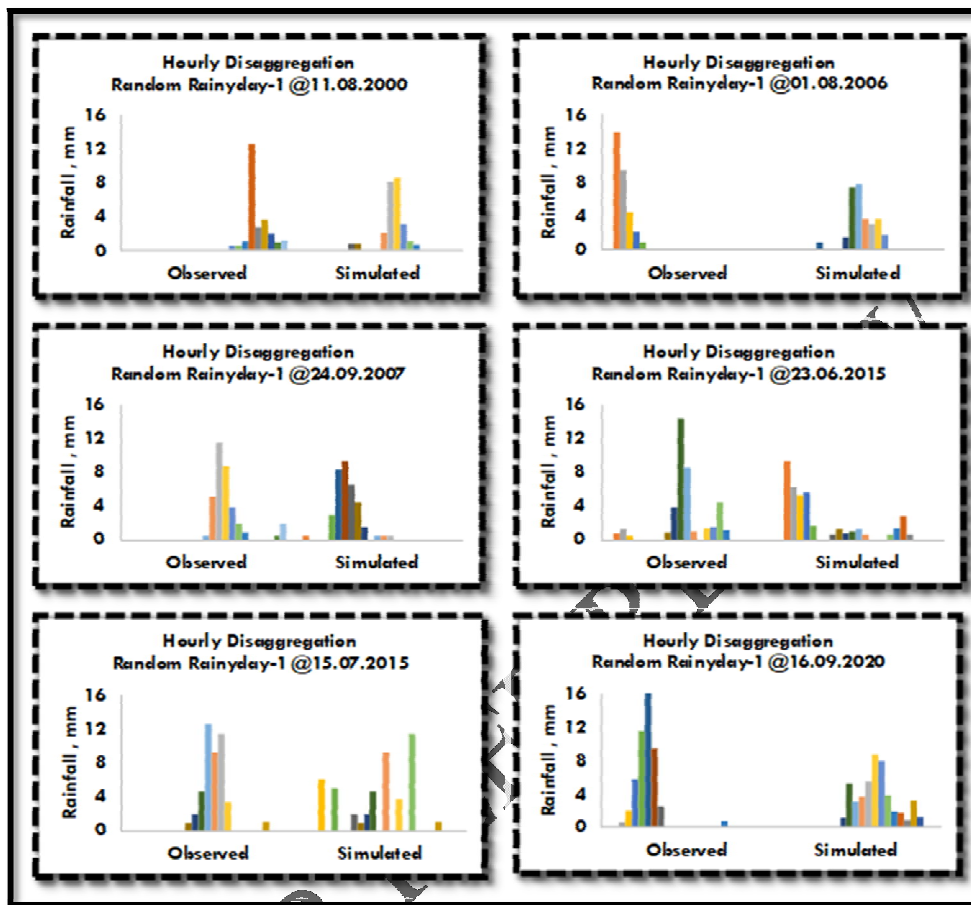


Fig. 4 Predictive performance of BLRPM used with its location specific parameters for randomly opted rainy days (6 nos) for a sample study location (Dakor)

So arrived results were meticulously judged and linked to compare the patterns of observed versus simulated values for identical set of conditions. Findings from this study revealed that contrast of equivalences or matching gets deteriorated as we shorten the time step. But the overall results very well established the fact that under ungauged situations where sub-daily rain records are scarce or non-existence; the model and approach adopted in present study could prove as a boon for all purposes in watershed and water-based planning and management interferences. Presentation of all such results are beyond scope, owing to paucity of space.

Comment [RT5]: Highlight limitations of the study

CONCLUSION

In the present work an extensive analysis of daily rainfall was carried out for longer time span of about 2 decades, encompassing 6 explicit random locations in middle Gujarat region of western India, where blended records of ground based as well as satellite-based precipitations existed. Valid magnitudes and ranges of location specific BLRPM parametrization remained

decidedly realistic capable to produce truthful and valid projections of sub-daily rainfall estimates using merely the daily rainfall values. Level of predictive performances was adequately satisfactory, except some dispersals in hourly temporal distribution of project rains. Ultimate findings and conceptual framework of present study established importance of rainfall disaggregation from daily to sub-daily for its shorter as well as many larger end benefits in NRM subject domains specifically the land, water, and watersheds.

Comment [RT6]: Check for grammatical mistakes and plagiarism
Add some more references

REFERENCES

- Christoph Ritschel (2017). BLRPM: Stochastic Rainfall Generator Bartlett-Lewis Rectangular Pulse Model. R package, version 1.0. <https://CRAN.R-project.org/package=BLRPM>
- El- Sayed, E. A. H. (2018). Development of synthetic rainfall distribution curves for Sinai area. *Ain Shams Engineering Journal*, 9(4), 1949–1957. <https://doi.org/10.1016/j.asej.2017.01.010>
- Gaur, M.L. 2022. Catchment scale forest water interfaces for pollution management. In: Yadav, S., Negm, A.M., Yadava, R.N. (eds) *Water Quality, Assessment and Management in India*. Earth and Environmental Science Library. Springer, Cham, pp. 71-111.
- Guntner, A., Olsson, J., Calver, A., & Gannon, B. (2001). Cascade-based disaggregation of continuous rainfall time series: the influence of climate. *Hydrology and Earth System Sciences*, 5(2), 145–164. <https://doi.org/10.5194/hess-5-145-2001>
- Huff, F. A. (1967). Time distribution of rainfall in heavy storms. *Water Resources Research*, 3(4), 1007–1019. <https://doi.org/10.1029/wr003i004p01007>
- Koutsoyiannis, D. (1992). A nonlinear disaggregation method with a reduced parameter set for simulation of hydrologic series. *Water Resources Research*, 28(12), 3175–3191. <https://doi.org/10.1029/92wr01299>
- Miao, Q., Yang, D., Yang, H., & Li, Z. (2016). Establishing a rainfall threshold for flash flood warnings in China's mountainous areas based on a distributed hydrological model. *Journal of Hydrology*, 541, 371–386. <https://doi.org/10.1016/j.jhydrol.2016.04.054>
- Muller T. H., & Sikorska, A. E. (2019). Does the complexity in temporal precipitation disaggregation matter for a lumped hydrological model? *Hydrological Sciences Journal*, 64(12), 1453–1471. <https://doi.org/10.1080/02626667.2019.1638926>
- Muller, T. H., Wallner, M., & Förster, K. (2018). Rainfall disaggregation for hydrological modeling: is there a need for spatial consistence? *Hydrology and Earth System Sciences*, 22(10), 5259–5280. <https://doi.org/10.5194/hess-22-5259-2018>
- Parmar, Urvashiben and Gaur Murari Lal. 2022. Judging spatio-temporal variability of key river water quality vis-à-vis land use changes in Mahi Basin. *Journal of Soil and Water Conservation* 21(4): 385-393, October-December 2022. ISSN: 022-457X. doi: 10.5958/2455-7145.2022.00049.2
- Rodriguez-Iturbe, I., de Power, B. F., & Valdes, J. B. (1987). Rectangular pulses point process models for rainfall: Analysis of empirical data. *Journal of Geophysical Research*, 92(D8), 9645. <https://doi.org/10.1029/jd092id08p09645>
- RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.