Including Uncertainty with Doppler-Radar Estimates

Valmir Bucaj[a] [1], Tyson Walsh[b], Michael Scioletti[c], Theodore Hromadka II[d]

[a,b,c,d] Department of Mathematical Sciences, United States Military Academy, West Point, NY, USA

Abstract

Applications that rely on Doppler radar estimates of precipitation include hydro-meteorology, engineering, floodplain management, and weather forecasting. In this article we provide a methodology to quantify, in a probabilistic sense, the uncertainty associated with Doppler radar estimates of precipitation and the propagation of such uncertainties in the rainfall-runoff model. Through multiple Monte-Carlo type simulations we demonstrate the variation in these calculations related to the uncertainty of estimation and a practical way in which the engineering management of reservoirs may use this methodology to assist them in decision-making.

Keywords: Radar Precipitation Uncertainty, Rainfall-Runoff Model, Monte-Carlo Simulations

1 Introduction

Managing water storage and flow is essential in hydro-meteorology, engineering, floodplain management, weather forecasting, and water storage reservoir routing studies. Current practices to do this rely on estimates of precipitation, which are inaccurate. Walsh et. al., see [26] attempts to quantify the uncertainty associated with the Doppler Radar WSR-88D by developing distributions of Doppler Radar values versus the actual precipitation measurements collected by rain gauges during the prediction timeline. The authors indicate there is a high degree of uncertainty associated with Doppler radar estimates of precipitation and that a point estimate prediction may not be appropriate. As a result, appropriate distributions of the candidates for the gauge precipitation measurements on the ground are built by Doppler radar estimate bands ranging in width based on number of observations.

Using current rainfall data, which we define as a coupled pair, namely Doppler Radar Estimated Precipitation (DREP) and Gauge Estimated Precipitation (GEP), it is possible to track storm cells as well as to predict precipitation quantities in order to asses possible flow runoff and make flow release strategy adjustments to reduce downstream flood risk. However, due to uncertainties associated with such Radar data-based analysis, further insight into the flood risk is possible by inclusion of the possible uncertainty with storm runoff estimates. The paper by Berne et. al., [4], provides a backdrop as to some of the issues currently encountered in using Doppler Radar for hydrologic studies. Because a modeling relationship is defined in

[1] Corresponding author. Email: Valmir.Bucaj@westpoint.edu

the transformation of precipitation to runoff process, the uncertainty associated with the Radar estimates cascades into the runoff estimates, producing distributions of runoff as a function of the distribution of the Radar estimate of precipitation versus the precipitation gauge readings.

Much work has been done in recent years to quantify the error associated with radar rainfall estimates. For example Rossa et. al in [23] provides a fairly comprehensive review of the progress made in quantifying the uncertainty associated with observed and forecasted precipitation and the way these uncertainties propagate in various hydrological models. For a more in-depth analysis we direct the reader to the work of Ciach et. al. [9], Villarini et. al. [25], Quintero et. al. [21], Rico-Ramirez et. al. [22] etc.

In this paper, we improve upon [26] and extend the analysis further by providing a methodology to determine the uncertainty associated with Doppler radar estimates of precipitation and then propagate it in computations of water runoff estimates. After presenting the methodology we run Monte-Carlo type simulations using the data compiled by [26] to demonstrate the variations in outcomes associated with the uncertainty.

Some of the main advantages of our methodology include its generality, data-driven nature, ease of implementation, applicability in a variety of hydrological and hydraulic models, and its usefulness as a tool in helping the engineering management of reservoirs in decision-making. On the flip side, one possible draw-back of this methodology is the fact that its accuracy depends on the number of the coupled data points (DREP, GEP). In other words, the larger the number of data-points, the more reliable and accurate the outcomes of the methodology, and vice versa. We will elaborate on this point further when we conduct Monte-Carlo type simulations in Section 3.

The statistical analysis, visualizations, and the code generated as part of this research were all done in Python. Specifically, some of the packages that we used for statistical analysis were: Pandas, Numpy, Stats package from SciPy etc.; for visualizations: Matplotlib, Seaborn, JoyPy etc.

The preliminary results of this paper were presented at the ASFPM and AWRA conferences, see [7], [5].

2 Methodology

In this section we present an algorithm which, in a probabilistic sense, quantifies the uncertainty associated with radar precipitation estimates and the propagation of such uncertainties in various hydorlogical and hydraulic models for flood forecasting. We will refer to this algorithm as the *HBWS algorithm*, after the authors.

Step 0: Data Pre-processing. We standardize each Doppler Radar estimate with associated rain gauge values, that reflect the actual precipitation collected, to have a mean of zero and a standard deviation of 1. We denote this coupled pair of (DREP, GEP) as $(\mathcal{D}, \mathcal{G})$. We then create a string, which mimics a Doppler

Radar prediction over time, comprised of K DREP measurements, $\{a_i\}_{i=1}^{K}$, in standard deviations.

Step 1: Probability Density Functions. Let $X_{\mathcal{G}}$ and $X_{\mathcal{D}}$ denote the random variables for $\mathcal{G}$ and $\mathcal{D}$, respectively, and let us denote their joint probability density function by $f_{X_{\mathcal{G}},X_{\mathcal{D}}}(g,d)$. For each a_i and a given *discretization scale* $\epsilon > 0$ we, formally, define the desired family of conditional probability density functions, $(f_{a_i,\epsilon})_{\epsilon>0}$, by

$$f_{a_i,\epsilon}(g) \stackrel{\text{def}}{=} f_{X_{\mathcal{G}}}\left(g|X_{\mathcal{D}} \in \mathcal{A}_{a_i,\epsilon}\right)$$

where $\mathcal{A}_{a_i,\epsilon} = [a_i - \epsilon, a_i + \epsilon]$. One should think of the $f_{a_i,\epsilon}$ as the probability density function for the subset of $\mathcal{G}$ whose $\mathcal{D}$ counterparts belong in the interval $\mathcal{A}_{a_i,\epsilon}$.

Assuming finite data sets, the ϵ is necessary, as otherwise the domain of $f_{a_i,\epsilon=0}$ may be empty or lack sufficient entries, which is problematic. Ideally ϵ is minimized for each $f_{a_i,\epsilon}$.

Step 2: Kernel Density Estimators. Let $\widehat{f}_{a_i,\epsilon}$ denote an *estimator* of the probability density function $f_{a_i,\epsilon}$ defined in *Step 1*, and let $\{g_k\}_{k=1}^{n_i}$, be such that $(d_k, g_k) \in (\mathcal{D} \cap \mathcal{A}_{a_i,\epsilon}, \mathcal{G})$, for $k = 1, \ldots, n_i$, where n_i is maximal. Then

$$\widehat{f_{a_i,\epsilon}}(g) \stackrel{\text{def}}{=} \frac{1}{n_i h\sqrt{2\pi}} \sum_{k=1}^{n_i} e^{-\left(\frac{g-g_k}{\sqrt{2}h}\right)^2},$$

where $h > 0$ is some appropriately chosen smoothing parameter. For a more elaborate discussion on kernel density estimates you may see [20].

Step 3: Distributions. For each a_i, the distribution of the probable daily gauge precipitation outcome readings, denoted by $\mathcal{V}_{a_i}$, is the one resulting from the kernel density estimate, $\widehat{f}_{a_i,\epsilon}$, in *Step 2.*

Step 4: GEP Realizations. For each a_i, the set of probable candidates for daily gauge precipitation readings on the ground, denoted by $\mathcal{B}_i$, is obtained via randomly sampling from the distribution $\mathcal{V}_{a_i}$, constructed in *Step 3.* Specifically, $\mathcal{B}_i = (\mathcal{O}_{a_i}^j)_{j=1}^{N}$, where each of these N randomly sampled outcomes, denoted by $\mathcal{O}_{a_i}^j$, is a candidate for an estimate of the gauge precipitation measurement on the ground corresponding to a_i.

Step 5: Strings of GEP realizations. To obtain *strings* of probable candidates for the gauge precipitation measurements on the ground we iteratively select an element from each $\mathcal{B}_i$, and cast them as such in a new list. We then obtain *strings* of the form $\mathcal{C}^{(j)} \stackrel{\text{def}}{=} \left(\mathcal{O}_{a_1}^j, \mathcal{O}_{a_2}^j, \ldots, \mathcal{O}_{a_K}^j\right)$, where each $\mathcal{C}^{(j)}$ represents a probable outcome for the gauge precipitation reading on the ground corresponding to the *string* of DREP measurements $\{a_1, a_2, \ldots, a_K\}$.

Step 6: Application Models. We cascade $\mathcal{C}^{(j)}$ into various application models, such as *rainfall-runoff models, water conservation, soil-strength analysis* etc.

In this paper, we will focus on the *rainfall-runoff* model, leaving the treatment of other types of models for future projects.

3 Practical Simulations

Using the same collection of paired data, $(\mathcal{D}, \mathcal{G})$, as in [26] we will run a Monte-Carlo type simulation demonstrating the practical use of the *HBWS algorithm* to the *rainfall-runoff* model. To our knowledge, this procedure provides a novel way to account for the propagated uncertainty associated with the radar estimates of precipitation in runoff estimates.

For demonstration purposes, for the rest of this paper, we will use the following *string* of $K = 13$, arbitrarily chosen, daily DREP measurements, in standard deviations,

$$\mathcal{I} \overset{\text{def}}{=} \{-0.3, 0.1, 0.4, 0.75, 1, 1.4, 1.9, 2.1, 2.3, 2.65, 3.0, 3.25, 3.8\}. \tag{1}$$

We use $\epsilon = 0.125$, which is minimized.

3.1 Distributions For GEP Measurements

Applying Steps 0 through 3 from the *HBWS algorithm*, we build the distributions for the gauge precipitation estimates corresponding to each of the DREP reading in (1). These distributions are displayed in Figure 1 below.

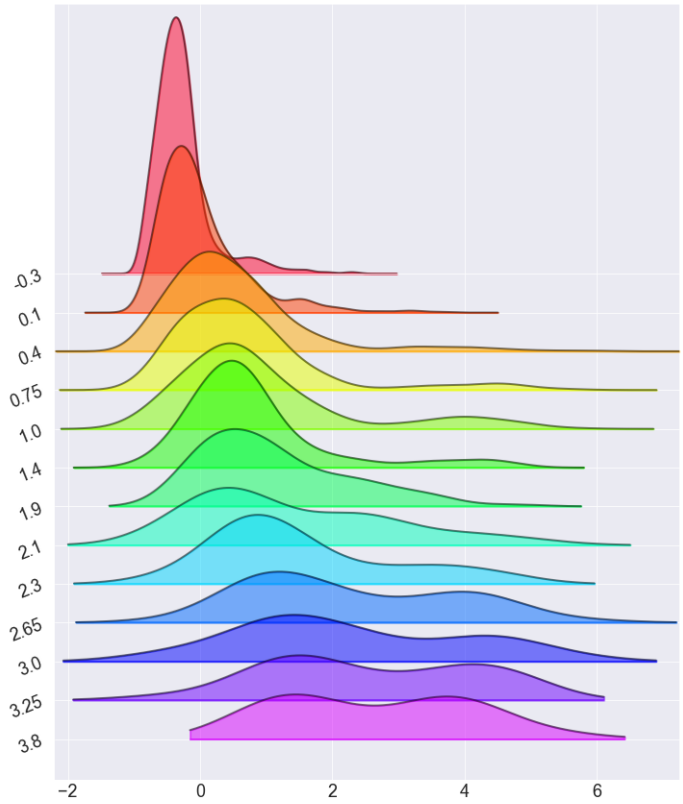

Figure 1: Distributions for GEP measurements (in standard deviations)

The computational results stemming out of this methodology may depend heavily upon choices made for the discretization scale, time scale, and other measured Doppler-Radar attributes. For illustration

purposes, we conduct a visual comparison of a few GEP distributions built with different discretization scales. Specifically, for the following subset of $\mathcal{I}$ values $\{-0.3, 1.4, 2.3, 3.25, 3.8\}$ we build a corresponding GEP kernel density estimate (see Step 2) for $\epsilon = [0.05, 0.1, 0.125, 0.25]$, respectively. These are illustrated in Figures 2 through 6 below.

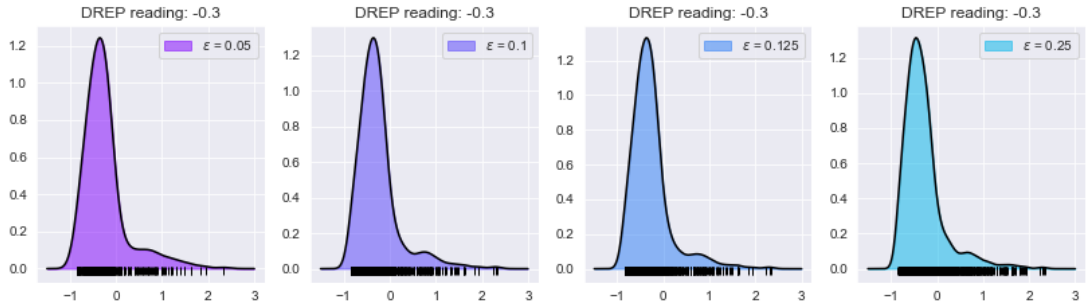

Figure 2: GEP marginal distributions for a DREP reading of -0.3 with different discretization scales

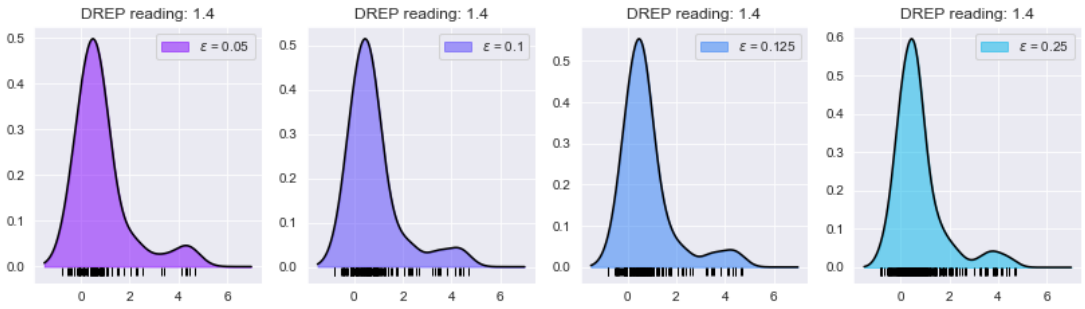

Figure 3: GEP marginal distributions for a DREP reading of 1.4 with different discretization scales

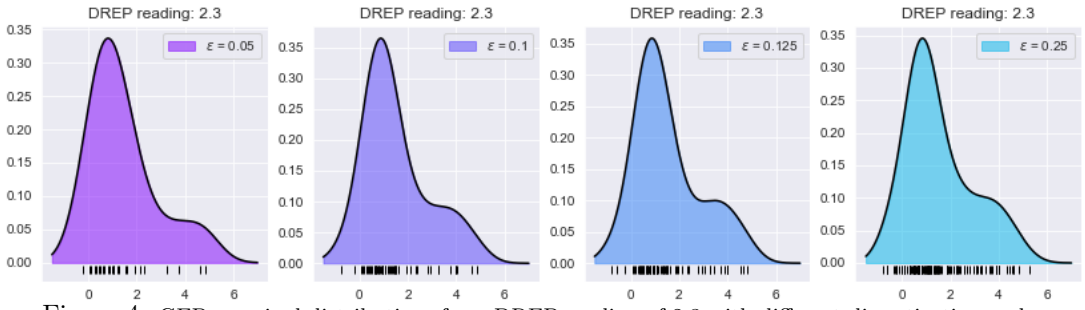

Figure 4: GEP marginal distributions for a DREP reading of 2.3 with different discretization scales

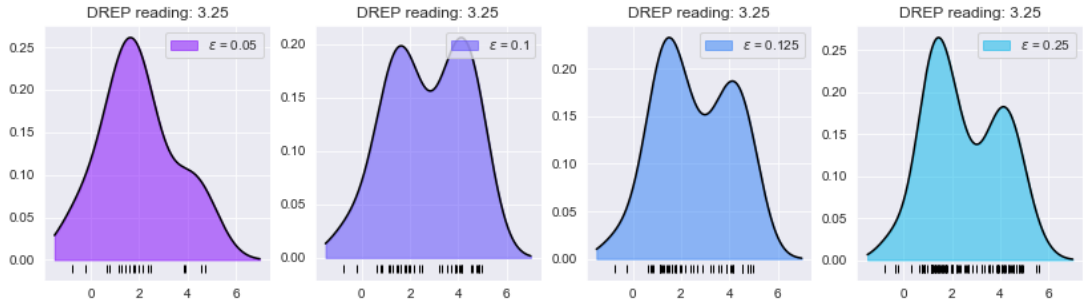

Figure 5: GEP marginal distributions for a DREP reading of 3.25 with different discretization scales

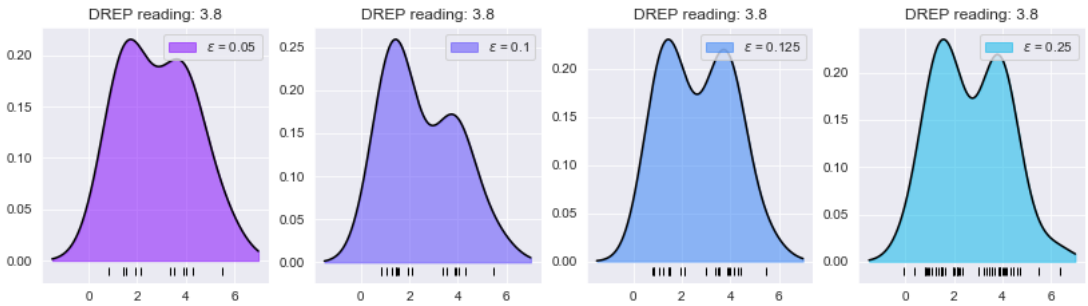

Figure 6: GEP marginal distributions for a DREP reading of 3.8 with different discretization scales

3.2 GEP realizations over thirteen Doppler-Radar daily measurements

Applying Steps 4 and 5 from the *HBWS algorithm*, for each DREP reading in $\mathcal{I}$, we generate $N = 50$ *strings* of GEP realizations. Then, we compute the $25^{th}, 50^{th}$ and 85^{th} percentiles for each of these generated probable outcomes of gauge precipitation measurements. Figure 7 gives a visual representation of these results.

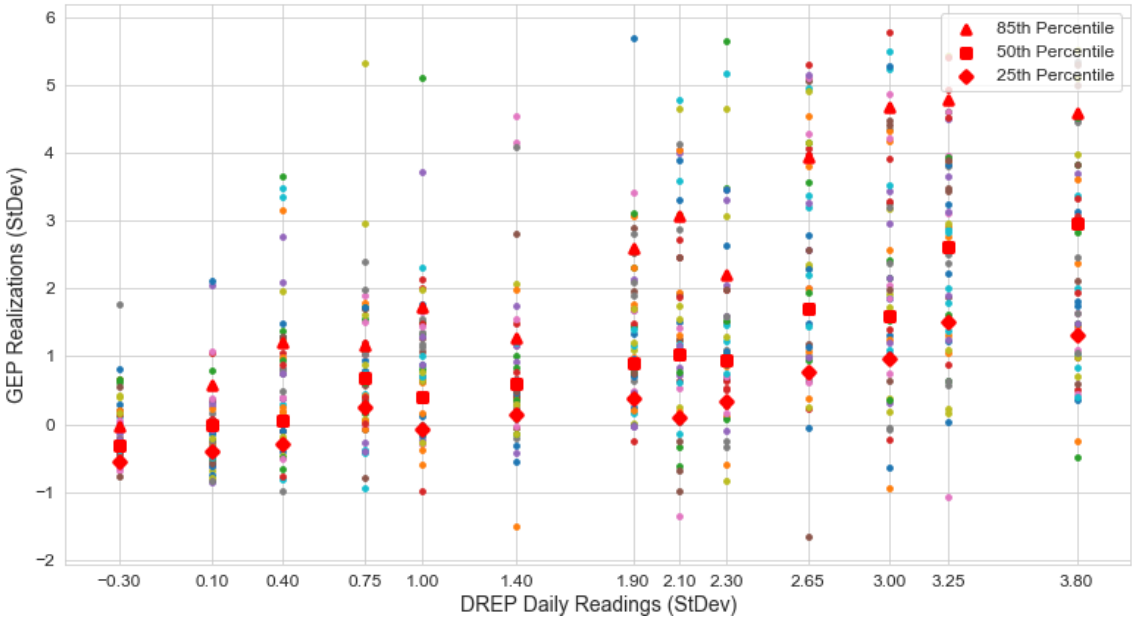

Figure 7: N=50 samples – corresponding to the 50 estimated GEP Realizations, each tracked over 13 Doppler-Radar daily measurements

For example, focusing on the Doppler Radar daily measurement of 0.1, these calculations infer that 85% of the probable outcomes of precipitation will be below 0.604, and similarly 50% will be below -0.047 while 25% will be below -0.306, bearing in mind that all these measurements are in terms of standard deviations. The complete list of percentiles is given in Table 1 below.

DREP/Percentiles	-0.3	0.1	0.4	0.75	1	1.4	1.9	2.1	2.3	2.65	3	3.25	3.8
25th Percentile	-0.517	-0.306	-0.010	0.070	-0.083	0.182	0.268	0.040	0.435	1.549	1.370	0.900	1.577
50th Percentile	-0.352	-0.047	0.498	0.675	0.591	0.548	0.997	0.476	1.164	2.574	1.962	2.423	2.849
85th Percentile	0.061	0.604	1.584	2.220	2.372	2.018	2.047	2.736	3.190	4.490	4.782	4.087	4.407

Table 1: Percentiles for the GEP realizations of the daily Doppler-Radar measurements (StDev)

3.3 Distributions of the Maximal Values of GEP Realizations

Of particular interest are *rare weather events* such as continuous, multi-day, rain storms, and even hurricanes as they directly affect areas susceptible to flooding and impact populated areas. Hence, it is necessary to understand the distribution of maximal precipitation estimates. These distributions can be highly sensitive to a particular string of daily Doppler-Radar measurements, e.g. one day of intense rain amidst a drought, and as such an effective method is to build these distributions for each specific string of daily measurements in the following manner.

First, applying Step 5 from the *HBWS algorithm* we generate N strings of probable outcomes of precipitation estimates on the ground. Second, from each $\mathcal{C}^{(j)}$, we collect the maximal value. Finally, using a similar procedure as in Steps 2 and 3 of the *HBWS algorithm* we build the desired distribution for the entire collection of these maximal values.

For $N = 1000$ and 10000, the distributions are depicted visually in Figures 8 and 9, respectively.

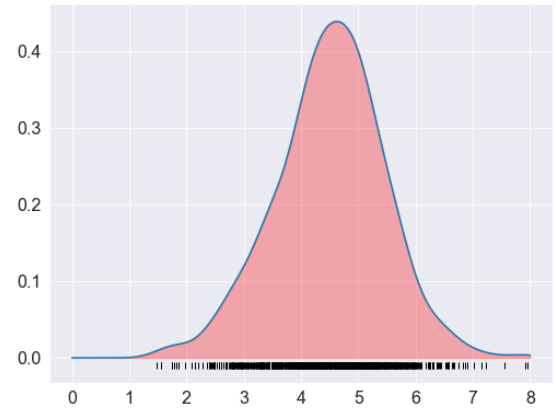

Figure 8: Distribution for N=1000 estimated maximal GEP realizations in standard deviations

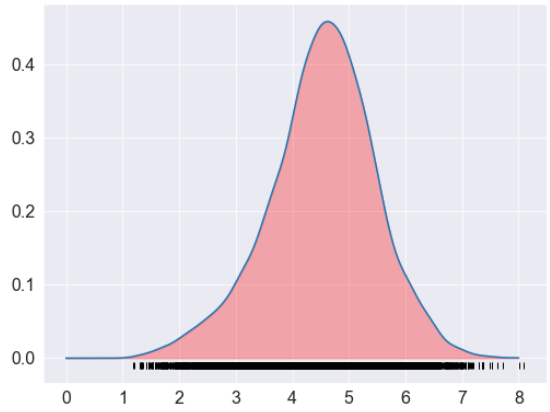

Figure 9: Distribution for N=10000 estimated maximal GEP realizations in standard deviations

3.4 Application: Runoff Model

Utilizing the generated *strings* of GEP realizations in Step 5 of the *HBWS algorithm*, example applications include determining the approximate amount of direct runoff from a rainfall event in a particular area. Runoff is of particular interest in urban populations or areas at risk for flooding. For a more elaborate discussion see [17].

To calculate the runoff amount, denoted by Q, we first convert each of the generated GEP realizations from standard deviations to inches via the equation $x_i = (sz_i + \bar{x})/25.4$, where the mean $\bar{x} = 23.88$ and sample standard deviation $s = 28.19$. For this application, we calculate the runoff via the well-known formula ,

$$Q = \frac{(P - 0.2S)^2}{(P + 0.8S)} \tag{2}$$

where P is the rainfall in inches, $S = \frac{1000}{CN} - 10$ is the potential maximum soil retention after runoff begins, and CN is the runoff curve number. In this simulation, we have chosen a $CN = 60$ as it represents a more average value for natural soil conditions. We refer the reader to [8], [18] for a more extensive discussion.

Figures 10 and 11 depict possible runoff estimates calculated from $N = 1$ and $N = 50$ generated *string(s)* of the GEP realizations, respectively. For a more elaborate discussion on the evolution of the RCN see [27].

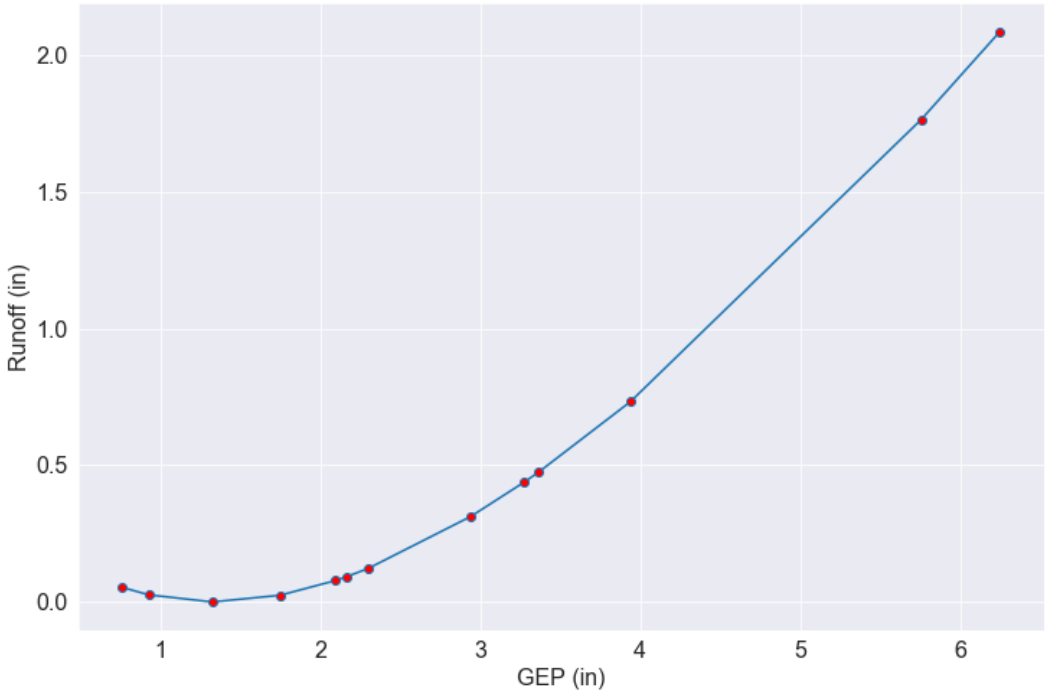

Figure 10: Runoff vs one *string* of GEP realizations

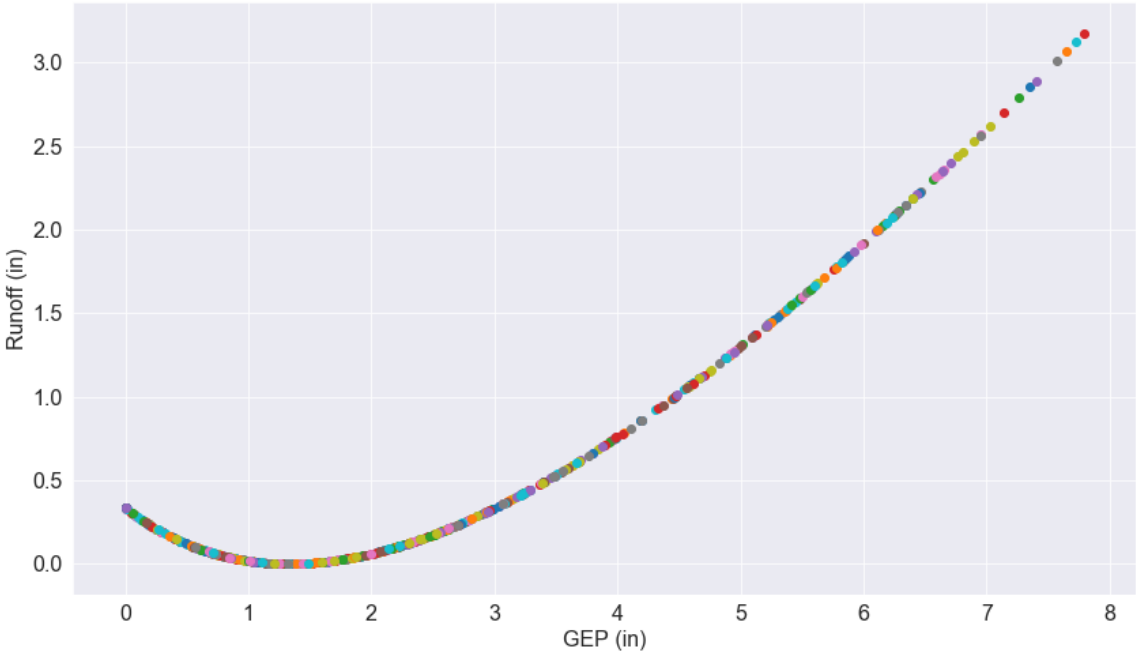

Figure 11: Runoff vs GEP realizations for N=50 generated *strings*

The graph in Figure 12 depicts runoff estimates, in inches, over the thirteen Doppler-Radar daily measurements $\mathcal{I}$. Note that the parabolic nature of the graphs is consistent with the Peak Discharge Calculator's formula for runoff Q and the scale difference between axes.

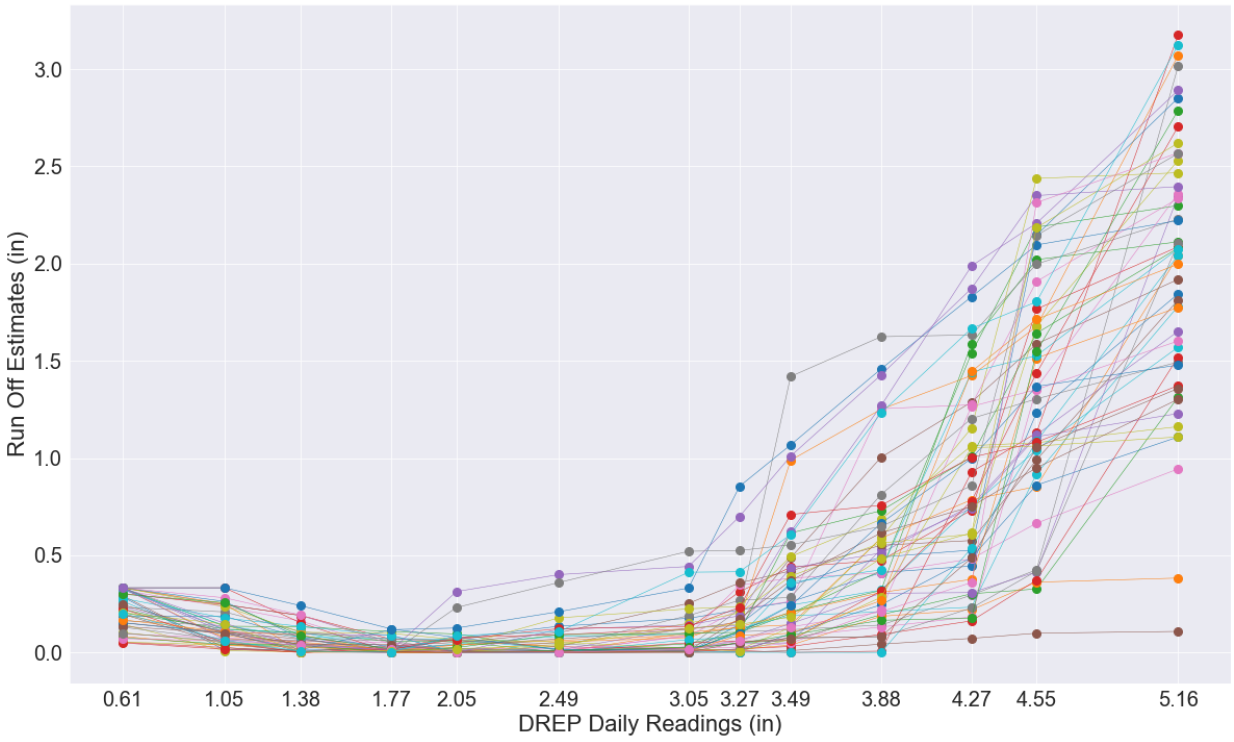

Figure 12: Runoff Estimates for N=50 simulations over 13 Doppler-Radar daily measurements in inches

4 Conclusions and Further Research

There is a high degree of uncertainty associated with Doppler-Radar predictions. This uncertainty cascades into any follow-up application model, such as *rainfall-runoff model, soil-strength analysis, water conservation analysis* etc. For this reason, it is of significant importance to quantify and incorporate this uncertainty in Doppler-Radar predictions into any of these follow-up applications and models.

The goal of this research is to provide a possible methodology which will accomplish precisely this task; that is, it attempts to quantify and most importantly incorporate the uncertainty associated with Doppler-Radar based-analysis into many follow-up applications and models of interest. The general methodology is described in Section 2 and a practical, concrete, simulation is conducted using the current published data of Doppler-Radar estimated precipitation versus Gauge estimated precipitation.

Further considerations such as quantifying and possibly reducing the uncertainty associated with the Doppler-Radar measurements of precipitation that results specifically from the variable $z-R$ relationship, mutual dependency in daily outcome distributions, applications of this methodology to specific severe weather events etc., are areas of future and ongoing research.

Acknowledgments

The authors would like to thank Dr. Prasada Rao for his role in collecting and compiling the data used for simulations in this paper.

References

[1] Austin P.M. Relation between Measured Radar Reflectivity and Surface Rainfall. *Mon. Weather Rev.*, 115(5):1053–1070, 1987.

[2] Baeck M.L , Smith J.A. Rainfall estimation by the WSR-88D for heavy rainfall events. *Weather Forecast.*, 13(2):416–436, 1998.

[3] Barnston A.G., Thomas J.L. . Rainfall measurement acuracy in FACE: a comparison of Gauge and Radar rainfalls. *J. Climmate App. Met.*, 22(12):2038–2052, 1983.

[4] Berne A., Krajewski,W. F. Radar for hydrology: Unfulfilled promise or unrecognized potential? *Adv. Water. Resour.*, 51:357–366, 2013.

[5] Bloor, C. AWRA 2019 Spring Specialty Conference. *Integrated Water Resources Management, Omaha, NE*, 2019.

[6] Brandes E.A., Vivekanandan J., Wilson J. W. A comparison of radar reflectivity estimates of rainfall from collocated radars. *J. Atmospheric Ocean. Technol.*, 16(9):1264–1272, 1999.

[7] Bucaj, V. Association of State Floodplain Managers' Annual Floodplain Management Conference. *Evaluating Complicated Hydrology, Cleveland OH*, 2019.

[8] Chow V., Maidment D., Mays L. *App. Hydrol.* McGraw-Hill Book Co., Singapore, 1988.

[9] Ciach G. J., Krajewski W.F., Villarini G. Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data. *J. Hydrometeorol.*, 8(6):1325–1347, 2007.

[10] Cunha L. K., Smith J. A., Baeck M. L., Krajewski W. F. An early performance evaluation of the NEXRAD dual-polarization radar rainfall estimates for urban flood applications. *Weather Forecast.*, 28(6):1478–1497, 2013.

[11] Dinku T., Anagnostou E. N., Borga M. Improving radar-based estimation of rainfall over complex terrain. *J. App. Met.*, 41(12):1163–1178.

[12] Fulton R. Hydrometerology group;s projects and plans for improving wsr88d rainfall algorithms and products.

[13] Gourley J.J., Maddox R. A., Howard K. W., Burgess D. W. An exploratory multisensor technique for quantitative estimation of stratiform rainfall. *J. Hydrometeorol.*, 3(2):166–180, 2002.

[14] Hromadka T.V., Rao P., Walsh T. Assessment of uncertainty in Doppler-Radar estimated precipitation. *Amer. Inst. Hydrol. Bull.*, 34(2), 2018.

[15] Jayakrishnan R., Srinivasan R., Arnold J. . Comparison of raingage and WSR88D stage III precipitation data over the Texas-Gulf basin. *J. Hydrol.*, 292(14):135–152, 2004.

[16] Klazura G. E., Thomale J. M. , Kelly D. S., Jendrowski P. A comparison of NEXRAD WSR-88D radar estimates of rain accumulation with gauge measurements for high and low reflectivity horizontal gradient precipitation events. *J. Atmospheric Ocean. Technol.*, 16(11):1842–1850, 1999.

[17] Lim K., Engel B., Muthukrishnan S., Harbor J. Effects of initial abstraction and urbanization on estimated runoff using CN technology. *J. Am. Water Resour. As.*, 42(3):629–643, 2007.

[18] LMNO Engineering Research and Software Ltd. Tr-55 peak discharge and runoff calculator.

[19] Morin E., Maddox R. A., Goodrich D. C., Sorooshian S. Radar Z-R relationship for summer monsoon storms in Arizona. *Weather Forecast.*, 20(4):672–679, 2005.

[20] Parzen E. On Estimation of a Probability Density Function and Mode. *Ann. Math. Statist.*, 33(3):1065–1076, 1962.

[21] Quintero F. , Sempere-Torres D. , Berenguer M. , Baltas E. A scenario-incorporating analysis of the propogation of uncertainty to flash flood simulations. *J. Hydrol.*, 460-461:90–102, 2012.

[22] Rico-Ramirez M.A. , Liguori S. , Schellart A.N.A. Quantifying radar-rainfall uncertainties in urban drainage flow modelling. *J. Hydrol.*, 528:17–28, 2015.

[23] Rossa, A., Liechti K. , Zappa M. , Bruen M. , Germann U. , Keil C. , Kraheh P. The COST 731 Action: a review on uncertainty propogation in advanced-meteorological forecast systems. *Atmos. Res.*, (100):150–167, 2011.

[24] Seo B., Dolan B., Krajewski W. F., Rutledge S. A., Petersen W. Comparison of single and dual polarization based rainfall estimates using NEXRAD data for the NASA Iowa flood studies project. *J. Hydrometeorol.*, 16(4):1658–1675, 2015.

[25] Villarini G., Krajewski W.F. Empirically based modelling of radar-rainfall uncertainties for c-band radar at different time-scales. *Quart. J. Roy. Meteorol. Soc.*, 135(643):1424–1438, 2009.

[26] Walsh T., Scioletti M., Rao P., Hromadka T.V., Minvale M. Assessment of uncertainty in Doppler-Radar estimates of precipitation for use in geosciene studies. *The Professional Geologist*, 56(1), 2019.

[27] Williams J. R., Kannan N., Wang X., Santhi C., Arnold J. G. Evolution of the SCS runoff curve number method and its applications to continuous runoff simulation. *J. Hydrol. Eng.*, 17(11), 2012.