Electronic Journal of Statistics
Vol. 7 (2013) 2794–2821
ISSN: 1935-7524
DOI: 10.1214/13-EJS862

Bayesian estimation of Huff curves

Nilabja Guha

Department of Mathematics and Statistics
University of Maryland Baltimore County
Baltimore, MD 21250
e-mail: nguha1@umbc.edu

Ishani Roy

University Corporation for Atmospheric Research
Boulder, CO 80307
e-mail: ishani1.roy@gmail.com

and

Anindya Roy

Department of Mathematics and Statistics
University of Maryland Baltimore County
Baltimore, MD 21250
e-mail: anindya@umbc.edu

Abstract: We propose a nonparametric Bayesian method for estimating regression functions that arise as cumulative distribution functions (cdfs) of a stochastically ordered family of distribution supported on [0,1]. The motivating example is estimation of Huff curves which are depth duration curves for heavy storm rainfall. The Bayesian methodology is compared with the linear programming based estimation method that is currently used by the National Oceanic and Atmospheric Administration (NOAA) for producing the Huff curves. The methodology is illustrated with the rainfall data from the rain gauge stations in California, US. Some limited simulation results are provided to illustrate the finite sample performance of the proposed estimator. We also establish consistency of the proposed method.

AMS 2000 subject classifications: Primary 62P12, 62G08; secondary 90C05.
Keywords and phrases: Storm frequency hyetograph, linear programming, stochastic ordering, Bayesian estimation, Dirichlet process, Bernstein polynomial.

Received September 2012.

1. Introduction

We propose a Bayesian nonparametric methodology to fit a family of monotone regression functions on [0,1]. The regression functions start from zero and reach one and do not intersect in between. Our motivating example comes from

hydrology where families of design-storm hyetographs are estimated simultaneously and the hyetographs form the family of non intersecting monotone regression functions. The most pertinent example is that of Huff curves [10] which are dimensionless curves measuring the depth-duration relationship of storms.

We propose Bayesian estimators of Huff curves that account for smoothness, monotonicity and also ensure that the fitted curves do not intersect at any point other than the end-points of the interval. While a motivating example is that of Huff curves, our methodology is applicable to a broader setup. Compared to smoothness and monotonicity, non-intersecting property is relatively harder to incorporate in a nonparametric estimation problem. Of course one could identify a flexible parametric class of distributions, stochastically ordered according to a partially ordered chain of parameter values and then estimate the best value of the chain from the data. However, in the parametric case, to bring in added flexibility one should consider multi-parameter distributions and finding a partially ordered chain of parameter values that fit the data well and make the corresponding family stochastically ordered can be an intractable problem. There have been considerable interest in estimating stochastically ordered distribution functions mostly in the context of density estimation. In the frequentist setup many methods, mainly based on empirical cdfs, have been proposed for estimation of a family of cdfs under stochastic order constraints; see [18, 8] and the references therein. Inclusion of stochastic order constraints can be achieved in a nonparametric Bayesian setup by specifying priors on classes of stochastically ordered cdfs. Most efforts have concentrated on the direct use of Dirichlet process for specifying priors on distribution functions; see [1, 9, 5] and [8]. In [15], two classes of priors were specified based on Pólya tree representation and Bernstein approximation of cdfs.

We propose a nonparametric Bayesian approach for estimation of Huff curves, viewing them as a family of smooth regression functions with monotonicity and order constraints. The methodology is used to estimate the temporal distribution of heavy storm precipitation in 14 regions in California. The model that we use is similar to the implicit model currently used in NOAA Huff curve estimation procedure. We use a Bernstein-Dirichlet prior in a regression function estimation context. Thus, the curve fitting exercise is fundamentally different from that of estimating distribution function where such priors have been used. We evaluate the performance of the nonparametric Bayesian procedure and compare with the existing linear programming approach. We also establish consistency of the Bayesian estimator.

2. Huff curves

Precipitation frequency atlases are of vital importance for hydrological engineering and management for different geographical regions. Agencies such as National Oceanic and Atmospheric Administration (NOAA) routinely publish such precipitation maps for different areas of the United States. The precipitation figures in such atlases are in turn used as input in hydrological designs by

the U.S. Army Corps of Engineers. One of the major components of hydrological design is watershed model. Key inputs to watershed models include temporal profiles of design storms which provide probabilistic/mechanistic description of typical storm intensity patterns for the geographical region.

There are several methods for generating design storms including the different types recommended by the National Resources Conservation Service (1986). Some other recommendations include [17, 4, 23, 21]. One of the main probabilistic approaches that is used by NOAA for generating temporal storm profiles is due to Huff, described in [10] and later developed and extended in [11]–[14]. Our statistical estimation problem is motivated by the problem of estimation of Huff curves.

Huff curves are graphical representations of empirical distribution of storm characteristics that can be used to depict the natural variability of storm precipitation patterns of a given region. The curves can be used as design-storm hyetographs which in turn are used by hydrologists for assessing drainage patterns of watershed regions [22]. The curves provide probabilistic relationship between dimensionless storm depth and dimensionless storm duration, computed from individual depth-duration relationship in terms of probability isopleths in increments of 10%. The main idea is that the precipitation depth during the entire duration of a storm is non uniform, and this non-uniformity can be used to classify storms and provide better estimates of run-offs based on these temporal characteristics of the storms. The depth-duration relationship is estimated independently for different durations and different iso-probabilities at equal probability increments. The underlying true curves are perceived to be coming from a family of curves, each of which are non-intersecting smooth monotonic curves ranging from 0 to 100. The data comprise the observed values of the curves at specific durations, ranging from 0 to 100% of the storm duration. By rescaling the axes to [0,1], the underlying true curves can then be thought of as cdfs on [0,1] arising out of a stochastically ordered family of distributions. The current software at NOAA that produces smoothed monotonic Huff curve estimates uses the linear programming based approach described by Bonta and Rao [3]. Previous NOAA publications have used moving average based smoothing of the observed depth-duration quantiles.

2.1. Huff curves for California region

Our motivating example is based on the precipitation values obtained from the California rain-gauge system. We will use the precipitation data from California to describe the construction of Huff curves, relate the precipitation analyses terminologies to the quantities in the mathematical model used in the methodology and also illustrate the performance of the proposed method.

California is a vast state with varied climatic conditions. The coastal regions see more rainfall than the semi arid regions in the east of the state. Typical behavior of storm systems developing across the state can vary substantially over different subregions within the state. This fact is borne out in the dras-

tically different Huff curve shapes across the fourteen climatic regions in California. The regions are categorized by NOAA atlas 14 and can be found in NOAA Atlas 14, vol 6: California, p. 225 (www.nws.noaa.gov/oh/hdsc/PF_documents/Atlas14_Volume6.pdf). The Huff curves published by NOAA are typically based on rainfall series with duration types 6 hours, 12 hours, 24 hours and 96 hours available from the precipitation frequency estimate at hdsc.nws.noaa.gov. For example, the 6 hour duration series gives the aggregate rainfall in consecutive 6 hour periods for the entire recording period. The time series used were prepared by standard NOAA protocol. For every precipitation observing station in the project area that recorded precipitation at least once an hour, the three largest precipitation accumulations were selected for each month. Since the Huff curves are related to storm events, a minimum threshold was applied to make sure only heavier precipitation cases were being captured. The precipitation with an average recurrence interval (ARI) of 2 years at each observing station for each duration was used as the minimum threshold at that station. In statistical terms that means that only observations bigger than the median precipitation at that station were selected. A minimum threshold of 25-year ARI (corresponding to the 96th quantile) was also tested. It was found to produce results similar to using a 2-year ARI minimum threshold. The 25-year ARI threshold was not used because it reduced the sample size making the estimates potentially unstable. Each storm event length was then standardized to the type of duration used. For example, for the 6 hours series, the storm lengths were converted to a six hour event and fractional storm durations were calculated as a percentage of the total storm time.

For our mathematical model we rescale the total storm duration to be one and denote the fractional storm times by x. Typically the x's will be equispaced points on the unit interval, [0,1], representing different fractional times at which the storm depths are evaluated. Each storm accumulations was also converted to a total accumulation of 100% which for our mathematical model we rescale to a total of one. The storm depth at a storm duration $x \in [0,1]$ is the proportion of the total precipitation reached by the storm by time x. Let $D(x)$ denote the storm depth at duration x. The depths are random quantities and the Huff curves are estimates of the deciles (10%–90%) of the distribution of D at different fractions x. Thus, a value on the 9th decile curve at a duration fraction x would mean that at least 90% of storms reach a depth equal to that value by the fractional time x of the total storm duration. More information about the data can be found in NOAA atlas 14, [16]. The data are converted into the observed deciles of total number of storm events for different duration fractions. Let $y_k(x_{i,k})$ denote the observed value of the kth decile, $k = 1, \ldots, 9$, observed at a duration $x_{i,k}$ where $i = 1, \ldots, n_k$, and n_k typically depends on the type of duration of the series (e.g. 6 hours, 12 hours). For the Huff curve examples, the observed fractional durations $x_{i,k}$ are same for each k, but we use the double suffix to allow for a more general observation design. The expected values of $y(x)$ are the unknown quantiles $F(q,x)$ of $D(x)$ that are to be estimated. That is, $P(D(x) \leq F(q,x)) = q$. The curves are estimated for $q \in \{0.1, 0.2, \ldots, 0.9\}$ and they are indexed as $F_k(x)$, $k = 1, \ldots, 9$.. The natural restrictions on the

unknown $F_k(x)$ are that $F_k(0) = 0, F_k(1) = 1$, and $F_k(x) > F_l(x')$ is either $1 \leq k < l \leq 9$, or $0 < x' < x < 1$.

The current methodology for estimating Huff curves that is used by NOAA is based on a linear programming algorithm described in [3]. The linear programming approach provides least squares polynomial fit to the observed points $y(x_{i,k})$ with the constraints that the fit at the observed points maintain the required restrictions. Because there are no restrictions on the curves that are put on points in between the observed values, the fit obtained from linear programming can potentially violate both types of restriction. Indeed, [3] commented on this feature, particularly for the behavior of the fit for very low or very high depths. The violation would require ad hoc adjustments to the fit when the values are used as input in design storms for hydrologic calculations.

The observed points generally maintain the restriction that the points along each isopleths are monotonic and at each time point the observations along the isopleths are ordered according to the decile values of their isopleths. Thus, a simple linear interpolator would maintain the natural restriction of monotonicity and non-intersecting property of this family of isopleths. However, the underlying isoproability curves, $(F_k(x))$, are generally perceived to be smooth and hence some degree of smoothing is required for estimation of the Huff curves. As an example of how the fit violates the restrictions, we look at the the Huff curves obtained using the linear programming algorithm to the set of first quartile storms (based on 24 hour duration series) for California precipitation regions 10 in Figure 1, top panel. The fit are obtained using polynomials of degree 12. The zoomed in versions of the lower-left corner of the plot clearly shows that the curves intersect. The upper corner violations are for extremely high storm depths and hence unlikely to cause any practical problems from the use of the Huff curves in hydrologic calculations. However, the violation in the lower corner are in around 2% storm depth and may pose problems if used unadjusted. Moreover, for other project areas where rain gauge system is more sparse the problem of intersecting curves may appear in moderate duration as well.

The primary goal of the present article is to provide a non-intersecting, smooth, monotonic fit to the empirical storm depth values. While there are many frequentist methods for estimating curves under monotonicity restriction, there are no existing frequentist procedure for estimating the family of curves under the restrictions mentioned above. Sieve maximum likelihood methods are possible to formulate for this particular problem, but we find it easier to follow a Bayesian scheme where restrictions can be naturally imposed via prior specification.

3. Bayesian estimation

In the formulation that is currently being used for estimation of the Huff curves, a family of regression models is assumed using the probability isopleths as the mean functions. Since the underlying depths and durations are normalized to 0–100% (and hence can be renormalized to 0–1 scale), the mean functions can

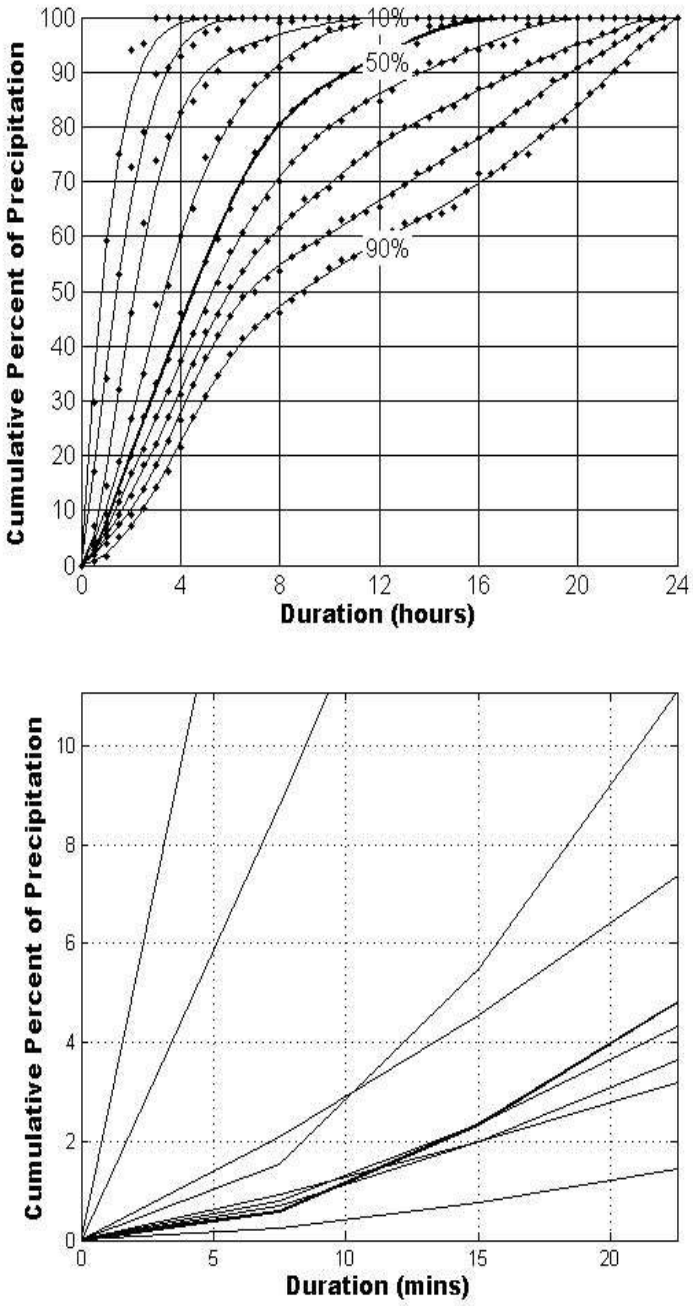

FIG 1. *Estimated Huff curves for California region 10 for 24 hour duration. The top plot shows the estimated Huff curves based on first quartile storms. The bottom row plot shows zoomed in versions of the lower corner of the top plot. The magnified plot shows that the estimated curves corresponding to different rain depth quantiles intersect.*

be thought of as a family of nine stochastically ordered cdfs. The design points in this formulation are all equi-spaced dimensionless storm durations that are same for each curve. While the observations used in the Huff curve calculations are monotonic and non-intersecting, for the modeling purposes we entertain the general scenario where the observations themselves (prior to any smoothing or isotonization) may not satisfy the monotonicity or the ordering restrictions and the design points may differ for the different curves.

3.1. Model

Consider K monotone curves namely $F_1, F_2, \ldots, F_K$ from $[0,1]$ to $[0,1]$ with $F_i(0) = 0$ and $F_i(1) = 1$ for each i. Suppose the curves satisfy the ordering $F_1 < F_2 < \cdots < F_K$, i.e, $F_i(x) < F_j(x)$ for all $x \in (0,1)$ whenever $i < j$. We build our methodology using regression setup but allow the set of design points to be different for different curves. The values that are observed need not satisfy the monotonicity property and the ordering of the underlying curves. Our goal is to propose monotone, smooth and consistent estimates for the curves that maintain ordering.

Suppose, for the kth curve we have n_k observations at the design points $x_{i,k}$'s. The data generating model is

$$y_{ik} \equiv y_k(x_{i,k}) = F_k(x_{i,k}) + \epsilon_{i,k}, \tag{3.1}$$

where the error $\epsilon_{i,k}$'s are assumed to be independent of each other with $E(\epsilon_{i,k}) = 0$. This regression model is implicitly assumed in the linear programming formulation of the Huff curve estimation problem. Under distributional assumption of the error, the model provides a pseudo likelihood for the underlying family of curves and can be conveniently used to formulate likelihood based estimation schemes. The estimand is the family of curves $\boldsymbol{F} = (F_1, \ldots, F_K)$. The error variances are nuisance parameters. Though we develop the methodology assuming independence of $\epsilon_{i,k}$'s it can be shown (Remark 2) that the proposed estimator remain consistent even under weak correlation between the observations.

To incorporate the monotonicity and order restriction, we approximate the curves by Bernstein polynomials and impose restrictions on the coefficients. Let

$$B^K_{m,\uparrow} = \left\{ b^k_m = \sum_{i=1}^{m} p_{i,m}(x) G_{k,m}(\frac{i}{m}), k = 1, \ldots, K \right\}, \tag{3.2}$$

where $G_{k,m}$'s are cdfs on $[0,1]$ with support on $\frac{i}{m}$'s with $1 \le i \le m$ and with the restriction $0 < G_{1,m}(x) < G_{2,m}(x) < \cdots < G_{K,m}(x) < 1$ for $x \in (0,1)$ and $p_{i,m}(x) = \binom{m}{i} x^i (1-x)^{(m-i)}$. Monotonicity and the ordering of the components of $B^K_{m,\uparrow}$ follow from the ordering of the $G_{k,m}$'s. Let $\boldsymbol{F}^* = (F^*_1, \ldots, F^*_K)$ be the true value of the functional parameter $\boldsymbol{F}$. We want to estimate $\boldsymbol{F}^*$ by $\boldsymbol{F}^*_m \in B^K_{m,\uparrow}$. As Bernstein approximation converges uniformly, for any $\boldsymbol{F}^*$ there will be sequences in $B^K_{m,\uparrow}$ that converge uniformly to $\boldsymbol{F}^*$. In this sense elements of $B^K_{m,\uparrow}$ can be used as approximate target parameters.

Following the implicit assumption in LP estimation [3], we use the working model $\epsilon_{i,k} \sim N(0, \sigma^2_{i,k})$ with the understanding that $0 \leq y_{i,k} \leq 1$. At point x, for the k th curve the variance function is denoted by $\sigma^2_{x,k} = \sigma^2_k(x)$. Let $\mathbf{\Delta}_K = \{G_{(k,m)}(\frac{i}{m})\}_{1\leq i\leq m, 1\leq k\leq K}$. Let $\underline{Y}_k$ be the n_k dimensional vector of the observed data points for the k th curve and $\mathbf{Y} = \{y_{i_k k}\}_{i_k=1,\ldots,n_k; k=1,\ldots,K}$. The likelihood under this model is given by

$$L(\mathbf{\Delta_K}) \propto \prod\nolimits_k \frac{1}{\sqrt{det(\Sigma_k)}} \exp^{-\frac{1}{2}(\underline{Y}_k - \mathbf{M_k})' \Sigma_k^{-1} (\underline{Y}_k - \mathbf{M_k})}, \tag{3.3}$$

where $\mathbf{M_k}_i = \sum_{l=1}^m p_{l,m}(x_{i,k}) G_{(k,m)}(\frac{l}{m})$. Here Σ_k is the covariance matrix for the errors on the k th curve. We have $\Sigma_{k,i,i} = \sigma^2_{i,k} = \sigma^2_k(x_{i,k})$ and $\Sigma_{k,i,j} = 0$, for $i \neq j$. Let $\mathbf{N} = (n_1, \ldots, n_K)$.

The normality in the working model for the $\epsilon_{i,k}$'s gives a natural least square interpretation for the fit. As discussed in Remark 3 in the appendix, the proposed Bayesian estimator continues to be a reasonable estimator even under mild model misspecification.

3.2. Prior specification

Let, $\delta_{k,i,m} = G_{k,m}(\frac{i}{m}) - G_{k,m}(\frac{i-1}{m})$. For each $G_{k,m}$'s, we assume a Dirichlet distribution on the increments $\delta_{k,i,m}$'s with the induced ordering mentioned earlier. Thus, the joint density of $G = (G_{1,m}, G_{2,m}, \ldots, G_{k,m})$ is given by,

$$\Pi_{G,m}(\mathbf{\Delta}_K) \propto \prod_{k=1}^{K} \prod_{i=1}^{m} \left(G_{k,m}\left(\frac{i}{m}\right) - G_{k,m}\left(\frac{i-1}{m}\right)\right)^{\frac{1}{m}-1} \mathbf{1}_{(G_{k,m}(\frac{i}{m}) < G_{l,m}(\frac{i}{m}), k<l)}, \tag{3.4}$$

where $G_{k,m}(0) = 0$ and $G_{k,m}(1) = 1$.

We make a few observations about the limiting nature of the sequence of priors as specified above. If $K = 1$, then as $m \to \infty$ this prior converges weakly to the Dirichlet process Π, that is DP$(1,\lambda[0,1])$, where λ is the Lebesgue measure on $[0,1]$. Let $P_1, P_2, \ldots, P_k$ be K random independent cdfs from that DP. Let R_m be the set of K tuples where the cdfs are stochastically ordered at $\frac{i}{m}$'s and let Π^K be the K fold distribution induced by K independent DP. Provided that $\lim \Pi^K(R_m)$ exists and greater than zero, we can have a well defined weak limit of $P_{\Pi_{G,m}}$, the probability measure corresponding to $\Pi_{G,m}$. Since it is not the focus of the present exercise, we do not investigate the existence and the positivity of the limit any further.

3.3. Posterior computation

We can write the posterior density as follows

$$\Pi_{\mathbf{N}}(\mathbf{\Delta}_K | \mathbf{Y}) \propto \prod\nolimits_k \frac{1}{\sqrt{det(\Sigma_k)}} \exp^{-\frac{1}{2}(\underline{Y}_k - \mathbf{M_k})' \Sigma_k^{-1} (\underline{Y}_k - \mathbf{M_k})} \times \Pi_{G,m}. \tag{3.5}$$

Gibbs sampling and Metropolis-Hastings(MH) algorithm can be used for the posterior sampling of $G_{k,m}(\frac{i}{m})$'s. We can generate posterior sample of $G_{k,m}(\frac{i}{m})$ given the others. Posterior mean can be used as an estimate of the $G_{k,m}(\frac{i}{m})$'s. The variance functions $\sigma^2_{x,k}$'s are chosen and the choice is discussed later. The steps can be written as follows

1. Start with initial values with, $G_{l,m}(\frac{i}{m}) < G_{j,m}(\frac{i}{m})$, for $l < j$, i.e the initial ordered curves.
2. Given k' th curves ($k' \neq k$) and $\{G_{k,m}(\frac{j}{m}),\ j \neq i\}$, we sample from the posterior distribution of $G_{k,m}(\frac{i}{m})$ by MH.
3. We continue until posterior chains of $G_{k,m}(\frac{i}{m})$'s converge.

Let $\mathbf{F}^1, \mathbf{F}^2, \ldots, \mathbf{F}^L$ be L samples from the posterior distribution of the K curve tuples. Then $\hat{\mathbf{F}} = L^{-1} \sum_{l=1}^{L} \mathbf{F}^l$ is the proposed estimator of the k curves. More details about the posterior sampling step are given in the appendix.

3.4. Posterior consistency

A desirable property for a Bayesian methodology is posterior consistency. Since the seminal work by Schwartz [19] there has been many important development in the theory of posterior consistency. Specifically researchers have developed techniques to show consistency under very general conditions on the statistical experiments; see [6, 2, 7] and the references therein.

In our case, we use the compactness of the range and the domain of the regression functions and show consistency using direct calculations based on Schwartz's approach. To prove consistency we need to make the following assumptions about the design points $x_{i,k}$ and the error distributions. Let $Q_{x,k}$ be the error distribution function at point x for the kth curve. Suppose, under the setup of (3.1) the following assumptions hold.

- A1: For a random design (RD) let H_k, be the distribution of the observation points for the kth curve. Let λ denote the Lebesgue measure on [0,1]. Assume that for each k, H_k is absolutely continuous with λ, i.e., $H_k \ll \lambda$. For a fixed design(FD), let $H_{k,n}$ be the empirical distribution of the observation points for kth curve. Assume that for each k, $H_{k,n}$ converges to H_k in distribution, and $H_k \ll \lambda$.
- A2: For any compact subset S_c of (0,1) and for $x_1, x_2 \in S_c$, given $\delta > 0$, there exists $\epsilon > 0$ s.t $|Q^{-1}_{x_1,k}(Q_{x_2,k}(z)) - z| \leq \delta$ for all k and z, whenever $|x_1 - x_2| \leq \epsilon$.
- A3: For each k, let $\sigma_{x,k}$ be strictly positive continuous functions of $x \in (0,1)$, and let $\int_0^1 \frac{1}{\sigma^2_{x,k}} \mathrm{d}x < \infty$.

In A1 the design points are assumed to become dense over the unit interval at a uniform rate. In A2 we assume smoothness of $Q_{x,k}$ by assuming that $Q_{x_1,k}$ and $Q_{x_2,k}$ are close if x_1 and x_2 are close. Since the variance functions are used as direct input, we make the assumption A3 regarding the choice of variance function.

Next, we define the appropriate topology under which consistency will be investigated. Given a particular variance function $\sigma_{x,k}$, the space of the mth degree Bernstein polynomials can be given the inner product defined by, $I_{\mathbf{F},\mathbf{G}} = \langle \mathbf{F}, \mathbf{G} \rangle = \sum_{k=1}^{K} \int \frac{F_k(x)G_k(x)}{\sigma_{x,k}^2} h_k(x)\mathrm{d}x$, for $\mathbf{F}, \mathbf{G}$ in $B_{m,\uparrow}^K$. Here $\mathbf{F} = (F_1, \ldots, F_K)$, $\mathbf{G} = (G_1, \ldots, G_K)$ and $h = (h_1, \ldots, h_K)$ are the sampling densities for the design points associated with the K curves. The inner product induces a metric defined by $d_{\sigma,h,K}^2(\mathbf{F}, \mathbf{G}) = \langle \mathbf{F} - \mathbf{G}, \mathbf{F} - \mathbf{G} \rangle$. From assumptions A1 and A3, the topology with respect to the metric $d_{\sigma,h,K}$ is same as the topology corresponding to the metric defined by the inner product $I_{\mathbf{F},\mathbf{G}}^* = \langle \mathbf{F}, \mathbf{G} \rangle_* = \sum_{k=1}^{K} \int F_k(x)G_k(x)\mathrm{d}x$. For notational convenience we will suppress the dependence of the metric on the variance function σ and the sampling densities h. For consistency we further need to assume some growth relationship between the degree of Bernstein approximation and the number of observations for each curve. First of all we need the degree of the Bernstein approximations to go to infinity. We can choose m large enough such that the class of proposed curves intersects the ϵ neighborhood of the true curve tuples (under metric d_K) and then increase the number of observations (depending on m) such that the posterior probability of that ϵ neighborhood approaches one. It is interesting to note that increasing number of observations and m at some appropriate rate (see A5) can ensure convergence. The issue is discussed in details in the appendix. For simplicity we make a slightly stronger assumption and assume that the sample size for each curve going to infinity at the same rate across the K curves. Such an assumption maybe helpful to establish rate of convergence of the posterior of the K curve tuple. Let $N = \max_{1 \le k \le K} n_k$. We make the following assumptions.

- A4: For each k and $\mathbf{N}$, we have $w_{k,N} = \frac{n_k}{N} > k_0 > 0$ and $N \to \infty$.
- A5: Let $m \to \infty$ and $\frac{N}{m} \to \infty$.

Let $\boldsymbol{F}^*$ be the true value of the K curve tuple and U_ϵ be an open neighborhood around $\boldsymbol{F}^*$ under d_K. Heuristically the probability that $G_{k,m}$'s are stochastically ordered at any knot point is $(K!)^{-1}$ and thus, we would expect $P_{\Pi_{G,m}}(B_{m,\uparrow}^K \cap U_{\frac{\epsilon k_0}{2}}) > O((K!)^{-m})$. If this result holds then A5 gives a sufficient condition on the rate of N and m for the conclusion of Theorem 1 to hold. The details of this result are given in the appendix.

The following consistency theorem can now be stated.

Theorem 1. *Consider model (3.1). If assumptions A1–A5 hold, then as $m, N \to \infty$, the posterior probability $P_{\mathbf{N}}(U_\epsilon \cap B_{m,\uparrow}^K | \mathbf{Y}) \to 1$ almost surely.*

Remark 1. *For fixed n_k and m, if $\sigma_{x,k} = \sigma v_{x,k}$ and σ approaches zero, then the Bayes estimator will approach the weighted least square solution in $B_{m,\uparrow}^K$, weighted by $v_{x_{i,k}}^{-1}$'s. For large values of m and $\min_{1 \le k \le K} n_k$, the weighted least squares solution, will be close to $\mathbf{F}^{perp_m}$, the closest point of $\mathbf{F}^*$ in $B_{m,\uparrow}^K$ in the metric d_K.*

TABLE 1
Square root of IMSE's for the Bayesian and the LP methods

	Polynomial degree								
	6			12			15		
	Bayes Estimator								
$\sigma_{x,k}$	N=20	N=30	N=60	N=20	N=30	N=60	N=20	N=30	N=60
0.10	0.028	0.026	0.023	0.034	0.027	0.019	0.037	0.030	0.021
0.05	0.025	0.023	0.021	0.026	0.021	0.016	0.029	0.023	0.018
0.02	0.024	0.022	0.020	0.024	0.020	0.016	0.026	0.022	0.016
	Linear Programming Estimator								
	0.052	0.036	0.022	0.062	0.046	0.021	0.095	0.051	0.020

4. Numerical example

Before presenting the Bayesian estimates of Huff curves for the California region, we report the findings of a limited simulation study done to evaluate the finite sample performance of the proposed estimator and to contrast the performance with that of the traditional linear programming based estimator. An example with five stochastically ordered cdfs is considered. We vary the Bernstein degree and also the choice of the variance function. The mean functions are given by

$$\begin{aligned} F_1(x) &= B(x;7,1), \\ F_2(x) &= B(x;5,1), \\ F_3(x) &= 0.5B(x;5,1) + 0.5B(x;1,5), \\ F_4(x) &= B(x;1,5), \\ F_5(x) &= B(x;1,7), \end{aligned}$$

where $B(x,a,b)$ denotes the value of the cdf of the $Beta(a,b)$ distribution at x. The observations are generated as $y_{x,k} \sim Beta(cF(x), c(1-F(x))$ with $c = 30$. However, for implementation of the Bayesian scheme we use the approximate normal likelihood described in the methodology section. Generating the observed data points for $N = 20, 30, 60$ equispaced grid points for each of these curves, we use the Bayesian methodology to estimate the ordered curve tuple, varying the degree and the variance function. The variance function is chosen to be constant $\sigma_{x,k} = \sigma$ for all x and k. The value of the constant is chosen from $\{0.02, 0.05, 0.1\}$. The Bernstein degree is varied over the set $m = 6, 12$ and 15. We compare the square root of the integrated mean squared error (IMSE) for different methods based on 100 Monte Carlo replications. For each Monte Carlo sample, the we draw 3000 MCMC posterior samples and use the final 1000 samples to compute the Bayes estimators. In all simulations the chain length seemed to provide adequate convergence. For each sample we also compute the traditional LP estimator for polynomial degree equal to $m = 6, 12$ and 15 and compute the IMSE for the LP estimator as well. The square root IMSE's for both estimators are presented in Table 1.

The IMSE's generally decrease with sample size, which is consistent with the theoretical results. For small number of observed data points, cases with higher degree have bigger IMSE, as the fit will be more prone to model errors. However,

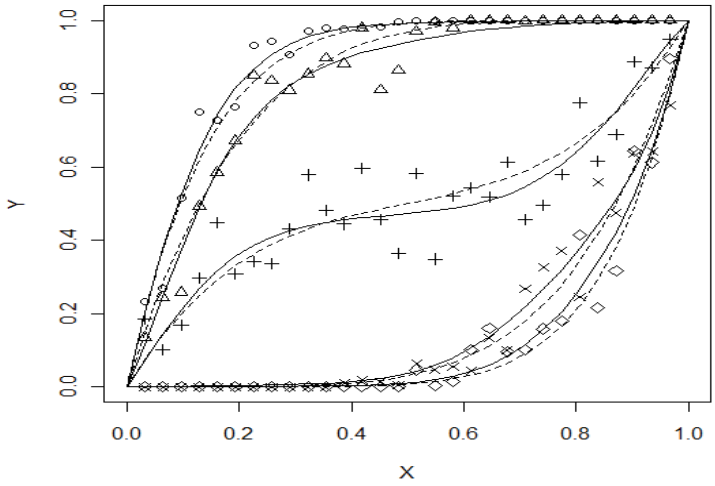

FIG 2. *One typical fit for the simulation example. The data points are given by* $\diamond$, $\times$, $+$, $\triangle$ *and* $\circ$ *for* F_1, F_2, F_3, F_4 *and* F_5 *respectively. Fitted and the true curves are plotted by solid and dashed lines, respectively. Here* $N = 30$,$m = 12$ *and* $\sigma_{x,k} = .05$.

as the number of observations increases, the IMSE decreases as expected. With small $\sigma_{x,k}$ we have smaller IMSE (see Remark 1). An example of a Bayesian fit is given in Figure 2. The IMSE for the Bayesian estimator is generally smaller, particularly when the degree is high compared to sample size. This is because the LP estimator is overfitting the data at the observed points.

5. Temporal analysis for California heavy storm precipitation

The current linear programming based estimation method for the Huff curves is essentially a least squares procedure, which maintains the two basic constraints, monotonicity and non-intersecting property, across the curves at the grid points used in the optimization procedure. The estimation accuracy is not only a function of the polynomial degree but also depends on the shape of the curves. While higher degree polynomials provide better fit (provided the algorithm converges), they generally also tend allow for greater number of violations of the constraints. While the polynomials maintain the constraints at the optimized grid, the curves can be non-monotonic or may intersect each other in between the grid points. We used a standard grid of 96 equally spaced points in the range 0 to 100 to evaluate the estimated curves and check for violations of constraints in between the grid points. To check for violations of monotonicity, we looked at pairwise differences of consecutive grid points along each curve and tabulated the number of violations as a percentage of total number of comparisons. Similarly, we

TABLE 2
Percentage violation of constraints

		degree = 6				degree = 12			
		duration (hours)				duration (hours)			
		monotonicity constraint							
region	quartile	6	12	24	96	6	12	24	96
10	q1	1	2	1	1	2	3	2	1
	q2	0	0	0	1	1	1	1	1
	q3	0	1	0	1	0	1	0	1
	q4	0	0	0	0	1	0	1	0
	all	0	0	0	0	0	1	1	1
14	q1	3	4	1	1	3	5	3	0
	q2	0	0	0	1	3	1	1	2
	q3	0	1	1	1	0	1	1	0
	q4	0	0	1	0	0	1	1	0
	all	1	1	1	1	2	3	2	1
		stochastic ordering constraint							
10	q1	2	1	1	0	1	1	1	0
	q2	1	1	1	0	1	1	0	0
	q3	1	1	1	0	1	1	0	0
	q4	1	1	0	0	1	1	0	0
	all	0	0	0	0	0	0	0	0
14	q1	2	1	0	0	2	2	0	0
	q2	0	1	1	1	1	1	1	1
	q3	1	1	1	0	1	0	0	0
	q4	1	1	0	0	1	1	1	0
	all	1	1	1	1	0	1	1	0

also checked for violation of non-intersecting property by checking pairwise differences over consecutive curves along each grid point and tabulated the total number of violations as a percentage of the total number pairs compared.

We used curves for all quartile storms and at all durations for regions 1, 10 and 14 to illustrate the potential pitfalls of the linear programming method. A jth quartile storm, ($j = 1, \ldots, 4$), is one that has the maximum depth during the jth quartile of the duration. The particular choices of the regions reflect the wide variability in precipitation type over the state. Region 1 is a coastal region near Oregon border and receives about 120 inches of rainfall every year. Region 10 is inland and has moderate amount of rainfall while region 14 is in the semi-arid region near the south-east border with Arizona. While there were no violations in region 1, the estimated curves in region 10 and 14 showed several violations. To evaluate the effect of the degree of the polynomial fit, we checked for polynomial degree equal to 6 (as used in NOAA Atlas Vol. 14) and for degree equal to 12. Table 2 provides the percentage violations for both type of violations for regions 10 and 14 for polynomial degree equal to 6 and 12. One can see that the percentage violations can be substantial. The percentage of violations is generally higher for the higher degree fit. The highest value reported is 5%, for a 12 degree fit to first quartile 12 hour storms in region 14. Apart from the fact that there are several instances of violations, it is also interesting to see that the pattern of violation changes spatially. The coastal region storms tend to be more uniform while those in the semi-arid regions have grater variability. These facts adversely affect the estimation accuracy in different ways across the

TABLE 3
Integrated L_2 error of linear programming method with polynomial degree equal to 6 (LP6) and 12 (LP12) and the Bayesian method (Bayes). Dur denotes the duration type of the rainfall time series and Q denote the quartile type of the storms. Each entry is multiplied by 100

		Region								
		1			10			14		
Dur	Q	LP6	LP12	Bayes	LP6	LP12	Bayes	LP6	LP12	Bayes
	q1	0.40	0.22	0.58	1.01	0.77	1.24	1.71	0.93	2.06
	q2	0.43	0.22	0.62	0.91	0.55	1.19	1.49	0.72	1.49
6h	q3	0.39	0.21	0.57	0.60	0.44	0.97	0.55	0.37	0.84
	q4	0.38	0.21	0.76	0.74	0.36	1.11	0.53	0.33	0.82
	all	0.22	0.16	0.53	0.57	0.45	0.81	0.87	0.53	1.33
	q1	0.51	0.33	0.63	1.44	0.80	1.27	3.33	1.10	2.09
	q2	0.64	0.41	0.78	0.68	0.44	0.67	1.05	0.49	0.80
12h	q3	0.42	0.24	0.53	0.96	0.60	0.89	1.36	0.46	0.89
	q4	0.60	0.36	0.67	0.91	0.57	0.99	1.10	0.68	0.99
	all	0.28	0.21	0.40	0.56	0.41	0.65	1.32	0.54	1.43
	q1	0.68	0.55	0.66	1.35	0.75	1.04	3.73	0.99	3.06
	q2	0.59	0.42	0.55	1.71	0.65	1.13	2.57	0.69	1.59
24h	q3	0.68	0.44	0.60	2.51	0.89	1.92	2.89	0.97	1.99
	q4	0.90	0.52	0.74	2.16	0.95	2.06	2.72	1.58	2.56
	all	0.48	0.34	0.47	0.71	0.55	0.64	1.08	0.60	0.85
	q1	1.14	0.80	0.98	1.37	0.93	1.18	6.27	2.09	4.64
	q2	1.48	0.86	1.08	3.97	1.70	2.59	7.52	2.88	5.24
96h	q3	1.69	1.03	1.31	4.72	2.00	3.59	4.72	2.10	3.60
	q4	1.29	0.99	1.09	2.04	1.06	1.71	3.95	1.80	3.05
	all	0.78	0.52	0.62	1.10	0.78	0.87	3.19	0.95	1.61

region, with the linear programming method facing greater challenge when there is high variability in the storm systems. The trade-off for using higher degree is of course better fit versus greater potential for violations of constraints. The reduction in the integrated L_2 distance between the observed and the fit,

$$d(\hat{F}, F) = \sqrt{(NK)^{-1} \sum_{i=1}^{N} \sum_{k=1}^{K} (\hat{F}(x_{i,k}) - y_k(x_{i,k}))^2},$$

obtained from higher degree fit is obvious from the values reported in Table 3. The table also gives the L_2 error for the Bayesian fit using 12 knot points for 12, 24 and 96 hour cases and 6 knots for the 6 hour case. The constant variance function $\sigma_{x,k} = 0.01$ was used. The squared error for the Bayesian generally falls in between the linear programming fit with 6 and 12 degree polynomial. Of course, the Bayesian fit maintains all the required restrictions. From the Figures 3–5, the Bayesian fit are very comparable with the linear programming fit in terms of accuracy while only the Bayesian fit maintain all the required constraints.

6. Selection of the degree of Bernstein basis

There are two unknown quantities that need to be determined during the estimation process. The variance function of the working normal model and the degree

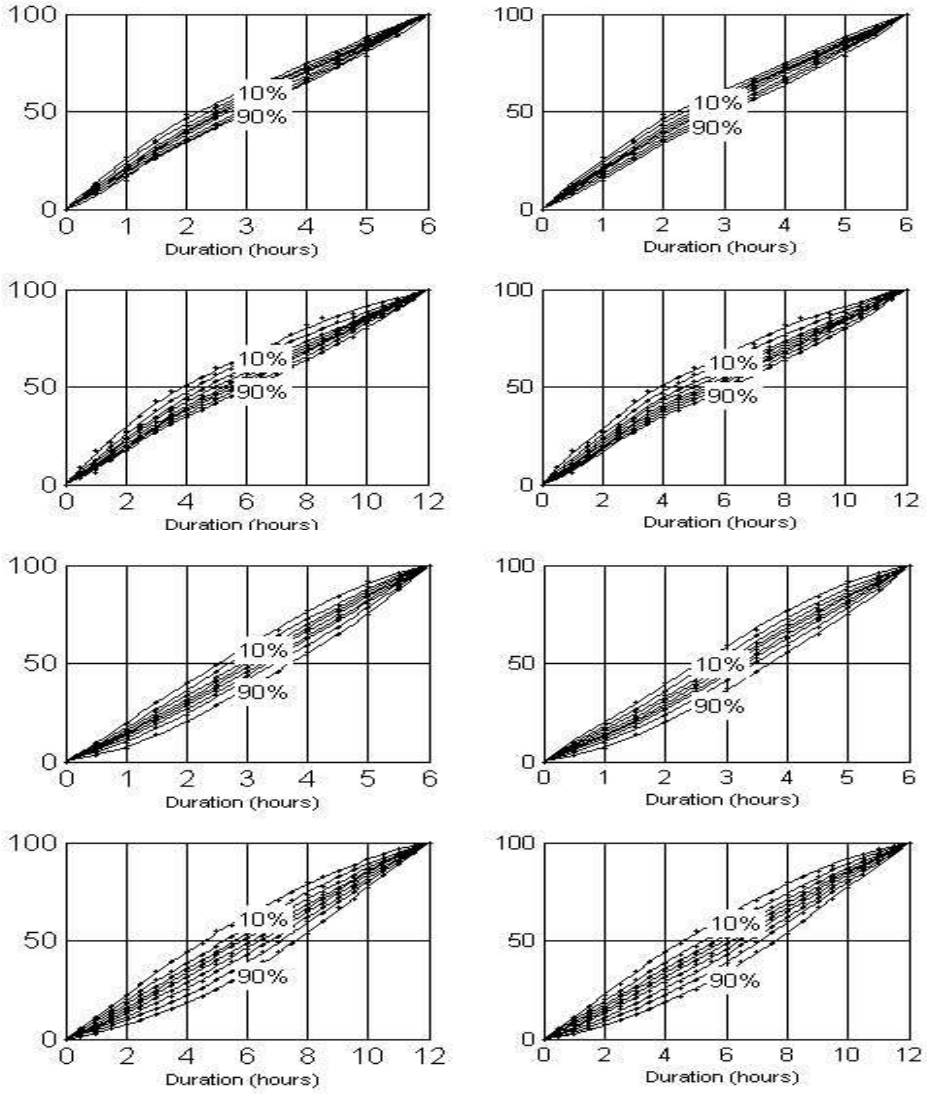

FIG 3. *The estimated huff curves for California region 1. The top two rows are estimated curves for first quartile storms for 6 hour and 12 hour durations, respectively. The bottom two rows are estimated curves for all the storms in the region for 6 hour and 12 hour durations, respectively. The left column shows the estimated curves using the linear programming algorithm. The second colum shows the curves estimated using the Bayesian algorithm. Cumulative percentage precipitation are given along y-axis.*

of Bernstein approximation. The Bayes estimator estimates $\boldsymbol{F}^*$ as a quantity in $B^K_{m,\uparrow}$ by the posterior mean $\mathbf{F}_{(m,\mathbf{N})}$ and the estimation performance depends on the degree m and the variance function $\sigma_{x,k}$'s. The effect of the variance

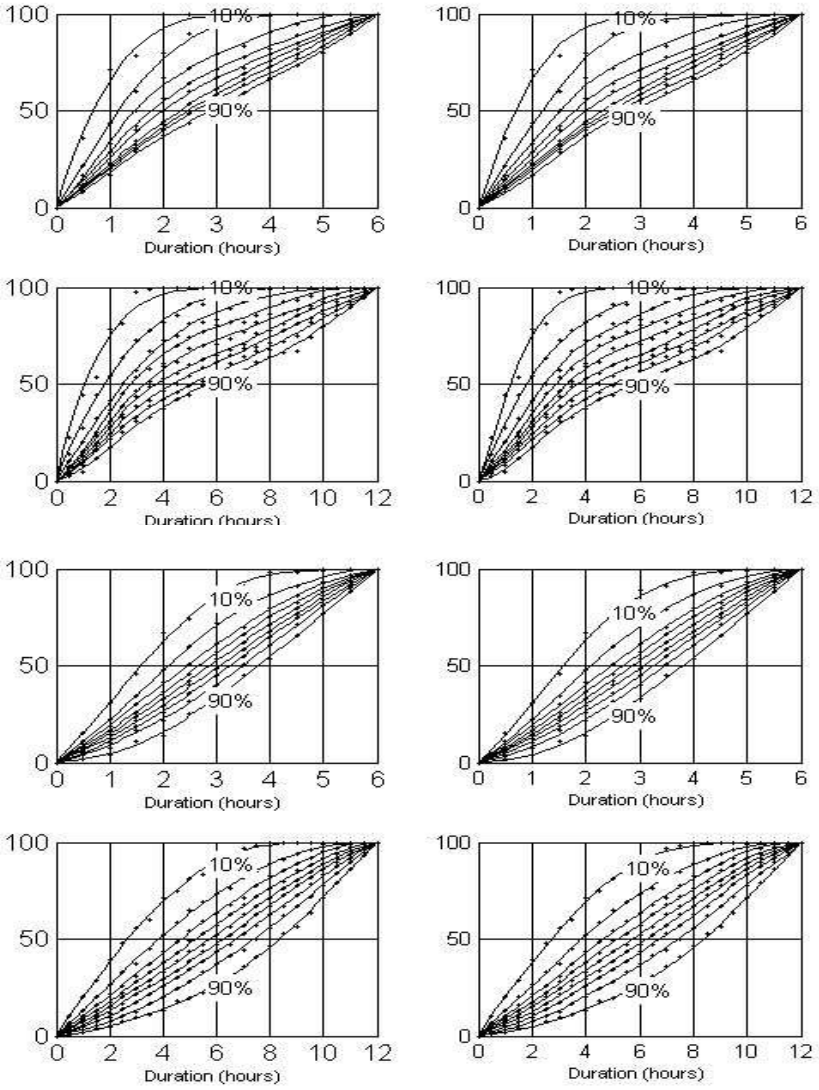

FIG 4. *The estimated huff curves for California region 10. The top two rows are estimated curves for first quartile storms for 6 hour and 12 hour durations, respectively. The bottom two rows are estimated curves for all the storms in the region for 6 hour and 12 hour durations, respectively. The left column shows the estimated curves using the linear programming algorithm. The second colum shows the curves estimated using the Bayesian algorithm. Cumulative percentage precipitation are given along y-axis.*

function is easily understood. When a small constant variance function is used the estimated fit goes to the least square solution in $B^K_{m,\uparrow}$. From the simulation, we see that the performance of the estimator does not depend critically on the

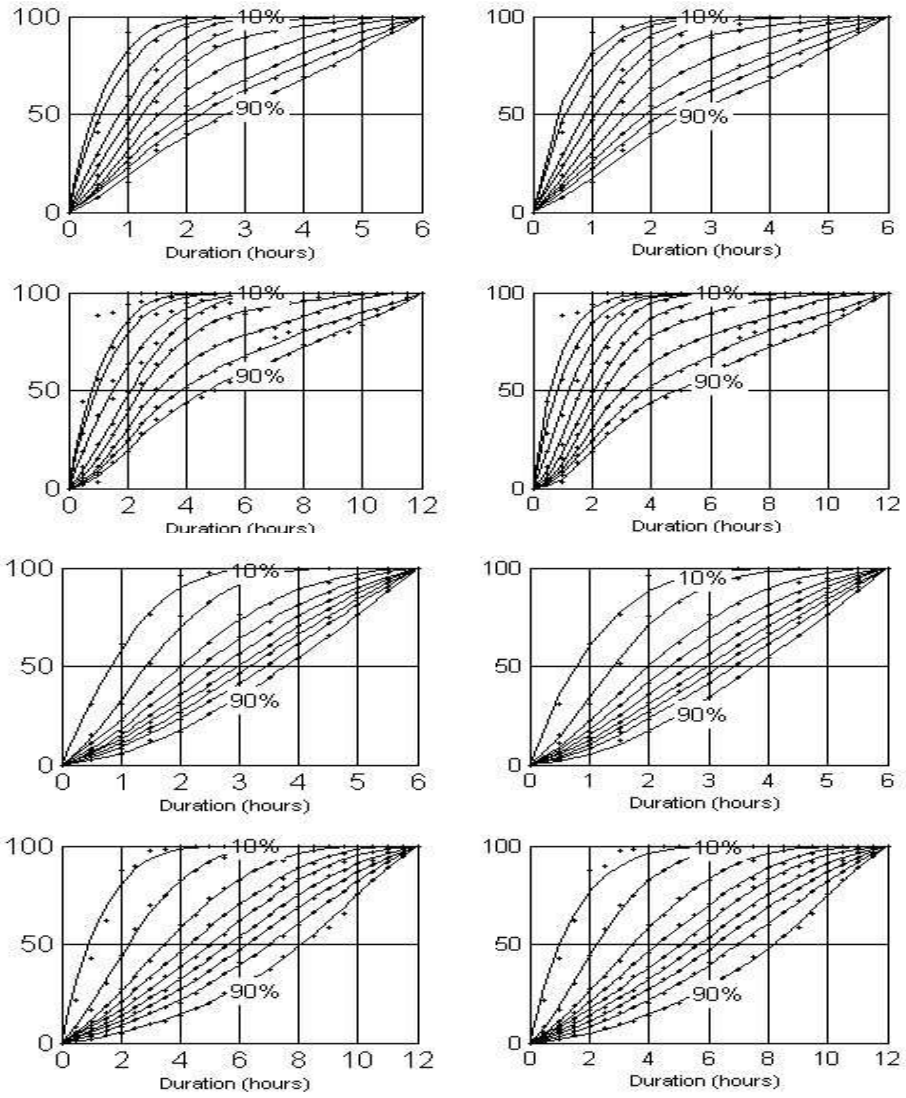

FIG 5. *The estimated huff curves for California region 14. The top two rows are estimated curves for first quartile storms for 6 hour and 12 hour durations, respectively. The bottom two rows are estimated curves for all the storms in the region for 6 hour and 12 hour durations, respectively. The left column shows the estimated curves using the linear programming algorithm. The second column shows the curves estimated using the Bayesian algorithm. Cumulative percentage precipitation are given along y-axis.*

variance function as long as a reasonably small value is chosen. The choice of the Bernstein degree is more involved. If a good fit in the least square sense is required then using more knots may decrease the fit error but the prediction per-

formance at an unobserved data point may not be good. This is similar to what happens in the traditional LP method when a high polynomial degree is chosen. In the Bayesian framework it makes sense to use the unknown degree as a parameter and specify a prior for the degree. However, the computation becomes prohibitive. One could choose the degree using Bayes factor and then proceed with the chosen model. Another option is to use a somewhat calibrated Bayes approach and use cross validation along with the MCMC sampling. Both options are computationally taxing. We define an approximate prediction error estimate based on generalized cross validation (GCV) and use it to select the degree.

The main idea is that a least squares fit without any constraints on the Bernstein coefficient will be still consistent for any true F^* but the least squares fit need not satisfy the monotonicity and stochastic ordering restrictions. Also the fitted least square curve tuples $\boldsymbol{F}_{m,ols} = (F_{1,m,ols}, \ldots, F_{K,m,ols})$ should be close to the posterior mean curve $\mathbf{F}_{(m,\mathbf{N})} = (F_{(1,m,\mathbf{N})}, \ldots, F_{(K,m,\mathbf{N})})$ which maintains the required constraints. We propose the approximate prediction error (APE)

$$\begin{aligned} APE^2(m) = K^{-1} \sum_{k=1}^{K} n_k \sum_{i=1}^{n_k} (n_k - m)^{-2} (y_{i,k} - F_{k,m,ols}(x_{i,k}))^2 \\ + K^{-1} \sum_{k=1}^{K} n_k^{-1} \sum_{i=1}^{n_k} (F_{(k,m,\mathbf{N})}(x_{i,k}) - F_{k,m,ols}(x_{i,k}))^2 \end{aligned}$$

where m is the Bernstein degree. The first term stands for the GCV estimate of prediction error in the least square fit and the second term gives the squared error differences between the least square fit and the posterior mean functions. The optimum m is chosen as the minimizer of the $APE(m)$. The added computation time for this procedure is much more reasonable compared to the full Bayesian approach.

A limited simulation is performed to evaluate the merit of the proposed selection procedure. The data generating mechanism is same as that of section 4 with $N = 20, 40, 60$ and $\sigma_{x,k} = 0.02, 0.05$ and 0.1 for each x and k. In figure 6, the prediction error (PE) from cross validation is compared with APE for $N = 20$, where PE and APE show similar pattern. The effect of the selection criterion on the IMSE can be explored as well. We generated 100 Monte Carlo samples. For each sample we selected the best model by the proposed method and looked at the IMSE. From Table 4 the APE based method performs well in terms of IMSE and can be used to select the degree. Given that the observation points are equi-spaced in unit interval, for small value of $N = 20$ relatively higher degree is selected to capture the shape. For larger N, we have more accuracy in terms of IMSE with some increase in the selected degree.

7. Discussion

The proposed Bayesian fit for the Huff curves for California rain gauge data matches those provided by NOAA in terms of accuracy maintaining the monotonicity and non-intersecting properties while the LP method may fail to enforce

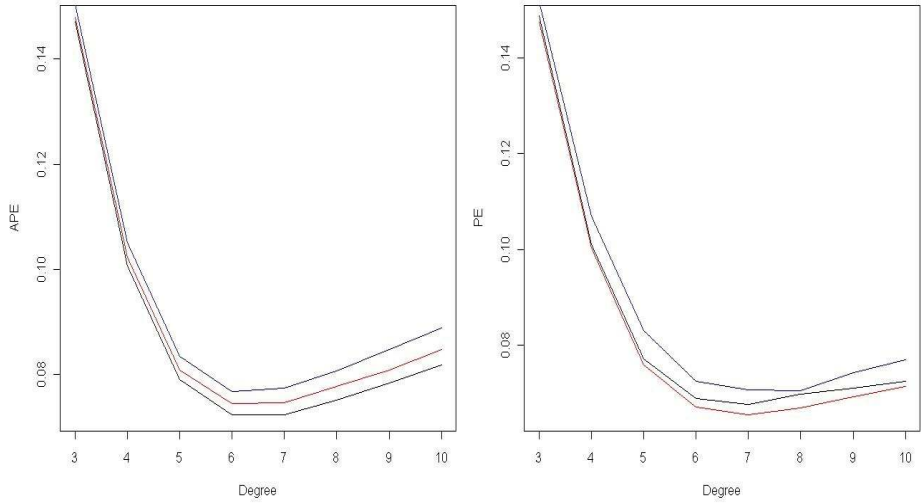

FIG 6. *The measures PE and APE are given for degree in* $m = 3, 4, \ldots 9, 10$ *with* $n_k = 20$ *for all* K*. The PE is estimated by leave one out cross validation. Respective values for* $\sigma_{x,k} = 0.1, 0.05$ *and* 0.02*, for each* x *and* k *are given in blue, red and black.*

TABLE 4
Square root MSE's under the knot selection criteria. Here $\bar{m}$ *denotes the average number of degree selected over all replications*

$\sigma_{x,k}$	N=20		N=40		N=60	
	rmse	$\bar{m}$	rmse	$\bar{m}$	rmse	$\bar{m}$
0.10	.027	6.53	.020	7.02	.016	7.30
0.05	.024	6.47	.016	6.91	.013	7.02
0.02	.022	6.55	.015	6.92	.012	7.05

the constraints. The contrast between the Bayesian methodology and the existing methodology for producing Huff curves can be more pronounced in regions with sparse rain gauge system and high variability of the observed data. Implementing the Bayesian method is computationally no more taxing than the linear programming method for the examples investigated. In fact, for the California data, we did not face any issue with convergence of chains with the Bayesian scheme while there were a few cases where the linear programming method did not converge.

The Bernstein degree for the Bayesian estimation is a tuning parameter and can be chosen by the proposed hybrid approach. The variance function $\sigma_{x,k}$ needs to be specified. While no extensive sensitivity analysis was done the results were robust to the choice of $\sigma_{x,k}$ in the simulation example. As mentioned in Remark 1, we can use small constant values for all $\sigma_{i,k}$ to get least squares type fit in $B^K_{m,\uparrow}$. For the data analysis part, we used $\sigma_{i,k} = 0.01$. One can assume some variance structure satisfying A3 as well. From Remark 2, the posterior

convergence holds for empirical processes, which is compatible with the observed data where observed data points converge to true mean function uniformly over any compact subset of the support, as the number of storms increases. The Bayesian fit can be viewed as a weighted combination of the equispaced knot values, where the weights are given by the Bernstein coefficients. One interesting case may arise, when we have functions whose shapes change rapidly in some subinterval of $[0, 1]$. To capture such cases, we can use local partition. We can partition the interval locally, around some knot and use Bernstein coefficients to split the weight of that knot at equispaced knot points in that interval.

It is possible to generalize this method for modeling a family of monotonic stochastically ordered curves starting from the same point. Such methods can be useful for Intensity-Duration-Frequency curves also produced by NOAA.

8. Appendix

In this section, we provide proofs of the stated results and also give details about posterior computation for the proposed method.

8.1. Proofs

We first state and prove some lemmas that are needed for the proof of Theorem 1. The following notation and constructions are used in the lemmas and also in the proof of Theorem 1.

Given $0 < \eta < 1$, define a class, $\mathbf{C}_\eta$, of piecewise linear continuous cdfs on [0,1] in the following manner: Let $l = \lceil \frac{1}{\eta} \rceil$. Partition the unit square, $[0, 1]^2$, into l^2 smaller squares based on equi-spaced grids along both axes. Any cdf in $\mathbf{C}_\eta$ be continuous piecewise linear such that the join points of the linear segments all belong to the l^2 grid. Let $I_i = [\frac{i-1}{l}, \frac{i}{l}]$, for $i = 1, \ldots, l$. For any strictly monotone cdf F on [0,1], define the approximation of F belonging to $\mathbf{C}_\eta$ as $F_c(\frac{i}{l}) = \frac{m}{l}$ where $\frac{m}{l} < F(\frac{i}{l}) < \frac{m+1}{l}$, and joined linearly in between the grid points.

Lemma 1. *For any cdf F and its approximation F_c in C_η, we have*

$$\sum_{i=1}^{l} \sup_{x \in I_i} |F(x) - F_c(x)| < 3.$$

Proof. From the definition of the approximation F_c, we have

$$\sup_{x \in I_i} |F(x) - F_c(x)| \leq F\left(\frac{i}{l}\right) - \left(F\left(\frac{i-1}{l}\right) - \eta\right).$$

Taking sum over i,

$$\sum_{i=1}^{l} \sup_{x \in I_i} |F(x) - F_c(x)| \leq \sum_{i=1}^{l} \left[F\left(\frac{i}{l}\right) - F\left(\frac{i-1}{l}\right)\right] + l\eta = 1 + l\eta.$$

Since $l\eta < 1 + \eta$, we have the result. □

Lemma 2. *Let y_{ik} denote the observations for the k th curve from 3.1 and let F_k^* be the true value. Suppose A1–A3 hold. Then,*

$$n_k^{-1}\sum_{i=1}^{n_k}\frac{(y_{ik}-F^*(x_{i,k}))(F^*(x_{i,k})-F_c(x_{i,k}))}{\sigma^2_{x_{i,k}}}\to 0$$

a.s. as $n_k\to\infty$, uniformly over $\mathbf{C}\eta$, where $F_c\in\mathbf{C}_\eta$ and $\epsilon_{i,k}$'s are independent.

Proof. Let $0<\delta<1$. Choose $0<a<b<1$ such that $H([a,b])>1-\delta$. Let $w_c(x)=\frac{(F^*(x)-F_c(x))}{\sigma^2_{x,k}}$. Given $\mathbf{C}_\eta$, we partition $[a,b]$ in $B_1,B_2,\ldots,B_{M+1}$ such that $B_1=[a,b_1]$, $B_2=[b_1,b_2],\ldots,B_{M+1}=[b_M,b]$, $\sup_{x_1,x_2\in B_l,c\in C}|w_c(x_1)-w_c(x_2)|<\frac{\delta}{2}$ and $|Q^{-1}_{x_1}(Q_{x_2}(z))-Q^{-1}_{x_3}(Q_{x_4}(z))|<\delta$, whenever $x_1,x_2,x_3,x_4\in B_l$, where l is between 2 and M. Let $\epsilon_{x_i}=y_{ik}-F^*(x_{i,k})$ and a_l be the midpoint of B_l. We denote $\epsilon^*_{x_{i,k}}=Q^{-1}_{a_l}(Q_{x_{i,k}}(\epsilon_{x_{i,k}}))$. Thus, for $2\le l\le M$, we have

$$\begin{aligned}&\left|n_k^{-1}\sum_{i=1}^{n_k}\frac{(y_{ik}-F^*(x_{i,k}))(F^*(x_i)-F_c(x_{i,k}))}{\sigma^2_{x_{i,k}}}\mathbf{1}_{x_{i,k}\in B_l}\right|\\&\quad\le\left|n_k^{-1}\sum_{i=1}^{n_k}\epsilon^*_{x_{i,k}}w_c(a_l)\mathbf{1}_{x_{i,k}\in B_l}\right|+3\delta p_l,\end{aligned}$$

where p_l is the proportion of $x_{i,k}$ in B_l. If $\epsilon_{x_{i,k}}$'s are independent, then $\epsilon^*_{x_i}$ are independent and identically distributed for $x_{i,k}$ in B_l. By the strong law of large number (SLLN) $n_k^{-1}\sum_{i=1}^{n_k}\epsilon^*_{x_{i,k}}w_c(b_l)\mathbf{1}_{x_{i,k}\in B_l}\to 0$ a.s.

For $l=1,M+1$

$$\lim_{n\to\infty}\left|n_k^{-1}\sum_{i=1}^{n_k}\frac{(y_{ik}-F^*(x_{i,k}))(F^*(x_{i,k})-F_c(x_{i,k}))}{{\sigma_{x_{i,k}}}^2}\mathbf{1}_{x_{i,k}\in B_l}\right|\le\delta\int_0^1\frac{1}{\sigma^2_{x,k}}\mathrm{d}x.$$

Because δ is arbitrary, the result follows. □

Before proving Theorem 1, we discuss the relative rate of increase of N and m in the context of providing lower bound of prior probability in a neighborhood of true parameter.

Lemma 3. *Let $G_{k,m}$'s for $k=1,\ldots,K$, are cdf's with support $\frac{i}{m}$'s where the increments of each $G_{k,m}$ follows Dirichlet$(\alpha_1,\ldots,\alpha_m)$ with $\alpha_1=\alpha_2=\cdots=\alpha_m>0$ and $G_{k,m}$'s are independent. Then*

$$P\left(G_{1,m}(x)<G_{2,m}(x)<\cdots<G_{K,m}(x)\text{ for all }x\in\left[\frac{1}{m},1\right)\right)>(K!)^{-m}.$$

Proof. We proceed by induction. Easy to see this result for $m=2$. Suppose it is true for $m-1$. Define, $\tilde{G}_{k,m}(x)=\frac{G_{k,m}(x)-G_{k,m}(\frac{1}{m})}{1-G_{k,m}(\frac{1}{m})}$. The increments of $\tilde{G}_{k,m}$'s follow independent Dirichlet distribution and $\tilde{G}_{k,m}$'s have $m-1$ support-points.

The stochastic ordering of $\tilde{G}_{k,m}$'s implies the stochastic ordering of $G_{k,m}$'s. Thus, conditioning $G_{k,m}$'s on $x = \frac{1}{m}$, from the induction hypothesis

$$P\left(G_{1,m}(x) < G_{1,m}(x) < \cdots < G_{K,m}(x) \text{ for all } x \in \left[\frac{1}{m}, 1\right)\right)$$
$$\geq (K!)^{-m+1}.(K!)^{-1} = (K!)^{-m}.$$

□

The following result can now be stated.

Lemma 4. *Under A1–A4 and (3.4)* $P_{\Pi_{G,m}}(B^K_{m,\uparrow} \cap U_{\frac{\epsilon k_0}{2}}) > p_0(K!)^{-m}$, *for some constant* $p_0 > 0$.

Proof. From the fact that $F_1, \ldots, F_K$ are smooth functions on compact set, given $\epsilon > 0$, there exist r many points $0 < x_1 < \cdots < x_r < 1$ and $1 < i_{m,1} < \cdots < i_{m,r} < m$ such that $\frac{i_{m,j}}{m} \to x_j$ and $G_{k,m}(\frac{i_{m,j}}{m}) \in I_{k,j}$ implies corresponding K curve tuple lies in $U_{\epsilon k_0/2}$ for all $m > m_0$. Here $I_{k,j} = (F_k(x_j) - \nu(\epsilon), F_k(x_j) + \nu(\epsilon))$, where $\nu(\epsilon) > 0$ depends on ϵ. Let $s_{m,r} = \{\frac{i_{m,1}}{m}, \ldots \frac{i_{m,r}}{m}\}$. Also,

$$P\left(G_{k,m}\left(\frac{i_{m,j}}{m}\right) \in I_{k,j}, \forall k \text{ and } G_{1,m}(x) < \cdots < G_{K,m}(x) \text{ for } x \in s_{m,r}\right) > p_0$$

for some $p_0 > 0$ as r is finite. Following the construction of Proposition 3, for $1 < j < r$ the quantities $\tilde{G}_{k,m,j}(x) = \frac{G_{k,m}(x) - G_{k,m}(\frac{i_{m,j}}{m})}{G_{k,m}(\frac{i_{m,j+1}}{m}) - G_{k,m}(\frac{i_{m,j}}{m})}$ for $\frac{i_{m,j}}{m} < x < \frac{i_{m,j+1}}{m}$ has $i_{m,j+1} - i_{m,j} - 1$ many support points and the increments of $\tilde{G}_{k,m,j}(x)$'s follows independent Dirichlet distribution. Ordering of the corresponding Dirichlet distribution imply the ordering of the curves. Let

$$R_{m.r} = \left\{G_{k,m}\left(\frac{i_{m,j}}{m}\right) \in I_{k,j} \text{ and } G_{1,m}(x) < G_{1,m}(x) < \cdots < G_{K,m}(x), x \in s_{m,r}\right\}.$$

Thus, conditioning on the values at $\frac{i_{m,j}}{m}$'s

$$P_{\Pi_{G,m}}(B^K_{m,\uparrow} \cap U_{\frac{\epsilon k_0}{2}})$$
$$> P\left(G_{1,m}(x) < \cdots < G_{K,m}(x) \forall x \in \left[\frac{1}{m}, 1\right); G_{k,m}\left(\frac{i_{m,j}}{m}\right) \in I_{k,j}, j = 1, \ldots, r\right)$$
$$= \int_{R_{m,r}} P(G_{1,m}(x) < \cdots < G_{K,m}(x) \forall x \in \left[\frac{1}{m}, 1\right) \Big| G_{k,m}\left(\frac{i_{m,j}}{m}\right) \in I_{k,j} \text{ and}$$
$$G_{1,m}(x) < G_{1,m}(x) < \cdots < G_{K,m}(x)) \text{ for } x \in s_{m,r}) d(.)$$
$$> p_0(K!)^{-m}.$$

Here the last integral is with respect to the joint density of $G_{k,m}(\frac{i_{m,j}}{m})$'s for $k = 1, \ldots, K$ and $j = 1, \ldots, r$. □

8.1.1. Proof of Theorem 1

Proof. Let $\mathbf{\Delta}^K = \{\Delta^1, \Delta^2, \ldots, \Delta^K\}$, where Δ^k is the parameter vector corresponding to the k th Bernstein cdf and let m be the number of knots. Let F_{Δ^k} be the curve corresponding to Δ^k. We denote $\sigma_k^2(x_{i,k}) = \sigma_{x_{i,k}}^2$. Thus, we can write the likelihood

$$L_{\mathbf{\Delta}^K} = L(\mathbf{\Delta}^K) = \exp\left(\sum_{k=1}^{K} \frac{n_k}{2}(-S_{e^k} - S_{d,\Delta^k} + S_{e,d,\Delta^k}) + c_{0,K}\right)$$

where

$$\begin{aligned} S_{e^k} &= n_k^{-1} \sum_{i=1}^{n_k} \left\{ \frac{(y_{ik} - F_k^*(x_{i,k}))}{\sigma_{x_{i,k}}} \right\}^2, \\ S_{d,\Delta^k} &= n_k^{-1} \sum_{i=1}^{n_k} \left\{ \frac{(F_k^*(x_{i,k}) - F_{\Delta^k}(x_{i,k}))}{\sigma_{x_{i,k}}} \right\}^2, \\ S_{e,d,\Delta^k} &= 2n_k^{-1} \sum_{i=1}^{n_k} \frac{(y_{ik} - F_k^*((x_{i,k}))(F_k^*((x_{i,k}) - F_{\Delta^k}((x_{i,k}))}{\sigma_{x_{i,k}}^2}, \end{aligned}$$

and $c_{0,K}$ is a constant.

Let $U_\epsilon = \{\mathbf{G} : d_K^2(\mathbf{F}^*, \mathbf{G}) < \epsilon\}$. For m large enough we can find $\delta = \delta(m) > 0$ such that $P_{\Pi_{G,m}}(B_{m,\uparrow}^K \cap U_{\frac{\epsilon k_0}{2}}) > \delta$, where $P_{\Pi_{G,m}}$ is probability measure associated with the density $\Pi_{G,m}$. The density $\Pi_{G,m}$ is denoted by Π for simplicity.

Let $A = (B_{m,\uparrow}^K \cap U_{\frac{\epsilon k_0}{2}})$, $B = (B_{m,\uparrow}^K \cap {U_\epsilon}^c)$ and S be the set of all elements in $B_{m,\uparrow}^K$, then

$$P_{\mathbf{N}}(B) = \frac{\int_B \Pi_{\mathbf{N}}(\mathbf{\Delta}^K|\mathbf{Y})d\mathbf{\Delta}^K}{\int_S \Pi_{\mathbf{N}}(\mathbf{\Delta}^K|\mathbf{Y})d\mathbf{\Delta}^K} \leq \frac{\int_B \Pi_{\mathbf{N}}(\mathbf{\Delta}^K|\mathbf{Y})d\mathbf{\Delta}^K}{\int_A \Pi_{\mathbf{N}}(\mathbf{\Delta}^K|\mathbf{Y})d\mathbf{\Delta}^K} = \frac{\int_B \Pi(\mathbf{\Delta}^K)L_{\mathbf{\Delta}^K}d\mathbf{\Delta}^K}{\int_A \Pi(\mathbf{\Delta}^K)L_{\mathbf{\Delta}^K}d\mathbf{\Delta}^K}.$$

Thus,

$$P_{\mathbf{N}}(B) \leq \frac{\sup_{\mathbf{\Delta}^K \in B} L_{\mathbf{\Delta}^K} \int_B \Pi(\mathbf{\Delta}^K) d\mathbf{\Delta}^K}{\inf_{\mathbf{\Delta}^K \in A} L_{\mathbf{\Delta}^K} \int_A \Pi(\mathbf{\Delta}^K) d\mathbf{\Delta}^K}. \tag{8.1}$$

The likelihood can be written as

$$L(\mathbf{\Delta}^K) = \exp\left(N \sum_{k=1}^{K} \frac{w_{k,N}}{2}(-S_{e^k} - S_{d,\Delta^k} + S_{e,d,\Delta^k}) + c_{0,K}\right).$$

For each of the k curves we have

$$S_{d,\Delta^k} = n_k^{-1} \sum_{i=1}^{n_k} \left\{ \frac{(F_k^*(x_{i,k}) - F_{c_{\Delta^k}}(x_{i,k}))}{\sigma_{x_{i,k}}} \right\}^2 + C_1,$$

and

$$S_{e,d,\Delta^k} = 2n_k^{-1}\sum_{i=1}^{n_k}\frac{(y_{ik}-F_k^*((x_{i,k}))(F_k^*((x_{i,k})-F_{c_{\Delta^k}}((x_{i,k}))}{\sigma^2_{x_{i,k}}} + C_2,$$

where $F_{c_{\Delta^k}}$ is the approximation of F_{Δ^k} in $\mathbf{C}_\eta$ for some $\eta > 0$.

Note that, $|C_1| \leq 3n_k^{-1}\sum_{j=1}^{l}\sup_{x_{i,k}\in I_j}|F_{\Delta^k}(x) - F_{c_{\Delta^k}}(x)|\sum_{x_{i,k}\in I_j}\frac{1}{\sigma^2_{x_{i,k}}}$ and $C_2 = \frac{2}{n_k}\sum_{i=1}^{n_k}\frac{(y_{ik}-F^*(x_{i,k}))(F_{\Delta^k}(x_{i,k})-F_c(x_{i,k}))}{\sigma^2_{x_{i,k}}}$.

Using Cauchy-Schwarz inequality we have,

$$|C_2|^2 \leq 4n_k^{-1}\sum_{i=1}^{n_k}\frac{(y_{ik}-F^*(x_{i,k}))^2}{\sigma^2_{x_{i,k}}}\left\{n_k^{-1}\sum_{j=1}^{l}\sup_{x_{i,k}\in I_j}|F_\Delta(x)-F_c(x)|\sum_{x_i\in I_j}\frac{1}{\sigma^2_{x_{i,k}}}\right\}.$$

The cross product terms in S_{e,d,Δ^k} go to zero a.s. by Lemma 2.

As $n_k^{-1}\sum_{x_{i,k}\in I_j}\frac{1}{\sigma^2_{x_{i,k}}} \to \int_{I_j}\frac{1}{\sigma^2_{x,k}}h_k(x)\mathrm{d}x = H(I_j,\sigma_{x,k})$, using the result from Lemma 1 and A3, we can choose η small enough such that $\sup_{j,k}H(I_j,\sigma_{x,k}) \leq \alpha\epsilon$, for any $\alpha > 0$. As $N \to \infty$, for some $\mathbf{\Delta}^K \in B$ we have $\sum w_{k,N}S_{d,\Delta^k} > \epsilon k_0$ and for $\mathbf{\Delta}^K \in A, \sum w_{k,N}S_{d,\Delta^k} < \frac{\epsilon k_0}{2}$. From Lemma 4, $\exp(N\frac{\epsilon k_0}{2})P_{\Pi_{G,m}}(A) \to \infty$ if $\frac{N}{m} \to \infty$, for any $\epsilon > 0$. Also, C_η is a finite set. Hence, $P_{\mathbf{N}}(B) \to 0$ a.s. □

Remark 2. *The conclusion of Lemma 2 and a critical part in the proof of Theorem 1 remain valid under weak correlation between the $\epsilon_{i,k}$'s as long as the strong law used in the proof holds. In case, where $y_{i,k}$'s are dependent quantities based on $n_{1,k}$ observations, such that $y_{ik} = y_{n_{1,k},i}$ converges to $F^*(x_{i,k})$ a.s. uniformly on $[a,b]$ as $n_{1,k}$ goes to infinity, then*

$$n_k^{-1}\sum_{i=1}^{n_k}\frac{(y_{ik}-F^*(x_i))(F^*(x_{i,k})-F_c(x_{i,k}))}{\sigma^2_{x_{i,k}}}\mathbf{1}_{x_{i,k}\in[a,b]} \to 0$$

a.s. as n_k and $n_{1,k}$ goes to infinity. This setup is compatible with the Huff curve example.

Remark 3. *The proof of Theorem 1 is for general error distribution in the setup of model (3.1). The normality assumption was used only for the sampling purpose.*

8.1.2. Proof of Remark 1

Proof. Let $d_{ols} = \inf_{F_k\in B^K_{m,\uparrow}}\{\sum_{k=1}^{K}\sum_{i=1}^{n_k}\frac{(y_{ik}-F_k(x_{i,k}))^2}{v^2_{x_{i,k}}}\}$. Let

$$U_{ols,\frac{\epsilon}{2}} = \left\{\mathbf{\Delta}^K \text{ such that } \sum_{k=1}^{K}\sum_{i=1}^{n_k}\frac{(y_{ik}-F_k(x_{i,k}))^2}{v^2_{x_{i,k}}} < d_{ols}+\frac{\epsilon}{2}\right\}.$$

Then $\Pi(U_{ols,\frac{\epsilon}{2}}) > \delta'$ for some $\delta' > 0$. As σ^2 goes to zero from equation (8.1)

$$P_n(U^c_{ols,\epsilon}) \leq {\delta'}^{-1} \exp\left(-\frac{\epsilon}{4\sigma^2}\right) \to 0. \qquad \square$$

8.2. Posterior sampling

Define $G_{0,m}(\frac{i}{m}) = 0$ and $G_{K+1,m}(\frac{i}{m}) = 1$ for all i in $1, \ldots, (m-1)$, $G_{K+1,m}(1) = G_{0,m}(1) = 1$ and $G_{K+1,m}(0) = G_{0,m}(0) = 0$. Let $a_{k,i} = \min\{G_{k,m}(\frac{i-1}{m}), G_{k-1,M}(\frac{i}{m})\}$ and $b_{k,i} = \max\{G_{k,m}(\frac{i+1}{m}), G_{k+1,m}(\frac{i}{m})\}$ for $k = 1, 2, \ldots, K$ and $i = 1, 2, \ldots, m-1$.

Thus,

$$\begin{aligned}
&\Pi_{\mathbf{N}}\left(G_{k,m}\left(\frac{i}{m}\right)\middle|\mathbf{Y}, \boldsymbol{\Delta}_{\mathbf{K}.i,k}\right) \\
&\quad\propto \prod\nolimits_k \frac{1}{\sqrt{det(\Sigma_k)}} \exp^{-\frac{1}{2}(\underline{Y}_k - \mathbf{M_k})'\Sigma_k^{-1}(\underline{Y}_k - \mathbf{M_k})} \\
&\quad\times \left(G_{k,m}\left(\frac{i}{m}\right) - G_{k,m}\left(\frac{i-1}{m}\right)\right)^{\frac{1}{m}-1} \\
&\quad\times \left(G_{k,m}\left(\frac{i+1}{m}\right) - G_{k,m}\left(\frac{i}{m}\right)\right)^{\frac{1}{m}-1} \mathbf{1}_{a_{k,i} < G_{k,m}(\frac{i}{m}) < b_{k,i}}.
\end{aligned}$$

Note that $\boldsymbol{\Delta}_{\mathbf{K}}$ denotes the set of parameters and $\boldsymbol{\Delta}_{\mathbf{K}.i,k}$ denotes the parameter set excluding the value at the i th knot point for the kth curve that is $G_{k,m}(\frac{i}{m})$. The sampling process can be stated as follows.

At 0th step, we start with initial values of $G_{k,m}$'s satisfying the restriction.

At each iteration, we update for kth curve for $k = K, K-1, \ldots, 1$. For each curve, we update at each knot points for $i = 1, \ldots, m-1$ given the rest of the parameters.

At the Lth iteration, for $k = K, \ldots, 1$ we repeat the following steps.

- Let $G^{(L-1)}_{k,m}(\frac{i}{m})$ be the value for the k th curve, at i th knot point in $(L-1)$ th iteration. The values $a^L_{k,i}, b^L_{k,i}$ are defined as, $a^L_{k,i} = \min\{G^{(L)}_{k,m}(\frac{i-1}{m}), G^{(L-1)}_{k-1,m}(\frac{i}{m})\}$ and $b^L_{k,i} = \max\{G^{(L-1)}_{k,m}(\frac{i+1}{m}), G^{(L)}_{k+1,m}(\frac{i}{m})\}$. Given the rest of the parameters we generate a proposed value of $G_{k,m}(\frac{i}{m})$, lets say $G^*_{k,m}(\frac{i}{m})$ by generating a random number x^* from $Beta(b_1, b_2)$ and setting $G^*_{k,m}(\frac{i}{m}) = a^L_{k,i} + x^*(b^L_{k,i} - a^L_{k,i})$. Let $x = \frac{G^{(L-1)}_{k,m}(\frac{i}{m}) - a^L_{k,i}}{b^L_{k,i} - a^L_{k,i}}$.
- Let $\mathbf{M_k}$ be the vector of fitted values at the design points at this step using $G_{k,m}(\frac{i}{m}) = G^{(L-1)}_{k,m}(\frac{i}{m})$ and $\mathbf{M^*_k}$ be the vector of the fitted values when $G^{(L-1)}_{k,m}(\frac{i}{m})$ is replaced by the proposed value $G^*_{k,m}(\frac{i}{m})$. Then the acceptance probability α can be written as

$$\alpha = \min\left\{\frac{\exp^{-\frac{1}{2}(\underline{Y}_k - \mathbf{M^*_k})'\Sigma_k^{-1}(\underline{Y}_k - \mathbf{M^*_k})}}{\exp^{-\frac{1}{2}(\underline{\mathbf{Y}}_\mathbf{k} - \mathbf{M_k})'\Sigma_k^{-1}(\underline{Y}_k - \mathbf{M_k})}} \frac{p^{(L)}(G^*_{k,m}(\frac{i}{m}), a^L_{k,i}, b^L_{k,i})}{p^{(L)}(G^{(L-1)}_{k,m}(\frac{i}{m}), a^L_{k,i}, b^L_{k,i})} \frac{q_{x,b_1,b_2}}{q_{x^*,b_1,b_2}}, 1\right\},$$

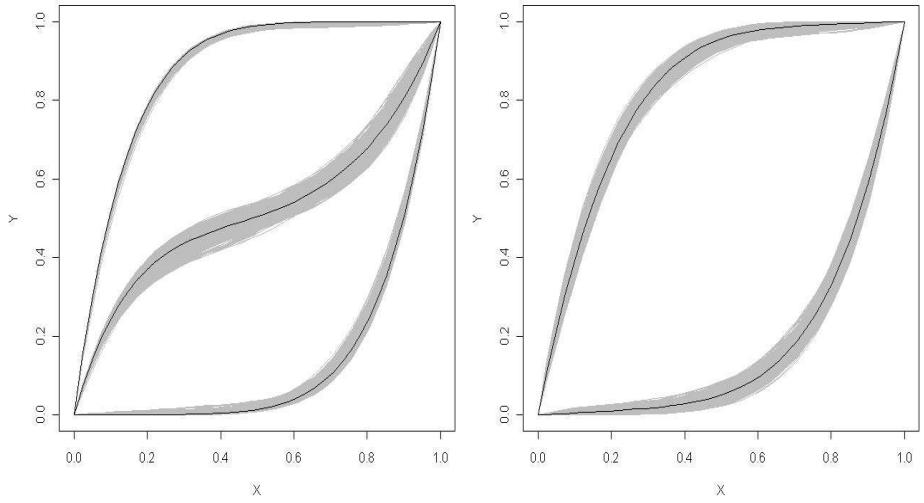

FIG 7. *An example of posterior curves and the mean curves with* $N = 40, m = 7$ *and* $\sigma_{x,k} = 0.05$*, in the setup of section 4. Left hand panel represents the posterior curves and the mean curves for* F_1, F_3, F_5 *and the right hand panel shows the posterior and mean curves for* F_2, F_4*. Fitted mean curves are given in black and the posterior curves are given in gray.*

where

$$p^{(L)}\left(G_{k,m}^{(L-1)}\left(\frac{i}{m}\right), a_{k,i}, b_{k,i}\right) = \left(G_{k,m}^{(L-1)}\left(\frac{i}{m}\right) - G_{k,m}^{(L)}\left(\frac{i-1}{m}\right)\right)^{\frac{1}{m}-1} \times \left(G_{k,m}^{(L-1)}\left(\frac{i+1}{m}\right) - G_{k,m}^{(L-1)}\left(\frac{i}{m}\right)\right)^{\frac{1}{m}-1} \mathbf{1}_{a_{k,i}^L < G_{k,m}(\frac{i}{m}) < b_{k,i}^L},$$

$$p^{(L)}\left(G_{k,m}^{*}\left(\frac{i}{m}\right), a_{k,i}, b_{k,i}\right) = \left(G_{k,m}^{*}\left(\frac{i}{m}\right) - G_{k,m}^{(L)}\left(\frac{i-1}{m}\right)\right)^{\frac{1}{m}-1} \times \left(G_{k,m}^{(L-1)}\left(\frac{i+1}{m}\right) - G_{k,m}^{*}\left(\frac{i}{m}\right)\right)^{\frac{1}{m}-1} \mathbf{1}_{a_{k,i}^L < G_{k,m}(\frac{i}{m}) < b_{k,i}^L},$$

and q_{x,b_1,b_2} is the density value of the *Beta* distribution with parameter b_1, b_2 at point x.
Thus, $G_{k,m}^{(L)}(\frac{i}{m}) = G_{k,m}^{*}(\frac{i}{m})$ with probability α and $G_{k,m}^{(L)}(\frac{i}{m}) = G_{k,m}^{(L-1)}(\frac{i}{m})$ with probability $1-\alpha$.

We tried the values, $b_1 = b_2 = b = 1, .8$ and $.6$ in the proposal density and the result were nearly identical in terms of fitted curves and IMSE values. An example for posterior fitted curves, along with the posterior mean curves is given in figure 7 in the context of the five-curve scenario considered in section 4.

Acknowledgments

We would like to thank the associate editor and the anonymous referee for their valuable inputs.

References

[1] Arjas, E. and Gasbarra, D. (1994). Nonparametric Bayesian inference from right-censored survival data, using the Gibbs sampler. *Statist. Sinica*, **4**, 505–524. MR1309427

[2] Barron, A., Schervish, M. and Wasserman, L. (1999). The consistency of posterior distributions in nonparametric problems. *Ann. Statist.*, **27**, 536–561. MR1714718

[3] Bonta, J. V. and Rao, A. R. (1988). Fitting equations to families of dimensionless cumulative hyetographs. *Transaction of the ASAE*, **31**, 756–760.

[4] Desbordes, M. and Raous, P. (1976). Un example de L'interet des etudes de sensibilite des modeles hydrolegiques. *La Houille Blanche*, **1**, 37–43 (in french).

[5] Gelfand, A. and Kottas, A. (2001). Nonparametric Bayesian modeling for stochastic order. *Ann. Inst. Statist. Math.*, **53**, 865–876. MR1880817

[6] Ghosal, S., Ghosh, J. K. and Ramamoorthi, R. V. (1999). Posterior consistency of Dirichlet mixtures in density estimation. *Ann. Statis.*, **27, 1**, 143–158. MR1701105

[7] Ghosal, S. and van der Vaart, A. (2007). Convergence rates of posterior distributions for noniid observations. *Ann. Statis.*, **35, 1**, 192–223. MR2332274

[8] Hoff, P. (2003). Bayesian methods for partial stochastic orderings. *Biometrika*, **90**, 303–317. MR1986648

[9] Hoff, P., Halberg, R., Shedlovsky, A., Dove, W. and Newton, M. (2001). Identifying carriers of a genetic modifier using nonparametric Bayes methods. In: Gatsonis, C., Kass, R., Carlin, B., Carriquiry, A., Gelman, A., Verdinelli, I., West, M. (Eds.), *Case Studies in Bayesian Statistics*. Springer, New York.

[10] Huff, F. A. (1967). Time distribution of rainfall in heavy storms. *Water Resources Research*, **3**, 1007–1019.

[11] Huff, F. A. (1990). Time distribution of heavy rainstorms in Illinois. *Illinois State Water Survey*, Circular 173, Champaign, III.

[12] Huff, F. A. and Angel, J. R. (1989). Rainfall distribution and hydroclimatic characteristic of heavy rainstorms in Illinois (Buletin 70). *Illinois State Water Survey*, Champaign, III.

[13] Huff, F. A. and Angel, J. R. (1992). Rainfall frequency atlas of the Midwest. *Illinois State Water Survey*, Champaign, III.

[14] Huff, F. A. and Vogel, J. L. (1976). Hydrometeorology of heavy rainstorms in Chicago and Northeastern Illinois. *Illinois State Water Survey Report of Investigation*, **82**, Champaign, III.

[15] Karabatsos, G. and Walker, S. (2007). Bayesian nonparametric inference of stochastically ordered distributions, with Pólya trees and Bernstein polynomials. *Statist. Prob. Lett.*, **77**, 907–913. MR2363440

[16] Perica, S., Dietz, S., Heim, S., Hiner, L., Maitaria, K., Martin, D., Pavlovic, S., Roy, I., Trypaluk, C., Unruh, D., Yan, F., Yekta, M., Zhao, T., Bonnin, G., Brewer, D., Chen, L., Parzybok, T. and Yarchoan, J. (2011). NOAA Atlas 14 Volume 6 Version 2.0, *Precipitation Frequency Atlas of the United States, California*. NOAA, National Weather Service, Silver Spring, MD.

[17] Pilgrim, D. H. and Cordery, I. (1975). Rainfall temporal patterns for design floods. *J. of the Hydraulics Division HYI*, 81–95.

[18] Robertson, T. and Wright, F. (1974). On the maximum likelihood estimation of stochastically ordered random variates. *Ann. Statist.*, **2**, 528–534. MR0388641

[19] Schwartz, L. (1965). On Bayes procedures. *Z. Wahrs.*, **4**, 10–26. MR0184378

[20] Soil Conservation Service (1986). Technical Release 55: Urban Hydrology for Small Watersheds. Department of Agriculture, Washington DC.

[21] Sutherland, F. R. (1982). An improved rainfall intensity distribution for hydrograph synthesis. Univ. of the Witwatersrand, Dept. of Civil engineering, Johannesburg, South Africa. Report No. 1/1983.

[22] Terstriep, M. L. and Stall, J. B. (1974). The Illinois Urban Drainage Area Simulator, ILLUDAS. ISWS B-58 NTIS#: PB237566/AS.

[23] Yen, B. C. and Chow, V. T. (1980). Design Hyetographs for small drainage structures. *J. of the Hydraulics Division*, ASCE 106 (HY6), 1055–1076.