This paper presents a design flood estimation method that combines a continuous hydrological model (DDD) with a stochastic event model (DDDEvent), using Bayes' theorem to derive runoff exceedance probabilities from precipitation exceedance probabilities. The method performs well in rainfall-dominated catchments and offers advantages such as parameter parsimony and ease of implementation, making it potentially valuable for flood estimation in ungauged basins. The topic addresses a clear engineering need, and the methodological approach is innovative. However, several important issues remain regarding theoretical rigor, sufficiency of validation, and delineation of applicability. The authors are encouraged to carefully revise the manuscript before reconsideration.

Major Comments

1. In Section 2.7, Bayes' theorem is used to derive runoff exceedance probabilities from the conditional distributions f(Q|P) and f(P|Q). However, because a uniformly distributed precipitation input (1–300 mm at 0.5 mm intervals) is used, the resulting f(P|Q=Q_T) is not a true conditional distribution but an empirical distribution influenced by the uniform marginal distribution assumption. The authors briefly defend this (lines 290–295), arguing that the uniform distribution is an acceptable approximation in the tail where the relative frequency gradient is small. However, this assumption lacks quantitative validation. The authors are encouraged to quantify the bias introduced by this approximation by comparing with conditional distributions weighted by a realistic precipitation marginal distribution (e.g., GEV), or to theoretically demonstrate that tail exceedance probabilities are insensitive to the shape of the marginal distribution.
2. The paper's title covers "natural and urban catchments," and the abstract claims the method works for catchments ranging from 0.06 to 1092 km². However, the results show that the method completely fails in snowmelt-dominated catchments (e.g., Tannsvatn with SimScore of -10.8), and performance is also poor in mixed rain-snow catchments (e.g., Kråkfoss with MAPE of 47%). While the authors acknowledge this limitation in the Discussion, the overall framing of the paper—particularly the abstract and conclusions—should more honestly delineate the method's boundaries. It is recommended to clearly state that the method is primarily applicable to rainfall-dominated catchments, and that inclusion of snowmelt contributions is future work.
3. The GEV distribution is a three-parameter distribution fitted directly to observed streamflow series, whereas DDDEvent does not use observed flood peaks for calibration—observed discharge is only used for calibrating the DDD model parameters. It is therefore expected that GEV would naturally show better fitting performance. The authors are encouraged to more clearly explain the different positioning of the two approaches, or to adopt a fairer comparison by, for example, performing FFA on long flow series generated by continuous DDD simulation as an additional benchmark, thereby isolating the contribution of the "event model + Bayes" innovation.
4. The paper tests the regionalized DDD model for PUB on only four rainfall-dominated catchments and claims that "PUB estimates are better than calibrated versions" in three of them. However, a sample size of four is too small to support the conclusion that the method "performs well in ungauged catchments." Furthermore, the precipitation time series used for PUB (52 years) differs from that used for the calibrated version (32 years), and the difference in record length itself may cause different results (as demonstrated by the Årdal catchment), confounding the PUB comparison. It is recommended to increase the number of PUB test catchments, or at least to weaken the related conclusions.

5. Urban catchment validation is limited by short data records. The three urban catchments (Sandsli, Lye2, Skivika) have relatively short observation records (2–3 years), and no comparison with an independent method (e.g., SCHADEX or other urban flood estimation methods) is provided. I suggest: Discussing the impact of short calibration periods on the estimation of S distributions (Sandsli uses only 2 years of data). If possible, adding a comparison with standard urban hydrology methods (e.g., Rational Method, SWMM).

Minor Comments

Line 407: "Fig.87" is a typo and should be "Fig. 8"

Line 415: PUB outperforms calibrated versions is misleading given the small sample size

Line 430: The physical interpretation and handling of initial S exceeding S_cap needs to be supplemented

Line 618: "Inclusion of the contribution of snow meltwater...", this sentence is incomplete.

Line 740: Steinbakk et al. reference format is inconsistent.

In addition, please further review the manuscript to ensure that the semantics are more clearly articulated.