Received: 28 November 2023 | Revised: 15 February 2024 | Accepted: 18 March 2024
DOI: 10.1002/sta4.676

ORIGINAL ARTICLE

WILEY

Using extreme value theory to evaluate the leading pedestrian interval road safety intervention

Nicola Hewett[1] | Lee Fawcett[1] | Andrew Golightly[2] | Neil Thorpe[3]

[1]School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne, UK
[2]Department of Mathematical Sciences, Durham University, Durham, UK
[3]Jacobs, Rotterdam House, Newcastle upon Tyne, UK

Correspondence
Nicola Hewett, School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne NE1 7RU, UK.
Email: nicola.hewett@ncl.ac.uk

Improving road safety is hugely important with the number of deaths on the world's roads remaining unacceptably high; an estimated 1.3 million people die each year as a result of road traffic collisions. Current practice for treating collision hotspots is almost always reactive: once a threshold level of collisions has been overtopped during some pre-determined observation period, treatment is applied (e.g., road safety cameras). Traffic collisions are rare, so prolonged observation periods are necessary. However, traffic *conflicts* are more frequent and are a margin of the social cost; hence, traffic conflict before/after studies can be conducted over shorter time periods. We investigate the effect of implementing the leading pedestrian interval treatment at signalised intersections as a safety intervention in a city in north America. Pedestrian-vehicle traffic conflict data were collected from treatment and control sites during the before and after periods. We implement a before/after study on post-encroachment times (PETs) where small PET values denote 'near-misses'. Hence, extreme value theory is employed to model extremes of our PET processes, with adjustments to the usual modelling framework to account for temporal dependence and treatment effects.

KEYWORDS
before-after analysis, bivariate threshold excess model, extreme value theory (EVT), leading pedestrian interval (LPI), post-encroachment time (PET), traffic conflicts

1 | INTRODUCTION

Around 1.3 million people are killed every year as a result of road traffic accidents (RTAs), and between 20 and 50 million people suffer from non-fatal injuries—around half of these are vulnerable road users such as pedestrians (WHO, 2020). The National Highway Traffic Safety Administration found that, in the United States, the second most common cause for injury-classed collisions is left-turning vehicles, and 22% of crashes at intersections involve vehicles turning left (NCSA, 2021). Furthermore, 23% of pedestrian crashes occurred at intersections.

With the safety of pedestrians in mind, safety treatments at signalised intersections have been investigated. 'The leading pedestrian interval (LPI) is one treatment that has been implemented at signalised intersections to permit pedestrians to begin crossing several seconds before the release of conflicting vehicle movements' (van Houten et al., 2000). This, theoretically, should reduce potential conflicts between pedestrians and vehicles. A significant reduction in potential conflicts will likely lead to a significant reduction in actual collisions. With such an intervention, there is the added benefit of treating potential road safety hotspots *proactively*. Standard road safety interventions are usually analysed *reactively*, once a threshold collision count has been over-topped during some pre-determined observation period. Collisions are rare events, and so prolonged observation periods are necessary, with much waiting for collisions to happen to evaluate a treatment using a standard before/after (BA) analysis (perhaps using Empirical Bayes methodology—see, for example, Fawcett & Thorpe, 2013; Fawcett et al., 2017; Hauer, 1980). Working with near-misses, as the LPI intervention does, means not having to wait for collisions to happen; the LPI is implemented by adjusting the signal-phasing and pedestrian interval to provide a walk display of several seconds before the adjacent vehicle green display, making this an efficient and low-cost safety measure.

BA studies are prominent in road safety intervention evaluations. Traditionally, RTA BA studies are focused on the reduction in frequency and/or severity of collisions from before the intervention, to after (see, e.g., Elvik, 2002, Hauer, 1997, Sayed & Sacchi, 2016). Traffic conflicts are generally defined as a situation in which a pedestrian and a vehicle approach each other in time and space to such an extent that they will collide if their movements remain unchanged. Traffic conflicts are more frequent than collisions and easy to observe and are of marginal social cost (Tarko, 2018). As such, traffic conflict BA studies can be conducted over shorter time periods. Furthermore, new technology has been adopted, such as automated video techniques (Saunier & Sayed, 2007) and traffic simulation (Wang et al., 2018), to allow for detecting and tracking moving objects based on their trajectories; conflict data are thus easily extracted.

Commonly, statistical inference is based on averages obtained from datasets, with procedures utilising the central limit theorem being employed—supporting the use of the standard Normal distribution or associated t distribution. As we shall discuss in the next section, the LPI intervention, and associated traffic conflict data on collision near-misses, justifies the use of the *extremal types theorem* instead. The associated extreme value distributions, introduced by Fisher and Tippett (1928), arise as limits for the distribution of maxima (or minima) in sequences of independent, identically distributed random variables. Typical applications occur in the environmental sciences to model, for example, extreme precipitation events or extreme wind speeds (see, e.g., Arun et al., 2023, Fawcett & Walshaw, 2006); for interested readers, we point to the classical reference of Gumbel (1958) and, more recently, the tutorial-style text of Coles (2001). In this paper, we focus on the lower tail of the data where small time values indicate a dangerous situation: a near-miss or collision between a pedestrian and left-turning vehicle.

The remainder of this paper is organised as follows. A brief description of the data is given in Section 2. The methodology of extreme value theory is explained in Section 3 including describing two methods to handle dependence and the inclusion of covariates. In Section 4, we outline the details of the Bayesian inference scheme, before considering the real data application in Section 4.5. Conclusions are drawn in Section 5.

2 | DATA

We have data from a national collision database, investigating the implementation a 5-s LPI at eight intersections in a north American city, to give pedestrians more time to cross before a left-turning car is released. At each intersection, one crosswalk has been treated. Data were collected from 1 March to 30 June 2018 (before period) and 1 August to 31 October 2018 (after period), with data coming from the same 12-h period each day (8:00–20:00) to mitigate potential confounders such as cyclic variation. Data spanning exactly the same time period are also available for a further seven intersections that have not been treated with the LPI intervention, for comparison purposes.

In this dataset, a conflict between a vehicle and a pedestrian is indicated using *post-encroachment time* (PET). PET is the time between the moment the first road user passes the conflicting point, t_1, and the moment the second user reaches that point, t_2. The positions of the vehicle and pedestrian are shown in Figure 1, where PET $= t_1 - t_2$. We have the minimum PET in 10-min intervals over the 12-h study period in each day, for all intersections. PETs <15 s were recorded. As PET $= t_1 - t_2$, if $t_1 = t_2$, then we have a collision between pedestrian and vehicle. Small PET values imply a near-miss, a value close to zero implying a dangerous situation. Our aim is to model extremely small PET values using extreme value theory (EVT) and through this modelling template investigate differences between extremes from periods before and after the LPI treatment was introduced.

In order to use standard methods from the EVT toolkit directly (designed for analysing 'large' extremes), we negate our series of PET values at each location, thus switching the focus from very small values to very large values to identify dangerous situations in our series. Figure 2 shows a time series plot of these negated PET values, and a plot of observations at neighbouring time points, for one of the intersections at which the LPI was applied. Note the apparent systematic decrease in PET values after the LPI implementation at this specific intersection; note also the clear temporal dependence between consecutive data values, persisting into the extremes of the process.

3 | METHODOLOGY

3.1 | Background

EVT has become popular in traffic conflict analyses. Zheng and Sayed (2019a) use a peaks over threshold approach including covariates in the scale parameter for crash estimation; furthermore, they also use EVT on block maxima from a traffic conflict BA study (Zheng & Sayed, 2019b). Wang et al. (2019) use bivariate EVT to predict annual crash frequencies at intersections. Fu et al. (2021) use hierarchical EVT modelling on traffic conflict extremes for crash estimation. Guo et al. (2020) analysed the effectiveness of LPI treatment at two signalised intersections in Vancouver,

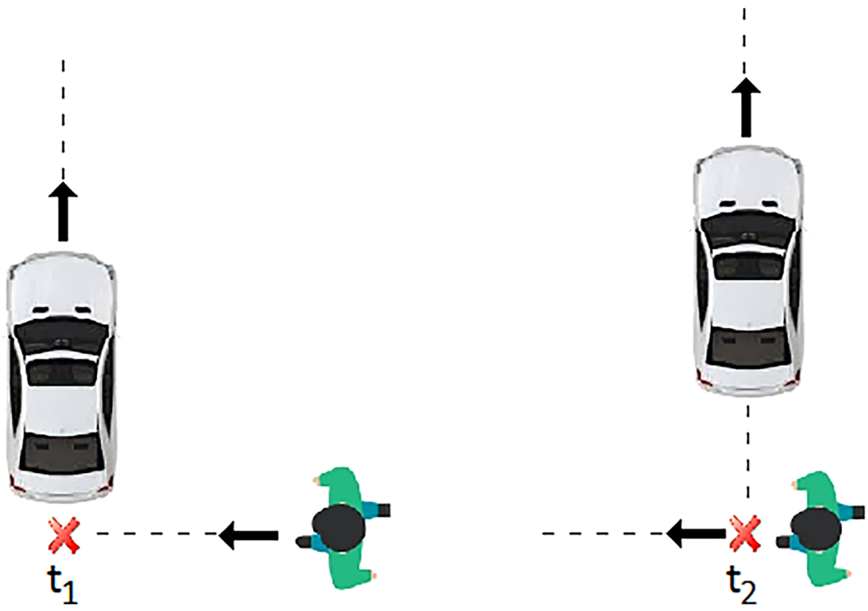

FIGURE 1 Post-encroachment times (PET), $PET = t_2 - t_1$

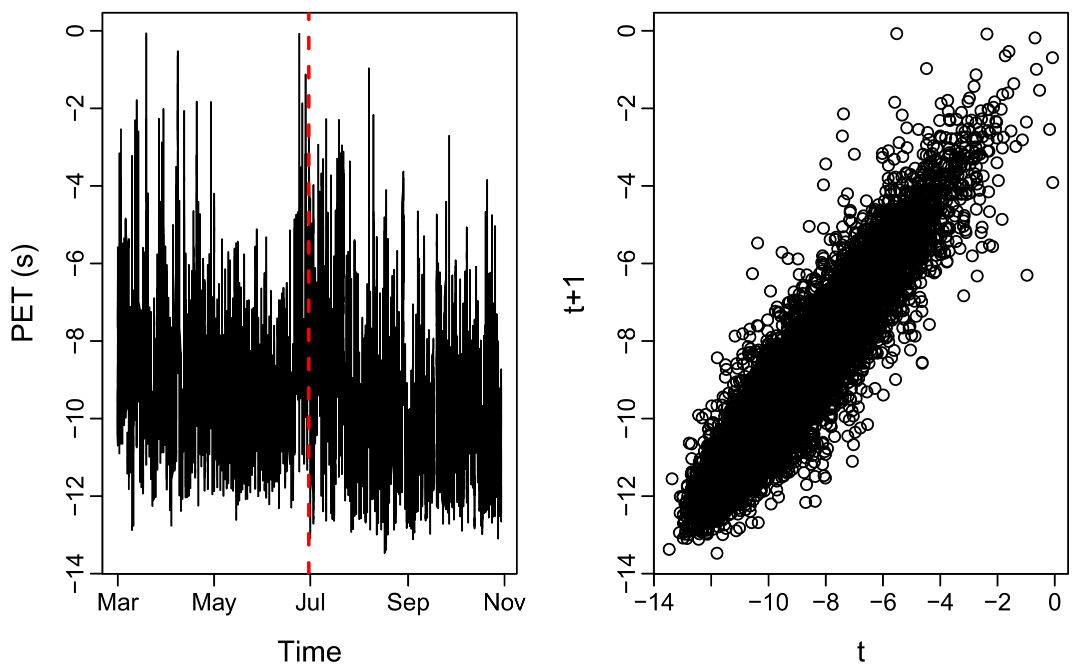

FIGURE 2 Left: time series plot of negated PET values (seconds) before and after intervention at Intersection 1, the vertical red line showing the start of the after period. Right: temporal dependence of observations at neighbouring time points for intersection 1 (t on x axis against $t+1$ on y axis)

Canada. They modelled the scale and shape parameters of the GPD as a function of a treatment indicator, a period indicator (before/after) and an interaction of the two variables where dependence between consecutive extremes was removed through a declustering approach. In this paper, we use EVT in an attempt to capture treatment effects from the LPI intervention described in Section 1. Our method aims to maximise data usage by performing a threshold-based analysis (as opposed to a block maxima analysis; see, e.g., Coles, 2001). We also allow for a dependence between consecutive extremes, as identified in Figure 2, through a first-order Markov chain structure, and a treatment effect through linear modelling of one of the parameters in the extreme value model used.

EVT provides a framework for estimating the probability of extreme events. The extremal type theorem gives limiting results for the extremes of our process, providing a range of techniques for modelling the tail behaviour of our random variable without any assumptions about the underlying distribution of the data itself. There are two main approaches: the so-called 'block maxima approach' and the 'threshold-based approach'. In the former, a series of observations are blocked into fixed time intervals (e.g., years, giving rise to an 'annual maxima analysis'), and the maximum observation of each block is extracted to create our series of extremes. In a threshold-based analysis, all observations exceeding a suitable, pre-defined threshold are classified as extreme. This study proposes the use of a threshold-based approach on the premise of including more data in the analysis than the block maxima approach, which can be wasteful through discarding all but the most extreme observations in each block. It is hoped that the inclusion of more datapoints will lend greater precision to the analysis, including reduced uncertainty in our estimated treatment effect.

3.2 | Classifying and modelling extremes

Consider a stationary sequence of random variables, $\{X_1,...,X_n\}$, each with the same distribution function (d.f.), F. Let M_n denote the maximum of the set, such that $M_n = \max\{X_1,...,X_n\}$. As $n \to \infty$, it is typically the case that

$$Pr(M_n \le x) \approx F^{n\theta}(x), \tag{1}$$

where $\theta \in (0,1)$ and is the *extremal index*; for further details, see Leadbetter and Rootzén (1988). As $\theta \to 0$, the process displays increasing dependence in its extremes. In the case of an independent process, $\theta = 1$, as $Pr(M_n \le x) = Pr(X_1, X_2, ..., X_n \le x) = Pr(X_1 \le x) \times Pr(X_2 \le x) \times ... = F^n(x)$. Traditional Extreme Value Theory (EVT) initially focused on the independence case, aiming to discover limiting models for F^n without reference to the marginal distribution F; small discrepancies in F could lead to substantial discrepancies in F^n.

In the study of extreme value theory, the Extremal Types Theorem provides a framework for the limiting behaviour of the maximum value M_n as $n \to \infty$ (see Fisher & Tippett, 1928; Gnedenko, 1943). Analogous to the Central Limit Theorem, where the sample mean converges to the population mean, the limiting distribution of M_n is degenerate, meaning the distribution converges to the upper endpoint of the distribution F with probability 1. The theorem asserts that if constants $a_n > 0$ and b_n exist, then

$$Pr\{(M_n - b_n)/a_n \le x\} \to G(x),$$

where G exists; it will conform to one of three distributions: Gumbel (I), Fréchet (II) or Weibull (III), with associated distribution functions,

$$\begin{aligned}
I: G(x) &= \exp\{-\exp(-x)\}, -\infty < x < \infty;\\
II: G(x) &= \begin{cases} 0 & x \le 0,\\ \exp(-x^{-\alpha}) & x > 0, \alpha > 0;\end{cases}\\
III: G(x) &= \begin{cases} \exp\{-(-x)^{\alpha}\} & x < 0, \alpha > 0,\\ 1 & x \ge 0.\end{cases}
\end{aligned}$$

Collectively, these are termed the extreme value distributions. For Gumbel and Fréchet types, the upper–endpoint of the limiting distribution G tends to ∞, while the Weibull type exhibits a finite upper limit. Importantly, the theorem neither guarantees the existence of a non-degenerate limit for M_n nor specifies which of the three types will apply if such a limit exists. Nevertheless, when a limit distribution does exist, it will align with one of the types defined by the Extremal Types Theorem, irrespective of the original distribution F.

3.2.1 | Block maxima: The generalised extreme value distribution

Independently, von Mises (1954) and Jenkinson (1955) introduced the Generalized Extreme Value (GEV) distribution, which unifies all three extreme value distributions (Gumbel, Fréchet and Weibull). As the limiting model for F^n, the d.f. of the GEV is given by

$$\mathcal{G}(y;\mu,\sigma,\xi) = \begin{cases} \exp\left[-(1+\xi(y-\mu)/\sigma)_+^{-1/\xi}\right], & \xi \ne 0\\ \exp[-\exp(-(y-\mu)/\sigma)], & \xi = 0\end{cases},$$

defined on $\{y : 1+\xi(y-\mu)/\sigma>0\}$, where $-\infty<\mu<\infty$; $\sigma>0$ and $-\infty<\xi<\infty$ are parameters of location, scale and shape, respectively; and $a_+ = \max(0,a)$. It is noteworthy that $F^{n\theta}$ also follows the GEV distribution with d.f. $\mathcal{G}(y;\mu^*,\sigma^*,\xi)$, where $\mu^*=\mu-\frac{\sigma}{\xi}\left(1-\theta^{\xi}\right)$, $\sigma^*=\sigma\theta^{\xi}$ when $\xi\neq 0$, and $\mu^*=\mu+\sigma\ln\theta$, $\sigma^*=\sigma$ when $\xi=0$. This is contingent on satisfying Leadbetter's $D(u_n)$ condition, as indicated in Leadbetter et al. (1983). Hence, for block maxima, in practical terms, short-range dependence can be ignored. In practice, the GEV is employed to model maximum values over a convenient calendar unit, often yearly. However, selecting an appropriate block size can present challenges. In our specific study, which spans only months, employing years as blocks would be unsuitable. Blocks must be sufficiently large for the limiting theory to apply, yet if too large, the available maxima for inference become sparse. Furthermore, this approach can be hugely wasteful of data, as it retains only the block maxima while discarding the rest of the data. We wish to retain as much information as possible on the extremes of the process; however, this approach may inadvertently eliminate other significant extremes that are not the absolute maximum within their respective blocks.

3.2.2 | Threshold excesses: The generalised Pareto distribution

Pickands (1975) showed that for an independent process ($\theta=1$) and large enough threshold, u, $(X-u\,|\,X>u)$ follows a Generalised Pareto Distribution (GPD) with d.f.

$$\mathcal{H}(y)=\begin{cases}1-(1+\xi y/\tilde{\sigma})^{1/\xi}, & \xi\neq 0\\ 1-\exp(-y/\tilde{\sigma}), & \xi=0\end{cases}, \tag{2}$$

defined on $y>0$, with scale and shape parameters $\tilde{\sigma}$ and ξ, respectively. Here, $\tilde{\sigma}$ is related to the parameters in the corresponding GEV for block maxima through $\tilde{\sigma}=\sigma+\xi(u-\mu)$. Threshold-based methods identify observations as extreme when they exceed a high threshold, commonly denoted as u. The GPD described in Equation (2) is then fitted to these threshold excesses. Section 3 will illustrate graphical diagnostics available for optimal threshold selection. Unlike GEV-based modelling of block maxima, powering F^n (denoted $\mathcal{H}$ for the GPD) by θ, as specified by Equation (1), does not result in another extreme value distribution incorporating the extremal index into its parameters. This necessitates careful scrutiny of extremal dependence, which is often more prevalent in consecutive threshold excesses compared to consecutive block maxima.

3.3 | Handling dependence

As discussed, our aim here is to maximise precision by including as much data in the analysis as possible; hence, we will proceed with a threshold-based approach to analysis. In this section, we describe two methods for handling temporal dependence present between consecutive threshold excesses: a declustering approach, leading to the commonly-used ‘peaks over threshold’ analysis of a filtered subset of extremes, and an approach that explicitly models the transition from one extreme to the other through a bivariate extreme value model. As the second plot in Figure 2 reveals, even above a high threshold, there appears to be dependence between consecutive PET values. At busier times of the day—perhaps early morning or late-afternoon—we might expect more pedestrians to be using the crosswalks at each of our intersections, and a greater number of vehicles turning into the intersections, perhaps resulting in a greater number of near-misses (with an associated clustering of small PET values) at these times. Clustering of extremes is common in many other applications of the threshold approach to extreme value modelling—for example, dependence between consecutive temperature or wind speed extremes and serial correlation in extremes obtained from financial time series. Ignoring such dependence will likely lead to under-estimated uncertainty measures (e.g., confidence intervals that are unrealistically narrow); see Barao and Tawn (1999) and Shi et al. (1992).

3.3.1 | Declustering

The goal of the declustering approach is to isolate a sequence of independent threshold excesses, thereby validating the approximation $\theta\approx 1$ in Equation (1). To achieve this, an auxiliary ‘declustering parameter’, denoted as κ, is selected. A cluster of threshold excesses is considered to have ended when at least κ consecutive observations fall below the threshold. This process is iteratively applied across the entire data series to identify clusters. Subsequently, the highest observation, or ‘peak’, is extracted from each identified cluster. The GPD is then fitted to the, hopefully independent, peak excesses. This technique aims to handle extremal dependence effectively by focusing only on the most extreme values within each cluster. This ‘peaks over threshold’ approach (POT; Davison & Smith, 1990) is a prevalent strategy for addressing clustered extremes. While straightforward to implement, the technique does face challenges, particularly concerning the optimal choice of the declustering parameter, κ. A low κ value risks insufficient separation between cluster peaks, undermining the assumption of independence. Conversely, a high κ value results in fewer clusters for inference, making the approach data-inefficient. Additionally, the sensitivity of parameter estimates to κ has been highlighted in previous studies (Fawcett & Walshaw, 2012). Given how commonplace POT analyses for threshold excesses have become, we will include results based on this approach as a baseline for comparing our results using a first-order extreme value Markov chain for modelling dependence.

3.3.2 | Modelling dependence: First-order extreme value Markov chain

To avoid declustering, we can account for dependence between consecutive extremes by assuming a first-order Markov structure. For example, we can assume the following joint density for our series of (negated) PET values at each intersection:

$$\frac{\prod_{i=1}^{n-1} g(y_i, y_{i+1}; \Theta)}{\prod_{i=2}^{n-1} g(y_i; \Theta)}, \tag{3}$$

where Θ is a generic parameter vector. In a threshold excess context, univariate contributions to the denominator in Equation (3) are given through the GPD (on differentiation of Equation (2)). Appealing to bivariate EVT, transformation from GPD to unit Fréchet margins (see, e.g., Coles, 2001) provides a range of models to use for contributions to the numerator in Equation (3), the most commonly used being the logistic family with d.f.:

$$H(y_i, y_{i+1}) = \exp\left\{-\left(y_i^{-1/\alpha} + y_{i+1}^{-1/\alpha}\right)^{\alpha}\right\}, \tag{4}$$

where H is a non-degenerate bivariate extreme value distribution. Here, $\alpha \in (0,1]$ controls the extent of extremal dependence in the process, with independence giving $\alpha = 1$ and $\alpha \to 0$ corresponding to increasing levels of extremal dependence. Differentiation of Equation (4), with careful censoring when one or both of (y_i, y_{i+1}) lie sub-threshold, gives pairwise contributions to the numerator in Equation (3). Interested readers are referred to Coles (2001) for further information and a more detailed discussion of bivariate EVT more generally. Where direct evaluation of Equation (1) is necessary—for example, when obtaining quantiles from the fitted distribution (often used as estimates of *return levels* in applications of EVT)—Fawcett and Walshaw (2012) provide an approximation to the extremal index θ based on the estimated logistic dependence parameter α.

3.4 | Including covariates

When the data admit non-stationarity—for example, trend or a dependence on covariates—we can attempt to incorporate this non-stationarity through linear modelling of the GEV or GPD parameters. Generally, we can write the extreme value parameters in the form $h(X^T\beta)$, where h is a specified function, β is a vector of parameters and X is the model vector. Recall that the GPD $(\tilde{\sigma}, \xi)$ arises from the GEV (μ, σ, ξ), where the GPD scale parameter is a function of the GEV location and shape parameters. Thus, attempting to model any trend in our threshold excesses is usually done through linear modelling of the scale parameter, $\tilde{\sigma}$. The PET data we are investigating have before/after time implications at each of the 15 intersections; hence, we attempt to capture the treatment effect through the following parameterisation of $\tilde{\sigma}$:

$$\tilde{\sigma}_t = \exp(\beta_0 + \beta_1 t), \tag{5}$$

to respect the positivity of $\tilde{\sigma}$ and where $t = 0$ in the before period and $t = 1$ in the after period. Hence, at each intersection, we might use the slope parameter β_1 as a proxy for our LPI treatment effect; an estimate of β_1 that might be deemed significantly different from zero might be indicative of a treatment effect at an intersection.

4 | APPLICATION

4.1 | Threshold selection

The *threshold stability property* of the GPD means that if it is a valid model for excesses over some threshold u_0, then it is valid for excesses over all thresholds $u > u_0$. Furthermore, for all $u > u_0$, $E[X - u | X > u]$ is a linear function of u. In practice, this expectation can be estimated empirically as the sample mean of the excesses over u. This leads to the mean residual life plot (MRL plot; see, e.g., Coles, 2001): a graphical procedure for identifying a suitably high threshold for modelling extremes via the GPD in which mean excesses over u are plotted against a range of values for u, and the optimal threshold is chosen at the lowest point above which we observe linearity in the plot. MRL plots are a standard diagnostic tool for threshold selection and have been constructed using PET data at each of our intersections to provide site-wise thresholds to identify extremes.

4.2 | Bayesian inference

Section 3 covers methodology for modelling extreme values, with a particular focus on threshold methods and handling temporal dependence. We choose the Bayesian paradigm within which to make inferences on these models. A typical statistical analysis might formulate the likelihood function for the assumed statistical model, maximising this with respect to the parameters in that model to obtain their *maximum likelihood estimates*: values that maximise the likelihood that the process described by the model produced the data that were actually observed. In a classical sense, these are sample-based estimates of fixed but unknown quantities. In the Bayesian paradigm, the likelihood is merely an ingredient in the inferential process; via Bayes Theorem, it is combined with the density of the *prior distribution* to provide a *posterior distribution* for the parameters of interest—in effect an update in our beliefs about the parameters after having observed some data, relative to our beliefs before observing these data. With a careful choice of prior distribution, in some cases, it is possible to obtain the precise form of posterior distribution analytically; for example, in the *conjugate* case, both the prior and posterior are from the same family of distributions, and the application of Bayes Theorem is trivial.

Crucially, the interpretation of the model parameters is different in the Bayesian setting: rather than being fixed (but unknown) constants, as in the classical setting, the parameters are now regarded as random variables, with a distribution (prior and posterior). This means that in the Bayesian setting, confidence intervals, for example, have a much more natural interpretation, with there being a probability of .95 that the parameter falls within the bounds of the 95% Bayesian confidence (or credible) interval. But the advantages of working within the Bayesian setting stretch beyond the interpretation of resulting confidence intervals. For example, by their very nature extremes are scarce, and being able to supplement an analysis with more information through the prior distribution has the potential to increase estimation precision. For example, Fawcett and Walshaw (2006) record up to an 82% reduction in uncertainty in estimated wind speed extremes when comparing Bayesian and frequentist approaches to inference. In the analysis of rainfall extremes, Walshaw and Smith (2003) noted an 84% reduction in the posterior standard deviation in their estimates when using informative priors formed through discussions with a hydrologist, compared to the corresponding standard errors from a maximum likelihood analysis. Also, when working with extreme value models, we do not need to worry about the regularity conditions surrounding maximum likelihood estimation of the shape parameter ξ, resulting in maximum likelihood estimates being unobtainable when $\xi < -0.5$—for full details, see Coles (2001).

Bayesian inference is commonplace in road safety BA studies. A popular method is *empirical Bayes*, which is used to account for regression to the mean effects (see, e.g., Fawcett & Thorpe, 2013; Hauer, 1980). More recently, *full Bayes* methods have been used to filter effects of regression to the mean from genuine treatment effects in road safety schemes (see, e.g., El-Basyouny & Sayed, 2012; Heydari et al., 2014). As Fawcett and Thorpe (2013) discuss, a full Bayes analysis can provide a more flexible inferential framework with a range of prior distributions beyond the conjugate case being available; it can also provide a more realistic assessment of uncertainty in estimated treatment effects.

4.3 | Prior specification

In the absence of any expert prior information regarding our model parameters, and for ease of computation, we adopt fairly uninformative, independent prior distributions. Recall that, for each intersection, we adopt a bivariate threshold excess model for negated PET values exceeding a threshold (identified and validated through use of an MRL plot). The margins are assumed GPD with a linear model enabling the scale parameter to vary between before and after periods via Equation (5); for the dependence between successive threshold excesses we adopt a logistic model with a parameter quantifying the degree of serial correlation present. Thus, our parameter vector at each intersection can be written as

$$\Theta = (\beta_0, \beta_1, \xi, \alpha)^T,$$

for which we set the following prior:

$$\pi(\Theta) = \pi(\beta_0)\pi(\beta_1)\pi(\xi)\pi(\alpha),$$

and where

$$\beta_0 \sim N(0,10), \quad \beta_1 \sim N(0,10), \quad \xi \sim N(0,100) \quad \text{and} \quad \alpha \sim U(0,1).$$

4.4 | MCMC sampling

It is often the case that the posterior distribution cannot be found analytically. For example, in the case of the GPD, there exists no conjugate prior specification for the model parameters, and so we cannot easily find the posterior distribution $\pi(\Theta|\mathbf{x})$; we cannot plot this distribution nor find its moments, etc. Fortunately, techniques to sample from the posterior have been developed and are now routinely used, thanks to their inclusion in many popular statistical software packages. The most common approach is to use a *Markov chain Monte Carlo* (MCMC) scheme (see, e.g., Casella & George, 1992). A simple MCMC algorithm is a Normal random walk Metropolis-Hastings (MH) scheme (Hastings, 1970). Although full details are omitted here, it is noted that—according to some careful 'tuning' of the algorithm through the choice of variance in the Normal random walk updating for the parameter vector Θ—it is possible to optimise the algorithm, with the sample returned forming a Markov chain whose stationary distribution is the posterior distribution $\pi(\Theta|\mathbf{x})$. Depending on the starting values chosen, an initial period B might be discarded as 'burn-in', to ensure the sample used is indeed from the stationary distribution; to minimise the effects of autocorrelation, it is also common to 'thin' the sample and use posterior draws from just every kth iteration.

In our MCMC scheme, we set initial values for all parameters to their prior means, using a simple MH random walk update to give successive draws

$$\left(\beta_0^{(j)},\beta_1^{(j)},\xi^{(j)},\alpha^{(j)}\right), \quad j=1,\ldots,10^5,$$

after thinning by every $k=10$ iterations to obtain sufficiently low autocorrelation between realisations and removing the first $B=2{,}000$ iterations as burn-in. A pilot run was implemented to help tune the scheme, with methods from Roberts and Rosenthal (2001) being used to quickly optimise the algorithm.

4.5 | Results

Figure 3 shows the posterior means and 95% credible intervals for β_1 over all sites, from analyses that (i) ignore dependence, (ii) filter out dependence through declustering and (iii) explicitly model the dependence via our first-order extreme value Markov chain model. The treated sites are denoted with a 'T' on the x axis. In our model, β_1 captures the treatment effect as the slope term in our linear predictor for the GPD scale parameter.

When we ignore dependence, we use all threshold excesses, and hence, our credible intervals are relatively narrow owing to the maximal use of extreme data. However, we are violating the assumption that consecutive threshold excesses are independent; as such, these credible intervals are likely to be unrealistically narrow. Declustering (here with $\kappa=10$) removes this dependence, but at a cost: reduced datasets, with Site 1, for example, now having just 107 threshold excesses post-declustering (from an original 17,467 raw excesses). We have done nothing to 'optimise' the declustering interval κ here, and it could be that our choice of κ is unnecessarily large resulting in a procedure that is wasteful of data

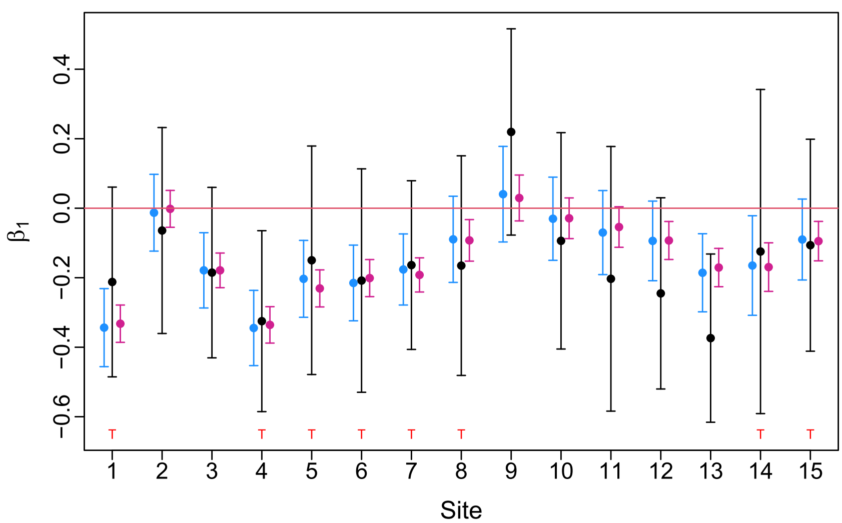

FIGURE 3 Posterior means and 95% CIs for β_1 over all sites from ignoring dependence (pink), declustering (grey) and the logistic model (blue). Treated sites denoted with 'T' above the x axis

(giving relatively wide credible intervals). However, even if we were to investigate an optimal choice of declustering interval κ, our extreme value Markov chain model for explicitly modelling dependence is probably a superior approach here, as it maximises data usage while also making some effort to capture the dependence in the series of threshold excesses.

As β_1 corresponds to the treatment effect, a value of $\beta_1 < 0$ shows a successful treatment as this indicates smaller (negated) PET values in the after period (in other words, larger values on the raw PET scale, meaning a move away from a near-miss/actual collision). Of course, there is uncertainty in our inference, so we look for 95% credible intervals that are wholly negative to identify a successful treatment effect. In the majority of the treated sites $\beta_1 = 0$ lies outside the range of the 95% credible intervals, hence, we conclude these sites have seen an improvement post-treatment, with increased PETs. Notice that, for some sites, treatment effects have not been identified under the POT approach using declustered extremes; the loss in precision, owing to the use of much smaller datasets, often results in wider credible intervals that include zero. Examples include Sites 1, 5–7 and 14—all sites that have been treated with the LPI intervention, and whose credible intervals for β_1 are wide enough to include zero under the POT approach (but under our extreme value Markov chain model these intervals are wholly negative).

The numerical summaries of the marginal posterior distributions for each parameter, including their mean and 95% credible intervals, are presented in Table 1 for Sites 1–3, along with the chosen threshold. The full results for all sites can be found in the appendix, specifically in Table A1. For all sites and analyses, the posterior mean of ξ is negative, meaning the GPD has a 'light tail' and is upper bounded. This implies that the distribution has a low probability of producing extreme events. Similarly, as seen in Figure 3 for the posterior summaries of β_1, the posterior credible intervals for parameters β_0 and ξ when using declustering are much wider than those from the logistic model. Furthermore, the declustering analysis results in the posterior mean estimates for ξ being more negative meaning we would expect an even lower probability of extreme events. Hence, including all of the data has hugely improved the parameter estimates in terms of uncertainty and has altered our interpretation on the occurrence of extreme events.

5 | DISCUSSION

An extreme value Markov chain was proposed to conduct a traffic conflict-based BA safety evaluation, modelling threshold excesses with the GPD marginally, and using a bivariate extreme value model for the temporal evolution of our extremes, at each intersection. Our approach combines traffic conflicts at different sites (treatment sites and control sites) and for different periods (the before period and the after period) to estimate potential treatment effects of the LPI intervention. Pedestrian traffic conflict data were collected from the treatment and control sites

TABLE 1 Numerical posterior summaries for model parameters at Intersections 1–3

Model	Site	Threshold, u		β_0	β_1	ξ
Ignore dependence	1	−5.6930	Mean	0.4108	−0.3324	−0.1809
			95% CI	(0.3253, 0.4963)	(−0.3861, −0.2786)	(−0.2360, −0.1259)
Declustering			Mean	0.9363	−0.2123	−0.3516
			95% CI	(0.5911, 1.2816)	(−0.4853, 0.0607)	(−0.5916, −0.1115)
Logistic			Mean	0.3981	−0.3436	−0.1649
			95% CI	(0.3015, 0.4948)	(−0.4559, −0.2314)	(−0.2221, −0.1077)
Ignore dependence	2	−4.8800	Mean	0.1058	−0.0019	−0.1373
			95% CI	(0.0147, 0.1968)	(−0.0549, 0.0512)	(−0.1949, −0.0797)
Declustering			Mean	0.6136	−0.0642	−0.2552
			95% CI	(0.2878, 0.9394)	(−0.3605, 0.2321)	(−0.4900, −0.0204)
Logistic			Mean	0.1297	−0.0129	−0.1454
			95% CI	(0.0251, 0.2344)	(−0.1234, 0.0975)	(−0.2082, −0.0826)
Ignore dependence	3	−5.9600	Mean	0.4687	−0.2308	−0.2119
			95% CI	(0.3967, 0.5408)	(−0.2841, −0.1775)	(−0.2538, −0.1700)
Declustering			Mean	0.9325	−0.1498	−0.3003
			95% CI	(0.6369, 1.2280)	(−0.4786, 0.1790)	(−0.5286, −0.0719)
Logistic			Mean	0.4695	−0.2033	−0.1910
			95% CI	(0.3858, 0.5533)	(−0.3139, −0.0926)	(−0.2368, −0.1452)

Note: For each parameter, we show the posterior mean and 95% credible interval from the three different models (ignoring dependence, declustering and logistic) and the chosen threshold.

during the before and after periods using automated computer vision analysis techniques. The treatment effects were measured through linear modelling in the scale parameter of the GPD; we avoided unnecessary data wastage, common in the POT approach to analysis, by accounting for temporal dependence and thus including all threshold excesses in the analysis.

Use of the bivariate logistic model to account for extremal dependence in our PET processes resulted in narrower credible intervals for our defined treatment effect than a declustering approach and was also more successful in correctly identifying sites that had been treated. The posterior distributions for our treatment effect parameter coincide for the model ignoring dependence and our extreme value Markov chain model, with similar posterior means. This is also true for the other parameters in our model. However, although ignoring dependence and fitting a GPD to all threshold excesses results in the narrowest credible intervals, when we ignore dependence our model is no longer valid as our processes clearly exhibit dependence, even in the extremes. Using a declustering approach, to remove dependence, results in wider credible intervals, which makes drawing conclusions about the treatment effect ambiguous. This approach also results in more negative posterior estimates of the shape parameter, which results in shorter tails in the fitted GPD, thus implying that extreme events are less likely.

Our findings suggest interesting possibilities for extending the modelling approach, particularly considering observed spatial dependence in the data. Incorporating methods from the spatial extremes toolbox could provide a valuable opportunity to account for this spatial dependence and potentially lead to a further increase in precision. As highlighted by Arun et al. (2023), while our current approach utilises stationary thresholds at the treated sites, it is important to acknowledge that the LPI treatment might influence these thresholds. Therefore, exploring dynamic thresholds that capture the potential changes induced by the treatment in both the before and after periods could result in improved model fit. Overall, our proposed approach showcases the potential for utilising extreme value Markov chain models in traffic conflict-based BA safety evaluations, and our findings provide valuable insights into the effectiveness of the LPI intervention, paving the way for further investigations in this area.

AUTHOR CONTRIBUTIONS

Nicola Hewett: Conceptualisation; formal analysis; investigation; methodology; software; visualisation; writing—original draft, review and editing. **Lee Fawcett:** Conceptualisation; methodology; software; writing—original draft, review and editing. **Andrew Golightly:** Methodology; writing—original draft, review and editing. **Neil Thorpe:** Writing—review and editing.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Nicola Hewett https://orcid.org/0009-0000-1519-7326

REFERENCES

Arun, A., Lyon, C., Sayed, T., Washington, S., Loewenherz, F., Akers, D., Ananthanarayanan, G., Shu, Y., Bandy, M., & Haque, M. M. (2023). Leading pedestrian intervals—Yay or nay? A before-after evaluation of multiple conflict types using an enhanced non-stationary framework integrating quantile regression into bayesian hierarchical extreme value analysis. *Accident Analysis & Prevention*, *181*, 106929. https://doi.org/10.1016/j.aap.2022.106929

Barao, M. I., & Tawn, J. A. (1999). Extremal analysis of short series with outliers: Sea-levels and athletics records. *Journal of the Royal Statistical Society: Series C: Applied Statistics*, *48*(4), 469–487. https://doi.org/10.1111/1467-9876.00166

Casella, G., & George, E. I. (1992). Explaining the Gibbs sampler. *The American Statistician*, *46*(3), 167–174. https://doi.org/10.1080/00031305.1992.10475878

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer Series in Statistics. Springer.

Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society: Series B: Methodological*, *52*(3), 393–425.

El-Basyouny, K., & Sayed, T. (2012). Measuring direct and indirect treatment effects using safety performance intervention functions. *Safety Science*, *50*(4), 1125–1132. https://doi.org/10.1016/j.ssci.2011.11.008

Elvik, R. (2002). The importance of confounding in observational before-and-after studies of road safety measures. *Accident Analysis & Prevention*, *34*(5), 631–635. https://doi.org/10.1016/S0001-4575(01)00062-8

Fawcett, L., & Thorpe, N. (2013). Mobile safety cameras: Estimating casualty reductions and the demand for secondary healthcare. *Journal of Applied Statistics*, *40*(11), 2385–2406. https://doi.org/10.1080/02664763.2013.817547

Fawcett, L., Thorpe, N., Matthews, J., & Kremer, K. (2017). A novel Bayesian hierarchical model for road safety hotspot prediction. *Accident Analysis & Prevention*, *99*, 262–271. https://doi.org/10.1016/j.aap.2016.11.021

Fawcett, L., & Walshaw, D. (2006). A hierarchical model for extreme wind speeds. *Applied Statistics*, *55*(5), 631–646.

Fawcett, L., & Walshaw, D. (2012). Estimating return levels from serially dependent extremes. *Environmetrics*, *23*(3), 272–283. https://doi.org/10.1002/env.2133

Fisher, R. A., & Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, *24*(2), 180–190. https://doi.org/10.1017/S0305004100015681

Fu, C., Sayed, T., & Zheng, L. (2021). Multi-type Bayesian hierarchical modeling of traffic conflict extremes for crash estimation. *Accident Analysis & Prevention*, *160*, 106309. https://doi.org/10.1016/j.aap.2021.106309

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics*, *44*, 423–453. https://doi.org/10.2307/1968974

Gumbel, E. J. (1958). *Statistics of extremes*. Columbia University Press. https://doi.org/10.7312/gumb92958

Guo, Y., Sayed, T., & Zheng, L. (2020). A hierarchical Bayesian peak over threshold approach for conflict-based before-after safety evaluation of leading pedestrian intervals. *Accident Analysis & Prevention*, *147*, 105772. https://doi.org/10.1016/j.aap.2020.105772

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97–109. https://doi.org/10.1093/biomet/57.1.97

Hauer, E. (1980). Bias-by-selection: Overestimation of the effectiveness of safety countermeasures caused by the process of selection for treatment. *Accident Analysis & Prevention*, *12*(2), 113–117. https://doi.org/10.1016/0001-4575(80)90049-4

Hauer, E. (1997). *Observational before/after studies in road safety: Estimating the effect of highway and traffic engineering measures on road safety*. Emerald Group Publishing Limited.

Heydari, S., Miranda-Moreno, L. F., & Liping, F. (2014). Speed limit reduction in urban areas: A before–after study using Bayesian generalized mixed linear models. *Accident Analysis & Prevention*, *73*, 252–261. https://doi.org/10.1016/j.aap.2014.09.013

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, *81*(348), 158–171. https://doi.org/10.1002/qj.49708134804

Leadbetter, M., Lindgren, G., Rootzén, H., & Nosko, V. (1983). *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer New York. https://doi.org/10.1007/978-1-4612-5449-2

Leadbetter, M. R., & Rootzén, H. (1988). Extremal theory for stochastic processes. *The Annals of Probability*, *16*(2), 431–478.

National Center for Statistics and Analysis. (2021). Traffic safety facts 2019: A compilation of motor vehicle crash data.

Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, *3*(1), 119–131.

Roberts, G. O., & Rosenthal, J. S. (2001). Optimal scaling for various metropolis-Hastings algorithms. *Statistical Science*, *16*(4), 351–367.

Saunier, N., & Sayed, T. (2007). Automated analysis of road safety with video data. *Transportation Research Record*, *2019*(1), 57–64. https://doi.org/10.3141/2019-08

Sayed, T., & Sacchi, E. (2016). Evaluating the safety impact of increased speed limits on rural highways in British Columbia. *Accident Analysis & Prevention*, *95*, 172–177. https://doi.org/10.1016/j.aap.2016.07.012

Shi, D., Smith, R. L., & Coles, S. G. (1992). Joint versus marginal estimation for bivariate extremes. Technical report, North Carolina State University. Dept. of Statistics.

Tarko, A. P. (2018). Surrogate measures of safety. In *Safe mobility: Challenges, methodology and solutions*. Emerald Publishing Limited. https://doi.org/10.1108/S2044-994120180000011019

van Houten, R., Retting, R. A., Farmer, C. M., & van Houten, J. (2000). Field evaluation of a leading pedestrian interval signal phase at three urban intersections. *Transportation Research Record*, *1734*(1), 86–92. https://doi.org/10.3141/1734-13

Von Mises, R. (1954). La distribution de la plus grande de n valeurs. In *Selected papers* (Vol. ii) (pp. 271–294). American Mathematical Society.

Walshaw, D., & Smith, E. (2003). *Modelling bivariate extremes in a region*. Newcastle University.

Wang, C., Xu, C., & Dai, Y. (2019). A crash prediction method based on bivariate extreme value theory and video-based vehicle trajectory data. *Accident Analysis & Prevention*, *123*, 365–373. https://doi.org/10.1016/j.aap.2018.12.013

Wang, C., Xu, C., Xia, J., Qian, Z., & Lu, L. (2018). A combined use of microscopic traffic simulation and extreme value methods for traffic safety evaluation. *Transportation Research Part C: Emerging Technologies*, *90*, 281–291. https://doi.org/10.1016/j.trc.2018.03.011

WHO. (2020). Road traffic injuries. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries. [accessed 24-May-2021].

Zheng, L., & Sayed, T. (2019a). Bayesian hierarchical modeling of traffic conflict extremes for crash estimation: A non-stationary peak over threshold approach. *Analytic Methods in Accident Research.*, *24*, 100106. https://doi.org/10.1016/j.amar.2019.100106

Zheng, L., & Sayed, T. (2019b). A full bayes approach for traffic conflict-based before–after safety evaluation using extreme value theory. *Accident Analysis & Prevention*, *131*, 308–315. https://doi.org/10.1016/j.aap.2019.07.014

AUTHOR BIOGRAPHIES

Nicola Hewett is a lecturer of Mathematical Sciences at Newcastle University, UK. She earned her PhD from the same university in 2023, and her research focuses on applying statistical analysis to improve road safety.

Lee Fawcett is a reader in Applied Statistics in the School of Mathematics, Statistics and Physics and an associate dean (global), at Newcastle University in the UK. Lee's primary research areas include environmental extremes and statistical methods for road safety. He also has a passion for global education initiatives, including setting up mobility opportunities for under-represented student groups.

Andrew Golightly is a professor of Statistics in the Department of Mathematical Sciences at Durham University. He received his PhD from Newcastle University in 2006 and was appointed to a lectureship in the same year. He was promoted to a senior lecturer in 2016 and a reader in 2019 before moving to Durham University in 2022 and taking up a chair in 2023. Andrew is particularly interested in computational Bayesian inference and regularly publishes in this area.

Neil Thorpe is a senior associate director at Jacobs in the UK and currently engaged in a range of transport planning projects in the UK and overseas. Prior to joining Jacobs in 2022, Neil was a senior lecturer in Transport and Head of the Future Mobility Group in the School of Civil Engineering and Geosciences at Newcastle University, from where he also received his PhD in 2005. His main research focus at Newcastle

University was the analysis of road traffic collision data to inform the evaluation of road safety countermeasures and to predict the location of road traffic collision hotspots.

How to cite this article: Hewett, N., Fawcett, L., Golightly, A., & Thorpe, N. (2024). Using extreme value theory to evaluate the leading pedestrian interval road safety intervention. *Stat*, *13*(2), e676. https://doi.org/10.1002/sta4.676

APPENDIX A.

TABLE A1 Numerical posterior summaries for model parameters at all intersections

Model	Site	Threshold, *u*		β_0	β_1	ξ
Ignore dependence	1	−5.6930	Mean	0.4108	−0.3324	−0.1809
			95% CI	(0.3253, 0.4963)	(−0.3861, −0.2786)	(−0.2360, −0.1259)
Declustering			Mean	0.9363	−0.2123	−0.3516
			95% CI	(0.5911, 1.2816)	(−0.4853,0.0607)	(−0.5916, −0.1115)
Logistic			Mean	0.3981	−0.3436	−0.1649
			95% CI	(0.3015, 0.4948)	(−0.4559, −0.2314)	(−0.2221, −0.1077)
Ignore dependence	2	−4.8800	Mean	0.1058	−0.0019	−0.1373
			95% CI	(0.0147, 0.1968)	(−0.0549,0.0512)	(−0.1949, −0.0797)
Declustering			Mean	0.6136	−0.0642	−0.2552
			95% CI	(0.2878, 0.9394)	(−0.3605, 0.2321)	(−0.4900, −0.0204)
Logistic			Mean	0.1297	−0.0129	−0.1454
			95% CI	(0.0251, 0.2344)	(−0.1234, 0.0975)	(−0.2082, −0.0826)
Ignore dependence	3	−5.9500	Mean	0.4687	−0.2308	−0.2119
			95% CI	(0.3967, 0.5408)	(−0.2841, −0.1775)	(−0.2538, −0.1700)
Declustering			Mean	0.9325	−0.1498	−0.3003
			95% CI	(0.6369, 1.2280)	(−0.4786, 0.1790)	(−0.5286, −0.0719)
Logistic			Mean	0.4695	−0.2033	−0.1910
			95% CI	(0.3858, 0.5533)	(−0.3139, −0.0926)	(−0.2368, −0.1452)
Ignore dependence	4	−5.7400	Mean	0.4632	−0.3357	−0.2066
			95% CI	(0.3821, 0.5442)	(−0.3881, −0.2834)	(−0.2502, −0.1630)
Declustering			Mean	0.9938	−0.3250	−0.3682
			95% CI	(0.6649, 1.3226)	(−0.5854, −0.0647)	(−0.6117, −0.1247)
Logistic			Mean	0.4792	−0.3447	−0.2071
			95% CI	(0.3865, 0.5719)	(−0.4529, −0.2364)	(−0.2630, −0.1513)
Ignore dependence	5	−5.9600	Mean	0.4687	−0.2308	−0.2119
			95% CI	(0.3967, 0.5408)	(−0.2841, −0.1775)	(−0.2538, −0.1700)
Declustering			Mean	0.9325	−0.1498	−0.3003
			95% CI	(0.6369, 1.2280)	(−0.4786, 0.1790)	(−0.5286, −0.0719)
Logistic			Mean	0.4695	−0.2033	−0.1910
			95% CI	(0.3858, 0.5533)	(−0.3139, −0.0926)	(−0.2369, −0.1452)
Ignore dependence	6	−6.7500	Mean	0.409	−0.2011	−0.1837
			95% CI	(0.3245, 0.4935)	(−0.2544, −0.1478)	(−0.2309, −0.1364)
Declustering			Mean	0.8900	−0.2082	−0.2974
			95% CI	(0.5393, 1.2407)	(−0.5297, 0.1133)	(−0.4987, −0.0962)

TABLE A1 (Continued)

Model	Site	Threshold, u		β_0	β_1	ξ
Logistic			Mean	0.4177	−0.2151	−0.1719
			95% CI	(0.3211, 0.5144)	(−0.3240, −0.1062)	(−0.2229, −0.1208)
Ignore dependence	7	−5.5300	Mean	0.3845	−0.1920	−0.1823
			95% CI	(0.3051, 0.4639)	(−0.2411, −0.1428)	(−0.2256, −0.1390)
Declustering			Mean	0.9984	−0.1637	−0.4098
			95% CI	(0.6883, 1.3084)	(−0.4064, 0.0791)	(−0.6374, −0.1822)
Logistic			Mean	0.3999	−0.1763	−0.2046
			95% CI	(0.3082, 0.4915)	(−0.2786, −0.0740)	(−0.2541, −0.1551)
Ignore dependence	8	−7.4020	Mean	0.3183	−0.0925	−0.0834
			95% CI	(0.2354, 0.4012)	(−0.1523, −0.0327)	(−0.1317, −0.0351)
Declustering			Mean	1.0146	−0.1652	−0.2931
			95% CI	(0.6750, 1.3542)	(−0.4812, 0.1508)	(−0.5060, −0.0801)
Logistic			Mean	0.3153	−0.0895	−0.1186
			95% CI	(0.2190, 0.4116)	(−0.2136, 0.0345)	(−0.1728, −0.0643)
Ignore dependence	9	−6.1300	Mean	0.1779	0.0295	−0.1706
			95% CI	(0.0718, 0.2841)	(−0.0366, 0.0955)	(−0.2325, −0.1088)
Declustering			Mean	0.8241	0.2195	−0.4150
			95% CI	(0.4972, 1.1510)	(−0.0773, 0.5163)	(−0.6876, −0.1424)
Logistic			Mean	0.2068	0.0404	−0.1358
			95% CI	(0.0859, 0.3276)	(−0.0971, 0.1779)	(−0.2066, −0.0650)
Ignore dependence	10	−5.1600	Mean	0.1200	−0.0289	−0.1340
			95% CI	(0.0279, 0.2122)	(−0.0874, 0.0296)	(−0.1901, −0.0778)
Declustering			Mean	0.6670	−0.0939	−0.2505
			95% CI	(0.3498, 0.9843)	(−0.4052, 0.2175)	(−0.4641, −0.0369)
Logistic			Mean	0.1327	−0.0303	−0.1072
			95% CI	(0.0286, 0.2368)	(−0.1500, 0.0894)	(−0.1741, −0.0402)
Ignore dependence	11	−6.5155	Mean	0.2028	−0.0541	−0.1249
			95% CI	(0.1095, 0.2960)	(−0.1124, 0.0042)	(−0.1661, −0.0837)
Declustering			Mean	0.7097	−0.2032	−0.1831
			95% CI	(0.3112, 1.1082)	(−0.5839, 0.1776)	(−0.4316, 0.0654)
Logistic			Mean	0.2102	−0.0701	−0.1239
			95% CI	(0.1033, 0.3171)	(−0.1910, 0.0508)	(−0.1809, −0.0668)
Ignore dependence	12	−6.2600	Mean	0.3638	−0.0929	−0.1961
			95% CI	(0.2757, 0.4518)	(−0.1476, −0.0382)	(−0.2467, −0.1455)
Declustering			Mean	1.0513	−0.2451	−0.3788
			95% CI	(0.6634, 1.4393)	(−0.5202, 0.0300)	(−0.6586, −0.0989)
Logistic			Mean	0.3778	−0.0941	−0.1688
			95% CI	(0.2774, 0.4783)	(−0.2086, 0.0205)	(−0.2255, −0.1121)
Ignore dependence	13	−5.0500	Mean	0.2562	−0.1708	−0.1470
			95% CI	(0.1641, 0.3482)	(−0.2259, −0.1157)	(−0.1919, −0.1021)
Declustering			Mean	0.9524	−0.3738	−0.3970
			95% CI	(0.6237, 1.2811)	(−0.6159, −0.1318)	(−0.6729, −0.1212)
Logistic			Mean	0.2456	−0.1858	−0.1634
			95% CI	(0.1394, 0.3518)	(−0.2982, −0.0734)	(−0.2260, −0.1008)
Ignore dependence	14	−10.4100	Mean	0.3635	−0.1695	−0.0788
			95% CI	(0.2895, 0.4374)	(−0.2394, −0.0996)	(−0.1207, −0.0368)

(Continues)

TABLE A1 (Continued)

Model	Site	Threshold, u		β_0	β_1	ξ
Declustering			Mean	0.7947	−0.1247	−0.0480
			95% CI	(0.4865, 1.1029)	(−0.5910, 0.3417)	(−0.2650, 0.1689)
Logistic			Mean	0.3666	−0.1649	−0.0572
			95% CI	(0.2836, 0.4496)	(−0.3082, −0.0216)	(−0.1056 −0.0087)
Ignore dependence	15	−5.1635	Mean	0.2499	−0.0947	−0.1548
			95% CI	(0.1612, 0.3386)	(−0.1514, −0.0380)	(−0.2156, −0.0940)
Declustering			Mean	0.6906	−0.1064	−0.2655
			95% CI	(0.3828, 0.9985)	(−0.4114, 0.1986)	(−0.4958, −0.0352)
Logistic			Mean	0.2289	−0.0901	−0.1454
			95% CI	(0.1274, 0.3304)	(−0.2066, 0.0264)	(−0.2074, −0.0833)

Note: For each parameter, we show the posterior mean and 95% credible interval from the three analyses (ignoring dependence, declustering and logistic) and the chosen threshold.