Scientific African 29 (2025) e02810

Contents lists available at ScienceDirect

Scientific African

journal homepage: www.elsevier.com/locate/sciaf

Bayesian prior elicitation for malaria modelling

Makwelantle Asnath Sehlabana [a],*, Daniel Maposa [a], Alexander Boateng [b], Sonali Das [c]

[a] University of Limpopo, Department of Statistics and Operations Research, Polokwane, 0727, South Africa
[b] Modern College of Business and Science, Department of Mathematics and Computer Science, Muscat, 112, Sultanate of Oman
[c] University of Pretoria, Department of Business Management, Pretoria, 0028, South Africa

ARTICLE INFO

Editor name: Ebenezer Bonyah

Keywords:
Analytic Hierarchy Process (AHP)
Expert knowledge
Statistical epidemiology
Climate factors
AHP weights

ABSTRACT

Subjective Bayesian methods, which incorporate expert knowledge into disease modelling, remain underutilised in epidemiology. This is despite the growth of knowledge in statistical approaches to disease analysis. Objective priors are often favoured for their simplicity. However, subjective Bayesian approaches can produce more informative models by using expert insights, such as those related to malaria transmission. This study focuses on translating expert knowledge into prior probability distributions through a process known as prior elicitation. Prior elicitation presents several challenges. Converting expert judgments into probability distributions is complex and often requires specialised tools. Established methods like the Sheffield approach are resource-intensive, requiring considerable time and cognitive effort from both experts and researchers. To address these limitations, this study makes a major contribution by integrating the Analytic Hierarchy Process (AHP) with statistical validation techniques to quantify expert knowledge into prior probability distributions. Expert knowledge was collected through questionnaires and structured as pairwise comparisons. These were quantified into AHP weights, representing the relative importance of environmental factors influencing malaria transmission. The weights were then fitted to various probability distributions and evaluated using goodness-of-fit tests. Results showed that the beta, gamma and normal distributions best represented the elicited expert knowledge. These findings suggest that beta, gamma and normal distributions are suitable as prior distributions in Bayesian models of malaria transmission. By simplifying the elicitation process and reducing technical complexity, this approach offers a practical framework for applying subjective Bayesian methods in epidemiology. Future research will compare these elicited priors with objective priors to evaluate their impact on model performance across domains.

Introduction

Malaria remains a major public health challenge, which makes accurate modelling essential in epidemiological research. Although statistical epidemiology offers valuable tools to understand disease dynamics, a key question arises: how well is prior knowledge leveraged to develop more accurate and informed models? Despite the vast amount of epidemiological data available, subjective Bayesian methods are underutilised in the field [1]. Frequentist methods, although widely applied, treat each analysis as independent, disregarding previous knowledge [2]. In contrast, Bayesian methods provide a powerful alternative by incorporating

* Corresponding author.
E-mail address: asnath.sehlabana@ul.ac.za (M.A. Sehlabana).

https://doi.org/10.1016/j.sciaf.2025.e02810
Received 7 March 2025; Received in revised form 7 May 2025; Accepted 10 June 2025

Available online 25 June 2025

prior knowledge such as expert insights on malaria transmission alongside current data, resulting in more informed statistical analyses [3].

The value of Bayesian approaches lies in the selection of prior distributions that best utilise available prior knowledge [4]. However, defining these distributions is not without challenges. Subjective Bayesian analysis, which involves the explicit use of expert knowledge, can be difficult because of the complexities of translating that knowledge into statistical terms. Consequently, many researchers turn to objective prior distributions, which are more general but less tailored to specific subject knowledge [5]. Subjective priors, also known as informative priors, offer an alternative by incorporating specific prior knowledge, allowing for broader inference beyond the sample data [6]. These priors are typically constructed through a process known as elicitation.

Elicitation refers to the process of converting expert knowledge and beliefs into a probability distribution known as a subjective prior distribution [7]. Experts, who have extensive knowledge through research or experience, can provide insight that helps inform these distributions. Methods for eliciting priors include expert elicitation, meta-analysis, and previous data analysis [8]. Despite the growing popularity of Bayesian methods, the field of prior elicitation remains underutilised in practice due to the complexity of the elicitation process and the challenges of generalising it across different domains [9]. The study of Mikkola et al. [10] identified several barriers to the broader adoption of prior elicitation, including the need for improved software integration, cost-effective evaluation techniques, and more streamlined processes.

Two primary approaches to elicitation have been distinguished: the predictive approach, in which experts provide estimates of the response variable at fixed values of explanatory variables, and the structural approach, in which experts focus on parameters of interest [11]. The predictive approach is often favoured because it does not require experts to have in-depth statistical knowledge, they only need to provide insights into observable quantities [5]. Elicitation methods can be further categorised into behavioural combinations, in which experts collaborate to reach consensus, and mathematical aggregation, where individual expert opinions are combined into a single prior distribution [12].

Most structural methods, such as the widely used Sheffield method, follow the behavioural combination approach, where experts undergo training, review evidence, and collaborate to form their opinions [13]. However, achieving consensus among experts with differing opinions can be challenging, and the logistical demands of gathering and training experts, especially in international contexts, make this method resource-intensive [14]. In light of these challenges, researchers often resort to alternative elicitation methods. For example, Scharzsenegger [15] employed data pooling on similar processes to estimate prior distributions, acknowledging that expert elicitation can be prohibitively costly.

In addition to logistical hurdles, experts can struggle to express their knowledge in statistical terms such as density functions, means, or standard deviations [16]. Effective elicitation methods should allow non statisticians to contribute their expertise without overwhelming them with technical requirements. However, many structural methods require researchers to design custom elicitation software, and experts may need additional training to use such tools, as seen in the studies by Garthwaite et al. [17] and Jansen et al. [18], where experts were required to graph their priors using specialised software. Examples of commonly used software include SHELF and MATCH [19]. In some cases, experts must also have skills in statistical modelling, mapping, or GIS, further increasing the burden on both researchers and experts [5,12].

Given the challenges of structural elicitation methods, researchers have increasingly turned to mathematical aggregation methods. For example, Lan et al. [20] employed the equal weighted linear pooling method to combine individual priors. Although this approach reduces the logistical burden, it introduces the challenge of selecting appropriate pooling rules. To address this, some studies use an analytic hierarchy process (AHP) to guide the elicitation of priors, which has become a common framework for constructing subjective prior distributions [21]. The strength of AHP lies in its ability to combine knowledge from multiple experts into normalised weights that can inform a single prior distribution.

Most AHP elicitation methods are not designed to be universally applicable. For example, O'Leary et al. [12] developed a model specifically tailored for logistic regression, while Van de Schoot et al. [8] proposed an approach to define priors and estimate models in the context of latent growth mixture modelling (LGMM). Similarly, Syed et al. [22] used AHP to summarise expert estimates of time to failure in the form of informative prior distributions, which were then applied to Bayesian updates for Weibull distribution. However, these approaches are often highly specialised and may not be suitable for broader applications. In some cases, researchers rely on conjugate prior distributions, calculated based on the mean and standard deviation of expert knowledge. For example, Cagno et al. [16] employed AHP to determine prior distributions for gas pipeline failures, while De Persis et al. [21] constructed independent beta prior distribution for model parameters. A crucial challenge with this approach arises when expert knowledge does not align with the assumptions of the selected distribution, typically a conjugate prior. In such cases, experts are often forced to adjust their priors to fit the distribution, or alternatively, use non-informative priors, such as the uniform prior, which can dilute the influence of expert knowledge.

The established methods for prior elicitation are often resource-intensive and demand considerable time and cognitive effort from both experts and researchers. To overcome these limitations, this study introduces a practical and efficient alternative that integrates the AHP with statistical validation techniques to quantify expert knowledge into prior probability distributions. The main contribution lies in the modification of a predictive approach that is straightforward, resource-efficient, and time-saving. Expert knowledge on malaria transmission is collected via structured questionnaires, and AHP is used to generate normalised weights reflecting the relative importance of environmental factors. These weights are then treated as prior information and fitted to several candidate probability distributions. Goodness-of-fit tests are applied to identify the distribution that best represents the derived AHP weights. By combining decision-making analysis with statistical model validation, this approach ensures that the selected prior distribution accurately captures expert knowledge, offering a robust foundation for subjective Bayesian models in malaria research.

Methods

Research ethics and data collection

The authors adhered to comprehensive research ethics in their data collection methods, ensuring voluntary participation, obtaining informed consent from the participants, and maintaining strict confidentiality, privacy and respect for all participants. All the methods used in this study are fully in compliance with the relevant guidelines and regulations. These methods were reviewed and approved by the Turfloop Research Ethics Committee (TREC), a registered entity under the South African National Health Research Ethics Council of South Africa (Registration Number: REC-0310111-031). Malaria transmission is influenced by various climate and environmental factors such as temperature, normalised difference vegetation index (NDVI), rainfall, relative humidity, and forest cover [23,24]. In this study, data were collected using a purposive sampling approach, facilitated by a questionnaire. This method was deemed appropriate due to the large and uncertain population size combined with the limited resources of the researcher and the workforce. Expert elicitation, which plays a central role in this study, often faces challenges related to the availability of experts and their experience. To address this, experts were recruited both nationally and internationally.

A literature search on malaria models developed for South Africa was conducted using Google Scholar, PubMed, Scopus, and Web of Science. We screened the retrieved articles to determine whether the risk factors investigated aligned with the objectives of our study. This step was crucial because the responses of malaria modelling experts to the questionnaire were expected to be influenced by the findings of their own research. Consequently, only experts with experience in South African malaria modelling research were considered for participation in this study. If an article met our inclusion criteria, we recorded the contact information of its authors in a database of potential participants. The questionnaires were then distributed to all eligible experts, 46 in total, along with a brief summary explaining the elicitation process and the purpose of the questionnaire. Participants were also encouraged to ask questions and seek clarification from researchers whenever needed. The questionnaire was distributed in both manual and electronic formats, providing flexibility in how respondents could participate.

Although data collection through surveys and questionnaires is often hindered by low response rates, this study successfully obtained four expert responses, two from national experts and two from international experts. In Bayesian analysis, even a small number of well-informed experts can provide valuable information that improves the accuracy of the model [25]. Thus, the four responses are considered sufficient for the purposes of this study. The questionnaire focused on identifying the climate and environmental factors that lead to the increase in malaria cases. Experts were asked to provide information on variables such as average maximum and minimum daily temperatures, NDVI, and rainfall. The expert responses were subsequently converted into prior probability distributions, as detailed in the following sections.

Study design

In this study, pairwise comparisons are used to gather expert knowledge about factors that lead to increased cases of malaria. These comparisons are presented in the form of a questionnaire asking experts to assess the relative likelihood of observable quantities related to the response variable such as the likelihood of certain events, making this a predictive approach to prior elicitation. In this study a modified version of the traditional mathematical aggregation method is used. Rather than fitting probability distributions to each expert's individual judgments and then using a pooling rule to combine these distributions into one, the proposed method aggregates expert judgments using AHP. The aggregated knowledge is then used to calculate normalised weights, also referred to as priorities, which reflect the influence of each environmental factor on the incidence of malaria. After calculating these weights, they are fitted to various probability distributions. Statistical tests are then applied to evaluate how well each probability distribution fits the normalised weights. The best fitting distribution is ultimately selected as the prior distribution representing the collective expert knowledge.

The process of aggregating judgments and calculating normalised weights is achieved using the SpiceLogic AHP wizard-based software, while R software packages are used to fit and evaluate probability distributions.

The analytical hierarchy process (AHP)

The AHP is a general method used to derive ratio scores from discrete and continuous paired comparisons [26]. These comparisons reflect the relative strength, preferences, feelings, and knowledge of experts. The AHP was developed by Saaty in 1971 and has since become widely used for decision making and prioritisation [26]. For example, Saaty [26] used the AHP to select the best college. In this study, the goal is to identify the climate and environmental factors such as maximum and minimum average daily temperatures, NDVI, and rainfall that contribute to increased malaria cases. These factors are assessed based on their intensity levels: low, medium, and high. See Table 1 for details. In the context of AHP, these intensity levels can be termed the criteria, while climate and environmental factors are considered the alternatives.

Table 1 provides an overview of the variables of interest, their descriptions, dataset codes, and data types. Intensity levels represent the lower, middle and upper quartiles calculated using South African malaria data, which include environmental and climate factors of interest. The low intensity level corresponds to the lower quartile of maximum and minimum average daily temperatures, rainfall, and NDVI data. The medium intensity level represents the median values of these variables, while the high intensity level corresponds to the upper quartile. This classification helps to analyse the relationship between malaria incidence and varying climatic and environmental conditions. The structure of the AHP for this study is described in Fig. 1, where the first level

Table 1
Description of Variables.

Variable	Description	Data set code	Data type
Maximum temperatures	This variable represents the average temperatures during the day. It is classified into three intensity levels: low (at most 24 °C), medium (between 24 °C and 28.45 °C), and high (at least 28.45 °C).	MaxT	Numeric
Minimum temperatures	This variable captures the average temperatures during the night. It is classified into three intensity levels: low (at most 9.23 °C), medium (between 9.23 °C and 16.53 °C), and high (at least 16.53 °C).	MinT	Numeric
Normalised differenced vegetation index (NDVI)	It is a widely used metric for quantifying the health and density of vegetation using sensor data. It is classified into three intensity levels: low (at most 0.32), medium (between 0.32 and 0.49), and high (at least 0.49).	NDVI	Numeric
Rainfall	This variable measures the amount of precipitation in each period. It is classified into three intensity levels: low (at most 8.81 mm), medium (between 8.81 mm and 89.50 mm), and high (at least 89.50 mm).	Rain	Numeric

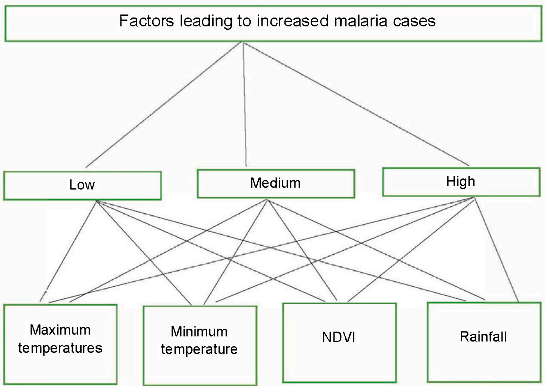

Fig. 1. AHP hierarchical structure for identifying climate and environmental factors influencing malaria cases.

represents the overall goal, the second level consists of the criteria (the intensity levels) and the third level represents the options or alternatives (the environmental factors).

The fundamental Saaty scale of the AHP is a numerical scale ranging from 1 to 9, which is used to facilitate pairwise comparisons. This scale allows experts to express their judgments about the relative importance of various factors. Unlike conventional measurements, the derived scale resembles probabilities, as it does not have a defined unit or an absolute zero. Table 2 presents the Saaty scale, illustrating how different values correspond to varying degrees of preference between the options being compared.

In this study, experts are asked to complete a questionnaire that facilitates pairwise comparisons between criteria (intensity levels) regarding specific options (risk factors). Each expert compares the options with one criterion at a time. The responses are then organised into matrices for AHP calculations. For instance, if we seek expert opinions on options related to a particular criterion, such as high intensity levels, the questions would be framed as follows: "Using the Saaty scale of 1 to 9, how much more or less does a high maximum average daily temperature contribute to increased malaria cases compared to heavy (high) rainfall?". Based on the findings of their own research, each expert responds to the questionnaire. Pairwise comparisons are conducted for each option in relation to the selected criterion, one at a time. An example of the resulting matrix for each expert is presented in Table 3.

The matrix in Table 3 represents pairwise comparisons of four climate and environmental factors (MaxT, MinT, NDVI, and Rain) at the high intensity level using the Saaty scale. Each element in the matrix indicates the relative importance of one factor compared to another based on expert judgment. In this example, the expert compares four factors, which means that they answered a total of six unique pairwise comparison questions:

- MaxT vs. MinT
- MaxT vs. NDVI
- MaxT vs. Rain
- MinT vs. NDVI
- MinT vs. Rain
- NDVI vs. Rain

Table 2
Saaty scale for pairwise comparisons in AHP.

Intensity of importance on an absolute scale	Definition	Description
1	Equal importance.	Two activities contribute equally to the objective.
3	Moderate importance of one over the other.	Experience and judgment moderately favour one activity over the other.
5	Strong importance.	Experience and judgment strongly favour one activity over the other.
7	Very strong importance.	An activity is strongly favoured, and its dominance demonstrated in practice.
9	Extreme importance.	The evidence favouring one activity over another is of highest possible order of affirmation.
2,4,6,8	Intermediate values between the two adjacent judgments.	Used when compromise is need.

Table 3
Pairwise comparisons matrix.

High intensity	MaxT	MinT	NDVI	Rain
MaxT	1	0.5	0.33	3
MinT	2	1	0.5	3
NDVI	3	2	1	4
Rain	0.33	0.33	0.25	1

The matrix should be reciprocal, meaning that if a_{ij} represents the comparison of factor i with factor j, then $a_{ji} = \frac{1}{a_{ij}}$. This ensures consistency in the pairwise comparisons. The diagonal elements are all equal to 1, since each factor is equally important compared to itself. A value greater than 1 (for example, $MinT/MaxT = 2$) indicates that the row factor is considered more important than the column factor. In this case, MinT is judged to be twice as important as MaxT. A value less than 1 (for example, $MaxT/MinT = 0.5$) means that the row factor is considered less important than the column factor. Here, MaxT is half as important as MinT. The NDVI factor has high relative importance, as it is rated 3 times more important than MaxT and twice as important as MinT. Rainfall is the least important factor in this comparison, as it has lower values relative to other factors (e.g., Rain/NDVI = 0.25, which means NDVI is 4 times more important than Rain). These comparisons help to derive priority weights, which are essential in decision making for malaria risk assessment. The next step involves aggregating the pairwise comparison matrices from all experts by calculating the weighted geometric mean for the pairwise comparisons of each expert. This approach ensures that the aggregated matrix reflects the collective judgments while preserving consistency. For example, if multiple experts provided different comparisons for MinT vs. MaxT, their individual values are combined using the geometric mean:

$$\tilde{a}_{ij} = \left(\prod_{k=1}^{n} a_{ij}^{w_k} \right)^{\frac{1}{\sum w_k}},$$

where:

- $\tilde{a}_{ij}$ is the aggregated pairwise comparison for criterion i vs. j,
- a_{ij}^{k} is the pairwise comparison value given by expert k,
- w_k is the weight assigned to expert k,
- n is the number of experts.

In this study, all experts are assigned equal weights. In addition to the geometric mean, the arithmetic mean can also be used. However, it tends to lead to similar results. Once all pairwise comparison matrices have been aggregated, the principal eigenvector method is used to determine the final option weights. The principal eigenvector method is a key component of the AHP, used to determine the relative priority vector of each criterion or option in a pairwise comparison matrix. A priority vector is a numerical ranking of the criteria and options that indicates the order of preference among them. This order should reflect intensity or cardinal preference, indicated by the ratio of numerical values, and thus is unique up to a multiplicative constant. The necessary conditions that the priority vector must satisfy are as follows:

i. It should belong to a ratio scale (that is, it should remain invariant under multiplication by a constant c).
ii. It must be invariant under hierarchical composition for its judgment matrix, ensuring that new priority vectors do not continuously arise from the same matrix.

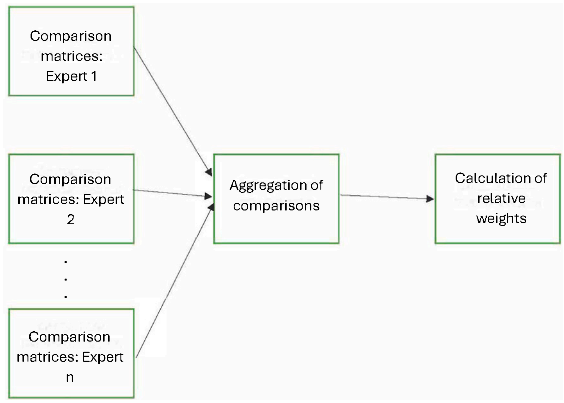

Fig. 2. Aggregation process for pairwise comparison matrices in AHP.

Consider a pairwise comparison matrix $\mathbf{A}$ for n criteria or options. Let $\vec{\boldsymbol{w}}$ be a priority vector, where each element a_{ij} in $\mathbf{A}$ represents the relative importance or preference of the criterion i over the criterion j. The priority vector $\vec{\boldsymbol{w}}$ must satisfy the relation:

$$\mathbf{A}\vec{\boldsymbol{w}} = c\vec{\boldsymbol{w}}. \tag{1}$$

Due to the need for invariance to produce a unique priority vector, $\vec{\boldsymbol{w}}$ must be the reciprocal eigenvector of $\mathbf{A}$ and c is its corresponding principal eigenvalue. This result is grounded in the following theorem:

Theorem 1 (*Perron–Frobenius Theorem*)**.** *For a given positive matrix* $\mathbf{A}$*, the only positive vector* $\vec{\boldsymbol{w}}$ *and the only positive constant c that satisfies* $\mathbf{A}\vec{\boldsymbol{w}} = c\vec{\boldsymbol{w}}$*, is a vector* $\vec{\boldsymbol{w}}$ *that is a positive multiple of the Perron vector (principal eigenvector) of* $\mathbf{A}$*, and the only such c is the Perron value (principal eigenvalue) of* $\mathbf{A}$.

The matrix $\mathbf{A}$ is typically reciprocal, meaning that:

$$a_{ij} = \frac{1}{a_{ji}} \text{and } a_{ii} = 1 \forall i.$$

The comparison matrix $\mathbf{A}$ is normalised so that the sum of its components is equal to 1. This normalisation is achieved by dividing each element a_{ij} of the matrix $\mathbf{A}$, by the total sum of all elements in its respective column:

$$a'_{ij} = \frac{a_i j}{\sum_{i=1}^{n} a_{ij}},$$

where a'_{ij} represents the normalised priorities for criterion i over criterion j and $\sum_{i=1}^{n} a_{ij}$ is the total sum of all elements in the jth column of $\mathbf{A}$. Normalising the matrix $\mathbf{A}$ ensures that they form valid proportions that sum to 1, while also eliminating reciprocity in the normalised matrix. The eigenvector method also assesses the consistency of the judgments within the pairwise comparison matrix. A consistency check is performed, using the consistency index (CI), defined as:

$$CI = \frac{c - n}{n - 1},$$

and the average random consistency index (RI) for a matrix of size n, defined as:

$$RI(n) = \frac{1}{N} \sum_{k=1}^{N} CI_k$$

where:

- N is the total number of generated random matrices.
- CI_k is the consistency index of the kth random matrix.

The consistency ratio (CR) is calculated as:

$$CR = \frac{CI}{RI}. \tag{2}$$

If the CR is below the threshold (usually 0.1), the judgments are considered reasonably consistent. In summary, Fig. 2 illustrates the steps involved in deriving the AHP weights. Similarly to the conclusion made by Saaty [26] when AHP was used to select the best college, the option with the highest priority value was chosen as the best. Likewise, in this study, the option (refer to Fig. 1 for options) with the highest priority value, is the one that contributes the most to increased malaria cases.

Fitting various probability distributions to the AHP weights

The purpose of this analysis is to identify the probability distribution that best fits the AHP weights, which will serve as the prior probability distribution for malaria data. To determine the most appropriate distribution, we fit various relevant probability distributions to the AHP weights and employ goodness-of-fit tests to assess which distribution aligns best with the AHP weights. The relevant distributions to consider include:

i. Beta distribution, which is suitable for modelling data in the interval [0,1] and can accommodate the variability within the AHP weights. The probability density function of beta distribution is given by:

$$f(x;\alpha,\beta)=\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, \tag{3}$$

where α and β are shape parameters.

ii. Dirichlet distribution, a multivariate generalisation of a beta distribution, is ideal for modelling proportions that sum up to 1. The Dirichlet distribution function can be expressed as:

$$f(x_1,x_2,\ldots,x_n;\alpha_n)=,\frac{\Gamma(\sum_{i=1}^{n}\alpha_i)}{\prod_{i}^{n}\Gamma(\alpha_i)}\prod_{i=1}^{n}(x_i)^{\alpha_i-1}, \tag{4}$$

where α_i's are the concentration parameters, and Γ is the gamma function.

iii. Gamma distribution, which is typically used for positive continuous data and can be used if the AHP weights are scaled appropriately. The gamma distribution function is given by:

$$f(x;\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}, \tag{5}$$

where α is the shape and β is the rate parameters.

iv. Although the normal distribution may not be the best fit for AHP weights constrained within the interval [0, 1], it is defined throughout the real line. This characteristic allows it to theoretically accommodate any dataset, including AHP weights. As a result, it is commonly employed in the literature as a prior distribution. A large standard deviation is used to create a flat (weakly informative) prior, indicating a lack of strong assumptions about the parameter's value. In contrast, a smaller standard deviation suggests high confidence that the parameter is close to the mean, representing strong prior information. The normal density function is given by:

$$f(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}},-\infty<x<\infty,\sigma>0, \tag{6}$$

where μ and σ represents the mean and standard deviation of the distribution.

v. The uniform distribution can be considered for its simplicity, especially when it is reasonable to assume that all values in the range [0,1] are equally likely. The probability density function for a uniform distribution is expressed as:

$$f(x;a,b)=\frac{1}{b-a},a\le x\le b, \tag{7}$$

where a and b presents the lower and upper bounds of the distribution.

Model evaluation and validation

Instead of relying on model selection criteria such as AIC or DIC, which prioritise likelihood and model complexity, the focus is on identifying a distribution that accurately represents the observed AHP weights. Therefore, for each fitted distribution, the Kolmogorov–Smirnov (KS) test is performed to evaluate the goodness of fit. The KS test compares the empirical cumulative distribution function (ECDF), $F_n(x)$, of the AHP weights to the theoretical cumulative distribution function (CDF), $F(x;\hat{\theta})$, of the fitted distribution. The KS test statistic is given by:

$$D_n=\sup_x\left|F_n(x)-F(x;\hat{\theta})\right|, \tag{8}$$

where D_n is the maximum absolute difference between the empirical and theoretical CDFs. Distributions are assessed based on their Kolmogorov–Smirnov (KS) test statistic and the corresponding $p-$ values. The distribution with the smallest statistic of the KS test and the highest $p-$ value is considered the best fit for the AHP weights. The KS test is particularly relevant and useful for many continuous distributions, including the beta and gamma distributions. However, for distributions such as the uniform distribution, alternative tests may be more appropriate. To provide a more comprehensive evaluation, the Cramér–von Mises (CvM) test is also used as a complementary goodness-of-fit test. The CvM test treats all parts of the distribution equally and is based on the squared differences between the empirical cumulative distribution function (ECDF) of the observed data and the theoretical cumulative distribution function (CDF) of the fitted distribution. The CvM test statistic, denoted as W^2, is given by:

$$W^2=\frac{1}{12n}+\sum_{i=1}^{n}\left(F(x_i)-\frac{2i-1}{2n}\right)^2 \tag{9}$$

Table 4
Expert variable priorities by intensity levels.

Option	Expert 1			Expert 2			Expert 3			Expert 4		
	Low	Medium	High	Low	Medium	High	Low	Medium	High	Low	Medium	High
MaxT	0.270	0.028	0.055	0.034	0.077	0.237	0.021	0.121	0.199	0.061	0.069	0.050
MinT	0.123	0.039	0.025	0.024	0.028	0.169	0.052	0.044	0.102	0.030	0.135	0.086
NDVI	0.192	0.029	0.039	0.098	0.039	0.035	0.063	0.062	0.048	0.021	0.097	0.145
Rain	0.071	0.115	0.014	0.155	0.054	0.050	0.062	0.086	0.141	0.086	0.190	0.030

Table 5
Expert AHP Weights.

Option	Expert 1	Expert 2	Expert 3	Expert 4
MaxT	0.353	0.348	0.341	0.180
MinT	0.187	0.220	0.197	0.251
NDVI	0.260	0.173	0.173	0.263
Rain	0.200	0.259	0.289	0.306

where n is the sample size and $F(x_i)$ is the CDF of the fitted distribution at data point x_i. The large value of W^2 indicates a poor fit. Furthermore, random samples are generated from fitted distributions and their behaviour is compared to empirical AHP weights using density plots and box plots. This approach helps validate whether fitted distributions effectively capture the characteristics of the empirical data. If the simulated data closely resemble the empirical data, it supports the adequacy of the selected distributions.

Bayesian model

A Bayesian statistical model consists of a parametric statistical model $f(y|\theta)$, which describes the sample data, and a prior distribution $\pi(\theta)$, which represents prior knowledge. Bayes' theorem updates the information on θ by incorporating the observed data y [27]. Inference in Bayesian models is based on the posterior distribution $\pi(\theta|y)$, which represents the updated belief about θ given y. It is defined as:

$$\pi(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\theta)\pi(\theta)\,d\theta}, \tag{10}$$

where the denominator is the marginal likelihood, acting as a normalising constant. Since this denominator is independent of θ, the posterior distribution is often written as:

$$\pi(\theta|y) \propto f(y|\theta)\pi(\theta). \tag{11}$$

This study utilises **SpiceLogic AHP software** to aggregate expert matrices and compute AHP weights, while **R software** is used to fit and evaluate the relevant probability distributions to the AHP weights.

Results and discussion

Experts individual knowledge

The pairwise comparison matrices for criteria (intensity climate conditions: low, medium, and high) and options (climate factors: MaxT, MinT, NDVI, and Rain) for each expert response have been specified with the relative priorities for each factor (criteria and options) calculated. However, the matrices are not included in the results presented in this study. Relative priorities, which represent the importance of options in different criteria; and final option priorities (AHP weights), which represent the overall importance of options, are presented in Tables 4 and 5.

Table 4 presents the variable priorities based on three intensity levels: Low, Medium, and High. The options considered include maximum temperature (MaxT), minimum temperature (MinT), normalised difference vegetation index (NDVI), and rainfall (rain). The priority values were calculated on the basis of expert judgment using the AHP method discussed in Section "Bayesian model". In Table 4, we observe that different priorities are calculated for the options based on expert domain knowledge and experience. For example, Expert 1 places the highest priority on MaxT in the low category, while Expert 2 assigns the highest priority to MaxT in the high category. Similarly, Expert 4 assigns the highest priority to Rain in the Medium category. These variations highlight the subjective nature of expert judgment in determining the relative importance of environmental factors.

Table 5 presents the Analytic Hierarchy Process (AHP) weights calculated for four options: Maximum Temperature (MaxT), Minimum Temperature (MinT), Normalised Difference Vegetation Index (NDVI), and Rainfall (Rain) based on each expert comparison matrices. These weights reflect the relative importance of each variable as perceived by the experts, irrespective of intensity levels of climate and environmental factors. The table reveals that Experts 1 and 2 attribute the highest importance to MaxT, with respective weights of 0.353 and 0.348, indicating a consensus that temperature is a key factor. Expert 3, on the other hand, assigns the highest weight to Rain (0.289), highlighting a different perspective where precipitation plays a more critical role.

Table 6
Aggregated variable priorities by intensity levels.

Option	Low	Medium	High
MaxT	0.076	0.076	0.134
MinT	0.058	0.059	0.097
NDVI	0.088	0.059	0.069
Rain	0.111	0.121	0.052

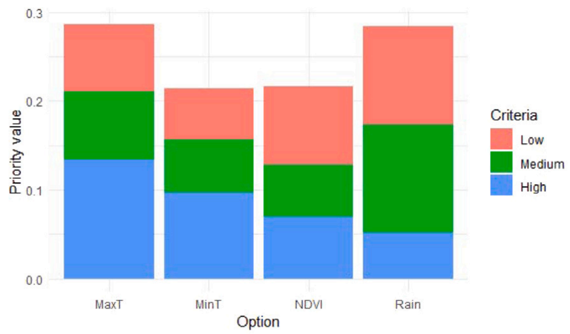

Fig. 3. Summary of contributions of climate factors to malaria transmission across different intensity levels, aggregated perspectives of all experts.

Table 7
Aggregated AHP weights.

Option	Value
MaxT	0.287
MinT	0.214
NDVI	0.215
Rain	0.283

Expert 4 places the most emphasis on Rain (0.306) and NDVI (0.263), suggesting a stronger consideration of vegetation and precipitation in their assessment. The variations in these weights suggest differing expert opinions, possibly influenced by their backgrounds and experiences. Understanding these perspectives allows for a more comprehensive assessment of environmental factors in decision-making and model development.

Aggregation of experts knowledge

The pairwise comparisons of each expert with respect to the climate intensity levels (low, medium, and high) are aggregated as weighted geometric means. Then, once all pairwise comparison matrices are generated by aggregation, the regular principal eigenvector method is used to calculate the final option priorities (AHP weights). Each cell of the final aggregated matrix is generated by the geometric mean of the corresponding cells of each member's comparison matrix. All judgments of experts were given equal priority.

Table 6 and Fig. 3 present the aggregated priority attributes, showing the overall importance of the different options (MaxT, MinT, NDVI and Rain) in the three criteria, low, medium, and high. The average temperature during the day (MaxT) emerges as the most influential factor in high intensity conditions with a priority value of 0.134, while rainfall is the most impactful factor in climate conditions of low (0.111) and medium (0.121) intensity. NDVI has the second highest importance under low-intensity conditions, but its influence decreases noticeably across the medium and high intensity levels. The average temperature during the night (MinT) ranks as the least influential factor under the low intensity level and the second largest under the high intensity level.

Table 7 and Fig. 4 present the aggregated AHP weights, highlighting the overall importance of each option (MaxT, MinT, NDVI, and Rain). The average temperature during the day (0.287) and rainfall (0.283) are almost equally prioritised, indicating that they are the most influential factors overall. Although NDVI (0.215) and the average temperature during the night (0.214) have lower AHP weights, they still contribute to malaria transmission, serving as secondary contributors compared to daytime temperature and rainfall.

Fitted models evaluation and validation

Table 8 presents the Kolmogorov–Smirnov goodness-of-fit statistics for the distributions fitted to the final AHP priorities. This table includes the estimated parameters of the candidate distributions, along with the p-values and test statistics (D) from the KS test. The gamma distribution provides the best fit, with the highest p-value (0.771) and the second-lowest test statistic value (0.298),

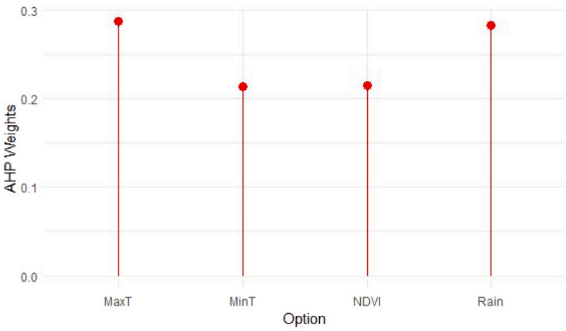

Fig. 4. Aggregated importance of climate factors in malaria transmission.

Table 8
Kolmogorov–Smirnov goodness-of-fit tests.

Probability Distribution	Estimated parameters	P-values	Test statistic (D)
$Beta(\alpha, \beta)$	(27.948; 83.957)	0.300	0.772
$Dirichlet(\alpha_1, \ldots, \alpha_n)$	(0.246; 0.265; 0.220; 0.269)	0.581	0.110
$Gamma(\alpha, \beta)$	(37.585; 150)	0.771	0.298
$Normal(\mu, \sigma)$	(0.250; 0.041)	0.761	0.303
$Uniform(a, b)$	(0.214; 0.287)	0.212	0.486

Table 9
Cramér-von Mises goodness-of-fit tests.

Probability distribution	Estimated parameters	P-values	Omega 2
$Beta(\alpha, \beta)$	(27.948; 83.957)	0.673	0.088
$Dirichlet(\alpha_1, \ldots, \alpha_n)$	(0.246; 0.265; 0.220; 0.269)	N/A	N/A
$Gamma(\alpha, \beta)$	(37.585; 150)	0.668	0.089
$Normal(\mu, \sigma)$	(0.250; 0.041)	0.672	0.088
$Uniform(a, b)$	(0.214; 0.287)	0.148	0.285

making it a strong candidate for modelling the data set. The normal and Dirichlet distributions also fit moderately well, with the p-values of 0.761 and 0.581 respectively. Overall, the gamma and normal distributions appear to be the most appropriate models for the data, whereas the beta and uniform distributions seem to be the least suitable.

Table 9 shows the Cramér–von Mises goodness-of-fit statistics for the distributions fitted to the final AHP priorities presented in Table 7. The table includes the estimated distribution parameters, the Cramér–von Mises test p-values, and the corresponding test statistics (D). The beta distribution shows the best fit closely followed by the normal distribution, with p-values of 0.673 and 0,672, respectively, and the smallest omega value (0.088), indicating a close alignment with the data. The gamma distribution also fits well, with a p-value of 0.668 and an omega value of 0.089. In comparison, the uniform distribution provides a weaker fit. There are no results for the Dirichlet distribution, likely because it lacks a cumulative distribution function in the traditional sense, or there may be insufficient data points to estimate its fit properly. Therefore, the Cramér–von Mises test may not be feasible in this case. However, a Monte Carlo simulation was used to generate random samples from the fitted distributions, using the estimated parameters to further validate the distributions. The results are presented using overlay density plots in Fig. 5 and box plots in Fig. 6, comparing the behaviour of the simulated data with the empirical data.

Fig. 5 shows the fitted distributions overlaid on the empirical density of the AHP weights, represented by the black curve. Among the fitted distributions, the beta (blue), gamma (green), and normal (light blue) distributions align most closely with the empirical density, particularly around the central peak and tails. However, none of these distributions perfectly captures the empirical distribution. In contrast, the uniform distribution (red) displays a sharp, narrow peak that does not resemble the empirical density. The Dirichlet distribution (purple) also diverges remarkably, failing to approximate the observed distribution of AHP weights.

The side-by-side box plots shown in Fig. 6 illustrate the spread and central tendencies of the AHP weights for each fitted distribution. Among the fitted distributions, the beta (blue), gamma (green), and normal (light blue) distributions effectively model the empirical AHP weights, with medians that closely align with the empirical values and reasonable ranges. In contrast, the Dirichlet (purple) distribution exhibits excessive variability, while the uniform (red) distribution fails to accurately capture the spread compared to the beta and gamma distributions. For a Bayesian model developed using any of the priors, in this case, the $\text{Beta}(27.95, 83.96)$ prior, given by:

$$\pi(\theta) = \frac{\Gamma(27.95 + 83.96)}{\Gamma(27.95)\Gamma(83.96)} \theta^{27.95-1}(1-\theta)^{83.96-1}, \quad 0 < \theta < 1, \tag{12}$$

where, θ represents the probability parameter, which is the unknown quantity being estimated.

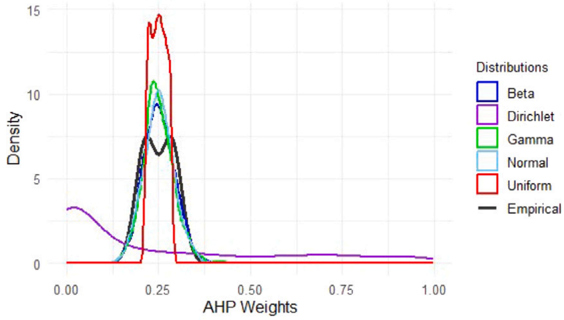

Fig. 5. Overlay density of fitted distributions and the empirical density of AHP weights.

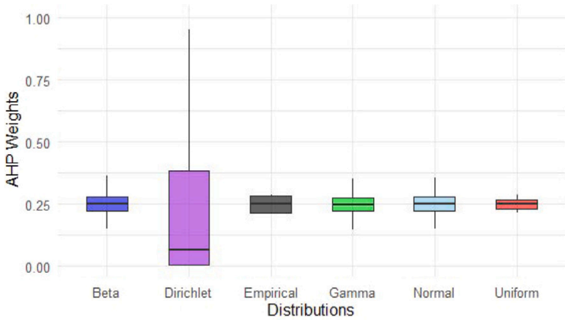

Fig. 6. Box plots for fitted distributions and the empirical AHP weights.

Consider that malaria counts Y follow a negative binomial distribution. The probability mass function of Y is given by:

$$P(Y|r,\theta) = \frac{\Gamma(Y+r)}{\Gamma(r)\Gamma(Y+1)}\theta^{r}(1-\theta)^{Y}. \tag{13}$$

where, θ represents the probability parameter, while r is the dispersion (or shape) parameter, which controls overdispersion in the count data. Combining the prior distribution in Eq. (12) and the likelihood function of Eq. (13), the posterior distribution is given by:

$$\pi(\theta|Y) \propto \theta^{27.95-1}(1-\theta)^{83.96-1}\prod_{i=1}^{n}\theta^{r_i}(1-\theta)^{Y_i}. \tag{14}$$

Thus, the posterior distribution derived in Eq. (14) defines the Bayesian Negative Binomial model, where inference is based on the posterior distribution of θ. This approach integrates prior knowledge through the Beta prior while updating beliefs about θ using observed malaria count data. The Bayesian framework provides a probabilistic interpretation of uncertainty, improving the robustness of the model for the analysis of malaria transmission. Furthermore, any other prior (as elicited in this study) can be used in Eq. (14) instead of the Beta prior, depending on the specific modelling requirements.

Conclusion and recommendations

Conclusion

This study presents a practical and resource-efficient approach for incorporating expert knowledge into Bayesian models of malaria transmission through a predictive elicitation framework. By using structured questionnaires and the AHP, expert judgments were systematically converted into quantitative weights reflecting the perceived importance of environmental and climatic factors. These AHP-derived weights were then statistically validated by fitting them to several candidate probability distributions. Goodness-of-fit analyses identified the beta, gamma, and normal distributions as the best representations of expert knowledge.

The findings confirm that average daily temperature and rainfall are perceived by experts as the most influential drivers of malaria transmission, followed by NDVI and minimum temperature. These results are consistent with previous studies by Kleinschmidt et al. [28] and Abiodun et al. [29], while also revealing contrasts with the work of Sehlabana et al. [30], who identified minimum temperature as the primary driver. These differences highlight the need for further region-specific research to understand the interplay of environmental factors.

A key innovation of this study lies in simplifying the prior elicitation process, overcoming logistical and cognitive challenges typically associated with structural methods. The integration of AHP with statistical validation not only strengthens the reliability of elicited priors but also aligns them with distributions commonly used in Bayesian modelling. For example, the gamma distribution is a conjugate prior for the Poisson model and the beta distribution for the negative binomial model, as noted by Robert et al. [27]. Normal priors are also widely applicable in various Bayesian settings.

In contrast to standard practices that rely on objective priors derived solely from data, this study emphasises the value of subjective priors informed by expert insights. This approach enhances model interpretability and contextual relevance. Overall, the methodology contributes a scalable and adaptable framework for improving the use of expert knowledge in statistical epidemiology and offers a foundation for broader application in public health modelling.

Recommendations

We recommend that researchers with limited resources consider adopting this approach as an alternative to the traditional structural method of prior elicitation. The method developed in this study is flexible and can be applied in various settings, such as Bayesian generalised linear models (GLMs) and Bayesian statistical studies in epidemiology. Additionally, this approach enables the exploration of a wide range of potential prior probability distributions, ensuring that the distributions that best capture expert knowledge are identified.

Limitations of the study and future research

This study was limited to a small number of experts who responded to the questionnaire. Although a larger sample could improve diversity and reduce individual biases, Bayesian methods prioritise the quality of expert judgment, as smaller but more knowledgeable groups of experts can provide more precise and reliable knowledge essential for defining informative priors distributions. While a sample size of four is considered sufficient for this study, future research may benefit from a larger expert panel to improve robustness and capture a wider range of perspectives.

In future research, the prior probability distributions identified in this study, which best capture the knowledge of experts, will be applied to the malaria modelling. These expert-based priors will also be compared with several objective prior probability distributions to evaluate their effect on model performance. In addition, more research is needed to explore the sensitivity of various prior elicitation methods to different sample sizes.

CRediT authorship contribution statement

Makwelantle Asnath Sehlabana: Collected the dataset, Developed the Methodology, Conducted the data analysis, Captured the results. **Daniel Maposa:** Collected the dataset, Structured the manuscript, Interpreted the results, Edited the article, Supervised the work carried out by Makwelantle Asnath Sehlabana. **Alexander Boateng:** Supervised the work carried out by Makwelantle Asnath Sehlabana. **Sonali Das:** Supervised the work carried out by Makwelantle Asnath Sehlabana.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

All authors reviewed and approved the final manuscript.

Data availability

The datasets generated and analysed during this study are not publicly available due to the consent agreement associated with data collection. This agreement specifies that only anonymised extracts of the dataset can be shared. However, the data sets can be obtained from the corresponding author upon reasonable request.

References

[1] L. Held, Bayesian methods in epidemiology, in: Handbook of Epidemiology, Springer, 2024, pp. 1–35.
[2] M. Goldstein, Subjective Bayesian analysis: principles and practice, Bayesian Anal. (2006).
[3] R. van de Schoot, et al., Bayesian statistics and modelling, Nat. Rev. Methods Prim. 1 (2021) 1.
[4] R.L. Winkler, The assessment of prior distributions in Bayesian analysis, J. Amer. Statist. Assoc. 62 (1967) 776–800.
[5] R. Denham, K. Mengersen, Geographically assisted elicitation of expert opinion for regression models, in: Bayesian Statistics 8: Proceedings of the Eighth Valencia International Meeting, 2007.
[6] J.R. Falconer, E. Frank, D.L. Polaschek, C. Joshi, Methods for eliciting informative prior distributions: A critical review, Decis. Anal. 19 (2022) 189–204.
[7] P. Mikkola, et al., Prior knowledge elicitation: The past, present, and future, 2021, arXiv preprint arXiv:2112.01380.
[8] R. van de Schoot, et al., Bayesian ptsd-trajectory analysis with informed priors based on a systematic literature search and expert elicitation, Multivar. Behav. Res. 53 (2018) 267–291.
[9] F. Bockting, S.T. Radev, P.-C. Bürkner, Simulation-based prior knowledge elicitation for parametric Bayesian models, Sci. Rep. 14 (2024) 17330.
[10] P. Mikkola, et al., Prior knowledge elicitation: The past, present, and future, Bayesian Anal. (2023) 1–33.
[11] J.B. Kadane, Predictive and structural methods for eliciting prior distributions, Bayesian Anal. Econ. Stat. 18 (1980).
[12] O'Leary R.A., et al., Comparison of three expert elicitation methods for logistic regression on predicting the presence of the threatened brush-tailed rock-wallaby petrogale penicillata, Environmetrics Off. J. Int. Environmetrics Soc. 20 (2009) 379–398.
[13] N. Dallow, N. Best, T.H. Montague, Better decision making in drug development through adoption of formal prior elicitation, Pharm. Stat. 17 (2018) 301–316.
[14] A. O'Hagan, Expert knowledge elicitation: subjective but scientific, Amer. Statist. 73 (2019) 69–81.
[15] R. Schwarzenegger, J. Quigley, L. Walls, Is eliciting dependency worth the effort? a study for the multivariate Poisson-gamma probability model, Proc. Inst. Mech. Eng. Part O J. Risk Reliab. 237 (2023) 858–867.
[16] E. Cagno, F. Caron, M. Mancini, F. Ruggeri, Using ahp in determining the prior distributions on gas pipeline failures in a robust Bayesian approach, Reliab. Eng. Syst. Saf. 67 (2000) 275–284.
[17] P.H. Garthwaite, S.A. Al-Awadhi, F.G. Elfadaly, D.J. Jenkinson, Prior distribution elicitation for generalized linear and piecewise-linear models, J. Appl. Stat. 40 (2013) 59–75.
[18] J.O.and others Jansen, Elicitation of prior probability distributions for a proposed Bayesian randomized clinical trial of whole blood for trauma resuscitation, Transfusion 60 (2020) 498–506.
[19] D. Azzolina, P. Berchialla, D. Gregori, I. Baldi, Prior elicitation for use in clinical trial design and analysis: a literature review, Int. J. Environ. Res. Public Heal. 18 (2021) 1833.
[20] J. Lan, et al., Remote, real-time expert elicitation to determine the prior probability distribution for Bayesian sample size determination in international randomised controlled trials: Bronchiolitis in infants placebo versus epinephrine and dexamethasone (biped) study, Trials 23 (2022) 279.
[21] C. De Persis, J.L. Bosque, I. Huertas, S.P. Wilson, Quantitative system risk assessment from incomplete data with belief networks and pairwise comparison elicitation, 2019, arXiv preprint arXiv:1904.03012.
[22] Z. Syed, O. Shabarchin, Y. Lawryshyn, A novel tool for Bayesian reliability analysis using ahp as a framework for prior elicitation, J. Loss Prev. Process. Ind. 64 (2020) 104024.
[23] S. Bhatt, et al., The effect of malaria control on plasmodium falciparum in africa between 2000 and 2015, Nature 526 (2015) 207–211.
[24] Brugueras S., et al., Environmental drivers, climate change and emergent diseases transmitted by mosquitoes and their vectors in Southern Europe: A systematic review, Environ. Res. 191 (2020) 110038.
[25] F. Sadia, S.S. Hossain, Contrast of Bayesian and classical sample size determination, J. Mod. Appl. Stat. Methods 13 (2014) 23.
[26] R.W. Saaty, The analytic hierarchy process—what it is and how it is used, Math. Model. 9 (1987) 161–176.
[27] C.P. Robert, et al., The Bayesian Choice: from Decision-Theoretic Foundations to Computational Implementation, vol. 2, Springer, 2007.
[28] I. Kleinschmidt, B. Sharp, G. Clarke, B. Curtis, C. Fraser, Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in Kwazulu Natal, South Africa, Am. J. Epidemiol. 153 (2001) 1213–1221.
[29] G.J. Abiodun, et al., A dynamical and zero-inflated negative binomial regression modelling of malaria incidence in Limpopo Province, South Africa, Int. J. Environ. Res. Public Heal. 16 (2019) 2000.
[30] M.A. Sehlabana, D. Maposa, A. Boateng, Modelling malaria incidence in the Limpopo Province, South Africa: comparison of classical and Bayesian methods of estimation, Int. J. Environ. Res. Public Heal. 17 (2020) 5016.