On the Estimation of the Size of a Target Population

Nicolas Valeyrie[†], Jérémie Houssineau[‡], Christopher Strode[†], Yan Pailhas[†]
[†] Centre for Maritime Research and Experimentation
Science and Technology Organization, NATO
19126 La Spezia, Italy
[‡]Division of Mathematical Sciences
Nanyang Technological University
Singapore 637371

***Abstract*—This paper is concerned with the Bayesian estimation with possibility theory of the number of targets in an area of interest when no prior information is available. The estimation relies solely on the number of detections that is reported every time the area is surveyed. It is assumed that detections can be true and false positives. Possibility theory is capable of quantifying the complete absence of prior information on the number of targets with the introduction of a prior possibility function that is equal to one everywhere on the set of natural numbers. This has no equivalent in probability theory. The prior possibility function is turned into a posterior possibility function from which a point estimate of the number of targets can be derived. Besides the point estimate, a meaningful notion of confidence interval can be derived from the posterior possibility and necessity associated with subsets of the natural numbers.**

I. Introduction

This paper is concerned with the Bayesian estimation of the number of targets in an area of interest when no prior information is available. The emphasis is on the quantification of the absence of information about the number of targets and its implications on the estimation process.

The scenario we envision is as follows. The area of interest is surveyed by a sensor by which we mean a sensing device, an imaging sonar for instance, and a detector, acting on sensor data and tuned to a certain type of target. We assume that the area is much larger than the sensor's field of view so it makes sense to speak of a survey of the area when the sensor is carried on a vehicle. We assume that the area contains a fixed number of targets if any at all. Thus targets appear in and disappear from the sensor's field of view during a survey but never from the area itself. A number of detections is reported after every survey, the detections being a collection of true and false positives depending on whether the detections originate from the targets. We assume that a true positive originates from a single target. In its most general form, the question is what can be said about the number of targets in the area from the number of detections when no prior information about the number of targets is available?

Possibility theory is capable of quantifying the complete absence of prior information on the number of targets via a particular representation that satisfies most criteria regarding an uninformative prior [1, 2]. In its most modern form, possibility theory has been applied successfully to sensor data fusion for Space Situational Awareness [3] and multi-target tracking [4] with an emphasis on the connections between possibility theory and probability theory besides matters strictly related to sensor data fusion and multi-target tracking. With probability theory, it is notoriously difficult to properly define a weakly informative prior. Paradoxes arise indeed when one uses improper prior probability distributions [5]. In particular, a prior probability distribution that is constant on a subset of the real line so as not to favour a value more than another can be argued to be informative since a simple change of variable may yield a prior probability distribution that is no longer constant [6]. Even when considering a probability distribution on a countable set such as the natural numbers, there are consequences for choosing a uniform prior probability distribution, as we will show in our numerical experiments. The problem we are concerned with has been addressed in [7] with probability theory and a prior on the number of targets that is constant up to a predetermined maximum number. Here, we generalise this approach to random true and false positive rates. This will be our baseline when we evaluate the performance of our estimation procedure based on possibility theory.

In Section II, we present elements of possibility theory. In Section III, we present the Bayesian inference on the number of targets with possibility theory. In particular, we present the point estimation of the number of targets along with the construction of confidence intervals on the number of targets. In Section IV, we present the Bayesian inference on the number of targets with probability theory. In Section V, we compare in simulation the performance of the point estimator derived from possibility theory to the performance of the one derived from probability theory.

II. Elements of possibility theory

A. Possibility functions, possibility and necessity of events

We present here elements of possibility theory. Let X (with realisations x) be an uncertain variable on a certain space $\mathbf{X}$. Uncertain variables are to possibility theory what random variables are to probability theory. The information we hold about X is quantified by a possibility function $F_X : \mathbf{X} \to [0,1]$ which is a function such that $\max_{x\in\mathbf{X}} F_X(x) = 1$. Events are subsets of $\mathbf{X}$ and the credibility about an event is quantified by two quantities: the possibility and the necessity of the event.

If $U \subset \mathbf{X}$ is a subset of $\mathbf{X}$ and U^c its complement in $\mathbf{X}$, the possibility $\pi_X(U)$ and the necessity $\nu_X(U)$ of the event U are

$$\begin{cases} \pi_X(U) = \max_{x \in U} F_X(x) \\ \nu_X(U) = 1 - \max_{x \in U^c} F_X(x). \end{cases} \tag{1}$$

The possibility and the necessity of U are real numbers in the interval $[0, 1]$ and may be interpreted in the following way [1, 2]:

- The closer the possibility $\pi_X(U)$ is to one the more possible U is. We would not be surprised if U happened.
- The closer the necessity $\nu_X(U)$ is to zero the less necessary U is. We would be surprised if U happened.
- The closer the possibility $\pi_X(U)$ is to zero the less possible U is. We would not be surprised if U did not happen.
- The closer the necessity $\nu_X(U)$ is to one the more necessary U is. We would be surprised if U did not happen.

B. Complete knowledge or absence of knowledge

Consider the extreme case where $F_X(x) = 1$ for every $x \in \mathbf{X}$. For any event $U \subset \mathbf{X}$, we have $\pi_X(U) = 1$ so the event is possible to the highest degree and we would not be surprised if the event happened. We also have $\nu_X(U) = 0$ so the event is unnecessary to the highest degree and we would not be surprised if the event did not happen. The possibility function F_X quantifies therefore the complete absence of knowledge about X.

Consider the other extreme case where $F_X(x) = 1$ if $x = x_0$ and $F_X(x) = 0$ if $x \neq x_0$ for some $x_0 \in \mathbf{X}$. We have $\pi_X(\{x_0\}) = 1$ and $\nu_X(\{x_0\}) = 1$ so the event $\{x_0\}$ is both possible and necessary to the highest degree. For any other event U such that $x_0 \notin U$, we have $\pi_X(U) = 0$ and $\nu_X(U) = 0$. The event U is both impossible and unnecessary to the highest degree. We deem it impossible if an event other than $\{x_0\}$ happened unless the event contains x_0. The possibility function F_X quantifies therefore complete knowledge about X.

C. Bayesian inference

If Y (with realisations y) is another uncertain variable on a certain space $\mathbf{Y}$, the information we hold jointly about X and Y is quantified by the joint possibility function $F_{X,Y} : \mathbf{X} \times \mathbf{Y} \to [0, 1]$ which is such that $\max_{(x,y) \in \mathbf{X} \times \mathbf{Y}} F_{X,Y}(x, y) = 1$. The marginal possibility function of X and the conditional possibility function of X given a realisation of $y \in \mathbf{Y}$ of Y are [3, 4]

$$F_X(x) = \max_{y \in \mathbf{Y}} F_{X,Y}(x, y) \tag{2}$$

and

$$F_{X|Y}(x|y) = \frac{F_{X,Y}(x, y)}{F_Y(y)}. \tag{3}$$

Realisations of Y may also belong to a subset V of $\mathbf{Y}$. The generalised conditional possibility function of X given realisations $y \in V$ of Y is

$$F_{X|Y}(x|V) = \frac{\max_{y \in V} F_{X,Y}(x, y)}{\max_{y \in V} F_Y(y)}. \tag{4}$$

Note that (4) reduces to (3) when V is a singleton. The general form of Bayes' theorem for possibility functions is therefore

$$F_{X|Y}(x|V) = \frac{\max_{y \in V} F_X(x) F_{Y|X}(y|x)}{\max_{(x',y) \in \mathbf{X} \times V} F_X(x') F_{Y|X}(y|x')}. \tag{5}$$

Equations (1) and (5) provide together the posterior possibility and necessity of events $U \subset \mathbf{X}$ conditioned on realisations $y \in V$ of Y, namely

$$\begin{cases} \pi_{X|Y}(U|V) = \max_{x \in U} F_{X|Y}(x|V) \\ \nu_{X|Y}(U|V) = 1 - \max_{x \in U^c} F_{X|Y}(x|V). \end{cases} \tag{6}$$

D. Change of variable

We consider the change of variable $Z = \zeta(X)$ for a general mapping $\zeta : \mathbf{X} \to \mathbf{Z}$ [4]. The possibility function of the uncertain variable Z is obtained from that of X according to $F_Z(z) = \max_{x \in \zeta^{-1}(z)} F_X(x)$ where $\zeta^{-1}(z)$ is the inverse image of $z \in \mathbf{Z}$ by ζ, that is $\zeta^{-1}(z) = \{x \in \mathbf{X} : \zeta(x) = z\}$. We observe that if we hold no information about X at all and $F_X(x) = 1$ everywhere on $\mathbf{X}$, we find that $F_Z(z) = 1$ everywhere on $\mathbf{Z}$. We hold therefore no information about Z at all either. In the context of Bayesian inference with possibility theory, this means that an uninformative prior possibility function remains uninformative under a change of variable. We remember that on the other hand, in the context of Bayesian inference with probability theory, a uniform probability distribution, considered thus uninformative, does not in general remains uniform under a change of variable [6]. We now consider the true and false positives reported after a survey. If X represents the number of true positives and Y the number of false positives, $Z = \zeta(X, Y) = X + Y$ represents the number of detections. The possibility function of Z is

$$F_Z(z) = \max_{(x,y) \,\in\, \mathbf{X} \times \mathbf{Y} \,:\, x \,+\, y \,=\, z} F_{X,Y}(x, y). \tag{7}$$

E. Additional notions

Independence for uncertain variables is somewhat different than independence for random variables. If there exists possibility functions $G_X : \mathbf{X} \to [0, 1]$ and $G_Y : \mathbf{Y} \to [0, 1]$ and such that

$$F_{X,Y}(x, y) = G_X(x) G_Y(y) \tag{8}$$

for every $(x, y) \in \mathbf{X} \times \mathbf{Y}$, we say that X and Y are independently described by $G_X(x)$ and $G_Y(y)$. This notion of independence for uncertain variables only implies that the information we hold about X is not related to Y and conversely, which is very different to the notion of independence for random variables. A meaningful notion of expectation has been derived [8] and, in keeping with previous notations, the expectation of X conditioned on realisations $y \in V$ of Y is $\arg\max_{x \in \mathbf{X}} F_{X|Y}(x|V)$. Note that the expectation is set-valued in general.

III. Inference with possibility functions

A. Framework

Let N (with realisations n) be the uncertain variable on $\mathbb{N}$ that represents the number of targets in the area of interest where $\mathbb{N}$ is the set of natural numbers. Let C_i (with realisations c_i) be the uncertain variable on $\mathbb{N}$ that represents the number of contacts after survey i where i is a discrete index that identifies the survey. We know that the contacts are a mixture of true and false positives but we know nothing about their proportions. Accordingly, let C_{T_i} and C_{F_i} (with realisations c_{T_i} and c_{F_i}) be the uncertain variables on $\mathbb{N}$ that represent the number of true and false positives. We have accordingly $C_i = C_{T_i} + C_{F_i}$. After $m \in \mathbb{N}^\star$ surveys, where $\mathbb{N}^\star$ is the set of strictly positive natural numbers, we write $\mathbf{C}_m$ for the collection $(C_1, \ldots, C_m)$ and $\mathbf{c}_m$ for the collection $(c_1, \ldots, c_m)$. The posterior possibility function of the number of contacts is

$$F_{N|\mathbf{C}_m}(n|\mathbf{c}_m) = \frac{F_N(n) F_{\mathbf{C}_m|N}(\mathbf{c}_m|n)}{\mathcal{E}(\mathbf{c}_m)} \tag{9}$$

where

$$\mathcal{E}(\mathbf{c}_m) = \max_{n\in\mathbb{N}} F_N(n) F_{\mathbf{C}_m|N}(\mathbf{c}_m|n) \tag{10}$$

is the evidence. We consider the conditional expectation

$$\hat{N}_m = \arg\max_{n\in\mathbb{N}} F_{N|\mathbf{C}_m}(n|\mathbf{c}_m) \tag{11}$$

as a point estimate of the number of targets. We remember that the expectation is set-valued in general. We choose the median of the set as the point estimate. If one is much concerned about the underestimation of the number of targets, one may choose instead the maximum element as the point estimate of the number of targets, which is not what we do here. We assume that conditioned on a number of targets the information we hold about the numbers of contacts does not relate one to any other. This is is the notion of independence detailed in section II and we write accordingly

$$F_{\mathbf{C}_m|N}(\mathbf{c}_m|n) = \prod_{i=1}^{m} F_{C_i|N}(c_i|n). \tag{12}$$

We remember that $C_i = C_{T_i} + C_{F_i}$ which, combined with the change of variable formula of Section II, leads to

$$F_{C_i|N}(c_i|n) = \max_{c_{T_i}+c_{F_i}=c_i} F_{C_{T_i},C_{F_i}|N}(c_{T_i}, c_{F_i}|n). \tag{13}$$

We assume once again that conditioned on a number of targets the information we hold about the number of true and false positives does not relate one to the other so that we write further

$$F_{C_i|N}(c_i|n) = \max_{c_{T_i}+c_{F_i}=c_i} F_{C_{T_i}|N}(c_{T_i}|n) F_{C_{F_i}}(c_{F_i}). \tag{14}$$

At this point we have to specify the prior possibility function of the number of targets, the conditional possibility function of the number of true positives given the number of targets, and that of the number of false positives.

B. Prior on the number of targets

For the prior possibility function on the number of targets, we consider two cases. In the first case, we assume that we hold no prior information on the number of targets, and we have accordingly

$$F_N(n) = 1 \quad \text{for all} \quad n \in \mathbb{N}. \tag{15}$$

In the second case, we assume that we are certain that the number of targets does not exceed a certain number $n_0 \in \mathbb{N}$, and we have accordingly

$$F_N(n) = \begin{cases} 1 & \text{if } 0 \le n \le n_0 \\ 0 & \text{if } n > n_0 \end{cases} \tag{16}$$

As expected, the prior possibility and necessity of the event $0 \le N \le n_0$ are both equal to one while the prior possibility and necessity of the event $N > n_0$ are both equal to zero. The motivation for considering a prior that imposes an upper bound on the number of targets as well as one that does not is related to the Monte Carlo simulations that will come later. The number of targets will be uniformly chosen at random in between zero and some upper bound, and it seems appropriate to consider a prior that captures some of this information.

C. Binomial-Poisson model

For the number of true and false positives, we consider a model derived from the Binomial and the Poisson probability distributions with uncertain parameters. We assume that the uncertain Binomial detection rate A (with realisations α) belongs to an interval $[\alpha_0, \alpha_1]$ where $0 \le \alpha_0 < \alpha_1 \le 1$ as opposed to the set of positive real numbers. We also assume that the uncertain Poisson parameter Λ (with realisations λ) belongs to an interval $[\lambda_0, \lambda_1]$ where $0 \le \lambda_0 < \lambda_1$. The prior possibility functions on A and Λ are equal to one on $[\alpha_0, \alpha_1]$ and $[\lambda_0, \lambda_1]$ respectively and zero elsewhere. We find that

$$\begin{cases} F_{C_{T_i}|N}(c_{T_i}|n) = \max\limits_{\alpha_0\le\alpha\le\alpha_1} \bar{\mathcal{B}}(n, c_{T_i}, \alpha) \\ F_{C_{F_i}}(c_{F_i}) = \max\limits_{\lambda_0\le\lambda\le\lambda_1} \bar{\mathcal{P}}(c_{F_i}, \lambda) \end{cases} \tag{17}$$

where $\bar{\mathcal{B}}$ and $\bar{\mathcal{P}}$ are the Binomial and Poisson possibility functions derived from the Binomial and Poisson probability distributions and detailed in the Appendix. The distributions are normalised by a multiplicative factor so that the corresponding possibility functions take their values in the interval $[0, 1]$. Note that Λ represents the average number of false positives in the area that is being surveyed and also its variance. Also note that $\bar{\mathcal{B}}(n, c_{T_i}, \alpha) \neq 0$ when $0 \le c_{T_i} \le n$ and $\bar{\mathcal{B}}(n, c_{T_i}, \alpha) = 0$ when $c_{T_i} > n$ for it is not possible to have more true positives than the number of targets.

D. Example

We consider here a scenario based on simulated data. We must therefore detail the probability distributions which the simulated data is based on as well as the possibility functions which the inference is based on. The true number of targets is not randomly chosen and set to $n^* = 15$. We assume we

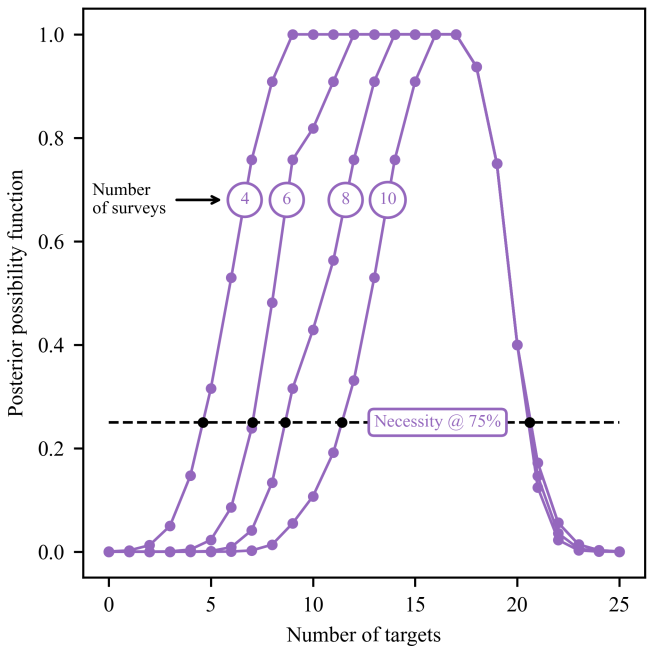

Fig. 1. Posterior possibility functions $F_{N|\mathbf{C}_m}(n|\mathbf{c}_m)$ on the number of targets.

hold no prior information on the number of targets so that the prior is given by (15). The Binomial detection rate is uniformly chosen at random between $\alpha_0 = 0.8$ and $\alpha_1 = 0.9$ and independently for every survey. The information conveyed by the prior possibility function on the Binomial detection rate – an uncertain rather than a random variable as far as the inference is concerned – is that it is somewhere in the interval $[\alpha_0, \alpha_1]$. The Poisson parameter is uniformly chosen at random between $\lambda_0 = 0$ and $\lambda_1 = 10$ and independently for every survey. The information conveyed by the prior possibility function on the Poisson parameter – once again an uncertain rather than a random variable as far as the inference is concerned – is that it is somewhere in the interval $[\lambda_0, \lambda_1]$. True and false positives are simulated independently according to the Binomial and Poisson discrete probability distributions and added together to produce the number of detections. One such simulation produced the following number of detections:

$$(c_1, \dots, c_{10}) = (17, 19, 15, 14, 19, 21, 16, 23, 25, 16). \quad (18)$$

Fig. 1 shows the posterior possibility functions on the number of targets given the number of detections as a function of the number of targets and the number of surveys. We chose $m \in \{4, 6, 8, 10\}$ rather than $m \in \{1, \dots, 10\}$ so as to make the figure as clear as possible. We first observe that for a given number of surveys, the posterior possibility function is maximum over a subset of the natural numbers. We remember that the point estimator of the number of targets $\hat{N}_m$ is chosen as the median that subset, which produces half-integer values when the subset is composed of an even number of elements. For $m \in \{4, 6, 8, 10\}$, we find that

$$(\hat{N}_4, \hat{N}_6, \hat{N}_8, \hat{N}_{10}) = (13, 14.5, 15.5, 16.5). \quad (19)$$

More information may be extracted from the posterior possibility functions. The posterior possibility and necessity of events $U \subset \mathbb{N}$ are

$$\begin{cases} \pi_m(U|\mathbf{c}_m) = \max_{n \in U} F_{N|\mathbf{C}_m}(n|\mathbf{c}_m) \\ \nu_m(U|\mathbf{c}_m) = 1 - \max_{n \in U^c} F_{N|\mathbf{C}_m}(n|\mathbf{c}_m). \end{cases} \quad (20)$$

Say we are interested in quantifying the credibility of the event $N = n^* = 15$, which is the true number of targets. We find that after four surveys, the possibility and necessity of the event are:

$$\begin{cases} \pi_4(N = 15|(17, 19, 15, 14)) = 1 \\ \nu_4(N = 15|(17, 19, 15, 14)) = 0. \end{cases} \quad (21)$$

The event $N = 15$ is both possible and unnecessary to the highest degree. We hold therefore little information on the event as explained in Section II. Careful examination of the posterior possibility functions shows that we hold as little information on the event after ten surveys. We must therefore broaden the set $\{15\}$ and, to this effect, we consider the subsets of the natural numbers for which the necessity is at 75%. The black dots along the dashed line in Fig. 1 indicate where are the extremities of these subsets. We find that after four surveys, it is 75% certain that the number of targets is in the subset $\{4, \dots, 21\}$ and in the subset $\{11, \dots, 21\}$ after ten surveys. It is clear that this type of reasoning may be used to derive confidence intervals on the number of targets.

IV. Inference with probability distributions

We now turn to the Bayesian inference of the number of targets with probability theory. The approach we present here is a generalisation of the approach presented in [7] to the case where the target detection rate and the Poisson parameter are random, both with a uniform prior on an interval[1]. In what follows, what was previously an uncertain variable is now a random variable. We keep our notations as much the same as before and expect the reader to tell whether a variable is an uncertain variable or a random variable from the context. The posterior probability distribution of the number of targets is

$$P_{N|\mathbf{C}_m}(n|\mathbf{c}_m) = \frac{P_N(n) \prod_{i=1}^m P_{C_i|N}(c_i|n)}{\mathcal{E}'(\mathbf{c}_m)} \quad (22)$$

where

$$\mathcal{E}'(\mathbf{c}_m) = \sum_{n=0}^{\infty} \left\{ P_N(n) \prod_{i=1}^{m} \left\{ P_{C_i|N}(c_i|n) \right\} \right\} \quad (23)$$

is the evidence. The sigma additivity of probability distributions together with the assumed statistical independence between the numbers of true and false positives leads to

$$P_{C_i|n}(c_i|n) = \sum_{c_{T_i}=0}^{c_i} \left\{ P_{C_{T_i}|N}(c_{T_i}|n) P_{C_{F_i}}(c_i - c_{T_i}) \right\} \quad (24)$$

[1]To be more specific, we use a Poisson probability distribution to model the number of false positives. In [7] the area of interest is divided into cells. The number of false positives thus follows a Binomial probability distribution with a certain false detection rate.

A uniform prior over the Binomial detection rate A and another over the Poisson parameter Λ leads to

$$\begin{cases} P_{C_{T_i}|N}(c_{T_i}|n) = (\alpha_1 - \alpha_0)^{-1} \displaystyle\int_{\alpha_0}^{\alpha_1} \mathcal{B}(n, c_{T_i}, \alpha) d\alpha \\ P_{C_{F_i}}(c_{F_i}) = (\lambda_1 - \lambda_0)^{-1} \displaystyle\int_{\lambda_0}^{\lambda_1} \mathcal{P}(c_{F_i}, \lambda) d\lambda \end{cases} \tag{25}$$

where $\mathcal{B}$ and $\mathcal{P}$ denote the Binomial and the Poisson probability distributions. The integrals can be evaluated numerically with the help of the incomplete Beta function and the incomplete Gamma function as detailed in the Appendix. We decide to use a proper prior probability distribution on the number of targets so that we have

$$P_N(n) = \begin{cases} 1 & \text{if } 0 \le n \le n_0 \\ 0 & \text{if } n > n_0 \end{cases} \tag{26}$$

for a certain $n_0 \in \mathbb{N}$. In keeping with the spirit of the point estimation procedure with possibility functions, the point estimator of the number of targets with probability distributions is the conditional expectation

$$\hat{N}'_m = \sum_{n=0}^{+\infty} \{ n \, P_{N|\mathbf{C}_m}(n|\mathbf{c}_m) \}. \tag{27}$$

At this point, we have three point estimation procedures of the number of targets from the number of contacts. The first two are based on possibility theory and differ only by the prior on the number of targets. One with and the other without an upper bound on the number of targets. These two procedures will be referred to as PS-UB and PS-NO-UB in the figures that follow. The third procedure is based on probability theory and imposes an upper bound on the number of targets. It will be referred to as PR-UB in the figures that follow.

V. Performance

We consider here the performance of the three point estimators. We use the bias, the standard deviation (STD) and the root mean squared error (RMSE) to perform the comparison. The unit for all is the number of targets, which is why we favour say, the standard deviation over the variance. They are related to one another according to the well-known decomposition [9, p.147]

$$(\text{RMSE})^2 = (\text{bias})^2 + (\text{STD})^2. \tag{28}$$

They are estimated with a Monte Carlo simulation. The number of Monte Carlo runs is set to one million. The number of targets is uniformly chosen at random between zero and $n_0 = 30$. The Binomial detection rate is uniformly chosen at random between $\alpha_0 = 0.8$ and $\alpha_1 = 0.9$ and independently for every survey. The Poisson parameter is uniformly chosen at random between $\lambda_0 = 0$ and $\lambda_1 = 10$ and independently for every survey. For every survey, the numbers of true and false positives are simulated independently according to the Binomial and Poisson probability distributions and added together to produce the number of detections. Fig. 2 shows the overall bias, STD and RMSE for the three point estimators. We observe that they are overall all nearly unbiased. Therefore, the STD and the RMSE are nearly indistinguishable. The point estimators based on possibility theory do worse than the one based on probability theory in terms of the RMSE but by no more than half a target. In particular, this is true for the one that does not impose an upper bound on the number of targets and thus rely on one fewer parameter. Fig. 3 and 4 show the bias as a function of the true number of targets when the number of surveys is set to two and to four. As expected, an increase in the number of surveys reduces the bias. All three estimators are bias towards zero which is a natural lower bound on the number of targets. There the number of targets is over-estimated. When one imposes an upper bound on the number of targets with a prior possibility function or probability distribution, the resulting estimator is biased towards this upper bound, in our case towards $n_0 = 30$. When one does not impose an upper bound on the number of targets, which is possible with a prior possibility functions, the bias is nearly equal to zero. If an increase in the RMSE by half a target is acceptable, we recommend using the estimator that does not rely on prior information on the number of targets.

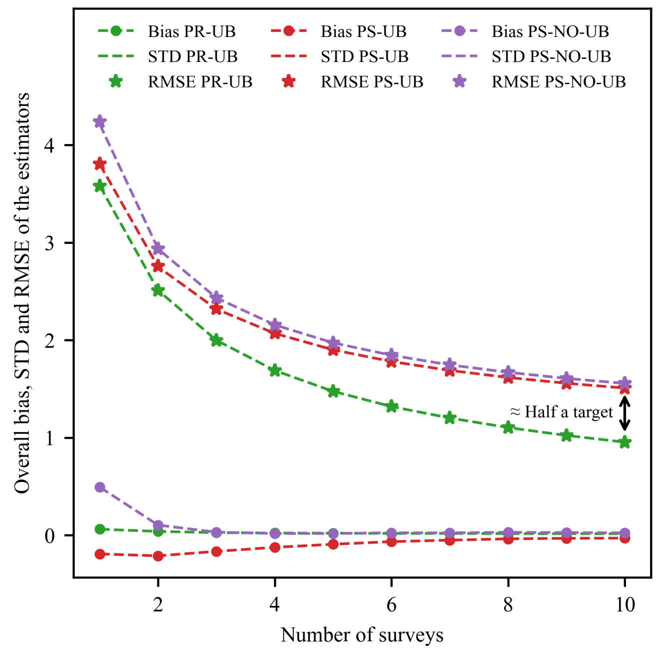

Fig. 2. Overall bias, standard deviation (STD) and root mean squared error (RMSE) of the three point estimators as a function of the number of surveys. Note that for all three point estimators, the STD is nearly equal to the RMSE since the bias is nearly equal to zero.

VI. Conclusion

This paper has been concerned with the Bayesian estimation of the number of targets in an area of interest when no prior information on the number of targets is available. We based our work on possibility theory for it is capable of quantifying the complete absence of information about the number of targets. We derived two point estimators this way – one that does not impose an upper bound on the number of targets and,

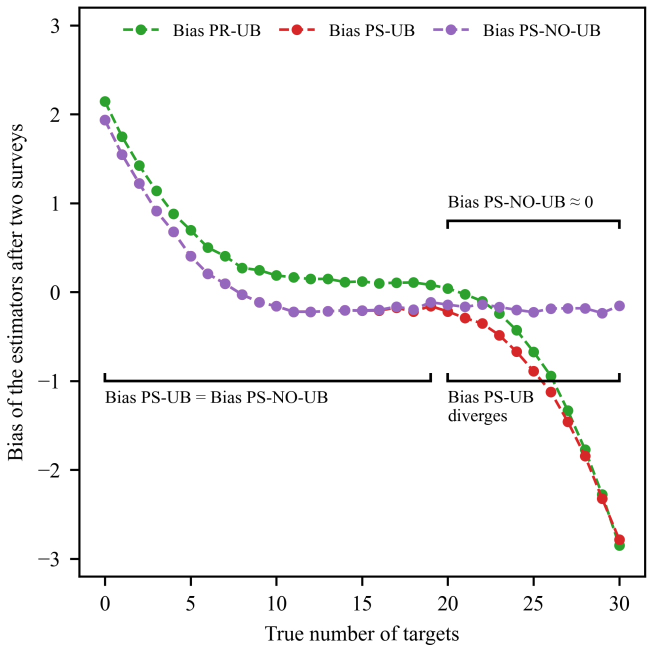

Fig. 3. Bias of the three point estimators as a function of the true number of targets when the number of surveys is equal to two. A positive (negative) bias corresponds to an over-estimation (under-estimation) of the number of targets.

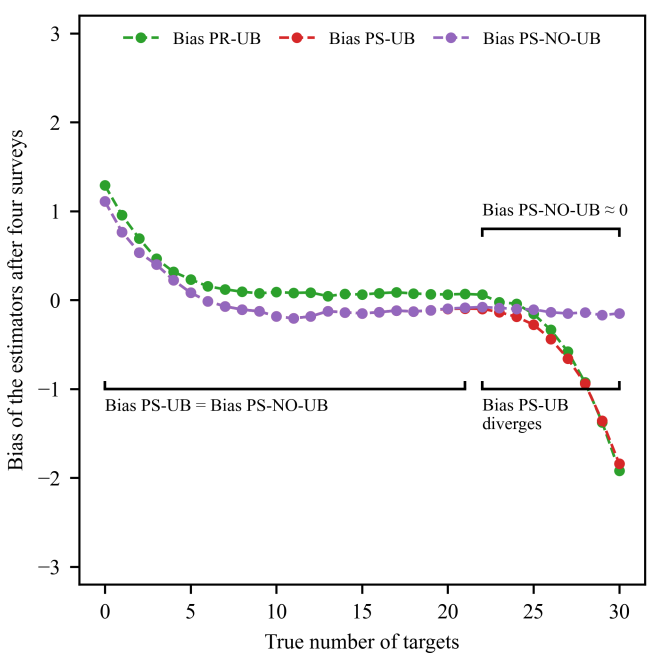

Fig. 4. Bias of the three point estimators as a function of the true number of targets when the number of surveys is equal to four. A positive (negative) bias corresponds to an over-estimation (under-estimation) of the number of targets.

for the sake of completeness, another that does. We compared the performance of the estimators to a that of a third based on probability theory. Note that the estimator based on possibility theory that does not impose an upper bound on the number of targets relies effectively on one fewer parameter than the other two – the upper bound itself. We showed, in simulation, that their overall performance in terms of the bias, standard deviation and root mean squared error are equivalent give or take half a target. The better performance goes to the estimator based on probability theory. In terms of the bias as a function of the true number of targets, we recommend using the estimator that does not impose an upper bound on the number of targets for its bias is towards the upper bound is nearly equal to zero.

Appendix

A. *Binomial and Poisson possibility functions*

The Binomial and Poisson possibility functions are derived from the Binomial and Poisson probability distributions with a suitable normalisation. The Binomial probability distribution $\mathcal{B}(n,k,\alpha)$ with parameters $n \in \mathbb{N}$ and $\alpha \in [0,1]$ is defined for $k \in \mathbb{N}$ by

$$\mathcal{B}(n,k,\alpha) = \begin{cases} C_n^k \alpha^k (1-\alpha)^{n-k} & \text{if } 0 \leq k \leq n \\ 0 & \text{if } k > n. \end{cases} \tag{29}$$

The Poisson probability distribution $\mathcal{P}(k,\lambda)$ with parameter $\lambda \in \mathbb{R}^+$ is defined for $k \in \mathbb{N}$ by

$$\mathcal{P}(k,\lambda) = e^{-\lambda}(\lambda^k/k!). \tag{30}$$

The Binomial and Poisson possibility functions are defined by

$$\begin{cases} \bar{\mathcal{B}}(n,k,p) = \mathcal{B}(n,k,p)/\max\limits_{k' \in \mathbb{N}} \mathcal{B}(n,k',p) \\ \bar{\mathcal{P}}(k,\lambda) = \mathcal{P}(k,\lambda)/\max\limits_{k' \in \mathbb{N}} \mathcal{P}(k',\lambda). \end{cases} \tag{31}$$

B. *Bayesian integration*

The Bayesian integration of the Binomial probability distribution over its parameter involves the incomplete Beta function. We have

$$\begin{aligned} \int_{\alpha_0}^{\alpha_1} \mathcal{B}(n,k,\alpha)d\alpha &= C_n^k\, B(k+1, n-k+1) \\ \times\ (I_{\alpha_1}(k+1,n-k+1) &- I_{\alpha_0}(k+1,n-k+1)) \end{aligned} \tag{32}$$

where B and I are the Beta function and the incomplete Beta function. The Bayesian integration of the Poisson probability distribution over its parameter involves the incomplete Gamma function. We have and

$$\int_{\lambda_0}^{\lambda_1} \mathcal{P}(k,\lambda)d\lambda = \frac{\Gamma(k+1)}{k!}\left(P_{\lambda_1}(k+1) - P_{\lambda_0}(k+1)\right) \tag{33}$$

where Γ and P denote the Gamma function and the incomplete Gamma function.

References

[1] D. Dubois and H. Prade. *Théorie des Possibilités*. Masson, 1985.

[2] D. Dubois and H. Prade. "Possibility Theory and Its Applications: Where Do We Stand?" In: *Springer Handbook of Computational Intelligence*. Springer, 2015.

[3] E. Delande, J. Houssineau, and J. Moriba. "A New Representation of Uncertainty for Data Fusion in SSA Detection and Tracking Problems". In: *21st International Conference on Information Fusion*. 2018, pp. 1–8.

[4] J. Houssineau. "A Linear Algorithm for Multi-Target Tracking in the Context of Possibility Theory". In: *IEEE Transactions on Signal Processing* 69 (2021), pp. 2740–2751.

[5] A.P. Dawid, M. Stone, and J.V. Zidek. "Marginalization paradoxes in Bayesian and structural inference". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 35.2 (1973), pp. 189–213.

[6] A. Ly et al. "A Tutorial on Fisher Information". In: *Journal of Mathematical Psychology* 80 (2017), pp. 40–55.

[7] Ø. Midtgaard et al. "Multi-phase performance evaluation for modern minehunting systems". In: SPIE, 2019.

[8] J. Houssineau, N.K. Chada, and E. Delande. *Elements of asymptotic theory with outer probability measures*. 2020. arXiv: 1908.04331.

[9] C.M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.