Proceedings of Machine Learning Research 215:438–449, 2023 ISIPTA 2023

Open World Dempster-Shafer Using Complementary Sets

Erik Skau EWSKAU@LANL.GOV
Los Alamos National Laboratory, Information Sciences Group

Cassandra Armstrong CLARMSTRONG@LANL.GOV
Los Alamos National Laboratory, Space Science and Applications Group

Duc P. Truong DPTRUONG@LANL.GOV
Los Alamos National Laboratory, Information Sciences Group

David Gerts DGERTS@LANL.GOV
Los Alamos National Laboratory, Space Science and Applications Group

Kari Sentz KSENTZ@LANL.GOV
Los Alamos National Laboratory, Information Sciences Group

Abstract

Dempster-Shafer Theory (DST) is a mathematical framework to handle imprecision and uncertainty in reasoning and decision making. One assumption of DST is that of a closed-world, or the assumption that all propositions are known *a priori*. In this work, we explore an alternative formulation of Dempster-Shafer that allows for the dynamic inclusion of new propositions. Specifically, we expand the framework to include the complement of every set of propositions. This adjustment enables an open-world interpretation that can support unspecified and dynamic propositions as we learn about the problem space. Including complementary sets distinguishes this from previous work in DST where the open world is attributed to the empty set. We demonstrate our open world Dempster-Shafer Theory on a variety of synthetic and real datasets.

1. Introduction

Dempster-Shafer Theory (DST) [26, 4] is a mathematical theory of evidence that generalizes probability theory to account for imprecise information. Dempster-Shafer theory is one of many possible generalizations of probability associated with nonadditive monotone measures [13, 30]. The set-theoretic interpretation of DST makes it particularly attractive for this work in complementary sets. The Dempster rule of combination has been reasonably criticized for counter-intuitive results in the face of conflict [33, 34] but still has relevance when a product intersection combination rule is applicable. The limited range of applicability of Dempster's rule has led to the development of many different combination rules ranging from conjunctive and disjunctive operations that afford an enhanced expressibility for data and information fusion. As a consequence, DST remains an important applied approach of imprecise probabilities in the data fusion community. Accordingly, DST and the related methods, Probability Bounds Analysis (PBA) [8], Generalized Evidence Theory (GET) [5], Transferable Belief Model (TBM) [27], can be found in a broad range of applications: target identification [3], image analysis [2], land cover classification [14], qualifying model predictions [11], risk analysis [9], validation and verification in scientific computing [18], to name a few.

The basis of DST is a *frame of discernment*, or collection of possible propositions one wishes to consider, sometimes called a sample space in probability theory. *Basic Probability Assignments* (BPA) are mappings that allocate mass to subsets, called *focal elements*, of the frame of discernment, indicating the degree of belief in the proposition or collection of propositions defined by the focal element. BPAs are typically represented with the letter m and satisfy,

$$\begin{aligned} m : 2^{\Omega} &\to [0,1] \\ \sum_{u\in 2^{\Omega}} m(u) &= 1 \ . \end{aligned} \tag{1}$$

In classical DST, BPAs had another constraint that $m(\emptyset) = 0$. In extensions of DST such as the Transferable Belief Model [27, 29], $m(\emptyset)$ can take on values in the interval $0 \le m(\emptyset) \le 1$ to represent conflict. From the BPAs, lower and upper probability bounds for each focal element can be derived, called the *belief* and *plausibility* respectively,

$$\begin{aligned} \mathrm{bel}_m(u) &= \sum_{\substack{v\in 2^{\Omega} \\ v\subseteq u \\ v\neq\emptyset}} m(v) \\ \mathrm{pl}_m(u) &= \sum_{\substack{v\in 2^{\Omega} \\ v\cap u\neq\emptyset}} m(v) \ . \end{aligned} \tag{2}$$

1.1. Open World Interpretation and DST

The original conception of DST assumes a closed world, namely, that the frame of discernment is known exhaustively *a priori* [26]. The closed world assumption poses a challenge to any theory or representation where something outside the sample space or frame of discernment can be encountered. There have been a variety of methods proposed to mitigate the closed world problem or to construct open world variants for DST. TBM and GET address the closed world problem by allowing mass in the empty set which can be used to represent a level of ignorance [27, 5]. This is a departure from the original DST concept where $m(\emptyset) = 0$. This approach is an improvement over DST for the representation of conflict, but conflates the situation where BPAs conflict with one another and when there is a missing proposition from the frame of discernment.

Another method, called extended open world [23], extends the frame of discernment with an "*" singleton that acts as a wildcard to represent all unspecified propositions. This approach cannot systematically add unspecified propositions to the frame of discernment. Another drawback of GET, TBM, and extended open world, is their inability to distinguish between a single missing proposition and multiple missing propositions without recomputation or adjustments of BPAs.

A third open world approach to DST focuses on the idea of *non-exhaustive frames of discernment*. In this there have been various methods proposed to identify when a frame of discernment is incomplete and incrementally adding new propositions [31, 16, 12, 32]. These suffer from having to re-compute BPAs and loses interpretability of the propositions.

A frame of discernment can also be extended by taking a Cartesian product with another frame of dissimilar focal elements [7, 3]. This allows for extensions of BPAs from one frame to a Cartesian product of frames by pairing each focal element with the vacuous component in its Cartesian product counterpart. While useful, this approach does not provide a meaningful way to extend a frame of discernment with another frame of similar focal elements. A frame of discernment can be reduced in size either with marginalization when it is a Cartesian product of frames, or by removing focal elements to find a suitable frame by balancing the aggregated uncertainty and the amount of conflict [25].

1.2. Complements and DST

In addition to the precedent of work in open world interpretations for DST, there is also precedent for work with complement sets in DST [22, 20]. These are concerned with the truth value of a proposition and employ the complement to represent what is *not true* in source reliability and probabilistic argumentation systems.

1.3. A New Approach

In the complementary DST approach proposed here, we introduce additional focal elements representing complements of sets into the basic probability assignments to represent belief of "what is not." With this we gain an open world interpretation while still explicitly quantifying the deficits of incomplete information. The complementary DST generalization provides the following capabilities:

- identify when a frame of discernment is incomplete,
- combine BPAs that do not share identical frames of discernment,
- systematically add propositions to the frame of discernment without needing to recompute BPAs.

In Section 2 we provide generalized definitions and theorems for complementary DST, and in Section 3 we demonstrate these capabilities on various synthetic and real data.

2. Complementary Dempster-Shafer Theory

There are several important concepts that we generalize for complementary DST: focal elements, BPAs, joining of BPAs, and belief and plausibility. While the proofs of all theorems are relatively straightforward from definitions and identities, they are included in supplementary material for completeness.

2.1. Complementary Basic Probability Assignments

The power of DST is the ability to assign a probability-like mass to a set of propositions (focal element) without having to ascribe a mass to the constituents of the set. We introduce complementary focal elements to represent elements that are not constituents of a set of propositions of interest.

Definition 1 Complementary focal elements *are members of the Cartesian product between the power set of a frame of discernment and the Boolean space,* $(u, a) \in 2^{\Omega} \times \mathbb{B}$.

We use the Boolean variable, a, in $(u, a) \in 2^{\Omega} \times \mathbb{B}$ as a flag to differentiate between the set u with $a = True$ and the complement of u with $a = False$. In other words, $(u, True)$ represents the hypothesis of u, while $(u, False)$ represents the hypothesis of everything other than u, including propositions not in the current frame of discernment.

Complementary focal elements on one frame of discernment, $(u, a) \in 2^{\Omega_1} \times \mathbb{B}$ can be evaluated to focal elements

Table 1: All complementary focal elements for frame of discernment $\Omega = \{A, B\}$.

$(\emptyset, True)$	$(\{A, B\}, False)$
$(\{A\}, True)$	$(\{B\}, False)$
$(\{B\}, True)$	$(\{A\}, False)$
$(\{A, B\}, True)$	$(\emptyset, False)$

on a different frame of discernment Ω_2 using the identities

$$\begin{aligned} (u, True)|_{\Omega_2} &= \Omega_2 \cap u \\ (u, False)|_{\Omega_2} &= \Omega_2 - u \ . \end{aligned} \tag{3}$$

For example, consider the frame of discernment consisting of the propositions $\Omega_2 = \{A, B\}$ and the complementary focal elements listed in Table 1. We see the evaluation of these complementary focal elements on the different frames of discernment including $\Omega_1 = \{A\}$, and $\Omega_3 = \{A, B, C\}$ in Figure 1. It is evident from Figure 1(a) that if propositions are removed in the evaluation, multiple complementary focal elements evaluate to the same sets. In Figure 1(b) it is clear that there are two ways to represent every set when evaluated on its natural frame of discernment. These two representations correspond to the set itself, and the complement of the complement of the set. Figure 1(c) demonstrates the real power of the complementary frames of discernment, the fact that complementary focal elements can evaluate to cover previously unknown propositions. Specifically, both complementary focal elements $(\{A\}, True)$ and $(\{B\}, False)$ evaluate to $\{A\}$ on the frame of discernment Ω_2. The compelling property of the set $(\{B\}, False)$ is that in the event of a new proposition, C, being added its evaluation changes from $\{A\}$ in Figure 1(b) to $\{A, C\}$ in Figure 1(c).

We use complementary focal elements to define complementary BPAs.

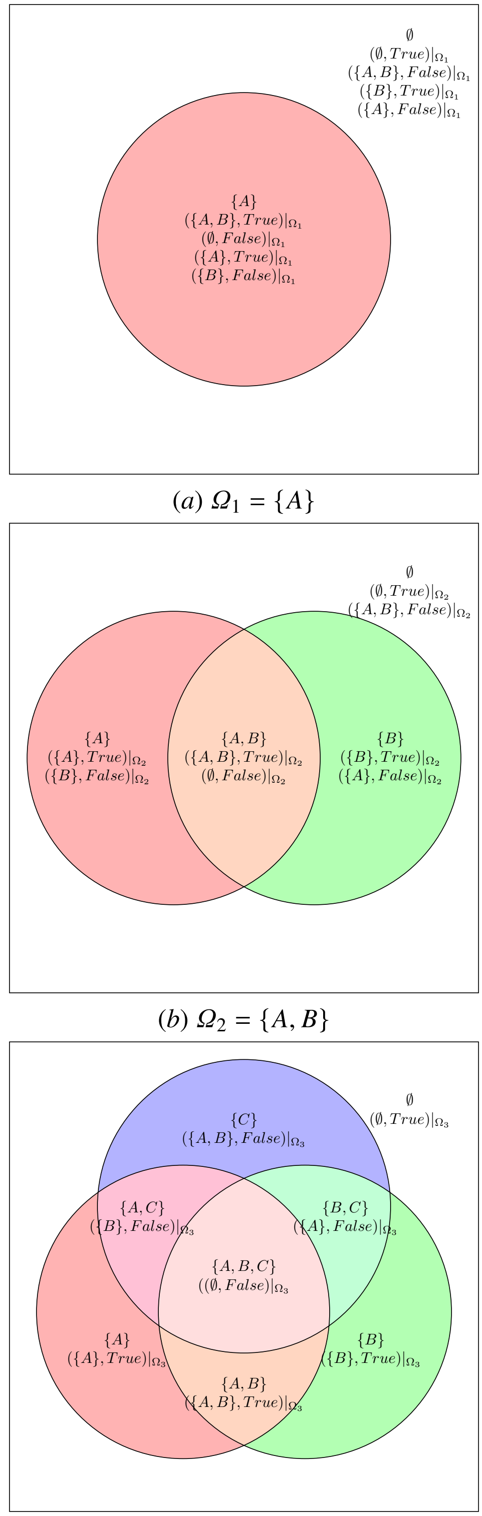
Definition 2 *A Complementary Basic Probability Assignment (CBPA) on a frame of discernment, Ω, is a bivariate mapping*

$$m : 2^{\Omega} \times \mathbb{B} \to [0, 1]$$

that satisfies $\sum_{x \in 2^{\Omega} \times \mathbb{B}} m(x) = 1$.

Since CBPAs are supported on complementary focal elements, they can be evaluated on different frames of discernments to construct traditional BPAs. A CBPA m defined with frame of discernment Ω_1 can be evaluated on Ω_2 resulting in a BPA defined by

$$\begin{aligned} m|_{\Omega_2} &: 2^{\Omega_2} \to [0, 1] \\ m|_{\Omega_2}(x) &= \sum_{\substack{y \in 2^{\Omega_1} \times \mathbb{B} \\ y|_{\Omega_2} = x}} m(y) \ . \end{aligned} \tag{4}$$



(a) $\Omega_1 = \{A\}$

(b) $\Omega_2 = \{A, B\}$

(c) $\Omega_2 = \{A, B, C\}$

Figure 1: Venn diagrams of all complementary focal elements from $\Omega_2 = \{A, B\}$ evaluated on different frames of discernment.

This reduces to,

$$m|_{\Omega}(x) = m((x, True)) + m((\Omega - x, False)) \quad (5)$$

for a CBPA m both defined and evaluated on the same frame of discernment Ω.

CBPAs offer different, more nuanced representations of conflict. For a given frame of discernment, Ω_1, both $(\emptyset, True)$ and $(\Omega_1, False)$ indicate a type of conflict in belief as they both evaluate to the empty set on Ω_1. However, when evaluated on a different frame of discernment, Ω_2, their meanings diverge where $(\emptyset, True)|_{\Omega_2} = \emptyset$ while $(\Omega_1, False)|_{\Omega_2} = \Omega_2 - \Omega_1$. This differentiation allows us to quantify conflict and missing propositions separately and explicitly. The mass in $(\emptyset, True)$ signifies the fundamental conflict between assertions. The mass in $(\Omega_1, False)$ is ignorance that can be resolved by adding one or more propositions missing from Ω_1. This approach differentiates between fundamental conflict and the ignorance that arises from a lack of information. Significant mass in $(\emptyset, True)$ is indicative of a fundamental conflict, while significant mass in $(\Omega_1, False)$ is indicative of a missing proposition.

The dual representations of the vacuous focal element are also noteworthy. For a given frame of discernment, Ω_1, the sets $(\emptyset, False)$ and $(\Omega_1, True)$ are identical in meaning as they both describe the whole space. When evaluated on another frame of discernment, Ω_2, their meanings again diverge with $(\emptyset, False)|_{\Omega_2} = \Omega_2$ and $(\Omega_1, True)|_{\Omega_2} = \Omega_2 \cap \Omega_1$. So mass in $(\emptyset, False)$ is a true vacuous belief containing no information in the belief of any set of propositions. In contrast, mass in $(\Omega_1, True)$ is not actually a vacuous statement, it is a statement in the confidence of the propositions in Ω_1.

2.2. Combining Complementary Basic Probability Assignments

There are numerous competing rules of combination in DST, GET, and TBM, each with their own advantages and disadvantages. To limit the scope, we concentrate our discussion on the TBM conjunctive join rule [6]. While our analysis concentrates on the conjunctive join, it is straightforward to modify our analysis to work with Dempster's rule of combination for DST [29], as well as various other rules such as Yager's rule or alternatively implement a disjunctive join. We first define the intersection rules between complementary focal elements with the identities,

$$\begin{aligned}
(u, True) \cap (v, True) &\coloneqq (u \cap v, True)\\
(u, True) \cap (v, False) &\coloneqq (u - v, True)\\
(u, False) \cap (v, True) &\coloneqq (v - u, True)\\
(u, False) \cap (v, False) &\coloneqq (v \cup u, False)\ .
\end{aligned} \quad (6)$$

These identities are best understood using a Venn diagram with overlapping u and v regions representing arbitrary subsets of an unknown global frame of discernment Ω as seen in Figure 2. In the figure, $(u, True)$ represents the set u while $(u, False)$ represents the complement of u with respect to the unknown global set Ω. When applied to the figure, the intersection identities become self-evident.

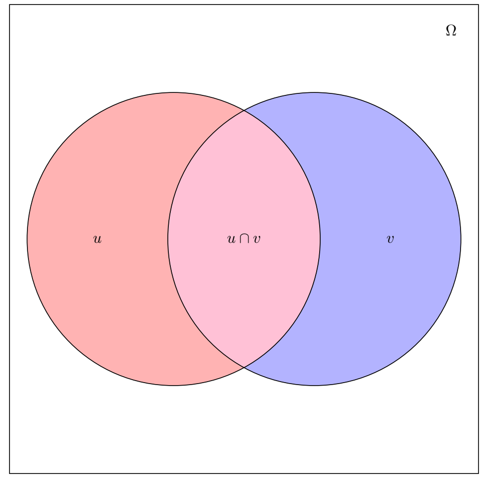


Figure 2: Venn diagram with subsets u and v within an unknown global frame of discernment Ω.

With the intersection rules specified we define the conjunctive join operators.

Definition 3 *Let m_1 and m_2 be CBPAs with frames of discernment Ω_1 and Ω_2 respectively. The* conjunctive join *$m_{1,2} = m_1 \oplus m_2$ is another CBPA with frame of discernment $\Omega_1 \cup \Omega_2$, defined by*

$$m_{1,2} : 2^{\Omega_1 \cup \Omega_2} \times \mathbb{B} \to [0, 1] \quad (7)$$

$$m_{1,2}(z) = \sum_{\substack{x \in 2^{\Omega_1} \times \mathbb{B} \\ y \in 2^{\Omega_2} \times \mathbb{B} \\ x \cap y = z}} m_1(x) \cdot m_2(y)\ . \quad (8)$$

This is essentially Smets' conjunctive join rule [27] applied to CBPAs. This is able to fuse CBPAs together even if they do not share a frame of discernment as the complementary focal elements and intersection rules can accommodate an unknown global frame of discernment. It is easy to see that the conjunctive join is also a CBPA by observing that every product $m_1(x) \cdot m_2(y)$ for $x \in 2^{\Omega_1} \times \mathbb{B}$ and $y \in 2^{\Omega_2} \times \mathbb{B}$ is included in exactly one summation, therefore the summation over z is also 1.

If the global frame of discernment is known *a priori*, then the evaluation of each constituent CBPA is equivalent to evaluating the joined CBPA at the end. This is stated in Theorem 4.

Theorem 4 *Let m_1 and m_2 be CBPAs with frame of discernments Ω_1 and Ω_2 respectively. For a new frame of discernment Ω, the evaluation distributes through the conjunctive join,*

$$m_1|_\Omega \oplus m_2|_\Omega = (m_1 \oplus m_2)|_\Omega \ .$$

While Theorem 4 is useful for understanding the relation between complementary DST and other methods like GET and TBM, it is practically of little use since it requires knowledge of an exhaustive frame of discernment beforehand.

Here, we additionally list the identities for the union of complementary focal elements,

$$\begin{aligned}(u, True) \cup (v, True) &:= (u \cup v, True)\\ (u, True) \cup (v, False) &:= (v - u, False)\\ (u, False) \cup (v, True) &:= (u - v, False)\\ (u, False) \cup (v, False) &:= (v \cap u, False)\ ,\end{aligned} \tag{9}$$

which constitute the core mechanism of the disjunctive join [6].

2.3. Belief and Plausibility

In the framework of DST, the degree to which evidence supports a proposition or set of propositions within a frame of discernment is bounded by the values $belief$ and $plausibility$ where $belief \le plausibility$. These measures can also be generalized to complementary focal elements for an open world frame of discernment.

To define belief, we first specify how to determine whether one complementary focal element is a subset of another. Letting $(u, a) \in 2^{\Omega_1} \times \mathbb{B}$ and $(v, b) \in 2^{\Omega_2} \times \mathbb{B}$ we define the logical subset operator for all pairwise combinations of the Boolean variables,

$$\begin{aligned}(u, True) \subseteq (v, True) &:= u \subseteq v\\ (u, True) \subseteq (v, False) &:= (u \cap v = \emptyset)\\ (u, False) \subseteq (v, True) &:= False\\ (u, False) \subseteq (v, False) &:= v \subseteq u\ .\end{aligned} \tag{10}$$

Figure 3 depicts the subset evaluation between the complementary focal elements with themselves for frame of discernment $\Omega = \{A, B, C\}$. The belief function quantifies the lowest possible degree to which the evidence supports a complementary focal element by summing the masses of all subsets of the complementary focal element that are not empty.

Definition 5 *The belief in complementary focal element y from CBPA m is defined as the following:*

$$\mathrm{bel}_m(y) = \sum_{\substack{x \in 2^\Omega \times \mathbb{B}\\ x \subseteq y\\ x \neq (\emptyset, True)}} m(x)\ . \tag{11}$$

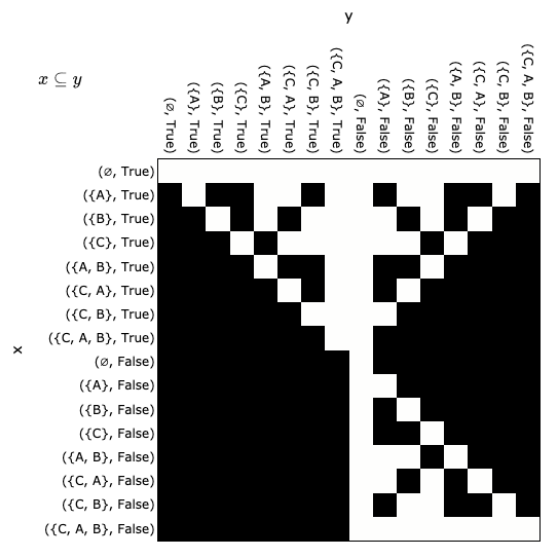


Figure 3: Subset property between complementary focal elements for frame of discernment $\Omega = \{A, B, C\}$.

We order three related belief computations in Theorem 6:

Theorem 6 *For a CBPA m with a frame of discernment Ω and for $u \in 2^\Omega$,*

$$\mathrm{bel}_m((u, True)) \le \mathrm{bel}_{m|_\Omega}(u) \le \mathrm{bel}_m((\Omega - u, False))\ .$$

Plausibility of a focal element is defined as the sum of all masses of focal elements that are not disjoint to that focal element, visualized in Figure 4. It follows that plausibility is the degree to which evidence does not contradict the proposition (or set of propositions) represented by the focal element or in other words, the degree to which the focal element proposition(s) are possible.

Definition 7 *The plausibility of complementary focal element y from CBPA m is defined as the following:*

$$\mathrm{pl}_m(y) = \sum_{\substack{x \in 2^\Omega \times \mathbb{B}\\ x \cap y \neq (\emptyset, True)}} m(x) \tag{12}$$

There is a relation of three plausibility computations described in Theorem 8.

Theorem 8 *For a CBPA m with a frame of discernment Ω and for $u \in 2^\Omega$,*

$$\mathrm{pl}_m((u, True)) = \mathrm{pl}_{m|_\Omega}(u) \le \mathrm{pl}_m((\Omega - u, False))\ .$$

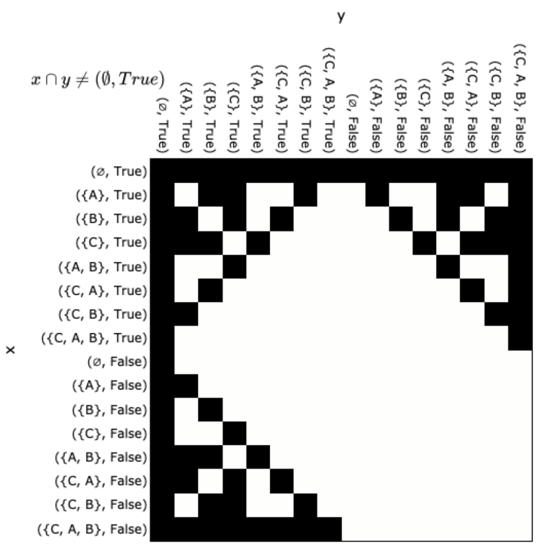


Figure 4: The disjoint property between complementary focal elements for frame of discernment $\Omega = \{A, B, C\}$.

We note that the difference between the relation $\text{bel}_m((u, True)) \leq \text{bel}_{m|_\Omega}(u)$ in Theorem 6 and $\text{pl}_m((u, True)) = \text{pl}_{m|_\Omega}(u)$ in Theorem 8 comes from the differences in behavior of subsets and intersections of complementary focal elements. Subsets of $(u, True)$ will only consist of $True$ components, while both $True$ and $False$ components can have nonempty intersections with $(u, True)$.

A relation between belief and plausibility computations can be seen in Theorem 9 and is an adaptation of previous relationships between belief and plausibility to CBPAs [4, 26, 28].

Theorem 9 *Belief and plausibility are related through the following equation:*

$$m((\emptyset, True)) + \text{bel}((u, \neg a) + \text{pl}((u, a)) = 1\ . \qquad (13)$$

3. Experiments

Here we demonstrate complementary DST on various problems. We illustrate Theorems 6, 8 and 9 on synthetic data. Additionally, we highlight our capabilities on the iris species dataset, and the land coverage identification problem from Sentinel-2 satellite data.

3.1. Synthetic

To demonstrate complementary DST, we create two CBPAs, m_1 with frame of discernment $\{A, B\}$, and m_2 with frame of discernment $\{B, C\}$, seen in the Tables 2(*a*) and 2(*b*). We apply the TBM conjunctive join to compute $m_1 \oplus m_2 = m_{1,2}$, and show the evaluation of $m_{1,2}$ on $\Omega = \{A, B, C\}$ in Tables 2(*c*) and 2(*d*).

Table 2: Nonzero values of CBPAs: m_1, m_2, $m_{1,2}$, and the evaluation of $m_{1,2}$ on $\Omega = \{A, B, C\}$.

(*a*) m_1

Element	Mass
$(\emptyset, False)$	0.3
$(\{A\}, False)$	0.4
$(\{B\}, False)$	0.2
$(\{A, B\}, True)$	0.1

(*b*) m_2

Element	Mass
$(\{B\}, False)$	0.4
$(\{C\}, True)$	0.3
$(\{B, C\}, True)$	0.2
$(\{B, C\}, False)$	0.1

(*c*) $m_{1,2}$

Element	Mass
$(\emptyset, True)$	0.03
$(\{A\}, True)$	0.05
$(\{B\}, True)$	0.02
$(\{B\}, False)$	0.20
$(\{C\}, True)$	0.31
$(\{A, B\}, False)$	0.16
$(\{B, C\}, True)$	0.14
$(\{B, C\}, False)$	0.05
$(\{A, B, C\}, False)$	0.04

(*d*) $m_{1,2}|_\Omega$

Element	Mass
$\emptyset$	0.07
$\{A\}$	0.1
$\{B\}$	0.02
$\{C\}$	0.47
$\{A, C\}$	0.2
$\{B, C\}$	0.14

In this example, the mass in $m_{1,2}((\emptyset, True))$ comes from the combination of the conflicting beliefs of $m_1((\{A, B\}, True) = 0.1$ and $m_2((\{C\}, True) = 0.3$. From the tables it is also clear that unlike GET and TBM, complementary DST differentiates between a missing proposition from the frame of discernment and fundamental conflict. Specifically, in the computation the classic conflict term $m_{1,2}|_\Omega(\emptyset) = m_{1,2}((\emptyset, True)) + m_{1,2}((\{A, B, C\}, False))$, we see that it is the sum of two masses. The mass, $m_{1,2}((\emptyset, True))$ represents the conflict of the CBPAs, and $m_{1,2}((\{A, B, C\}, False))$ represents the conjecture that a proposition is missing.

The belief and plausibility intervals of $m_{1,2}$ and $m_{1,2}|_\Omega$ where $\Omega = \{A, B, C\}$ are shown in Figure 5. Three intervals are shown for each focal element $u \in 2^\Omega$: $[\text{bel}_m((u, True)), \text{pl}_m((u, True))]$ in red, $[\text{bel}_{m|_\Omega}(u), \text{pl}_{m|_\Omega}(u)]$ in green, and $[\text{bel}_m((\Omega - u, False)), \text{pl}_m((\Omega - u, False))]$ in blue. Our interpretation of these intervals is that red represents the possible probabilities of a specific focal element being true, green represents these probabilities assuming that the current frame of discernment is complete, and blue represents the feasible

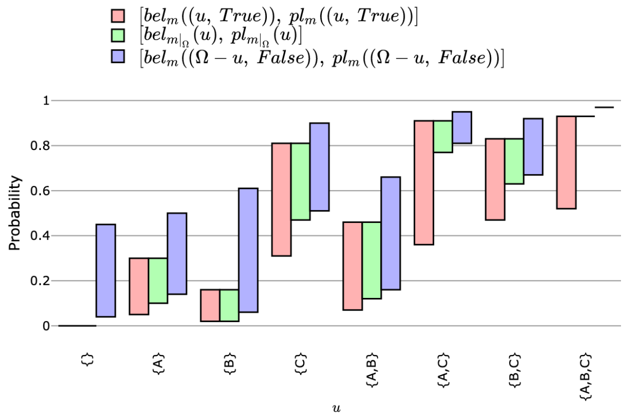
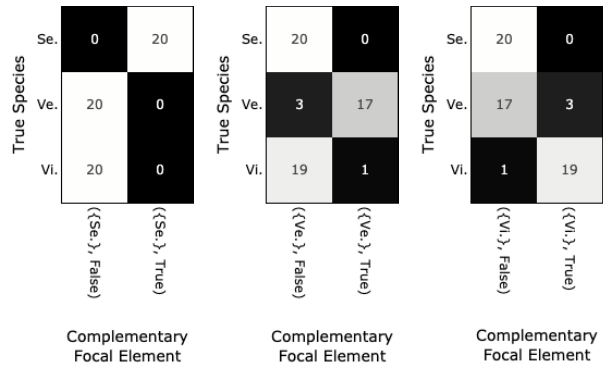


Figure 5: An example of belief and plausibility for the complementary approach to Dempster-Shafer Theory.



$(a)\ m_{Se}$ $(b)\ m_{Ve}$ $(c)\ m_{Vi}$

Figure 6: Confusion matrices of CBPAs for each class

probabilities that the complement of each focal element is not true. Some insight can be gained from comparing the red and blue intervals. For example, with $\Omega = \{A, B, C\}$ an apparent inconsistency of high confidence in $(\{C\}, False)$ and simultaneously low confidence is $(\{A, B\}, True)$ can be reconciled with a missing proposition D. In Figure 5, we can clearly see the relationship between beliefs and plausibilities described in Theorems 6 and 8. Additionally, we can see the relation between belief and plausibility described in Theorem 9 with the value of $m_{1,2}((\emptyset, True))$.

3.2. Iris Species Classification

The iris dataset is a well known labeled dataset consisting of 3 classes, each containing 50 samples [10]. Each sample is a collection of 4 iris measurements: the sepal length, sepal width, petal length, and petal width, and is associated with one of the three species of iris: Setosa, Versicolor, or Virginica. We abbreviate Setosa, Versicolor, and Virginica as Se, Ve, and Vi, respectively.

To highlight the ability of CBPAs to work with incomplete frames of discernment, we construct CBPAs from One-Class Support Vector Machines (OneClassSVM) trained individually on each class from the iris dataset. No classifier was trained on samples from more than one class. We use 60%, 40% train-test split to get 30 train and 20 test samples of each class. On the training data, we fit a OneClassSVM with the parameters, $\nu = 0.1$, for each class of iris individually. For each of the test samples, indexed by i, and for each class, c, we predict the class membership as a Boolean

$$p_c^i = \begin{cases} 1 \text{ if sample i predicted to belong to class } c \\ 0 \text{ if sample i predicted to not belong to class } c \ . \end{cases}$$

For each prediction we construct a CBPA with the values,

$$m_c^i((\{c\}, True)) = p_c^i \tag{14}$$

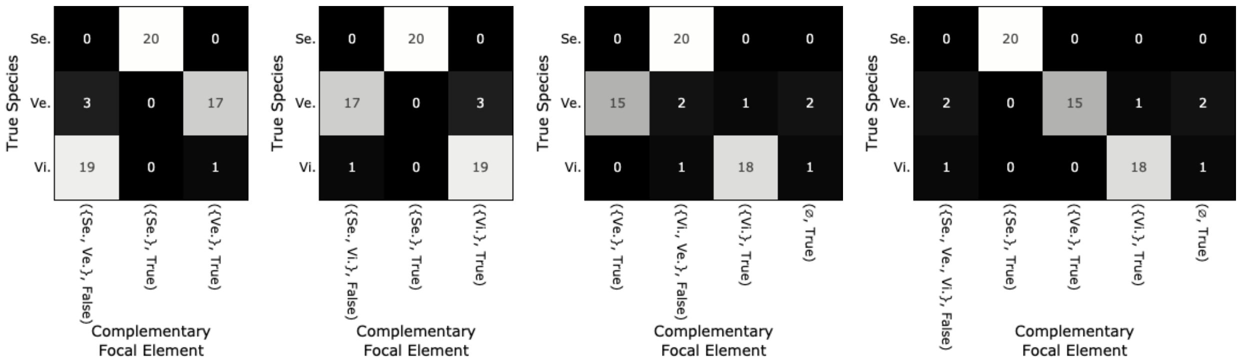
$$m_c^i((\{c\}, False)) = 1 - p_c^i \ . \tag{15}$$

We elected to assign mass to $(\{c\}, False)$ rather than $(\emptyset, False)$ to represent the confidence that our training and testing datasets come from the same distribution and the training dataset should cover all cases of the class.[1]

Figure 6 depicts confusion matrices showing the agglomerated results over the test samples for each class. The OneClassSVM trained on Setosa in Figure 6(a) was able to perfectly place all test samples. The two classifiers independently trained on Versicolor in Figure 6(b) and Virginica in Figure 6(c) performed perfectly on Setosa but both miscategorized three test samples of Versicolor and one test sample from Virginica.

In Figure 7 we see the confusion matrix results of different conjunctive join combinations of the CBPAs seen in Figure 6. While the results in Figures 7(a) and 7(b) are mostly straightforward, the combination $m_{Versicolor,Virginica}$ in Figure 7(c) makes an important distinction between ignorance and conflict with Versicolor and Virginica OneClassSVMs. With the 20 Versicolor test samples, 15 were correctly classified, 2 were put into $(\{Versicolor, Virginica\}, False)$ meaning neither classifier claimed those samples as their species (*ignorance*), 1 sample was misidentified as Virginica, and 2 samples were assigned to the conflict element, $(\emptyset, True)$, meaning that both classifiers claimed those samples as their species. Similarly for the Virginica samples, 18 were correctly placed, 1 neither classifier claimed (*ignorance*), and 1 both classifiers claimed (*conflict*).

[1] If one does not have confidence that the training and testing datasets come from the same distributions, one could make the argument that a negative result does not indicate that the sample is not from the training class, and therefore the mass should be assigned to a truly vacuous statement, $(\emptyset, False)$.



(a) $m_{Se,Ve}$ (b) $m_{Se,Vi}$ (c) $m_{Ve,Vi}$ (d) $m_{Se,Ve,Vi}$

Figure 7: Confusion matrices of conjunctive joins of CBPAs.

The addition of a new proposition, $\{Setosa\}$, to the frame of discernment $\{Versicolor, Virginica\}$is demonstrated in the transition between Figures 7(*c*) and 7(*d*). The migration of the 20 Setosa samples from the complementary focal element $(\{Versicolor, Virginica\}, False)$ in Figure 7(*c*) to the complementary focal element $(\{Setosa\}, True)$ in Figure 7(*d*) does not require any modification or recalculation of the previous CBPAs. Additionally, all samples still not claimed by any classifier appear in the newly introduced complementary focal element $(\{Setosa, Versicolor, Virginica\}, False)$. Mass in the complementary focal element for the known frame of discernment can indicate the need for more sophisticated classifiers or that the current frame of discernment is incomplete. In either event, this is a quantified indication of the known unknowns.

3.3. Sentinel-2 Land Cover Classification

Here, we illustrate our complementary DST on the task of land cover classification of Sentinel-2 imagery. The Sentinel-2 constellation from the Copernicus program is operated by the European Space Agency [21]. Each Sentinel-2 satellite carries a multi-spectral instrument for acquiring high spatial resolution optical imagery. The level 1C products contain Top Of Atmosphere (TOA) reflectances of 13 bands summarized in Table 3. In this experiment, we use a scene from Anchorage, Alaska, seen in Figure 8, taken by Sentinel-2A on Aug. 2, 2021 which has been downsampled to a 180m resolution.

Remote sensing subject matter experts take advantage of different reflectance properties of materials to develop simple methods to accurately identify different land covers. Some of the most successful methods rely on the normalized difference,

$$\mathrm{ND}(X_1, X_2) = \frac{X_1 - X_2}{X_1 + X_2}\ . \tag{16}$$

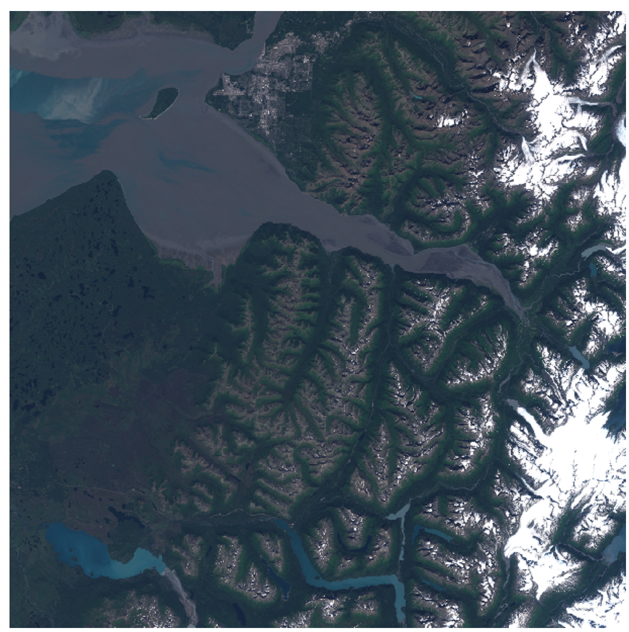
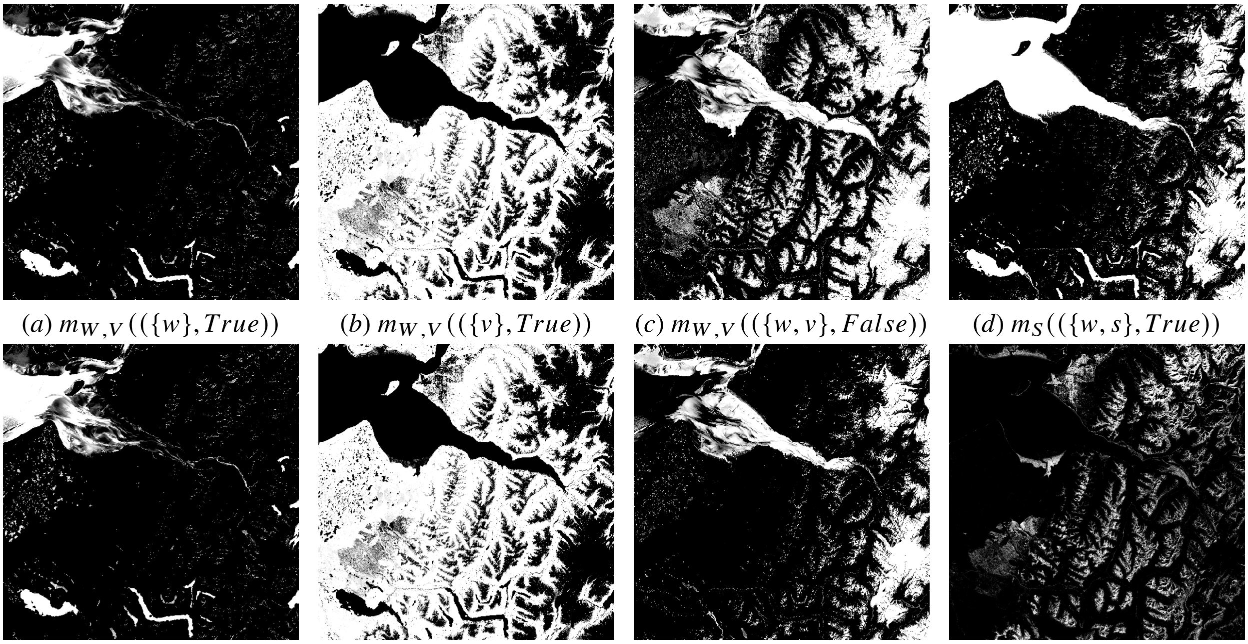


Figure 8: True color image of Anchorage, Alaska downsampled to 180m resolution taken by Sentinel-2A on Aug. 2, 2021.



(a) $m_{W,V}((\{w\},True))$ (b) $m_{W,V}((\{v\},True))$ (c) $m_{W,V}((\{w,v\},False))$ (d) $m_S((\{w,s\},True))$

(e) $m_{W,V,S}((\{w\},True))$ (f) $m_{W,V,S}((\{v\},True))$ (g) $m_{W,V,S}((\{s\},True))$ (h) $m_{W,V,S}((\{w,v,s\},False))$

Figure 9: Visualization of masses for specific focal elements from CBPAs constructed from Sentinel-2 imagery. Black represents a mass of 0, white of 1. There was little mass in the conflict focal elements, $(\emptyset, True)$, so they are omitted for space.

Table 3: Sentinel-2 level 1C band names, descriptions, approximate wavelengths, and resolutions.

Band	Description	Wav. (nm)	Res. (m)
B1	Coastal Aerosol	442	60
B2	Blue	492	10
B3	Green	559	10
B4	Red	665	10
B5	Veg. red edge	704	20
B6	Veg. red edge	740	20
B7	Veg. red edge	781	20
B8	NIR	833	10
B8A	Narrow NIR	864	20
B9	Water vapour	944	60
B10	SWIR - Cirrus	1375	60
B11	SWIR	1612	20
B12	SWIR	2194	20

The normalized differences between different bands for each pixel provides a value between -1 and 1 whose relative weight is an indicator of the presence of one or multiple materials. The Normalized Difference Water Index (NDWI) [17] and Normalized Difference Vegetation Index (NDVI) [19] are used to indicate water and vegetation respectively. The Normalized Difference Snow Index (NDSI) [24] can indicate snow, clouds, and water depending on selected thresholds [1, 15]. We abbreviate NDWI, NDVI, and NDSI as W, V, and S, and water, vegetation, and snow, as w, v, and s.

Table 4: The normalized difference method names, material classes, bands, and thresholds used to compute CBPAs.

Method	Materials	X_1	X_2	T_l	T_u
NDWI	$\{water\}$	B3	B8A	0.3	0.5
NDVI	$\{vegetation\}$	B8A	B4	0.2	0.5
NDSI	$\{water, snow\}$	B3	B11	-0.1	0.2

We apply the normalized difference methods to generate CBPAs and use our complementary DST to do data fusion between CBPAs. Using the values in Table 4, we create

CBPAs by clipping the values,

$$m_{Method}((Materials, True)) = \frac{\mathrm{ND}(X_1, X_2) - T_l}{T_u - T_l}$$
$$m_{Method}((Materials, False)) = 1 - \frac{\mathrm{ND}(X_1, X_2) - T_l}{T_u - T_l}, \tag{17}$$

to [0, 1] for each different normalized difference method. Notice that the frames of discernment of NDVI do not have any common elements with either NDWI or NDSI. Additionally, the frame of discernment of NDWI is a subset of NDSI's. These differences in frames of discernment pose no problem in joining these CBPAs together as they do not need to share a common frame of discernment.

In Figures 9(*a*)-9(*c*) we see the complementary focal element values from $m_{W,V}$ for $(\{water\}, True)$, $(\{vegetation\}, True)$, and $(\{water, vegetation\}, False)$ respectively. The identification of water and vegetation is relatively straightforward. The significant presence in $(\{water, vegetation\}, False)$ indicates that the frame of discernment $\{water, vegetation\}$ is insufficient to describe this scene. Comparison to the TCI suggests that the pixels in $(\{water, vegetation\}, False)$ are correctly categorized as neither $water$ nor $vegetation$, and that perhaps a proposition $\{snow\}$ is missing.

We systematically add the $snow$ proposition with the conjunctive joining of the m_S CBPA shown in Figure 9(*d*) to $m_{W,V}$. The resulting CBPA, $m_{W,V,S}$, is displayed in Figures 9(*e*)-9(*h*). We see insignificant modifications to the $water$ and $vegetation$ masses in Figures 9(*e*) and 9(*f*). The masses in Figure 9(*c*) have been distributed between snow in Figure 9(*g*), and still unknown propositions in Figure 9(*h*).

The significant mass in Figure 9(*h*) suggests the frame of discernment $\{water, vegetation, snow\}$ is still incomplete. Visual comparison with the TCI suggests that $urban$, $rocks$, and $minerals$ may be suitable additions to the frame of discernment. Using our complementary DST framework, subject matter experts could devise CBPAs from new sources of evidence that can further help to identify the unclassified features without needing to update all previous CBPAs with the new propositions.

Frequently more refined material classifications are sought than that of $\{water, vegetation, snow\}$. In practice, imperfect and sensitive thresholding systems are implemented to attempt identification of individual materials. With CBPAs, if various methods allows for discernment between different groups of materials, then the CBPAs can be joined together to provide a consistent picture. This provides the advantage of being able to choose less sensitive thresholds to discriminate groups of materials, rather than individual materials. Additionally, the joined CBPAs automatically distinguish between conflict and pixels that are left unidentified.

4. Conclusion

In this work we propose a new approach for open-world DST based on complements of sets that we call complementary DST. We provide the generalized definitions and theorems necessary to characterize complementary DST. Experimental results indicate that complementary DST can identify when a frame of discernment is incomplete, join together CBPAs that are not defined on the same frame of discernment, and systematically added propositions to a frame of discernment. Complementary DST can distinguish between *conflict* in sources of information and *ignorance* of a possible outcome. Conflict can be resolved by replacing inconsistent information where ignorance can be resolved by considering new propositions from an open world.

Acknowledgments

We would like to thank our colleagues Dax Gerts, Judith Cohn, and Galina Rogova who provided useful feedback for framing this discussion. We would like to thank the reviewers for their insightful comments. The research presented in this article was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number 20220799DI.

References

[1] Mario Arreola-Esquivel, Carina Toxqui-Quitl, Maricela Delgadillo-Herrera, Alfonso Padilla-Vivanco, Gabriel Ortega-Mendoza, and Anna Carbone. Non-binary snow index for multi-component surfaces. *Remote Sensing*, 13(14):2777, 2021.

[2] Isabelle Bloch. Some aspects of Dempster-Shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account. *Pattern Recognition Letters*, 17:905–919, 1996.

[3] Francois Delmotte and Philippe Smets. Target identification based on the transferable belief model interpretation of Dempster-Shafer model. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 34(4):457–471, 2004.

[4] Arthur P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Statistics*, 28:325–339, 1967.

[5] Yong Deng. Generalized evidence theory. *Applied Intelligence*, 43(4):530–543, 2015.

[6] Didier Dubois and Henri Prade. On the combination of evidence in various mathematical frameworks. *Reliability data collection and analysis*, pages 213–241, 1992.

[7] A. Appriou F. Janez. Theory of evidence and non-exhaustive frames of discrenment: Plausibilities correction methods. *International Journal of Approximate Reasoning*, 18:1 – 19, 1998.

[8] S. Ferson, V. Kreinovich, L. Ginzburg, and K. Sentz. Constructing probability boxes and Dempster-Shafer structures. Technical report, Sandia National Lab., Albuquerque, NM, United States, 2003.

[9] Scott Ferson, Mark Burgman, Lloyd Goldwasser, Scott Ferson, and Lev Ginzburg. Variability and measurement error in extinction risk analysis: the northern spotted owl on the Olympic Peninsula. *Quantitative methods for conservation biology*, pages 169–187, 2000.

[10] Ronald A Fisher. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2): 179–188, 1936.

[11] J. C. Helton, J.D. Johnson, W.L. Oberkampf, and C.B. Storlie. A sampling-based strategy for the representation of epistemic uncertainty in model predictions with evidence theory. Technical report, Sandia National Laboratories, SAND2006-5557, 2006.

[12] Wen Jiang, Yue Chang, and Shiyu Wang. A method to identify the incomplete framework of discernment in evidence theory. *Mathematical Problems in Engineering*, 2017, 2017.

[13] G. J. Klir. *Uncertainty and Information*. Wiley, 2006.

[14] Huapeng Li, Shuqing Zhang, Yan Sun, and Jing Gao. Land cover classification with multi-source data using evidential reasoning approach. *Chinese Geographical Science*, 21:312–321, 2011.

[15] Muyi Li, Xiufang Zhu, Nan Li, and Yaozhong Pan. Gap-filling of a MODIS Normalized Difference Snow Index product based on the similar pixel selecting algorithm: a case study on the Qinghai–Tibetan Plateau. *Remote Sensing*, 12(7):1077, 2020.

[16] Fan Liu and Yong Deng. Determine the number of unknown targets in open world based on elbow method. *IEEE Transactions on Fuzzy Systems*, 29(5):986–995, 2020.

[17] Stuart K McFeeters. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International journal of remote sensing*, 17(7):1425–1432, 1996.

[18] W. L. Oberkampf and C. J. Roy. *Verification and Validation in Scientific Computing*. Cambridge University press, 2010.

[19] Nathalie Pettorelli, Jon Olav Vik, Atle Mysterud, Jean-Michel Gaillard, Compton J Tucker, and Nils Chr Stenseth. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in ecology & evolution*, 20(9):503–510, 2005.

[20] Frédéric Pichon, Didier Dubois, and Thierry Denœux. Quality of information sources in information fusion. *Information Quality in Information Fusion and Decision Making*, pages 31–49, 2019.

[21] Copernicus Sentinel-2 (processed by ESA). Sentinel-2 MSI level-1c TOA reflectance, 2021. URL https://doi.org/10.5270/s2_-742ikth.

[22] Haenni R. and S. Hartmann. Modelling partially reliable information sources: a general approach based on Dempster-Shafer theory. *Information Fusion*, 7(4): 361–379, 2006.

[23] Cyril Royère, Dominique Gruyer, and Véronique Cherfaoui. Data association with believe theory. In *Proceedings of the Third International Conference on Information Fusion*, volume 1, pages TUD2–3. IEEE, 2000.

[24] Vincent V Salomonson and I Appel. Estimating fractional snow cover from MODIS using the normalized difference snow index. *Remote sensing of environment*, 89(3):351–360, 2004.

[25] Johan Schubert. Constructing and evaluating alternative frames of discernment. *International Journal of Approximate Reasoning*, 53(2):176–189, 2012.

[26] Glenn Shafer. *A mathematical theory of evidence*, volume 42. Princeton university press, 1976.

[27] Philippe Smets. The combination of evidence in the transferable belief model. *IEEE Transactions on pattern analysis and machine intelligence*, 12(5): 447–458, 1990.

[28] Philippe Smets. The nature of the unnormalized beliefs encountered in the transferable belief model. In *Uncertainty in artificial intelligence*, pages 292–297. Elsevier, 1992.

[29] Philippe Smets. Analyzing the combination of conflicting belief functions. *Information fusion*, 8(4): 387–412, 2007.

[30] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, 1991.

[31] Yang Wenran, Li Xinde, and Deng Yong. A clustering based method to complete frame of discernment. *Chinese Journal of Aeronautics*, 2022.

[32] Y. Tang X. Zhou. A note on incomplete information modeling in the evidence theory. *IEEE Access*, 7: 166410 – 166414, 2019.

[33] L.A. Zadeh. Review of Books: A Mathematical Theory of Evidence. *AI Magazine*, 5:81–83, 1984.

[34] L.A. Zadeh. A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination. *AI Magazine*, 7:85–90, 1986.