WCCI 2012 IEEE World Congress on Computational Intelligence
June, 10-15, 2012 - Brisbane, Australia FUZZ IEEE

Sugeno Fuzzy Integral Generalizations for Sub-Normal Fuzzy Set-Valued Inputs

D. T. Anderson
Mississippi State University
Electrical and Computer Engineering
Mississippi State, MS 39759
Email: anderson@ece.msstate.edu

T. C. Havens
Michigan State University
Computer Science and Engineering
East Lansing, MI 48824
Email: havenst@gmail.com

C. Wagner
University of Nottingham
School of Computer Science
Nottingham, United Kingdom
Email: christian.wagner@nottingham.ac.uk

J. M. Keller
University of Missouri
Electrical and Computer Engineering
Columbia, MO 65202
Email: kellerj@missouri.edu

M. F. Anderson
Starkville, MS 39759
Email: melissafanderson@gmail.com

D. J. Wescott
Texas State University
Department of Anthropology
San Marcos, TX 78666
Email: dwescott@txstate.edu

***Abstract*—In prior work, Grabisch put forth a direct (i.e., result of the Extension Principle) generalization of the Sugeno *fuzzy integral* (FI) for *fuzzy set* (FS)-valued normal (height equal to one) integrands and number-based *fuzzy measures* (FMs). Grabisch's proof is based in large on Dubois and Prade's analysis of functions on intervals, fuzzy numbers (thus normal FSs) and fuzzy arithmetic. However, a case not studied is the extension of the FI for sub-normal FS integrands. In prior work, we described a real-world forensic application in anthropology that requires fusion and has sub-normal FS inputs. We put forth an alternative non-direct approach for calculating FS results from sub-normal FS inputs based on the use of the number-valued integrand and number-valued FM Sugeno FI. In this article, we discuss a direct generalization of the Sugeno FI for sub-normal FS integrands and numeric FMs, called the *Sub-normal Fuzzy Integral* (SuFI). To no great surprise, it turns out that the SuFI algorithm is a special case of Grabisch's generalization. An algorithm for calculating SuFI and its mathematical properties are compared to our prior method, the *Non-Direct Fuzzy Integral* (NDFI). It turns out that SuFI and NDFI fuse in very different ways. We assert that in some settings, e.g., skeletal age-at-death estimation, NDFI is preferred to SuFI. Numeric examples are provided to stress important inner workings and differences between the FI generalizations.**

***Index Terms*—Sugeno fuzzy integral; SuFI; NDFI; fuzzy set valued integrands; extension principle**

I. Motivation

To date, the vast majority of research in the field of *fuzzy integrals* (FIs) is predominately focused on topics involving numeric integrands and numeric *fuzzy measures* (FMs). However, a few works have appeared regarding the generalization of both the integrand as well as the FM with respect to *fuzzy sets* (FSs) [1–9]. While these generalizations can be applied to a variety of cases, they have not yet been specifically applied to sub-normal FS integrands. In this work, we discuss an application for which sub-normal FS-valued inputs exist and need to be fused. We demonstrate and discuss the impact of using a direct (i.e., result of the Extension Principle [10]) generalization of the FI, which we call the *Sub-normal Fuzzy Integral* (SuFI), in relation to our prior non-direct approach, the *Non-Direct Fuzzy Integral* (NDFI) [11, 12]. While we investigate SuFI and NDFI in the context of skeletal anthropology, the analysis and methods put forth herein are not restricted to age-at-death estimation. One could also imagine many other applications where fusion of FS inputs is necessary.

Age-at-death estimation of an individual skeleton is important to forensic and biological anthropologists for identification and demographic analysis. It has been shown that current individual aging methods are often unreliable because of skeletal variation and taphonomic factors [11]. Previously, we introduced the NDFI algorithm as an alternative way to estimate adult skeletal age-at-death [11]. In particular, focus was placed on the production of numeric [11], graphical [11, 12] and linguistic descriptions of age-at-death [12]. The NDFI algorithm takes as input multiple age-range intervals representing age-at-death estimations from different methods. It also takes into account the accuracies of these methods as well as the condition of the bones being examined. Advantages of NDFI, relative to related work in forensic anthropology, are that it does not require a skeletal population and it produces additional information (numeric, graphical and linguistic) that can assist an investigator. A formal description of NDFI is included in Section VI.

In [12], we presented a way to measure the uncertainty present in a FS produced by NDFI. Specifically, we demonstrated a way to generate linguistic descriptions in order to establish domain standardization for the goal of assisting forensic and biological anthropologists. To achieve this goal, we extracted features from FSs, introduced fuzzy class definitions for age-at-death FSs, and we put forth an *ordered weighted average* (OWA) contrast operator to measure specificity in age-at-death FSs [12].

Before proceeding, it is important to highlight the following. The NDFI algorithm is not a direct extension in the same regard as Grabisch's [1, 2] (or even the SuFI algorithm put forth herein). Initially, we were not aware of higher order extensions to the FI. NDFI was put forth to address a specific problem in anthropology. Indirectly, it provided a *work around* for the lack of a direct extension. Following the publication of NDFI, we discovered that Grabisch had previously shown how one can formally extend the FI for FSs, specifically normal, integrands and numeric FMs. This article is a review of Grabisch's work, the NDFI algorithm, development and demonstration of a direct extension of the FI for the case of sub-normal FS integrands (SuFI), and most importantly, discussion and comparison between these approaches.

The remainder of the article is organized as follows. First, we review Sugeno's number-based (integrand and FM) FI. Grabisch's generalization is then discussed for interval-valued as well as normal FS integrands. Next, we investigate a direct generalization for the case of sub-normal FS integrands (SuFI), followed by a review of NDFI. In closing, we present and remark on a few examples and important properties of these different generalizations.

II. Sugeno's Number-Based Fuzzy Integral

The fusion of information using the *classical* FI (Sugeno or Choquet) has a rich history. Much of the theory and several applications can be found in [2, 13]. With respect to this problem, we consider a finite set of sources of information $X = \{x_1, ..., x_E\}$ and a function that maps X into some domain (initially $[0,1]$) that represents the partial support of a hypothesis from the standpoint of each source of information. Depending on the problem domain, X can be a set of experts, sensors, features, pattern recognition algorithms, etc. In our prior work [11, 12], X is a set of methods that help determine the age-at-death of a person from their skeletal remains (e.g., Todd's method for the pubic symphysis [14]). The hypothesis is usually thought of as an alternative in a decision process or a class label in pattern recognition. In age-at-death analysis, a hypothesis is that the individual died at a specific age. Both Sugeno and Choquet integrals take partial support for the hypothesis from the standpoint of each source of information and fuse it with the (perhaps subjective) worth (or reliability) of each subset of X in a non-linear fashion. This worth is encoded into a FM [15]. Initially, the function $h : X \rightarrow [0,1]$ and the FM $g : 2^X \rightarrow [0,1]$ took real number values in $[0,1]$. Certainly, the output range for both function and FM can be (and have been) defined more generally, but it is convenient to think of them in the unit interval for confidence fusion.

More formally, for a finite set X, a FM is a function $g : 2^X \rightarrow [0,1]$, such that

1. $g(\phi) = 0$ and $g(X) = 1$;
2. If $A, B \subseteq X$ with $A \subseteq B$, then $g(A) \leq g(b)$.

Note, that if X is an infinite set, a third condition guaranteeing continuity is required, but this is a moot point for finite X. Given a finite set X, a FM g and a function h, the (*numeric*) Sugeno FI of h with respect to g is

$$\int_S h \circ g = \bigvee_{i=1}^{E} \left(h(x_{(i)}) \wedge g(x_{(1)}, ..., x_{(i)})\right), \tag{1}$$

where X has been sorted so that

$$h(x_{(1)}) \geq h(x_{(2)}) \geq ... \geq h(x_{(E)}). \tag{2}$$

This finite realization of the actual definition highlights that the Sugeno integral represents the best pessimistic agreement between the objective evidence in support of a hypothesis (the h function) and the (perhaps) subjective worth of the supporting evidence (the FM g). The FM can be specified using only the densities via the Sugeno λ-FM [15] or it can be learned from training data, e.g. [16].

III. FS-Valued Normal Integrands

Sometimes numbers are not sufficient to represent the uncertainty in a situation. With respect to fusion by fuzzy integration, this uncertainty can exist in the partial support function and/or in the FM. Extensions of both Sugeno and Choquet integrals to the case where the partial support function outputs are fuzzy numbers (normal convex fuzzy subsets of the reals, $\Re$, called $FN(\Re)$) are direct results of the Extension Principle [10]. They are computable from level set representations using the Decomposition Theorem and methods from [1, 2]. Interval logic and arithmetic operations make the extension possible in a practical sense. This works because the theory that shows that the level sets of the generalized fuzzy integral reduce to the fuzzy integrals of the endpoints of the intervals that form the level cuts of fuzzy numbers.

Let $I(\Re^+) = \{\bar{u} \subset \Re^+ | \bar{u} = [u^l, u^r], u^l \leq u^r\}$ be the set of all closed intervals over the positive reals. Dubois and Prade showed that if a function φ is continuous and non-decreasing, then when defined on intervals it produces an interval whose endpoints are equal to the function values on the lower bound and upper bound of the individual intervals, i.e., $\varphi(\bar{u}) = [\varphi(u^l), \varphi(u^r)]$ [3]. Dubois and Prade also showed how φ extends to FS inputs, specifically normal, convex FSs (fuzzy numbers). The Choquet and generalized Sugeno FIs are continuous, non-decreasing functions. Grabisch leveraged these properties and Dubois's and Prade's work in order to extend the Choquet and generalized Sugeno integrals as follows. Let $\bar{H} : X \rightarrow I(\Re^+)$ denote the interval-valued partial support function. Additionally, let $\bar{H}_i = \bar{H}(x_i) = [\bar{H}_i^l, \bar{H}_i^r]$ denote the i^th interval (where $\bar{H}_i^l$ and $\bar{H}_i^r$ are the left and right interval endpoints respectively). The generalized interval Sugeno FI is defined as

$$\int_I \bar{H} \circ g = [\int_S \bar{H}^l \circ g, \int_S \bar{H}^r \circ g]. \tag{3}$$

Now, let $H : X \rightarrow FN(\Re)$ denote the FS partial support function and $H_i = H(x_i)$ the i^{th} FS. Additionally, let $[H_i]_\alpha = [(H_i)_\alpha^l, (H_i)_\alpha^r]$ for $0 \leq \alpha \leq 1$. The generalized Sugeno FI for normal FS integrands is

$$\int_{NFI} H \circ g = \bigcup_{\alpha \in [0,1]} \alpha [\int_{NFI} H \circ g]_\alpha \tag{4}$$

$$= \bigcup_{\alpha \in [0,1]} \alpha [\int_I H_\alpha \circ g], \tag{5}$$

which can be efficiently calculated on a computer in terms of α-cut interval operations (eq. 3). Algorithm 1 is the computational method to calculate the FI for FS (normal) integrands and number-based FMs.

IV. SuFI

Grabisch's extension covers a wide range of scenarios one might encounter in practice. However, it does not address the case of sub-normal FS integrands. Grabisch's proof is based on the fuzzy arithmetic for (normal) fuzzy numbers work of Dubois and Prade. This might explain in part the lack of formal procedure to date for sub-normal FSs. However, in our anthropology work we have sub-normal FSs and therefore a motivation to study such cases.

The purpose of this section is to study the direct extension of the Sugeno FI for sub-normal FS integrands. Using the Extension Principle, we see that

$$(\int H \circ g)(y) = \sup_{(z_1,...,z_E) \in S_y} \{(H_1)(z_1) \wedge ... \wedge (H_E)(z_E)\}, \tag{6}$$

$$S_y = \{(z_1, ..., z_E) | z_1, ..., z_E \in \Re, \int h_{(z_1,...,z_E)} \circ g = y\}, \tag{7}$$

where $h_{(z_1,...,z_E)}$ is the partial support function,

$$h(x_1) = z_1, ..., h(x_E) = z_E. \tag{8}$$

The first observation is that the extended function is bounded (the height of the resultant FS) by a constant β,

$$\text{Height}(\int H \circ g) \leq \beta, \tag{9}$$

with respect to

$$(H_1)(z_1) \wedge ... \wedge (H_E)(z_E) \leq \beta, \tag{10}$$

or specifically,

$$\beta = \min\{\text{Height}(H_1), ..., \text{Height}(H_E)\}, \tag{11}$$

where, for a fuzzy set A, with membership function μ_A,

$$\text{Height}(A) = \max_{x_i \in \Re}(\mu_A(x_i)), \tag{12}$$

and A is called *normal* if $\text{Height}(A) = 1$.

Equations 9-11 show that the height of the FI result is based on the t-norm of the heights of the FSs in our partial support function. Therefore, the result is sub-normal if at least one input is sub-normal. While this turns out to have a *drastic* impact for sub-normal FS integrands (mathematically as well as conceptually), it is not problematic for normal FS integrands (i.e., $\beta = 1$). In the case of at least one sub-normal FS input, the level-cuts for all $\beta < \alpha \leq 1$ are the source of the problem. If one attempts to use the level cut/interval representation and Decomposition Theorem approach of Grabisch, Dubois and Prade, then different level cuts possess different numbers of inputs/information sources. For example, consider equation 5 and any $\alpha > \beta$. H^l_α and H^r_α are number-valued partial support functions for the left and right endpoints of the intervals of the FS partial support function H at level cut α. However, there exists at least one j such that $(H_j)_\alpha = \phi$. Additionally, g is the FM for all X. While it might be *natural* to attempt to interpret and perform calculation using only the valid subset of inputs (whose α-cuts are not ϕ), such an approach leads to FSs that fail the vertical line test (see Fig. 1). What one can extract from the Extension Principle is

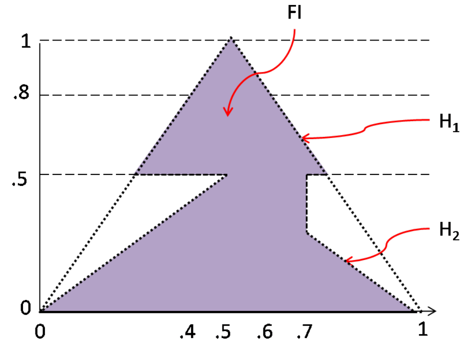


Fig. 1. Example showing the violation of the vertical line test if one attempts to use only the non-empty set of information sources at level cuts greater than β. Example is for $H_1 = [0, 0.5, 1]$ with $Height(H_1) = 1$ and $H_2 = [0, 0.5, 1]$ with $Height(H_2) = 0.5$ (two triangular membership functions) and $g^1 = g(x_1) = g^2 = .7, g(X) = 1$. The resulting FS is shown in purple.

$$[\int H \circ g]_{\alpha > \beta} = \phi. \tag{13}$$

This leads us to a definition of SuFI,

$$\int_{SuFI} H \circ g = \bigcup_{\alpha \in [0,1]} \alpha [\int_{SuFI} H \circ g]_\alpha \tag{14}$$

$$= (\bigcup_{a \in [0,\beta]} a[\int_{SuFI} H \circ g]_a) \cup (\bigcup_{b \in (\beta,1]} b[\int_{SuFI} H \circ g]_b), \tag{15}$$

$$= \bigcup_{a \in [0,\beta]} a[\int_{SuFI} H \circ g]_a, \tag{16}$$

which, like Grabisch's NFI, can be efficiently calculated in terms of interval-valued FI operations,

$$[\int_{SuFI} H \circ g]_\alpha = [\int_S H^l_\alpha \circ g, \int_S H^r_\alpha \circ g]. \tag{17}$$

In fact, this is what Grabisch showed (equation 4) when $\beta = 1$ (i.e., each input FS is normal). Algorithm 2 is the way to calculate SuFI.

Algorithm 1 Computation of the NFI algorithm

Input the fuzzy measure g $\triangleright$ *use the Sugeno λ-fuzzy measure, learn g from data or manually specify g*
Input partial support function H $\triangleright$ *i.e., $H(x_e) = H_e$ and $H(x_e) \in FN(\Re)$*
Select B α-cuts, A = $\{\alpha_1 = 1/B, \alpha_2 = 2/B, ..., \alpha_B = 1\}$
for each $\alpha_i \in A$ **do**
  Calculate $[\int_{NFI} H \circ g]_{\alpha_i} = [\int_S H^l_{\alpha_i} \circ g, \int_S H^r_{\alpha_i} \circ g]$ $\triangleright$ *the number-based (integrand and measure) fuzzy integral*
end for

Algorithm 2 Computation of the SuFI algorithm

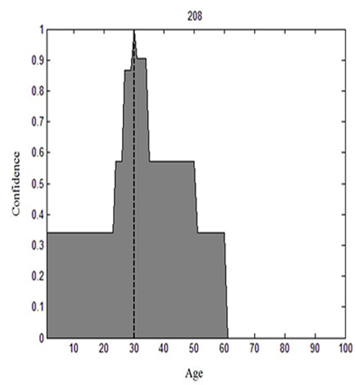
Input the fuzzy measure g $\triangleright$ *use the Sugeno λ-fuzzy measure, learn g from data or manually specify g*
Input partial support function H $\triangleright$ *i.e., $H(x_e) = H_e$ and $H(x_e)$ is a sub-normal FS*
Calculate $\beta = minimum\{Height(H_1), ..., Height(H_E))\}$ $\triangleright$ *Mininum height of partial support FSs*
Select B α-cuts, A = $\{\alpha_1 = \beta/B, \alpha_2 = (2\beta)/B, ..., \alpha_B = \beta\}$
for each $\alpha_i \in A$ **do**
  Calculate $[\int_{SuFI} H \circ g]_{\alpha_i} = [\int_S H^l_{\alpha_i} \circ g, \int_S H^r_{\alpha_i} \circ g]$ $\triangleright$ *the number-based (integrand and measure) fuzzy integral*
end for

V. Interpretation of SuFI

We assert that SuFI is an extremely *limiting* generalization. Consider a sensor fusion scenario in which three different sources (e.g., radar, infrared and visual spectrum) are being aggregated using the FI. Imagine that one of the sources, for example radar, turns out to very unreliable. Now, consider that the radar is assigned a very small density, e.g., 0.1, relative to 1 and 0.8 for the infrared and visual spectrum sources. If infrared and visual spectrum both have a FS input of *near* 1 (e.g., a triangular membership function $[0.9, 1, 1.1]$) and radar has a FS input *near* 0 (e.g., $[-0.1, 0, 0.1]$), the SuFI algorithm outcome is devastating to the fusion result. Intuitively, we would expect that because the radar has very little relative worth, i.e., a density value of 0.1, that the radar decision would influence the decision result very little. However, the height of the resultant set is bounded by β, which is 0.1 in this scenario. The point is, SuFI provides a way to calculate a result; however, this result is not intuitively pleasing in some circumstances. For the provided sensor fusion example, one should intuitively ignore the radar input based on the SuFI algorithm result. Next, we review the NDFI algorithm.

VI. NDFI

In [11], we present an alternative, non-direct way of generating FS results from the number-based (integrand and FM) FI for sub-normal FS inputs (called the NDFI). Our age-at-death NDFI procedure takes interval-valued inputs, e.g., 'method 1 says that the skeleton is between the ages of 20 to 35 at the time of death'. We also have information, namely correlation coefficients, representing the *reliability* of each aging method. Lastly, we have a $[0, 1]$ value indicating the quality of each bone found. Each aging method is based on, and ultimately bounded by, the quality of these remains. The membership function for method i with respect to its interval-valued input and corresponding bone quality value, q_i, is



Skeleton	208
Chronological Age	30 yrs
Sex	Female
Sugeno Age Range	30 yrs
Pubic Symphysis	27-30 yrs
Auricular Surface	30-34 yrs
Ectocranial Suture – Vault	24-60 yrs
Ectocranial Suture - Lateral-Anterior	< 50 yrs

Fig. 2. Example age-at-death skeletal estimation fusion result (skeleton 208 from the Terry Anatomical Collection) for the NDFI algorithm [11, 12]. The true age-at-death is 30 years. The sex is female. Four different aging methods were used. Information about the FM, anthropological details and a wider range of rich examples can be found in [11, 12].

$$\mu_{A_i}(x) = \begin{cases} q_i, & \text{if } v_{i,l} \leq x \leq v_{i,r} \\ 0, & \text{otherwise,} \end{cases} \tag{18}$$

where μ_{A_i} is the membership function and $[v_{i,l}, v_{i,r}]$ are the first/left (l) and last/right age (r) in the age interval for aging method i (e.g., the interval $[10, 15]$ years). This is the sub-normal FS input we have been discussing. It is worth noting here that we are exploring ways to fuzzify the individual aging methods. At the moment, the fuzzy sets have only 0 and q_i membership values. The NDFI algorithm is formally described in Algorithm 3. Figure 2 is a result of the NDFI algorithm for skeleton 208 from the Terry Anatomical Collection [11, 12].

NDFI is based on the idea of multiple hypothesis testing. A single hypothesis is: 'the skeleton was age k at death (a specific age, not range)'. The (classical) Sugeno integral is repeatedly applied, once for each possible age using the respective accuracy, range and quality information. Every age,

Algorithm 3 NDFI algorithm in the context of skeletal age-at-death estimation for forensic anthropology [11, 12]

Input fuzzy measure g ▷ *use the Sugeno λ-fuzzy measure, learn g from data or manually specify g*
Input bone quality weathering values, $\{q_1, ..., q_E\}$ ▷ *Where $q_e \in [0, 1]$*
Input age-at-death intervals for each aging method, $\{\bar{v}_1, ..., \bar{v}_E\}$ ▷ *Where $\bar{v}_e$ is an age interval, e.g., $\bar{v}_e = [5, 20]$ years*
Discretized the output domain, $D = \{d_1, ..., d_{|D|}\}$ ▷ *e.g., $D = \{1, 2, ..., 110\}$*
Initialize the (FS) result to $R(d_i) = 0$
for each $d_i \in D$ (i.e., each discrete age) **do**
  for each $h_{i,e} \in \{h_{i,1}, ..., h_{i,E}\}$ **do** ▷ *Calculate the partial support function h_i at d_i*
    if $d_i \geq v_{e,l}$ and $d_i \leq v_{e,r}$ **then** ▷ *Where l and r are the left and right endpoints, e.g., $[v_{e,l}, v_{e,r}]$*
      $h_{i,e} = q_i$ ▷ *Age method e indicates possible age-at-death, use bone quality q_e*
    else
      $h_{i,e} = 0$ ▷ *Age method e indicates not a possible age-at-death, so no support in the hypothesis*
    end if
  end for
  Set $R(d_i) = \int_S h_i \circ g$ ▷ *Fuzzy membership at d_i is the number-valued (integrand and measure) FI of h_i with g*
end for

in discrete one year increments from 1 to 110 is tested. The age indicators provide input based on if the age tested is in their respective interval. The h values are a function (t-norm) of the quality, a $[0, 1]$ value, and if the aging method indicates that the age tested falls in the age interval, either a 0 for false or 1 for true. Again, the result of this procedure is a collection of (age tested, FI result) pairs, which is a FS defined over the age domain. In this respect, we were able to address sub-normal FSs. Refer to [11, 12] for more details regarding the application of NDFI to skeletal age-at-death estimation.

It is trivial to verify that NDFI results in FS outputs that pass the vertical line test. The NDFI algorithm generally produces sub-normal and non-convex results, Grabisch's extension (for the case of normal FSs) produces normal, convex results, and SuFI produces sub-normal, convex results. Additionally, both Grabisch's extension and SuFI produce FSs between the min and max with respect to the partial support function. The NDFI algorithm also generates FSs between the min and max, however only in regions between the min and max that are covered by at least one of the inputs.

The difference between NDFI and SuFI is apparent with respect to $(\int H \circ g)(y)$. At y, the SuFI calculation is

$$\sup_{(z_1,...,z_E)\in S_y} \{(H_1)(z_1) \wedge ... \wedge (H_E)(z_E)\}, \tag{19}$$

while NDFI is

$$\int h_y \circ g. \tag{20}$$

The Extension Principle route is all number-based FIs whose result is y and a t-norm of the membership degrees of the FS inputs at those locations. The NDFI algorithm is a number-based FI at y. The NDFI and SuFI approaches fuse the information in very different ways. The NDFI algorithm integrates *vertically* while SuFI integrates *horizontally*. In the next section, we look at numeric examples and argue that both methods have utility. Namely, the "correct approach" is problem dependent.

VII. Comparison of SuFI and NDFI

Upon beginning this investigation, the underlying question was: what is the direct method of extending the FI for sub-normal FS integrands and does it produce a better or the same result as NDFI? The short answer is no, SuFI does not produce the same result as NDFI. Also, we assert that it is unfortunately not simple to declare one approach as definitively better than the other. Each approach has its own respective advantages and disadvantages. These pros and cons are illustrated through the following numeric examples.

A. Example 1: Normal FSs

Consider the example in Fig. 3. This scenario contains two inputs $X = \{x_1, x_2\}$ with partial support function H. The two FS inputs are characterized by the triangular membership functions $\mu_{H_1} = [0, 0.2, 0.4]$ and $\mu_{H_2} = [0.6, 0.8, 1]$. The reliability of these sources is given by the fuzzy measure, $g^1 = g(x_1) = 0.5, g^2 = g(x_2) = 0.5, g(\{x_1, x_2\}) = g(X) = 1$.

SuFI produces a result which, although technically a FS, is the singleton 0.5, with a membership of 1. If the fuzzy measure is changed to $g^1 = 1, g^2 = 1, g(X) = 1$, then SuFI produces the triangular FS $[0.6, 0.8, 1]$ with height of 1 as the result. Note that this is exactly equal to μ_{H_2}.

NDFI produces very different results, shown in Fig. 4(a). View (a) shows the NDFI algorithm result for FM 1. The result is two triangles, $[0, 0.2, 0.4]$ and $[0.6, 0.8, 1]$, both with heights of 0.5. For FM2, shown in view (b), the result is the same; however, each triangle has a height of 1. A possible downside of NDFI is that for this very straight-forward example, the result is a non-convex (and for FM1, sub-normal) FS.

This example could be considered as the combination (e.g., average and maximum) of two FNs, with linguistic representations of 'about 0.2' and 'about 0.8'. If the Choquet FI was employed, then one could more easily interpret the aggregation for a given FM (e.g., OWA if sets of information sources of equal cardinality have equal measure value, average if all densities sum to one and are equal, etc.) [13]. Intuitively, we expect the output to look like the inputs: in this case,

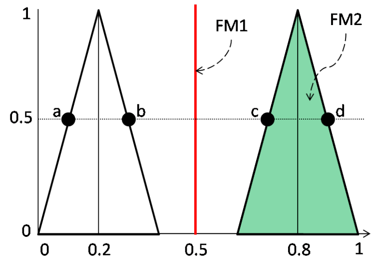
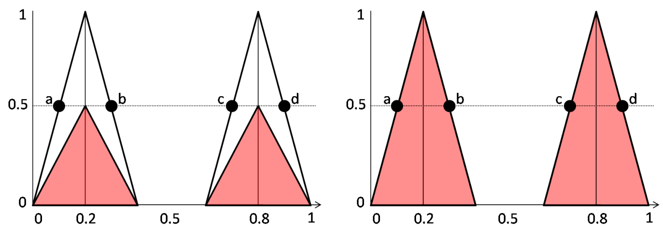


Fig. 3. Illustration of a FS integrand and interval endpoints used to compute SuFI at $\alpha = 0.5$. The results for two FMs are provided, red ($g^1 = 0.5, g^2 = 0.5, g(X) = g(\{x_1, x_2\}) = 1$) and green ($g^1 = 1, g^2 = 1, g(X) = 1$). The two FS are characterized by the triangular membership functions $\mu_{H_1} = [0, 0.2, 0.4]$ and $\mu_{H_2} = [0.6, 0.8, 1]$.

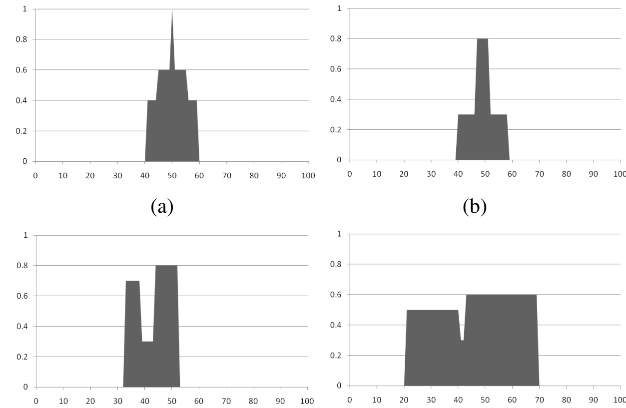


(a) (b)

Fig. 4. Illustration of a FS integrand and interval endpoints used to compute NDFI at $\alpha = 0.5$. Case (a) is for FM 1 ($g^1 = 0.5, g^2 = 0.5, g(X) = 1$) while case (b) is for FM 2 ($g^1 = 1, g^2 = 1, g(X) = 1$). The results for these FMs is shown in red. The two FS are characterized by the triangular membership functions $\mu_{H_1} = [0, 0.2, 0.4]$ and $\mu_{H_2} = [0.6, 0.8, 1]$.

a triangular FS with the linguistic interpretation of something like 'about 0.5' or 'about 0.8' (depending on the FM). The NDFI algorithm, again depending on the selection of FM, produces a result that is differently shaped from each of the inputs. In contrast, SuFI produces outputs that look very much like the inputs, namely triangular FSs, which would be easily interpreted. However, the downfall of SuFI is that if *any* of the inputs are sub-normal FSs then the output will have a maximum membership of the minimum-height sub-normal FS, even if the respective reliability (g) of that sub-normal input is 0-valued (which intuitively means that we should ignore that input; it has no worth in the solution to the FI). Hence, both GAFI and SuFI have their respective drawbacks.

In contrast, for age-at-death estimation in anthropology, we desire a *restricted* result. That is, Anthropologists indicate that one should be careful to not produce ages *outside* of intervals indicated by the individual aging methods. For example, if one method reports $[10, 20]$ and another method reports $[60, 100]$ (which, for most practical cases is unlikely), we do not want to produce an age interval such as $[40, 50]$. In addition to fusing the inputs, we would like to have a way to discover that there is disagreement among the sources and we would like to find the age(s) that are the most confident. That is, we would like to take into account the agreement between sources, the method's confidences and our confidences in the sources. If one input has a low height, we do not want the



(a) (b)

(c) (d)

Fig. 5. Interpretation of resultant FS in age-at-death estimation using NDFI [11, 12]. Categories identified by Anthropologists include: (a) specific age (aging method have come together and agree on a single age-at-death), (b) age interval (there is agreement between the sources but no single definitive age), (c) disagreement (there is disagreement between the methods, thus multiple plateaus) and (d) inconclusive (so much disagreement or general lack of confidence that it is difficult to conclude anything).

FI result to be ultimately limited by this amount. In [11], our objective was to find a way to fuse the various information (FS inputs, bone quality values and numeric values representing the 'worth' of the information sources) and then analyze the result. The result was the introduction of NDFI. In [11], we calculated a single age-at-death number (e.g., died at age 20). We identified FS features and created fuzzy class definitions to assist with interpreting the FS results [12]. We also measured the confidence and specificity of the resultant FSs. The four anthropological FS categories are shown in Fig. 5. These categories represent: specific age (aging method have come together and agree on a single age-at-death), age range (agreement between the sources but no single definitive age), disagreement (there is disagreement between the methods, thus multiple plateaus) and inconclusive (so much disagreement or lack of confidence that it is difficult to conclude anything).

B. *Example 2: Sub-Normal FSs*

Consider the example in Fig. 6(a). This scenario contains two inputs $X = \{x_1, x_2\}$ with partial support function H. The two FS inputs are characterized by the triangular membership functions $\mu_{H_1} = [0, 0.2, 0.4]$ and $\mu_{H_2} = [0.6, 0.8, 1]$, and the FM is $g^1 = 1, g^2 = 0, g(X) = 1$ (i.e., no worth is assigned to the second information source). However, in this example let the height of μ_{H_2} be 0.01 (sub-normal FS).

The SuFI algorithm results in the trapezoidal membership function $[0, 0.002, 0.398, 0.4]$ with height 0.01. Note, this result is different in *shape* than the input. That is, the inputs are triangular while the result is a trapezoid. While the second source is completely un-trustworthy ($g^2 = 0$), it has substantially impacted the result. The resultant height is so low that intuitively one should ignore the result. However, for this second experiment NDFI produces a more pleasing

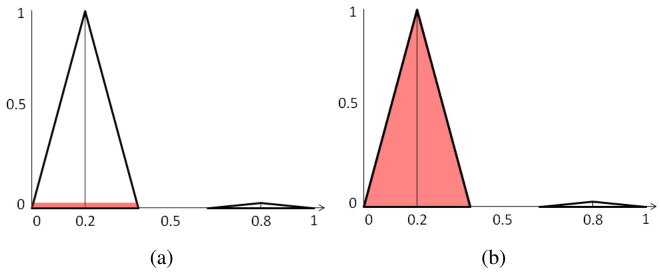


Fig. 6. Results for the FM $g^1 = 1, g^2 = 0, g(X) = 1$. Case (a) is for SuFI and (b) is for NDFI. The two FS are characterized by the triangular membership functions $\mu_{H_1} = [0, 0.2, 0.4]$ and $\mu_{H_2} = [0.6, 0.8, 1]$, with $Height(H_1) = 1$ and $Height(H_2) = 0.01$.

result. That is, a single triangle of height 1 at $[0, 0.2, 0.4]$ and no support (height 0.01) in $[0.6, 0.8, 1]$ (shown in Fig. 6(b)).

C. Example 3: Age-at-Death Estimation

Next, we consider a case from our prior skeletal age-at-death estimation work [11]. This example (Table I) consists of eight aging methods. Each remain (bone) is associated with a skeletal quality value of less than one, i.e., $Height(H_i) \leq 1$. From an Anthropological standpoint, looking at the agreement between these aging methods, we would expect a result close to the true age-at-death (which is 38). Specifically, we expect a *narrow* interval (not a single age-at-death because the inputs are all interval-valued with width greater than 1) that includes the age 38. The input FSs have heights (their confidence) equal to their respective quality of bone. Additionally, the fusion procedure (SuFI and NDFI) is expected to fuse this information with respect to the reliability of the aging methods. In this work, as well as in our previous work, the Sugeno λ-FM is used to build the entire FM from the densities. Figure 7 is the result of SuFI and NDFI. Note, with respect to the SuFI, the inputs are first scaled from $[0, 110]$ to $[0, 1]$ (division by 110), the SuFI algorithm is run and the results are then scaled back to $[0, 110]$ (multiplication by 110).

The following observations are made with respect to SuFI and NDFI. First, the inputs are trapezoids and the output of SuFI is a trapezoid. Specifically, the output is sub-normal and convex and its shape is that of the inputs. In comparison, the output of NDFI is sub-normal and non-convex and its shape does not resemble that of the individual inputs. Second, the interval $[37, 39]$ has the most agreement among the inputs. That is, each age method reports these ages. However, we do not desire an overly simple procedure that just counts the number of times that an age is agreed upon by the aging methods followed by a selection of an interval that has a maximum score. It is very likely that multiple intervals could exist. Additionally, we would like to take into consideration the reliability of each aging method. This is the motivation for taking a generalized Sugeno FI approach. That said, SuFI returns a single (and very wide or non-specific at that) interval, $[37, 76]$. While the SuFI algorithm output does include the true age-at-death, it includes to many other ages as well.

TABLE I
INPUT FOR EXAMPLE 3 FROM OUR PRIOR AGE-AT-DEATH WORK [11]

Aging Method	**Quality**	**Age Range**	g^i
Pubic Symphysis	0.6	35-39	.57
Auricular Surface	0.8	35-39	.72
Ectocranial Sutures - vault	0.2	24-75	.59
Ectocranial Sutures - lateral	0.5	23-63	.59
Sternal Rib Ends	0.5	33-42	.75
Endocranial Sutures	0.4	35-39	.51
Proximal Humerus	0.3	37-86	.44
Proximal Femur	0.7	25-76	.56

In comparison, the NDFI algorithm result indicates a single maximum *plateau* of $[35, 39]$, which for example 3 is a single interval associated with the highest membership degree (see [11] and [12] for a formal definition of maximum *plateau*). However, in some cases, such as those discussed in [11, 12], multiple plateau's can exist. To summarize example 3, both NDFI and SuFI include the true age in their result, however NDFI indicates a fewer number of possible ages. The SuFI result is a wide (that is, non-specific) interval that is of little-to-no use for age-at-death estimation. It reports that the true age-at-death is one of 40 possible ages. However, the NDFI algorithm result is more specific, i.e., the true age-at-death is one of 5 possible ages (according to the maximum plateau).

As discussed in our prior work [11, 12], NDFI is loaded with a wealth of additional information. Using our FS approach to linguistically describe generalized FI produced FSs, the following can be concluded (which is not available in the SuFI algorithm output). First, the *shape* of the resultant FS informs us about the nature of the agreement. That is, the result is of type *interval* (one of many possible ages), however it is not very wide and could potentially be considered as type *specific* (a single age-at-death). Additionally, in [12] we defined a linguistic variable to interpret the confidence of the output decision. For example 3, NDFI reports that the fused result is of *moderate* confidence (the maximum plateau has a height of 0.72), while SuFI (of height 0.2) is of *very low* confidence (and most likely should be ignored).

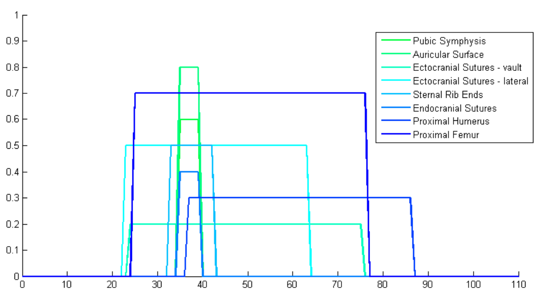
VIII. CONCLUSION

In closing, we investigated different generalizations of the Sugeno *fuzzy integral* (FI). We reviewed existing number, interval, and *fuzzy set* (FS)-valued integrand extensions to the Sugeno FI. One problem is that current FS-valued solutions require normality. However, we highlight an age-at-death application from anthropology that has sub-normal FS inputs. To address this problem, we proposed a generalization for sub-normal FS integrands (the SuFI). The advantages and shortcomings (summarized in Table II) of SuFI and our prior approach, NDFI, are discussed and shown using numeric examples and cases from skeletal age-at-death estimation.

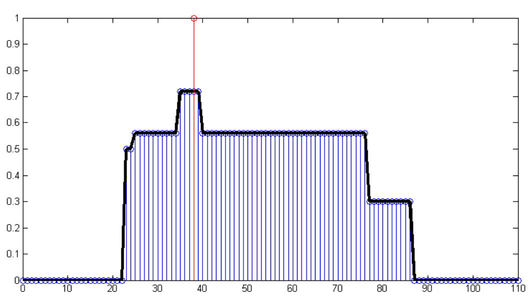
Our general goal is to develop a solid understanding of the unrestricted extension of FIs with respect to both the integrand and FMs. This article is a first step in that line of work. On a final note, we will explore a quality definition of a type-2

TABLE II
IMPORTANT PROPERTIES OF SUFI AND NDFI.

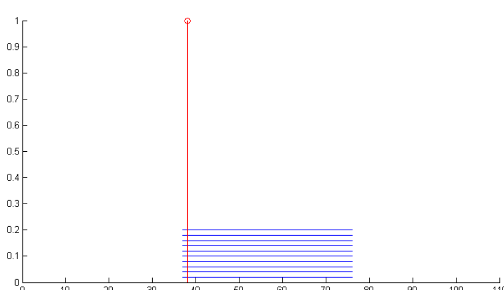
Property	SuFI	NDFI
Height($\int H \circ g$)	Height of lowest FS (i.e., minimum of Height(H_1), ..., Height(H_E))	Depending on the FM, anywhere between 0 and the maximum FS Height
Range of $\int H \circ g$	$\int H \circ g$ can be between the minimum and maximum of input FSs (the integrand)	$\int H \circ g$ can be between the minimum and maximum of the input FSs, however the result is *restricted* to regions between the minimum and maximum that are covered by at least one of the inputs
Approach to $(\int H \circ g)(y)$	Extension Principle (Equation 18)	(Sugeno) FI at y (Equation 19)
Resulting shape of $\int H \circ g$	Can be different from that of the inputs, e.g., for triangular shaped sub-normal FS inputs can obtain a trapezoidal shaped output. In general, sub-normal (if any input is sub-normal) and convex	In general, will be sub-normal and non-convex



(a)



(b)



(c)

Fig. 7. Results found for the inputs, bone quality values and Sugeno λ-FM for the densities reported in Table I. In (a), the input FSs are shown. In (b), the NDFI algorithm result is shown. The x-axis is the range $[0, 110]$. In (c), the SuFI algorithm result is shown for 11 α-cuts.

fuzzy measure (FM). One of the motivating reasons for this article was a better understanding of the behavior of the FI for the case of sub-normal FS integrands as it relates to type-2 extensions if one approached it from the standpoint of the FI with respect to a collection of embedded type-1 FSs.

REFERENCES

[1] M. Grabisch, H. Nguyen, and E. Walker, *Fundamentals of uncertainty calculi, with applications to fuzzy inference*. Kluwer Academic, Dordrecht, 1995.

[2] T. M. M. Grabisch and M. Sugeno, *Fuzzy Measures and Integrals: Theory and Applications*. Physica-Verlag, Heidelberg, 2000.

[3] D. Dubois and H. Prade, *Fuzzy Numbers: An Overview*. CRC Press, 1987.

[4] S. Auephanwiriyakul, J. Keller, and P. Gader, "Generalized choquet fuzzy integral fusion," *Information Fusion*, vol. 3, pp. 69–85, 2002.

[5] C. Guo, D. Zhang, and C. Wu, "Fuzzy-valued measures and generalized fuzzy integrals," *Fuzzy Sets and Systems*, vol. 97, pp. 255–260, 1998.

[6] C. Wu, S. Wang, and M. Ma, "Generalized fuzzy integrals i," *Fuzzy Sets and Systems*, vol. 57, pp. 219–226, 1993.

[7] ——, "Generalized fuzzy integrals ii," *Fuzzy Sets and Systems*, vol. 70, pp. 75–87, 1995.

[8] D. Zhang and Z. Wang, "Fuzzy integrals of fuzzy-valued functions," *Fuzzy Sets and Systems*, vol. 54, pp. 63–67, 1993.

[9] D. Zhang and C. Guo, "Generalized fuzzy integrals of set-valued functions," *Fuzzy Sets and Systems*, vol. 76, pp. 365–373, 1995.

[10] L. Zadeh, "The concept of linguistic variable and its application to approximate reasoning," *Information science*, vol. 8, pp. 199–249, 1975.

[11] M. Anderson, D. Anderson, and D. Wescott, "Estimation of adult skeletal age-at-death using the sugeno fuzzy integral," *Journal of Physical Anthropology*, vol. 142, pp. 30–41, 2010.

[12] D. Anderson, M. Anderson, J. Keller, and D. Wescott, "Linguistic description of adult skeletal age-at-death estimations from fuzzy integral acquired fuzzy sets," in *IEEE International Conference on Fuzzy Systems (FUZZ)*, 2011, pp. 2274–2281.

[13] H. Tahani and J. Keller, "Information fusion in computer vision using the fuzzy integral," *IEEE Transactions System Man Cybernetics*, vol. 20, pp. 733–741, 1990.

[14] T. Todd, "Age changes in the pubic bone: I: the male white pubis," *American Journal of Physical Anthropology*, vol. 3, pp. 285–334, 1920.

[15] M. Sugeno, "Theory of fuzzy integrals and its applications," *Ph.D. thesis*, vol. Tokyo Institute of Technology, 1974.

[16] D. Anderson, J. Keller, and T. Havens, "Learning fuzzy-valued fuzzy measures for the fuzzy-valued sugeno fuzzy integral," in *International conference on information processing and management of uncertainty*, 2010, pp. 502–511.