The WOWA Operator: A Review

Vicenç Torra

Abstract. The WOWA operator (Weighted OWA) was proposed as a generalization of both the OWA and the Weighted mean. Formally, it is an aggregation operator that permits the aggregation of a set of numerical data with respect to two weighting vectors: one corresponding to the one of the weighted mean and the other corresponding to the one of the OWA. In this chapter we review this operator as well as some of its main results.

1 Introduction

Aggregation operators [20, 21] permit us to combine data provided from several sources and return a single datum that is of better quality and, therefore, gives more accurate information. Several aggregation operators have been defined in the literature. Differences in the operators are based on the differences between the data, and the properties of these data.

A common classification of the operators is related to the nature of the data. In this way, we can distinguish between numerical data, categorical data, and also between data in other terms as e.g. partitions, (fuzzy) clusters, dendrograms, sequences.

In this chapter we will review some results about the WOWA operator. This operator, introduced in [13] and [14], was defined for numerical data.

From a practical point of view, this operator was defined to encompass in a single operator the advantages of the weighted mean and of the OWA operator. Informally, the weighted mean permits us to weight the information sources, and the OWA permits us to represent a compensation degree, or, alternatively, to give importance to the data according to their values.

Vicenç Torra
IIIA-CSIC, Campus UAB s/n, 08193 Bellaterra,
Catalonia, Spain
e-mail: vtorra@iiia.csic.es

R.R. Yager et al. (Eds.): Recent Developments in the OWA Operators, STUDFUZZ 265, pp. 17–28.
springerlink.com

From a mathematical point of view, the operator is a generalization of both the weighted mean and the OWA. That is, particular parametrizations of the operator lead to either the weighted mean or the OWA. In addition, it has also been proven [15] that the operator is a particular case of the Choquet integral [4].

In this paper we review some of the main results on this operator. In Section 2 we review the WOWA operator as well as other aggregation operators that are related to the WOWA. In Section 3 we discuss some generalizations of the operator. In Section 4 we review some learning approaches for the parameters of this operator.

2 The WOWA Operator and Other Aggregation Operators

Aggregation operators are functions that combine N different data into a single datum. We use $\mathbb{C}$ from $\mathbb{C}$onsensus or $\mathbb{C}$ombination to represent them. Then, in general, it is assumed that the aggregation of $a_1,\ldots,a_N$ in a given domain D is $\mathbb{C}(a_1,\ldots,a_N)$, also in this domain D. That is, $\mathbb{C} : D^N \to D$.

In some cases it is useful to represent in an explicit way where the data come from. That is, which is the information source that has supplied each data. We will use $X = \{x_1,\ldots,x_N\}$ to represent the set of information sources. Then, we will use f to represent the relationship between x_i and the supplied value a_i. That is, $f(x_i) = a_i$ represents that x_i supplies a_i. Using this notation, we have that the aggregation of the data supplied by the information sources in X is $\mathbb{C}(f(x_1),\ldots,f(x_N))$, or, with an abuse of notation, $\mathbb{C}(f)$.

There exist a few different definitions on what an aggregation function is. In general it is usual to require monotonicity and unanimity or idempotency (for at least a few elements in the domain). We consider aggregation operators as functions $\mathbb{C}$ satisfying:

- **Unanimity or idempotency:** $\mathbb{C}(a,\ldots,a) = a$ for all a in D
- **Monotonicity:** $\mathbb{C}(a_1,\ldots,a_N) \geq \mathbb{C}(a'_1,\ldots,a'_N)$ when $a_i \geq a'_i$

Some require unanimity only in the boundaries of D. In particular, if $D = [0,1]$, unanimity is only required for 0 and 1. So, $\mathbb{C}(0,\ldots,0) = 0$ and $\mathbb{C}(1,\ldots,1) = 1$. This is the case of [2]. In this case, t-norms and t-conorms are aggregation functions. In this case, the term *mean operators* is used to name functions that satisfy unanimity for all a in D.

In addition, in some circumstances the symmetry condition is also required to aggregation operators. This property, which is formalized below, implies that there is no distinguished data.

- **Symmetry:** For any permutation π on $\{1,\ldots,N\}$ it holds that

$$\mathbb{C}(a_1,\ldots,a_N) = \mathbb{C}(a_{\pi(1)},\ldots,a_{\pi(N)})$$

2.1 Arithmetic Mean, Weighted Mean and OWA Operator

The arithmetic mean, the weighted mean and the OWA operator are some of the most well known aggregation operators. Both the weighted mean and the OWA combine a set of data with respect to a weighting vector. The arithmetic mean does not include any parameter.

The weighting vector in the weighted mean permits us to take into account some *a prior* knowledge, following artificial intelligence jargon. about the information sources. We give below the definitions of the weighting vector, the arithmetic mean and the OWA operator.

Definition 1. Let $A = (a_1, \dots, a_N)$ be N data in $\mathbb{R}$. Then, we define a weighting vector, the arithmetic mean ($AM : \mathbb{R}^N \to \mathbb{R}$) of A, and the weighted mean (WM) of A with respect to a weighting vector as follows:

- A vector $v = (v_1 \dots v_N)$ is a *weighting vector* of dimension N if and only if $v_i \in [0,1]$ and $\sum_i v_i = 1$.
- AM is an *arithmetic mean*, if $AM(a_1, ..., a_N) = (1/N) \sum_{i=1}^{N} a_i$.
- WM is the *weighted mean* with respect to a weighting vector $\mathbf{p}$, if $WM_{\mathbf{p}}(a_1, ..., a_N) = \sum_{i=1}^{N} p_i a_i$.

The OWA (Ordered Weighting Averaging) operator has a definition similar to the one of the weighted mean. It is as follows:

Definition 2. [22, 23] Let $\mathbf{w}$ be a weighting vector of dimension N; then, a mapping OWA: $\mathbb{R}^N \to \mathbb{R}$ is an *Ordered Weighting Averaging (OWA) operator* of dimension N if

$$OWA_{\mathbf{w}}(a_1, ..., a_N) = \sum_{i=1}^{N} w_i a_{\sigma(i)},$$

where $\{\sigma(1), ..., \sigma(N)\}$ is a permutation of $\{1, ..., N\}$ such that $a_{\sigma(i-1)} \geq a_{\sigma(i)}$ for all $i = \{2, ..., N\}$ (i.e., $a_{\sigma(i)}$ is the ith largest element in the collection $a_1, ..., a_N$).

The weighted mean and the OWA operator are similar operators as both are a linear combination of the values with respect to the weights. Nevertheless, the ordering step that takes place in the OWA operator makes a fundamental difference. This difference makes different the interpretation of the weights in both operators.

In the weighted mean, the weight is attached to the information source. Due to this, weights correspond to the importance of the information sources. E.g., when the data correspond to sensors, the weight might correspond to the reliability of the corresponding sensor; and when the data correspond to the evaluation of some criteria (or experts) in a multicriteria decision making problem, the weights correspond to the importance of the criteria (or of the experts).

In contrast, in the OWA operator, the weight is attached to the data, with respect to its relative position. Due to this, weights permit us to give more importance to e.g. low values, central values or high values. For example, we can give more importance to small distances (e.g., if we want to avoid a collision of a robot, is more importance a nearer object than a farther one), or permit some compensation (e.g., if a bad evaluation of a criteria can be compensated with a good one – or in the extreme case, if a single good criteria can override all the others).

The degree of compensation in the OWA operator is measured with the *orness* degree. This is a measure that evaluates in what extent the outcome of the aggregation is near to the maximum of the data being aggregated. The larger the outcome, the larger the orness and the larger the compensation. Note that the maximum compensation corresponds to assigning the largest value to the output of the function.

The orness is formally defined below and, in fact, the definition is valid for all aggregation operators $\mathbb{C}$ and for all parameterizations P. It results that for some of the operators, the orness does not depend on the particular parameterization selected, while for others the orness depends on the particular parameterization used. The weighted mean is an example of the former (i.e., the orness of the weighted mean is independent of the parameter used), and the OWA is an example of the latter (i.e., different parameters give different orness for the OWA).

Definition 3. Let $\mathbb{C}$ be an aggregation operator with parameters P; then, the *orness* of $\mathbb{C}_P$ is defined by

$$orness(\mathbb{C}_P) := \frac{AV(\mathbb{C}_P) - AV(\min)}{AV(\max) - AV(\min)}. \quad (1)$$

The orness of the aggregation operators reviewed above is as follows:

- $orness\ (AM) = 1/2$
- $orness(WM_{\mathbf{p}}) = 1/2$
- $orness(OWA_{\mathbf{w}}) = \frac{1}{N-1}\sum_{i=1}^{N}(N-i)w_i$

From the orness of the OWA we can infer that its maximum orness is 1 when $w_1 = 1$ and $w_i = 0$ for all $i \neq 1$ (note that in this case the OWA corresponds to the maximum), and that the minimum ornes is 0 when $w_N = 1$ and $w_i = 0$ for all $i \neq N$ (note that in this case the OWA corresponds to the minimum).

2.2 *The WOWA Operator*

Due to the fact that in some applications it is of interest to assign weight to information sources and also to the compensation degree (or the relative importance of values), a generalization of both weighted mean and OWA was proposed. This generalization is the WOWA operator. WOWA operator stands for Weighted Ordered Weighted Averaging operator. Formally, it is also a linear combination of values $a_1, \ldots, a_N$ with respect to weights. Nevertheless, these weights are computed taking into account two weighting vectors. One of the weighting vectors has the interpretation of the ones in the weighted mean, and the other has the interpretation of the ones in the OWA. We use here **p** to represent the weights with the interpretation used in the weighted mean, and **w** to represent the weights with the interpretation used in the OWA. Note that although we use here different letters **w** and **p**, both weighting vectors have the same mathematical properties.

Definition 4. [13, 14] Let **p** and **w** be two weighting vectors of dimension N; then, a mapping WOWA: $\mathbb{R}^N \to \mathbb{R}$ is a *Weighted Ordered Weighted Averaging (WOWA) operator* of dimension N if

$$WOWA_{\mathbf{p},\mathbf{w}}(a_1,...,a_N) = \sum_{i=1}^{N} \omega_i a_{\sigma(i)},$$

where σ is defined as in the case of OWA (*i.e.*, $a_{\sigma(i)}$ is the ith largest element in the collection $a_1,...,a_N$), and the weight ω_i is defined as

$$\omega_i = w^*(\sum_{j\leq i} p_{\sigma(j)}) - w^*(\sum_{j<i} p_{\sigma(j)}),$$

with w^* being a nondecreasing function that interpolates the points

$$\{(i/N, \sum_{j\leq i} w_j)\}_{i=1,...,N} \cup \{(0,0)\}.$$

The function w^* is required to be a straight line when the points can be interpolated in this way.

As stated above, this definition uses an interpolation method to build a function from the points in the set $\{(i/N, \sum_{j\leq i} w_j)\}_{i=1,...,N} \cup \{(0,0)\}$. The original definition used the interpolation method described in [17]. Other interpolations approaches have been used as e.g. linear interpolation. A discussion on the effects of different interpolation methods is given in [19].

For details on the WOWA operator, and about the meaning of the function see [20, 21].

Some extensions have been defined for this operator. One of them, the Linguistic WOWA (L-WOWA) operator [14], was given to deal with categorical data. L-WOWA operator can be seen as an extension of the L-OWA, in the same way that the WOWA operator is an extension of the OWA operator.

2.3 The Choquet Integral

The Choquet integral is an operator that generalizes the WOWA operator, as proven in [15]. As the WOWA generalizes the arithmetic mean, the weighted mean and the OWA, it can be said that all these functions belong to the same family of operators.

The basis of this integral, in comparison with the other mentioned operators, is that now the *weights* (or importances) are not of a single information source but to a set of them. While in the weighted mean, we have p_i as the weight of information source x_i, we can not consider the weight of e.g. the set $\{x_1, x_4\}$. Formally, in the weighted mean we have weights $p : X \to [0,1]$ such that $\sum_{x_i \in X} p(x_i) = 1$, and we use the notation $p_i = p(x_i)$. Thus, $p_i = p(x_i)$ is the importance of information source x_i.

In contrast, in the case of the Choquet integral we use functions μ over subsets of X. Then, $\mu(\psi)$ for $\psi \subseteq X$ is the importance of the elements in ψ taken together.

As in the case of the weighting vectors, $\mu(\psi) \in [0,1]$. These functions are known as fuzzy measures, and we review them below.

Definition 5. A fuzzy measure μ on a set X is a set function $\mu : \wp(X) \to [0,1]$ satisfying the following axioms:

(i) $\mu(\emptyset) = 0$, $\mu(X) = 1$ (boundary conditions)
(ii) $A \subseteq B$ implies $\mu(A) \leq \mu(B)$ (monotonicity)

That is, μ are set functions that satisfy monotonicity. Monotonicity means that the larger the set, the larger the measure, or, equivalently, the larger the set of criteria, the larger their importance. In addition, the maximum importance (equal to 1) is achieved for the whole set of criteria, and the minimum importance (equal to 0) is achived for the empty set.

Choquet integrals permit to aggregate values taking into account the importance expressed in the measures. The aggregation corresponds to the integral of a function with respect to the measure. The function corresponds to the data to be aggregated as expressed above with the expression $\mathbb{C}(f)$.

Definition 6. [4] Let μ be a fuzzy measure on X; then, the *Choquet integral* of a function $f : X \to \mathbb{R}^+$ with respect to the fuzzy measure μ is defined by

$$(C)\int f d\mu = \sum_{i=1}^{N} [f(x_{s(i)}) - f(x_{s(i-1)})]\mu(A_{s(i)}), \tag{2}$$

where $f(x_{s(i)})$ indicates that the indices have been permuted so that $0 \leq f(x_{s(1)}) \leq \cdots \leq f(x_{s(N)}) \leq 1$, and where $f(x_{s(0)}) = 0$ and $A_{s(i)} = \{x_{s(i)}, \ldots, x_{s(N)}\}$.

When no confusion exists, we can use $CI_\mu(a_1, \ldots, a_N) = (C)\int f d\mu$, where, $f(x_i) = a_i$, as before. There are alternative expressions for the Choquet integral that are equivalent to the one given above. The next proposition presents one of them.

Proposition 1. *Let* μ *be a fuzzy measure on* X*; then, the Choquet integral of a function* $f : X \to \mathbb{R}^+$ *with respect to* μ *can be expressed as*

$$(C)\int f d\mu = \sum_{i=1}^{N} f(x_{\sigma(i)})[\mu(A_{\sigma(i)}) - \mu(A_{\sigma(i-1)})], \tag{3}$$

where $\{\sigma(1), \ldots, \sigma(N)\}$ *is a permutation of* $\{1, \ldots, N\}$ *such that* $f(x_{\sigma(1)}) \geq f(x_{\sigma(2)}) \geq \cdots \geq f(x_{\sigma(N)})$*, where* $A_{\sigma(k)} = \{x_{\sigma(j)} | j \leq k\}$ *(or, equivalently,* $A_{\sigma(k)} = \{x_{\sigma(1)}, \ldots, x_{\sigma(k)}\}$ *when* $k \geq 1$ *and* $A_{\sigma(0)} = \emptyset$*).*

As stated above, the WOWA operator is a particular case of the Choquet integral. In particular, a WOWA operator corresponds to a Choquet integral with respect to a distorted probability. Distorted probabilities are a particular type of fuzzy measure. All Choquet integral with respect to this type of fuzzy measures are equivalent to a WOWA operator, and all WOWA operators with weights **p** and **w** are equivalent to a Choquet integral with respect to the distorted probability constructed from **p** and the function w^* constructed using the interpolation method in the definition of the WOWA (Definition 4).

In the next two definitions, we review the definition of a distorted probability.

Definition 7. Let $P : 2^X \to [0,1]$ be a probability distribution. Then, we say that a function f is strictly increasing with respect to P if and only if

$$P(A) > P(B) \text{ implies } f(P(A)) > f(P(B))$$

At this point it is relevant to state that as we suppose that X is a finite set, when there is no restriction on the function f, a strictly increasing function f with respect to P can be regarded as a strictly increasing function on $[0,1]$. Note that with respect to increasingness only the points in $\{P(A)|A \in 2^X\}$ are essential, the others are not considered by $f(P(A))$.

Definition 8. [1, 3] Let μ be a fuzzy measure. We say that μ is a distorted probability if there exists a probability distribution P and a strictly increasing function f with respect to P such that $\mu = f \circ P$.

3 Generalizations of the WOWA Operator

In a recent paper [11], we introduced an extension of distorted probabilities. This was motivated by the fact that distorted probabilities is only a small fraction [7, 11] of all possible fuzzy measures. m-dimensional distorted probabilities permits us, with an appropriate value m, to represent all fuzzy measures.

These measures, together with m-symmetric ones, permit us to naturally extend WOWA and OWA operators into m-dimensional WOWA and m-dimensional OWA. The m-dimensional ones with $m = |X|$ are equivalent to a Choquet integral with an unconstrained fuzzy measure. That is, a Choquet integral with an arbitrary fuzzy measure. Definitions and results are reviewed in this section.

3.1 m-Dimensional Distorted Probabilities

We start defining m-dimensional distorted probabilities, and then review two basic properties.

Definition 9. [11] Let $\{X_1, X_2, \cdots, X_m\}$ be a partition of X; then, we say that μ is an at most m-dimensional distorted probability if there exists a function f on $\mathbb{R}^m$ and probabilities P_i on $(X_i, 2^{X_i})$ such that:

$$\mu(A) = f(P_1(A \cap X_1), P_2(A \cap X_2), \cdots, P_m(A \cap X_m)) \quad (4)$$

where f on $\mathbb{R}^m$ is strictly increasing with respect to each variable.

We say that an at most m-dimensional distorted probability μ is an m-dimensional distorted probability if μ is not an at most $(m-1)$-dimensional.

The fact that all fuzzy measures can be represented as m-dimensional distorted probabilities follows from the next proposition, that is trivial from the above definition.

Proposition 2. *[11] Every fuzzy measure is an at most m-dimensional distorted probability with* $m = |X|$.

Note that for $n = |X|$, we are considering the following partition of X: $\{X_1 = \{x_1\}, \ldots, X_n = \{x_n\}\}$. So, $f(a_1, \ldots, a_n) = \mu(A)$ when $a_i = 1$ if and only if $x_i \in A$.

To complete the properties of these fuzzy measures, we have the following proposition that states that m-dimensional distorted probabilities define a family of measures with increasing complexity with respect to m. This means that increasing the value of m, the number of measures being representable increases. The following proposition establishes this property.

Proposition 3. *[11] Let* $\mathscr{M}_k$ *be the set of all fuzzy measures that are k-dimensional distorted probabilities and let* $\mathscr{M}_0$ *be the empty set. Then* $\mathscr{M}_{k-1} \subset \mathscr{M}_k$ *for all* $k = 1, 2, \ldots, |X|$.

3.2 *m-Symmetric Fuzzy Measures*

Symmetric fuzzy measures are those measures where the measure of a set depends only on the number of elements in the set. That is, $\mu(A) = f(|A|)$ for a function f ($|\cdot|$ stands for the cardinality of a set). It has been proven that an OWA operator corresponds to a Choquet integral with respect to the following symmetric fuzzy measure: $\mu(A) = \sum_{i=1}^{|A|} w_i$.

The concept of symmetric fuzzy measure has been extended to m-symmetric fuzzy measures [9, 8]. The definition is based on the *set of indifference*. Such set is defined by elements that do not affect the value of the measure. That is, the elements of a set are indistinguishable with respect to the fuzzy measure.

Definition 10. [9, 8] Given a subset A of X, we say that A is a set of indifference if and only if:

$$\forall B_1, B_2 \subseteq A, |B_1| = |B_2|,$$

$$\forall C \subseteq X \setminus A \quad \mu(B_1 \cup C) = \mu(B_2 \cup C)$$

In the case of $m = 2$, we have the following definition. Below is the general one.

Definition 11. [9, 8] Given a fuzzy measure μ, we say that μ is an at most 2-symmetric fuzzy measure if and only if there exists a partition of the universal set $\{X_1, X_2\}$, with $X_1, X_2 \neq \emptyset$ such that both X_1 and X_2 are sets of indifference. An at most 2-symmetric fuzzy measure is 2-symmetric if X is not a set of indifference.

Definition 12. [9, 8] Given a fuzzy measure μ, we say that μ is an at most m-symmetric fuzzy measure if and only if there exists a partition of the universal set $\{X_1, \ldots, X_m\}$, with $X_1, \ldots, X_m \neq \emptyset$ such that $X_1, \ldots X_m$ are sets of indifference.

It is clear from this definition that all fuzzy measures are m-symmetric for a large enough value m. This is stated in the next proposition.

Proposition 4. *Every fuzzy measure μ is an at most n-symmetric fuzzy measure for $n = |X|$.*

Definition 13. [9, 8] Given two partitions $\{X_1, \ldots, X_p\}$ and $\{Y_1, \ldots, Y_r\}$ on the finite universal set X, we say that $\{X_1, \ldots, X_p\}$ is coarser than $\{Y_1, \ldots, Y_r\}$ if the following holds:

$$\forall X_i \exists Y_j \text{ such that } Y_j \subseteq X_i$$

Definition 14. [9, 8] Given a fuzzy measure μ, we say that μ is m-symmetric if and only if the coarsest partition of the universal set in sets of indifference contains m non empty sets. That is, the coarsest partition is of the form: $\{X_1, \ldots, X_m\}$, with $X_i \neq \emptyset$ for all $i \in \{1, \ldots, m\}$.

It is known that symmetric fuzzy measures are a particular case of distorted probabilities. This is in relation to the fact that OWA operators are a particular case of WOWA operators. This relationship can also be established between m-symmetric fuzzy measures and m-dimensional distorted probabilities.

Proposition 5. *[10] Let μ be an m-symmetric fuzzy measure with respect to the partition $\{X_1, \ldots, X_m\}$. Then, μ is an m-dimensional distorted probability.*

The reversal of this proposition is not true, as it is the case for 1-dimensional ones, where the OWA and the WOWA are also not equivalent. The next proposition characterizes one case in which m-dimensional distorted probabilities are m-symmetric fuzzy measures.

Proposition 6. *[10] Let μ be an m-dimensional distorted probability. If, $p_i(x_j) = p_i(x_k)$ for all $x_j, x_k \in X_i$ and for all $i = 1, \ldots, m$, then μ is an m-symmetric fuzzy measure.*

3.3 *m*-Dimensional OWA and *m*-Dimensional WOWA

The definition of m-dimensional operators relies on the well known fact that OWA operators are equivalent to Choquet integrals with respect to symmetric fuzzy measures. On the basis of this fact, m-symmetric fuzzy measures permit us to define the corresponding generalization of the OWA operator. This is defined below.

Definition 15. [10] The m-dimensional OWA is defined as the Choquet integral with respect to an m-symmetric fuzzy measure.

In a similar way, WOWA operators are equivalent to Choquet integrals with respect to distorted probabilities [15]. Therefore, a Choquet integral with an m-dimensional probability can be seen as a generalization of the WOWA operator. We give this definition below.

Table 1 Data for learning

u_{c_1}	u_{c_2}	$\dots$	uc_N	$R_{\mathbb{C}}$	$u_{\mathbb{C}}$
a_1^1	a_2^1	$\dots$	a_N^1	p^1	b^1
a_1^2	a_2^2	$\dots$	a_N^2	p^2	b^2
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
a_1^M	a_2^M	$\dots$	a_N^M	p^M	b^M

Definition 16. [10] The m-dimensional WOWA is defined as the Choquet integral with respect to an m-dimensional distorted probability.

Definitions 15 and 16 permits us to establish the following result that is a corollary of Proposition 5.

Corollary 1. *[10] An m-dimensional OWA is a particular case of an m-dimensional WOWA. In other words, a Choquet integral with respect to an m-symmetric fuzzy measure is a particular case of a Choquet integral with respect to an m-dimensional distorted probability.*

Thus, the same relationship that holds for OWA and WOWA, also holds for m-dimensional OWA and m-dimensional WOWA.

4 Learning Parameters for the WOWA Operator

Learning parameters for the WOWA operator corresponds to the process of determining its weighting vectors **p** and **w**. This problem was considered in [18] under a supervised environment. That is, it is assumed that we have a set of examples for which both the input data and the output data are known. Table 1 represents this situation. Under this assumption, we select the weights **p** and **w** that minimize the difference between the expected output and the real output.

Assuming that the difference between the expected outcome and the actual outcome is computed using the Euclidean distance, the problem can be formalized as follows.

$$\begin{array}{l} \text{Minimize } D_{WOWA}(\mathbf{p} = (p_1, \dots, p_N), \mathbf{w} = (w_1, \dots, w_N)) = \\ \qquad\qquad \sum_{j=1}^{M} (\sum_{i=1}^{N} WOWA_{\mathbf{p},\mathbf{w}}(a_1^j, \dots, a_N^j) - b^j)^2 \\ Subject\ to \\ \qquad \sum_{i=1}^{N} p_i = 1 \\ \qquad \sum_{i=1}^{N} w_i = 1 \\ \qquad p_i \geq 0 \\ \qquad w_i \geq 0 \end{array} \tag{5}$$

To solve this problem, [18] used an hybrid approach that bootstrapped from the optimal solution obtained for the weighted mean and the OWA operator (following [16]), and then applied the gradient descent as proposed in [5, 6]. This hybrid approach is needed because the optimization problem in the case of the WOWA is not quadratic and there is no easy way to compute its optimal solution.

These results assume that data is complete, that is, there is no missing data. In the case of such data, we developed an approach based on genetic algorithms. Our approach, as well as some experiments, is reported in [12].

Finally, we have also considered the process of learning m-dimensional distorted probabilities. An approach for this type of problems is described in [11].

5 Summary

In this chapter we have described our main results about the WOWA operator, and some of its extensions. In particular, we have described m-dimensional WOWA operators. In addition, we have presented a short overview about the process of learning the weights of this operator. [20, 21] presents more details and some examples.

Acknowledgements. Partial support by the Generalitat de Catalunya (2005 SGR 00446 and 2005-SGR-00093) and by the Spanish MEC (projects ARES – CONSOLIDER INGENIO 2010 CSD2007-00004 – and eAEGIS – TSI2007-65406-C03-02) is acknowledged.

References

1. Aumann, R.J., Shapley, L.S.: Values of Non-Atomic Games. Princeton University Press, Princeton (1974)
2. Beliakov, G., Pradera, A., Calvo, T.: Aggregation Functions: A Guide for Practitioners. Springer, Heidelberg (2007)
3. Chateauneuf, A.: Decomposable measures, distorted probabilities and concave capacities. Mathematical Social Sciences 31, 19–37 (1996)
4. Choquet, G.: /54) Theory of capacities. Ann. Inst. Fourier 5, 131–295 (1953/1954)
5. Filev, D., Yager, R.R.: Learning OWA operator weights from data. In: Proc. of the 3rd IEEE Int. Conf. on Fuzzy Systems (IEEE WCCI), vol. 1, pp. 468–473 (1994)
6. Filev, D.P., Yager, R.R.: On the issue of obtaining OWA operator weights. Fuzzy Sets and Systems 94, 157–169 (1998)
7. Honda, A., Nakano, T., Okazaki, Y.: Distortion of fuzzy measures. In: Proc. of the SCIS/ISIS conference (2002)
8. Miranda, P., Grabisch, M.: p-symmetric fuzzy measures. In: Proc. of the IPMU 2002 Conference, pp. 545–552 (2002)
9. Miranda, P., Grabisch, M., Gil, P.: p-symmetric fuzzy measures. Int. J. of Unc., Fuzz. and Knowledge Based Systems 10, 105–123 (2002)
10. Narukawa, Y., Torra, V.: On n-dimensional distorted probabilities and p-symmetric fuzzy measures. In: IPMU 2004, Perugia, Italy, July 4-9, vol. 2, pp. 1279–1284 (2004) (ISBN 88-87242-54-2)
11. Narukawa, Y., Torra, V.: Fuzzy measure and probability distributions: distorted probabilities. IEEE Trans. on Fuzzy Systems 13(5), 617–629 (2005)
12. Nettleton, D., Torra, V.: A comparison of active set methods and genetic algorithm approaches for learning weighting vectors in some aggregation operators. Int. J. of Intel. Syst. 16(9), 1069–1083 (2001)

13. Torra, V.: Weighted OWA operators for synthesis of information. In: Proc. of the 5th IEEE Int. Conf. on Fuzzy Systems, pp. 966–971 (1996)
14. Torra, V.: The weighted OWA operator. Int. J. of Intel. Syst. 12, 153–166 (1997)
15. Torra, V.: On some relationships between the WOWA operator and the Choquet integral. In: Proc. of the IPMU 1998 Conference, pp. 818–824 (1998)
16. Torra, V.: On the learning of weights in some aggregation operators. Mathware and Soft Computing 6, 249–265 (1999)
17. Torra, V.: The WOWA operator and the interpolation function W*: Chen and Otto's interpolation method revisited. Fuzzy Sets and Systems 113(3), 389–396 (2000)
18. Torra, V.: Learning weights for Weighted OWA operators. In: Proc. of the IEEE Int. Conf. on Industrial Electronics, Control and Instrumentation (IECON 2000), pp. 2530–2535 (2000)
19. Torra, V., Lv, Z.: On the WOWA operator and its interpolation function. Int. J. of Intel. Systems (2009) (in press)
20. Torra, V., Narukawa, Y.: Modeling decisions: information fusion and aggregation operators. Springer, Heidelberg (2007)
21. Torra, V., Narukawa, Y.: Modelització de decisions: fusió d'informació i operadors d'agregació. UAB Press (2007)
22. Yager, R.R.: On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Trans. on Systems, Man and Cybernetics 18, 183–190 (1988)
23. Yager, R.R., Kacprzyk, J. (eds.): The Ordered Weighted Averaging Operators: Theory and Applications. Springer, Heidelberg (1997)