

Fitting ST-OWA operators to empirical data

Gleb Beliakov

Deakin University
Australia
gleb@deakin.edu.au

Luigi Troiano

RCOST - University of Sannio
82100 Benevento - Italy
troiano@unisannio.it

Abstract

The OWA operators gained interest among researchers as they provide a continuum of aggregation operators able to cover the whole range of compensation between the minimum and the maximum. In some circumstances, it is useful to consider a wider range of values, extending below the minimum and over the maximum. ST-OWA are able to surpass the boundaries of variation of ordinary compensatory operators. Their application requires identification of the weighting vector, the t-norm, and the t-conorm. This task can be accomplished by considering both the desired analytical properties and empirical data.

Keywords: Aggregation operators, OWA, t-norms, ST-OWA.

1 Introduction

The *Ordered Weighted Averaging* (OWA) operator [19] gained interest among researchers due to its property of providing a parameterized aggregation operation that includes the min, max and arithmetic mean. Due to this ability, OWA operators have been experimented in many problems regarding decision making, information retrieval and information fusion [6, 18, 20].

An OWA operator is defined as

$$M_{[\mathbf{w}]}(a_1, \dots, a_n) = \sum_{i=1}^n w_i a_{(i)} \quad (1)$$

where $a_{(1)}, \dots, a_{(n)}$ is a non-increasing permutation of arguments a_1, \dots, a_n , so that $a_{(i)} \geq a_{(j)} \forall i < j$.

The weighting vector $\mathbf{w} = (w_1, \dots, w_n)$ provides the parametrization of OWA operators. Weights are such

that

$$w_i \in [0, 1] \quad (2)$$

$$\sum_{i=1}^n w_i = 1 \quad (3)$$

Some notable examples of the weighting vector are:

- $M_{[1,0,\dots,0]}(a_1, \dots, a_n) = \max_{i=1..n} a_i$
- $M_{[1/n,\dots,1/n]}(a_1, \dots, a_n) = \frac{1}{n} \sum_{i=1}^n a_i$
- $M_{[0,\dots,0,1]}(a_1, \dots, a_n) = \min_{i=1..n} a_i$

The main property of OWA operators is that they are *internal* (i.e., compensatory, averaging) operators, as

$$\min_{i=1..n} a_i \leq M_{[\mathbf{w}]}(a_1, \dots, a_n) \leq \max_{i=1..n} a_i \quad (4)$$

for any weighting vector \mathbf{w} . This property is useful in many applications as it allows to compensate lower inputs with higher inputs. Nevertheless, in some circumstances it can be limitative because (i) there is an interaction between the arguments, or (ii) some reinforcement of higher and lower scores is required together with their compensation.

To overcome these limitations, the families of T-OWA, S-OWA and ST-OWA have been proposed and investigated. We can rewrite Eq.(1) as

$$M_{[\mathbf{w}]}(a_1, \dots, a_n) = \sum_{i=1}^n w_i \min(a_{(1)}, \dots, a_{(i)}) \quad (5)$$

This suggests that Eq.(1) can be generalized by using a generic t-norm T [11], as

$$M_{T[\mathbf{w}]}(a_1, \dots, a_n) = \sum_{i=1}^n w_i T(a_{(1)}, \dots, a_{(i)}) \quad (6)$$

This is the T-OWA operator introduced in [21]. In this case, the operator is provided with an additional

parameter which is a t-norm T . Different t-norms can be used. Usual choices include the minimum (\min), the product (T_P), the Lukasiewicz t-norm (T_L), and the drastic t-norm (T_D). Other possibilities entail the use of parametric t-norms, such as Schweizer and Sklar's t-norms ($T_{SS(p)}$) and Yager's t-norms ($T_{Y(v)}$) [11].

Eq.(1) can also be rewritten as

$$M_{[\mathbf{w}]}(a_1, \dots, a_n) = \sum_{i=1}^n w_i \max(a_{(i)}, \dots, a_{(n)}), \quad (7)$$

which leads to the definition of S-OWA operators [21] as

$$M_{S[\mathbf{w}]}(a_1, \dots, a_n) = \sum_{i=1}^n w_i S(a_{(i)}, \dots, a_{(n)}). \quad (8)$$

The additional parameter is provided by a t-conorm S . The usual choices are the maximum (\max), the probabilistic sum (S_P), the Lukasiewicz t-conorm (S_L), and the drastic t-conorm (S_D). Also in this case, it is possible to use parametric t-conorms such as Schweizer and Sklar's t-conorms ($S_{SS(p)}$) and the Yager's t-conorm ($S_{Y(v)}$).

The S-OWA and T-OWA operators extend the range of values provided by the ordinary OWA operators, as depicted in Fig.1.

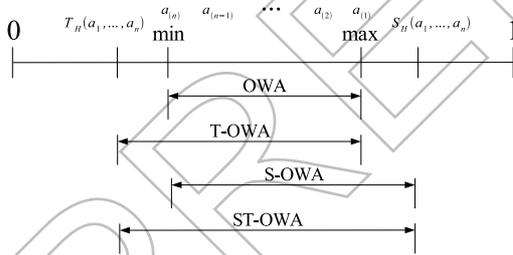


Figure 1: Range of values

In particular, the T-OWA extends the range of values below the minimum, whilst the S-OWA extends the range above the maximum. To extend the range in both directions (see Fig.1), the ST-OWA has been defined as [17]

$$M_{ST[\mathbf{w}]}(a_1, \dots, a_n) = \sum_{i=1}^n w_i ((1 - \sigma)T(a_{(1)}, \dots, a_{(i)}) + \sigma S(a_{(i)}, \dots, a_{(n)})), \quad (9)$$

where $\sigma \in [0, 1]$ is the OWA attitudinal character, as described in the following section. In this case, the t-norm T and the t-conorm S provide the two parameters additional to the weighting vector \mathbf{w} .

An important issue for practical applicability of ST-OWA operators is the identification of parameters, which aims at properly choosing the t-norm, the t-conorm and the weighting vector. This task can be performed by considering both the desired analytical properties and empirical data. In particular, when parametric t-norms are considered, the identification can be translated into an optimization problem aimed at finding the parameter(s) of a t-norm (t-conorm) and the OWA weighting vector that best fit the empirical data. The remainder of this paper is devoted to the solution of this problem. Section 2 provides an overview of the main analytical properties of ST-OWA operators. Section 3 deals with the identification problem. Section 4 describes an example of an application. Section 5 briefly outlines conclusions and future work.

2 Properties

Different weighting vectors entail different emphasis to higher and lower input values. This is described by the *attitudinal character*, also known as the *orness measure*, which is a function of weights defined as

$$AC(\mathbf{w}) = \sum_{i=1}^n w_i \frac{n-i}{n-1} \in [0, 1]. \quad (10)$$

In particular

$$AC(\mathbf{w}) = \begin{cases} 1 & \text{if } w_1 = 1, w_{i \neq 1} = 0, \\ 0.5 & \text{if } w_i = 1/n, \\ 0 & \text{if } w_n = 1, w_{i \neq n} = 0. \end{cases} \quad (11)$$

The attitudinal character itself can be computed by using OWA operators.

Proposition 1.

$$AC(\mathbf{w}) = M_{[\mathbf{w}]}(1, \frac{n-2}{n-1}, \dots, \frac{1}{n-1}, 0). \quad (12)$$

As discussed in [21], Prop.1 suggests a way for computing the attitudinal character of T-OWA operators as

$$AC(\mathbf{w}, T) = M_{T[\mathbf{w}]}(1, \frac{n-2}{n-1}, \dots, \frac{1}{n-1}, 0). \quad (13)$$

In a similar way we can compute the attitudinal character of S-OWA operators as

$$AC(\mathbf{w}, S) = M_{S[\mathbf{w}]}(1, \frac{n-2}{n-1}, \dots, \frac{1}{n-1}, 0), \quad (14)$$

and for ST-OWA operators, as

$$AC(\mathbf{w}, S, T) = M_{ST[\mathbf{w}]}(1, \frac{n-2}{n-1}, \dots, \frac{1}{n-1}, 0). \quad (15)$$

Moreover, for any choice of T and S , we get

$$AC(\mathbf{w}, S, T) = (1 - \sigma)AC(\mathbf{w}, T) + \sigma AC(\mathbf{w}, S), \quad (16)$$

where $\sigma = AC(\mathbf{w})$.

T-norms and t-conorms can be compared with respect to the aggregated value they provide.

Definition 1. Given two t-norms T_1 and T_2 , T_1 is stronger than T_2 iff

$$T_1(x, y) \geq T_2(x, y) \quad \forall x, y \in [0, 1]. \quad (17)$$

In this case we write $T_1 \geq T_2$.

It can be easily proven that $T_D \leq T_L \leq T_P \leq \min$. The same definition can be applied to t-conorms as well. In this case, it results in $\max \leq S_P \leq S_L \leq S_D$. There is a kind of symmetry between t-norms and t-conorms, due to the duality. With respect to parametric families of t-norms (t-conorms), we note that the values of the parameter provide a natural ordering for some families (although this is not true in general). For instance, with respect to Yager's t-norm and t-conorm, defined as

$$T_{Y(v)}(x, y) = 1 - \min(1, ((1-x)^v + (1-y)^v)^{\frac{1}{v}}), \quad (18)$$

$$S_{Y(v)}(x, y) = \min(1, (x^v + y^v)^{\frac{1}{v}}), \quad (19)$$

increasing values of the parameter v entail stronger t-norms (weaker t-conorms). It is obvious that t-conorms are stronger than t-norms.

The order relation between two t-norms (t-conorms), leads to ordering of T-OWA (S-OWA) operators. Indeed

Proposition 2. Given two t-norms (or t-conorms) R_1 and R_2 , such that $R_1 \geq R_2$, it holds

$$M_{R_1[\mathbf{w}]}(a_1, \dots, a_n) \geq M_{R_2[\mathbf{w}]}(a_1, \dots, a_n) \quad (20)$$

for any weighting vector \mathbf{w} , and any arguments a_1, \dots, a_n . Then we write

$$M_{R_1} \geq M_{R_2}. \quad (21)$$

For any weighting vector \mathbf{w} , it holds

$$M_{T_D} \leq M_{T_L} \leq M_{T_P} \leq M \leq M_{S_P} \leq M_{S_L} \leq M_{S_D}. \quad (22)$$

For any, weighting vector \mathbf{w} , if $T_1 \leq T_2$ ($S_1 \geq S_2$) then

$$\begin{aligned} AC(\mathbf{w}, T_1) &\leq AC(\mathbf{w}, T_2), \\ (AC(\mathbf{w}, S_1) &\geq AC(\mathbf{w}, S_2)). \end{aligned} \quad (23)$$

Proposition 3. For any t-norm T and t-conorm S ,

$$AC(\mathbf{w}, T) \leq AC(\mathbf{w}) \leq AC(\mathbf{w}, S). \quad (24)$$

An important property of the OWA operators (as well as any aggregation operator), is their *monotonicity* with respect to arguments.

Proposition 4. Given two vectors of n arguments, a_1, \dots, a_n and b_1, \dots, b_n , such that $a_i \geq b_i \forall i = 1..n$, it holds

$$M_{[\mathbf{w}]}(a_1, \dots, a_n) \geq M_{[\mathbf{w}]}(b_1, \dots, b_n) \quad (25)$$

for any weighting vector \mathbf{w} .

Duality of t-norms and OWA operators can be translated to the ST-OWA setting. Recall that the dual weighting vector $\hat{\mathbf{w}}$ is obtained from \mathbf{w} using $\hat{w}_i = w_{n-i+1}$. Similarly to ordinary OWA operators, we can easily prove the following proposition.

Proposition 5. Given a pair of weighting vectors $\hat{\mathbf{w}}$ and \mathbf{w} , if $\hat{\mathbf{w}}$ is dual to \mathbf{w} , and S is dual to T , then $M_{ST[\hat{\mathbf{w}}]}$ and $M_{ST[\mathbf{w}]}$ are dual.

Definition 2. Given an aggregation operator $R_{\mathbf{w}}$ which depends on the weighting vector \mathbf{w} , and given a vector $\mathbf{a} \in [0, 1]^n$, its variation range at \mathbf{a} is the interval $r(\mathbf{a}) = [\inf_{\mathbf{w}} R_{\mathbf{w}}(\mathbf{a}), \sup_{\mathbf{w}} R_{\mathbf{w}}(\mathbf{a})]$. The aggregation operator $R_{\mathbf{w}}$ has a greater variation range than $S_{\mathbf{w}}$ if $s(\mathbf{a}) \subset r(\mathbf{a})$ for all $\mathbf{a} \in [0, 1]^n$.

If the dual pair is $\langle \min, \max \rangle$ we get the ST-OWA with the minimal variation range: it corresponds to the case of ordinary OWA operators. If the dual pair is $\langle T_D, S_D \rangle$, we get the ST-OWA with maximum variation. In the case of a dual pair $\langle T, S \rangle$, the attitudinal character is

$$\begin{aligned} AC(\mathbf{w}, S, T) &= (1 - \sigma)AC(\mathbf{w}, T) + \sigma(1 - AC(\hat{\mathbf{w}}, T)) \\ &= (1 - \sigma)(1 - AC(\hat{\mathbf{w}}, S)) + \sigma AC(\mathbf{w}, S) \end{aligned} \quad (26)$$

In particular,

$$AC(\mathbf{w}, S, T) = \begin{cases} 0, & \sigma = 0 \\ 1/2, & \sigma = 1/2 \\ 1, & \sigma = 1 \end{cases} \quad (27)$$

Example 1. Let us consider $\mathbf{a} = [0.9, 0.7, 0.8, 0.9, 0.8]$. The range of variation by varying $\sigma \in [0, 1]$ for an ordinary OWA is

$$\langle \min, \max \rangle: [0.700, 0.900]$$

If we choose a different pair $\langle T, S \rangle$, the range of variation at \mathbf{a} becomes

$$\langle T_P, S_P \rangle: [0.363, 0.999]$$

$$\langle T_L, S_L \rangle: [0.100, 1.000]$$

$$\langle T_D, S_D \rangle: [0.000, 1.000]$$

In particular, if we choose $\mathbf{w} = [0.25, 0.25, 0.5, 0, 0]$, then

$$\sigma = AC(\mathbf{w}) = 0.687$$

and

$$AC(\mathbf{w}, S_P, T_P) = 0.738$$

$$AC(\mathbf{w}, S_L, T_L) = 0.718$$

$$AC(\mathbf{w}, S_D, T_D) = 0.824$$

3 Operator identification

In this section we consider various instances of the problem of fitting parameters of ST-OWA to empirical data. We assume that there is a set of input-output pairs $\mathcal{D} = \{(\mathbf{x}_k, y_k)\}, k = 1, \dots, K$, with $\mathbf{x}_k \in [0, 1]^n$, $y_k \in [0, 1]$ and n is fixed. Our goal is to determine parameters S, T, \mathbf{w} which fit the data best.

3.1 Identification with fixed S and T

In this instance of the problem we assume that both S and T have been specified. The issue is to determine the vector \mathbf{w} . For S-OWA and T-OWA, fitting the data in the least squares sense involves a solution to a quadratic programming problem (QP)

$$\begin{aligned} \text{Minimize} \quad & \sum_{k=1}^K \left(\sum_{i=1}^n w_i S(x_{(i)k}, \dots, x_{(n)k}) - y_k \right)^2 \\ \text{s.t.} \quad & \sum_{i=1}^n w_i = 1, w_i \geq 0, \end{aligned} \quad (28)$$

and similarly for the case of T-OWA. We note that the values of S at any \mathbf{x}_k are fixed (do not depend on \mathbf{w}). This problem is very similar to that of calculating the weights of standard OWA operators from data [2, 4, 7, 16], but involves a fixed function $S(x_{(i)k}, \dots, x_{(n)k})$ rather than just $x_{(i)k}$.

If an additional requirement is to have a specified value of $AC(\mathbf{w}, S)$ and $AC(\mathbf{w}, T)$, then it becomes just an additional linear constraint, which does not change the structure of QP problem (28).

Next, consider fitting ST-OWA. Here, for a fixed value of $AC(\mathbf{w}) = \sigma$, we have the QP problem

$$\begin{aligned} \text{Minimize} \quad & \sum_{k=1}^K \left(\sum_{i=1}^n w_i ST(\mathbf{x}_k, \sigma) - y_k \right)^2 \\ \text{s.t.} \quad & \sum_{i=1}^n w_i = 1, w_i \geq 0, \\ & AC(\mathbf{w}) = \sigma, \end{aligned} \quad (29)$$

where

$$ST(\mathbf{x}, \sigma) = (1 - \sigma)T(x_{(1)}, \dots, x_{(i)}) + \sigma S(x_{(i)}, \dots, x_{(n)}).$$

However σ may not be specified, and hence has to be also found from the data. In this case, we present a bi-level optimization problem, in which at the outer level nonlinear (possibly global) optimization is performed with respect to parameter σ , and at the inner level the problem (29) with a fixed σ ,

$$\text{Minimize}_{\sigma \in [0, 1]} [F(\sigma)], \quad (30)$$

where $F(\sigma)$ denotes the solution to (29).

We note that the inner quadratic problem (29) has a unique global minimum, which is easily found by using standard numerical methods, see [8, 12]. If a global optimization method is applied to the outer problem with respect to σ , then the globally optimal solution to (30) with respect to both σ and \mathbf{w} will be obtained. Numerical solution to the outer problem with just one variable can be performed by a number of methods, including grid search, multistart local search, or Pijavski-Shubert method [9, 13].

3.2 Identification of T-OWA and S-OWA

Consider now the problem of fitting parameters of the parametric families of participating t-norm and t-conorm, simultaneously with \mathbf{w} and σ . With start with S-OWA, and assume a suitable family of t-norms T has been chosen, e.g., Yager t-conorms $S_{Y(v)}$ parameterized with v . We will rely on efficient solution to problem (28) with a fixed S (i.e., fixed v). We set up a bi-level optimization problem

$$\text{Minimize}_{v \in [0, \infty]} [F_1(v)],$$

where $F_1(v)$ denotes solution to (28).

The outer problem is nonlinear, possibly global optimization problem, but because it has only one variable, its solution is relatively simple. We recommend Pijavski-Shubert deterministic method [13]. Identification of T is performed analogously.

The advantage of using bi-level optimization is that the nonlinear parameter (v) is separated from the vector of weights, which is found by solving a standard QP. Hence for the nonlinear problem we have just one variable, and for multivariate problem with respect to \mathbf{w} we have a special structure, accommodated by efficient QP algorithms. Since the inner QP problem is convex, it has a unique global minimum, and the whole problem with respect to all parameters is solved to global optimality.

Next consider fitting ST-OWA operators. Here we have three parameters: the two parameters of the participating t-norm and t-conorm, which we will denote by v_1, v_2 , and σ as in Problem (29). Of course, T and S may be chosen as dual to each other, in which case we have to fit only one parameter $v = v_1 = v_2$. To use the special structure of the problem with respect to \mathbf{w} we again set up a bi-level optimization problem analogously to (30).

$$\text{Minimize}_{\sigma \in [0, 1], v_1, v_2 \geq 0} [F(\sigma, v_1, v_2)], \quad (31)$$

where $F(\sigma, v_1, v_2)$ is the solution to QP problem

$$\begin{aligned} \text{Minimize} \quad & \sum_{k=1}^K \left(\sum_{i=1}^n w_i ST(\mathbf{x}_k, \sigma, v_1, v_2) - y_k \right)^2 \quad (32) \\ \text{s.t.} \quad & \sum_{i=1}^n w_i = 1, w_i \geq 0, \\ & AC(\mathbf{w}) = \sigma, \end{aligned}$$

and

$$\begin{aligned} ST(\mathbf{x}, \sigma, v_1, v_2) = \\ (1 - \sigma)T_{Y(v_1)}(x_{(1)}, \dots, x_{(i)}) + \sigma S_{Y(v_2)}(x_{(i)}, \dots, x_{(n)}). \end{aligned}$$

Solution to the outer problem is complicated because of the possibility of numerous local minima. One has to rely on methods of multivariate global optimization [9, 14]. One such (deterministic) method is the Cutting Angle Method (CAM) developed in [1, 3, 15]. It allows one to solve efficiently global optimization problems in up to 10 variables.

3.3 Least absolute deviation problem

Besides the least squares approach, fitting to the data can be performed by using the Least Absolute Deviation (LDA) criterion [5], by replacing the sum of squares in (28) and (29) with the sum of absolute values. It is argued that the LDA criterion is less sensitive to outliers.

In this case the optimization problems are converted to linear programming (LP) problems by introducing auxiliary non-negative variables r_k^+, r_k^- , such that

$$r_k^+ - r_k^- = \sum_{i=1}^n w_i ST(\mathbf{x}_k, \sigma) - y_k,$$

and

$$r_k^+ + r_k^- = \left| \sum_{i=1}^n w_i ST(\mathbf{x}_k, \sigma) - y_k \right|.$$

This conversion is well known, see [5]. The counterparts of problems (28) and (29) become LP problems, which are easily solved by the simplex method. The outer nonlinear optimization problems do not change.

3.4 Preservation of ordering of the outputs

In [10] it was argued that fitting the numerical outputs is not as important as preserving the ordering of the outputs. The empirical data usually comes from human subjective evaluation, and people do not reliably express their preference on a numerical scale. In contrast, people are very good at ranking the alternatives. Therefore, the authors of [10] argued that fitting methods should aim at preserving the order of empirical output values. They showed that various methods

of fitting the numerical values do not preserve this ordering.

Without loss of generality, we assume that the outputs are ordered as $y_1 \leq y_2 \leq \dots \leq y_K$. The condition for order preservation is

$$M_{ST[\mathbf{w}]}(\mathbf{x}_k) \leq M_{ST[\mathbf{w}]}(\mathbf{x}_{k+1}), \text{ for all } k = 1, \dots, K - 1. \quad (33)$$

Because $M_{ST[\mathbf{w}]}$ depends on \mathbf{w} linearly for a fixed σ , (33) is a system of linear inequalities, which does not change the structure of the QP or LP problems.

For example, problem (28) will have additional $K - 1$ linear constraints for $k = 1, \dots, K - 1$:

$$\sum_{i=1}^n w_i (S(x_{(i)k+1}, \dots, x_{(n)k+1}) - S(x_{(i)k}, \dots, x_{(n)k})) \geq 0.$$

Problem (29) will have the constraints

$$\sum_{i=1}^n w_i (ST(\mathbf{x}_{k+1}, \sigma) - ST(\mathbf{x}_k, \sigma)) \geq 0.$$

4 Numerical experiments and examples

As a first step of testing the correctness and suitability of the mentioned algorithms for determination of ST-OWA parameters, we have generated random data $\mathbf{x}_k, k = 1, \dots, 20$ and the values of y_k , computed by a model S-, T- and ST-OWA aggregation operators in 3-5 variables (i.e., with fixed weights and fixed participating t-norm and t-conorm from Yager, Hamacher and Frank families). Then we used these data to calculate:

- The weighting vector \mathbf{w} of S-, T- and ST-OWA operators with known parameter(s) $v(v_1, v_2)$.
- The weighting vector and the unknown parameter v of S- and T-OWA operators.
- The weighting vector and the unknown parameter v of S- and T-OWA operators with a given attitudinal character.
- The weighting vector and the unknown parameters v_1, v_2 of ST-OWA operators with a given attitudinal character σ .
- The weighting vector and the unknown parameters v_1, v_2 and σ of ST-OWA operators.

In the test cases we also included the limiting cases of ST-OWA being a pure t-norm, t-conorm or OWA. All our experiments were successful, the correct values of the parameters used in the models used to simulate the data have been found. The computing time for S-, T-

and ST-OWA operators with given v (v_1, v_2) was below 1 sec (Pentium IV 2 GHz workstation), and when parameters v, v_1, v_2, σ also had to be fitted to the data, the computing time was < 10 sec. These experiments show the robustness and efficiency of the proposed algorithms.

The algorithm robustness is provided by the non-sensitivity of precision to structural attributes, such as the number of criteria n or by the number of input values K . Indeed, the non-sensitivity of precision to the number of criteria n , guarantees the algorithm to be able to converge also with a growing dimension of the input space. In order to verify this hypothesis we considered, for each $n = 3, 10, 20$ random samples, each made of 20 n -ples of input value; the input values were aggregated assuming random value of the weighting vector and t-norm/t-conorm parameters. For each sample we collected the RMSE and coefficients between the observed and calculated values. ANOVA procedure showed the null hypothesis (i.e. $H_0 = \text{there at least two samples whose RMSE, or correlation coefficient, are statistically different}$) can be rejected ($p < 0.01$). This resulted for T-, S- and ST-OWA with respect to Frank, Hamacher and Yager norms. Similarly, we assumed $n = 3$ and for $K = 10, 20, \dots, 100$ we considered 20 random samples. Again, ANOVA procedure showed that RMSE and correlation coefficients are not statistically different ($p < 0.01$).

An interesting property that emerged from experimentation was the ability of the algorithm to compensate noise, as depicted by Table 1.

Run	Noise RMSE	Approx. RMSE
1	0.05312	0.00955
2	0.06228	0.01205
3	0.05512	0.00996
4	0.06213	0.01160
5	0.05315	0.00979
6	0.04991	0.01285
7	0.06202	0.01091
8	0.05437	0.01123
9	0.05621	0.01176
10	0.05947	0.01227

Table 1: Noise compensation

For this test we assumed $n = 3$ and $K = 20$. For each input 3-ple we computed the aggregated value using the Hamacher t-norm, given the weighting vector and the t-norm parameter. Then, we added to the aggregated value uniformly distributed errors in the range $[-0.1, 0.1]$. The resulted value was given to the algorithm in order to identify the ST-OWA operator best fitting the data. Table 1 shows the result on this experimentation on a sample of 10 runs. In column 2 we

report the RMSE affecting the expected values, while in column 3 the RMSE between the expected and observed values. We can note that the two series are statistically different (the Wilcoxon matched-pairs signed-ranks test on 50 runs confirmed this hypothesis). Therefore, the algorithm behaves in such a way as to identify parameters that are different from the original ones, but that provides an ST-OWA operator able to better approximate the expected values. Similar conclusions can be obtained for S- and T-OWA operators, and for Yager and Frank t-norms.

Next we took the actual experimental data (two data sets of 20 data) from [22, 23] (also replicated in [10]). For both data sets, the root mean square error of approximation were $RMSE_1 = 0.0148$ and $RMSE_2 = 0.0105$, which compares favorably with the fitted standard OWA ($RMSE_1 = 0.015$, $RMSE_2 = 0.011$) and Zimmermann's γ operators ($RMSE_1 = 0.0151$, $RMSE_2 = 0.0105$). The correlation coefficients (between the observed and calculated values) has also been higher ($Corr_1 = 0.985$, $Corr_2 = 0.974$). While the gain in numerical accuracy is not very impressive, we note that for the two data sets, we used different participating t-norm and t-conorm in γ -operators (max/min and T_P/S_P , by trial and error process) and reported the best result, while for ST-OWA operators the parameters were found automatically. We also note that both data sets have only two inputs.

Finally, we used three empirical data sets with 4 inputs, collected by the second author, described as follows. A group of students ($K = 41$) was asked to provide their numerical evaluation of the quality of three objects: public broadcast TV programs, University and town they live in. First they provided an overall score (y) and then scores with respect to four criteria (x_1, \dots, x_4), namely, quality of programs, advertisement pressure, sports, news (for TV); curriculum, potential for personal growth, quality of labs, other services (University); public events, criminality, cleanliness and services to young people (town). There were no missing data. There were clear outliers (e.g., respondent #14 has provided scores like $\mathbf{x} = (23, 25, 23, 10)$, $y = 85$ despite the overall averaging tendency for the group), however outliers were not removed for the analysis.

For the three data sets we obtained $RMSE_1 = 0.022$, $RMSE_2 = 0.018$ and $RMSE_3 = 0.021$, with correlation between computed and observed values 0.76, 0.82, 0.79 respectively. These values were the best we obtained among OWA, γ -operators (using min/max and T_P/S_D), various families of t-norms, uninorms, generalized means and Choquet integrals, although weighted generalized means gave close results. This shows the potential of ST-OWA operators in fitting experimental data.

Calculations were performed using the AOTool software available from <http://www.deakin.edu.au/~gleb/aotool.html>.

5 Conclusions and future work

We have studied the problem of identification of ST-OWA aggregation operators from empirical data. ST-OWA provide a broader range of behavior, from conjunctive via averaging to disjunctive, within one class of aggregation operators, parameterized with T , S and weighting vector \mathbf{w} . They are useful in the situations where it is unknown a priori which type of behavior is consistent with the data.

We have presented several mathematical programming formulations of the problem of identifying ST-OWA parameters from empirical data. An interesting feature of these problems is that they are set as nonlinear optimization problems with respect to just one or three parameters, with the vector of weights identified by standard quadratic or linear programming methods. This, of course, greatly reduces computational costs.

Our future line of research is to replace OWA with more general Choquet integral type construction, as well as to consider generalized OWA and nonparametric families of continuous t -norms.

References

- [1] G. Beliakov. Geometry and combinatorics of the cutting angle method. *Optimization*, 52:379–394, 2003.
- [2] G. Beliakov. How to build aggregation operators from data? *Int. J. Intelligent Systems*, 18:903–923, 2003.
- [3] G. Beliakov. The cutting angle method - a tool for constrained global optimization. *Optimization Methods and Software*, 19:137–151, 2004.
- [4] G. Beliakov. Learning weights in the generalized OWA operators. *Fuzzy Optimization and Decision Making*, 4:119–130, 2005.
- [5] P. Bloomfield and W.L. Steiger. *Least Absolute Deviations. Theory, Applications and Algorithms*. Birkhauser, Boston, Basel, Stuttgart, 1983.
- [6] B. Bouchon-Meunier. *Aggregation and Fusion of Imperfect Information*. Physica-Verlag, Heidelberg, 1998.
- [7] D. Filev and R. Yager. On the issue of obtaining OWA operator weights. *Fuzzy Sets and Systems*, 94:157–169, 1998.
- [8] E. M. Gertz and S. J. Wright. Object-oriented software for quadratic programming. *ACM Trans. Math. Software*, 29:58–81, 2003.
- [9] R. Horst, P. Pardalos, and N. Thoai. *Introduction to Global Optimization*. Kluwer Academic Publishers, Dordrecht, 2nd edition, 2000.
- [10] U. Kaymak and H.R. van Nauta Lemke. Selecting an aggregation operator for fuzzy decision making. In *3rd IEEE Intl. Conf. on Fuzzy Systems*, volume 2, pages 1418–1422, 1994.
- [11] E.P. Klement, R. Mesiar, and E. Pap. *Triangular Norms*. Kluwer, Dordrecht, 2000.
- [12] C. Lawson and R. Hanson. *Solving Least Squares Problems*. SIAM, Philadelphia, 1995.
- [13] S.A. Pijavski. An algorithm for finding the absolute extremum of a function. *USSR Comput. Math. and Math. Phys.*, 2:57–67, 1972.
- [14] J. Pinter. *Global Optimization in Action: Continuous and Lipschitz Optimization—Algorithms, Implementations, and Applications*. Kluwer Academic Publishers, Dordrecht; Boston, 1996.
- [15] A.M. Rubinov. *Abstract Convexity and Global Optimization*. Kluwer Academic Publishers, Dordrecht; Boston, 2000.
- [16] V. Torra. OWA operators in data modeling and reidentification. *IEEE Trans. on Fuzzy Syst.*, 12:652–660, 2004.
- [17] L. Troiano and R.R. Yager. On some properties of mixing OWA operators with t -norms and t -conorms. In *Proc. EUSFLAT 2005*, volume Vol.II, pages 1206–1212, Barcellona, Spain, 2005.
- [18] R. Yager and J. Kacprzyk, editors. *The Ordered Weighted Averaging Operators. Theory and Applications*. Kluwer, Boston, 1997.
- [19] R.R. Yager. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Transactions on Systems, Man and Cybernetics*, 18:183–190, 1988.
- [20] R.R. Yager. Families of OWA operators. *Fuzzy Sets and Systems*, 59:125–148, 1993.
- [21] R.R. Yager. Extending multicriteria decision making by mixing t -norms and OWA operators. *Int. J. Intelligent Systems*, 20:453–474, 2005.
- [22] H.-J. Zimmermann. *Fuzzy Set Theory - and its Applications*. Kluwer, Boston, 1996.
- [23] H.-J. Zimmermann and P. Zysno. Latent connectives in human decision making. *Fuzzy Sets and Systems*, 4:37–51, 1980.

PREPRINT