Estimating the Spanish Energy Demand Using Variable Neighborhood Search

Jesús Sánchez-Oro[1(✉)], Abraham Duarte[1], and Sancho Salcedo-Sanz[2]

[1] Department of Computer Science, Universidad Rey Juan Carlos, Móstoles, Spain
{jesus.sanchezoro,abraham.duarte}@urjc.es
[2] Department of Signal Processing and Communications,
Universidad de Alcalá, Madrid, Spain
sancho.salcedo@uah.es

Abstract. The increasing of the energy demand in every country has lead experts to find strategies for estimating the energy demand of a given country for the next year. The energy demand prediction in the last years has become a hard problem, since there are several factors (like economic crisis, industrial globalization, or population variation) that are not easy to control. For this reason, it is interesting to propose new strategies for efficiently perform this estimation. In this paper we propose a metaheuristic algorithm based on the Variable Neighborhood Search framework which is able to perform an accurate prediction of the energy demand for a given year. The algorithm is supported in a previously proposed exponential model for estimating the energy, and its input is conformed with a set of macroeconomic variables gathered during the last years. Experimental results show the excellent performance of the algorithm when compared with both previous approaches and the actual values.

Keywords: Energy demand · VNS · Metaheuristics · Estimation

1 Introduction

The estimation of the energy demand for a given country has become a relevant problem that needs to be tackled each year in every country [4]. Furthermore, the energy demand is constantly increasing, mainly due to the global industrialization experienced in the last decades. It is worth mentioning that approximately 50 % of the total energy demand of a given country is performed by the industry, which clearly shows the importance of this sector in the energy demand estimation. The variation in the population as well as the globalization are two additional key factors in the estimation of the energy demand of a country.

This problem has been previously addressed using different heuristic approaches. Most of the previous works have been devoted to predict the energy demand in developing countries like Turkey [1,4,7–9,14–16]. However, several algorithms have been proposed to perform a similar task in China [17–19], Iran [12], Korea [5], and Spain [13], among others. The algorithms proposed in these

O. Luaces et al. (Eds.): CAEPIA 2016, LNAI 9868, pp. 341–350, 2016.
DOI: 10.1007/978-3-319-44636-3_32

works base their effectiveness on the use of socio-economic variables. This fact can be illustrated with the Genetic Algorithm (GA) proposed in [1], which introduces a model based on four macroeconomic variables: Gross Domestic Product (GDP), population, import and export size. These variables are used again in a Multilayer Perceptron Neural Network [5] proposed for estimating the energy demand in Korea.

More bio-inspired heuristic algorithms have been used for solving this task. For instance, Particle Swarm Optimization (PSO) has been used as a standalone algorithm [9,16] and hybridized with other bio-inspired algorithms, as Ant Colony Optimization (ACO) and GA, for estimating the energy demand in several countries [8,17–19]. Several prediction models based on logarithmic and exponential functions have been proposed [12], supported by a GA for energy demand estimation in Iran. Finally, the prediction of the energy demand in Spain has been recently tackled with a Harmony Search algorithm combined with Extreme Learning Machine [13].

This work is intended to solve the estimation of the Spanish energy demand one-year-ahead, selecting the most relevant set of socio-economic variables, among the ones used in previous works. The prediction is performed by using a previously proposed exponential model [13], where a Variable Neighborhood Search (VNS) algorithm is responsible of selecting the best parameters for the model.

The remainder of the paper is structured as follows: Sect. 2 presents the exponential model used for estimating the energy demand. The algorithmic proposal is introduced in Sect. 3, while the experiments performed to test the VNS method are described in Sect. 4. Finally, the conclusions derived from the experiments are discussed in Sect. 5.

2 Prediction Model

The estimation of the energy demand in a given country can be represented by different models that try to fit a function to the real energy demand. We can find in the literature several models, based on different mathematical functions, but most of works concludes that the exponential model fits better to the real energy demand of the country [12]. Furthermore, in this paper we focus on estimating the energy demand of Spain, where the best results have been found by using the exponential model [13]. For that reason, our algorithmic proposal is intended to optimize this model.

The exponential function to estimate the energy demand $\hat{E}(t)$ for a given year t is defined based on the values X_i of the considered macroeconomic variables in the previous year $t-1$. In mathematical terms,

$$\hat{E}(t) = \epsilon + \sum_{i=1}^{|X|} \alpha_i X_i (t-1)^{\beta_i}$$

where ϵ, α_i and β_i are the parameters that need to be fitted in order to improve the energy demand prediction. Specifically, ϵ represents a bias in order to obtain a better fit, while α_i, β_i represent the weight of the macroeconomic variable X_i.

Unlike the original model [13], we do not fix the macroeconomic variables that are included in the model. Therefore, this model requires two different optimizations: (i) select the best subset of macroeconomic variables, and (ii) find the best values for the model parameters.

A solution $\mathcal{M}$ for the problem is given by both the subset of macroeconomic variables and the values for the parameters of the model. Let us define $X^\star$ as the set of all available macroeconomic variables. The solution is divided in two well-differenced sets: $X = \{X_1, X_m\} \in X^\star$ for the selected macroeconomic variables and $W = \{\epsilon, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m\}$ for the bias and variable weights.

For the sake of clarity, we describe the evaluation of an example solution by using the proposed representation. We consider the solution $\mathcal{M} = (X, W)$, with $X = \{ME1, ME3, ME4\}$ and $W = \{4.1, 0.4, 0.1, 0.2, -0.05, 0.3, -0.1\}$. In this solution, we have selected the macroeconomic variables $X_1 = ME1, X_2 = ME3, X_3 = ME4$, where each one (*ME1, ME3, ME4*) represents a different value (i.e., GDP, import and export size, etc.). In order to estimate the energy demand for a given year t, we first need to obtain the variable values for the previous year $t-1$. For this example, we set $X_1 = 0.1, X_2 = 0.4, X_3 = 0.2$. Notice that for each variable X_i (see for instance $\alpha_1 = 0.4, \beta_1 = 0.1$ for X_1) we provide a pair of coefficients α_i, β_i, plus the first bias coefficient ϵ (4.1 in the example).

The estimation of the energy demand for the current year is performed as follows:

$$\begin{aligned}\hat{E}(t) &= \epsilon + \alpha_1 \cdot X_1(t-1)^{\beta_1} + \alpha_2 \cdot X_2(t-1)^{\beta_2} + \alpha_3 \cdot X_3(t-1)^{\beta_3} \\ &= 4.1 + 0.4 \cdot 0.1^{0.1} + 0.2 \cdot 0.4^{-0.05} + 0.3 \cdot 0.2^{-0.1} \\ &= 4.98\end{aligned}$$

Each solution is evaluated using an objective function to analyze the similarity of the solution when compared with the real energy demand. In this work we minimize the mean squared error (MSE) between prediction and real value. The MSE is defined as follows:

$$f(\mathcal{M}) = \sum_{j=1}^{n} \left(E(j) - \hat{E}(j)\right)^2$$

where n is the number of samples used to test the quality of the solution. We use the same error measure previously considered in the related works [1,13] in order to ease the comparison among different proposals. Therefore, the objective is to find the solution $\mathcal{M}^\star$ with the minimum f-value. In mathematical terms,

$$\mathcal{M}^\star = \arg\min_{\mathcal{M} \in \mathbb{M}} f(\mathcal{M})$$

where $\mathbb{M}$ is the set of all possible solutions for the problem.

It is worth mentioning that all variable values and parameters are normalized in a traditional manner. Specifically, the value of each macroeconomic variable and each α, β parameter is normalized in the range $[-1, 1]$. However, the bias parameter considers a larger range $[-5, 5]$, as recommended in previous works [13], in order to obtain better results.

3 Variable Neighborhood Search

Variable Neighborhood Search (VNS) is a metaheuristic framework designed for solving hard optimization problems. VNS relies on systematic changes of neighborhood structures for finding high quality solutions, without guaranteeing their optimality. Several VNS strategies have been proposed during the study of this metaheuristic. Specifically, we can emphasize Variable Neighborhood Descent (VND), Reduced VNS (RVNS), Basic VNS (BVNS), Skewed VNS (SVNS), General VNS (GVNS), Variable Neighborhood Decomposition Search (VNDS), and Variable Formulation Search (VFS) among others (see [6] for a thoroughly review of the metaheuristic and its variants). This methodology has lead to several successfull research in recent years [3,10,11].

This paper is focused on the Basic VNS (BVNS) variant, which combines deterministic and stochastic changes of neighborhood during the search. Algorithm 1 shows the pseudocode of BVNS. The algorithm requires two parameters: an initial solution $\mathcal{M}$ and the maximum number of iterations. The procedure starts by exploring the neighborhoods from the first one (step 1). For each neighborhood, it randomly perturbs the current solution (step 4), with the shake procedure described in Sect. 3.1. Then, the perturbed solution is improved with the local search method presented in Sect. 3.2 (step 5). Finally, BVNS selects the next neighborhood to be explored (step 6). Specifically, if the search has found a new best solution, i.e., $f(\mathcal{M}'') < f(\mathcal{M})$, the method starts again from the first neighborhood ($k = 1$). Otherwise, it continues with the next neighborhood ($k = k+1$). These three steps are repeated until the maximum predefined neighborhood $k_{\max}$ is reached (steps 3–7). The VNS framework usually considers two different stopping criterion: time horizon or number of iterations. In this work the algorithm is executed a fixed number of iterations (steps 1–8).

Algorithm 1. Basic VNS($\mathcal{M}$, iterations)

1: **for** $i \in 1 \ldots$ iterations **do**
2: $k \leftarrow 1$
3: **while** $k \leq k_{\max}$ **do**
4: $\mathcal{M}' \leftarrow$ `Shake`$(\mathcal{M}, k)$
5: $\mathcal{M}'' \leftarrow$ `LocalSearch`$(\mathcal{M}')$
6: $k \leftarrow$ `NeighborhoodChange`$(\mathcal{M}, \mathcal{M}'', k)$
7: **end while**
8: **end for**

The initial solution for the BVNS algorithm can be generated either with a specific constructive procedure or with a random method. We have decided to use a random procedure since we do not have *a priori* information about the problem that can be useful during the constructive phase. Therefore, for the initial solution we select a subset of macroeconomic variables (each one is selected with a probability of 0.5). Then, we select a random bias in the range $[-5, 5]$ and, for each variable, random values for its α and β parameters. The resulting solution (improved with the local search method) is used as the initial solution for the BVNS algorithm.

The representation of the solution selected splits it into two parts: the variables selected and the associated weights. Regarding the BVNS algorithm, each of the main procedures are devoted to modify a different part of the solution. In particular, the shake procedure focuses in the selected variables, while the local search method is intended to improve the values of their ϵ, α and β parameters.

3.1 Shake

In the context of VNS, the shake procedure consists of perturbing the incumbent solution in the corresponding neighborhood in order to diversify the search. Regarding the energy demand estimation problem, the shake procedure is intended to modify the macroeconomic variables selected in the solution. Given a solution $\mathcal{M}$, X represents the set of selected macroeconomic variables ($X \subseteq X^\star$). The current neighborhood k represents the perturbation size.

The shake procedure selects k variables from $X^\star$ at random. Then, for each selected macroeconomic variable X_i, the method includes it in the solution ($X \leftarrow X \cup \{X_i\}$) if it is not yet included ($X_i \notin X$). Otherwise ($X_i \in X$) the shake procedure removes that variable from the solution ($X \leftarrow X \setminus \{X_i\}$).

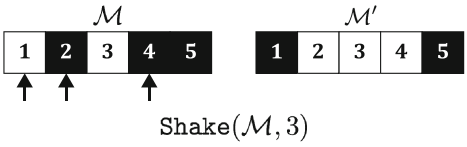

Fig. 1. Example of performing the shake procedure in a solution $\mathcal{M}$ with $k = 3$.

Figure 1 represents the shake procedure executed over a solution $\mathcal{M}$ where the macroeconomic variables selected are represented in with solid black color. Before executing the shake procedure, $\mathcal{M}$ includes the variables 2, 4, and 5. Since the neighborhood in the example is 3, the method selects three variables at random (1, 2 and 4, highlighted with an arrow). The first variable selected, is included in the new solution $\mathcal{M}'$, since it did not originally belong to $\mathcal{M}$. Variables 2 and 4 are removed from the new solution $\mathcal{M}'$ since they were already included in $\mathcal{M}$. Therefore, the resulting solution after executing the shake procedure only contains variables 1 and 5.

3.2 Local Search

Unlike the shake procedure, the local search procedure proposed in this paper is intended to improve a solution by modifying the value of the bias and the weights associated to each macroeconomic variable. Since these values belong to the real domain (i.e., $\{\epsilon, \alpha_i, \beta_i\} \in \mathbb{R}, \forall i \in 1 \dots |X|$), the local search considers global optimization techniques. One of the most extended strategy for improving solutions in global optimization is the Line Search method, which has lead to several successful works in the recent years [2].

The Line Search method iterates over every parameter of the solution (including ϵ, α and β), trying to improve the quality of the solution by modifying one parameter at a time.

The improvement of a parameter $w \in \{\epsilon, \alpha_1 \dots \alpha_{|X|}, \beta_1, \dots, \beta_{|X|}\}$ whose available range is $[MIN, MAX]$ starts by performing a discretization of the search space in h different points. Then, the method iterates over each generated point, assigning its value to the current weight, and evaluating the resulting solution. If the new solution outperforms the previous best one, it is updated, and the search starts again.

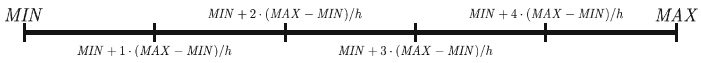

Fig. 2. Discretization of the search space with $h = 5$

Figure 2 shows the discretization of the search space for a parameter w in the range $[MIN, MAX]$. Since w belongs to the real domain, the search would need to explore infinite points in order to perform an exhaustive search for the best value. Let h be a parameter of the Line Search method. The algorithm generates h different values uniformly distributed through the variable domain. Therefore, in order to generate h discrete values, the gap between each pair of consecutive values is evaluated as $(MAX - MIN)/h$.

The search then traverses the set of values generated following a random order. If the method finds an improvement in a given value, then it is updated in the solution and the search starts again (first improvement strategy). Once all values generated have been evaluated without finding an improvement (local optimum), the search continues with the next parameters of the solution. It is worth mentioning that the next parameter to be explored is selected at random. The Line Search method ends when no improvement is found in any parameter, returning the best solution found during the search.

4 Computational Experiments

The problem addressed in this paper is completely focused in solving an actual problem. Therefore, the data used in the experiments needs to be real, in order

to analyze the applicability of the proposed algorithm to real-world data, which often include high level of noise. In particular, we have used the same data used in a previous work [13], related with the energy demand estimation in Spain. Specifically, the data contains information about several macroeconomic variables that were analyzed from 1980 to 2011.

For each year, the dataset contains information about 14 different macroeconomic variables, namely: Gross Domestic Product (GDP), population, export, import, energy production (kilotonne of oil equivalent, kTOE), electricity power transport (kW/h), electricity production (kW/h), GDP per unit of energy use, energy imports net (% of use), fossil fuel consumption (kW/h), electric power consumption (kW/h), CO_2 total emissions (Mtons), unemployment rate, and diesel consumption in road (kTOE).

The data have been split in two different sets in order to avoid overtraining. The training set contains data for 15 years, while the test set contains data for the remaining 16 years. Both sets have been generated at random, but considering the years related with the economic crisis (2010 and 2011) in the test set, with the aim of analyzing if the algorithm is able to predict these behaviors. It is worth mentioning that we have considered the same error measure, MSE, for the preliminary experimentation, in order to have comparable results.

Both, VNS (feature selection) and ELM (training) are executed only once. After that, a decision maker will only need to feed the trained algorithm in order to immediately obtain the energy demand prediction associated to the input data used, without requiring computing time. The computing time for training is 10 s per VNS iteration and 5 s for ELM.

The main objective of the first experiment is to select the best parameters for the BVNS algorithm proposed. In particular, it requires two different parameters: the number of iterations and the maximum neighborhood number to explore ($k_{\max}$). Notice that the maximum neighborhood to be explored represents the maximum number of variations in the macroeconomic variables that will be performed. On the one hand, it is easy to see that both, quality and computing time, increase with the number of iterations, so we have selected 25 in order to maintain reasonable computing times. On the other hand, we have tested the following values for $k_{\max} = \{2, 4, 6, 10\}$. Table 1 shows the results obtained for each variant in terms of the mean squared error obtained in the test set when considering the parameters selected by the algorithm in the training set. It is worth mentioning that the h parameter for the Line Search method has been experimentally set to 1000 during the whole experimentation.

Notice that those macroeconomic variables selected in all variants are highlighted in bold, since they seem to be relevant features. In particular, the common macroeconomic variables are import, energy imports net, and CO_2 total emissions. From these results we can extract that $k_{\max} = 4$ presents the best results in terms of MSE, emerging as the best BVNS variant. It can be seen that the MSE increases with the size of the maximum neighborhood. However, the worst results are obtained when considering $k_{\max} = 2$, which may reveal that it is a

small maximum neighborhood to be used in this kind of problems. Therefore, BVNS with $k_{\max} = 4$ is selected for the remaining experiments.

The next experiment is intended to compare the results obtained by our proposal when compared with the best previous works. Table 2 shows the results obtained by our BVNS compared with a Harmony Search algorithm isolated (HS) and combined (HS+ELM) with a neural network [13], and, finally, a Genetic Algorithm (GA) [1]. The bad performance of Genetic Algorithm can be partially motivated by its design, in which the algorithm is only able to consider 4 joint macroeconomic variables. The remaining algorithms are more versatile, since they consider a variable number of macroeconomic variables, increasing the explored region of the search space. We can also see an improvement of 0.20 % in HS+ELM with respect to HLM isolated, which reveals the relevance of the neural network in the algorithm. Finally, our BVNS emerges as the best algorithm, improving the combination of HS with ELM.

The final experiment is intended to test how the best results obtained by BVNS fit with the real energy demand values for each year available in the dataset. Figure 3 presents these results in a line graph where the x-axis represents the year evaluated and the y-axis indicates the total energy demand, measured in kilotonne of oil equivalent (kTOE). The line with triangle markers for each year represents the real measure, while the one with rectangle markers represents the estimation given by BVNS.

Table 1. Performance of BVNS with the proposed values of $k_{\max}$. Common features selected in all cases are highlighted in bold font.

$k_{\max}$	MSE (%)	Selected variables
2	3.40	{**4**, 7, **9**, **11**, 13, 14}
4	2.10	{3, **4**, 8, **9**, **11**, 13}
6	2.59	{1, 3, **4**, 5, 8, **9**, **11**}
10	2.86	{2, 3, **4**, 6, 7, **9**, **11**, 12}

Table 2. Results obtained by BVNS when compared with the best previous algorithms: Harmony Search (HS), Harmony Search with neural network (HS+ELM), and Genetic Algorithm (GA).

Algorithm	MSE (%)	Selected variables
BVNS	2.10	{3, 4, 8, 9, 11, 13}
HS+ELM	2.16	{1, 2, 3, 7, 8, 9, 12}
HS	2.36	{1, 2, 3, 7, 8, 9, 12}
GA	2.89	{1, 2, 3, 4}

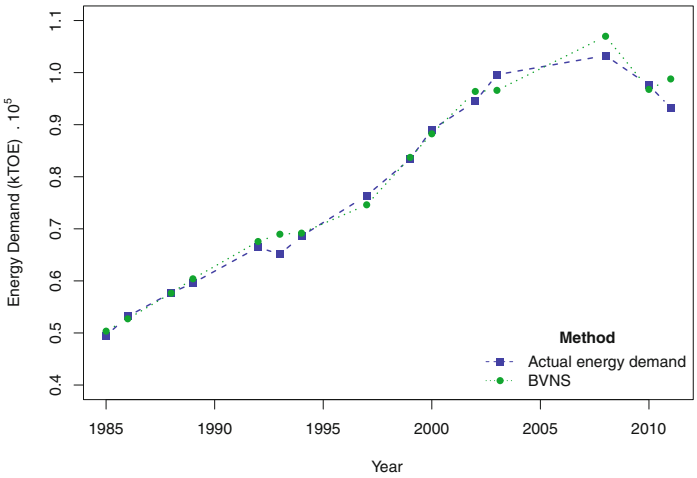

Fig. 3. Comparison of the BVNS prediction when compared with the real values

5 Conclusions

This paper has introduced a new metaheuristic algorithm, based on the Variable Neighborhood Search framework, for estimating the energy demand of a country. The experiments has been performed using a dataset which contains macroeconomic information about Spain for the last years. The thoroughly experimentation presented has been used for selecting the best parameters for the algorithms, as well as for comparing the best results obtained with respect to the actual energy demand of the years in which the data is available. The obtained results have shown the superiority of our proposal, reducing the error obtained when compared to the best previous approach, based on a Harmony Search method. We do believe that this algorithm can become a precise tool for helping decision makers.

Acknowledgments. This work has been partially supported by the projects TIN2014-54583-C2-2-R and TIN2015-65460-C2-2-P of the Spanish Ministerial Commission of Science and Technology (MICYT), and by the Comunidad Autónoma de Madrid, under project numbers S2013ICE-2933_02 and S2013ICE-2894.

References

1. Ceylan, H., Ozturk, H.K.: Estimating energy demand of Turkey based on economic indicators using genetic algorithm approach. Energy Convers. Manag. **45**(1516), 2525–2537 (2004)
2. Duarte, A., Martí, R., Glover, F., Gortázar, F.: Hybrid scatter tabu search for unconstrained global optimization. Ann. Oper. Res. **183**(1), 95–123 (2009)

3. Duarte, A., Pantrigo, J.J., Pardo, E.G., Sánchez-Oro, J.: Parallel variable neighbourhood search strategies for the cutwidth minimization problem. IMA J. Manag. Math. **27**(1), 55–73 (2016)
4. Ediger, V.S., Akar, S.: ARIMA forecasting of primary energy demand by fuel in Turkey. Energy Policy **35**(3), 1701–1708 (2007)
5. Geem, Z.W., Roper, W.E.: Energy demand estimation of South Korea using artificial neural network. Energy Policy **37**(10), 4049–4054 (2009)
6. Hansen, P., Mladenović, N.: Variable neighborhood search: principles and applications. Eur. J. Oper. Res. **130**(3), 449–467 (2001)
7. Kankal, M., Akpinar, A., Kömürcü, M.I., Özsahin, T.S.: Modeling and forecasting of Turkeys energy consumption using socio-economic and demographic variables. Appl. Energy **88**(5), 1927–1939 (2011)
8. Kiran, M.S., Özceylan, E., Gündüz, M., Paksoy, T.: A novel hybrid approach based on particle swarm optimization and ant colony algorithm to forecast energy demand of Turkey. Energy Convers. Manag. **53**(1), 75–83 (2012)
9. Kiran, M.S., Özceylan, E., Gündüz, M., Paksoy, T.: Swarm intelligence approaches to estimate electricity energy demand in Turkey. Knowl.-Based Syst. **36**, 93–103 (2012)
10. Sánchez-Oro, J., Sevaux, M., Rossi, A., Martí, R., Duarte, A.: Solving dynamic memory allocation problems in embedded systems with parallel variable neighborhood search strategies. Electron. Notes Discrete Math. **47**, 85–92 (2015)
11. Sánchez-Oro, J., Pantrigo, J.J., Duarte, A.: Combining intensification and diversification strategies in VNS. An application to the Vertex Separation problem. Comput. Oper. Res. **52**(Part B), 209–219 (2014)
12. Piltan, M., Shiri, H., Ghaderi, S.F.: Energy demand forecasting in Iranian metal industry using linear and nonlinear models based on evolutionary algorithms. Energy Convers. Manag. **58**, 1–9 (2012)
13. Salcedo-Sanz, S., Muñoz-Bulnes, J., Portilla-Figueras, J., Del Ser, J.: One-year-ahead energy demand estimation from macroeconomic variables using computational intelligence algorithms. Energy Convers. Manag. **99**, 62–71 (2015)
14. Toksari, M.D.: Estimating the net electricity energy generation and demand using the ant colony optimization approach: case of Turkey. Energy Policy **37**(3), 1181–1187 (2009)
15. Toksari, M.D.: Ant colony optimization approach to estimate energy demand of Turkey. Energy Policy **35**(8), 3984–3990 (2007)
16. Ünler, A.: Improvement of energy demand forecasts using swarm intelligence: the case of Turkey with projections to 2025. Energy Policy **36**(6), 1937–1944 (2008)
17. Yu, S., Wei, Y.M., Wang, K.: A PSOGA optimal model to estimate primary energy demand of China. Energy Policy **42**, 329–340 (2012)
18. Yu, S., Zhu, K.: A hybrid procedure for energy demand forecasting in China. Energy **37**(1), 396–404 (2012)
19. Yu, S., Zhu, K., Zhang, X.: Energy demand projection of China using a path-coefficient analysis and PSOGA approach. Energy Convers. Manag. **53**(1), 142–153 (2012)

Theorem 1. *The following properties hold for the family* $\{\models_a : a \in [0,1]\}$ *of graded entailment relations on* $Fm_3 \times Fm_3$ *induced by a* $*$*-similarity relation* S *on* Ω*:*

(i) Nestedness: if $\varphi \models_a \psi$ *and* $b \leq a$*, then* $\varphi \models_b \psi$
(ii) $\models_1$ *coincides with* $\models$*, while* $\models \subsetneq \models_a$ *if* $a < 1$*. Moreover, if* $\psi \not\models \bot$*, then* $\varphi \models_0 \psi$ *for any* φ*.*
(iii) Positive-preservation: $\varphi \models_a \psi$ *iff* $\varphi^+ \models_a \psi^+$
(iv) $*$*-Transitivity: if* $\varphi \models_a \psi$ *and* $\psi \models_b \chi$ *then* $\varphi \models_{a*b} \chi$
(v) Left-OR: $\varphi \vee \psi \models_a \chi$ *iff* $\varphi \models_a \chi$ *and* $\psi \models_a \chi$
(vi) Restricted Right-OR: for all $w \in \Omega$*,* $\overline{w} \models_a \varphi \vee \psi$ *iff* $\overline{w} \models_a \varphi$ *or* $\overline{w} \models_a \psi$
(vii) Restricted symmetry: for all $w, w' \in \Omega$*,* $\overline{w} \models_a \overline{w'}$ *iff* $\overline{w'} \models_a \overline{w}$
(viii) Consistency preservation: if $\varphi \not\models \bot$ *then* $\varphi \models_a \bot$ *only if* $a = 0$
(ix) Continuity from below: If $\varphi \models_a \psi$ *for all* $a < b$*, then* $\varphi \models_b \psi$

Conversely, for any family of graded entailment relations $\{\vdash_a : a \in [0,1]\}$ *on* $Fm_3 \times Fm_3$ *satisfying the above properties, there exists a* $*$*-similarity relation* S *such that* $\vdash_a = \models_a$ *for each* $a \in [0,1]$*.*

Proof (Sketch). The proof follows the same steps than the one of [4, Theorem 1] in the case of a classical propositional setting. The key points to take into account here are:

– it is easy to check that, for any formula $\varphi \in Fm_3$, φ^+ is logically equivalent in Ł$_3$ to the disjunction

$$\bigvee_{w \in \Omega : w(\varphi) = 1} \overline{w}.$$

– $(\varphi \vee \psi)^+$ is logically equivalent to $\varphi^+ \vee \psi^+$.
– for every $w, w' \in \Omega$, $\overline{w} \models_a \overline{w'}$ iff $S(w, w') \geq a$.

For the converse direction, the latter property is used to define the corresponding similarity S for a family of consequence relations $\{\vdash_a : a \in [0,1]\}$ satisfying (i)–(ix) as $S(w, w') = \sup\{a \in [0,1] \mid \overline{w} \vdash_a \overline{w'}\}$. □

Taking into account Lemma 4, a sort of dual characterization for $\models^C_a$, that we omit, can easily be derived from the above one for $\models_a$. On the other hand, the above properties also indirectly characterize $\models^{\leq}_a$ in the sense that, in our finite setting, $\models_a$ (and thus $\models^C_a$ as well) can be derived from $\models^{\leq}_a$ as well as the following lemma shows.

Lemma 5. *For any* $\varphi, \psi \in Fm_3$*, we have that* $\varphi \models_a \psi$ *iff for every* $w \in \Omega$ *such that* $w(\varphi) = 1$ *there exists* $w' \in \Omega$ *such that* $w(\psi) = 1$ *and* $\overline{w} \models^{\leq}_a \overline{w'}$*.*

Proof. It directly follows from properties (iv) and (v) of Theorem 1, by checking that, for every $w \in \Omega$, $\overline{w} \models^{\leq}_a \overline{w'}$ iff $\overline{w} \models_a \overline{w'}$.

However, admittedly, the resulting characterization of $\models^{\leq}_a$ we would obtain using this lemma is not very elegant.

4 A Logic to Reason About Graded Consequences $\models_a$, $\models_a^C$ and $\models_a^{\leq}$

In this section we will define a Boolean (meta) logic LAC3 to reason about the graded entailments $\models_a$, $\models_a^C$ and $\models_a^{\leq}$. The idea is to consider expressions corresponding to $\varphi \models_a \psi$, $\varphi \models_a^C \psi$ and $\varphi \models_a^{\leq} \psi$ as the concerned objects of our logic, and then to use Theorem 1 to devise a complete axiomatics to capture the intended meaning of such expressions.

To avoid unnecessary complications, we will make the following assumption: all $*$-similarity relations S will take values in a finite set G of $[0, 1]$, containing 0 and 1, and $*$ will be a given *finite* t-norm operation on G, that is, $(G, *)$ will be a finite totally ordered semi-group. In this way, we keep our language finitary and avoid the use of an infinitary inference rule to cope with Property (ix) of Theorem 1.

Our logic will be a two-tired logic, where at a first level we will have formulas and semantics of the 3-valued Łukasiewicz logic Ł$_3$ and at the second level we will have propositional classical logic CPC.

We start by defining the syntax of LAC3, with two languages:

- Language $\mathcal{L}_0$: built from a finite set of propositional variables $Var = \{p, q, r, \ldots\}$ and using Ł$_3$ connectives $\neg, \wedge, \vee, \rightarrow$. Other derived connectives are $\oplus$ and $\otimes$, defined as in Sect. 2. We will use $\top$ and $\bot$ as abbreviations for $p \rightarrow p$ and $\neg(p \rightarrow p)$ respectively, and φ^+ and φ^- as abbreviations of $\varphi \otimes \varphi$ and $(\neg\varphi)^+$ respectively.
- Language $\mathcal{L}_1$: atomic formulas of $\mathcal{L}_1$ are only of the form $\phi \succ_a^P \psi$, where ϕ, ψ are $\mathcal{L}_0$-formulas and $a \in G$, and compound $\mathcal{L}_1$-formulas are built from atomic ones with the usual Boolean connectives $\neg, \wedge, \vee, \rightarrow$.[2]
  Moreover, we will be using $\phi \succ_a^C \psi$ and $\phi \succ_a \psi$ as abbreviations of $\neg\psi \succ_a^P \neg\phi$ and $(\phi \succ_a^P \psi) \wedge (\phi \succ_a^C \psi)$ respectively.

The semantics is given by similarity Kripke models $M = (W, S, e)$ where W is a finite set of worlds, $S : W \times W \rightarrow G$ is a $*$-similarity relation, and $e : W \times Var \mapsto \{0, \frac{1}{2}, 1\}$ is a 3-valued evaluation of propositional variables in every world, which is extended to arbitrary $\mathcal{L}_0$-formulas using Ł$_3$ truth-functions. For every formula $\varphi \in \mathcal{L}_0$, we define: $[\varphi]_M : W \rightarrow \{0, 1/2, 1\}$ such that $w \mapsto e(w, \varphi)$, $[\varphi^+]_M = \{w \in W \mid e(w, \varphi) = 1\}$, and $[\varphi^-]_M = \{w \in W \mid e(w, \varphi) = 0\}$.

Each similarity Kripke model $M = (W, S, e)$ induces a function $e_M : \mathcal{L}_1 \rightarrow \{0, 1\}$, which is a (Boolean) truth evaluation for $\mathcal{L}_1$-formulas defined as follows:

- for atomic $\mathcal{L}_1$-formulas:
  $e_M(\phi \succ_a^P \psi) = 1$ if $[\phi^+]_M \subseteq ([\psi^+]_M)^a$, i.e., if $\min_{w \in [\phi^+]_M} \max_{w' \in [\psi^+]_M} S(w, w') \geq a$;
  $e_M(\phi \succ_a^P \psi) = 0$ otherwise.
- for compound formulas, use the usual Boolean truth functions.

[2] Although we are using symbols $\wedge, \vee, \neg, \rightarrow$ for both formulas of $\mathcal{L}_0$ and $\mathcal{L}_1$, it will be clear from the context when they refer to Ł$_3$ or when they refer to Boolean connectives.

Note that, by definition, $e_M(\phi \succ_a^C \psi) = 1$ iff $e_M(\neg\phi \succ_a^P \neg\psi) = 1$, and $e_M(\phi \succ_a \psi) = 1$ iff $e_M(\phi \succ_a^P \psi) = 1$ and $e_M(\phi \succ_a^C \psi) = 1$.

In the next lemma we list some useful properties of e_M.

Lemma 6. *The following conditions hold:*

- $e_M(\phi \succ_a^C \psi) = 1$ *iff* $[\psi^-]_M \subseteq ([\phi^-]_M)^a$
- $e_M(\phi \succ_a \psi) = 1$ *iff* $[\phi^+]_M \subseteq ([\psi^+]_M)^a$ *and* $[\psi^-]_M \subseteq ([\phi^-]_M)^a$
- $e_M(\phi \succ_1 \psi) = 1$ *iff* $[\phi]_M \leq [\psi]_M$
- $e_M((\phi \succ_1 \psi) \wedge (\psi \succ_1 \phi)) = 1$ *iff* $[\varphi]_M = [\psi]_M$, *iff* $[\phi \leftrightarrow \psi] = W$.

Now we define the notion of logical consequence in LAC3 for $\mathcal{L}_1$-formulas.

Definition 3. *Let $T \cup \{\Phi\}$ be a set of $\mathcal{L}_1$-formulas. We say that Φ logically follows from T, written $T \models_{LAC3} \Phi$, if for every similarity Kripke model $M = (W, S, e)$, if $e_M(\Psi) = 1$ for every $\Psi \in T$, then $e_M(\Phi) = 1$ as well.*

Finally we propose the following axiomatization of LAC3.

Definition 4. *The following are the axioms for LAC3:*

(A1) *Axioms of CPC for $\mathcal{L}_1$-formulas*
(A2) $\phi \succ_1^P \psi$, *where* ϕ, ψ *are such that* $\phi \models \psi$
(A3) $\neg(\top \succ_1^P \rightarrow \bot)$
(A4) $(\phi \succ_a^P \psi) \rightarrow (\phi \succ_b^P \psi)$, *where* $a \leq b$
(A5) $(\phi \succ_1^P \psi) \rightarrow (\phi^+ \wedge \neg\psi^+ \succ_1^P \bot)$
(A6) $\neg(\psi \succ_1^P \bot) \rightarrow (\phi \succ_0^P \psi)$
(A7) $(\phi \succ_a^P \bot) \rightarrow (\phi \succ_1^P \bot)$
(A8) $\neg(\overline{w} \succ_1^P \bot) \wedge (\overline{w} \succ_a^P \overline{w'}) \rightarrow (\overline{w'} \succ_a^P \overline{w})$, *for* $w, w' \in \Omega$
(A9) $(\phi \succ_a^P \chi) \wedge (\psi \succ_a^P \chi) \rightarrow (\phi \vee \psi \succ_a^P \chi)$
(A10) $(\overline{w} \succ_a^P \phi \vee \psi) \rightarrow (\overline{w} \succ_a^P \phi) \vee (\overline{w} \succ_c^P \psi)$
(A11) $(\phi \succ_a^P \psi) \wedge (\psi \succ_b^P \chi) \rightarrow (\phi \succ_{a*b}^P \chi)$
(A12) $(\phi \succ_a^P \psi) \leftrightarrow (\phi^+ \succ_a^P \psi^+)$

The only rule of LAC3 is modus ponens. The notion of proof defined from the above axioms and rule will be denoted $\vdash_{LAC3}$.

Finally, we have the following soundness and completeness theorem for LAC3.

Theorem 2. *For any set $T \cup \{\Phi\}$ of $\mathcal{L}_1$-formulas, it holds that $T \models_{LAC3} \Phi$ if, and only if, $T \vdash_{LAC3} \Phi$.*

Proof. One direction is soundness, and it basically follows from Theorem 1. As for the converse direction, assume $T \nvdash_{LAC3} \Phi$. The idea is to consider the graded expressions $\phi \succ_a^P \psi$ as propositional (Boolean) variables that are ruled by the axioms together with the laws of classical propositional logic CPC. Let Γ be the set of all possible instantiations of axioms (A1)–(A12). Then it implies that Φ does not follow from $T \cup \Gamma$ using CPC reasoning, i.e. $T \cup \Gamma \nvdash_{CPC} \Phi$. By completeness of CPC, there exists a Boolean interpretation v such that $v(\Psi) = 1$

for all $\Psi \in T \cup \Gamma$ and $v(\Phi) = 0$. Now we will build a $*$-similarity Kripke model M such that $e_M(\Psi) = 1$ for all $\Psi \in T$ and $e_M(\Phi) = 0$. To do that we take Ω and define $S : \Omega \times \Omega \to G$ by

$$S'(w, w') = \max\{a \in G \mid v(\overline{w} \succ_a^P \overline{w'}) = 1\}.$$

By axioms (A2), (A8) and (A11), S is a $*$-similarity. Note that, by definition and Axiom (A4), $S(w, w') \geq a$ iff $v(\overline{w} \succ_a^P \overline{w'}) = 1$. Finally we consider the model $M = (\Omega, S, e)$, where for each $w \in \Omega$ and $p \in Var$, $e(w, p) = w(p)$. What remains is to check that $e_M(\Psi) = v(\Psi)$ for every LAC3-formula Ψ. It suffices to show that, for every $\phi, \psi \in \mathcal{L}_0$ and $a \in G$, we have $e_M(\phi \succ_a^P \psi) = v(\phi \succ_a^P \psi)$, that is, to prove that

$$v(\phi \succ_a^P \psi) = 1 \quad \text{iff} \quad \min_{w \in [\phi^+]_M} \max_{w' \in [\psi^+]_M} S(w, w') \geq a.$$

First of all, recall that for every ϕ, Ł$_3$ proves the equivalence $\phi^+ \leftrightarrow \vee_{w \in \Omega : w(\varphi) = 1} \overline{w}$, and by axioms (A12), (A9) and (A10), we have that LAC3 proves

$$\phi \succ_a^P \psi \leftrightarrow \bigwedge_{w \in \Omega : w(\phi) = 1} \bigvee_{w' \in \Omega : w'(\psi) = 1} \overline{w} \succ_a^P \overline{w'}.$$

Therefore, $v(\phi \succ_a^P \psi) = 1$ iff for all w in Ω such that $w(\phi) = 1$, there exists w' such that $w'(\psi) = 1$ and $v(\overline{w} \succ_a^P \overline{w'}) = 1$. But, as we have previously observed, $v(\overline{w} \succ_a^P \overline{w'}) = 1$ holds iff $S(w, w') \geq a$. In other words, we actually have $v(\phi \succ_a^P \psi) = 1$ iff $\min_{w \in [\phi^+]_M} \max_{w' \in [\psi^+]_M} S(w, w') \geq a$. This concludes the proof. □

5 Conclusions and Future Work

We have presented an approach towards considering graded entailments between vague concepts (or propositions) based on the similarity between both the prototypes and counterexamples of the antecedent and the consequent. This approach is a natural generalization of the Łukasiewicz's three-valued consequence ($\models^{\leq}$) that preserves truth-degrees. The provided axiomatization is for the operators $\succ_a^P$, which are based on prototypes only, while the operators $\succ_a$, based on both prototypes and counterexamples, can be naturally obtained as a derived operators in the system. To derive a complete axiomatic system directly for the operators $\succ_a$ is an issue under current investigation. Besides, we leave other interesting issues for further research. First, in this paper, we have assumed $app(\omega, \alpha)$ to be a three-valued concept, and to define $\models_a^{\leq}$ from $\models_a$ and $\models_a^C$ we have used a *conjunctive* aggregation of the two aspects of similarity, similarity among prototypes and similarity among counterexamples. Another approach could be to let $app(\omega, \alpha)$ to admit itself a finer distinction by defining $app^*(\omega, \alpha) = S(w, [\alpha]^+) \odot (1 - S(w, [\alpha^-]))$ with $S(w, [\alpha^+]) = \max_{\omega' \in [\alpha^+]} S(\omega, \omega')$ and analogusly for S$(w, [\alpha^-])$. Then the extent to which α entails β can be defined based on the relationship of $app^*(\omega, \alpha)$ and $app^*(\omega, \beta)$ considering all possible situations ω. This direction

seems to have lots of challenges as $\odot$ might not be as simple as a conjunctive operation; also different notions of consequence can be worth exploring in the line of [5,6,8].

Acknowledgments. The authors are thankful to the anonymous reviewers for their helpful comments. Esteva and Godo acknowledge partial support of the project TIN2015-71799-C2-1-P (MINECO/FEDER).

References

1. Bou, F., Esteva, F., Font, J.M., Gil, A., Godo, L., Torrens, A., Verdú, V.: Logics preserving degrees of truth from varieties of residuated lattices. J. Logic Comput. **19**(6), 1031–1069 (2009)
2. Cignoli, R., D'Ottaviano, I.M.L., Mundici, D.: Algebraic Foundations of Many-Valued Reasoning. Trends in Logic, vol. 7. Kluwer, Dordrecht (1999)
3. Ciucci, D., Dubois, D., Lawry, J.: Borderline vs. unknown: comparing three-valued representations of imperfect information. Int. J. Approx. Reason. **55**(9), 1866–1889 (2014)
4. Dubois, D., Prade, H., Esteva, F., Garcia, P., Godo, L.: A logical approach to interpolation based on similarity relations. Int. J. Approx. Reason. **17**, 1–36 (1997)
5. Dutta, S., Bedregal, B.R.C., Chakraborty, M.K.: Some instances of graded consequence in the context of interval-valued semantics. In: Banerjee, M., Krishna, S.N. (eds.) ICLA. LNCS, vol. 8923, pp. 74–87. Springer, Heidelberg (2015)
6. Esteva, F., Godo, L., Rodríguez, R.O., Vetterlein, T.: Logics for approximate and strong entailments. Fuzzy Sets Syst. **197**, 59–70 (2012)
7. Ruspini, E.H.: On the semantics of fuzzy logic. Int. J. Approx. Reason. **5**, 45–88 (1991)
8. Vetterlein, T.: Logic of prototypes and counterexamples: possibilities and limits. In: Alonso, J.M., et al. (eds.) Proceedings of the IFSA-EUSFLAT 2015, pp. 697–704. Atlantis Press (2015)
9. Vetterlein, T.: Logic of approximate entailment in quasimetric spaces. Int. J. Approx. Reason. **64**, 39–53 (2015)
10. Vetterlein, T., Esteva, F., Godo, L.: Logics for approximate entailment in ordered universes of discourse. Int. J. Approx. Reason. **71**, 50–63 (2016)