Modelling perception and attitudes towards renewable energy technologies

Fernando Ribeiro

University of Minho, Centre for Industrial and Technology Management,

Campus Azurem, 4800-058 Guimaraes PORTUGAL

fernandor@dps.uminho.pt

Paula Ferreira (CORRESPONDING AUTHOR)

University of Minho, Centre for Industrial and Technology Management,

Campus Azurem, 4800-058 Guimaraes PORTUGAL

tel: +351253510760

paulaf@dps.uminho.pt

Madalena Araújo

University of Minho, Centre for Industrial and Technology Management,

Campus Azurem, 4800-058 Guimaraes PORTUGAL

mmaraujo@dps.uminho.pt

Ana Cristina Braga

University of Minho, Algoritmi Research Centre,

Campus Azurem, 4800-058 Guimaraes PORTUGAL

mmaraujo@dps.uminho.pt

Modelling perception and attitudes towards renewable energy technologies

Abstract

While renewable energy technologies (RET) increase their share in power generation systems worldwide, some questions remain open, namely those concerning the opinion of the populations on new projects of these technologies. Given the long period of planning and large capital sums required by RET and, in some cases, the fact of being subsidized, it is desirable for decision-makers to acknowledge the public opinion and at least perceive if the opinions are rooted on biased perceptions. In this paper we propose a methodology for public perception and awareness assessment, involving an initial phase of data collection by means of a survey, followed by a phase of regression models construction resulting in predictive models of expected perceptions and attitudes towards RET. The models were translated in a free and easy to use computational Excel application and its usefulness was demonstrated for the case of four electricity RET in Portugal: hydro, wind, biomass and solar.

Keywords

Renewable energy technologies; public opinion; ordered logistic regression; binary logistic regression; excel simulation tool

1 – Introduction

Renewable energy technologies (RET) are increasing their importance worldwide. This is especially true within the European Union, where institutional strategies like the EUSDS (European Union Sustainable Development Strategy) will monitor the next decades' development, based on economic, ecologic and social criteria. Concrete objectives like the ones established under the 2020 European climate & energy package envisage a rise in renewable energy consumption during this decade, which imposes the question of public acceptance of RET. The European public opinion has been generally supportive of renewable energy (Eurobarometer, 2012), but the possibility to please all the population has to be discarded, given not only the number of citizens but, more importantly, the unequal distribution of impacts generated by the proximity to the RET infrastructures (Ribeiro et al., 2013). Given the disperse character of some RET, visual and noise amenities affect mostly residents of rural areas, and this might induce a negative attitude due to local proximity. It is important for decision-makers to acknowledge the public opinion, because projects facing resistance may see their completion delayed (Cavallaro and Ciraolo, 2005).

It must be acknowledge that acceptance studies should go beyond the evaluation of overall public opinion recognizing the importance of the proximity effect and the perception towards benefits and costs that may explain public attitudes. Bertsch et al (2016) highlighted that transition towards RES-based energy systems is largely perceived positively in general but locally can be confronted with a lack of public acceptance. The authors conducted a nationally representative survey for Germany and concluded on the importance of local acceptance related to landscape modification and demonstrated also the importance of age and education in relation to acceptance. Bertsch et al (2017) implemented a survey in Ireland and concluded that in general people feel positively disposed towards RET but found also reluctance amongst people to have these technologies located close to their places of residence. Both these studies and Ribeiro et al (2013) clearly show the importance of local perception and of the assessment of the socio-demographic variables that can rule the local and national opposition.

In this paper, we propose a methodology to contribute to predict the public opinion over RET, supported on a survey for data collection complemented with statistical models. The methodology implementation is demonstrated for the Portuguese case, resourcing to the results of a survey implemented in Portugal and addressing hydro, wind, biomass and solar power previously detailed in Ribeiro et al. (2014). The Portuguese case is particularly interesting as the energy generated from RET has been increasing over the last years and remains a key objective for the European Commission energy policy (European Commission, 2014). In 2015 RET contributed for the generation of 47% of the total electricity demand in Portugal, which was 49 TWh that year. It is worth mentioning that 2015 was a dry year, meaning that rainfall values were well below the annual average and consequently reduced considerably RET share. In fact, in 2014, which was a wet year (rainfall above the average), the RET share reached 62% (REN, 2015).

We have created a visual and easy-to-use interface, linked to statistical models, which allows simulating the answer of a certain respondent (of a certain age, gender and educational degree) about a given technology. The NIMBY (Not In My BackYard) effect is also assessed, along with willingness to pay more for the technology, the perceptions of how it contributes for sustainable development, and also the probability of that respondent not acknowledging the technology. In this paper we use the term "NIMBY" as an attitude of being generally supportive of a technology but at the same time showing a negative attitude if it is implemented near one's residence (Jones and Eiser, 2009).

The aforementioned statistical models are generated resorting to regression methods, which are employed when the objective is to describe the relationship between a response variable and one or more explanatory variables (Hosmer and Lemeshow, 2000). In the present study we will characterize public opinion concerning renewable energy technologies, recurring to surveys further presented in section 2. As such, the outcomes will use ordered categories (ordered logistics regression) such as "totally agree",

"agree", "neither agree nor disagree", "disagree" and "totally agree" and binary categories (only two possible outcomes) such as "yes" and "no".

Different methods have been used in the literature to evaluate determinants of renewables acceptance and related topics frequently supported on statistical tools. Meta-analysis regression was used to integrate literature results and provide a quantitative assessment to estimate for example willingness to pay for RET and explain its heterogeneity (Ma et al, 2015; Bigerna and Polinori, 2015). Surveys were conducted at regional, local and national scale and the results are frequently analyzed by???? statistical tests (Bertsch et al, 2016; Karytsas and Theodoropoulou, 2014 and Ribeiro et al, 2013) and regression models with particular emphasis on logistic regression as it allows to predict a response or explain it according, for instance, to the socio-economic and geographic characteristics of the respondents described by nominal, ordinal and interval scales.

Logistic regression (discrete outcome variable) has been employed in many fields, ranging from biomedical research, business and finance, criminology, ecology, engineering, health policy, to linguistics, among others (Hosmer and Lemeshow, 2000, page ix).

In the past, ordinal logistic regression models (discrete outcome variable, with more than two possible values) were used to analyze household electricity consumption classes in Brazil (Fuks and Salazar, 2008), in Sweden to assess the importance of environmental attitudes in households' energy savings (Martinsson et al., 2011), on public opinion on natural gas drilling on two different counties in the USA (Kriesky et al., 2013). Binary logistic models were used to study factors that affect consumer acceptance of electrical vehicles in China (Zhang et al, 2011), in Greece to assess the opinion on different energy issues (Nikolau et al., 2012). In Greece, a study using binary logistic regression models shows that middle aged males are more likely to be willing to pay for a stay in a hotel which uses renewable energy (Kostakis and Sardianou, 2012). More recently, Bertsch et al (2017) analyzed how people's views of energy-related technologies are explained by socio-demographic characteristics, national energy policy preferences and technology-specific factors using also ordinal logistic regression models.

The contribution of this paper is then twofold: firstly a methodology supported on surveys and statistical models based on regression methods is proposed for RET public perception and awareness assessment; secondly the translation of these models in an easy-to-use interface was demonstrated for the case of Portugal and allowing to relate perception and attitudes with socio-economic characteristics of the population. We particularly seek to contribute to demonstrate the implementation potential and usefulness of these models to support energy decision making in the future. Whilst the application here is in Portugal, the proposed methodology is highly transferable to other contexts and in particular to countries with high reliance on RET for electricity generation.

The remainder of the paper is as follows: in section 2, we summarize the survey implementation and main results, in section 3 we introduce the methodology used for ordered logistic regression and binary logistic regression. Section 4 contains the obtained models along with the created Excel interface for simulating responses, section 5 presents the discussion and validation of the results, and section 6 draws conclusions and points directions for future work.

2 – Survey to assess public opinion

The survey aimed at studying the differences of public opinion towards the four technologies (hydro, wind, biomass and solar) between regions where RET plants are already operating and regions where RET plants are absent. Therefore, four different surveys exist, each to be applied in two samples consisting of distinct regions, totaling eight cases. The surveys were conducted by phone during May and June of 2012.

Three thousand and forty seven (3047) results were collected, which represented about 380 results for each case, ensuring a 95% confidence degree with a 5% margin of error, as detailed in Ribeiro et al (2014).

Each survey was divided in six sections and the respondent was firstly introduced to the technology to be addressed. The first section acted as a filter, and the questionnaire would count as valid for the respondents that passed on this filter question. The second section is about acceptance of the technology in the country, in the municipality, or near the respondent's residence. For the sake of this analysis, the municipality level encompassed a large urban administrative division and surrounding rural territory and small communities such as smaller towns and villages (in Portuguese "concelho"). For the high proximity effect, the analysis concerned the parish (in Portuguese "freguesia"), which is the smallest administrative subdivision of municipality. The third section evaluates the perception of economic impact of the given technology, while the forth and the fifth sections evaluate the environmental and social impacts. Finally, socio-demographic information such as educational level and age, besides gender, are collected. SPSS software was used for the statistical analysis of the results and modelling. The full questionnaire is available on Ribeiro et al (2013).

Table 1 presents the possible answers and how they were coded in SPSS. When asking the respondent, the "no answer" option was excluded, to force the respondent to another answer, however, if upon insistence no answer was given, a "no answer" was accepted. The "no answer" was coded as zeros in SPSS in order to assign each and all of them as missing values and avoid counting them in means and other indicators retrieved in statistical tests.

The main results of the study indicate that the Portuguese are well aware of the technologies assessed in the study, being hydro power the most acknowledged one. Also, the respondents are mostly in favor of new projects for all the four technologies and this is particularly evident for wind power plants. The case with least support technology is hydro power but even so gathering 77% of positive attitudes towards it. As for the NIMBY effect, this does not seem to be a major issue among Portuguese population. Solar and wind power are less prone to NIMBYism, but in the municipalities with biomass power plants evidence of some NIMBY attitude was found. It was found however that extreme NIMBYism in the biomass case increases with age and is higher among people with lower educational levels. Solar power is perceived as the technology contributing more for sustainable development, including cost, environmental impacts and contribution to social development perception. Only a small fraction of respondents perceive the renewable technologies as contributing to increase the electricity bill. Additional information on the results of the survey can be found in Ribeiro et al. (2014), including the statistical tests and graphical representation of the results.

Table 1: Variables encoded in SPSS.

Variable name	Type	Values	Note
Technology	Nominal	{1="Hydro", 2="Wind", 3="Biomass", 4="Solar"}	Information supplied by the survey implementer
Municipality_has_technology	Nominal	{0="no", 1="yes"}	Information supplied by the survey implementer
Accept_country	Ordinal	{0="no answer", 1="totally disagree", 2="tend to disagree", 3="tend to agree", 4="totally agree"}	Respondents acceptance towards RET in the country
Accept_municipality	Ordinal	{0="no answer", 1="totally disagree", 2="tend to disagree", 3="tend to agree", 4="totally agree"}	Respondents acceptance towards RET in the municipality
Accept_parish	Ordinal	{0="no answer", 1="totally disagree", 2="tend to disagree", 3="tend to agree", 4="totally agree"}	Respondents acceptance towards RET in the parish
NIMBY	Interval		Computed as the difference between Accept_country and Accept_parish
Perception_economy	Ordinal	{0="no answer", 1="greatly reduces bill", 2="slightly reduces bill", 3="does not alter bill", 4="slightly increases bill", 5="greatly increases bill"}	Respondents perception towards RET impact on the electricity bill
Perception_environment	Ordinal	{0="no answer", 1="greatly protects the environment", 2="slightly protects the environment", 3="no impact", 4="slightly endangers environment", 5="greatly endangers environment"}	Respondents perception towards RET impact on the environment
Perception_social	Ordinal	{0="no answer", 1=" greatly develops local populations", 2=" slightly develops local populations", 3="no impact", 4="slightly harms local populations", 5="greatly harms local populations"}	Respondents perception towards RET impact on the local population development
WTP (Willingness-to-Pay)	Nominal	{0="not WTP more", 1="WTP more"}	Equals 1 in the case that "perception_economy" is equal to 4 or 5, AND "accept_country" is equal to 3 or 4. Equals 0 in other cases.
Education	Ordinal	{0="no answer", 1="primary school", 2="4th grade", 3="9th grade", 4="12th grade", 5="university degree"}	Academic level of the respondents
Age	Interval		Age of the respondents
Gender	Nominal	{1="female", 2="male"}	Gender of the respondents

3 – Methodology

3.1 – Methodology for ordinal logistic regression

Having in mind the objectives of the present study, we propose a methodology consisting of four main phases, presented in Figure 1. The ordinal logistic regression models, or simply "ordinal models", were used to predict answers in five cases: economic impact, environmental impact, social impact, acceptance of the technology in the country and NIMBYism. The methodology follows Garson (2012) approach.

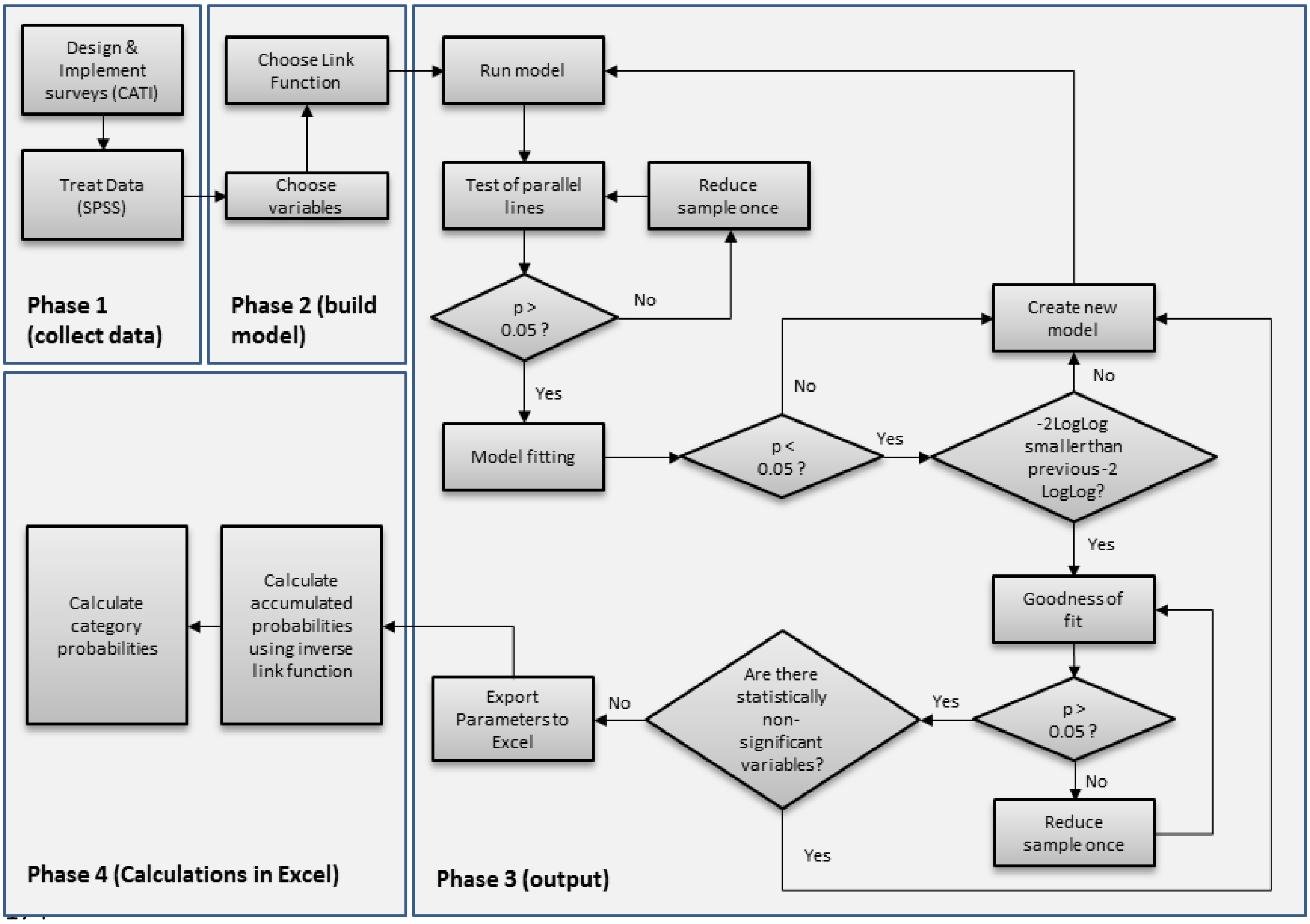

Figure 1 – Methodology for building ordinal logistic regression models.

The first block ("Phase 1") consists of data collection. It begins with designing the questionnaires to implement, along with the choice for collecting the answers. For the present study we contacted a company specialized in computer assisted telephone interviewing (CATI) and they performed 3047 structured interviews. Then, it became necessary to organize the data in order to use statistical software to build the models. Organizing the data involved coding variables, eliminating errors and coding the missing values to avoid their use in the models, among other tasks. We opted for the software IBM® SPSS 21®.

The Phase 2 is about building the model. Firstly it is necessary to determine the dependent variable (i.e. the variable to predict). As already mentioned, five variables are predicted: economic impact, environmental impact, social impact, acceptance of the technology in the country and NIMBYism. The first three variables are predicted using the list of independent variables "technology", "municipality has technology", "age", "gender" and "educational level". The attitude towards new power plants in the country and the NIMBYism used the same variables plus the perceived economic, environmental and social impacts.

The continuous variable "age" was inserted as covariate, and the others, nominal and ordinal variables, were inserted as "factors". The options were kept as default, with the exception of "output" and "link function". It is necessary to ensure that SPSS performs the Test of Parallel Lines, to be analyzed later in the third phase (output). The link function depends on the distribution of the dependent variable. "Logit" functions were considered for economic, environmental and social impact, given that they follow approximately a normal distribution. "Complementary log-log" functions were used for predicting "acceptance" and "NIMBYism", because these variables follow a distribution where the higher categories ("agreement" and "positive NIMBYism" respectively) are more frequent (Garson, 2012: 12). Besides looking at the distribution of the dependent variable, the best model will present a lower -2LogLog value in the output "model fitting". We tested different functions and confirmed the function corresponding to the lowest -2LogLog for every case.

The output of the model is interpreted in the third block. The first output is the "Test of the parallel lines", also called "proportionality of odds", and should not be statistically significant ($p > 0.05$). If the test is statistically significant it doesn't mean the model is impossible to use, due to a large sample size, because even small differences in slopes will be found significant (Garson, 2012: 15). The test is considered very conservative, and for particularly large samples it nearly always results in rejection, according to Allison (1999) and Clogg and Shihadeh (1994). As a result, every time the parallel lines test was significant, we re-ran the model after programming SPSS to choose a random sample of 5% (152 cases) out of the original 3040. If the test was significant once again, it would be recommended to perform multinomial regression. However, re-running the model with a smaller random sample always resulted in a non-significant test of parallel lines.

After the test of parallel lines, the Model Fitting table must be analyzed. Values to be retained in this phase are the "-2 Log Likelihood (final)" and the result of the significance test. Basically, at this stage, SPSS tests whether the generated model predicts the dependent variable significantly better than a null (intercept-only) model. If this is the case, the significance test indicates that $p < 0.05$. None of the created models had any problems in this test. A new model would have to be created if this test was non-significant. It is necessary to keep the value of "-2 Log Likelihood (final)", because if new models are created, they can be compared under this value, following the rule that the better model is the one with lower "-2 Log Likelihood (final)", as stated above.

The next table to evaluate is the goodness of fit, where a well-fitting model is non-significant on the Pearson and Deviance tests. For large samples, the results are significant for even small differences or when there are continuous independent variables Garson (2012: 16) as "age" in our case. Rerunning for a random 5% sample (of 152 cases), no test is significant anymore for any of the models.

Finally, SPSS gives as an output the Parameters Estimates. It is necessary to check whether the variables are considered statistically significant. To the continuous variable "age", only one parameter estimate is calculated. If p is lower than 0.05, then the variable "age" should enter the model. For the nominal or ordinal variables, one parameter estimate is calculated for each category. If any of those parameters is significant ($p < 0.05$), then the variable should enter the model. If, on the other hand, one variable has no significant parameter estimates, the model should be rebuilt and re-run. These parameters are aggregated in the array presented as β in the "Phase 4".The table also calculates parameter estimates for every category of the dependent variable, which will be indicated in "Phase 4" as α_k. The model is ready to be used when all the variables possess statistically significant parameters estimates.

The fourth block ("Phase 4") aims at calculating the probabilities of answers in categories. This calculation is performed in hidden Excel spreadsheets, and the final information is presented in the interface for the user. The calculation happens in two steps: firstly the accumulated probability, then the categorical probability.

For achieving the results for the accumulated probability it is necessary to perform the calculations according to the link function that was used when creating the model. The goal is to calculate, for example, how would a resident in a municipality without biomass, 42 year old and female, with education level corresponding to 12 years secondary school level react to a new biomass power plant in the country. The answer would be, for example, 38% probabilities that the respondent will “totally disagree” or “slightly disagree”. This probability is $P(Y_j \leq k | \mathbf{X})$, and it is calculated using Equation (1), where k is the class of the dependent variable to predict, $\mathbf{X}$ is the array of the independent variables values (respondent’s characteristics, technology to assess, among others; see Table 2 for each model specification), α_k and β_j are the parameter estimates calculated in Phase 3 (Marôco, 2011: 762).

$$Link\left\{(Y_j \leq k \mid \boldsymbol{X})\right\} = \alpha_k - \mathbf{X}^{*}\beta_j \qquad (1)$$

As already stated above, in our case we used two different link functions, logit and complementary log-log. Equation 2 presents the logit function, which after some arrangement results in Equation 3, which in turn allows calculation of accumulated probabilities for the category k.

$$Logit\left\{(Y_j \leq k \mid \boldsymbol{X})\right\} = \ln\left(\frac{P(Y_j \leq k \mid \mathbf{X})}{1 - P(Y_j \leq k \mid \mathbf{X})}\right) = \alpha_k - \mathbf{X}^{*}\beta_j \qquad (2)$$

$$P\{Y \leq k\} = \frac{1}{1 + e^{-(\alpha_k - X^{*}\beta_j)}} \qquad (3)$$

Equation 4 presents the complementary log-log function, which can be transformed in Equation 5 and allows calculation of accumulated probabilities for the category k.

$$Cloglog\left\{(Y_j \leq k \mid \boldsymbol{X})\right\} = \ln\left(-\ln\left(1 - P\left[Y_j \leq k \mid \mathbf{X}\right]\right)\right) = \alpha_k - \mathbf{X}^{*}\beta_j \qquad (4)$$

$$P\{Y \leq k\} = 1 - e^{-e^{(\alpha_k - X^{*}\beta_j)}} \qquad (5)$$

Obviously, the last category, K, has an accumulated probability of 100% to happen, since it encloses all the possible categories. To calculate the probability of each category to occur, it is then necessary to use Equations 6, 7 and 8. For $Y_j = 1$, the probability is the accumulated probability itself, since it only includes one category. For the intermediate categories achieved by subtracting the accumulated probability of k and $k-1$, and for the last category, K, it is necessary to subtract 1 and the accumulated probability of $K-1$.

$$P\left\{Y_j = 1\right\} = Link\left(\alpha_1 - x_j\beta\right) \qquad (6)$$

$$P\left\{Y_j = k\right\} = Link\left(\alpha_k - x_j\beta\right) - Link\left(\alpha_{k-1} - x_j\beta\right) \qquad (7)$$

$$P\left\{Y_j = K\right\} = 1 - Link\left(\alpha_{k-1} - x_j\beta\right) \qquad (8)$$

These numbers are then integrated into dynamic plots, which are presented to the user. Details of the interface, along with print screens are presented further in Section 4.

3.2 – Methodology for binary logistic regression

Binary logistic regression was used to build two models: one to predict whether the respondent is aware of the technology or not, and the other to predict whether the respondent is willing to pay more for it. In comparison with the ordered logistic regression, the process for binary logistic regression in SPSS is much simpler, mainly because the program employs iterative methods when building the model. This means that SPSS automatically removes the non-significant variables and creates a new model, contrarily to what happened in ordered logistic regression, and also because there are no such tests as the test of parallel lines which could invalidate the model. Figure 2 describes the methodology used for the binary logistic regression models.

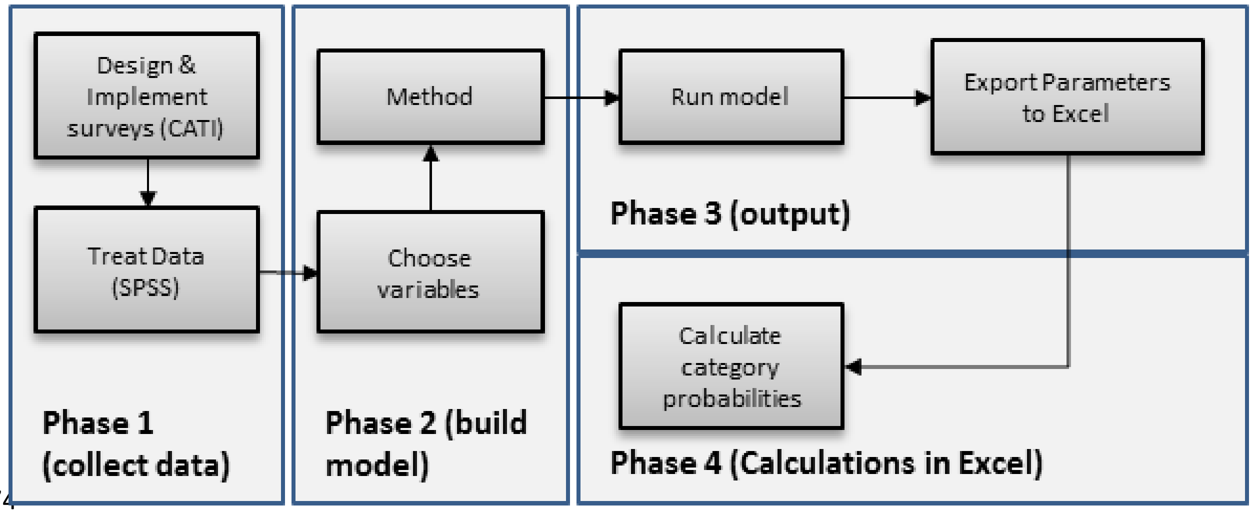

Figure 2 – Methodology for binary logistic regression models.

Phase 1 follows the same process as for ordered logistic regression.

Phase 2, where the model is built, deals with choice of "selection variable" (the dependent variable, the one which we want to predict), and the covariates (independent variables). Covariates are "technology", "municipality_has_technology", "age", "gender" and "education". It is then necessary to define which are categorical, among these, i.e. all excepting "age". It is asked to define the reference category, and it was decided to choose the first category as reference. This influences the parameter estimates presented further in next section, although it is not perceived by the user.

It is then necessary to choose the stepwise method. Among the possibilities, for both cases we chose Forward:LR. Basically the model is built from scratch in the first iteration, and in every following iteration one new independent variable is added. "LR" refers to likelihood ratio, a model fit calculation, which is compared in each iteration, allowing to conclude if the inclusion of the iteration's variable increases the model fit. According to Hosmer and Lemeshow (2000), research has shown that this method presents the best statistical properties. For other options, we used the SPSS default: probability for stepwise entry was 5%, and for removal was 20%, classification cutoff 0.5 and maximum iterations were 20.

In Phase 3 the model is ran and parameters exported to excel. These parameters are shown in the next section.

The fourth phase concerns the probability calculation. The calculation of the probability is relatively straightforward. Taking into account the table with parameter estimates β for the independent variables calculated by SPSS and presented in Table 6 of the next section, to calculate the probability of the independent variable Y_j assuming the value "yes" (for example, "respondent acknowledges technology"), coded in SPSS with the value "1", the probability is calculated in two steps, as follows:

$$a = \sum(\alpha + \beta_1 + \beta_2 + \cdots + \beta_i) \quad (9)$$

$$P\{Y_j = 1\} = \frac{e^a}{1+e^a} \quad (10)$$

where α is a constant parameter and β_k is the parameter which corresponds to the i^{th} independent variable.

For calculating the independent variable Y_j assuming the value "0", i.e. "the respondent does not acknowledge the technology", the probability is the complementary of the previous one.

$$P\{Y_j = 0\} = 1 - P\{Y_j = 1\} \quad (11)$$

4 – Logistic regression models for predicting public opinion

In this section we present the models obtained from SPSS. They allow obtaining the responses (dependent variables) predicted by given respondent's characteristics (independent variables) as explained in the previous sections.

Table 2 – Summary for ordinal logistic regression models tests and variables included.

Dependent variable	Independent variables	Link function	Test of parallel lines Sig.	Model fitting Sig.	Goodness of fit Pearson sig.	Goodness of fit Deviance sig.	Statistically non-significant variables
Perception of economic impact	Technology, Municipality has technology, age, gender, education	Logit	~0.000* / 0.212**	~0.000*	0.022* / 0.348**	1	-
Perception of environmental impact	Technology, education	Logit	~0.000* / 0.250**	~0.000*	0.000* / 0.739**	0.000* / 0.999**	Municipality has technology, age, gender
Perception of social impact	Technology, education	Logit	~0.000* / 0.232**	~0.000*	0.000* / 0.960**	0.000* / 0.878**	Municipality has technology, age, gender
Acceptance	Technology, education, age, perception_eco, perception_env, perception_soc	Complementary Log-log	~0.000* / 0.689**	~0.000*	1*	1*	Municipality has technology, gender
NIMBY	Technology, municipality has technology, age, education, perception_env	Complementary Log-log	~0.000* / 0.271**	~0.000*	0* / 0.603**	1*	Perception_eco, Perception_soc, gender

Values with * were obtained using the entire sample, while values with ** were obtained for a sample of 5% (see Section 3.1 for more details).

Taking into account the procedure described in the previous section it was found that the estimated models are well fitting.

Table 3 – Summary for binary logistic regression models and independent variables included.

Dependent variable	Independent variable	Stepwise method	Statistically non-significant variables
Acknowledges_technology	technology, municipality_has_technology, age, gender, education	Forward:LR	-
WTP	technology, municipality_has_technology, gender, education	Forward:LR	age

The fit of binary logistic regression models using the stepwise selection methodology, revealed that only age variable is non-significant in the case of WTP.

Table 4 –Parameter estimates for the perception of economic, environmental and social impact models, using ordinal logistic regression.

Parameter estimates						
Perception of economic impact	α_1= -1.911	β_{age}= 0.009	$\beta_{tech.=1}$= 1.112	$\beta_{mun.\_has\_tech.=0}$= -0.218	$\beta_{educ.=1}$= 0.495	$\beta_{gen.=1}$= -0.187
	α_2= 0.625		$\beta_{tech.=2}$= 0.662	$\beta_{mun.\_has\_tech.=1}$= 0	$\beta_{educ.=2}$= 0.292	$\beta_{gen.=2}$= 0
	α_3= 1.913		$\beta_{tech.=3}$= 0.144		$\beta_{educ.=3}$= 0.040	
	α_4= 3.278		$\beta_{tech.=4}$= 0		$\beta_{educ.=4}$= -0.009	
					$\beta_{educ.=5}$= 0	
Perception of environmental impact	α_1= -2.513	β_{age}= 0	$\beta_{tech.=1}$= 1.094	$\beta_{mun.\_has\_tech.=0}$= 0	$\beta_{educ.=1}$= 0.495	$\beta_{gen.=1}$= 0
	α_2= -1.128		$\beta_{tech.=2}$= 0.284	$\beta_{mun.\_has\_tech.=1}$= 0	$\beta_{educ.=2}$= 0.292	$\beta_{gen.=2}$= 0
	α_3= 0.559		$\beta_{tech.=3}$= 0.680		$\beta_{educ.=3}$= 0.0404	
	α_4= 2.628		$\beta_{tech.=4}$= 0		$\beta_{educ.=4}$= -0.009	
					$\beta_{educ.=5}$= 0	
Perception of social impact	α_1= -1.836	β_{age}= 0	$\beta_{tech.=1}$= 0.195	$\beta_{mun.\_has\_tech.=0}$= 0	$\beta_{educ.=1}$= 0.489	$\beta_{gen.=1}$= 0
	α_2= 0.502		$\beta_{tech.=2}$= 0.284	$\beta_{mun.\_has\_tech.=1}$= 0	$\beta_{educ.=2}$= 0.071	$\beta_{gen.=2}$= 0
	α_3= 2.201		$\beta_{tech.=3}$= 0.488		$\beta_{educ.=3}$= -0.058	
	α_4= 3.641		$\beta_{tech.=4}$= 0		$\beta_{educ.=4}$= -0.017	
					$\beta_{educ.=5}$= 0	

α give the estimated log-odds of intercept for the reference group
β are the ordered log-odds (logit) regression coefficients. Standard interpretation of the ordered logit coefficient is that for a one unit increase in the predictor, the response variable level is expected to change by its respective regression coefficient in the ordered log-odds scale while the other variables in the model are held constant.

Just as Likert scale have 5 points, there are four logit equations to predict the log-odds of

- Code 2 vs code 1
- Code 3 vs code 1
- Code 4 vs code 1
- Code 5 vs code 1

So, α gives the estimated log-odds of intercept for the reference group, i.e, when Technology = “solar”, Education="university degree", sex = “male”, municipality has technology= “yes”. For example, considering the perception of economic impact the estimated log-odds of code 2 versus code 1 in this group is −1.911; the estimated log-odds of code 3 versus code 1 is 0.625; and so on.

Considering a significance level of 5%, Table 4 shows the estimating coefficients in each model considered. The negative coefficients reveals that the lower value of independent variable are assign to higher ratings in dependent variable. For example, for the perception of economic impact, women (code 1) are less likely to assign higher ratings than men, populations are more likely to assign higher ratings to hydro (code 1), wind (code 2) or biomass (code 3) technology than to solar technology (code 4), people whose municipality do not have technology are less likely to assign higher ratings than the others, people with less education (less than 9th grade) are more likely to assign higher ratings than people with university degree (code 5), by other hand people with 12th grade (code 4) are less likely to assign higher ratings than people with university degree (code 5), and older people are more likely to assign higher ratings than the youngers.

In what concerns the perception of environmental impact and perception of social impact, the variables, “municipality has technology”, “age” and “gender” do not appear to be related to the rating. As such, these perceptions seem to be explained mainly from the previous contact with the technologies and education.

Taking into account the estimated coefficients (β) described in Table 5, for the acceptance of new power plants in the country, hydro (code 1), wind (code 2) or biomass (code 3) technology are less likely to be assigned with higher ratings in acceptance than for solar technology (code 4), older people are more likely to assign higher ratings than the youngers. The ratings of perception of economic, environmental and social impact are directly related with the ratings of acceptance as demonstrated in the last three columns of the table. Variables, “municipality has technology” and “gender” do not appear to be related to the rating of acceptance in the country.

Table 5 – Parameter estimates for the models of acceptance and NIMBYism, using ordinal logistic regression.

	Parameter estimates								
Acceptance of new plants in the country	α_1= -1.296	β_{age}= 0.009	$\beta_{tech.=1}$= -0.629	$\beta_{mun.\_has\_tech.=0}$= 0	$\beta_{educ.=1}$ = 0.625	$\beta_{gender=1}$= 0	$\beta_{percept\_eco=1}$= 1.379	$\beta_{percept\_env=1}$= 1.300	$\beta_{percept\_soc=1}$= 1.507
	α_2= -0.257		$\beta_{tech.=2}$= -0.015	$\beta_{mun.\_has\_tech.=1}$= 0	$\beta_{educ.=2}$= 0.134	$\beta_{gender=2}$= 0	$\beta_{percept\_eco=2}$= 0.529	$\beta_{percept\_env=2}$= 0.804	$\beta_{percept\_soc=2}$= 0.998
	α_3= 1.253		$\beta_{tech.=3}$= -0.526		$\beta_{educ.=3}$= 0.073		$\beta_{percept\_eco=3}$= 0.060	$\beta_{percept\_env=3}$= 0.584	$\beta_{percept\_soc=3}$= 0.568
			$\beta_{tech.=4}$= 0		$\beta_{educ.=4}$= 0.033		$\beta_{percept\_eco=4}$= 0.034	$\beta_{percept\_env=4}$= 0.387	$\beta_{percept\_soc=4}$= 0.456
					$\beta_{educ.=5}$= 0		$\beta_{percept\_eco=5}$= 0	$\beta_{percept\_env=5}$= 0	$\beta_{percept\_soc=5}$= 0
NIMBYism	α_1= -6.899	β_{age}= 0.005	$\beta_{tech.=1}$= -0.629	$\beta_{mun.\_has\_tech.=0}$= -0.097	$\beta_{educ.=1}$= 0	$\beta_{gender=1}$= 0	$\beta_{percept\_eco=1}$= 0	$\beta_{percept\_env=1}$= -0.381	$\beta_{percept\_soc=1}$= 0
	α_2= -4.330		$\beta_{tech.=2}$= -0.015	$\beta_{mun.\_has\_tech.=1}$= 0	$\beta_{educ.=2}$= 0	$\beta_{gender=2}$= 0	$\beta_{percept\_eco=2}$= 0	$\beta_{percept\_env=2}$= -0.197	$\beta_{percept\_soc=2}$= 0
	α_3= -2.324		$\beta_{tech.=3}$= -0.526		$\beta_{educ.=3}$= 0		$\beta_{percept\_eco=3}$= 0	$\beta_{percept\_env=3}$= -0.372	$\beta_{percept\_soc=3}$= 0
	α_4= 0.463		$\beta_{tech.=4}$= 0		$\beta_{educ.=4}$= 0		$\beta_{percept\_eco=4}$= 0	$\beta_{percept\_env=4}$= -0.198	$\beta_{percept\_soc=4}$= 0
	α_5= 1.020				$\beta_{educ.=5}$= 0		$\beta_{percept\_eco=5}$= 0	$\beta_{percept\_env=5}$= 0	$\beta_{percept\_soc=5}$= 0
	α_6= 1.485								
	α_7= 0								

α give the estimated log-odds of intercept for the reference group

β are the ordered log-odds (logit) regression coefficients. Standard interpretation of the ordered logit coefficient is that for a one unit increase in the predictor, the response variable level is expected to change by its respective regression coefficient in the ordered log-odds scale while the other variables in the model are held constant.

The variable NYMBYism is coded as an interval one obtained from the difference between the variables “Accept_country” and “Accept_parish”, both of them ordinal as detailed in Ribeiro et al (2013). To allow for this calculation, it was assumed that the scale assigned to the ordinal values possess equal intervals, meaning that the distance between 1 and 2 was the same that between 3 and 4 in the scale presented in Table 1.

For the NYMBYism, the results in Table 5 reveal that hydro (code 1), wind (code 2) or biomass (code 3) technology are less likely to be assigned with higher ratings than solar technology (code 4), older people are more likely to assign higher ratings than the youngers, people whose municipality do not have RET technology are less likely to assign higher ratings than the others. The ratings of perception of environmental impact are inversely related with the ratings of NYMBYism. Variables, “perception of economic impact”, “perception of social impact” and “gender” don’t appear to be related to the rating.

Table 6 describes the parameter for the binary logistic regression models of acknowledgement and willingness to pay. The variable “WTP” is coded as binary indicating also a trend for “yes” and “no” derived from the survey results as described in Ribeiro et al (2013) and as such no evidence of the monetary value assigned to this inferred WTP can be provided as this would be out of the scope of the conducted survey. For this study, WTP represents then an index of relative preferences stated by the respondents. In general the positive estimates of coefficients indicate that an increase of one unit in independent variable, contributes more to the result =1 of dependent variable, the negative estimates indicates the opposite.

For example, for the age, β=0.009 indicates that the probability of acknowledge of technology is greater for the oldest people when compared with the younger ones. The negative estimate in technology indicates that the probability of acknowledge of technology is greater for hydro (reference group) when compared to wind (β=-0.732) or biomass (β=-2.897) or solar (β=-1.537). If the municipality has technology (β=0.708) it contributes to the probability of acknowledge of technology.

The positive estimate of education reveals that probability of acknowledge of technology increases for the most graduate levels when compared with the group with primary school. Males have higher probability of acknowledge of technology when compared with females (β=0.627).

Table 6 – Parameter estimates for the binary logistic regression models of acknowledgement and willingness to pay.

	Parameter estimates					
Acknowledges_ technology	$\alpha = 1.306$	$\beta_{age} = 0.009$	$\beta_{technology=1} = 0$	$\beta_{mun.\_has\_tech.=0} = 0$	$\beta_{education=1} = 0$	$\beta_{gender=1} = 0$
			$\beta_{technology=2} = -0.732$	$\beta_{mun.\_has\_tech.=1} = 0.708$	$\beta_{education=2} = 0.927$	$\beta_{gender=2} = 0.627$
			$\beta_{technology=3} = -2.897$		$\beta_{education=3} = 1.525$	
			$\beta_{technology=4} = -1.537$		$\beta_{education=4} = 1.766$	
					$\beta_{education=5} = 2.063$	
WTP	$\alpha = -1.089$	$\beta_{age} = 0$	$\beta_{technology=1} = 0.000$	$\beta_{mun.\_has\_tech.=0} = 0$	$\beta_{education=1} = 0$	$\beta_{gender=1} = 0$
			$\beta_{technology=2} = -0.221$	$\beta_{mun.\_has\_tech.=1} = 0.289$	$\beta_{education=2} = -0.088$	$\beta_{gender=2} = 0.229$
			$\beta_{technology=3} = -0.899$		$\beta_{education=3} = -0.428$	
			$\beta_{technology=4} = -0.415$		$\beta_{education=4} = -0.802$	
					$\beta_{education=5} = -0.604$	

α give the estimated constant parameter of logit

β are the estimated logit regression coefficients for the independent variables

For willingness to pay, the variable “age” does not appear to be related with it. The negative estimate in technology indicates that the probability for willingness to pay is greater for hydro (reference group) when compared with any other technology. The negative estimate of education reveals that probability for willingness to pay is less for the most graduate levels when compared with the group with low academic background.

4.2 – The excel tool

The main objective of the excel tool was to make an easy to use way of disseminating results and facilitate their interpretation[1]. By using the tool, the information becomes more refined than doing statements such as "acceptance increases with age, decreases with educational level and is greater among males", because it allows simulation of real cases of respondents. It is then more attractive to characterize expectations and acceptance for population with particular characteristics since each individual is answered as a specific case, instead of deriving from average conclusions, such as the statements above.

The tool is constituted by an interface with three sheets, one of them being for introduction, a second for a help file, while the other is the interface where the user introduces and retrieves data. The plots and cells change almost immediately according to the inputs of the user. Several sheets of calculations, where the model information is presented, were hidden from the user to avoid confusion in the usability of the tool.

For demonstration purposes, Figure 3 presents a print screen for a real case simulation for wind power. The case corresponds to a 58-years-old male respondent with 12th degree level of education, living in a municipality where wind power is implemented. The models predict that there is 98% of probability of acknowledging this technology. The most probable category for acceptance of new wind power plants in the country is "totally agrees" (78.2%), and there is 65.2% probability of presenting no NIMBYism. There is also a high probability for unwillingness to pay more (83%). As for the most sustainable development perceptions, a person with these characteristics is expected to believe that wind power can contribute to slightly reduce the electricity bill, that it has no environmental impacts and that it slightly develops the local population.

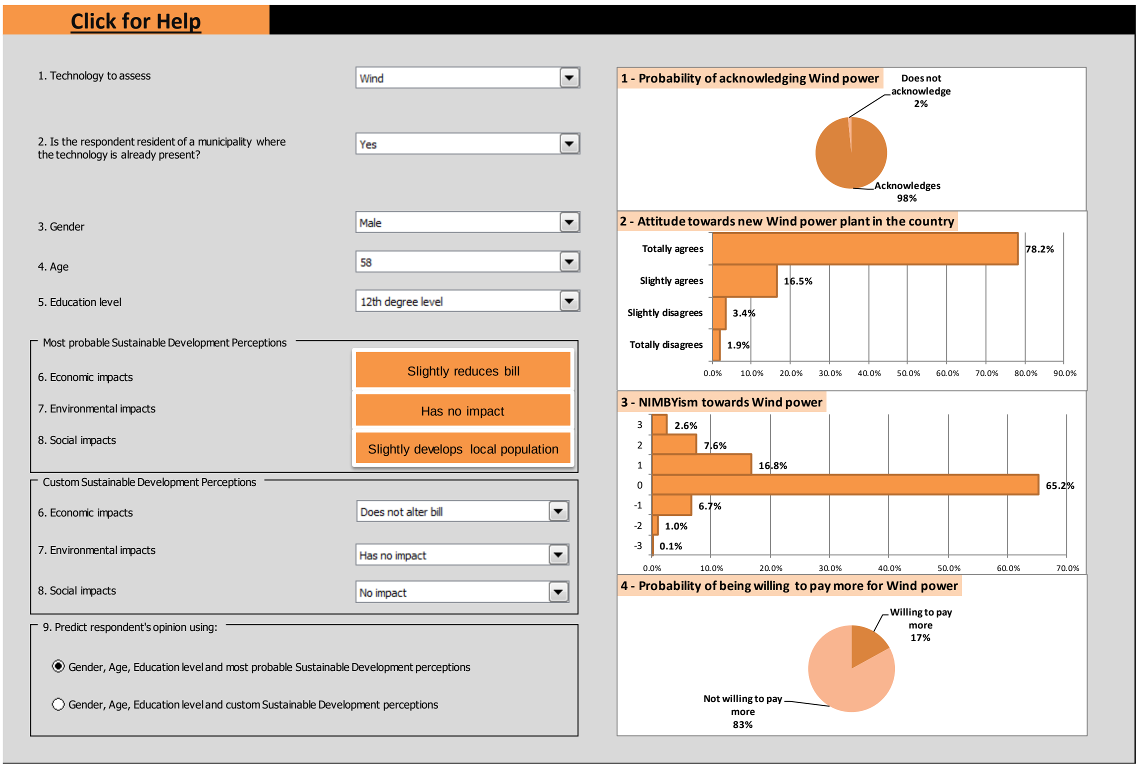

Figure 3 – Interface of the Excel tool for a real case for wind power.

[1] The tool is available online for download in http://sepp.dps.uminho.pt/results.html

On the excel tool, the required user inputs are (1) the technology, (2) whether the respondent lives in a municipality where the technology exists, (3) gender, (4) age and (5) educational level. After entering the first five inputs, the program already calculates the most probable perceived economic, environmental and social impacts and presents the graphs for probability of acknowledging the technology, acceptance of the technology, probability of NIMBYims and willingness to pay.

Additionally, if the user has already access to information about the perceived economic, environmental and social impacts of the individual, he can opt to include this as input to the model and obtain the corresponding new results on technology acknowledgment, attitudes, NIMBY and willingness to pay. As such, the optional inputs of the model are (6) perception of economic impact, (7) perception of environmental impact and (8) perception of social impact.

5 –Discussion

In order to validate the models it is necessary to realize how much they improve the capacity of prediction over proportional random classification (Marôco, 2011: 783). The calculation of proportional random classification is done by equation 12:

$$Random\ Prediction = 100 \times \left(\left(\frac{Cases_{i=1}}{Total\ cases}\right)^2 + \left(\frac{Cases_{i=2}}{Total\ cases}\right)^2 + \cdots + \left(\frac{Cases_{i=k}}{Total\ cases}\right)^2 \right) \tag{12}$$

where “total cases” are all the valid results (excluding “no answers”) concerning the variable predicted by the model and *k* is the number of categories adopted by the predicted variable.

The model correct prediction is the ratio between correct guesses made by the model and the verified answers (excluding “no answers”):

$$Model\ correct\ prediction = 100 \times \frac{correct\ guesses}{total\ answers} \tag{13}$$

Table 7 – Correct models classification: proportional classification versus ordinal regression models.

Variable predicted by the model	Proportional classification	Model correct prediction	Model improvement
Acceptance	43,80%	59,29%	15,49%
NIMBY	51,32%	71,64%	20,32%
Economic impact	27,00%	38,22%	11,22%
Environmental impact	27,90%	42,66%	14,75%
Social impact	32,11%	44,62%	12,51%

From Table 7 we can conclude that the new models perform between 10% and 20% better than the proportional classification model.

For the binary logistic regression models, the validation can be done with the aid of ROC curves. According to Hosmer and Lemeshow (2000), “the area under a ROC curve, which ranges from zero to one, provides a measure of the model’s ability to discriminate between those subjects who experience the outcome of interest versus those who do not”. As a result, models which have ROC=0.5 suggest no discrimination at all; for ROC varying between 0.7 and 0.8, Hosmer and Lemeshow (2000) consider acceptable discrimination; for ROC varying between 0.8 and 0.9 consider excellent discrimination, and above 0.9 it is outstanding discrimination (however, this last category is extremely unusual).

Using SPSS to perform the analysis of ROC curves for both “acknowledgement of technology” and “willingness-to-pay”, we obtained Figures 4 and 5, respectively. The area under the ROC curves for the acknowledgement model was 0.799 (for a 95% confidence interval, the lower limit of the area is 0.78 and the higher limit is 0.818). For the willingness-to-pay model the area is 0.635 (for a 95% confidence interval, the lower limit of the area is 0.609 and the higher limit is 0.661).These results suggest that the acknowledgement model performs acceptable to excellent discrimination. While the willingness-to-pay model does not reach the “acceptable” level, it is however statistically significantly better than a random model, given that the lower interval is higher than 0.5, which would be the area under the ROC curve for a random model.

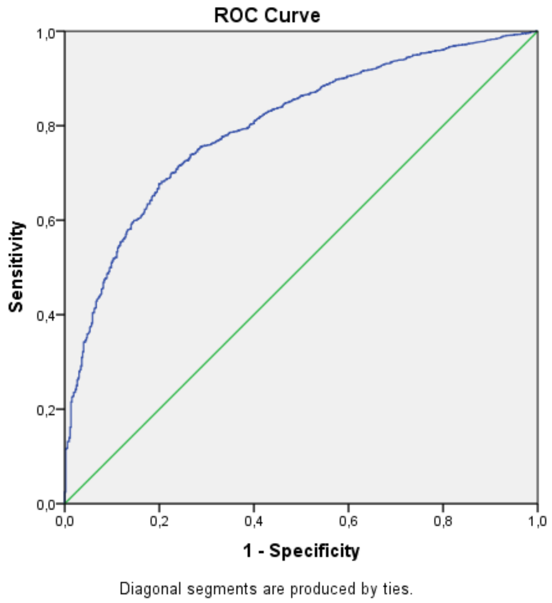

Figure 4 – ROC curve for the model “acknowledgement”.

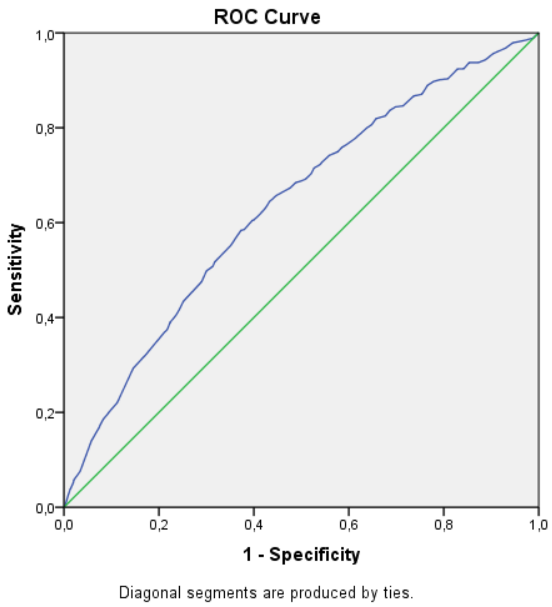

Figure 5 -ROC curve for the model “willingness-to-pay”.

6 – Conclusion

It is important for decision-makers to acknowledge public opinion towards RET, as sustainability evaluation must go beyond the economic, technological and environmental dimensions. The social assessment should include not only the evaluation of social indicators but also, the public perceptions and acceptance of population as fundamental key variables for central and local policy makers and for energy sector investors. Neglecting this social dimension can constrain the effective development of RET and threaten the concretization of energy policy objectives.

In the present paper a new methodology is proposed such that, based on respondent's gender, educational level and age and proximity to a given renewable energy technology, allows the prediction of several expected typical outcomes from one person, namely: the technology acknowledgement; he/her opinion towards new power plants and also their NIMBY effect; sustainable development perspectives (economic, environmental and social) and willingness to pay more for the technology. In a first phase, we collected more than 3000 completed and validated survey questionnaires, which were then used to generate the models for Portugal. These models were of two kinds: ordered logistic regression and binary regression. The former were used in five cases (acceptance, NIMBYism, economic, environmental and social perspectives) and the latter in two cases (acknowledgement and willingness to pay).

The proposed approach aimed to go further than a straightforward statistical analysis of the results, showing how the results of the surveys can be used for inference of acceptance towards RET. It should however be underlined that the model outputs, although being statistically valid, are prone to changes in perceptions and unexpected events that may lead to different views. As such, the model allows to assess overall trends on attitudes towards RETs and even to establish the socio-economic and geographical factors that can be determinant for these attitudes, but the interpretations' should be made with caution as acceptance, rejection and perception cannot be fully explained by quantitative basis and depend on ever changing external factors and moments. Nevertheless a better understanding of the variables affecting this outcome and their relative importance represent relevant information for investors and policy makers that can better recognize the social dimension when designing policies, incentives and promotion measures matching the public interests and concerns and as such contributing significantly for the project acceptance.

The models development implied an evaluation of the independent variables statistical significance for explaining the dependent variables. It was shown that education is particularly relevant for justifying economic, environmental and social perceptions and these ones are also significant variables for the acceptance of the technologies. On the opposite, the gender issues seem to have a minor role on the acceptance and NIMBY but impact the WTP. The results demonstrate the usefulness and quality of the models for predicting behaviors and attitudes towards renewable technologies and the main drivers of these perceptions.

It should be underlined that although the results obtained from the prediction models are specific for Portugal, the proposed models can easily be adapted to other countries or regions and should be regularly updated as perceptions and attitudes may change over time. This will require significant resources for collecting data from different countries but is deemed to be a valuable effort aimed to go beyond traditional technical evaluation of renewable energy potential and allowing to include in these studies the social acceptance and public engagement as a key aspects for the successful development of sustainable energy systems.

Further research should also address the development of new methodologies using revealed or stated preferences techniques (Menegaki, 2008) for the valuation of the WTP and to use this information to draw policy implications for instruments for environmental and energy policy. Moreover, the justification for the results obtained may go much beyond the obvious socio-economic and geographical variables and other aspects should be considered (Huijts et al, 2012), including in particular the respondents attitude towards risk that can play a major role on each respondent willingness to accept new RET projects.

References

Allison, P.D., Logistic Regression using the SAS system: Theory and application. Cary, NC: SAS Institute, 1999

Bertsch, V, Hall, M, Weinhardt, C, Fichtner, W, Public acceptance and preferences related to renewable energy and grid expansion policy: Empirical insights for Germany, Energy, Volume 114, 2016, Pages 465-477.

Bertsch, V, Hyland, M, Mahony, M, What drives people's opinions of electricity infrastructure? Empirical evidence from Ireland, Energy Policy, Volume 106, July 2017, Pages 472-497.

Cavallaro, F, Ciraolo, L, A multicriteria approach to evaluate wind energy plants on an Italian island, Energy Policy, Volume 33, Issue 2, 2005, Pages 235-244

Clogg, C, Shihadeh, E.S., Statistical models for ordinal variables. ThousandOaks, California: Sage Publications, 1994.

European Commission, 2012. Standard eurobarometer 78 / Autumn 2012.

European Commission, 2014. A policy framework for climate and energy in the period from 2020 to 2030. Brussels, 22.1.2014 COM(2014) 15 final

Fuks, M, Salazar, E, Applying models for ordinal logistic regression to the analysis of household electricity consumption classes in Rio de Janeiro, Brazil, Energy Economics, Volume 30, Issue 4, 2008, Pages 1672-1692

Garson, D, Ordinal Regression. In: Blue Book Series. Statistical Associates Publishing

Hosmer, D, Lemeshow, S, Applied Logistic Regression. In: Wiley Series in Probability and Statistics. Wiley-Interscience; 2000.

Huijts, MA, Molin, EJE, Steg, L, Psychological factors influencing sustainable energy technology acceptance: A review-based comprehensive framework, Renewable and Sustainable Energy Reviews, Volume 16, Issue 1, 2012, Pages 525-531.

Jones, C, Eiser, J, Identifying predictors of attitudes towards local onshore wind development with reference to an English case study, Energy Policy, Volume 37, Issue 11, 2009, Pages 4604-4614

Karytsas, S, Theodoropoulou, H, Socioeconomic and demographic factors that influence publics' awareness on the different forms of renewable energy sources, Renewable Energy, Volume 71, 2014, Pages 480-485.

Kostakis, I, Sardianou, E, Which factors affect the willingness of tourists to pay for renewable energy?, Renewable Energy, Volume 38, Issue 1, 2012, Pages 169-172

Kriesky, J, Goldstein,B. D., Zell, K, Beach, S, Differing opinions about natural gas drilling in two adjacent counties with different levels of drilling activity, Energy Policy, Volume 58, 2013, Pages 228-236

Marôco, J, Análise estatística com o SPSS Statistics (in Portuguese). 5th edition, Report Number, 2011.

Martinsson, J, Lundqvist, l, Sundström, A, Energy saving in Swedish households. The (relative) importance of environmental attitudes, Energy Policy, Volume 39, Issue 9, 2011, Pages 5182-5191

Menegaki, A, Valuation for renewable energy: A comparative review, Renewable and Sustainable Energy Reviews, Volume 12, Issue 9, 2008, Pages 2422-2437.

Nikolaou, I, Vitouladitis, H, Tsagarakis, K, The willingness of hoteliers to adopt proactive management practices to face energy issues, Renewable and Sustainable Energy Reviews, 16, Issue 5, June 2012

PNBEPH, 2011. Programa Nacional de Barragens com Elevado Potencial Hidroeléctrico - Memória. (in Portuguese)

REN-Redes Energéticas Nacionais, 2015. Dados Técnicos 2015/Techncial Data 2015.

Ribeiro, F, Ferreira, P, Araújo, M, Sustainability assessment of electricity production using a logic models approach, Renewable and Sustainable Energy Reviews, Volume 28, 2013, Pages 215-223

Ribeiro, F; Ferreira, P, Araújo, M, Braga, A. C., Public opinion on renewable energy technologies in Portugal, Energy, Volume 69, 2014, Pages 39-50.

Zhang, Y, Yu, Y, Zou, B, Analyzing public awareness and acceptance of alternative fuel vehicles in China: The case of EV, Energy Policy, Volume 39, Issue 11, 2011, Pages 7015-7024