

# Development of a Procedure to Analyze Customers' Choice of Renewable Energy Heating Technologies: Application in Ireland

Van-Nguyen Dinh, Biswajit Basu, and Matthew Kennedy

**Abstract**—In this paper, a procedure to analyze customers' choice of renewable energy technologies (RETs) using artificial neural networks is proposed. The relationships between input data such as the investment phases, grant amounts, date received, property age and total installation corresponding to each consumer and the target data consisting of RETs such as solar, biomass and heat pump are explored and each percentage RET choice are estimated for all phases together and for each phase by using several neural network models developed in this paper. Case studies of domestic dwelling heating in Ireland with the recently published data are analyzed. Through the proposed procedure and the case studies, the following applications are proposed (i) validation of the implemented Irish governmental Greener Homes Scheme and related customers subsidiary policies being enforced, (ii) forecasting customers choices in the future renewable energy schemes if the data on grant, time, property characteristics, expected installation and policy are given, and (iii) transferring and deployment of technologies to developing and emerging economies.

**Index Terms**—Artificial neural network, customers' choice, heating, renewable energy technology.

## I. INTRODUCTION

The global warming from the emissions of greenhouse gases is a major concern and the demand in energy is ever increasing. Australia has recently introduced a carbon pricing policy scheme [1]. In attempting to reduce carbon emissions and enhance security of supply, the Irish Government established the Greener Homes Scheme [2]. This programme commenced in 2006 with the intention to facilitate an enhanced deployment of renewable energy technology (RET) within the residential sector and so directly develop a sustainable heat market. In framing the GHS, the Government of Ireland was acutely aware of the significant levels of CO<sub>2</sub> emissions (25%) and primary energy consumption (40%) within the building stock across the European Union [3].

The key barrier to technology take-up across the GHS was identified as available finance and so the GHS key attribute involved the provision of financial support of up to 30% of the cost of the technology being developed within homes. However, in analyzing data of over 31,500 applications, a number of significant deviations in technology take-up across

the phases of the GHS were observed that extended consumer choice beyond economic factors and determinants [4], [5]. The attitudes relating to energy under variables that go beyond cost and include comfort levels, the ease of technological installation and environmental concerns were characterized [6]. In determining the key drivers for low carbon technological choice and selection and an increased penetration of low carbon solutions within residences, the consumer behavior is found to remain an increasingly important tool in assessing the viability of options [7].

Consumer choice analysis is a pattern recognition problem where the inputs including phases, grant amounts, dates received, and installation costs are associated with different classes of RETs. The four best known approaches for pattern recognition are: (i) template matching, (ii) syntactic matching, (iii) statistical classification, and (iv) artificial neural networks (ANNs) [8]. Template matching is one of the simplest and earliest approaches to pattern recognition. It is computationally demanding and its rigid template has a number of disadvantages, for instance, it would fail if the patterns are distorted, viewpoints change, or large intra-class variations among the patterns [9]. Syntactic approach adopts a hierarchical perspective where a pattern is viewed as being composed of simple sub-patterns which are themselves built from yet simpler sub-patterns. That approach may yields a combinatorial explosion of possibilities to be investigated, demanding large training sets and very large computational efforts [9]. Statistical classification methods have severe limitations when extended to spaces of many dimensions: a phenomenon as the curse of dimensionality [10].

ANNs are very good at pattern recognition problems [10]. An ANN with enough elements (called neurons) can classify any data with high accuracy. They are particularly well suited for complex decision boundary problems over many variables. The biogas generation profiles of waste water treatment plants were recently predicted by using ANN [11]. Therefore neural networks are chosen for solving the RET classification problem in this paper.

## II. ARTIFICIAL NEURAL NETWORK

Artificial neural networks are inspired by the neural architecture of the brain, as shown in Fig. 1a where there are ten billion neurons in the human brain; each connected to around 10,000 other neurons. Each neuron receives electrochemical inputs from other neurons at the dendrites. If sum of inputs is sufficiently powerful to activate the neuron, it transmits an electrochemical signal along the axon, and passes

Manuscript received April 5, 2014; revised June 30, 2014.

V.-N. Dinh and B. Basu are with Department of Civil, Structural and Environmental Engineering, School of Engineering, Trinity College Dublin, Dublin 2, Ireland (e-mail: nguyendv@tcd.ie, nguyendhhh@yahoo.com, basub@tcd.ie).

M. Kennedy is with Sustainable Energy Authority of Ireland, Wilton Park House, Dublin, Ireland (email: matthew.kennedy@seai.ie).

this signal to the other neurons whose dendrites are attached [12]. In an artificial neuron shown in Fig. 1b, a set of inputs  $i_n$  are multiplied by weights  $w_n$  and the sum of these weighted inputs is calculated. If weighted sum is greater than the neuron's threshold  $\beta$  then the neuron fires ( $O_1 = 1$ ). Otherwise the neuron does not fire ( $O_1 = 0$ ).

The development of an ANN model involves three stages, (i) training, (ii) validating and (iii) testing. The main strength of the ANN approach is its potential to deliver more accurate prediction results in comparison to other quantitative models. However, ANN does lack a theoretical background concerning the explanatory capabilities and the extraction of learning processes can be time consuming to model. The connection weights in the ANN networks cannot be interpreted or used to identify the relationships between dependent and independent variables and there are no formal techniques for non-linear methods to test the relative relevance of the independent variables and to carry out the variable selection process [13].

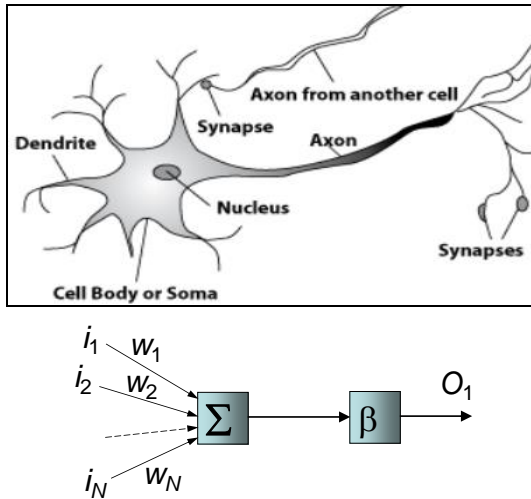


Fig. 1. Example of a neuron; (a) biological [12] and (b) artificial.

### III. IRISH GREENER HOMES SCHEME

The Greener Homes Scheme was established by the Irish Government in 2006 with the intention to facilitate an enhanced deployment of RETs within the residential sector and to directly develop a sustainable heat market. Over 31,500 technology installations under the GHS were obtained across three phases. Throughout its lifetime, the GHS resulted in the installation of over 26,500 low carbon heating technologies, with cumulative expenditure of nearly €67 million. A database was developed where data was organized into three distinct 'calendar' phases in which the GHS was delivered from March 2006 to December 2010: (i) Phase 1 (04/2006 - 09/2007), (ii) Phase 2 (09/2007 - 09/2008) and (iii) Phase 3 (09/08 - 12/10) as shown in Fig. 2. To ensure sufficient data were available on measure or technology type (including air source heat pump, ground source heat pump, vertical ground collector heat pump, well to water heat pump, biomass boiler, biomass stove, biomass stove with backboiler, wood gasification boiler, solar tube, solar flat plate and solar thermal), property type, grant type and amount and property age, only grant beneficiaries were included [5].

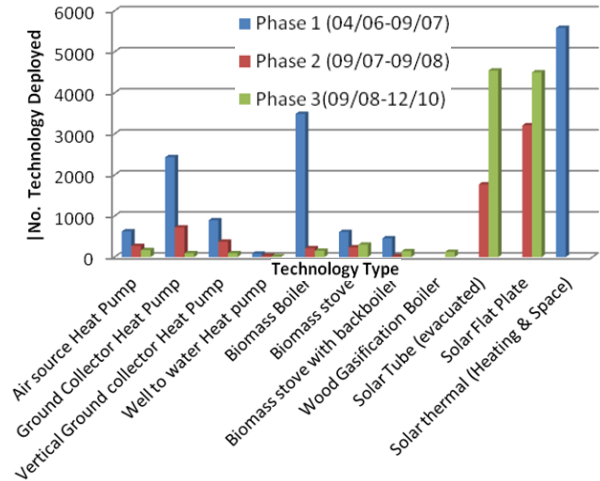


Fig. 2. RET by phase as % of total RET [5].

### IV. A PROPOSED PROCEDURE TO ANALYZE CUSTOMERS' CHOICE OF RETs BY USING ANN

#### A. Data Description

Data used in this procedure consists of several quantitative variables. The number of data lines corresponds to the number of customers. In the present case of Irish GHS data, there are five quantitative variables, namely phase, grant amount, date received, property age, and installation cost and  $N = 31,560$  data lines (i.e. consumers). Other social and non-quantitative variables such as property type, fuel displacement and natural gas availability if exists can be considered in a more comprehensive model. The data are cleaned to remove applicants that had accepted the grant offer but had not drawn down the offer within the programme timescale [5].

#### B. Data Preparation

The data are prepared into two matrices, the input matrix  $\mathbf{P}$  and the target matrix  $\mathbf{T}$ . The input matrix  $\mathbf{P}$  consists of five rows and  $N$  columns where the first, second, third, fourth and fifth rows represent the input data in 'phase', 'grant amount', 'date received', 'property age', and 'total installation cost' respectively.  $N$  is number of data lines i.e. the number of dwellings. The 'grant amount', 'property age' and 'total installation' are already in numerical values. The 'phase' and 'date received' data is converted to numerical forms. Values of 1, 2 and 3 denote numerically the GHS phases 1, 2 and 3. The date received is first converted to the absolute day. The minimum day is then found and assumed to be the first day. The target matrix  $\mathbf{T}$  consists of three rows and  $N$  columns where each row represents one of the technology types: bioenergy, solar, or geothermal. If the technology type is chosen, its value is 1, otherwise it is 0, so each row of the target matrix should contain two 0 and one 1. The outputs of the procedure are the estimated percentage of choices of the three RETs. An overall estimation of errors can also be calculated.

#### C. Data Division for ANN Schematic Model

In order to train data (i.e. find coefficients for the model) and to validate the data (i.e. prove they found coefficients accurate enough) within the model, the input data and the

target data are divided into three groups: training, validating and testing. In terms of the GHS, the data were collected at various time periods and locations, they cannot be divided into three sub-blocks for training, validating and testing as the implications of data in each sub-block may significantly differ from that of the full dataset. A continuous division of data is therefore applied.

The ratio per 10 data columns for training, validating and testing denoted as  $r_{\text{train}}$ ,  $r_{\text{vald}}$  and  $r_{\text{test}}$ , respectively are empirically specified first. In each data row of the input matrix **P** and the target matrix **T**, the data indices for training, validating and testing are continuously grouped as  $[1, \dots, r_{\text{train}}]$ ,  $[r_{\text{train}} + 1, \dots, r_{\text{vald}} + r_{\text{train}}]$ , and  $[r_{\text{vald}} + r_{\text{train}} + 1, \dots, r_{\text{vald}} + r_{\text{train}} + r_{\text{test}}]$ , respectively. That grouping scheme is continuously repeated for every  $[r_{\text{vald}} + r_{\text{train}} + r_{\text{test}}]$  range.

The ANN schematic model is developed in this paper and shown in Fig. 3. The computational units are grouped into three main layers: the first layer is the input layer, hidden layer(s) in the middle and the output layer being the last layer. Determining the optimal internal error within the model (i.e. the appropriate number of hidden nodes and layers) is challenging as action should not hinder the network's learning or its efficiency. This approach, therefore, involves significant trial and error.

It should be noted that with a sufficiently large single hidden layer, every continuous function can be represented. However, at least two layers are needed to be able to represent every discontinuous function. The number of neurons in each hidden layer is up to the designer. Except for purely linear networks, the more neurons in a hidden layer, the more powerful the network is. If a linear mapping needs to be represented linear neurons should be used. However, linear networks cannot perform any nonlinear computation. The use of a nonlinear transfer function makes a network capable of storing nonlinear relationships between input and output. A multilayer recurrent architecture using the standard back-propagation learning algorithm was applied. A back-propagation neural network together with the ANN toolbox of MATLAB [14] is used in this paper.

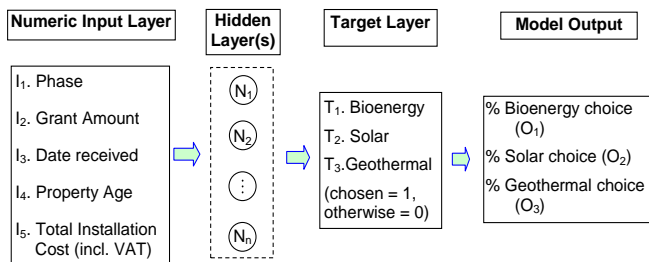


Fig. 3. Schematic ANN model for customer choice of energy technology.

#### D. Training the ANN Models

The development of ANN models involves three stages, (i) training, (ii) validating and (iii) testing. Errors were averaged across the three technologies and the model was tested with an independent set of data. The training phase involves in calculating the neural network weights (by randomly initializing connection weights), introducing a set of data inputs and actual outputs to the network, calculating the network output, and comparing the network output to the actual output and calculated error.

The fastest training function in MATLAB [14] is generally the 'trainlm' function representing Levenberg-Marquardt algorithm. This is the default training function for the feed-forward neural networks. However, that method tends to be less efficient for large networks (with thousands of weights), since it requires more memory and more computation time for these cases. Also, 'trainlm' performs better on function fitting (nonlinear regression) problems than on pattern recognition problems.

When training pattern recognition networks, the 'trainrp' function (representing resilient back-propagation algorithm) is a good choice. In defining the learning algorithm, attempting to seek predictive accuracy and minimising the network total mean squared error; the ANN adjusts the connection weights by propagating the error backward to determine how to best update the interconnection weights between individual neurons [13]. The memory requirement of the 'trainlm' is relatively small, yet it is much faster than standard gradient descent algorithms [14]. The 'trainlm' function is therefore used for training the ANN models developed in this paper.

### V. APPLICATION IN IRELAND

#### A. All-Phase Models

Several ANN models are developed across all 31,560 data lines (i.e. across all three phases of the GHS). The good models obtained are presented in Table I where the output results of RET choice percentage based on the model, overall error and numbers of iterations are displayed. There is a very minor differentiation across technology choice percentages for both solar and geothermal across the three simulations. The differential is 0.5% for solar and 0.3% for geothermal. In determining the most suitable configuration of the ANN model, the three hidden layer (7, 6 and 4 neurons) network, presented in Fig. 4, requires the least number of iterations (276). This makes it the most suitable model configuration to employ to successfully forecast consumer technology choice. The confusion matrix and performance curves of this model are shown in Fig. 5 and Fig. 6, respectively. In the confusion matrix, the diagonal cells show the number of cases that are correctly classified and the off-diagonal cells present the misclassified cases. The blue cell in the bottom right displays the total percent of correctly classified cases (95%) and the total percent of misclassified cases or total errors (5.0%). The overall percentage correct (95%) reflects the accuracy of the model in predicting consumer choice. Thus, the three hidden layer (7, 6 and 4 neurons) model is the best ANN network in this case as it can make predictions with 95% accuracy (overall error range of 5.0%) on the choice of certain RET over other alternatives (based on certain social conditions).

There are several interesting findings as follows:

- Consumer choice of RET reflects the data for solar (62%) and geothermal (19%).
- The model simulation highlights the ANN's forecast ability; in using a similar type input matrix it can accurately forecast consumer choice of RET.
- When the complexity of the model is increased, the overall error is generally reduced. The application of

such hidden layers implicitly improves the accuracy of the model.

- The number of neurons in the final hidden layer should be larger than, or equal to, that of the target layer.

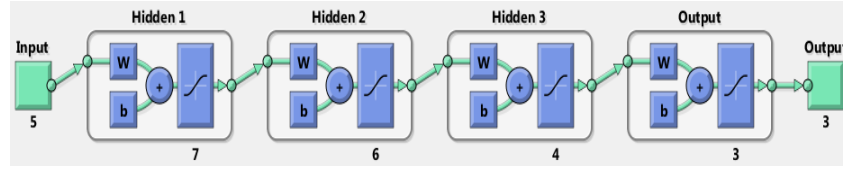


Fig. 4. Three hidden layers (7 + 6 + 4 neurons) network.

TABLE I: GOOD ANN MODELS AND PREDICTED % RET CHOICE FOR ALL PHASES

ANN Model (nr = neuron)	Bioenergy (%)	Solar (%)	Geothermal (%)	Error (%)	No. of iterations
1 hidden layer (7 neurons)	12.7	61.9	19.0	6.4	1909
2 hidden layers (7 + 4 nrs)	13.0	62.1	18.5	6.5	741
3 hidden layers (7 + 6 + 4 nrs)	14.9	61.4	18.7	5.0	276

Class: 1 = Bioenergy, 2 = Solar, 3 = Geothermal

Output Class	1	2	3	
1	1178 14.9%	12 0.2%	27 0.3%	96.8% 3.2%
2	295 3.7%	4843 61.4%	5 0.1%	94.2% 5.8%
3	43 0.5%	10 0.1%	1477 18.7%	96.5% 3.5%
	77.7% 22.3%	99.5% 0.5%	97.9% 2.1%	95.0% 5.0%
	1	2	3	
	Target Class			

Fig. 5. Confusion matrix of the 3-hidden layer (7 + 6 + 4 neurons) model.

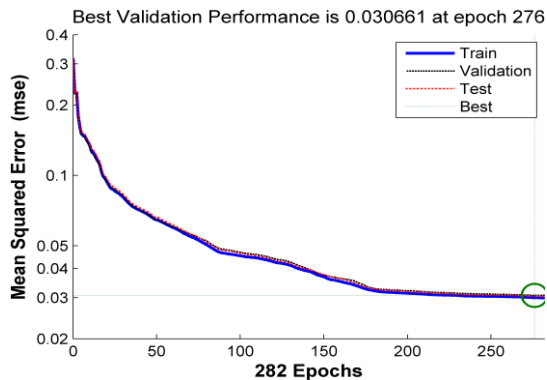


Fig. 6. Performance curves of the 3-hidden layer (7 + 6 + 4 neurons) model.

### B. Each Phase (Less-Varying Data) Models

As the all-phase data analysis (Section V.A) displayed a large data differential and variation across the Greener Homes Scheme three phases whereas the data variation in each phase is expected to be less than that in the all-phases case. Thus this section aims at training, validating and testing the ANN models for each phase. The good models obtained for Phase 1 (first 14,192 data), Phase 2 (next 7,239 data) and for Phase 3

(last 10,129 data) are shown in Table II-Table IV, respectively.

The analysis results from the three phase exhibit Irish market of renewable energy for residential heating. The bioenergy technologies were introduced at a large scale in Phase 1 together with the availability of woodchip and wood pellet supplies. Solar (tubes and panels) was commenced in Phase 2 with strong grant support, low maintenance and reducing installation costs and became a competing technology to bioenergy. By Phase 3, the Irish building regulations made renewable technologies compulsory in new build. This change, coupled with the global cost reduction, highlighted solar energy attractiveness. Whereas geothermal technologies require high maintenance compared with the alternatives and are expensive in terms upfront costs and unable to produce the required grade of heat.

TABLE II: ANN MODELS AND PREDICTED % RET CHOICE FOR PHASE 1 (FIRST 14,192 DATA)

ANN Model (nr = neuron)	Bioenergy (%)	Solar (%)	Geothermal (%)	Error (%)	No. of iterations
1 hidden layer (7 nrs)	23.8	39.9	24.4	11.7	955
2 hidden layers (7 + 4 nrs)	25.9	37.8	29.0	7.3	255
3 hidden layers (7 + 6 + 4 nrs)	25.6	37.8	29.6	7.0	220

TABLE III: ANN MODELS AND PREDICTED % ET CHOICE FOR PHASE 2 (MIDDLE 7,239 DATA)

ANN Model (nr = neuron)	Bioenergy (%)	Solar (%)	Geothermal (%)	Error (%)	No. of iterations
1 hidden layer (7 nrs)	7.4	70.1	18.4	4.1	204
2 hidden layers (7 + 4 nrs)	7.0	69.3	19.7	4.0	221
3 hidden layers (7 + 6 + 4 nrs)	8.1	68.8	19.2	3.9	132

TABLE IV: ANN MODELS AND PREDICTED % ET CHOICE FOR PHASE 3 (LAST 10,129 DATA)

ANN Model (nr = neuron)	Bioenergy (%)	Solar (%)	Geothermal (%)	Error (%)	No. of iterations
1 hidden layer (7 nrs)	0.0	88.2	3.7	8.1	115
2 hidden layers (7 + 4 nrs)	1.3	88.8	3.0	7.0	163
3 hidden layers (7 + 6 + 4 nrs)	4.0	87.9	3.4	4.7	197



## VI. CONCLUSION

A procedure to analyze customers' choice of renewable energy technologies (RETs) using artificial neural networks has been developed. The relationships between input data such as the investment phases, grant amounts, date received, property age and total installation corresponding to each consumer and the target data consisting of renewable energy heating technologies such as solar, biomass and heat pump have been explored and each percentage RET choice has been estimated for all phases together and for each phase by using several developed neural network models. The developed procedure has been applied in to Irish Government's Greener Homes Scheme (GHS). The following conclusions are drawn:

- The procedure validates the implemented governmental GHS scheme and the related customer subsidiary policies. It could forecast customers' choices in the future renewable energy schemes if the data on grant, time, property characteristics, expected installation and policy are given.
- If the model was implemented during the GHS scheme, it could provide recommendations and solutions to re-allocate the fund (grants) to each technology type. For example, if the model was done by the end of Phase 1, the fund amount and number of granted consumers would be allocated more to the solar energy technology. It might also provide recommendations to adjust the duration of the following phase. Better estimation of demand and behavior of customers would also be obtained.
- The usefulness of the results is to demonstrate the impact of low carbon heat policies to induce technological policy development within different jurisdictions, to suggest the factors that influence consumers' decision-making regarding their choice of low carbon technologies. The procedure can be a useful tool to frame and evaluate energy policy in similar socio-economic situations.

## ACKNOWLEDGMENT

This work was conducted as part of the INDICATE project (grant agreement No.: 608775) and the UMBRELLA project (grant agreement No.: 314343) under the European Union's 7th Framework Programme for research. The authors are grateful for the support.

## REFERENCES

- [1] C. A. Gibson, M. A. Meybodi, and M. Behnia, "How carbon pricing impacts the selection and optimization of a gas turbine combined heat and power system: an Australian perspective," *Journal of Clean Energy Technologies*, vol. 3, no. 1, January 2015.
- [2] SEAI (Sustainable Energy Authority of Ireland), *Greener Homes Scheme Statistics 2006-2011*, 2013.
- [3] IEA (International Energy Agency), *World Energy Outlook*, 2012.
- [4] M. Kennedy and B. Basu, "Overcoming barriers to low carbon technology transfer and deployment: An exploration of the impact of projects in developing and emerging economies," *Renewable and Sustainable Energy Reviews*, vol. 26, pp. 685-693, 2013.
- [5] M. Kennedy and B. Basu, "A study on the implementation of renewable heating technologies in the domestic sector in Ireland with implications on consumers' decision-making," *Journal of Cleaner Production*, vol. 44, pp. 133-142, 2013.
- [6] W. F. Van Raaij and T. M. M. Verhallen, "Patterns of residential energy behaviour," *Journal of Economic Psychology*, vol. 4, pp. 85-106, 1983.
- [7] A. Faiers, M. Cook, and C. Neame, "Towards a contemporary approach for understanding consumer behaviour in the context of domestic energy use," *Energy Policy*, vol. 35, pp. 4381-4390, 2007.
- [8] J. K. Basu, D. Bhattacharyya, and T. H. Kim, "Use of artificial neural network in pattern recognition," *International Journal of Software Engineering and Its Applications*, vol. 4, no. 2, April 2010.
- [9] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4-37, 1999.
- [10] C. M. Bishop. (April 2013). Neural network: A pattern recognition perspective. Technical Report NCRG/96/001, Aston University, UK. [Online]. Available: <http://www.ncrg.aston.ac.uk>
- [11] P. Kianmehr, W. Mansoor, and F. A. Kfoury, "Prediction of biogas generation profiles in wastewater treatment plants using neural networks," *Journal of Clean Energy Technologies*, vol. 2, no. 3, July 2014.
- [12] Neuralpower. (May 2013). [Online]. Available: <http://www.neuralpower.com>
- [13] C. Gan, V. Limsombunchai, M. Clemes, and A. Weng, "Consumer choice prediction: artificial neural networks versus logistic models," *Journal of Social Sciences*, vol. 1, issue 4, 2005.
- [14] *MATLAB and Neural Network Toolbox*, the MathWorks, Inc., Natick, United States, 2011.



**Van-Nguyen Dinh** obtained a bachelor degree from the National University of Civil Engineering, Hanoi, Vietnam in 1999 then completed his master and Ph.D. degrees in structural engineering in 2003 and 2008, respectively, at the Asian Institute of Technology, Thailand. His Ph.D. thesis won the MSIWF research award (Japan). He had taught at Konkuk University, Seoul and was a Marie-Curie experienced researcher at PLAXIS BV, the Netherlands and Trinity College Dublin, Ireland. Currently he is a research fellow at Trinity College Dublin. He is a member of Editorial Board of Progress and Communication in Sciences, an international journal.

His current research interests include urban energy efficiency, resource and consumption forecasting, modelling and vibration control of offshore and floating renewable energy devices, earthquake engineering, and structure and infrastructure engineering.



**Biswajit Basu** is a professor in the School of Engineering, Trinity College Dublin, Ireland. He obtained his Ph.D. degree from the Indian Institute of Technology, Kanpur, India, in 1997. He received a number of paper and research awards including the Dr. K.S. Krishnan Research Fellow Award. Currently, he is an associate editor of Journal of Structural Engineering (ASCE), an editor of IEEE Transactions on Sustainable Energy and also a member of editorial boards of other three international journals.

His current research interests include renewable energy technologies, smart grid, wind energy, offshore wind turbines, wave energy, structural health monitoring, passive, active and semi-active vibration control, random vibrations, wind and earthquake engineering, time-frequency and wavelet analysis, non-stationary processes, traffic flow theory and control.



**Matthew Kennedy** is the low carbon technology manager at Sustainable Energy Authority of Ireland. He completed his Ph.D. degree at Trinity College Dublin, Ireland in 2013. He is a member of the Technology Executive Committee of the UN, a member of the Advisory Board of the International Climate Technology Centre and Network (CTC/N) and the vice chair of the International Energy Agency (IEA) Renewable Energy Technology Deployment (RETD) Implementing Agreement.

His specialties include low carbon energy technology development, sustainable energy RD&D, energy technology roadmap development and modelling, technology transfer, international climate change and energy negotiations, government-industry relations, energy RD&D spending data and analysis.