



SDI Review Form 1.6

PART 1:

Journal Name:	British Journal of Economics, Management & Trade
Manuscript Number:	2013_BJEMT_4185
Title of the Manuscript:	Technoeconomic Projections with Artificial Neural Networks Using an Ensemble of Sparsely-sampled Bootstrapped Data

General guideline for Peer Review process is available in this link:

(<http://www.sciencedomain.org/page.php?id=sdi-general-editorial-policy#Peer-Review-Guideline>)

- This form has total 7 parts. Kindly note that you should use all the parts of this review form.



SDI Review Form 1.6

PART 2: Review Comments

	Reviewer's comment	Author's comment <i>(if agreed with reviewer, correct the manuscript and highlight that part in the manuscript. It is mandatory that authors should write his/her feedback here)</i>
<u>Compulsory</u> REVISION comments	Line 88: A blank space between the words “where” and “in” is required. Lines 152-154: In general terms, the output layer of an ANN is able to have more than one output-target value. It is not mandatory to have only one output unit. Please, check. Lines 161-165: Did the Authors test other kinds of ANN such as Time Lagged Recurrent networks (TLRNs) or Generalized Feedforward Networks (GFNs) or Recurrent Networks (RNs) etc? How did they decide to develop the MLP ANN which in any case is a universal approximator? Please, clarify.	- Corrected (Line 207:Revised MS) - Yes, more than one output is feasible. Necessary correction is made. (Lines 249-251: Revised MS) - The authors acknowledge the existence of the variety in ANN architectures. However, the present study is objectively devoted more to indicate the feasibility of using ANN-based approach to exercise forecasting with limited/sparse data. As such the study is kept restricted to a simple ANN and the efficacy of the proposal on forecasting with sparse data is illustrated. Nevertheless per Reviewer's suggestion, possibility of using different ANN architectures (in lieu of MLP) without any loss of generality is mentioned in the revised MS with relevant references posted: – See Lines 259 through 275



SDI Review Form 1.6

	4. Lines 186-187: Have the Authors tested other MLP architecture structures such as of two or more hidden layers? Also, how was the number of 10 hidden neurons (criteria) chosen in the hidden layer? e.g. the number of neurons must be a little bit higher for generating a configuration of the decision areas, complex enough, for a given problem, but if the number of neurons is too big, then the number of connections becomes too big and there is the risk for the balance of these connections not to be calculated only using available examples. In this way it is possible that the network generates noise (<i>Hornik, K., Stinchcombe, M., White, H., - Neural Networks, 2, 359, 1990</i>). According to the available literature, the number of hidden neurons should be between the size of the input layer and the size of the output layer (should be about 2/3 the size of the input layer, plus the size of the output layer). For example. Shibata and Yusuke (2009) proposed that with respect to the ANN learning stability, a rough guideline of appropriate number of hidden neurons $N^{(h)}$ and learning rate η are introduced as $N^{(h)} = \sqrt{N^{(i)} \cdot N^{(o)}}$ and $\eta = 32 / \sqrt{N^{(i)} \cdot N^{(o)}}$ where $N^{(i)}$ and $N^{(o)}$ are the number of input and output neurons respectively (<i>Katsunari Shibata and Yusuke Ikeda, "Effect of number of hidden neurons on learning in large-scale layered neural networks," in ICROS-SICE International Joint Conference, Fukuoka, Japan, 2009, pp. 5008-5013</i>). Please clarify.	4. Again, the study is focused to address the feasibility of ANN-based forecasting with sparse data using an appropriate bootstrap algorithm. Hence, ANN structural optimization is not attempted so as to limit the scope of the study to the objective posed. However, definitely as pointed out by the Reviewer concerns exist as regard to ANN selection (in terms of its architecture and operation etc.) regardless of the objective for which the network is used. Therefore, an elaboration on relevant issues as pointed out by the Reviewer is duly included as can be seen in: Revised MS: Lines 334-475 with appropriate references posted.
--	--	--



SDI Review Form 1.6

	5. Lines 262-272: The total number of bootstrap samples B is ranging between 50 and 200. Which is the appropriate method to choose the optimum B number? Also, according to the Authors, bootstrapping involves “a sample substituted with a replacement”. Because of the replacement process, each bootstrap sample may include multiple copies of some observations and no copies of other observations from the original set. Is this a supervised or a random process? How can the repetition of the same observations-prototypes affect the ANN’s training? Does this lead to the overtraining of the developed ANN model and to the memorization of the problem or not? Please clarify. 6. Line 285: A relevant reference for the Nyquist-Shannon criterion is requested. 7. Line 362: A blank space between the words “and” and “its” is required. 8. Line 371: A relevant reference of the sampling theorem and Nyquist sampling rate is requested. 9. Lines 424-426: Why is this necessary? Do the authors mean that their data is the result of a compatible mathematical algorithm which correlates the generated electric power with wind speed and not the observed wind park values of the aforementioned reference [4]?	5. Justifiable details on B are presented in Revised MS: Lines 502-541 See also revised Fig. 7 Details as regard to concerns expressed about: involves “a sample substituted with a replacement” are explained in: Revised MS – Lines 557 -622 6. Relevant references included: [43]-[49] Revised MS 7. Corrected 8. Relevant references included: [43]-[49] Revised MS 9. The wind park values in Ref [4 old] /Revised [20] are fitted (empirically) to deduce a compatible algorithm rigorously with a huge data set and adopted for ANN applications <i>in lieu</i> of using raw data as
--	--	---



SDI Review Form 1.6

	Please clarify. 10. Page 11, Fig. 4: Authors have to answer-clarify the following questions-remarks: - According to reference [4] (Li et al, 2001; page 277; Figure 2) the wind speed seems to range up to 50mph. How was the curve of power P (kW) for wind speed values greater than 50 mph plotted; - How were the subsets A and B selected (criteria)? Have the Authors examined the case of B being the “ex post” and A the “ex ante” subsets? Do they believe that they will have similar results and same forecasting accuracy? - As it seems, for wind speeds above 60mph (~26.5m/s) the generated power is constant (about 280 kW). In Fig. 2, [4] (Li et al, 2001; page 277; Figure 2) at 500 KW the output curve is leveled and is at its maximum value. The current generated is proportional to the generator shaft torque. The shaft is connected through a gear train to the hub of the turbine. The output power is limited by controlling the torque produced by the turbine blades. This is accomplished by using ailerons at the tip of the blades to reduce the blade lift. This is necessary during high winds in order to protect the equipment. How difficult is to forecast a constant value	such. The present study uses this algorithm on $f(.)$ and $g(.)$ in normalized form. 10. - Drafting error in original MS is now corrected in Fig. 4 limited to 50 mph - The <i>ex ante</i> regime (for forecasting) normally commences at the cessation of ex post regime where in the data is available a priori. The forecast is supposedly to be done just from the end of ex post. This is indicated in Section 5 at large: - The accuracy of forecasting however relies on the extent to which the test ANN is trained with the ensemble of data (bootstrapped or otherwise) adopted during training. Specifically, the forecast trend would follow the extensiveness and stochastics of the <i>ex post</i> details that prevail just prior to and in the vicinity of (<i>ex post</i>)-to-(<i>ex ante</i>) transition as observed by Neelakanta and Yassin in [3]. Further, as indicated in [3], such features of forecast details seen ahead in the <i>ex ante</i> frame (with constancy or jagged variations) are true only to an extent within a cone-of-forecast. The method of constructing
--	---	---



SDI Review Form 1.6

	(B, ex ante subset) even if using an ensemble set of sparse and insufficient sampled-data? IV. Are the greater than 60mph (~26.5m/s) wind speed values actual-observed values? Which is their rate (%) of occurrence in relation to the total wind speed data set (even for places like the 6100-foot elevation at the Fort Davis, Texas, USA)? V. Actually, most commercial wind turbines do not operate for wind speed values higher than 25m/s (or approximately 50mph). Hence the utilization of so high wind speeds in order to check the reliability of the proposed algorithm is quite peculiar. Please comment. 11. Page 12, Fig. 5: According to the proposed equation of $f(v)$ in reference [4] (Li et al, 2001; page 279; equation 1) the values of $f(v)$ range between 0 and 3.957. In the present Fig. 5 (upper graph) the values of function $f(v)$ range between 0 and 1.0. Why did this happen? 12. Lines 489-495. Caption of Fig. 6: I. How was the k windows' range (8 mph)	relevant progressive error-bar on forecast values is detailed in [3]. The remarks as above is included in Section 7: lines 960-982 Revised MS (iv) Greater than 60mph wind speed values indicated in the paper are projected forecasts based on Fig. 2 and $f(.) \times g(.)$ functional relation available in Ref [20]. We have no additional data to justify the computed values (v) The present study offers a method to project the ex post details in to an ex ante regime and details of Ref. [20] were adopted as exemplar data to illustrate the proposed methodology. We have no other details to support, extend or negate the exemplar data used. See Lines: 969-978 Revised MS 11. The present study considers using normalized version of $f(.) \times g(.)$. This is explicitly mentioned in Line 847 Revised MS 12.(I to III)
--	---	--



SDI Review Form 1.6

	selected? Please clarify. II. The referred windows' range ($k \geq 6$) is not correct (see Table 2). Please check. III. In the caption of Fig. 6, why is it referred that the number of windows is $k=13$? According to Table 2 the number is equal to $k=12$. Please, check. 13. Lines 500-506: The description of functions $f(v)$ and $g(\theta)$ is not appropriate. It is very difficult for the reader to understand. He/she has to read the reference paper [4] (Li et al, 2001) firstly. A better description is required. 14. Lines 515-517: See specific remark #10 (II & III). Please clarify. 15. Page 20, Fig. 8 The scale of the left vertical axis (kW) is incorrect. The numbers have to be multiplied by 10 in order to agree with Figure 4 respectively. Please, check it. 16. Lines 542-543 and Fig. 8: How interesting and how difficult is the prognosis of the generated power when the wind speed is greater than 60 mph ($\sim 26.5\text{m/s}$)? As it seems (Fig. 4 and Fig. 8) the generated power is constant. Please clarify.	There were typographical errors on k ranges. Now, corrected in the revised MS; $k = 1$ to 5 (ex post) and $k = 6-13$ (ex ante). See in Tables 1 and 2. 13. Lines 828-839 Revised MS Rewritten description is presented. 14. Relevant revisions made in the rewritten text on Remark # 10 holds water. 15. The correction is made 16. It will be of interest to determine the prognostic aspects of generated power when the wind speed is in the vicinity of 60 mph or higher. As per the details of Fig. 2 in [20], the generated power remains steady at wind speeds greater than 60 mph; and it appears that most commercial wind turbines currently do not operate for wind speed values higher than 50 mph. As such, specific considerations in elucidating the
--	--	---



SDI Review Form 1.6

	17. Lines 620-622: Reference [4] is not written correctly. The correct form is as follow: <i>Li S, Wunsch DC, O'Hair EA, Geisselmann MG. Using neural networks to estimate wind turbine power generation, IEEE Trans. on Energy Conversion. 2001; 16 (3):276-282.</i>	wind-turbine performance at 50 mph or more <i>via</i> forecasting and interpreting the results could be an open-question for future research if relevant field details are available. The present study, however is confined to the forecast results as presented in Fig 8(f) using the available details of Figs. 2, 7 and 8 in [20] viewed in terms of relevant functional relation $f(.) \times g(.)$. Lines: 969-978 Revised MS 17. Correction is made as indicated. -----
--	---	---



SDI Review Form 1.6

Minor REVISION comments		
Optional/General comments	The paper 2013_BJEMT_4185 deals, according to the authors, with a quite interesting scientific subject, i.e. the use of artificial neural networks (ANN) that can be designed exclusively for ex ante forecasting in a technoeconomic context using an ensemble set of sparse and insufficient sampled-data availed ex post. For that purpose, they developed and trained ANN with bootstrap-enhanced, scarcity-recovered sparse data on wind-power generation statistics and associated plant economics. That is, the ANN output obtained depicts forecast projections on the productivity of electric power generation in the ex ante regime. However, the economic part of the paper title is missing! According to the Authors, results indicate that the efficacy of the proposed method in the context of technoeconomics is illustrated with a practical real-world system wherein, sparsity recovery is done and relevant bootstrapped data sample-set is effectively applied to a test ANN for robust forecast purposes. The specific study, despite its innovation and scientific interest, may be considered for publication only after major revision by the authors. Recapitulating and taking into consideration the above mentioned indicative comments, I believe that the manuscript under evaluation, although examining a challenging topic, needs substantial revision before it is considered for potential publication in the British Journal of Economics, Management & Trade.	In view of the concern that the economic part of the paper title is missing, more stress on relevant considerations are presented in the Introduction part of the revised MS. Lines 31-145 The authors hereby duly acknowledge the critical review and constructive criticisms offered by the Reviewer I. Thank you!