System Modeling and Prediction of Respiratory Diseases using Machine Learning

Sampa Paul[1], Indrajit Naskar[2], A. K. Pal[3]
Heritage Institute of Technology, Kolkata, India.
[1]*sampa.paul.vlsi@gmail.com* [2]indrajit.naskar@heritageit.edu[3]arabindakumarpal@gmail.com

Abstract

In the first part of this paper a human respiratory model and their responses are presented and in the second part machine learning algorithm is used to predict the respiratory diseases from the datasets. This endeavor proposes a simple representation of the mathematical model of the human respiratory system comprises of nasal cavity, trachea, bronchi and alveolar sacs. Ranging from a very common disease called bronchitis to highly perilous diseases like emphysema are considered in this venture and the variation of responses are observed with input pressure. In the conventional computation and control the system modeling is very much essential for analysis, whereas, machine learning algorithms use computational methods to learn information directly from data without relying on a predetermined equation as a model. Logistic regression algorithm is applied to find natural patterns in respiratory data that generate insight and help to make better decision and prediction of respiratory diseases. The Logistic regression model makes a clear decision boundary between the respiratory diseases.

Keywords: *Bronchitis, Emphysema, Logistic regression, Mathematical model of the human respiratory system.*

1. Introduction

The human respiratory system can be broadly classified into four sections beginning with the nasal cavity. In this model the trachea-bronchial tree is divided into 24 generations (Table 1), where generation '0' is the trachea, the generation from 1 to 19 is counted for bronchi and the generation 20 to 23 corresponds to alveolar sacs. Starting from the trachea, considering each branch of a given generation divides into two identical daughters; therefore generation 'n' has 2n branches [1-5]. Each section have their characteristics Resistance, Inertance and Compliance due to the obstruction offered by the section, obstruction to the change in airflow through the section and expansion or contraction of the concerned section respectively [6-8]. These characteristics are used to develop the physical model of respiratory system as shown in Fig. 1. Validation of the proposed model is done in real time by Spirometer and model is verified through MATLAB System Identification Toolbox.

Accurate modeling of respiratory system is very difficult and thus the response observed from such modeling is dubious [9-12]. Machine learning approach uses only data to analyze the system performance and therefore it gives accurate information about the system [13]. In this paper Logistic Regression algorithm is applied to distinguish different respiratory diseases by creating decision boundary.

2. Modeling

2.1 Respiratory System

Table 1. Geometrical dimensions of the morphological model of the respiratory system [5]

Generation Number	Number of airways per generation	Airways diameter in cm	Length in cm	Total airways Area in cm^2	Avg. air flow velocity in cm/s	Resistance in cm of H_2O/ ltr./s Calculated using formula	Inertance in cm of H_2O/ ltr./s^2 Calculated using formula	Compliance in cm of H_2O/ ltr./s Calculated using formula
z	n(z)	d=2r	l	s	u	$8\mu l/\pi r^4$	$\rho l/s$	$ls/\rho n(z)u^2$
0	1	1.800	12.00	2.540	197.00	0.0860	0.0059	0.06311
1	2	1.220	4.76	2.330	215.00	0.0080	0.0025	0.0964
2	4	0.830	1.90	2.130	236.00	0.0075	0.0011	0.0145
3	8	0.560	0.76	2.000	251.00	0.0072	0.0004	0.0024
4	16	0.450	1.27	2.480	202.00	0.0145	0.0006	0.0038
5	32	0.350	1.07	3.110	161.00	0.0167	0.0004	0.0032
6	64	0.280	0.90	3.960	126.00	0.0172	0.0002	0.0028
7	128	0.230	0.76	5.100	98.00	0.0159	0.0001	0.0025
8	256	0.186	0.64	6.950	72.00	0.0157	0.0001	0.0026
9	512	0.154	0.54	9.560	52.00	0.0141	7.0290e-05	0.0029
10	1024	0.130	0.46	13.400	37.00	0.0118	4.2718e-05	0.0035
11	2048	0.109	0.39	19.600	26.00	0.0101	2.4761e-05	0.0044
12	4096	0.095	0.33	28.800	17.00	0.0074	1.4258e-05	0.0064
13	8192	0.082	0.27	44.500	11.00	0.0054	7.5502e-06	0.0097
14	16384	0.074	0.16	69.400	7.20	0.0024	2.8689e-06	0.0105
15	32768	0.050	0.13	117.00	4.30	0.0048	1.4145e-06	0.0206
16	65536	0.049	0.11	225.00	2.20	0.0022	6.1943e-07	0.0638
17	131072	0.040	0.09	300.00	1.70	0.0020	3.8576e-07	0.0591
18	262144	0.038	0.08	543.00	0.92	0.0011	1.9021e-07	0.1632
19	524288	0.036	0.07	978.00	0.51	0.0005	8.9067e-08	0.4034
20	1048576	0.034	0.07	1740.0	0.29	0.0003	5.0062e-08	1.1099
21	2097152	0.031	0.07	2730.0	0.18	0.0002	3.1907e-08	2.2600
22	4194304	0.029	0.67	5070.0	0.99	0.0016	1.6444e-07	0.6640
23	8388608	0.025	0.07	7530.0	0.66	0.0001	1.2394e-08	0.1241

Table 2. Calculated values of R, L, C, RC and LC for different segments of human respiratory system design.

Different section of human respiratory system	R (H_2O/ ltr./s)	L (H_2O/ ltr./s^2)	C (ltr./cm of H_2O)	RC	LC
NASAL CAVITY	16.332700	0.0200000000	0.1320	2.156000	0.0027000000
TRACHEA	0.086000	0.0059000000	0.0631	0.005400	0.0003700000
BRONCHI	0.008700	0.0002929000	0.0461	0.000402	0.0000144000
ALVEOLI	0.000550	0.0000000647	1.0396	0.000571	0.0000000672

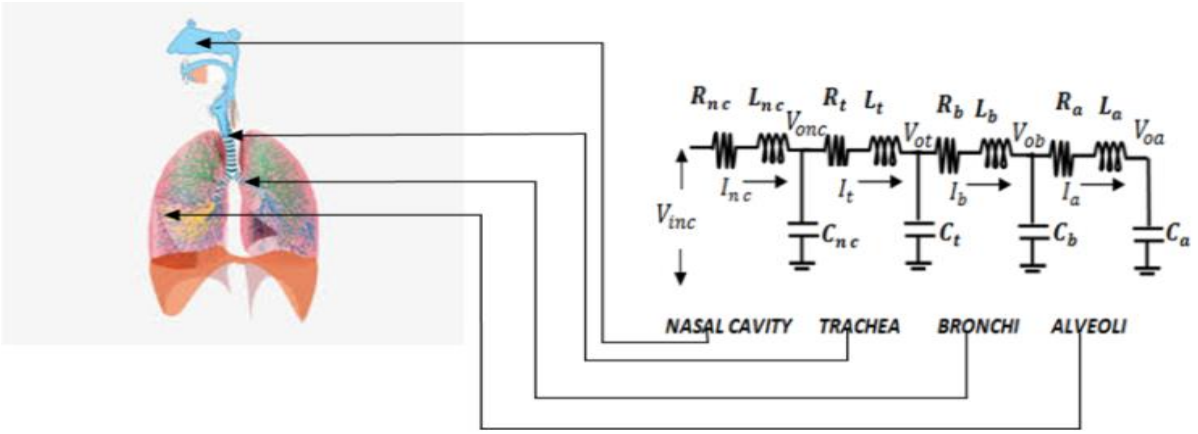

Figure 1. Electrical equivalent circuit diagram of the human respiratory system based on R, L and C model.

The Transfer Function (TF) of each section is determined by considering Fig. 1 which comes out to be in this form after executing Laplace transform, with all initial conditions to zero [1/(s^2LC+sRC+1)], where R, L, and C are the characteristic values of Resistance, Inductance and Capacitance of the section concerned [1-5]. The necessary calculations of R, L, and C to reflect it for the different generations of the respiratory system are shown in Table 1. Using Table 1 for 'generation 0', generation 1 to 19, and generation 20 to 23 correspond to the trachea, bronchi, and alveolar sacs respectively [10, 11]. The derived values of R, L, and C are tabulated in Table 2. The transfer functions for the different parts of the respiratory system are modelled by putting R, L, and C values in the developed electrical equivalent model as depicted in Fig. 1[12].

The individual transfer functions of the nasal cavity (TF_N), trachea (TF_T), bronchi (TF_B), and alveoli (TF_A) are derived from the circuit diagram of Fig.1 and corresponding components values presented in Table 2 as follows:

$$TF_N = \frac{1}{0.0027S^2+2.156S+1} \quad (1)$$

$$TF_T = \frac{1}{0.00037S^2+0.0054S+1} \quad (2)$$

$$TF_B = \frac{1}{0.0000144S^2+0.000402S+1} \quad (3)$$

$$TF_A = \frac{1}{0.0000000672S^2+0.000654S+1} \quad (4)$$

The resultant respiratory model TF_M is derived from the product of the individual models,

$$TF_M = TF_N * TF_T * TF_B * TF_A \quad (5)$$

3. Validation of the Respiratory Model

The derived model is validated in real time by spirometry test in the laboratory and the model transfer function (TF_R) is derived using MATLAB System Identification Toolbox [14-16].

$$TF_R = \frac{0.0030984}{\{(1.7s+1)(0.109s+1)(0.001s+1)\}} \quad (6)$$

Comparison of TF_R and TF_M with breathing (inhalation for 1.5 sec. and exhalation for 2.5 sec.) input signal is presented in Fig.2.

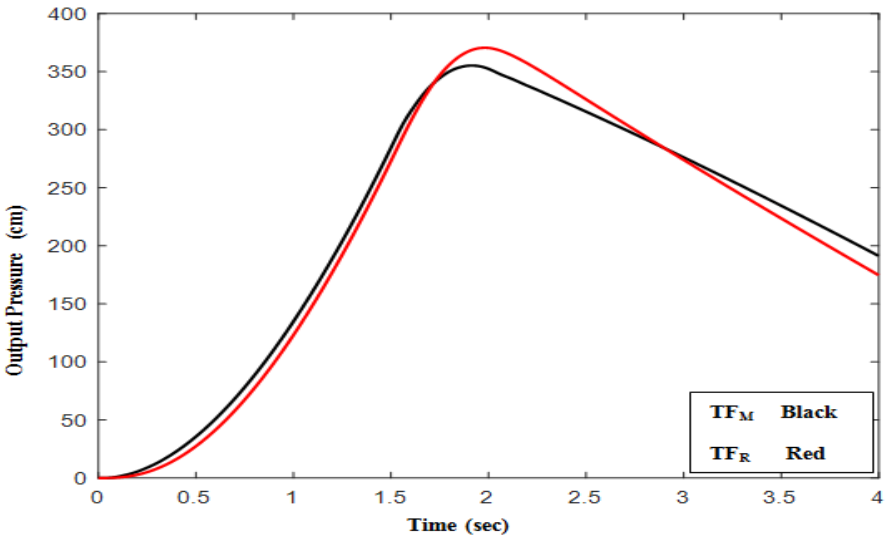

Figure 2.Model validation

4 Results

4.1 System Identification

TF_M, transfer function of healthy respiratory model is varied from its original value in diseased person. For instance, transfer function of bronchi (TF_B) in Bronchitis and transfer function of alveoli (TF_A) in Emphysema are not remaining same.

Due to bronchitis, the resistance of bronchial section is increased by around 2000 times and the bronchial model is changed to equation 7 from its initial model presented in equation 3.

$$TF_B' = \frac{1}{(0.0000144s^2+0.72s+1)} \qquad (7)$$

In case of emphysema, the capacitance of the alveoli is increased by more than 2000 times and the diseased alveoli model is modified to

$$TF_A' = \frac{1}{0.0003s^2+3s+1} \qquad (8)$$

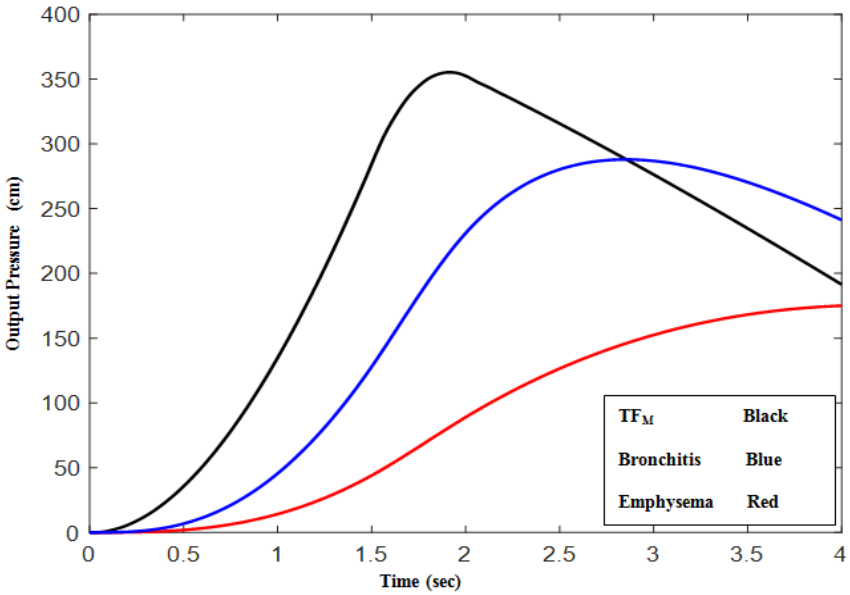

Figure 3. Comparative study of healthy and diseased Respiratory model

4.2 Machine Learning-Logistic Regression

The basic fundamental of logistic regression is to execute a decision boundary for a binary classification application. Logistic regression can also be used for multi-class classification.

Logistic regression algorithm is applied in the data sets (each of 250 nos. as shown in figure 4.a) to predict the disease and to make decision boundary as shown in the figures 4 (b, c, and d). For the iteration of 500 nos. training accuracy achieved 86.2% and each of the decision boundary is marked by the linear equation as shown in figures 4 (b, c, and d).

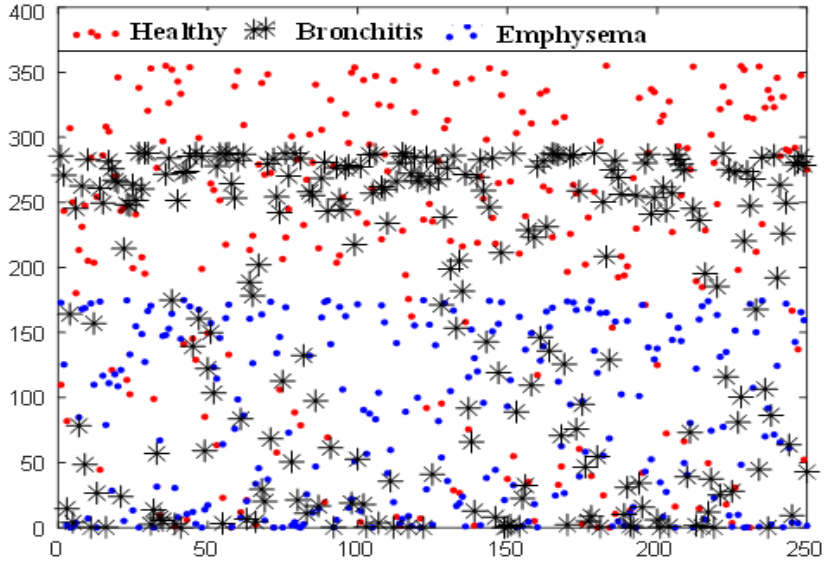

Figure 4. (a) Original Data plot

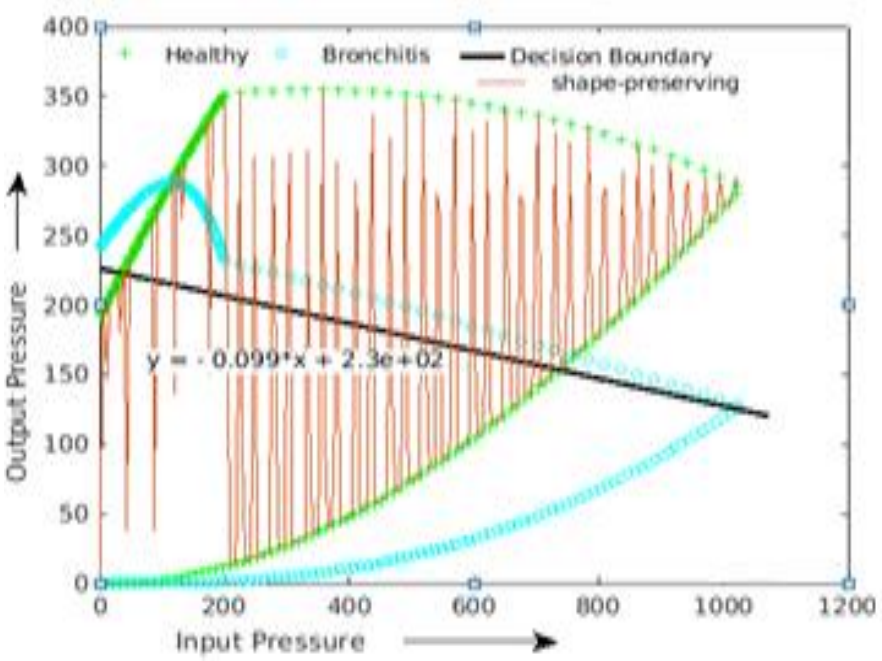

Figure 4. (b) Decision boundary between healthy and bronchitis

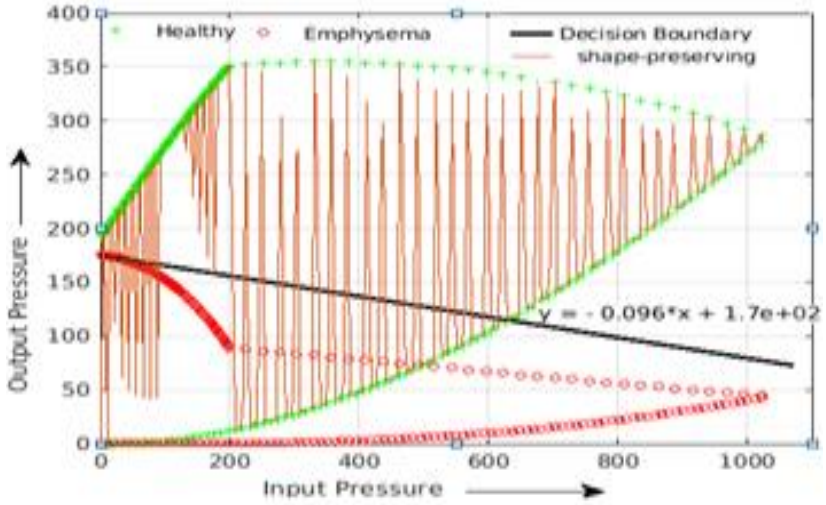

Figure 4. (c) Decision boundary between healthy and emphysema

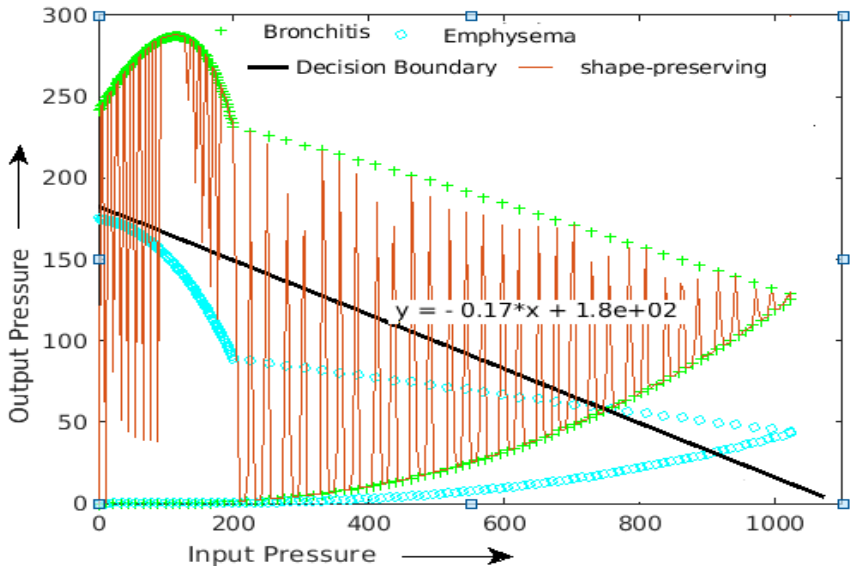

Figure 4. (d) Decision boundary between bronchitis and emphysema

Conclusion

It is practically impossible to attain a model that will exactly match with the real respiratory system given the fact that it is not viable to consider the numerous contributing parameters and precise quantitative as well as qualitative analysis of the human respiratory system. The model developed is validated with a real human respiratory system and it matches up to 95.63% to the model of the real respiratory system. Spirometer and MATLAB System Identification Toolbox are utilized to validate the theoretical model of the respiratory system. The logistic regression used to predict the diseases, gave a distinct separation between the data sets (diseases). It helped to predict the diseases by decision boundary and its corresponding mathematical equations.

References

1. *Fleur T. Tehrani, Mathematical model of the human respiratory system in chronic obstructive pulmonary disease, V7(6) , Online ISSN 2053-3713, pp. 139-145, (2020), doi: 10.1049/htl.2020.0060*
2. *Ibrahim M. Mehedi, Heidir S. M. Shah,1 Ubaid M. Al-Saggaf, Rachid Mansouri, and Maamar Bettayeb, Fuzzy PID Control for Respiratory Systems, Hindawi, Journal of Healthcare Engineering V 2021, Article ID 7118711, pp. 1-6, (2021), doi.org/10.1155/2021/7118711*
3. *J. Fernandez de Canete, D. Cuesta, A. Luque , J. Barbancho Physical modelling and computer simulation of the cardio-respiratory system based on the use of a combined electrical analogy, Taylor and Francis, pp. 453-488, (2021), doi.org/10.1080/13873954.2021.1977335*
4. *P. Ghafarian, H. Jamaati, and S. M. Hashemian, A Review on Human Respiratory Modeling, Tanaffos, V15(2), pp. 61-69, (2016).*
5. *B. Sul, Z. Oppito, S. Jayasekera, et al., Assessing airflow sensitivity to healthy and diseased lung conditions in a computational fluid dynamics model validated in vitro, Journal of Biomechanical Engineering, V140, pp. 1-49, (2018).*
6. *Riley, Corey (2017), A Mathematical Model of Cardiovascular and Respiratory Dynamics in Humans with Transposition of the Great Arteries, Rose-Hulman Undergraduate*

Mathematics Journal: V18(1), Article 13, https://scholar.rose-hulman.edu/rhumj/vol18/iss1/13, pp. 1-32, (2017).

7. *Martin Sébastien, and Maury Bertrand, Modeling of the oxygen transfer in the respiratory process, ESAIM: Mathematical Modeling and Numerical Analysis, EDP Sciences, V47 (04), pp. 935-960, (2013).*
8. *I. Jablonski, and J. Mroczka, A forward model of the respiratory system during airflow interruption, Metrology and Measurement Systems, V16 (2), pp. 219-232, (2009).*
9. *A. K Pal, Indrajıt Naskar, and Sampa Paul, A Fuzzy-Based Modified Gain Adaptive Scheme for Model Reference Adaptive Control, © Springer Nature Singapore Pvt. Ltd., Information and Decision Sciences, Advances in Intelligent Systems and Computing 701, pp. 315-324, (2018).*
10. *Boston: T. Martonen.: Medical applications of computer modeling: the respiratory system. (2001).*
11. *Lin S. L, Guo N. R, and Chiu C. C.: Modeling and simulation of respiratory control with Lab View. Journal of Medical and Biological Engineering, vol.32, no.1, 51-60. (2012).*
12. *Indrajıt Naskar and A. K Pal, :Automatic Indoor Air Quality Control of COVID-19 Patient Facilities Using Type-2 Fuzzy Controller: Journal of Physics, (2021).*
13. *Sekiya S, Sakaino S, and Tsuji T.: Linear logistic regression for estimation of lower limb muscle activations. In: IEEE Trans. on Neural Systems and Rehabilitation Engg., vol.27, no.3, pp. 523-532. (2019).*
14. *A. K Pal, R.K.Mudi, and R. R. De Maity, A Non-Fuzzy Self-Tuning Scheme of PD-Type FLC for Overhead Crane Control, Advances in Intelligent Systems and Computing book series(AISC), Springer,Vol.199, pp.35-42,(2013).*
15. *A. K Pal and J. Chakraborty, Design a Fuzzy Logic Controller with a Non-fuzzy Tuning Scheme for Swing up and Stabilization of Inverted Pendulum, Intelligent Computing, Communication and Devices, springer, pp.221-230,(2015).*
16. *A. K Pal , Indrajit Naskar, Sampa Paul and J Chakraborty, Auto-adaptive Fuzzy Controllers with Automatic Setpoint Adjustment, Intelligent Electrical Systems: A Step towards Smarter Earth, CRC Press, pp.181-189, (2021).*