



标题

Weighted Distance Based Outlier Factor Identifying and Its Application in Raw Wind Data Preprocessing





问题

- Generally, wind data preprocessing has not been specifically discussed in published literatures, nor the results or the validity of the existing preprocessing methods.
- The existing methods may not distinguish abnormal data in some situations.





贡献

- This paper presents an effective data-driven preprocessing algorithm.

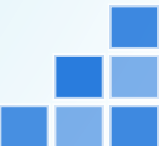




Figure 1

II. Raw wind data properties

	A	B	C	D	E	F	G	H	I	J	K	L
1	Time	Wind Speed (m/s)	Wind Power (kW)	Time	Wind Speed (m/s)	Wind Power (kW)	Time	Wind Speed (m/s)	Wind Power (kW)	Time	Wind Speed (m/s)	Wind Power (kW)
2	2012/11/21 12:00:00	5.7183	26312.8832	2013/1/31 20:00:00	4.2569		2013/2/25 15:00:00	10.0899	120150.5204	2013/3/19 8:00:00	0.0000	133676.0860
3	2012/11/21 12:15:00	6.0156	26147.7455	2013/1/31 20:15:00	4.8924		2013/2/25 15:15:00	10.0899	120150.5204	2013/3/19 8:15:00	11.1669	139648.4616
4	2012/11/21 12:30:00	6.3129	25990.6895	2013/1/31 20:30:00	5.3696		2013/2/25 15:30:00	10.0899	120150.5204	2013/3/19 8:30:00	12.7173	144727.4801
5	2012/11/21 12:45:00	6.6101	25853.8309	2013/1/31 20:45:00	5.8469		2013/2/25 15:45:00	10.0899	120150.5204	2013/3/19 8:45:00	13.3746	146251.4669
6	2012/11/21 13:00:00	6.6334	24884.7223	2013/1/31 21:00:00	6.3242		2013/2/25 16:00:00	10.0899	120150.5204	2013/3/19 9:00:00	14.0320	146400.7267
7	2012/11/21 13:15:00	6.5380	21272.0099	2013/1/31 21:15:00	7.2177		2013/2/25 16:15:00	10.0899	120150.5204	2013/3/19 9:15:00	14.2953	146283.9564
8	2012/11/21 13:30:00	6.4426	17661.7509	2013/1/31 21:30:00	8.1778		2013/2/25 16:30:00	10.0899	120150.5204	2013/3/19 9:30:00	13.5837	145707.5461
9	2012/11/21 13:45:00	6.3358	14005.8410	2013/1/31 21:45:00	8.8348		2013/2/25 16:45:00	10.0899	120150.5204	2013/3/19 9:45:00	12.8721	144580.4527
10	2012/11/21 14:00:00	5.8614	10983.4513	2013/1/31 22:00:00	8.2454		2013/2/25 17:00:00	10.0899	120150.5204	2013/3/19 10:00:00	12.1604	143073.4727
11	2012/11/21 14:15:00	5.3870	9180.8378	2013/1/31 22:15:00	7.9584		2013/2/25 17:15:00	10.0899	120150.5204	2013/3/19 10:15:00	11.9071	141696.8812
12	2012/11/21 14:30:00	4.9126	7389.7938	2013/1/31 22:30:00	8.9334		2013/2/25 17:30:00	10.0899	120150.5204	2013/3/19 10:30:00	11.9346	141118.6402
13	2012/11/21 14:45:00	4.4210	5712.7818	2013/1/31 22:45:00	7.3949		2013/2/25 17:45:00	10.0899	120150.5204	2013/3/19 10:45:00	11.9622	141338.6991
14	2012/11/21 15:00:00	3.8992	4453.0291	2013/1/31 23:00:00	12.3388		2013/2/25 18:00:00	10.0899	120150.5204	2013/3/19 11:00:00	11.9897	142351.9899
15	2012/11/21 15:15:00	3.3774	3680.1959	2013/1/31 23:15:00	11.1609		2013/2/25 18:15:00	10.0899	120150.5204	2013/3/19 11:15:00	11.9341	143078.8249
16	2012/11/21 15:30:00	2.8657	2946.6053	2013/1/31 23:30:00	9.8401		2013/2/25 18:30:00	10.0899	120150.5204	2013/3/19 11:30:00	11.8274	143652.7763
17	2012/11/21 15:45:00	2.7097	2253.0959	2013/1/31 23:45:00	9.1116		2013/2/25 18:45:00	10.0899	120150.5204	2013/3/19 11:45:00	11.7207	144406.3521

Figure 1. Sample of raw wind data





Figure 2

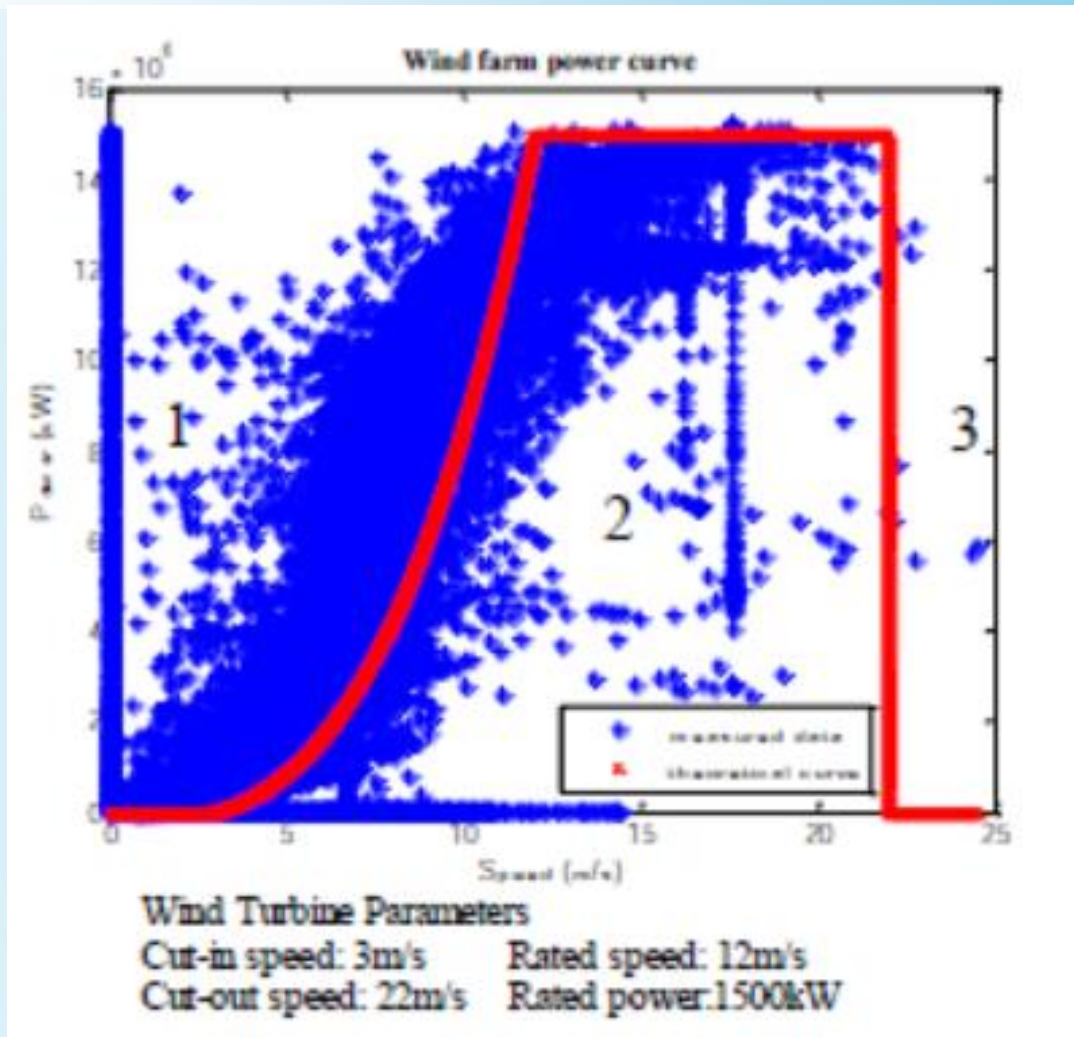


Figure 2. Wind farm power curve





Section III

III. Wind power curve properties

$$P^I = \begin{cases} 0, & \text{if } 0 < V < V_{ci} \\ \frac{V^3 - V_{ci}^3}{V_r^3 - V_{ci}^3} P_r, & \text{if } V_{ci} \leq V < V_r \\ P_r, & \text{if } V_r \leq V < V_{co} \\ 0, & \text{if } V_{co} \leq V \end{cases} \quad (1)$$





Section IV

IV. LOF Algorithm

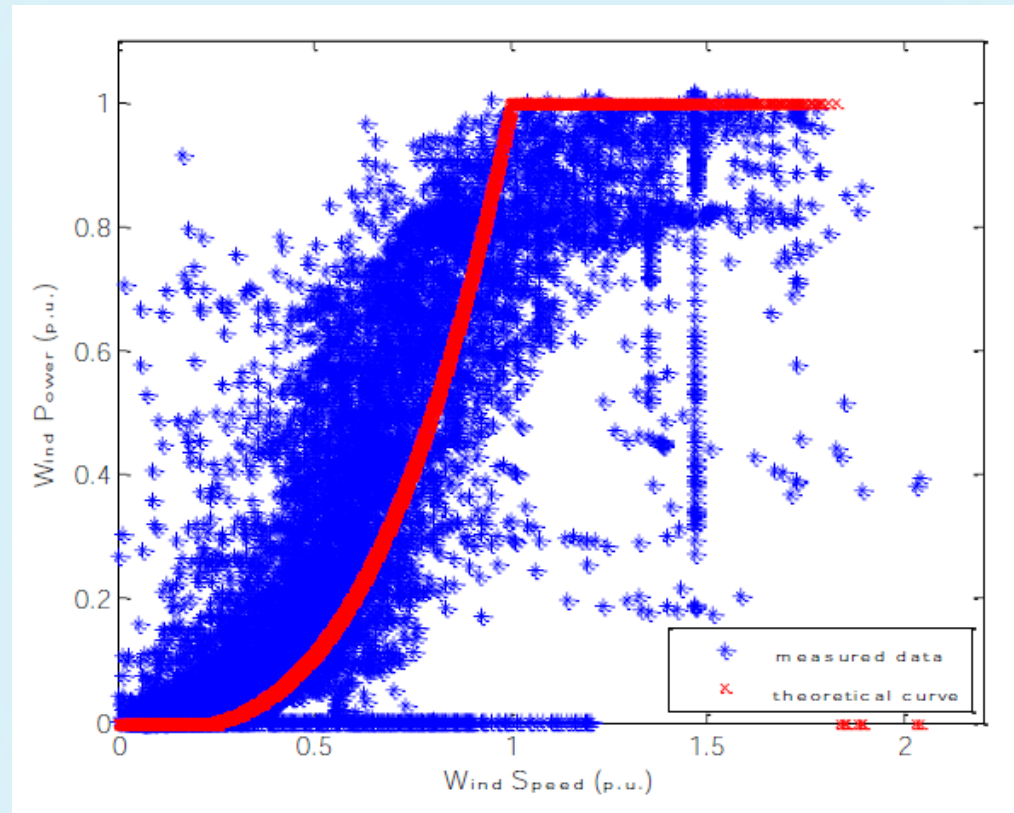


Figure 3. Test data in subset TD





Figure 4

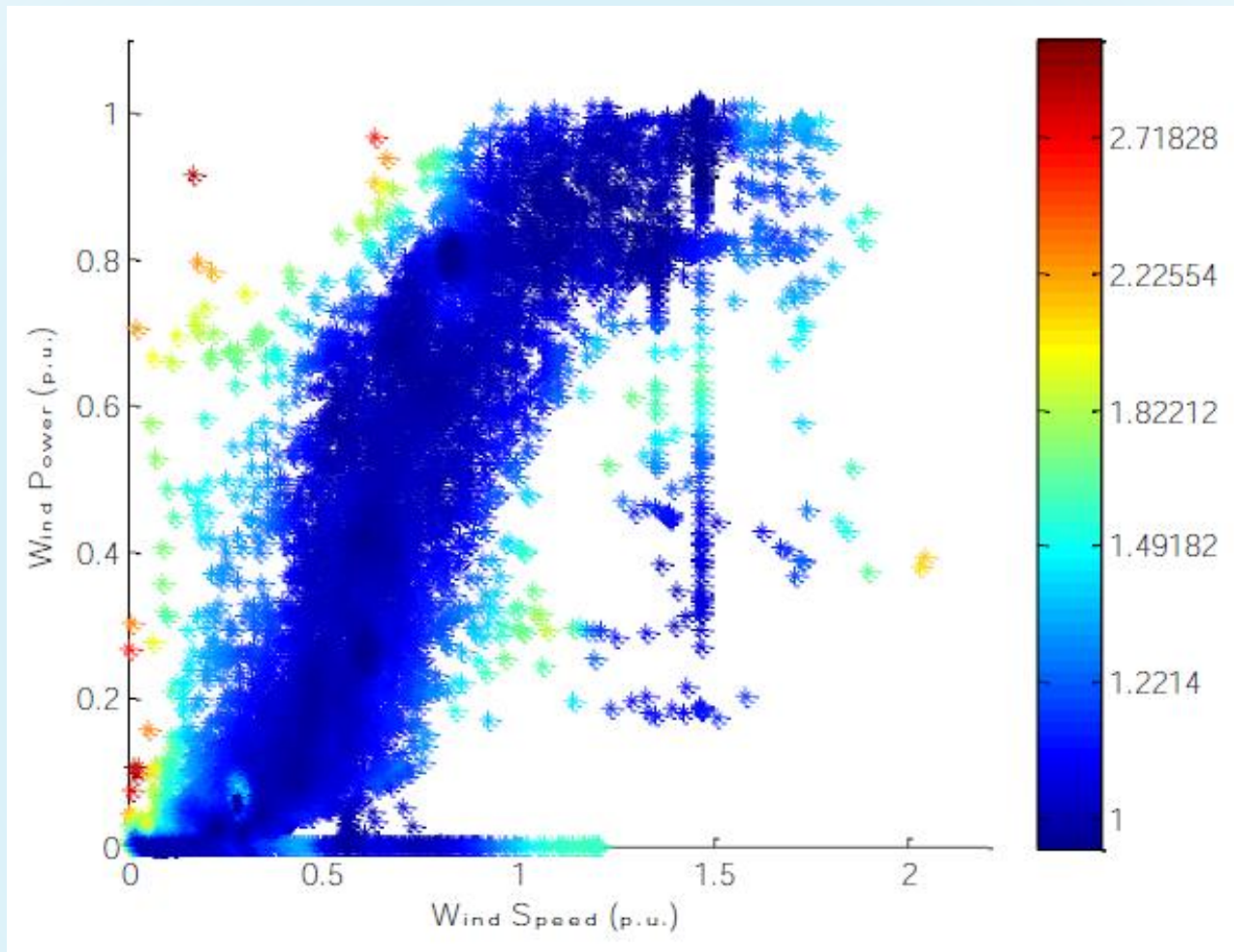


Figure 4. Outlier factors by LOF





Figure 5

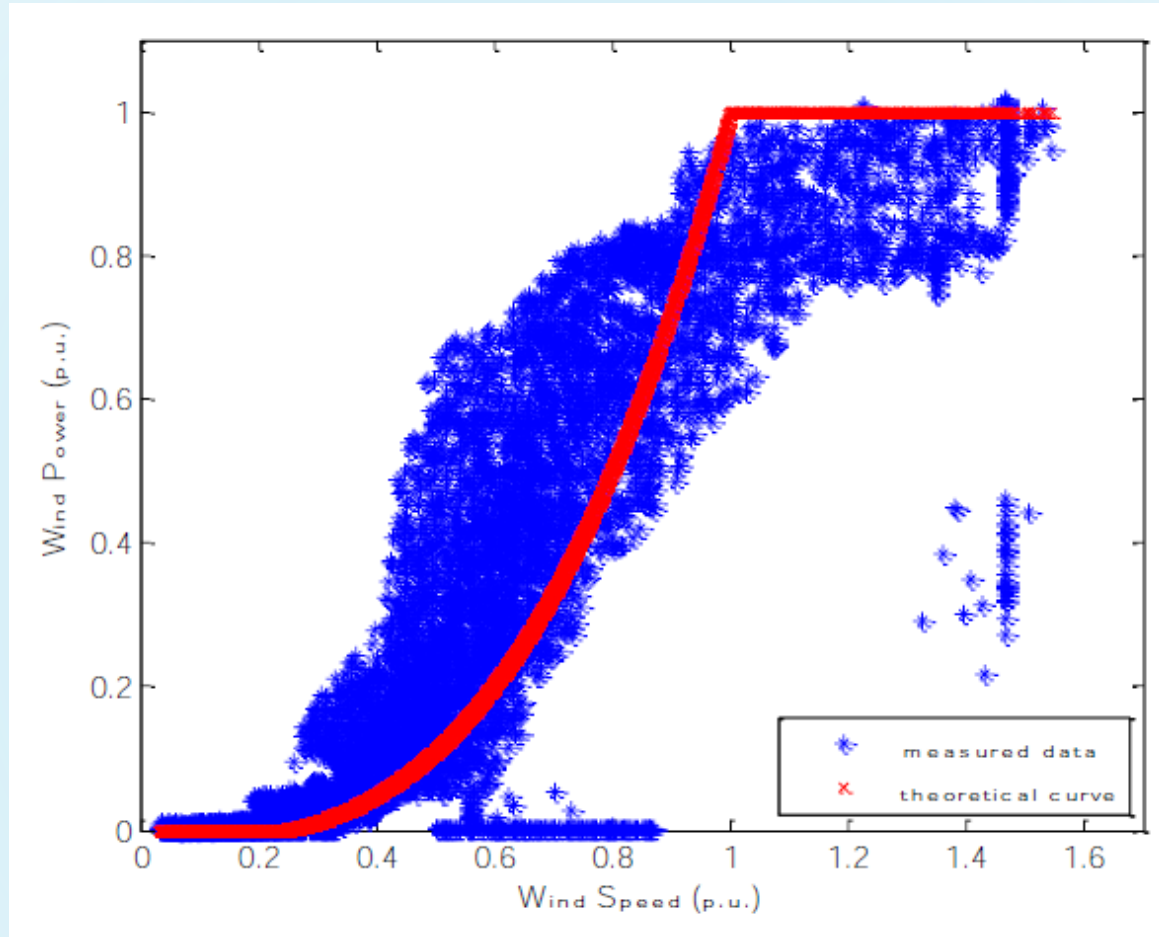


Figure 5. Data after outlier detection by LOF





Section V

V. WDOF Algorithm

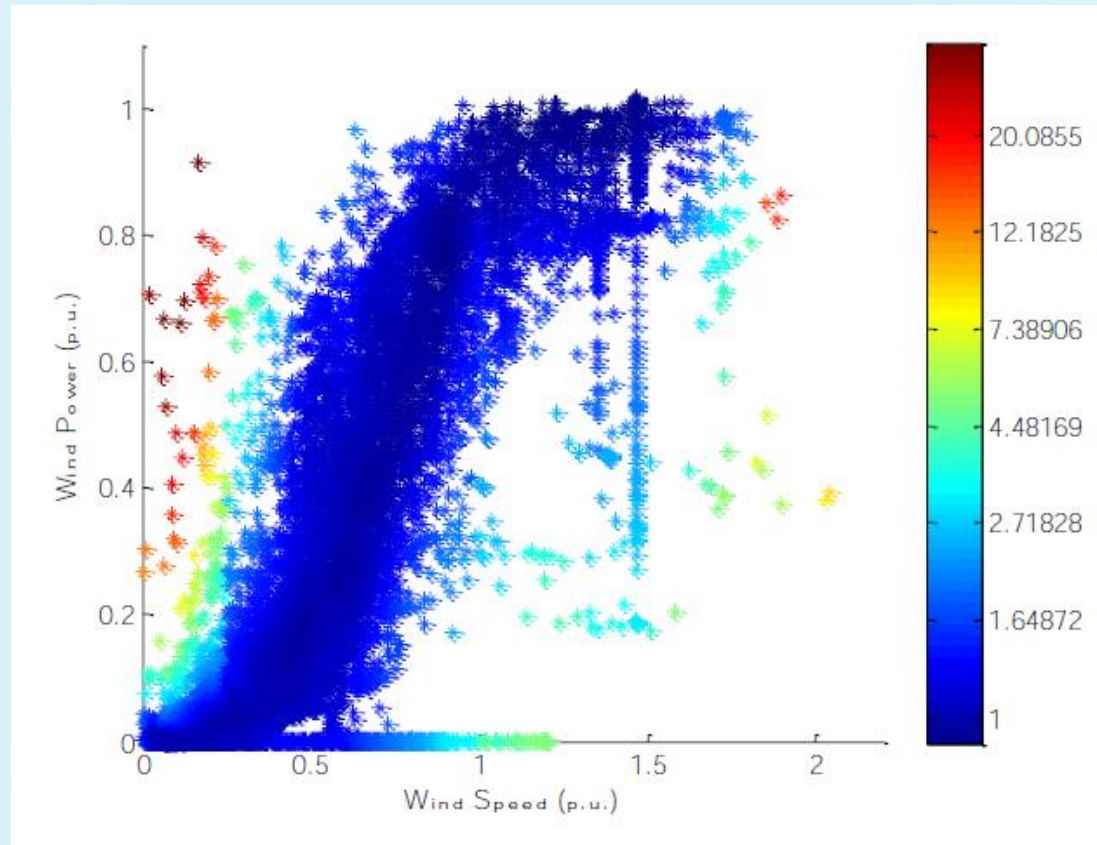


Figure 6. Outlier factors by WDOF

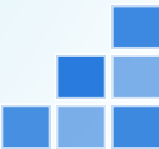




Figure 7

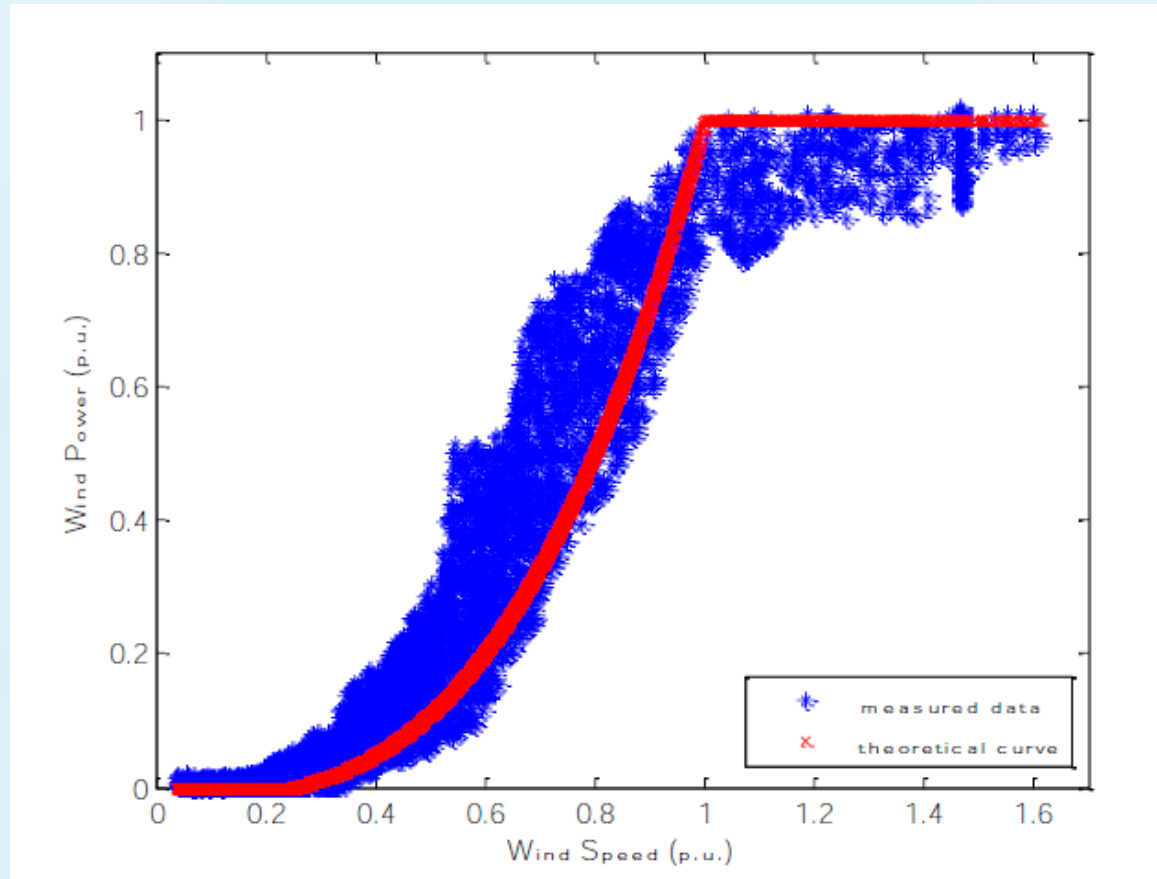


Figure 7. Data after outlier detection by WDOF

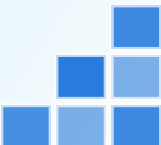




Table I

Table I Comparison of LOF and WDOF

TABLE I. COMPARISON OF LOF AND WDOF

Algorithm	Time (s)	Accuracy	Resolution
LOF	19.38	85.14%	2.19
WDOF	24.96	97.86%	32.1

