Temperature, pressure, relative humidity and rainfall sensors early error detection system for automatic weather station (AWS) with artificial neural network (ANN) backpropagation

P Wellyantama\(^1\) and S Soekirno\(^1\)

\(^1\)Physics Department, University of Indonesia, Depok, West Java, Indonesia

E-mail: pradawellyantama@gmail.com

Abstract. To improve the quality and quantity of meteorological data over Indonesia, Meteorology Climatology and Geophysics Agency of Indonesia (BMKG) is continuously developing automatic weather observations. BMKG has 63 units Automatic Weather Station (AWS) and 165 units Automatic Weather Observation System (AWOS) both inside and outside the BMKG Station environment. To make the control of sensor conditions easier, especially for temperature, pressure, relative humidity, and rainfall sensors, an additional system is needed to monitor and warn when problems occur with these sensors. The correlation among weather parameters data is the key to monitoring the sensor condition, these data are going to be trained and tested with the Artificial neural network (ANN) method. Then, the sensor condition (normal or error indicated) can be well detected based on AWS’s data. The quality improvement of automatic weather station data is expected to increase the utilization of the data.

1. Introduction

Indonesia is a very large archipelago country with an area of about 1,919,440 km\(^2\), Indonesia has 17,508 islands and a long coastline of about 81,000 km [1]. In Indonesia, weather information has an important role both, to plan and to operate daily life in various sectors. From the construction development, economy, social, transportation, tourism, health, etc. In the construction development sector for buildings, airports and ports require information about wind direction, wind speed, and tides, in the economic sector, the analysis of inflation in a region requires information on wave height, the tourism sector requires weather forecast data, temperature, humidity, wave height, and the land, sea, and air transportation sector requires weather information data, air pressure, wave height, and significant weather maps. The Meteorology Climatology and Geophysics Agency (BMKG) has 183 Meteorological Stations that observe and provide weather information spread across Indonesia. Weather observations are carried out manually or by using human power to observe weather parameters using conventional weather instruments and there are also automatic observations using digital weather instruments. Of the 183 meteorological stations, 62 use fully automatic observation, and the rest use a conventional instrument. BMKG has 63 units meteorological AWS (automatic weather station) and 165 units AWOS (automatic weather observation system) spread throughout Indonesia, both inside and outside of the Meteorological Station zone. Some digital instruments usually unnoticed if there is a problem with the values generated by the sensor, if they are not compared to other instruments or if there is no event that validates the value. This makes the
calibration process plays an important role to maintain data quality. BMKG always calibrates the equipment every 6 months, but between the 6 months it does not rule out the possibility of potential problems in measuring values, especially for electronic or digital equipment. The eligibility conditions for meteorological instruments adhere to the regulations of the World Meteorological Organization (WMO) CIMO no.8 of 2014, where the measurement tolerances are: (1) temperature: maximum of 0.3°C, (2) humidity: maximum 3%, (3) air pressure: maximum of 0.3 hPa, (4) maximum wind speed of 0.5 m/s, (4) wind direction: maximum 5°, (5) rainfall: maximum 5%, (6) sun radiation: maximum 5%.

To make the control of sensor conditions easier, especially temperature, pressure, humidity, and rainfall sensors, we need a system that can monitor and detect when problems occur with these sensors. The correlation among weather parameters is the key to controlling the sensor conditions to be trained and tested using the ANN backpropagation method. This ANN system design works by learning the correlation and pattern of each sensor data during the training phase. In the testing phase, the condition of the test data will be predicted. If any sensor outputs a value that is unusual or different from the pattern studied by ANN, the system will give a warning indicating sensor failure. With better quality weather observation data, it will improve the quality of providing weather information, so that the use of weather information becomes more accurate and useful. In a study [2] entitled Machine Learning-Based Calibration of Low-Cost Air Temperature Sensors Using Environmental Data, and research [3] entitled Temperature error correction based on BP neural network in meteorological wireless sensor network, they tried to approach a calibration using software and models, but only limited to the temperature sensor. In this study, we try to do the same approach, but for more sensors. The next approach to sensor error detection is studied based on the correlation pattern among sensors, this was done in a study [4] entitled Soft Sensors for Instrument Fault Accommodation in Semiactive Motorcycle Suspension Systems. The detection of a condition in classification has been carried out in a research conducted by [5] entitled Intelligent Multi-Sensor Control Device for Recognition of Gas-Air Mixture Samples with the Use of Artificial Neural Networks, which classifies and detection odors with electronic noses using ANN. From the researches above, the ANN model has good results, so this paper will try to apply the ANN-BP method for an early detection approach for error indication of more than one sensor on AWS in a result that is classified as error or normal.

2. Method

![Figure1. Schematic of the AWS sensor condition early detection system.](image-url)
The design of the early detection system begins with the design of the ANN backpropagation model, the model is built with pattern recognition in training data on observations of weather parameters, temperature, humidity, pressure, and rain at Tanjung Priok Maritime Station for 4 years, from 2017 until 2020. The data training is carried out using Rstudio software. The composition of data for training is 80% data. The training is carried out so that the network can recognize the patterns generated from the input and output pairs. The data input consists of weather parameters, temperature, humidity, pressure, and rain, and the output is a label of the sensor's condition, normal or an error indication.

After the model produces the best accuracy in training and testing data, then the ANN model is used to estimate and to detect the condition of the AWS sensors, especially pressure, temperature, humidity, and rain sensors. The details of the research steps are:

2.1. Preprocessing data
Before the data was processed using ANN, the data were compiled and conditioned, with a composition of ± 50% actual data and ± 50% in the form of synthetic data. The synthetic data mean the actual data that has been added and subtracted in value according to WMO CIMO regulation no.8 2014 to obtain data in the form of damaged sensor label values.

2.2. ANN Design
ANN design is done by determining the amount of input data used in training, the number of hidden layers used and the number of outputs desired. The data used as input are temperature, humidity, pressure, and rain observation data at the Tanjung Priok Maritime Meteorological Station from 2017 to 2020, with details of the network architecture as follows:
Figure 4. ANN architecture Air pressure.

Figure 5. ANN rainfall architecture.

Figure 6. Research algorithm.
2.3. **Pattern Recognition (training).**

In the training process, the maritime meteorological station’s conventional weather observation data for 4 years are arranged into 2 output conditions, namely: (1) Output conditions "sensor in normal conditions", where all input values are the original values of weather observations for the past 4 years. (2) The output condition is "problematic sensor", where all input values are added and also subtracted from the value that exceeds the tolerance limits of the CIMO World Meteorological Organization (WMO) No. 8 of 2014, where the measurement tolerance is as follows: (a) temperature: maximum 0.3°C, (b) Humidity: maximum 3%, (c) Air pressure: maximum of 0.3 hPa , (d) Rainfall: maximum 5%.

The input and output data during the training are in the form of: (1) Input temperature, humidity, and pressure data, the output temperature sensor label indication is damaged or normal, (2) Input temperature, humidity, and pressure data, the output humidity sensor label indicates damaged or normal. (3) Temperature, humidity, and pressure data input, the output pressure indication is damaged or normal. (4) Temperature, humidity, and rain data input, output rain label indication of damage or normal in all rain categories except 1-3mm rain which has additional pressure data input.

2.4. **Testing and estimation**

Data testing is carried out aimed to determine whether the network can recognize patterns of training data from the input data provided. If the resulting error value has reached the target, the resulting output can be used as estimation data. The model validation value is obtained from the accuracy coefficient with the following value interpretation:

| Table 1. The relation between accuracy coefficient and interpretation [6] |
|---------------------------------------------------------------|
| Coefficient Interval | Relation Level |
|----------------------|----------------|
| 00.0 - 20 %          | Very low       |
| 20.1 - 40.0 %        | Low            |
| 40.1 - 60.0 %        | Moderate       |
| 60.1 - 80.0 %        | High           |
| 80.1 - 100 %         | Very high      |

The estimation is done after the pattern recognition process is carried out by the network when the training is complete and the model has been tested with good accuracy values. Input data consist of AWS Tanjung Priok’s temperature, humidity, pressure, and rain data and the output is a classification of sensor conditions (a) Normal, or (b) The temperature sensor is indicated as damaged, or (c) The humidity sensor is indicated as damaged, or (d) The pressure sensor is indicated as damaged, or (e) The rain sensor is indicated to be damaged.

3. **Result and Discussion**

3.1. **Test result**

3.1.1. **Temperature Sensor.**

After the data training was carried out, then testing was carried out with the remaining 20% of the data, with the target data being the previously known sensor conditions. In the testing temperature sensor conditions, obtained a very high accuracy value is 99%, false negative (prediction is “normal”, which it should “error indication”) value is 0.6% and false positive (prediction is “error indication”, which it should “normal”) value is 0.7%, with the graph of the independent variable contribution as follows:
The figure above shows the intensity of the contribution of the independent variable in training and testing for the output temperature sensor condition label, where the highest contribution is the value of the temperature sensor itself.

3.1.2. Humidity Sensors.
After the data training was carried out, testing was carried out with the remaining 20% of the data, with the target data being in the form of previously known conditions. In testing the temperature sensor conditions, a very high accuracy value was obtained 94.6%, false negative (prediction is “normal”, which it should “error indication”) value is 1.96%, and false positive (prediction is “error indication”, which it should “normal”) of 4.17%, with a graph of the independent variable contribution as follows:

The figure above shows the intensity of the contribution of the independent variable in training and testing for the output humidity sensor condition label, where the highest contribution is the value of the humidity sensor itself.

3.1.3. Pressure Sensor.
After the data training was carried out, testing was carried out with the remaining 20% of the data, with the target data being in the form of previously known conditions. In testing the temperature sensor conditions, obtained a very high accuracy value of 100%, false negative (prediction is “normal”, which it should “error indication”) value is 0% and false positive (prediction is “error
indication”, which it should “normal”) value is 0%, with a contribution graph independent variable as follows:

![Figure 9](image)

**Figure 9.** Contribution of the independent variable, pressure sensor label output.

The figure above shows the intensity of the contribution of the independent variable in training and testing for the output Pressure sensor condition label, where the highest contribution is the value of the Pressure sensor itself.

3.1.4. Rain Sensor.
After the data training was carried out, testing was carried out with the remaining 20% of the data, with the target data being in the form of previously known conditions. In testing the temperature sensor conditions, obtained a very high accuracy value on average of 82%, an average false negative (prediction is “normal”, which it should “error indication”) value is 0.47% and false positive (prediction is “error indication”, which it should “normal”) value is 16.49% with details: (a) Rainfall 1-3 mm, the test accuracy is 77%, false-negative 1.89%, and false-positive 20.75%. (b) Rainfall 3-20 mm testing accuracy is 82%, false-negative 0%, and false-positive 18.03%. (c) Rainfall 20-50 mm has 82% accuracy testing, 0% false-negative and 17.65% false-positive. (d) Rainfall above 50 mm has 91% accuracy testing, false-negative 0%, and false-positive 9.52%. With the graph of the independent variable contribution as follows:

![Figure 10](image)

**Figure 10.** Contribution of the independent variable, 1-3mm, and 3-20mm rain sensor label output.
The Figure above shows the intensity of the contribution of the independent variable in training and testing for the output rain sensor condition label, where the highest contribution is the value of the rain sensor itself.

3.2. Estimation Results

After the training and data testing process, based on the high accuracy results above, the sensor condition estimation process is carried out. The data to be estimated is the latest AWS Tanjung Priok data on October 16 - 18, 2020 with the following results:

Table 2. The estimation results of the AWS Tanjung Priok sensor condition label.

| Data num. | Date       | Time (UTC) | Temp | Humid | Press   | Rain | Estimated of sensor condition labels |
|-----------|------------|------------|------|-------|---------|------|--------------------------------------|
| 1         | 16/10/2020 | 00:00:00   | 28.5 | 71.9  | 1009.5  | 0.00 | normal                              |
| 2         | 16/10/2020 | 03:00:00   | 31   | 66.7  | 1009.1  | 0.00 | normal                              |
| 3         | 16/10/2020 | 06:00:00   | 32.8 | 61.7  | 1006.1  | 0.00 | normal                              |
| 4         | 16/10/2020 | 09:00:00   | 30.6 | 68.5  | 1005.1  | 0.00 | normal                              |
| 5         | 16/10/2020 | 12:00:00   | 25.3 | 88.2  | 1008.2  | 51.60| normal                              |
| 6         | 16/10/2020 | 15:00:00   | 26.4 | 84.5  | 1008.8  | 1.2  | normal                              |
| 7         | 16/10/2020 | 18:00:00   | 27.2 | 78.3  | 1007.4  | 0.00 | normal                              |
| 8         | 16/10/2020 | 21:00:00   | 26.9 | 82.9  | 1006.9  | 0.00 | normal                              |
| 9         | 17/10/2020 | 00:00:00   | 27.2 | 78.7  | 1008.7  | 0.00 | normal                              |
| 10        | 17/10/2020 | 03:00:00   | 30.3 | 67.5  | 1009.5  | 0.00 | normal                              |
| 11        | 17/10/2020 | 06:00:00   | 31.5 | 65.2  | 1007.3  | 0.00 | normal                              |
| 12        | 17/10/2020 | 09:00:00   | 31.3 | 60.3  | 1005.8  | 0.00 | normal                              |
| 13        | 17/10/2020 | 12:00:00   | 29.7 | 69.9  | 1008.3  | 0.00 | normal                              |
| 14        | 17/10/2020 | 15:00:00   | 28.6 | 75.1  | 1009.2  | 0.00 | normal                              |
| 15        | 17/10/2020 | 18:00:00   | 28.2 | 74    | 1007.8  | 0.00 | Error Indication for Pressure sensor |
| 16        | 17/10/2020 | 21:00:00   | 27.6 | 74.2  | 1007.9  | 0.00 | Error Indication for Pressure sensor |
| 17        | 18/10/2020 | 00:00:00   | 27.9 | 74.8  | 1009.7  | 0.00 | normal                              |
| 18        | 18/10/2020 | 03:00:00   | 31   | 67.5  | 1009.9  | 0.00 | normal                              |
| 19        | 18/10/2020 | 06:00:00   | 31   | 67.4  | 1007.5  | 0.00 | normal                              |
| 20        | 18/10/2020 | 09:00:00   | 30.7 | 67    | 1006.3  | 0.00 | normal                              |
Based on the model obtained from training and tested with previous data, and used to estimate the AWS Tanjung Priok sensor data for 16-18 October 2020, it was found that almost all were in normal condition. 2 conditions indicated that the pressure sensor had an error, on October 17 at 18.00 and 21.00 UTC, which can be seen at those 2 times the pressure value suddenly decreased significantly, but other weather parameters were still in conditions not much different from the previous time.

4. Conclusion
The sensor condition, especially temperature, humidity, pressure, and rain on AWS Tanjung Priok can be estimated using the ANN backpropagation method, where the accuracy results between the model output and the target during training and testing show very high values. Based on this model, the estimation results of the AWS Tanjung Priok sensor conditions on 16-18 October 2020 are almost all in normal conditions, 2 conditions indicated that the pressure sensor had an error, on October 17 at 18.00 and 21.00 UTC, this can be seen at the 2 times the pressure value decreased significantly, but other weather parameters are still not much different from the previous time. Based on the results of this estimation, it is hoped that it can serve as a warning to the nearest Maritime Meteorological Station so that checks can be carried out as soon as possible and if damage occurs, replacement or repair of sensor hardware can be carried out so that the quality of AWS data can always be maintained.

Acknowledgment
This research was supported by the grant of PITTA (Publikasi Internasional Terindeks Untuk Tugas Akhir Mahasiswa) of Universitas Indonesia under the contract number NKB-1005/UN2.RST/HKP.05.00/2020. We would like to acknowledge the Indonesian Agency for Meteorology Climatology and Geophysics for supporting data and facilities.

References
[1] Dahuri R 2004 Pengelolaan Sumber Daya Wilayah Pesisir dan Lautan Secara Terpadu, Edisi Revisi (Jakarta: Pradnya Paramita)
[2] Yamamoto K, Togami T, Yamaguchi N, Ninomiya S 2017 Machine Learning-Based Calibration of Low-Cost Air Temperature Sensors Using Environmental Data. Sensors 17(6) 1290
[3] Wang B 2017 Temperature error correction based on BP neural network in meteorological wireless sensor network. Int. J. Sensor Networks 23(4)
[4] Capriglione D, et al 2020 Soft Sensors for Instrument Fault Accommodation in Semiactive Motorcycle Suspension Systems IEEE transactions on instrumentation and measurement 69 (5)
[5] Kulagin V P, et al 2017 Intelligent Multi-Sensor Control Device for Recognition of Gas-Air Mixture Samples with the Use of Artificial Neural Networks IEEE
[6] Sugiyono 2008 Metode Penelitian Kunatifatif Kualitatif dan R&D ( Bandung: Alfabeta)