UMT PRESS
Journal of Mathematical Sciences and Informatics
Journal Homepage: https://journal.umt.edu.my/index.php/jmsi
eISSN: 2948-3697
DOI: http://doi.org/10.46754/jmsi.2026.06.001

VISUALISATION TECHNIQUES TO DETECT ABNORMAL SEA SURFACE TEMPERATURE DATA IN THE INDIAN OCEAN

SHARIFAH SAKINAH SYED ABD MUTALIB, NORIZAN MOHAMED*, MAHARANI ABU BAKAR AND WAN NURAINI FAHANA WAN NASIR

Faculty of Computer Science and Mathematics, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia.

**Corresponding author: norizan@umt.edu.my*

ARTICLE INFO

Article History:
Received: 13 May 2025
Revised: 28 September 2025
Accepted: 25 November 2025
Published: 15 June 2026

Keywords:
Sea surface temperature, climatic factors, abnormal data, Chernoff faces, distance-distance plot.

ABSTRACT

Sea Surface Temperature (SST) is the temperature of the water near an ocean's surface. It plays a critical role in the interaction between the Earth's surface and its atmosphere. The factors affecting sea surface temperature are crucial to understanding climate change, while SST itself is influenced by climatic factors such as humidity, air temperature, wind speed, and radiation. In practice, multivariate, anomalous abnormal data cannot be avoided. SST data is no exception to these data inconsistencies and may have multivariate or abnormal data. Therefore, a technique which visualises the abnormal data is crucial. In this study, Chernoff faces, and distance-distance plots are used as visualisation techniques to detect abnormal sea surface temperature data. Chernoff faces is a graphical representation of multivariate data, while a distance-distance plot is a graphical representation to reveal abnormal data. Based on both analyses, the abnormal data are clearly visible of climate factors of SST. Additionally, an Artificial Neural Network (ANN) with an autoencoder was used to detect outliers by applying a threshold to reconstruction errors. The Artificial Neural Network (ANN) model with the autoencoder and Chernoff faces method are effective techniques for detecting outliers in SST datasets.

2020 Mathematics Subject Classification:

Introduction

Climate change is a naturally occurring process that can influence life on Earth in both direct and indirect ways [1]. Its potential future effects include heavier rainfall, stronger tropical storms, extended forest fire seasons, and droughts in certain regions. One of the key elements that contributes to climate change is Sea Surface Temperature (SST). SST data serves as a valuable indicator for identifying various oceanic phenomena such as monsoons, the Indian Ocean Dipole (IOD), and the *El Nino El Nino* Southern Oscillation (ENSO), which include both the *La Nina* and the *El Nino* events [2]. These phenomena play a significant role in shaping SST patterns across the Pacific and Indian Oceans [3]. In tropical regions, SST typically range

between 27°C and 29°C, while in subtropical areas, it falls between 15°C and 20°C. As depth increases, seawater temperature gradually decreases, stabilising at between 2°C and 4°C beyond depths of 1,000 m.

Early fluctuations in Sea Surface Temperature (SST) are comparable in scale to historical warming trends, but their accuracy is influenced by differences in measurement techniques, instrument calibration, data processing and storage protocols. These inconsistencies introduce biases that limit the reliability of early SST data in climate research. However, adjusting SST values within ocean models has enabled atmospheric simulations to reflect historical variations more accurately in North Atlantic hurricanes, highlighting the potential for improved storm forecasting [4]. In Indonesia, climatic conditions are strongly shaped by geographic location, particularly SST patterns in the Indian Ocean. These patterns are often referred to as the Indian Ocean Dipole (IOD) or Dipole Mode Index (DMI) [2]. SST can be analysed using various climate-related factors, including air temperature, precipitation, wind speed, solar radiation intensity, and humidity.

Patil and Ravichandran [5] proposed a hybrid approach for predicting Sea Surface Temperature (SST) by combining numerical modelling with neural network techniques. Their study focused on six locations within the Indian Ocean and examined SST data across three-time scales: Daily, weekly, and monthly. At each time interval, they calculated the discrepancy between numerical estimates and actual SST measurements, generating an error time series. This error series was then forecast using neural networks, and the predicted errors were added to the numerical estimates to produce SST predictions for five future time steps. The accuracy of this method was evaluated using various statistical error metrics, which demonstrated its visibility and outstanding performance. Among the neural models tested, the wavelet neural network based on the 'Meyer wavelet with discrete approximation' yielded the most satisfactory results compared to other wavelet types. Overall, the combination of numerical and neural methods yielded highly accurate SST forecasts, significantly enhancing the correlation between predicted and observed values, with minimal exceptions.

Yu *et al.* [6] proposed a deep learning-based framework for forecasting sea surface temperatures (SST) and introduced the Data-driven Graph Convolution (DGC) model. This architecture combines multiple Gated Recurrent Unit (GRU) layers with a Convolutional Neural Network (CNN) and a fully connected layer, enabling it to effectively capture the complex and non-linear characteristics of SST time series data. The researchers used Optimum Interpolation Sea Surface Temperature (OISST) data and conducted experiments using randomly selected points from the East China Sea and the Yellow Sea. On the Yellow Sea dataset, the model achieved strong predictive performance, with a Root mean square error (RMSE) of 0.3637°C, Root Mean Squared Prediction Error (RMSPE) of 1.7382%, Mean Absolute Percentage Error (MAPE) of 1.2915%, and Mean Absolute Error (MAE) of 0.2673°C. These metrics indicate that the DGC network outperforms traditional approaches and other existing models. Notably, the DGC network addresses a key limitation of standard GRU models, which tend to produce higher errors near peak SST values. Overall, the DGC model demonstrates superior accuracy, enhanced performance, and greater stability with regard to SST prediction.

In a recent study, Cheng *et al.* [7] applied a back-propagation neural network (BPNN) technique to estimate subsurface temperatures in the North Pacific Ocean (NPO) at 16 depth

levels ranging from 30 to 1,000 m. The model used satellite-derived sea surface parameters — sea surface height (SSH), surface height temperature (SST), surface height salinity (SSS), and surface height wind (SSW) —as input variables, with the corresponding subsurface temperatures serving as outputs. To enhance prediction accuracy, the researchers introduced Sea Surface Velocity (SSV) as an additional input. Their results showed that the BPNN model could effectively estimate subsurface temperatures, achieving mean square errors of 0.868 with four inputs and 0.802 when SSV was included. The average coefficients of determination were 0.952 and 0.967, respectively. They concluded that input of the SSV in addition to the SSH, SST, SSS and SSW, therefore, had a positive impact on the BPNN model and helped to improve the accuracy of the estimations.

Miftahuddin *et al.* [8] made use of a Generalised Additive Model (GAM) and link function for the detection of missing SST data in the Indian Ocean. After the amputation procedure, the missing values of the climate data were predicted using a Generalised Additive Model (GAM) while a link function tied the model's response variables to the model's prediction variables. Since the SST and all of the independent variables in this analysis lacked data, the researchers used the GAM to predict SST in the Indian Ocean using the median as a measurement for amputation. GAM had an Akaike Information Criterion (AIC) of 20.48 before amputation. This study has missing data, as evidenced by the lowest AIC score and the time series plot of SST and all independent variables. GAM's AIC was 1,500.88 when the amputation was applied. As there was no missing data after amputation, a high AIC value is meaningless for detecting missing data. A comparison of SST data before and after amputation revealed a smoother model fit post-processing. The study concluded that using the median was the most effective strategy for handling missing values, and that GAM combined with amputation provided a reliable method for reconstructing incomplete SST datasets.

Miftahuddin *et al.* [8] also explored rainfall prediction using ARIMA mixed models, recognising that rainfall is influenced by sea surface temperature. Their goal was to identify the most effective Autoregressive Integrated Moving Average (ARIMA) configurations when combined with Autoregressive Conditional Heteroskedasticity (ARCH) and Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models. The study analysed rainfall data and the number of rainy days in the Aceh Barat district of Indonesia from January 2008 to December 2017. Their findings indicated that the ARIMA(2,1,0)-GARCH(1,3) model provided the best results for forecasting rainfall, while the ARIMA (2,0,2) model was the most suitable for predicting the number of rainy days. These results highlighted the potential of combining ARIMA with GARCH techniques to improve the accuracy of climate-related time series predictions.

In 2022, Miftahuddin *et al.* [9] examined the factors influencing air temperature in the Indian Ocean, particularly near the Aceh region, using three statistical models: The Linear Model (LM), the Generalised Linear Model (GLM), and the Generalised Additive Model (GAM). Their analysis identified four key variables that significantly affected air temperature at the 5% significance level: Rainfall (X1), relative humidity (X2), sea surface temperature (X3), and wind speed (X4). All three models were validated through standard diagnostic tests, including checks for normality, multicollinearity, heteroscedasticity, and autocorrelation, confirming their suitability for analysis. Among them, the GAM model, which used adaptive smoothers produced the lowest AIC values, indicating that it provided the best model fit.

However, for visualisation purposes, the GAM model with a P-spline basis was found to be more effective. The study concluded that the GAM model with a P-spline basis was the most appropriate and reliable for modelling air temperature in the region.

Data visualisation is a powerful tool for representing data in a graphical format, making it easy to understand with minimal effort [10, 11]. Effective data visualisations enable efficient examination of large datasets and provide insights that allow for inferences based on observed relationships within the data. For this reason, visualisations are widely used in data mining and exploration, information retrieval, and the analysis of tactical and strategic intelligence [11].

Data is typically represented as vectors in a multidimensional space, with each dimension corresponding to a distinct attribute of the data. A one-dimensional dataset can be shown on a single axis as a set of points, a two-dimensional dataset as a planar curve, and a three-dimensional dataset as a spatial surface. However, datasets with more than three dimensions cannot be conventionally represented as planar curves or spatial figures. Despite this, visualising this type of dataset is crucial for data analysis because it provides a direct and intuitive view of the data. Several approaches exist for visualising datasets with more than three dimensions, one of which is using glyphs. Glyphs are graphical shapes or symbols that convey meaning. Among these, Chernoff's faces are the most complex, representing data through symbolic portrait figures [10].

Chernoff, an American statistician, introduced the idea of using faces to represent multidimensional data in 1973, motivated by the fact that facial features and emotions are impressive and easily discernible by the human viewer [10, 12]. This method displays multivariate data in the form of a human face, representing various numerical or binary variables in a dataset through the size and curvature of the face, the position and size of the eyes, the length of the nose, the orientation and curvature of the mouth, and so on. Each data point corresponds to a different Chernoff face. If two samples are almost the same, their corresponding Chernoff faces will be very similar; if two data points are very different, their corresponding Chernoff faces will look very different [10].

Moriss *et al.* [13] conducted a user study to test the effectiveness and pre-attentiveness of various features of Chernoff faces. Their study revealed that perceiving specific features such as eye size, individual faces, eyebrow slant, and the combination of eyebrow slant with eye size is a serial process. Additionally, they found that with longer viewing times, eye size and eyebrow slant were the most accurately perceived features. Hence, they concluded that these initial results suggest Chernoff faces may not offer a significant advantage over other iconic visualisation techniques for multidimensional information visualisation.

Gerhardt [14] applied the Chernoff faces approach to represent a multi- Key Performance Indicator (KPI) system as a unit. His study showed that when differences between compared companies are very small, the Chernoff faces can appear almost identical, which can be problematic. He limited the Chernoff faces to seven metrics due to the prominence of the most salient facial features but suggested that incorporating less salient facial features could address this limitation. By presenting KPIs in the form of Chernoff faces, the biological limitation of humans to hold three to five KPIs in iconic memory can be bypassed. He concluded that dynamic Chernoff faces provide significant added value, as they allow more information to be perceived simultaneously by humans.

Wyatt [15] developed a modification of face diagrams called face charts, which are less reminiscent of actual human faces yet are still realistic compared to two traditional diagramming methods: The bar chart and the star plot. Although they used a small sample of test respondents, the study generated strongly suggestive evidence that face charts enable respondents to better understand ten-dimensional data than star plots or bar charts. He concluded that face charts are superior to other methods traditionally used by previous researchers to visualise complicated data.

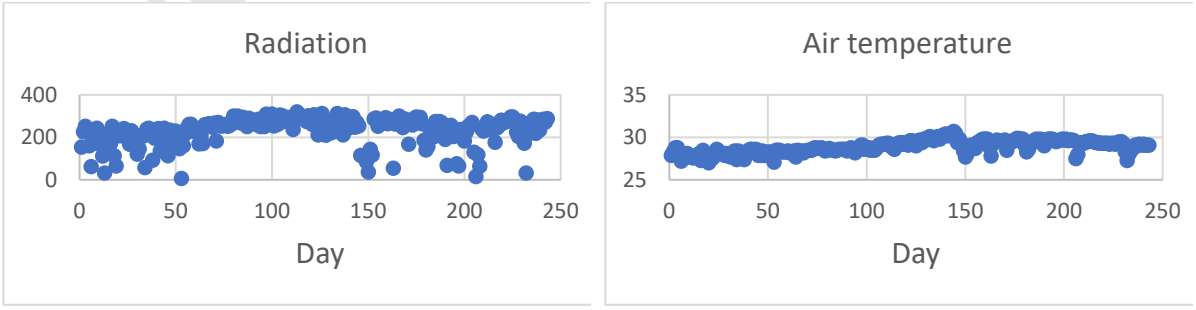
In this study, Chernoff faces is used as a visualisation method to identify abnormal SST data. Additionally, a distance-distance plot and Artificial Neural Networks are employed to support the Chernoff faces' findings in detecting abnormalities within the SST dataset.

Climate change which is influenced by factors like Sea Surface Temperature (SST), significantly impacts global and regional weather patterns such as monsoons, *El Nino*, and the Indian Ocean Dipole. SST varies by location and oceanic depth and plays a key role in atmospheric phenomena, including forecasting hurricanes. Various methods have been developed to predict SST, including numerical models combined with neural networks, deep learning architectures like GRU and CNN, and statistical models such as GAM and ARIMA. Additionally, data visualisation techniques like Chernoff faces are used to represent complex, multidimensional climate data visually, aiding in the identification of anomalies. Detecting outliers is a critical step in data analysis, statistics, and machine learning because anomalous data points can significantly distort results and lead to misleading conclusions. This study employs Chernoff faces together with Artificial Neural Networks to detect abnormal SST data, providing both visual and computational analysis tools.

Methodology

Sea Surface Temperature Data

The data for this study is taken from the sample of daily data of the ocean dataset from 1 December 2018 to 31 July 2019. The data consists of 243 observations. The data mainly explores the relationships between SST and other variables like radiation, air temperature, wind speed, rainfall, and humidity. Figure 1 shows five independent variables and one dependent variable. The description of each variable in the dataset is shown in Table 1.



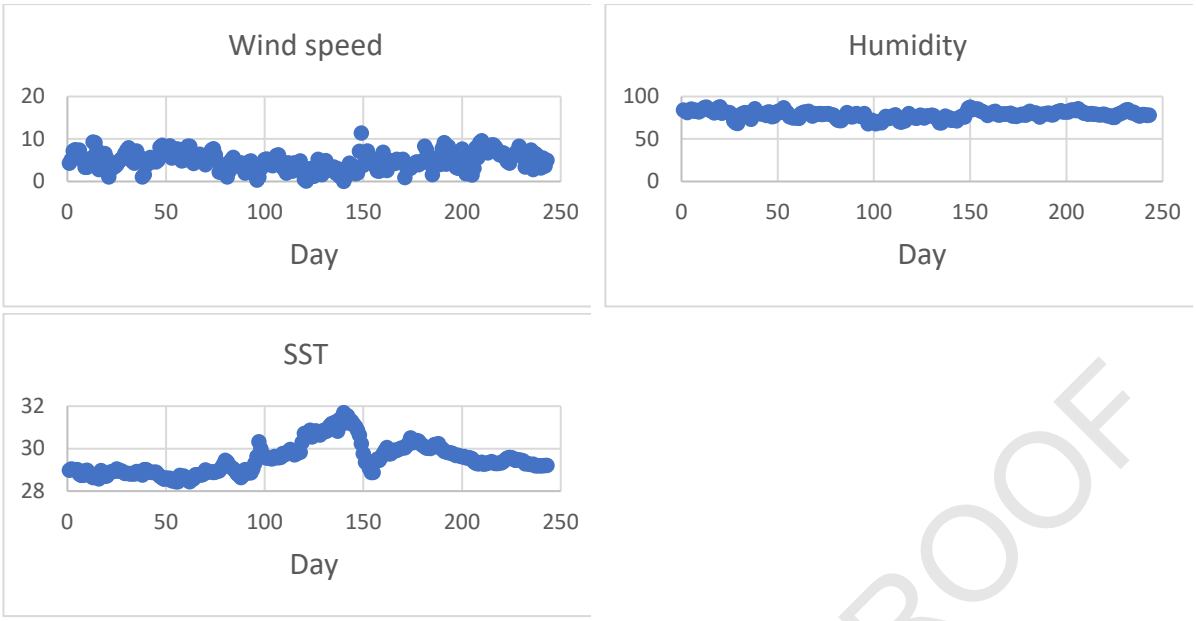


Figure 1: Ocean dataset

Table 1: Description of variables

Variables	Description
x_1	Radiation, W/m^2
x_2	Air temperature, ^0C
x_3	Wind speed, m/s
x_4	Humidity,
y	SST, ^0C

Data Visualisation Techniques

Chernoff faces, distance-distance plots, and ANN are used as visualisation techniques in this study to detect abnormal sea surface temperature data. Chernoff faces was developed by Chernoff [16] to represent multivariate data. The *p*-dimensional observations are presented as two-dimensional faces whose characteristics are determined by the measurement of the *p* variables. Chernoff faces can handle up to 18 variables [16]. Figure 2 shows the example of Chernoff faces for 3-dimensional data. As can be seen from Figure 2, each observation depicts different faces. The presentation of the faces also makes it easy to detect any abnormalities present in the data.

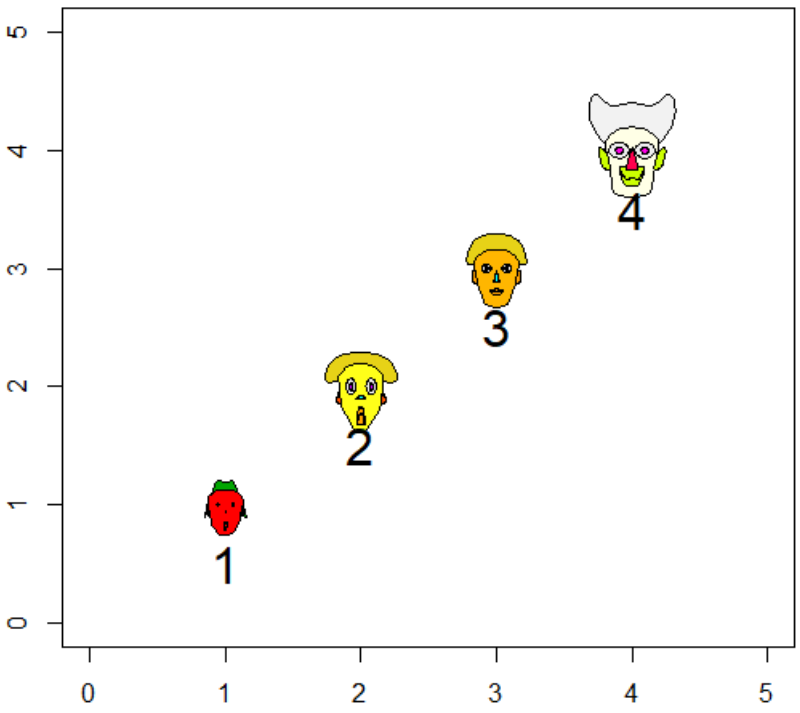


Figure 2: Examples of Chernoff faces

The visualisations used to detect abnormal observations is enhanced further using a Distance-Distance (DD) plot. Mahalanobis Distance (MD) is plotted versus Robust Distance (RD) in the DD plot. MD has been used as one of the methods to detect abnormal observations for multivariate data and is given in Equation (1) [17],

$$d_i(\mathbf{x}) = \sqrt{(\mathbf{x}_i - \bar{\mathbf{x}})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})}, \quad i = 1, 2, \ldots, n \qquad (1)$$

However, MD has been affected by masking and swamping effects in detecting abnormal observations. Thus, in this study, the Fast Minimum Covariance Determinant (FMCD) estimator is used as the robust estimator of the mean and covariance matrix. RD is given in Equation (2) as,

$$d_i(\mathbf{x}) = \sqrt{(\mathbf{x}_i - \bar{\mathbf{x}}_{FMCD})' \boldsymbol{\Sigma}_{FMCD}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}_{FMCD})}, \quad i = 1, 2, \ldots, n \qquad (2)$$

The cut-off value used to detect abnormal observations is $\sqrt{\chi^2_{p,0.975}}$. If $d_i(\mathbf{x}) > \sqrt{\chi^2_{p,0.975}}$, $\mathbf{x}_i$ is an outlier [18].

The next approach for the visualisation abnormal sea surface temperature data is using Artificial Neural Network (ANN). ANN has been developed as a generalisation of biological, mathematical models of the nervous system. The first wave of interest in neural networks (also known as connectionist models or parallel distributed processing) emerged after the introduction of simplified neurons [19]. The primary processing elements of neural networks are called artificial neurons, or just neurons or nodes. Inside a simplified mathematical model of neurons, its effects in the synapse are represented by a connection weight modulating the effect of the associated input signal, and the transfer function represents non-linear features

exhibited by neurons. Neuronal impulses were then calculated as the weighted sum of the input signals, which is varied by a transfer function [20]. Artificial learning ability neurons are achieved by adjusting the weight as appropriate to the selected learning algorithm.

The dataset includes four independent variables which are air temperature, radiation, wind speed and humidity. These independent variables will be used as the input for the model. An autoencoder was executed alongside ANN to detect the outliers. Figure 3 shows an example of the autoencoder model.

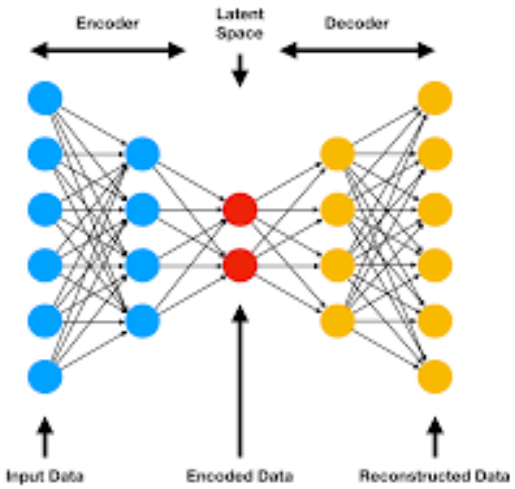


Figure 3: Autoencoder Model

The encoder and decoder procedures were applied in the autoencoder to learn a simplified version of the data. The encoding layers are filled with an input layer of four neurons which refer to our independent variables. The model's performance was optimised through training. Outliers can be identified by using a threshold on the reconstruction error. Python was the software environment that was used to create the autoencoder model, with the software's TensorFlow and Keras libraries making easier to build and train the models.

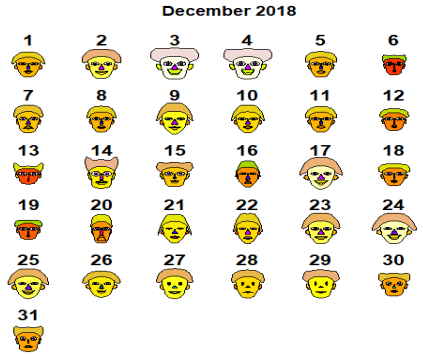
Results

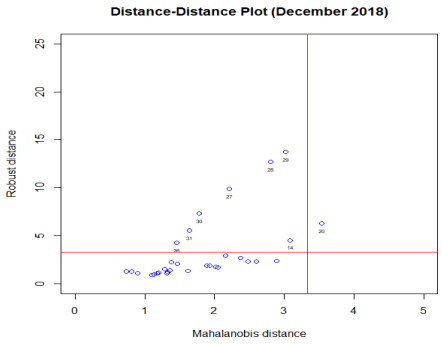
In this study, we employed multiple techniques to detect outliers in Sea Surface Temperature (SST) data, including Chernoff faces, Distance-Distance (DD) plots with Robust Distance (RD) and Mahalanobis Distance (MD), and an Artificial Neural Network (ANN) autoencoder model.

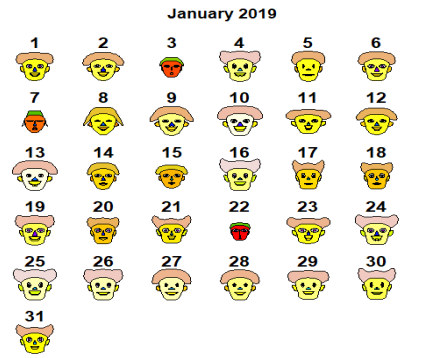
Figure 4 shows Chernoff faces and DD plots of SST data from December 2018 until July 2019. A comparison can be made when using both visualisation techniques. Observations that have distinct differences from other observations can be seen clearly in DD plots. However, for DD plots, RD can detect more abnormal observations than MD.

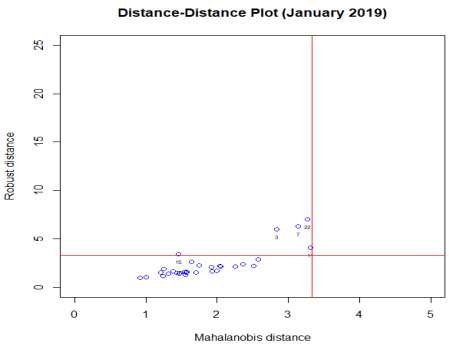
Observation 20 was identified as abnormal data for December 2018. The face features for this data have been seen to have different features than the rest of the observations. Additionally, RD found observations 14, and 26 to 31 to be abnormal. Similar results were observed in other months, where RD detected more abnormal data than MD. The RD findings are more in line with Chernoff faces than the MD findings.

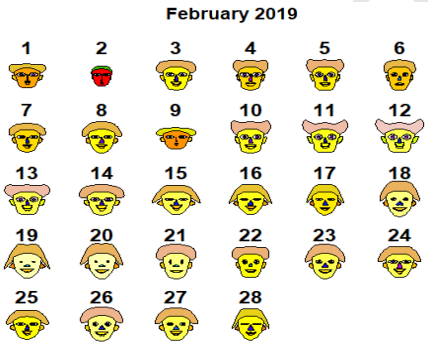
For January 2019, observations 3, 7, and 22 have different faces. The RD also found these observations to be anomalous. RD also found observations 11 and 15 to be abnormal. However, the MD failed to find these observations anomalous. The same results can be seen in March 2019 where the MD failed to detect the abnormal data although the faces of observations 1, and 4 to 6 as well as 11, 13, 21, and 31 differed from other observations.

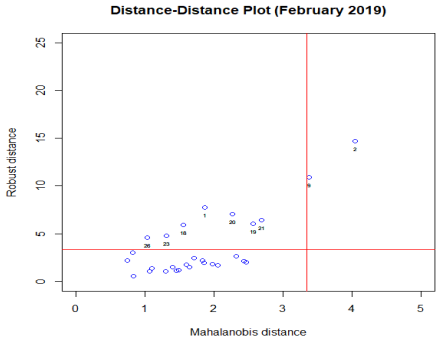


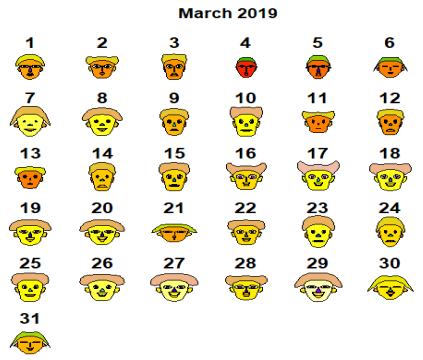


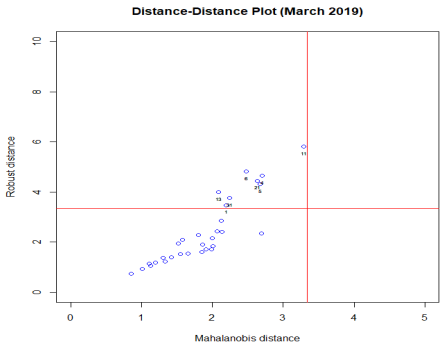












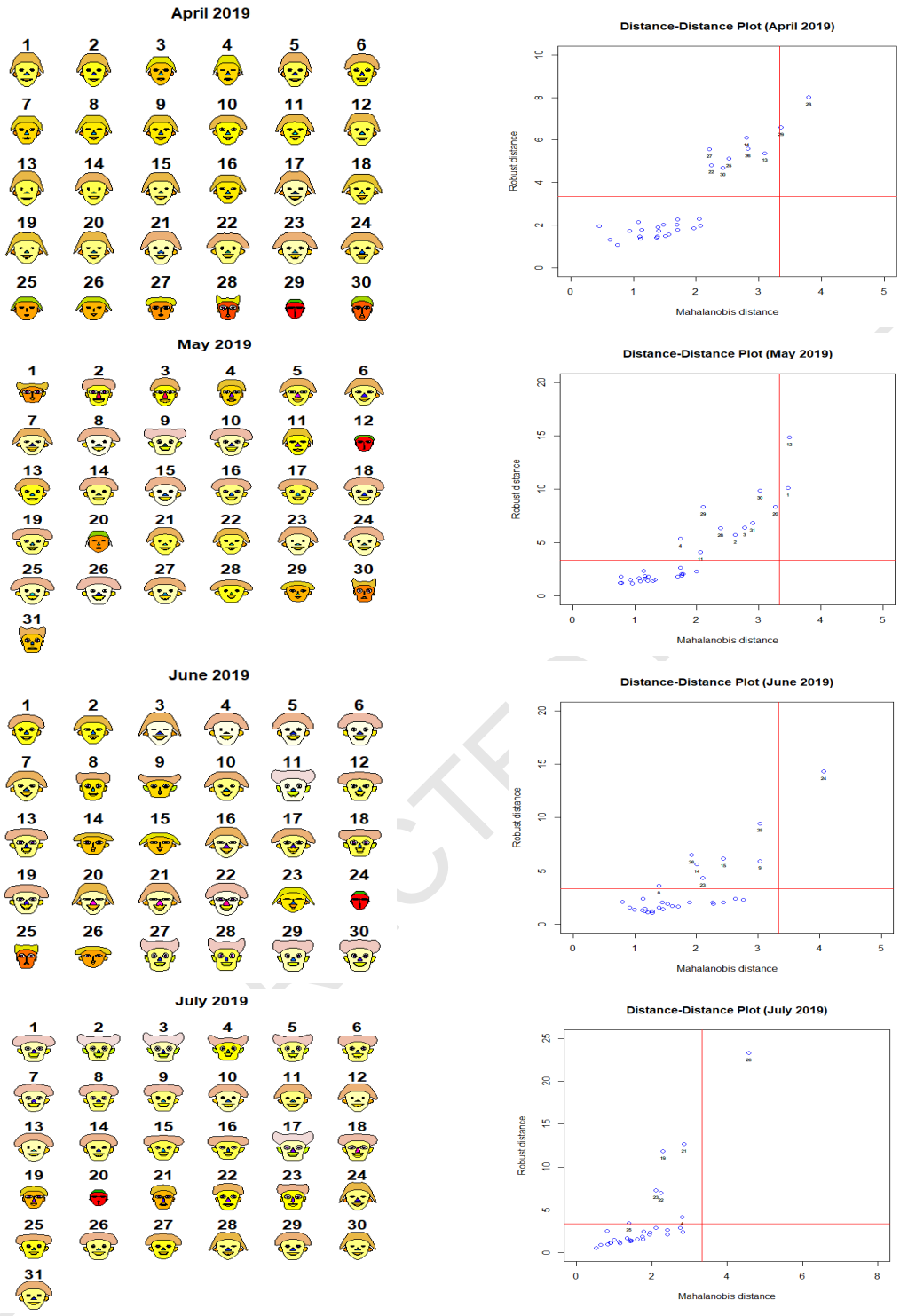
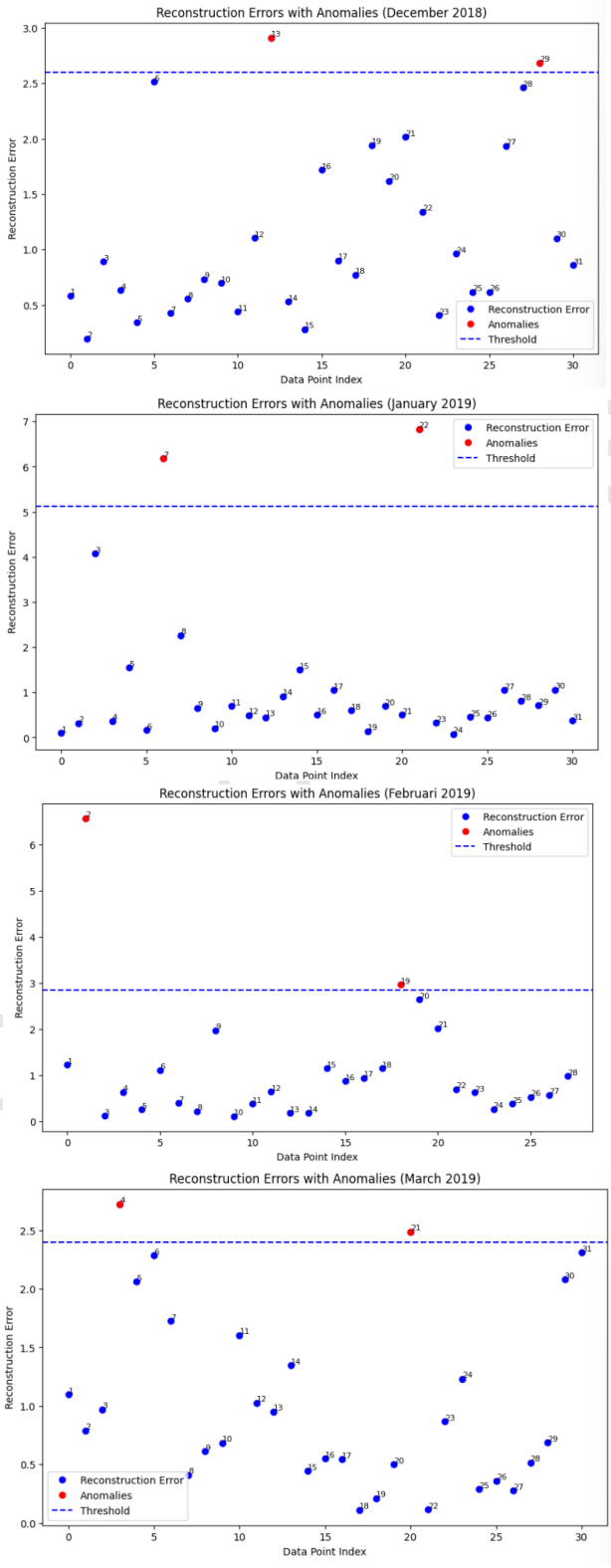


Figure 4: Visualisation of SST data using Chernoff faces and Mahalanobis distance

The ability of the Artificial Neural Network (ANN) to detect outliers with four independent variables are displayed in Figure 5. The autoencoder helped to detect outliers (anomalies) because it was designed to strip input data down to its essential features, then reconstruct (decode) the original input from this compressed representation.

The outliers were found by applying a threshold to the reconstruction error. If the reconstruction errors exceeded the threshold it was classified as an outlier. The graph of the

reconstruction errors was plotted to illustrate the outliers. The threshold line was set at the 95th percentile of the reconstruction errors and the data sets which lie above the threshold line are considered outliers.



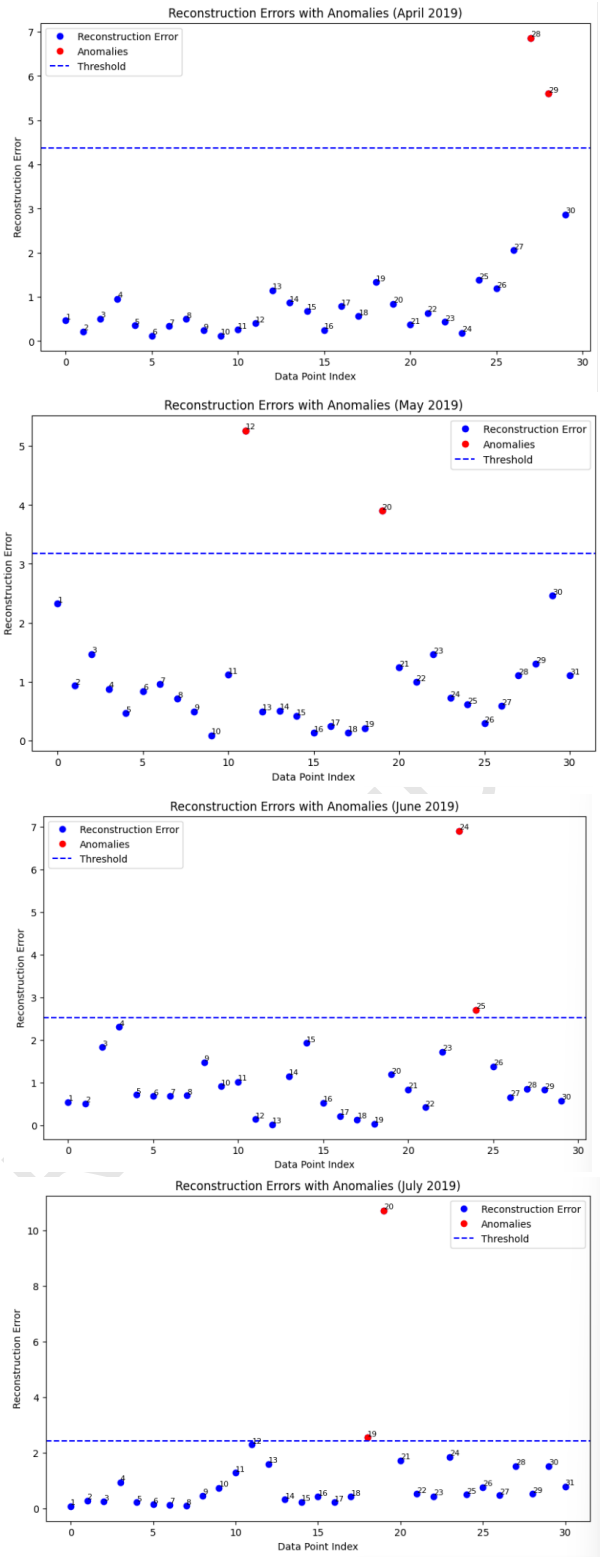


Figure 5: Visualisation of SST data using ANN

Discussion

In this study, an ANN Model with an autoencoder and the Chernoff faces method were used to detect outliers in a dataset using four independent variables. Distance-distance plots were used to visualise the outliers for Chernoff faces, where robust distance and Mahalanobis distance

were used as cut-off values. if the reconstruction errors exceeded the threshold for the ANN model it was classed as an outlier.

The Chernoff faces and DD plots served as complementary visualisation tools for identifying abnormal patterns in the SST dataset. Chernoff faces allowed for an intuitive, multidimensional representation of data points through facial features, making it easier to spot anomalies at a glance. DD plots, on the other hand, provided a quantitative visualisation of relationships between data points, helping to further highlight deviations or clusters of outliers. These visualisation techniques are primarily exploratory tools that assist in the initial detection of abnormal SST values. The ANN technique is then applied as a predictive and validation model to rigorously identify and confirm these anomalies detected visually. Unlike the visualisation methods, ANN provides a computational, data-driven approach to anomaly detection, offering objective classification and prediction capabilities.

With regard to the comparisons between these techniques, the visualisation methods (Chernoff faces and DD plots) and ANN serve different but complementary purposes in the present study. A visual comparison was also made by plotting the Chernoff faces and autoencoder results for the dataset. For December 2018, the robust distance and autoencoder method agreed that day 29 was an outlier. Mahalanobis distance identified day 20 as the outlier, and day 13 was also detected as the outlier by the autoencoder. Both the robust distance and autoencoder methods, agree that day 7 and 22 are an outlier for January 2019. Next, the autoencoder method was detected on days 2 and 19 as an outlier supported with Mahalanobis distance for day 2 and robust distance for day 19 in February 2019.

In March 2019, the autoencoder method and robust distance both agreed that days 4 and 21 were the outliers. However, in April 2019, all three methods identified the same outlier, specifically days 28 and 29. Day 12 and day 20 were found to be the outliers for both the robust and autoencoder methods in May 2019, while Mahalanobis was only detected for day 12. In June 2019, both methods, robust distance and autoencoder method notice those days 24 and 25 as the outliers; however, Mahalanobis distance just identified day 24 as the outliers. Lastly, in July 2019, days 19 and 20 were found to be the outliers for robust and autoencoder, but Mahalanobis distance only identified day 20 as an outlier. We then summarise in Table 2.

Table 2: Detection of outliers from December 2018 to July 2019

Month	Methods		
	Robust Distance	Mahalanobis Distance	ANN
December 2018	Day 29	Day 20	Days 29 and 13
January 2019	Days 7 and 22	-	Days 7 and 22
February 2019	Day 19	Day 2	Days 2 and 19
March 2019	Days 4 and 21	-	Days 4 and 21
April 2019	Days 28 and 29	Days 28 and 29	Days 28 and 29
May 2019	Days 12 and 20	Day 12	Days 12 and 20
June 2019	Days 24 and 25	Day 24	Days 24 and 25
July 2019	Days 19 and 20	Day 20	Days 19 and 20

Table 2 summarises the outliers detected by each method for monthly SST data from December 2018 to July 2019. Notably, there is frequent agreement between the robust distance and ANN autoencoder methods, indicating that both techniques identify similar abnormal observations. Mahalanobis distance, while still useful, detected fewer anomalies, which may reflect its sensitivity to multivariate normality assumptions and susceptibility to masking effects in the presence of outliers.

Conclusions

The article presents a study on the application of visualisation techniques to detect abnormal Sea Surface Temperature (SST) data in the Indian Ocean. We utilise Chernoff faces and distance-distance plots to identify anomalies in multivariate SST data. The study emphasises the significance of SST in climate change research and its interaction with various climatic factors.

The results indicate that the Robust Distance (RD) in distance-distance plots is more effective than the Mahalanobis Distance (MD) at detecting abnormal observations, with findings more aligned with Chernoff faces. Additionally, outliers can be detected using an Artificial Neural Network (ANN) with an autoencoder by applying a threshold to the reconstruction error.

Authors' Contributions

All of the authors contributed to the study conception and design. All authors read and approved the final manuscript. Introduction: Norizan Mohamed and Wan Nuraini Fahana Wan Nasir; Methodology: Sharifah Sakinah Syed Abd Mutalib and Maharani Abu Bakar; Result: Sharifah Sakinah Syed Abd Mutalib and Maharani Abu Bakar; Discussion and Conclusion: Norizan Mohamed, Wan Nuraini Fahana Wan Nasir, and Maharani Abu Bakar.

Acknowledgements

The authors would like to thank all reviewers for their comments and suggestions to improve this manuscript. This research was not funded by any grant.

Conflict of Interest Statement

The authors declare no conflict of interest.

References

[1] Abraham, A. (2005). Artificial neural networks. In Sydenham, P., & Thorn, R. (Eds.), *Handbook of measuring system design*. John Wiley & Sons. https://doi.org/10.1002/0471497398.mm421

[2] Chan, D. (2021). Combining statistical, physical, and historical evidence to improve historical sea-surface temperature records. *Harvard Data Science Review*, *3*(1). https://doi.org/10.1162/99608f92.edcee38f

[3] Cheng, H., Sun, L., & Li, J. (2021). Neural network approach to retrieving ocean subsurface temperatures from surface parameters observed by satellites. *Water*, *13*(3), Article 388. https://doi.org/10.3390/w13030388

[4] Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*, *68*(342), 361–368. https://doi.org/10.1080/01621459.1973.10482434

[5] Gerhardt, E. (2019). Visualization of multi key performance indicators by dynamic Chernoff faces. *CEUR Workshop Proceedings*, *2413*, 1–10.

[6] Hadi, A. S., Rahmatullah Imon, A. H. M. M., & Werner, M. (2009). Detection of outliers. *Wiley Interdisciplinary Reviews: Computational Statistics*, *1*, 57–70. https://doi.org/10.1002/wics.6

[7] IPCC. (2007). *Climate change 2007: The physical science basis*. IPCC.

[8] Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis* (5th ed.). Prentice Hall.

[9] Lee, M. D., Butavicius, M. A., & Reilly, R. E. (2003). Visualizations of binary data: A comparative evaluation. *International Journal of Human-Computer Studies*, *59*(5), 569–602. https://doi.org/10.1016/S1071-5819(03)00082-X

[10] Miftahuddin, M., Mohamed, N., Abu Bakar, M., Shima, N., Syamsuddin, F., & Setiawan, I. (2021). *Prediction of rainfall using ARIMA mixed models*.

[11] Miftahuddin, M., Sitanggang, A., Mohamed, N., & Abu Bakar, M. (2022). Modelling Indian Ocean air temperature using additive model. *Journal of Mathematics and Sciences with Informatics*, *2*, 23–36. https://doi.org/10.46754/jmsi.2022.06.003

[12] Morris, C., & Ebert, D. (2000). An experimental analysis of the effectiveness of features in Chernoff faces. In *Proceedings of SPIE*. https://doi.org/10.1117/12.384865

[13] Patil, K., Deo, M. C., & Ravichandran, M. (2016). Prediction of sea surface temperature by combining numerical and neural techniques. *Journal of Atmospheric and Oceanic Technology*, *33*(8). https://doi.org/10.1175/JTECH-D-15-0213.1

[14] Raciborski, R. (2009). Graphical representation of multivariate data using Chernoff faces. *Stata Journal*, *9*(3), 374–387. https://doi.org/10.1177/1536867X0900900302

[15] Saji, N., Goswami, B., Vinayachandran, P., & Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature*, *401*, 360–363. https://doi.org/10.1038/43854

[16] Schott, F. A., Xie, S.-P., & McCreary, J. P., Jr. (2009). Indian Ocean circulation and climate variability. *Reviews of Geophysics*, *47*(1). https://doi.org/10.1029/2007RG000245

[17] Sharma, A., & Dey, S. (2012). A document-level sentiment analysis approach using artificial neural network and sentiment lexicons. *ACM SIGAPP Applied Computing Review*, *12*(4), 67–75. https://doi.org/10.1145/2432546.2432552

[18] Song, R., Zhao, Z., & Wang, X. (2010). An application of the V-system to the clustering of Chernoff faces. *Computers & Graphics*, *34*(5), 529–536. https://doi.org/10.1016/j.cag.2010.06.003

[19] Wyatt, R. (2008). Face charts: A better method for visualizing complicated data. In *Proceedings of MCCSIS 2008 — IADIS Multi Conference on Computer Science and Information Systems: Computer Graphics and Visualization; Gaming; Designing for Engaging Experience and Social Interaction* (pp. 51–58). IADIS.

[20] Yu, X., Shi, S., Xu, L., Liu, Y., Miao, Q., & Sun, M. (2020). A novel method for sea surface temperature prediction based on deep learning. *Mathematical Problems in Engineering*, *2020*, Article 6387173. https://doi.org/10.1155/2020/6387173