Journal of Statistics and Its Application, Likelihood Vol 2 No 1, June 2026

Bayesian Fitting of Univariate Skew Normal Distribution on Padua Temperature Dataset

Didik Bani Unggul [1*]

[1] Institut Teknologi Sepuluh Nopember, Indonesia, email: 7003241003@student.its.ac.id

Abstract

Data-driven distributional exploration is an important issue in selecting appropriate statistical methods. Due to the presence of skewness that deviates from the normal distribution, the annual average temperature data of Padua cannot be adequately fitted using a symmetric distribution such as the normal distribution. To address this issue, this study aims to fit a skew normal distribution on Padua temperature dataset. Under a Bayesian approach, we specified priors and constructed posterior distributions as the basis for inference. The sampling results and diagnostics indicate that the chains satisfy the convergence criteria. The posterior means for the three skew normal parameters are 12.05, 1.04, and 2.26, corresponding to the location, scale, and shape parameters, respectively. Lastly, this study not only relaxes the assumption of normality but also provides a modeling foundation for Padua temperature data that can support prediction, forecasting, mapping, and other advanced data analysis purposes.

Keywords: *Bayesian, Skew Normal, Temperature Data Modeling*

Introduction

Statistics has an important role in environmental studies by helping researchers identify patterns, measure relationships, and support evidence-based decisions. Various statistical and analytical approaches, including forecasting (Hasnain et al., 2022; Marinov et al., 2022), predictive modeling (Abbas et al., 2024), spatial analysis (Megahed et al., 2025), and regional clustering (Unggul et al., 2023), can be used to understand environmental change, estimate future conditions, and group areas with similar environmental or infrastructure-related characteristics. Therefore, statistical approaches are flexible and can be applied to various cases.

Padua provides an interesting case for distributional temperature modeling. It is a medium-sized city in North-East Italy, a region recognized as one of the most industrialized and developed parts of the country (Noro et al., 2015). However, this industrialization, driven by urban expansion and infrastructure development, has also brought environmental challenges. Studies indicate that Padua's urban area is among the most severely affected by soil sealing (Todeschi et al., 2022), a process in which natural soil is permanently covered with impermeable artificial materials (Piero et al., 2017). Additionally, urbanization has contributed to the urban heat island (UHI) effect, where city temperatures rise significantly compared to surrounding rural areas (Battista et al., 2020; Nuruzzaman, 2015). To develop effective mitigation strategies, a thorough understanding of Padua's temperature patterns is essential. However, this first requires identifying the most suitable probability distribution for the temperature data, as this will determine the statistical methods used in subsequent analyses.

One important challenge in modeling temperature data is the possibility that the data distribution is not symmetric. If asymmetric data are forced to follow a normal distribution, the resulting parameter estimates, uncertainty measures, and predictive conclusions may be biased or less representative. Therefore, a more flexible distribution is required to capture the empirical characteristics of temperature data.

The skew normal distribution is one of the most popular distribution for handling asymmetric data. Azzalini (1985, 1986) introduced the skew normal distribution as a class of distributions that

https://doi.org/10.30605/likelihood.1054

includes the normal distribution as a special case while allowing skewness through an additional shape parameter. This characteristic makes the skew normal distribution particularly attractive because it maintains a close connection with the normal distribution but relaxes the restrictive assumption of symmetry. More recently, Ye et al. (2024) also discussed the theoretical development and applications of skew-normal models, showing that this distribution remains relevant in modern statistical modeling.

The application of skew normal models has expanded across various fields. In risk analysis and insurance, Bernardi et al. (2012) used skew mixture models to represent loss distributions under a Bayesian framework. Bansal et al. (2008) develop empirical and hierarchical Bayesian methodologies for the skew normal populations. These studies support the use of the skew normal distribution for empirical data that exhibit asymmetry, including environmental and temperature-related datasets.

In addition to the choice of distribution, the inferential framework is also important. Bayesian analysis provides a coherent approach for combining prior knowledge with observed data through the posterior distribution. For the Padua temperature dataset, prior information from climatic classification can help guide the specification of parameter distributions, especially for the location parameter. Furthermore, computational tools such as Stan and the No-U-Turn Sampler allow efficient posterior sampling for models with complex likelihood structures (Carpenter et al., 2017; Hoffman & Gelman, 2014).

Based on these considerations, this study aims to fit the annual average temperature data of Padua using the univariate skew normal distribution under a Bayesian framework. The study focuses on estimating the location, scale, and shape parameters and evaluating whether the skew normal distribution can adequately represent the non-normal characteristics of the data. By doing so, this research contributes to the statistical modeling of asymmetric climate data and provides a distributional foundation for future analyses, including prediction, forecasting, climate risk mapping, and urban environmental planning.

Method

Univariate Skew Normal Distribution

The skew normal distribution, first introduced, has attracted significant attention and is often the primary choice for relaxing normality assumptions (Azzalini, 1985). This distribution is a special case of a broader class developed by Azzalini (1985), defined by a probability density function (pdf) that incorporates components of the standard normal distribution function (Azzalini, 1986). A random variable X following a skew normal distribution, or $X \sim \mathrm{SN}(\xi, \omega, \alpha)$, has a pdf given by Equation (1),

$$f_X(x|\xi,\omega,\alpha) = \begin{cases} \dfrac{2}{\omega}\phi\left(\dfrac{x-\xi}{\omega}\right)\Phi\left(\alpha\left(\dfrac{x-\xi}{\omega}\right)\right) & ,-\infty < x < \infty \\ 0 & ,\text{otherwise} \end{cases} \tag{1}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively. The location, scale, and shape parameters are given by $-\infty < \xi < \infty$, $\omega > 0$, and $-\infty < \alpha < \infty$. It can be seen that the value of $\alpha = 0$ indicates that X will be normally distributed with ξ as the mean parameter and ω as the standard deviation. The mean and variance values of X are shown by equations (2) and (3). Visualizations of the pdf graph for several parameter values are shown in Figure 1.

$$E(X) = \xi + \omega\left(\frac{\alpha}{\sqrt{1+\alpha^2}}\right)\sqrt{\frac{2}{\pi}} \tag{2}$$

$$Var(X) = \omega^2\left(1 - \frac{2}{\pi}\left(\frac{\alpha}{\sqrt{1+\alpha^2}}\right)^2\right) \tag{3}$$

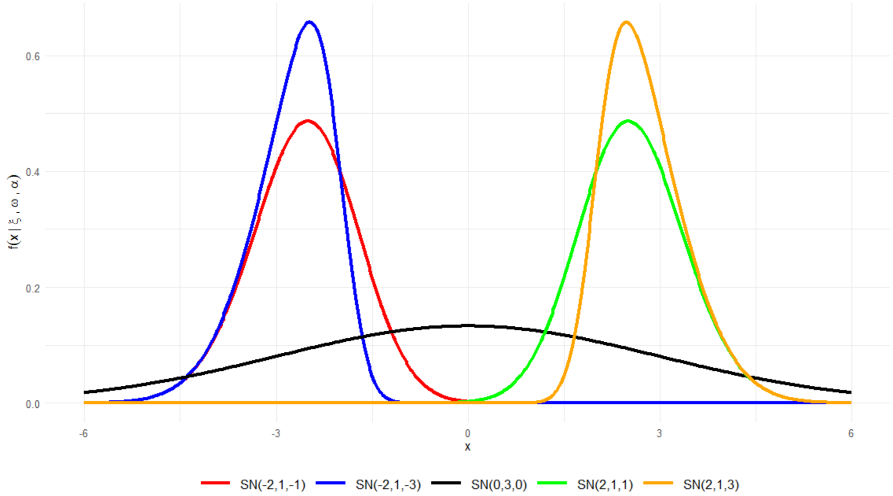

Figure 1. Visualization of the pdf graph of $SN(\xi,\omega,\alpha)$ *for several parameter values.*

Bayesian Analysis

The Bayesian approach has gained widespread popularity among statisticians over the past 50 years (Van De Schoot et al., 2021). This framework treats parameters as random variables with probability distributions. Prior knowledge and the researcher's rational beliefs about the parameters are expressed by the prior distribution (Hoff, 2009). Through Bayes' theorem, this prior is updated to a posterior distribution by incorporating information from observed data via the likelihood function. For illustration, consider a parameter set $\boldsymbol{\theta}$ with prior distribution $p(\boldsymbol{\theta})$. The prior is updated to the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$ through the likelihood $p(\boldsymbol{x}|\boldsymbol{\theta})$, where $\boldsymbol{x}$ represents the observed data. As shown in Equation (4), the posterior distribution is proportional to the product of the prior and likelihood.

$$p(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{p(\boldsymbol{\theta})\,p(\boldsymbol{x}|\boldsymbol{\theta})}{p(\boldsymbol{x})} \propto p(\boldsymbol{\theta})\,p(\boldsymbol{x}|\boldsymbol{\theta}) \tag{4}$$

Once the posterior distribution is obtained, samples are drawn from the joint posterior distribution of all model parameters using a predetermined number of iterations. The sampling is conducted via the No-U-Turn Sampler (NUTS) algorithm (Hoffman & Gelman, 2014), implemented in R using Stan (Carpenter et al., 2017). These posterior samples are then used for statistical inference, including point estimates (e.g., posterior mean) or credible intervals for each parameter.

Results and Discussion

Descriptive Statistics

Table 1. Descriptive statistics for Padua temperature data

Min.	1st Quartile	Median	3rd Quartile	Max.	Mean	Variance
11.3	12.3	12.8	13.2	15.6	12.8	0.51

In this section, we present Table 1, which contains the descriptive statistics of the annual average temperature data in Padua. The data have a minimum value of 11.3 and a maximum of 15.6, with a mean and median both equal to 12.8. The first quartile is 12.3 and the third quartile is 13.2, indicating that half of the data fall between these values. We also obtain a variance of 0.51.

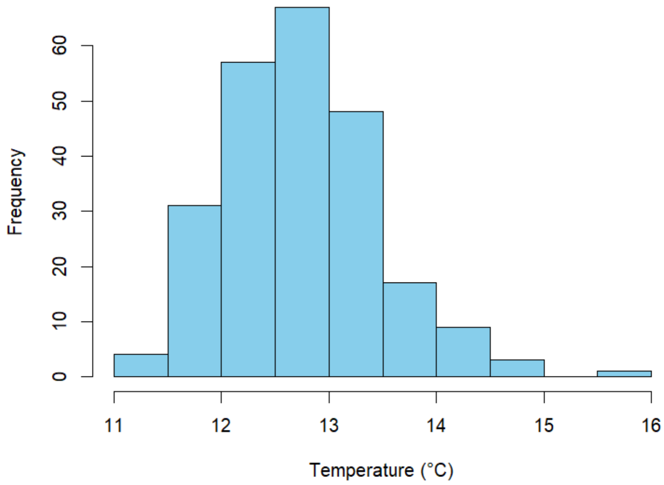

Figure 2. Histogram of Padua annual average temperature data.

Next, we provide the histogram of the data in Figure 2. It can be seen that there is a tendency toward skewness, with a heavier tail on the right side. Shapiro-wilk normality test was then conducted to assess the normality of the data. The test returned a p-value of 0.001259. Since the p-value is below the threshold of 0.05, we reject the null hypothesis that the data come from a normal distribution. This condition confirms the need to use a non-normal distribution for fitting the data. Therefore, the following discussion will focus on modeling the data using a skew-normal distribution to accommodate the observed non-normal characteristics.

Prior and Hyperparameter Specification

In the skew-normal distribution, there are three parameters: the location (ξ), scale (ω), and shape (α). Each requires a prior distribution that encodes our beliefs about their respective characteristics. The prior specification for location parameters is based on the information that the domain of ξ is all real numbers, similar to normal distribution. On the other hand, the restriction that the scale parameter must be positive justifies the use of the inverse gamma distribution. Note that this prior is specified for ω^2, though we will analyze ω directly during inference. Finally, domain suitability also motivates the choice of a normal prior for the parameter α, as it accommodates the unbounded nature of skewness effects. The probability density functions for these prior distributions are given in Equations (5), (6), and (7).

We assign $\xi \sim N\left(\mu_\xi, \sigma_\xi^2\right)$, then

$$p(\xi) = \frac{1}{\sqrt{2\pi\sigma_\xi^2}} \exp\left\{-\frac{1}{2\sigma_\xi^2}\left(\xi - \mu_\xi\right)^2\right\} \tag{5}$$

with $-\infty < \xi < \infty$,$-\infty < \mu_\xi < \infty$, and $\sigma_\xi^2 > 0$.

We have $\omega^2 \sim InvGamma(\gamma, \delta)$, then

$$p(\omega^2) = \frac{(\delta)^\gamma}{\Gamma(\gamma)} (\omega^2)^{-\gamma-1} \exp\left(-\frac{\delta}{\omega^2}\right) \tag{6}$$

with $\omega^2 > 0, \gamma > 0$, and $\delta > 0$.

Finally, $\alpha \sim N(\mu_\alpha, \sigma_\alpha^2)$ so that we get

$$p(\alpha) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left\{-\frac{1}{2\sigma_\alpha^2}(\alpha - \mu_\alpha)^2\right\} \tag{7}$$

with $-\infty < \alpha < \infty$,$-\infty < \mu_\alpha < \infty$, and $\sigma_\alpha^2 > 0$.

Based on the three equations above, several hyperparameters must be carefully specified in advance, including μ_ξ, σ_ξ^2, γ, δ, μ_α, and σ_α^2. To guide this selection, we refer to existing climatic classifications. Fratianni and Acquaotta (2017) previously categorized regions in Italy based on climatic and geographic characteristics. Their results show that Padua belongs to the Sub-continental temperate zone, with an average annual temperature between 10°C and 14°C. This climatic data informs the hyperparameter choices for the location parameter (ξ) prior. Specifically, we use a normal distribution with a mean of 12 and a standard deviation of 10 to represent this prior knowledge. For the scale parameter, an Inverse Gamma(0.001, 0.001) distribution is set as the prior for ω^2. As for the shape parameter α, its prior distribution is assigned hyperparameters with a mean of 1 and a standard deviation of 100. The mean of 1 reflects the right-skewed pattern observed in the data histogram, suggesting that the α parameter distribution is likely centered around positive values. The relatively large standard deviation helps account for uncertainty in specifying the prior distribution.

Posterior Construction

The posterior distribution is constructed by combining the likelihood function with prior distributions using Equation (4). The likelihood follows the probability density function given in Equation (1), representing the joint probability of observing the data $x_1, x_2, \ldots, x_N$ given parameters ξ, ω, and α. Under the assumption of independent priors, the joint posterior distribution takes the form given in Equation (7).

$$p(\xi, \omega^2, \alpha | x_1, x_2, \ldots, x_N) \propto \left[\prod_{i=1}^{N} f_X(x_i|\xi, \omega, \alpha)\right] \times p(\xi) \times p(\omega^2) \times p(\alpha)$$

$$p(\xi, \omega^2, \alpha | \boldsymbol{x}) \propto \left[\prod_{i=1}^{N} \frac{2}{\omega} \phi\left(\frac{x_i - \xi}{\omega}\right) \Phi\left(\alpha\left(\frac{x_i - \xi}{\omega}\right)\right)\right] \times \frac{1}{\sqrt{2\pi\sigma_\xi^2}} \exp\left(-\frac{1}{2\sigma_\xi^2}(\xi - \mu_\xi)^2\right)$$

$$\times \frac{(\delta)^\gamma}{\Gamma(\gamma)} (\omega^2)^{-\gamma-1} \exp\left(-\frac{\delta}{\omega^2}\right) \times \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(-\frac{1}{2\sigma_\alpha^2}(\alpha - \mu_\alpha)^2\right) \tag{7}$$

Posterior Inference and Diagnostics

The joint posterior distribution derived in Equation (7) is analytically intractable due to the complex parameter interactions in the skew normal likelihood. To overcome this, we implement a Markov Chain Monte Carlo (MCMC) method for posterior sampling. Specifically, we use the No-U-Turn Sampler (NUTS) implementation in Stan for this task. Four parallel chains were run for 10,000 iterations each, with the first 5,000 iterations discarded as burn-in. Convergence was verified through trace plot inspection and the values of $\hat{R}$.

ISSN: 3089-9729

Table 2. Posterior summary table for all parameter

Parameter	mean	sd	2.5%	25%	50%	75%	97.5%	$\hat{R}$
ξ	12.05	0.11	11.86	11.98	12.05	12.12	12.28	1.00
ω	1.04	0.09	0.87	0.98	1.04	1.10	1.22	1.00
α	2.26	0.65	1.17	1.81	2.20	2.64	3.74	1.00

Table 2 summarizes the posterior distributions for all parameters, including measures of central tendency (mean, median), uncertainty (standard deviation, credible intervals), and convergence diagnostics (Rhat). The location parameter ξ has a posterior mean of 12.05 (SD = 0.11) and a 95% credible interval [11.86, 12.28]. The scale parameter ω has a mean estimate of 1.04 (SD = 0.09), with its 95% credible interval [0.87, 1.22] properly constrained to positive values, as required for scale parameters. Lastly, the parameter α has posterior mean 2.26 (SD = 0.65) with 95% credible interval [1.17, 3.74]. All $\hat{R}$ values equal to 1 confirm MCMC conver-gence.

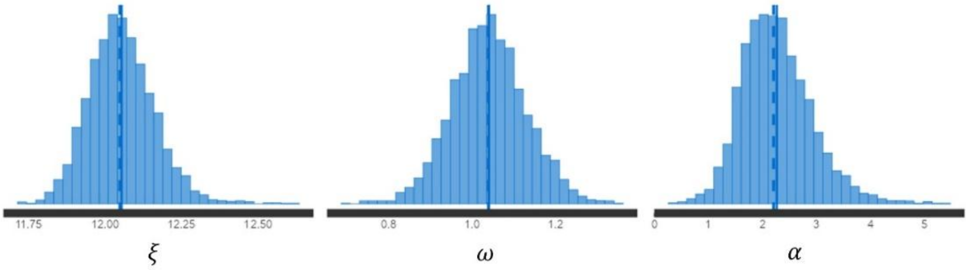

Figure 3. Posterior histograms of all sampled parameters.

For better perspective, Figure 3 presents the posterior histograms of all parameters, allowing a visual assessment of their distributional shapes. Vertical solid lines represent posterior means, while dashed lines indicate medians. The posterior histograms show that the parameters ξ and ω have fairly symmetric, bell-shaped distributions, indicating balanced uncertainty around their central estimates. In contrast, the histogram for α leans slightly to the right, showing that higher values occur more often in the samples.

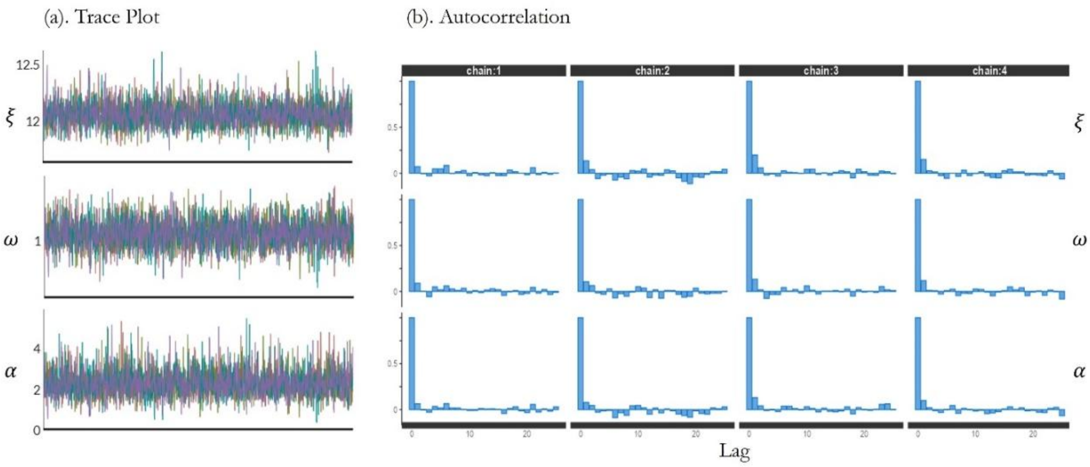

Figure 4. Diagnostic plots of (a) trace plots and (b) autocorrelation functions (ACFs) for sampled parameters.

Next, we also present Figure 4 to provide complementary visual diagnostics of sampling quality. This figure displays the trace plots and autocorrelation for all sampled parameters across four chains. The trace plots in panel (a) display good mixing across chains, with values fluctuating steadily around a central value, suggesting that the sampling process is stable. Panel (b) shows that the autocorrelations drop quickly. These visual diagnostics support the earlier $\hat{R}$ values from Table 2, confirming that the MCMC chains have converged properly.

Conclusion

This study modeled the annual average temperature data of Padua using a Bayesian univariate skew normal distribution. The descriptive analysis, histogram, and Shapiro–Wilk normality test indicated that the data deviate from normality and show a right-skewed pattern. Therefore, the skew normal distribution was considered appropriate because it can accommodate asymmetry while retaining the normal distribution as a special case. Bayesian inference was implemented using RStan with the No-U-Turn Sampler. The analysis yielded posterior means of 12.05, 1.04, and 2.26 for parameter of location (ξ), scale (ω), and shape (α), respectively. Convergence diagnostics including $\hat{R}$, trace plots, and ACF plots, confirmed the reliability of the MCMC sampling process. Overall, the Bayesian skew normal model provides a suitable distributional foundation for Padua temperature data and can support further climate-related analysis, forecasting, and environmental planning.

Acknowledgment

-

References

Abbas, F., Cai, Z., Shoaib, M., Iqbal, J., Ismail, M., Arifullah, Alrefaei, A. F., & Albeshr, M. F. (2024). Machine Learning Models for Water Quality Prediction: A Comprehensive Analysis and Uncertainty Assessment in Mirpurkhas, Sindh, Pakistan. Water, 16(7), 941. https://doi.org/10.3390/w16070941

Azzalini, A. (1985). A Class of Distributions Which Includes the Normal Ones. Scandinavian Journal of Statistics, 12(2), 171–178. JSTOR.

Azzalini, A. (1986). Further Results on a Class of Distributions Which Includes the Normal Ones. Statistica, 46(2), 199–208.

Bansal, N. K., Maadooliat, M., & Wang, X. (2008). Empirical Bayes and Hierarchical Bayes Estimation of Skew Normal Populations. Communications in Statistics - Theory and Methods, 37(7), 1024–1037. https://doi.org/10.1080/03610920701713195

Battista, G., Evangelisti, L., Guattari, C., De Lieto Vollaro, E., De Lieto Vollaro, R., & Asdrubali, F. (2020). Urban Heat Island Mitigation Strategies: Experimental and Numerical Analysis of a University Campus in Rome (Italy). Sustainability, 12(19), 7971. https://doi.org/10.3390/su12197971

Bernardi, M., Maruotti, A., & Petrella, L. (2012). Skew mixture models for loss distributions: A Bayesian approach. Insurance: Mathematics and Economics, 51(3), 617–623. https://doi.org/10.1016/j.insmatheco.2012.08.002

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A Probabilistic Programming Language. Journal of Statistical Software, 76(1). https://doi.org/10.18637/jss.v076.i01

Fratianni, S., & Acquaotta, F. (2017). The Climate of Italy. In M. Soldati & M. Marchetti (Eds.), Landscapes and Landforms of Italy (pp. 29–38). Springer International Publishing. https://doi.org/10.1007/978-3-319-26194-2_4

Hasnain, A., Sheng, Y., Hashmi, M. Z., Bhatti, U. A., Hussain, A., Hameed, M., Marjan, S., Bazai, S. U., Hossain, M. A., Sahabuddin, M., Wagan, R. A., & Zha, Y. (2022). Time Series Analysis and Forecasting of Air Pollutants Based on Prophet Forecasting Model in Jiangsu Province, China. Frontiers in Environmental Science, 10, 945628. https://doi.org/10.3389/fenvs.2022.945628

Hoff, P. D. (2009). A First Course in Bayesian Statistical Methods. Springer New York. https://doi.org/10.1007/978-0-387-92407-6

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research, 15(1), 1593–1623.

Marinov, E., Petrova-Antonova, D., & Malinov, S. (2022). Time Series Forecasting of Air Quality: A Case Study of Sofia City. Atmosphere, 13(5), 788. https://doi.org/10.3390/atmos13050788

Megahed, H. A., Farrag, A. E.-H. A., Mohamed, A. A., Darwish, M. H., AbdelRahman, M. A. E., El-Bagoury, H., D'Antonio, P., Scopa, A., & Saad, M. A. A. (2025). GIS-based modeling and analytical approaches for groundwater quality suitability for different purposes in the Egyptian Nile Valley, a case study in Wadi Qena. Frontiers in Water, 7, 1502169. https://doi.org/10.3389/frwa.2025.1502169

Noro, M., Busato, F., & Lazzarin, R. M. (2015). Urban heat island in Padua, Italy: Experimental and theoretical analysis. Indoor and Built Environment, 24(4), 514–533. https://doi.org/10.1177/1420326X13517404

Nuruzzaman, Md. (2015). Urban Heat Island: Causes, Effects and Mitigation Measures - A Review. International Journal of Environmental Monitoring and Analysis, 3(2), 67. https://doi.org/10.11648/j.ijema.20150302.15

Piero, M., Angelo, B., Antonello, B., Amedeo, D., Carlo, D. M., Michela, I., Giuliano, L., Florindo, M. A., Paolo, P., Simona, V., & Fabio, T. (2017). Soil Sealing: Quantifying Impacts on Soil Functions by a Geospatial Decision Support System. Land Degradation & Development, 28(8), 2513–2526. https://doi.org/10.1002/ldr.2802

Todeschi, V., Pappalardo, S. E., Zanetti, C., Peroni, F., & Marchi, M. D. (2022). Climate Justice in the City: Mapping Heat-Related Risk for Climate Change Mitigation of the Urban and Peri-Urban Area of Padua (Italy). ISPRS International Journal of Geo-Information, 11(9), 490. https://doi.org/10.3390/ijgi11090490

Unggul, D. B., Ainy, K. N., & Jannah, R. (2023). Profiling the Inequality of School Water, Sanitation, and Hygiene Facilities Among Indonesian Regions Using Cluster Analysis. JURNAL KESEHATAN LINGKUNGAN, 15(1), 27–36. https://doi.org/10.20473/jkl.v15i1.2023.27-36

Van De Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. Nature Reviews Methods Primers, 1(1), 1. https://doi.org/10.1038/s43586-020-00001-2

Ye, R., Zhu, W., Qi, Y., & Luo, K. (2024). Skew-Normal Model Theories and Their Applications (1st ed.). CRC Press. https://doi.org/10.1201/9781032711478