Science & Technology Development Journal – Science of The Earth & Environment 2025, 9(2):1129-1138

Research Article

Open Access Full Text Article

Application quantile regression method to forecast heat waves in hanoi with 5-day and 10-day forecasts

Pham Thi Minh[1,*], Truong Ngoc Vy[1], Trinh Minh Ngoc[2], Tran Thi Hong Tuong[3], Phan Vu Hoang Phuong[1], Huynh Thi Minh Suong[4]

ABSTRACT

The paper presents the results of applying the quantile regression method to forecast heat waves in Hanoi. The study uses the maximum temperature (Tx) in the months of May, June, July, August and September in 20 years (1999-2018) observed at the Hanoi station and the 5-day and 10-day forecast data of the sub-seasonal to seasonal forecast model (S2S). The above data series is used to build the forecast equations on different quantiles (q10, q25, q50, q75 and q90) and the linear regression equation. The results of the equation testing show that the linear regression equation and the equations at quantiles 10, 25, 50, 75 and 90 are statistically significant. The study evaluates the quality of the equations based on the data series of heat waves in 3 years 2019-2021 with a total of 45 waves corresponding to 179 days (samples) based on the statistical indexes ME, MAE, RMSE and relative error. The results show that the 5-day and 10-day forecasts, the 75th and 90th percentile regression equations improve by over 20% compared to the error of the S2S model and improve the error by about 6% compared to the linear regression method. In addition, the Tx forecast error values of the 75th and 90th percentile models for 5-day and 10-day periods are both below 2.0^{o}C. Thus, the 75th and 90th percentile regression models can be applied to forecast Tx operations for the Hanoi area.

Key words: Key word: Heat waves, quantile regression, linear regression, Heat waves forecasting

[1]*Department of Meteorology, Hydrology and Water Resources, Ho Chi Minh University of Natural Resources and Environment, Vietnam*

[2]*VNU Hanoi university of science, Vietnam*

[3]*Department of Information Technology, Ho Chi Minh University of Natural Resources and Environment, Vietnam*

[4]*Department of Applied Science, Ho Chi Minh University of Natural Resources and Environment, Vietnam*

Correspondence

Pham Thi Minh, Department of Meteorology, Hydrology and Water Resources, Ho Chi Minh University of Natural Resources and Environment, Vietnam

Email: ptminh@hcmunre.edu.vn

History

- Received: 12-11-2024
- Revised: 29-03-2025
- Accepted: 04-10-2025
- Published Online: 29-11-2025

DOI :

https://doi.org/10.32508/stdjsee.v9i2.794

INTRODUCTION

Heat is a special type of weather that often occurs during the summer months. Heat is a sunny, hot day with a high average daily temperature and is characterized by the highest temperature of the day. According to TCNV2021 standard, A day in a locality is considered hot when the highest temperature of the day (denoted as Tx) reaches $35^oC \leq Tx < 37^oC$, is considered extremely hot when $37^oC \leq Tx < 39^oC$ and is considered an especially extremely hot day when $Tx \geq 39^oC$. Heat is one of the dangerous hydrometeorological phenomena, directly affecting production industries, human health, causing unfavorable conditions for socio-economic activities and people's lives. The first and most obvious impact is that labor productivity will decrease when the temperature increases. When it is hot, people tend to slow down, cognitive functions are impaired, and the risk of work accidents increases. A study by the International Labor Organization predicts that by 2030, the world will lose an equivalent of 2% of total working hours each year due to excessive heat or workers having to slow down to adapt to hot weather (https://thesaigontimes.vn/troi-nong-anh-huong-the-nao-den-kinh-te/). In particular, prolonged heat can cause dehydration, exhaustion, heat stroke to the human body when exposed to high temperatures for a long time and can lead to death. Typically, the heat waves in 2018, 2017 and 2015 caused thousands of deaths and agricultural damage in South Korea, Japan and Pakistan[1]; Heat waves in 2003, 2010 and 2009 caused great economic and human losses in Europe, Russia and Australia[2–4]. In Vietnam, the heat wave in early summer 2017, this heat wave started on May 31, with the peak falling on June 3-4. The highest temperature measured at 13:00 (June 3) at most locations was above 40 degrees, such as Vinh Yen (Vinh Phuc) 40.4 degrees, Huu Lung (Lang Son) 40.3 degrees, Son Dong (Bac Giang) 40.2 degrees, Bac Ninh 40.3 degrees, Lang 40.2 degrees, Chi Linh (Hai Duong) 41.0 degrees, at Ha Dong station (Hanoi) recorded a temperature of 41.5-42 degrees, the highest in the past 46 years (http://www.baomoi.com/ly-giai-nguyen-nhan-cac-tinh-mien-bac-va-ha-noi-thanh-chaolua). Previous studies on heat waves in Vietnam mainly investigated the causes of heat waves for a typical heat wave[5–8] and forecasted heat waves using numerical models, statistical models, and machine learning models; In which, the post-model statistical cor-

Cite this article : Minh P T, Ngoc Vy T, Ngoc T M, Hong Tuong T T, Hoang Phuong P V, Minh Suong H T. **Application quantile regression method to forecast heat waves in hanoi with 5-day and 10-day forecasts**. *Sci. Tech. Dev. J. - Sci. Earth Environ.* 2025; 9(2):1129-1138.

rection method was implemented since the early 1970s [9,10] with the purposes of: 1) improving the quality of rainfall forecasts from numerical models, 2) forecasting for points that are not directly forecasted from numerical models, and 3) applying to the downscaling problem. Regarding machine learning models, in Vietnam, some studies use artificial neural networks to forecast heat waves and the number of hot days [11–14]. The results show the feasibility of machine learning in forecasting heat waves and the number of hot days in Vietnam. In 2002, Nguyen Viet Lanh and Duong Van Thai [15] used the US SAS software to build an equation to forecast maximum temperature and heat waves for Hanoi in May, June and July when this area is affected by the western low pressure. Authors Vo Van Hoa and Tran Hong Thai 2016 [16] also applied the quantile method to improve the quality of seasonal forecasts of severe cold, severe cold and heat waves from ECMWF forecast data. The results of applying quantile methods have significantly reduced the error in identifying the above extreme climate phenomena compared to the method of identifying based on current operational indicators and direct seasonal forecast data from ECMWF. In the world, some authors have used linear regression models and machine learning models to forecast extreme temperatures and temperatures [17–23]. Most of these studies indicate that the Tx forecasting skill is low and therefore cannot be suitable for heat wave forecasting except for the study of Khan et al. 2019 using forest quantile regression to forecast heat waves in Pakistan. Meanwhile, S2S data from the ECMWF global model: includes 12 forecast members (forecast members) around the world contributing to the database. This data is generated twice a week and each forecast product has a forecast horizon of 46 days, with a resolution of $0.125^\circ \times 0.125^\circ$ with 241 grid points. Therefore, if using S2S data to forecast heat waves for the Hanoi area, it will save costs because this is a freely available data set and can be applied to local stations where technology and high-performance computer systems are limited.

STUDYING AREA

Hanoi is currently located from $20^\circ 53'$ to $21^\circ 23'$ North latitude and $105^\circ 44'$ to $106^\circ 02'$ East longitude, bordering the provinces of Thai Nguyen - Vinh Phuc in the North; Ha Nam - Hoa Binh in the South; Bac Giang - Bac Ninh - Hung Yen in the East and Hoa Binh - Phu Tho in the West. The city lies at the heart of the Red River Delta, surrounded by fertile agricultural land, making it an important center for both urban and rural activities

Hanoi is the capital, a centrally-run city, and one of the two special-class cities of Vietnam. It is the largest city (in terms of area) in Vietnam and serves as the political center and one of the two important economic, cultural, and educational hubs of the country. Hanoi is located to the northwest of the center of the Red River Delta, with a terrain that includes the central plain area and hilly regions in the north and west of the city. With an area of 3,359.82 km^2 and a population of 8.4 million people, it is the second most populous city and has the second-highest population density among the 63 provincial-level administrative units of Vietnam, though its population distribution is uneven. Meanwhile, intense heat can reduce human health and productivity. As such, heatwaves and extreme heat events in Hanoi will have a direct impact on the social life of its residents. Accurately forecasting heatwaves can allow for early warnings to help citizens prepare for response measures, thereby minimizing the damage caused by extreme heat. Therefore, this study selects the Hanoi area to apply the quantile regression method for forecasting heatwaves in this region (Figure 1). And Tx monitoring data were taken at Lang meteorological station (Figure 1).

DATA AND METHOD

Data

The study uses observed Tx data from the Lang meteorological station during the months of May, June, July, August, and September over a 20-year period (1999-2018) as the dependent variable. The independent variable is the 5-day and 10-day Tx forecast data from the sub-seasonal to seasonal (S2S) forecasting model developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). These data are used in constructing linear regression equations and quantile regressions at the q10, q25, q50, q75, and q90 levels. And Tx data used to evaluation of the distribution characteristics of heat waves by year and by month in 20 years is the Tx observation during the day of 12 months in a year.

To evaluate the performance of the forecasting equations, 45 heatwaves that occurred in the Hanoi area from 2019 to 2021 were analyzed, encompassing a total of 179 hot days (test data series). The observed data were provided by the National Center for Hydrometeorological Data, while the 5-day and 10-day forecast data from the S2S model were historical reforecasts (Reforecast) of maximum daily temperatures (Tx). These data are part of the S2S dataset from the ECMWF, available in GRIB format with a resolution of 0.125 x 0.125 degrees of latitude and longitude. The data were downloaded from the following

Table 1: List of station used in the study

Station	Longitude (East)	Latitude (North)	Period	Area
Lang	105.80	21.02	1999 - 2018	Hanoi capital, Red River Delta

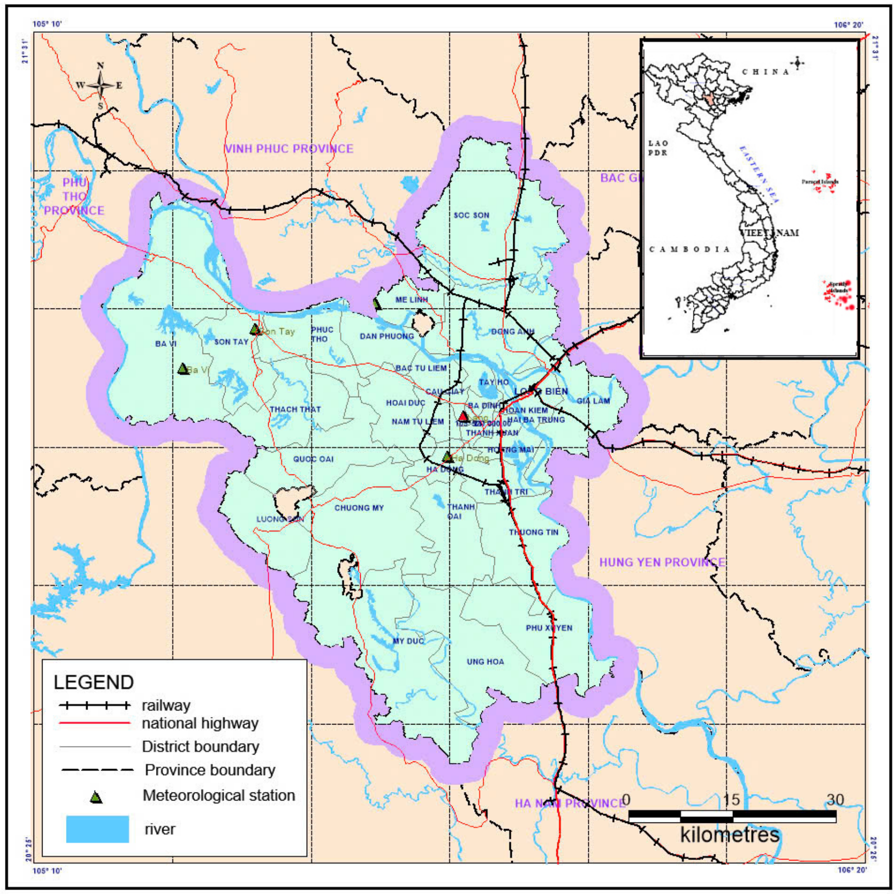

Figure 1: Study area and location of Lang meteorological station (red) for data collection.

website: https://apps.ecmwf.int/datasets/data/s2s-reforecasts-daily-averaged-ecmf/levtype=sfc/type=cf/.

Methodology

Quantile regression was first introduced by Koenker & Bassett in 1978 and is an extension of ordinary least squares (OLS) regression. Instead of estimating the parameters of the mean regression function as done in the OLS method, quantile regression estimates the regression parameters at each quantile of the dependent variable, minimizing the sum of the absolute differences of the regression function at the given quantile of the dependent variable. This means that, rather than determining the marginal effect of the independent variable on the mean of the dependent variable, quantile regression helps identify the marginal effect of the independent variable on the dependent variable at each quantile level, denoted by τ. The detailed theory of quantile regression is presented in [24].

To evaluate the experimental results in this study, we used several statistical evaluation metrics, such as mean error (ME), mean absolute error (MAE), root mean square error (RMSE) [25], and relative error.

Mean Error (ME):

$$ME = \frac{1}{n}\sum_{i=1}^{n}(F_i - O_i) \quad (1)$$

Where:

- n is the total number of forecasts;
- F_i is the forecast value for the i-th instance;
- O_i is the observed value for the i-th instance.

Mean Absolute Error (MAE) :

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|F_i - O_i| \quad (2)$$

Where:

- n is the total number of forecasts;
- F_i is the forecast value for the i-th instance;
- O_i is the observed value for the i-th instance.

Root Mean Square Error (RMSE):

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(F_i - O_i)^2} \quad (3)$$

Where:

- n is the total number of forecasts;
- F_i is the forecast value for the i-th instance;
- O_i is the observed value for the i-th instance.

RMSE is used to represent the average magnitude of the error. Similar to MAE, RMSE does not reflect the bias between forecast and observed values. The optimal value for RMSE is RMSE = 0, meaning that the forecast values are equal to the observed values at every point in the evaluation space (Wilks, 1997)[25].

Relative Error:

The relative error is used to evaluate the degree of error between the forecast results from the linear regression model, the quantile regression model, and the S2S model compared to observations:

$$STD = \left|\frac{F_i - O_i}{O_i}\right| \times 100\% \quad (4)$$

Where:

- $SSTD$ is the relative error;
- F_i is the forecast value for the i-th instance;
- O_i is the observed value for the i-th instance.

RESULTS AND DISCUSSION

Characteristics of heat distribution by month and by year in the period 1999-2018 in Hanoi

A study of heatwaves in the Hanoi area from 1999 to 2018 (Figure 2). The figure shows that 179 heatwaves occurred at the Hanoi station, with a total of 756 hot days. On average, there are about 8 to 9 heatwaves per year at the Hanoi station. In 1999 and 2000, the number of heatwaves was low, with only about 4 per year, but from 2001 onwards, the number of heatwaves tended to increase, ranging from 6 to 8 per year. The number of years with more than 10 heatwaves is quite high, with 16 heatwaves occurring in 2015 at the Hanoi station

Regarding the monthly heatwave statistics (Figure 3), heatwaves begin to occur from March and last until September, with the majority concentrated between May and September, peaking in July. Starting around March, heatwaves are recorded at the Hanoi station, but they are not significant, averaging only 1 to 2 days of heatwaves until April. April, the transitional month between spring and summer, sees a rapid increase in the number of hot days, with the most significant rise occurring in June and July. From August to September and into October, which are transitional months into winter, weak cold air masses begin to affect the Hanoi area, leading to a significant decrease in the number of hot days.

Thus, according to 20 years of statistics (1999-2018) at the Hanoi station, significant heatwave events occurred from May to September. Additionally, the calculation of some statistical characteristics (MEAN, STD, and PV95) for the observed and forecasted Tx data (5-day and 10-day forecasts) from the S2S model at the Hanoi station during 179 heatwave events (1999-2018) is shown in Figure 4. From Figure 4, it is clear that the mean value of observed Tx is approximately 9°C higher than the 5-day and 10-day forecasted Tx from the S2S model. The standard deviation between observations and forecasts does not differ significantly, meaning the degree of fluctuation around the mean value of Tx in the 5-day and 10-day forecast data is quite similar to that of the observed data. Meanwhile, the near-end value of Tx arranged in ascending order (95th percentile) in the observed Tx data series reaches nearly 40°C, while the 5-day and 10-day forecasted Tx data series both have values of 30°C. The evaluation of forecast errors for the 5-day and 10-day Tx forecasts at the Hanoi station, based on the 179 heatwave events with a total of 756 heatwave days, is presented through the ME, MAE, RMSE, and relative error indices (Figure 5). The results show that ME is always less than 0 in the 5-day and 10-day forecasts, indicating that the S2S model consistently underestimates Tx compared to observations. Meanwhile, the error value reflected by the MAE index reaches up to 10°C in both the 5-day and 10-day forecasts, with an error range of approximately 10°C. The relative error of the S2S model is around 24.7%. These statistical results show that directly using forecast data from the S2S model to predict Tx for the Hanoi station is not feasible. Therefore, in the next section, the study develops heatwave forecasting equations for Hanoi using linear regression and quantile regression methods.

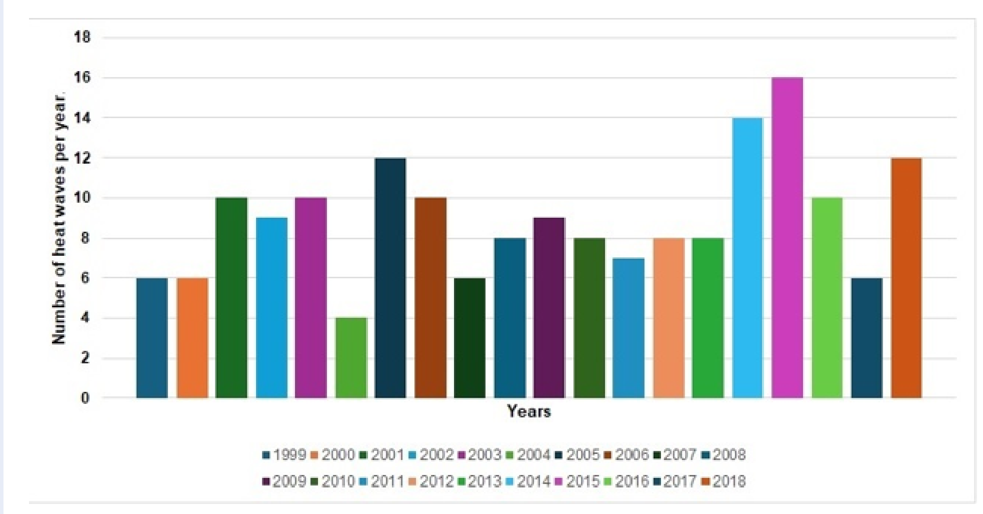

Figure 2: Distribution of heat waves during the year in Hanoi area from 1999-2018.

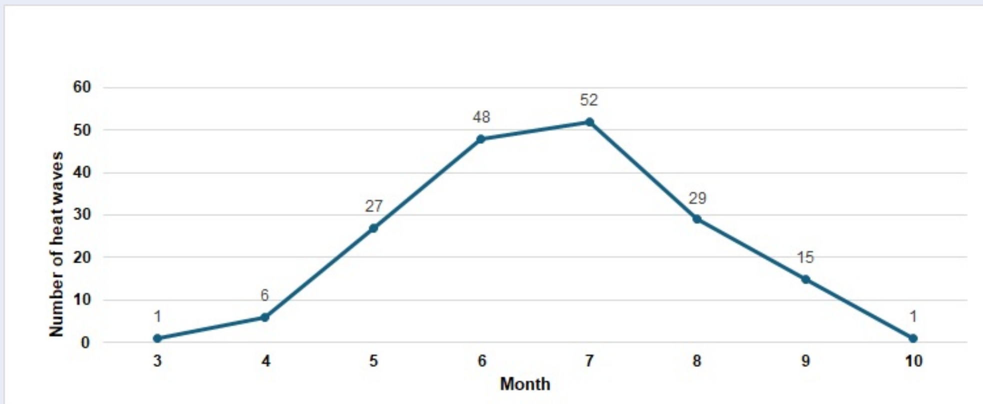

Figure 3: Distribution of heat waves by month in Hanoi area during the period 1999-2018.

Development of heat wave forecasting equation for Hanoi station

The study builds a forecasting equation based on the Tx data set observed at Hanoi station and 5-day and 10-day forecasting data from the seasonal forecasting product of the European Center for Medium and Long-Range Forecasting from May to September in 20 years (1999-2018). The slope of the linear regression equation and the slope and coefficient of freedom of the equations on the 10th, 25th, 50th, 75th and 90th percentiles (KH: q10, q25, q50, q75 and q90) are tested with a statistical significance level of 5%. The results of the 5-day and 10-day Tx forecasting equations are listed in Table 2 and Table 3. The results show that all the linear regression equations, q10, q25, q50, q75 and q90 and the median equation all satisfy the allowable statistical significance level or these equations can all be used for forecasting.

The next section evaluates the quality of the above equations through a test dataset of 45 heat waves corresponding to 179 hot days (2019-2021).

Results of evaluating the equations on the test data series

Using the equations in Table 1 and Table 2 to forecast heat waves in the test data series from 2019 to 2021 (including 45 heat waves corresponding to 179 hot days). Then the study used evaluation indicators including: ME, MAE, MRSE and relative error. The results are shown in Figure 6 and Figure 7.

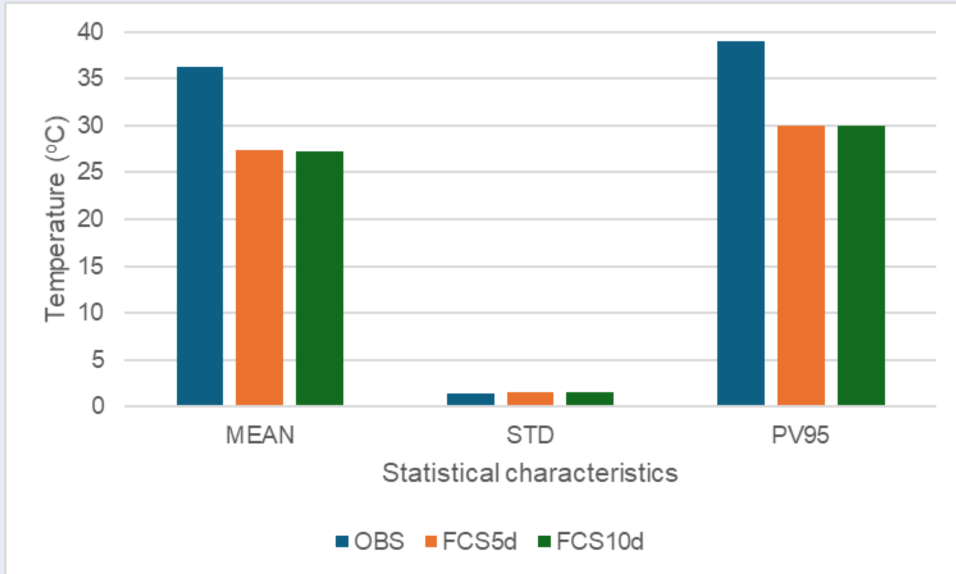

Figure 4: Mean, standard deviation and 95th percentile of the observed Tx data series, forecasted from the seasonal model (S2S) for 5-day and 10-day forecast in the period 1999-2018 at Hanoi station.

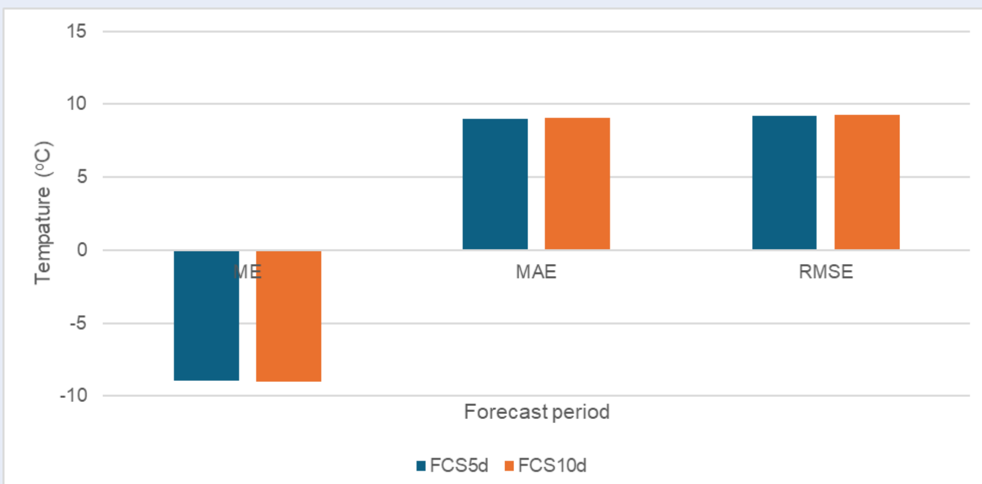

Figure 5: 5-day and 10-day Tx forecast errors of the seasonal forecasting model (S2S) during heat waves in the period 1999-2018.

For the 5-day forecast period, the ME error of the S2S model, the linear regression model and the quantile regression model all have negative values (Fig. 6a). In other words, all of the above models forecast Tx lower than the observation. The MAE error is the average absolute error of forecasting Tx, reflecting the value of the forecast error. From Figure 6b, it can be seen that the error of the S2S model is about 10^{o}C, the error of the q90 and q75 quantile models is below 2^{o}C, while q50 and OLS (linear regression) are below 4^{o}C, and q10 is 7^{o}C. The RMSE error reflects the Tx error amplitude, from Figure 6c, it can be seen that the error amplitude of q75 and q90 is smaller than the error amplitude of the S2S model and linear regression, with a value of about 2^{o}C, while q10, q25 and q50 have Tx error amplitudes smaller than the Tx error amplitude of the S2S model but larger than the error amplitude of the OLS model, except for the q50 quantile. The relative error of the q75 and q90 quantile regression models improves by 20-25% compared to the S2S

Table 2: Linear regression equation and quantile regression corresponding to the confidence level at Hanoi station for 5-day forecast.

Method	Equation	P Slope	P Free coefficients	Notes
OLS	y = 0.4231x + 22.1826	Less than 10^{-4}		receive
Median	y = 0.375x + 23.725		Less than 10^{-4}	receive
q10	y = 0.5x + 16.8			receive
q25	y = 0.375x + 21.925			receive
q50	y = 0.375x + 23.725			receive
q75	y = 0.4333x + 23.5667			receive
q90	y = 0.5x + 23			receive

Table 3: Linear regression equation and quantile regression corresponding to the confidence level at Hanoi station for 10-day forecast.

Method	Equation	P Slope	P Free coefficients	Notes
OLS	y = 0.3891x + 23.0731	Less than 10^{-4}		receive
Median	y = 0.34x + 24.6398		Less than 10^{-4}	receive
q10	y = 0.4167x + 18.9			receive
q25	y = 0.3x + 23.8999			receive
q50	y = 0.34x + 24.6398			receive
q75	y = 0.4x + 24.4			receive
q90	y = 0.48x + 23.68			receive

model, and improves by 4-6% compared to the OLS model. The remaining quantiles improve the error by 10 to 16% compared to the S2S model, however, these quantiles do not improve the error compared to OLS except quantile q50 (Figure 6d).

For the 10-day forecast horizon, the results of calculating the ME MAE, MRSE and relative forecast errors Tx of the S2S, OLS, q10, q25, q50, q75 and q90 models have similar results to those of the 5-day forecast horizon. Specifically, the S2S model, the linear regression model and the quantile regression model all forecast Tx with a lower bias than the observation (showing a negative ME value (Figure 7a). Evaluating the 10-day Tx forecast error value of the models through the MAE index, the results show that the model error of the quantile model q90 and q75 is less than 2^oC, smaller than the Tx forecast error of the S2S and OLS models (linear regression), especially q50 has an error equivalent to OLS with a value below 4^oC, while q10, q25 have error values smaller than the S2S model, but larger than OLS (Figure 7b). The RMSE error reflects the Tx error amplitude, from Figure 7c it can be seen that the error amplitude of q75 and q90 is smaller than the error amplitude of the S2S model and linear regression, with a value of about 2^oC, while q10, q25 and q50 has a smaller Tx error margin than the S2S model but larger than the OLS model, except for the q50 quantile. The relative errors of the q50, q75 and q90 quantile models improve by 2% to 6% compared to OLS, and improve the error by about 17-25% compared to the S2S model. The q10 and q25 models improve the error compared to the S2S model by about 10 to 13%, but q10 and q25 do not improve the error compared to OLS (Figure 7d).

CONCLUSIONS

a/ The distribution of heat waves by month and by year in the period 1999-2018 is quite suitable for the climate conditions in the Hanoi area. Moreover, the study has shown that the Tx forecast error at Hanoi station for 5 days and 10 days of the S2S model is about 10^oC.

b/ The study builds linear regression equations and regression equations on quantiles q10, q25, q50, q75 and q90. These equations all ensure statistical reliability of 95%.

c/ The results of the evaluation of the equations based on the test data series show that the 5-day and 10-day forecast error of the q90 and q75 quantile models is below 2^oC. And the 5-day forecast relative error of the

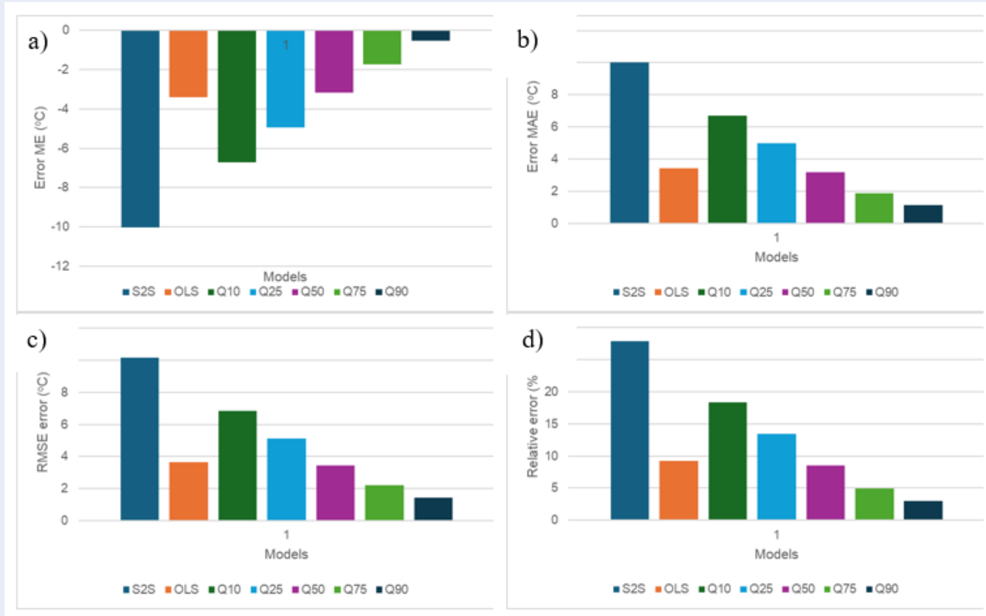

Figure 6: 5-day Tx forecast error (ME, MAE, RMSE and RE –Relative error) of the S2S model and quantile regression models at Hanoi station for 45 heat waves in the period 2019-2021.

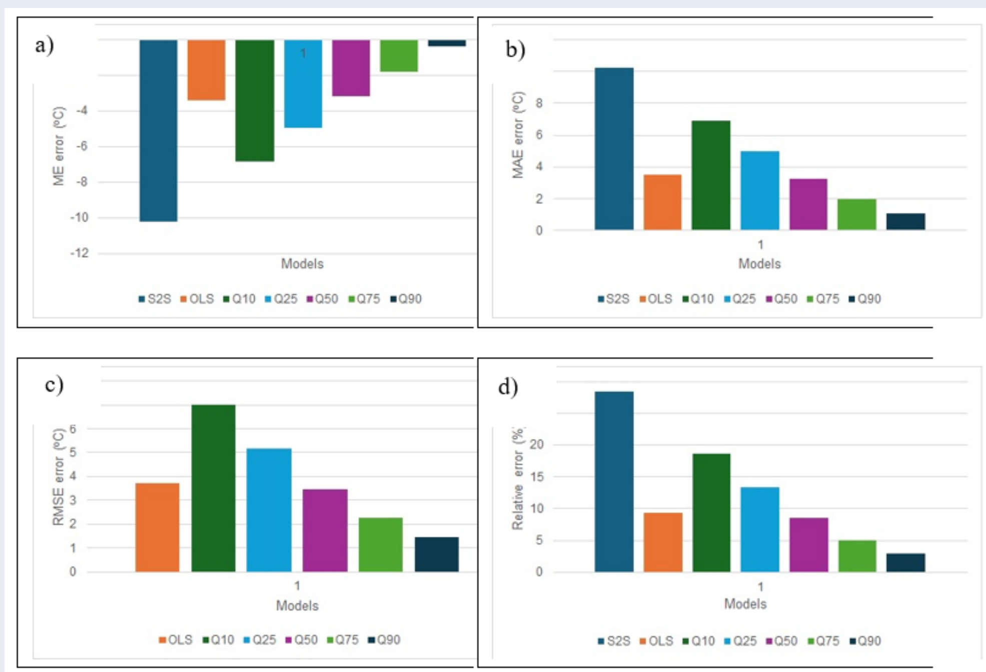

Figure 7: 10-day Tx forecast error (ME, MAE, RMSE and RE – Relative error) of the S2S model and quantile regression models at Hanoi station for 45 heat waves in the period 2019-2021.

q75 and q90 quantile regression models improves by 20-25% compared to the S2S model, and improves by 4-6% compared to the OLS model. The the 10-day forecast relative errors of the q50, q75 and q90 quantile models improve by 2% to 6% compared to OLS, and improve the error by about 17-25% compared to the S2S model. Therefore It is possible to apply the quantile regression equations q75 and q90 in the Tx forecasting service for 5 days and 10 days at Hanoi station.

COMPETING INTERESTS

The authors declare that there are no conflicts of interest related to the publication of this article.

AUTHORS' CONTRIBUTIONS

Authors Pham Thi Minh, Truong Ngoc Vy, Trinh Minh Ngoc, Tran Thi Hong Tuong, Phan Vu Hoang Phuong and Huynh Thi Minh Suong jointly performed the steps and tested the results of this study.

ACKNOWLEDGMENTS

This research has been supported by a grant for the student research project (SV.2024.11) from Ho Chi Minh University of Natural Resources and Environment, 236B Le Van Sy street, Ward 1, Tan Binh District, Ho Chi Minh City.

REFERENCES

1. Khan N, Shahid S, Ismail T, Ahmed K, Nawaz N. Trends in heat wave related indices in Pakistan. Stoch Env Res. 2018;.
2. Bono D, Peduzzi A, Kluser P, Giuliani S, G. Impacts of Summer 2003 Heat Wave in Europe. Heat Wave in Europe. 2004;1:2–5.
3. Dole R, Hoerling M, Perlwitz J, Eischeid J, Pegion P, Zhang T, et al. Was there a basis for anticipating the 2010 Russian heat wave. Geophys Res Lett. 2011;38:1–5.
4. Siddaway J, Petelina S. Transport and evolution of the 2009 Australian Black Saturday bushfire smoke in the lower stratosphere observed by OSIRIS on Odin. J Geophys Res. 2011;116:1–9.
5. Mau ND, Thang NV, Hieu NT. What causes heat wave during 1-6 june 2017 in the northern vietnam. Journal of Climate Change Scientce. 2017;2:1–7.
6. Hang PM, Dung TT, Quang ND. The impact of asiatic low and southwest pacific sub-tropical high to the heatwaves activity over the northern central from 2010 to 2015. Journal of hydro-meteorology. 2017;2:44–52.
7. Lanh NV, Huong C. Hot dry weather and weather patterns causing it in the highlands. Journal of hydro-meteorology. 2016;665:6–11.
8. Lanh NV. Hot dry weather and weather patterns causing it in Vietnam. Journal of hydro-meteorology. 2010;597:8–13.
9. Klein WH. An objective method of predicting quantitative precipitation in the Tennessee and Cumberland Valleys. In: Proc of the First Statistical Conf. Amer Meteor Soc; 1968. p. 20–28.
10. Glahn HR, Lowry DA. The use of model output statistics (MOS) in objective weather forecasting. Journal of applied meteorology. 1972;11:1203–1211.
11. Lam HP, Dien NH. Forecasting maximum temperature for the northern plains of Vietnam using artificial neural networks. Journal of hydro-meteorology. 2008;571:20–23.
12. Minh LT, Lam HP, Đat TT, Thanh VT, Nga LT. Automatically correction for forecasts city temperature from the ifs model output. Journal of hydro-meteorology. 2018;693:41–47.
13. Chinh TH, Lam HP, Thang VV, Kien TB. Predictability of hot day number over Vietnam by the artificial neural network. Journal of hydro-meteorology. 2022;741:98–105.
14. Khiem MV, Lam HP, Huong NV, Mai HT, Van TN, Hoa NT, et al. Predictability of hot day number over Vietnam by the artificial neural network. Journal of hydro-meteorology. 2022;736:64–78.
15. Lanh NV, Thai DV. Forecast of maximum temperature and heat for Hanoi in May, June and July when the area is affected by low pressure in the west. Journal of hydro-meteorology. 2002;502:14–19.
16. Hoa VV, Thai TH. Application of quantile methods to improve cold surge and heat wave phenomena forecast skill based on seasonal prediction dataset of ecmwf. Journal of hydro-meteorology. 2016;671:32–38.
17. Gao M, Wang B, Yang J, Dong W. Are peak summer sultry heat wave days over the Yangtze-Huaihe River basin predictable. J Clim. 2018;31:2185–2196.
18. Hjort J, Suomi J, Käyhkö J. Spatial prediction of urban-rural temperatures using statistical methods. Theor Appl Climatol. 2011;106:139–152.
19. Liu X, Yuan S, Li L. Prediction of temperature time series based on wavelet transform and support vector machine. JCP. 2012a;7:1911–1918.
20. Naing W, Htike ZZ. Forecasting of monthly temperature variations using random forests APRN. J Eng Appl Sci. 2015;10:10109–10112.
21. Zhu Z, Li T. Extended-range forecasting of Chinese summer surface air temperature and heat waves. Clim Dyn. 2018;50:2007–2021.
22. Brennan A, Cross PC, Creel S. Managing more than the mean: Using quantile regression to identify factors related to large elk groups. J Appl Ecol. 2015;52:1656–1664.
23. Khan N, Shahid S, Juneng L. Prediction of heat waves in Pakistan using quantile regression forests. Atmospheric Research. 2019;221:1–11.
24. Hao L, Naiman DQ, Naiman DQ. Quantile regression. SAGE LONDON. 2007;149:1–126.
25. Daniel W, S. Statistical Methods in the Atmospheric Scienes. vol. 59; 1997. p. 255–255.

Tạp chí Phát triển Khoa học và Công nghệ – Khoa học Trái đất và Môi trường 2025, 9(2):1129-1138

Bài nghiên cứu

Áp dụng phương pháp hồi quy phân vị dự báo nắng nóng tại Hà Nội với hạn dự báo 5 ngày và 10 ngày

Phạm Thị Minh[1,*], Trương Ngọc Vy[1], Trịnh Minh Ngọc[2], Trần Thị Hồng Tường[3], Phan Vũ Hoàng Phương[1], Huỳnh Thị Minh Sương[4]

TÓM TẮT

Bài báo trình bày kết quả ứng dụng phương pháp hồi quy phân vị để dự báo nắng nóng cực đoan tại Hà Nội. Nghiên cứu sử dụng số liệu nhiệt độ tối cao (Tx) trong các tháng 5, 6, 7, 8 và 9 của 20 năm (1999–2018) quan trắc tại trạm Hà Nội, cùng với số liệu dự báo 5 ngày và 10 ngày của mô hình dự báo từ hạn dưới mùa đến hạn mùa (S2S). Chuỗi số liệu trên được sử dụng để xây dựng các phương trình dự báo tại các phân vị khác nhau (q10, q25, q50, q75 và q90) và phương trình hồi quy tuyến tính. Kết quả kiểm định phương trình cho thấy phương trình hồi quy tuyến tính và các phương trình tại các phân vị 10, 25, 50, 75 và 90 đều có ý nghĩa thống kê. Nghiên cứu đánh giá chất lượng các phương trình dự báo dựa trên chuỗi số liệu nắng nóng của 3 năm 2019–2021 với tổng cộng 45 đợt, tương ứng 179 ngày (mẫu), thông qua các chỉ số thống kê ME, MAE, RMSE và sai số tương đối. Kết quả cho thấy với dự báo 5 ngày và 10 ngày, các phương trình hồi quy ở phân vị 75 và 90 giúp cải thiện trên 20% so với sai số của mô hình S2S, đồng thời giảm sai số khoảng 6% so với phương pháp hồi quy tuyến tính. Ngoài ra, giá trị sai số dự báo Tx của các mô hình phân vị 75 và 90 trong các giai đoạn 5 ngày và 10 ngày đều nhỏ hơn 2,0°C. Do đó, các mô hình hồi quy tại phân vị 75 và 90 có thể được ứng dụng trong nghiệp vụ dự báo Tx cho khu vực Hà Nội.

Từ khoá: Sóng nhiệt, Hồi qui phân vị, Hồi qui tuyến tính, Dự báo sóng nhiệt

[1]*Khoa Khí tượng Thuỷ văn và Tài Nguyên nước, trường Đại học Tài nguyên và Môi trường thành phố Hồ Chí Minh, Việt Nam*

[2]*Trường Đại học Khoa học Tự nhiên Hà Nội, Việt Nam*

[3]*Khoa công nghệ thông tin, trường đại học Tài nguyên và Môi trường tp Hồ Chí Minh, Việt Nam*

[4]*Khoa Khoa học ứng dụng, trường đại học Tài nguyên và Môi trường tp Hồ Chí Minh, Việt Nam*

Liên hệ

Phạm Thị Minh, Khoa Khí tượng Thuỷ văn và Tài Nguyên nước, trường Đại học Tài nguyên và Môi trường thành phố Hồ Chí Minh, Việt Nam

Email: ptminh@hcmunre.edu.vn

Lịch sử

- Ngày nhận: 12-11-2024
- Ngày sửa đổi: 29-03-2025
- Ngày chấp nhận: 04-10-2025
- Ngày đăng: 29-11-2025

DOI :

https://doi.org/10.32508/stdjsee.v9i2.794

Trích dẫn bài báo này: Thị Minh P, Ngọc Vy T, Minh Ngọc T, Hồng Tường T T, Hoàng Phương P V, Minh Sương H T. **Áp dụng phương pháp hồi quy phân vị dự báo nắng nóng tại Hà Nội với hạn dự báo 5 ngày và 10 ngày**. *Sci. Tech. Dev. J. - Sci. Earth Environ.* 2025; 9(2):1129-1138.