Quantum Information and Computation, Vol. 23, No. 3&4 (2023) 0181–0194

THE EFFECT OF SUPERPOSITION AND ENTANGLEMENT ON HYBRID QUANTUM MACHINE LEARNING FOR WEATHER FORECASTING

BEŞİR OĞUR[a]
Department of Computer Science , Maltepe University
Marmara Education Village, Maltepe, İstanbul, 34857, Turkey

İHSAN YILMAZ[b]
Department of Computer Science , Çanakkale Onsekiz Mart University
Terzioğlu Campus, Prof.Dr. Sevim Buluç Street, Çanakkale, 17100, Turkey

Received April 11, 2022
Revised December 29, 2022

Recently, proposed algorithms for quantum computing and generated quantum computer technologies continue to evolve. On the other hand, machine learning has become an essential method for solving many problems such as computer vision, natural language processing, prediction and classification. Quantum machine learning is a new field developed by combining the advantages of these two primary methods. As a hybrid approach to quantum and classical computing, variational quantum circuits are a form of machine learning that allows predicting an output value against input variables. In this study, the effects of superposition and entanglement on weather forecasting, were investigated using a variational quantum circuit model when the dataset size is small. The use of the entanglement layer between the variational layers has made significant improvements on the circuit performance. The use of the superposition layer before the data encoding layer resulted in the use of less variational layers.

Keywords: Quantum Computing, Machine Learning, Weather Forcasting, Variational Quantum Circuit, Hybrid Quantum-Classic Neural Network

1 Introduction

The main objective is to solve the optimal solution in the shortest time with the least cost, when solving natural phenomena or real problems in daily life. Solution methods developed in counterpart to the degree of difficulty arising from the nature of the problem are accepted as the scale of success. This approach has emerged as a fundamental goal in the evolution of computational methods, as well as in all developing fields of science and technology.

Quantum mechanics, which emerged to explain the behavior of subatomic particles, which cannot be understood by classical physics laws, has has emerged as a powerful approach again by proposing an alternative computation method to the conventional computers having the Von Neumann architecture which cannot meet the needs in terms of their ability to solve complex problems and perform faster processing.

In general, computation is defined as a physical process. Quantum computing is based on the fundamental postulates of quantum mechanics, while classical computation is built on the

[a]besirogur@gmail.com
[b]iyilmaz@comu.edu.tr

rules of classical physics [1, 2]. Classical computers, which are getting smaller and faster in terms of hardware and processing capabilities, are still insufficient in solving problem types with NP-Hard complexity class. On the other hand, the advantages obtained thanks to the principles of entanglement and superposition have made quantum computers more attractive. Developed quantum algorithms [3, 4, 5] and quantum processors exceeding the 100 qubit limit are important steps for the development of quantum technologies [6].

Machine learning is a modeling that can learn using patterns and relationships between data. Classical computers have the ability to solve basic algebraic operations and problems that can be expressed algorithmically at a much faster rate than the human brain. In contrast, classical computer algorithms in problems such as estimation, classification, association, pattern recognition, etc. are quite inefficient in subjects that require thinking like the human brain. The most important difference between machine learning and classical computer programming is that it can perform the learning process through examples without the need for algorithms with strictly defined rules [7, 8, 9].

The algorithms developed for machine learning, the rapid increase in the processing power of the classical computer, and the exponentially increasing datasets, the application of machine learning to many areas such as classification, regression, clustering were provided. The neural networks, one of the important methods, give very efficient results in estimation problems in machine learning. Weather forecasting is an example of an estimation problem.

Nowadays, weather forecasts are made on hourly, daily, weekly and monthly time scales and play an important role in society. Serious weather events such as hurricanes, extreme heat waves, and tornadoes occur each year, resulting in thousands of deaths and property damage. As the frequency and intensity of extreme weather events increase, weather forecasting becomes even more important. In real life, weather forecasting is a complex process. It is difficult for humans to make the necessary calculations for forecasting and use them to predict future weather when faced with complex data. However, the forecast performance of classical algorithms used for weather forecasting is directly proportional to the size of the dataset.

Many classical approaches can be used when the dataset size is small. As an alternative to these approaches, many types of research are carried out in quantum machine learning, which is the intersection of quantum computers and machine learning [10, 11, 12]. Hybrid quantum-classical machine learning, which is considered a type of quantum machine learning and is created by combining classical machine learning tools and quantum circuits, provides significant advantages over classical machine learning [13, 14, 15]. Furthermore, It is known that hybrid quantum-classical algorithms developed according to the quantum variational circuit approach produce efficient results in terms of fault tolerance limits for practical uses. This hybrid computation method provides important opportunities to use the capabilities of quantum computing already developed in solving practical problems [16]. Additionally, some related studies in the literature highlight the use of quantum computers for weather forecasting. Frolov [17] addressed the fundamental limitations that do not allow the development of numerical weather and climate forecasting models to be met, and the applicability of quantum computations and quantum computers to solve numerical weather and climate forecasting problems. Safari A. and Ghavifekr A. A. [18], compared Quantum Neural Networks (QNN) method with other artificial neural network-based techniques in weather forecasting.

As a result of the comparison, they concluded that quantum technology and QNN have the potential to be combined with other techniques leading to accurate models with fast processing and execution speed. Quantum neural networks were used as weather forecasting tools in [19, 20]. Experimental results reveal that quantum computing-based models have better prediction performance than other machine learning-based prediction models implemented on classical computers.

Based on the above analysis, this study aims to investigate the effect of two important quantum properties, entanglement and superposition, on hybrid quantum-classical neural network performance on small-size training data for weather forecasting. In addition, the results of using different number of variational layers on the quantum circuit on weather forecast performance are compared.

2 Background

2.1 Variational Quantum Circuit

It is difficult to integrate the nonlinear mathematical structure of artificial neural networks with the linearly functioning quantum computing architecture. However, algorithms developed with a variational quantum circuit solve this integration problem by combining a quantum circuit, which prepares quantum states and is guided by a limited number of qubits, with a classical optimization algorithm [11, 21]. In Figure 1, the working principle of the quantum variational algorithm is given schematically.

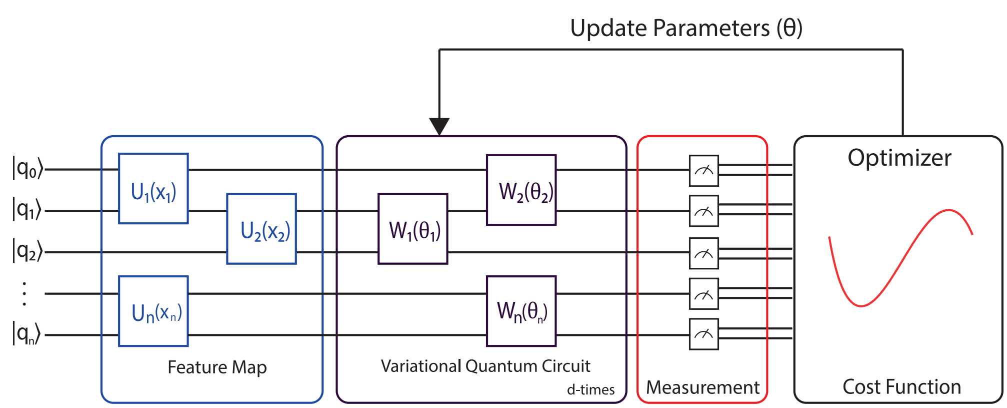

Fig. 1. The working principle of the variational quantum algorithm [22].

The proposed method with the variational quantum circuitry is similar to the artificial neural network (ANN) structure by optimizing the expected value obtained from the quantum processing unit in a classical processing unit, as an application of the hybrid quantum-classical approach [13, 14, 23].

Quantum algorithms developed with variational quantum circuitry are efficient in terms of error tolerance thanks to their parameter update feature. This advantage makes the variational quantum circuit approach prime candidates for implementation on near-term quantum processors [15, 16]. The working principle of the variational circuit algorithm consists of the following steps [13, 21, 24]:

1. $\{x_i\}$ input data is encoded into a quantum state.

$$\{x_i\} \rightarrow |\Psi(x)\rangle \tag{1}$$

2. A new state is obtained by applying rotation operators (unitary quantum operators) that can take parameters to the prepared quantum state.

$$|\Psi_{out}(x_i;\theta)\rangle = U(\theta)|\Psi(x_i)\rangle \tag{2}$$

3. In variational quantum circuits, Pauli operators are generally used as observables to measure the expected value of some observables selected on the circuit. The expected value of a $\hat{B}$ observable on the circuit;

$$y(x_i;\theta) = \langle\Psi U^{\dagger}(\theta)\hat{B}U(\theta)|\Psi\rangle \tag{3}$$

4. The expected value obtained from the quantum processing unit with classical calculation methods is given as an input to a cost function, and the cost function is minimized with a classical optimization algorithm. The cost value determined for this process is expressed over the θ parameter, according to the desired values of $y\left(x_i;\theta\right)$ and $f\left(x_i\right)$. This iteration is continued until the desired error level is achieved.

$$Cost(f(x_i), y(x_i;\theta)) \tag{4}$$

3 Material and Method

In this study, the effects of superposition and entanglement in a five-qubit quantum variational circuit model were investigated in weather forecasting, which is a regression problem. In this framework, variational layers were used in the range 1-8 in five different experiments on the use of superposition, entanglement, and rotation layers in different places. The characteristics of the dataset used and the details of the experiments are given in this section.

3.1 Dataset

The dataset includes hourly weather data for the Istanbul-Kireçburnu/Sarıyer region for the years 2016-2020. In the dataset, there are 43829 data in total, including total precipitation amount, vapor pressure, relative humidity, cloudiness, current pressure, and temperature. The characteristics of the dataset are given in Table 1.

3.2 Preprocessing

Attribute data (Total amount of precipitation , Vapor Pressure, Relative Humidity, Amount of cloudiness , Actual Pressure) were standardized using the following formula:

$$\hat{x} = \frac{x-\mu}{\sigma} \tag{5}$$

where, x : attribute data, μ:: attribute mean, σ: attribute standard deviation, $\hat{x}$: standardized attribute data.

3.3 Experimental Implementations

The five-qubit variational quantum circuit used for weather forecasting consists of five stages. First, all qubits are initialized in the initial state $|0\rangle$. Then, optionally, all qubits are placed in an equal amplitude superposition (this was investigated in Experiment 3). Then, feature mapping is performed by embedding the five features of the weather dataset into the α angles of the $R_y(\alpha)$) rotation gates with the data coding circuit consisting of $R_y(\alpha)$ rotation gates. Then, N variational layers are added and finally, the measurement is performed for the expected value on the basis of $\langle Z\rangle$. With $|\psi\rangle$ being the quantum state before the measurement, the expected value is calculated as follows:

$$\langle Z\rangle = \langle\psi|Z|\psi\rangle = \int dx\psi^*(x)Z\psi(x) \tag{6}$$

where $\psi(x) = \langle x|\psi\rangle$. The spectral decomposition of Z is

$$\langle Z\rangle = \langle\psi|Z|\psi\rangle = \sum_i z_i\langle\psi|\psi_i\rangle\langle\psi_i|\psi\rangle = \sum_i z_i|\langle\psi_i|\psi\rangle|^2 \tag{7}$$

with

$$Z = \sum_i z_i|\psi_i\rangle\langle\psi_i|. \tag{8}$$

where $Z|\psi_i\rangle = z_i|\psi_i\rangle$. The obtained measurement results are fed to the output layer. Although the output layer consists of a neuron, the initial weights have a normal distribution. In addition, in the hybrid quantum-classical models used in the study, all parameters are trained by the back propagation method and the mean square error function was used as the error function. Then, the estimation result is obtained with the models obtained as a result of the training.The basic quantum variational circuit model is given in Figure 2. A detailed representation of the circuits used in the variational layer is given in Figure 3.

In the first experiment performed in this study, the entanglement layer types 1 circuit and rotation layer given in Figure 3 were used in the variational layer in the quantum variational circuit architecture given in Figure 2. The entanglement layer types 1 circuit indirectly entangles all qubits by entanglement transfer. The parameters of the rotation gates in the rotation layer are trainable by a classical optimization algorithm. The circuit architecture used in this experiment is given in Figure 4.

In the second experiment, the effect of entanglement was investigated by removing the entanglement sublayer in the variational layer used in the first experiment. The circuit architecture used in this experiment is given in Figure 5.

Table 1. Weather dataset properties

	Total amount of precipitation (kg/m^2)	Vapor pressure (hPa)	Relative humidity (%)	Amount of cloudiness (8 okta)	Actual pressure (hPa)	Temperature (°C)
Mean	0,09	14,43	79,33	3,89	1008,8	15,41
Standard Deviation	0,71	6,43	14,35	2,74	7,99	7,61
Variance	0,51	41,36	205,86	7,58	63,83	57,84
Mod	0	9,3	99	7	1007,8	23,2
Median	0	13	81	4	1008,3	15,4

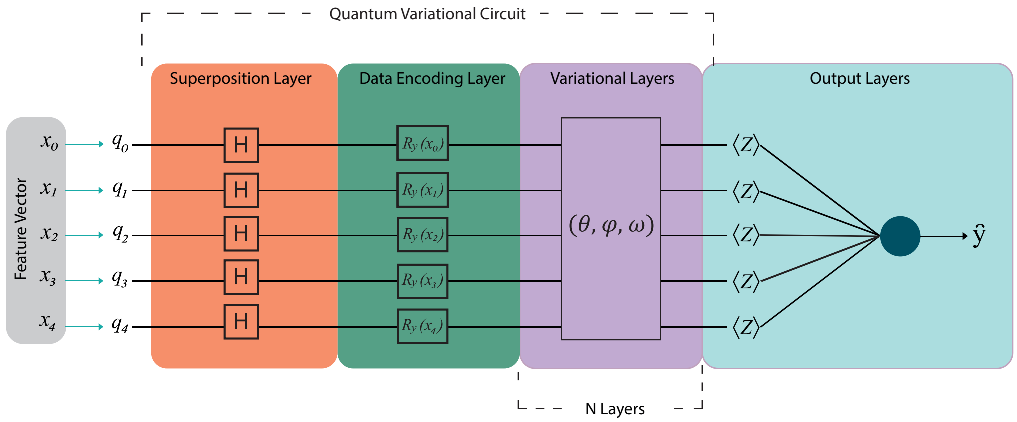

Fig. 2. Basic quantum variational circuit model diagram.

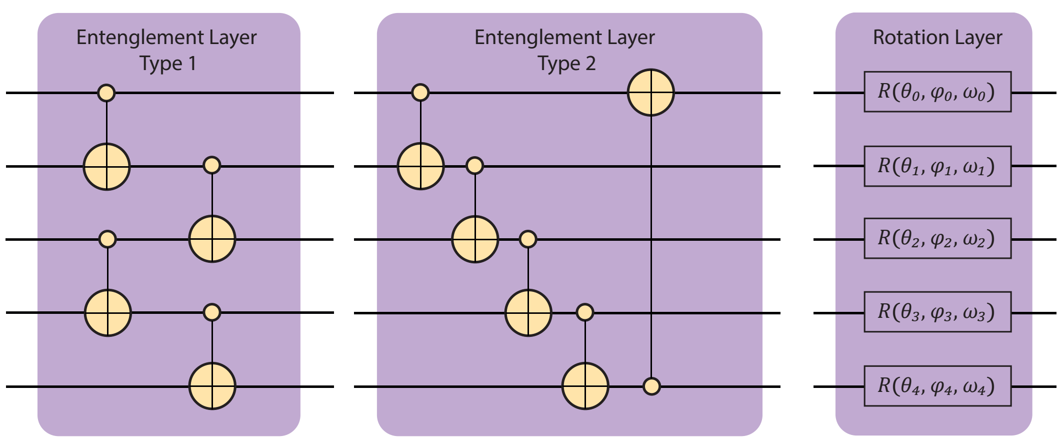

Fig. 3. Detailed representation of the different entanglement and rotation sublayers used in the variational layer.

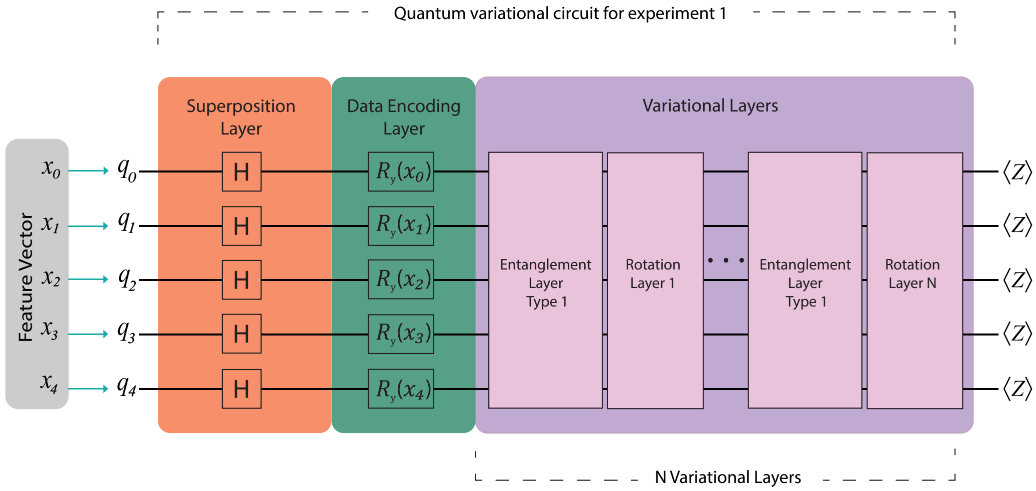

Fig. 4. Quantum variational circuit architecture used in experiment 1.

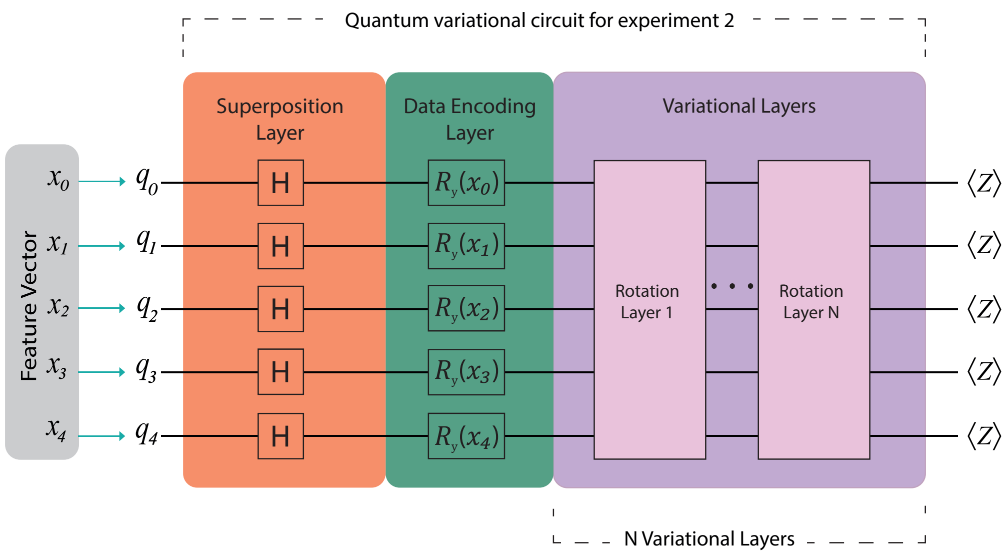

Fig. 5. Quantum variational circuit architecture for experiment 2.

In the third experiment, the effect of removing the superposition layer used in the first experiment and using the Hadamard gate before the data encoding layer was investigated. The circuit architecture used in this experiment is given in Figure 6.

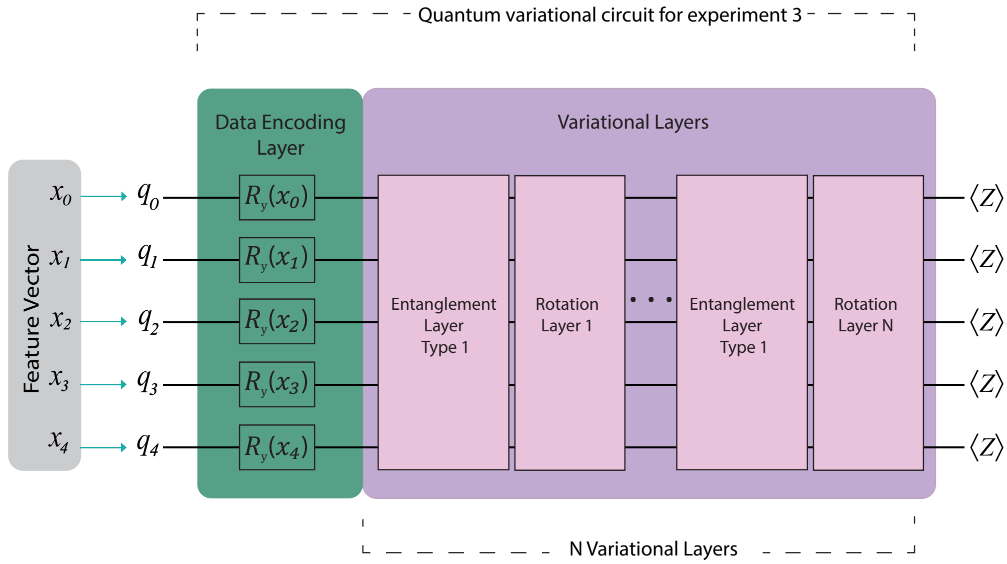

Fig. 6. Quantum variational circuit architecture for experiment 3.

In the fourth experiment, the entanglement layer type 2 circuit given in Figure 3 was used instead of the entanglement layer type 1 in the variational layer used in the first experiment. The entanglement layer types 2 circuitry entangles all qubits directly with each other. In this experiment, the effect of two different entanglement sublayer relative to each other was investigated. The circuit architecture used is given in Figure 7.

In the last experiment, in the variational layer in the variational circuit model given in Figure 2, a variational sublayer array consisting of rotation layer-entanglement layer type 2 layer was defined. In this experiment, the effect of a different variational layer structure was investigated. The circuit architecture used is given in Figure 8.

3.4 Experiment procedure

Pennylane quantum computing simulator [23] was used in the study. In all experiments performed, 9% of the dataset was reserved for training, 1% for validation, and 90% for testing. The Adam stochastic gradient algorithm was used as the classical optimization algorithm. Initially, it started with a large learning rate of 0.1 and since there was no reduction in the validation error for 3 training periods, it was dynamically reduced by a factor of 0.1 and the minimum learning rate was set to 0.00001. The number of training periods (epoch) was taken as 30 for all experiments and the mean square error function was used as the error function.

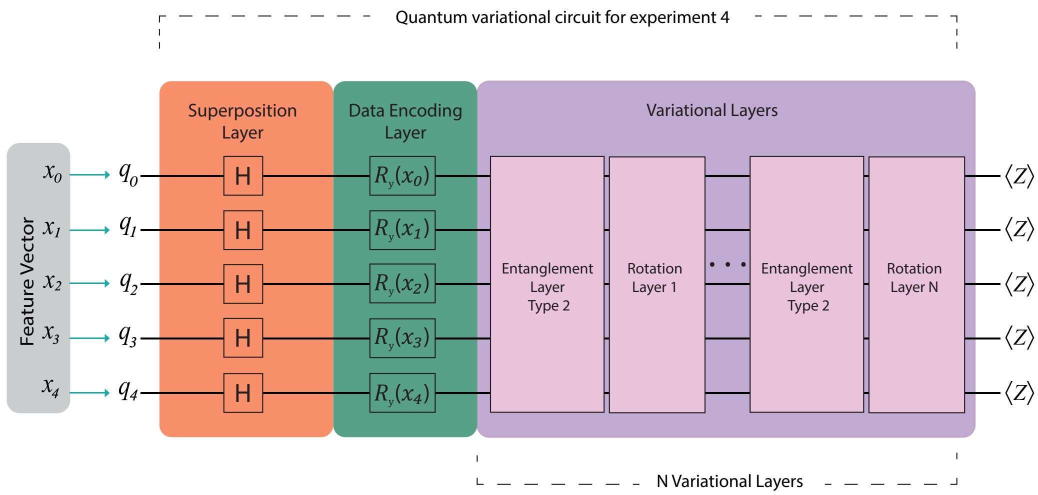

Fig. 7. Quantum variational circuit architecture for experiment 4.

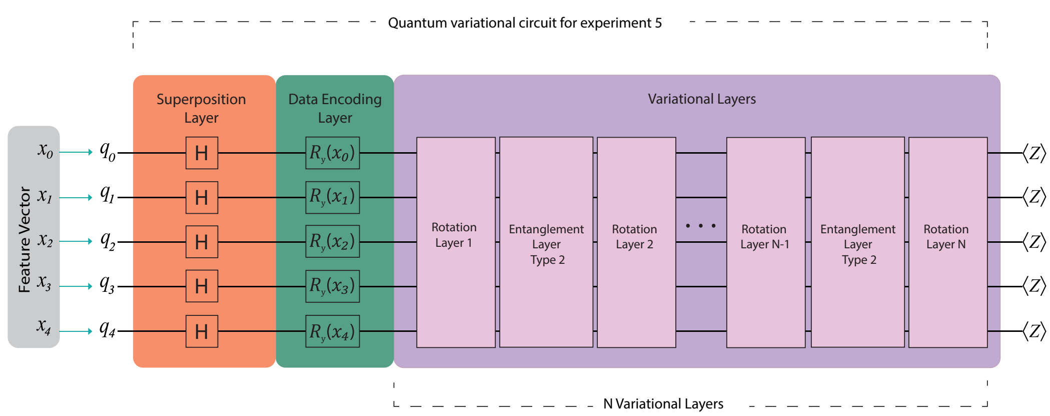

Fig. 8. Quantum variational circuit architecture used in experiment 5.

4 Discussion

The variation of the training and validation stages of the 1, 2, 3, 4 and 5th experiments carried out in the study according to the number of training periods is given in Figure 9-13. In addition, the test results of each experiment are shown in Table 2.

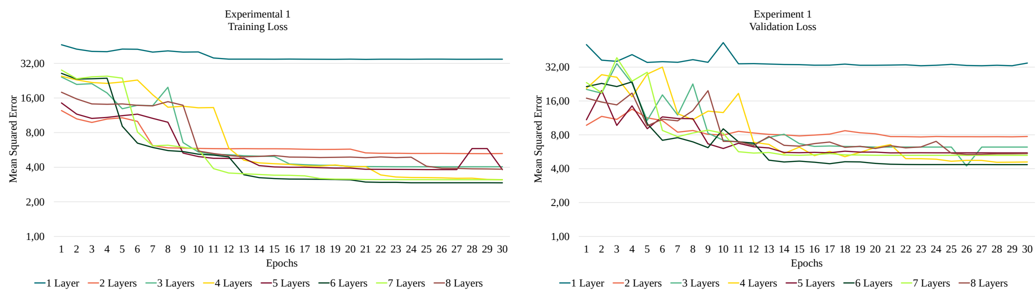

Fig. 9. Variation of training loss and validation loss for Experiment 1 with respect to the number of training periods.

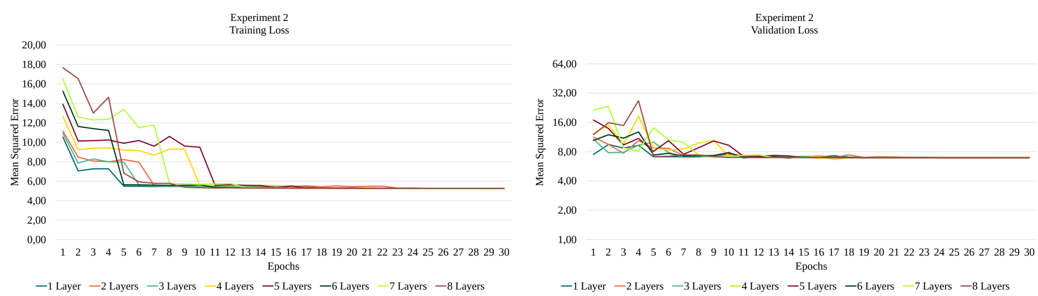

Fig. 10. Variation of training loss and validation loss for Experiment 2 with respect to the number of training periods.

It can be seen that the training and validation errors decrease as the number of variation layers increases in experiment 1 from Figure 9. Looking at Table 2 for Eexperiment 1, there is no regular increase or decrease in test results as the number of layers increases. However, the lowest error was obtained in the 6-layer model. This error is the lowest error value obtained among all experiments performed. The increase in error values after the 6th layer indicates that the distance of the model to the data points increases, as can be understood from the R^2 score.

As the number of layers increased in experiment 2, the training and validation error decreased until the 11th training period, but no improvement was observed thereafter. As can be seen in Table 2, the increase in the number of layers did not cause a downward effect on the test errors. In this context, it is observed that the use of the entanglement layer according to Table 2 has a great effect on the test results.

As the number of layers increased in experiment 3, a decrease in training and validation error was observed, but when compared to experiment 1, it was observed that this decrease

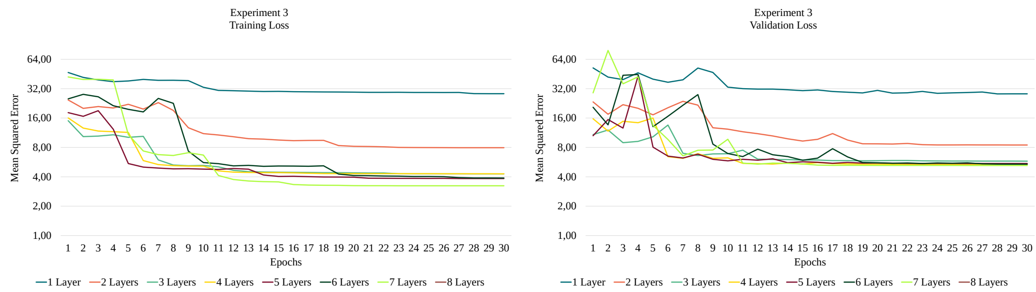

Fig. 11. Variation of training loss and validation loss for Experiment 3 with respect to the number of training periods.

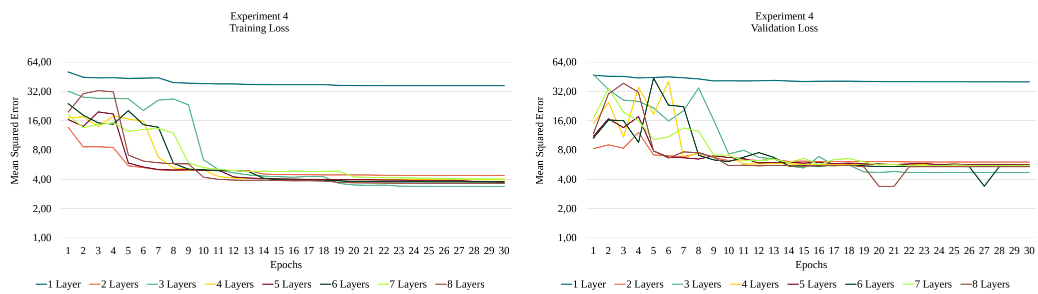

Fig. 12. Variation of training loss and validation loss for Experiment 4 with respect to the number of training periods.

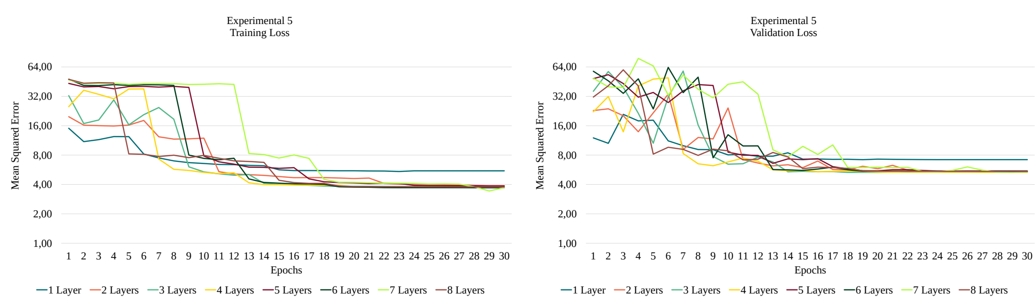

Fig. 13. Variation of training loss and validation loss for Experiment 5 with respect to the number of training periods.

was not a remarkable decrease. However, when the test error results in Table 2 are compared, it can be said that experiment 1 showed higher performance with lower error up to the eighth layer compared to experiment 3. In this context, it has been determined that the use of the superposition layer before the data coding layer allows us to use less variational layers.

In experiment 4, as the number of layers increases, the training and validation error is similar to experiment 1. However, looking at Table 2, it is seen that the error results obtained in experiment 1 are generally lower.

In experiment 5, on the other hand, as the number of layers increased, there was a decrease in the training and validation errors, but according to the test results in Table 2, it is seen that it has a higher error result when compared to experiments 1 and 2.

5 Conclussion

As quantum computer technologies progress, quantum machine learning develops in parallel and is applied to new areas. It is known that by using fast processing power of quantum computing on variational quantum circuit algorithms, which is a hybrid modelling, better results are obtained than classical machine learning. In this study, the effect of two important quantum properties, entanglement and superposition, on hybrid quantum-classical machine learning performance using small-size training data for weather forecasting is investigated. The use of the entanglement layer between the variational layers has made significant improvements on the circuit performance (see Table 2). In terms of fault tolerance performance, the effect of the indirect (transfer) entanglement layer is better than the direct entanglement effect (see Table 2). The use of the superposition layer before the data encoding layer enables the use of fewer variational layers (see Table 2). It is suitable for near-term quantum technologies in terms of computational complexity and fault tolerance.

Data and code availability

https://drive.google.com/file/d/1u9HmH9KEVW2gmTetD5DEl5rVB2HThky1

Acknowledgements

This work was produced within the framework of Beşir Oğur's doctoral thesis. We would like to thank anonymous referees for valuable suggestions.

References

1. Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.
2. Emmanuel Desurvire. *Classical and quantum information theory: an introduction for the telecom scientist*. Cambridge university press, 2009.
3. David Deutsch. Quantum theory, the church–turing principle and the universal quantum computer. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 400(1818):97–117, 1985.
4. Lov K Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 212–219, 1996.
5. Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41(2):303–332, 1999.
6. B Jack Copeland. The church-turing thesis. 1997.

7. Gérard Dreyfus. *Neural networks: methodology and applications.* Springer Science & Business Media, 2005.
8. Howard Demuth and Mark Beale. Neural network toolbox user's guide. 2000.
9. Ethem Alpaydin. *Introduction to machine learning.* MIT press, 2020.
10. Maria Schuld and Francesco Petruccione. *Supervised learning with quantum computers*, volume 17. Springer, 2018.
11. Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.
12. Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. An introduction to quantum machine learning. *Contemporary Physics*, 56(2):172–185, 2015.
13. Marcello Benedetti, Erika Lloyd, Stefan Sack, and Mattia Fiorentini. Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, 4(4):043001, 2019.
14. Maria Schuld, Ville Bergholm, Christian Gogolin, Josh Izaac, and Nathan Killoran. Evaluating analytic gradients on quantum hardware. *Physical Review A*, 99(3):032331, 2019.
15. Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 2021.
16. John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 2018.
17. AV Frolov. Can a quantum computer be applied for numerical weather prediction? *Russian meteorology and hydrology*, 42(9):545–553, 2017.
18. Ashkan Safari and Amir Aminzadeh Ghavifekr. Quantum neural networks (qnn) application in weather prediction of smart grids. In *2021 11th Smart Grid Conference (SGC)*, pages 1–6. IEEE, 2021.
19. Gurwinder Singh and VP Singh. *Quantum neural network application for weather forecasting.* PhD thesis, 2009.
20. Dmytro Zubov, Francesco Volponi, and Mahdi Khosravy. D-wave quantum computing ising model: A case study for the forecasting of heat waves. In *2015 International Conference on Control, Automation and Information Sciences (ICCAIS)*, pages 149–152. IEEE, 2015.
21. Maria Schuld and Nathan Killoran. Quantum machine learning in feature hilbert spaces. *Physical review letters*, 122(4):040504, 2019.
22. Pinaki Sen, Amandeep Singh Bhatia, Kamalpreet Singh Bhangu, and Ahmed Elbeltagi. Variational quantum classifiers through the lens of the hessian. *Plos one*, 17(1):e0262346, 2022.
23. Ville Bergholm, Josh Izaac, Maria Schuld, Christian Gogolin, M Sohaib Alam, Shahnawaz Ahmed, Juan Miguel Arrazola, Carsten Blank, Alain Delgado, Soran Jahangiri, et al. Pennylane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*, 2018.
24. Kosuke Mitarai, Makoto Negoro, Masahiro Kitagawa, and Keisuke Fujii. Quantum circuit learning. *Physical Review A*, 98(3):032309, 2018.

Table 2. Mean square error, R^2 score, and standard deviation results on test data for all experiments.

# Layer	MSE	R^2 Score	Standard Deviation
		Experiment 1	
1	33.7237	0.4171	46.7110
2	5.4962	0.9050	7.2614
3	4.2573	0.9264	7.3165
4	3.242	0.9440	7.3420
5	4.0139	0.9295	7.3265
6	3.0793	0.9468	7.3356
7	3.3926	0.9414	7.3463
8	4.1031	0.9291	7.4942
		Experiment 2	
# Layer	MSE	R^2 Score	Standard Deviation
1	5.3956	0.9067	7.2827
2	5.3896	0.9068	7.2925
3	5.402	0.9066	7.2985
4	5.4459	0.9059	7.2886
5	5.4136	0.9064	7.2869
6	5.3996	0.9067	7.2958
7	5.4203	0.9063	7.2854
8	5.4148	0.9064	7.2896
		Experiment 3	
# Layer	MSE	R^2 Score	Standard Deviation
1	28.3449	0.5101	48.5700
2	8.4131	0.8546	7.1925
3	4.4781	0.9226	7.2786
4	4.4672	0.9228	7.3231
5	4.084	0.9294	7.3505
6	4.0592	0.9298	7.3525
7	3.6647	0.9367	7.3556
8	3.99	0.9311	7.3448
		Experiment 4	
# Layer	MSE	R^2 Score	Standard Deviation
1	36.3256	0.3721	42.1010
2	4.5752	0.9209	7.3308
3	3.6155	0.9375	7.3252
4	4.1919	0.9275	7.4157
5	4.0221	0.9305	7.3322
6	4.0032	0.9308	7.3472
7	4.238	0.9267	7.3423
8	3.9611	0.9315	7.3467
		Experiment 5	
# Layer	MSE	R^2 Score	Standard Deviation
1	5.6709	0.9020	4.2380
2	4.1363	0.9285	7.3147
3	4.0384	0.9302	7.3366
4	3.9693	0.9314	7.3562
5	4.096	0.9292	7.3530
6	4.0213	0.9305	7.3064
7	3.9637	0.9315	7.4283
8	4.0156	0.9306	7.3475