JESTSA JOURNAL
JOURNAL OF ENGINEERING ADVANCES AND TECHNOLOGIES FOR SUSTAINABLE APPLICATIONS
Volume 1, Issue 1 (January, 2025)
Print ISSN: 3062-5629 Online ISSN: 3062-5637
Open Access
https://jeats.journals.ekb.eg/

Weather Feature Selection for Robust and Optimized Energy Load Prediction

Mohsen Tavakolian, Hamidreza Zareipour
Department of Electrical and Software Engineering, University of Calgary, Calgary, AB, Canada
mohsen.tavakolian@ucalgary.ca

doi: 0.21608/jeatsa.2025.427790
Received 29-09-2024
Revised 16-10-2024
Accepted: 1-11-2024
Published: Jan-2025

Print ISSN: 3062-5629
Online ISSN: 3062-5637

Abstract- This study explores the impact of weather features on short to medium electricity load prediction across diverse geographical locations. Using hourly load data, we evaluated the effectiveness of several feature selection methods, including Mutual Information (MI), Principal Component Analysis (PCA), Lasso, and Heatmap correlation. We benchmarked these feature selection methods with a hybrid deep learning model to investigate the impact of choosing the correct multiple weather features instead of temperature. For this purpose, we practiced different combinations of temperature, relative humidity, dew point, air pressure, and wind speed benchmarked with the base case of single feature (temperature). The comparison was performed based on the load prediction accuracy improvement. The hybrid Artificial Neural Network (ANN) and temporal setup was implemented to predict energy loads across four different lead times (1, 6, 12, and 24 hours ahead) to not only study the feature selection methods, but also its behavior at different lead time predictions. Moreover, this study inspected the dynamic behavior of weather features selection by location to explore the need for location-specific feature engineering. All steps and theories were examined by a real-world dataset from a location of interest and the result was visualized across the geographical extent, offering insights into the spatial variability of feature importance. Future work will investigate the development of lead-time-specific models to further improve load prediction accuracy. This research highlights the importance of an in-depth inspection of weather feature selection and its dynamic behavior for enhancing energy load forecasting models.

Keywords- *Energy load prediction, Feature selection, Weather parameters, Machine learning, Artificial Neural Network (ANN), Temporal models.*

I. INTRODUCTION

Effective energy load forecasting is essential for optimizing grid operations, particularly as renewable energy sources become more integrated into power systems. Accurate load prediction helps ensure the stability and efficiency of energy grids. Traditionally, temperature has been the primary weather variable used in forecasting models due to its significant impact on energy demand [1]. The concept of degree-days, for instance, is widely used to measure the influence of temperature on electricity consumption [2-4]. However, the relationship between weather and energy demand is more complex than temperature alone, and other meteorological factors can also play a crucial role in predicting loan [5].

This research aims at enhancing the reliability and precision of energy load prediction tools through machine learning techniques. As energy systems evolve, the need for accurate forecasting becomes increasingly critical, especially as we move towards more sustainable and decarbonized grids. The ability to predict energy demand accurately is not just a technical challenge but a necessary step for enabling smarter grid operations and integrating renewable energy sources effectively.

Our approach begins by examining various feature selection methods and benchmarking them with a hybrid machine learning model composed of the temporal nature of load and weather data with artificial neural network (ANN) structure. This investigation aims to identify optimal and practical methods and to study the impact of different combinations of weather features, i.e. temperature, relative humidity, air pressure, dew point, and wind speed, on the reliability and robustness of load prediction. The study also explores the potential correlation between weather feature selection and geological location, as well as its sensitivity to the lead time range.

This paper represents one of the initial steps in improving the accuracy of conventional models for energy load prediction by utilizing real-world data. By digging deeper into weather-related input parameters and their temporal

JESTSA JOURNAL
JOURNAL OF ENGINEERING ADVANCES AND TECHNOLOGIES FOR SUSTAINABLE APPLICATIONS
Volume 1, Issue 1 (January, 2025)
Print ISSN: 3062-5629 Online ISSN: 3062-5637
Open Access
https://jeats.journals.ekb.eg/

depth, we aim to uncover insights that can lead to more accurate predictions, particularly over extended periods.

The findings presented here serve as the foundation for future work. Upcoming publications will test our models on various datasets, refine the methodology to accommodate the growing complexity of energy markets, and address challenges posed by sustainability goals, decarbonization efforts, and extreme weather conditions. This work ultimately seeks to contribute to the development of more reliable, scalable, and adaptable forecasting tools that can meet the demands of future energy systems.

II. LITERATURE SURVEY AND BACKGROUND

The field of energy load forecasting has long recognized the importance of weather variables, primarily focusing on temperature due to its direct influence on heating and cooling demands. However, recent studies have started to shed light on the significance of incorporating additional weather features to improve prediction accuracy. This evolving understanding is exemplified in the work of Maia-Silva et al. 2020 [6], who emphasized that while temperature is a crucial factor, it alone is insufficient to capture the full spectrum of climate sensitivity in energy demand. They highlighted the critical role of humidity in residential space cooling demand, demonstrating that neglecting this factor can lead to significant underestimations of future demand, especially under warming scenarios. This finding underscores the necessity of a more comprehensive approach to weather data inclusion, one that accounts for humidity and other climatic stressors.

Building on this notion, Beccali et al. 2007 [7] introduced a forecasting model based on an Elman artificial neural network, focusing on the suburban area of Palermo, Italy. Their work reinforced the importance of the Humidex index, which integrates outdoor air temperature and humidity to evaluate household electricity consumption influenced by HVAC appliances. Despite their advancements, their study was geographically limited and did not provide a detailed comparison of various weather features across different locations.

Mirasgedis et al. 2004 [8] took a further step by utilizing both primitive (relative humidity) and derived (heating and cooling degree-days) meteorological parameters to analyze electricity consumption over a decade. While their medium-term demand estimation highlighted the relevance of these weather variables, their approach was constrained by the use of monthly and yearly average consumption data, lacking the granularity needed for more precise forecasting.

Friedrich et al. 2013 [9] explored the decomposition of electricity load into weather-independent and weather-dependent portions in Abu Dhabi, UAE, identifying key weather drivers such as temperature, humidity, solar irradiance, and wind speed. Their regression model effectively segregated these influences but was again limited to a single location, raising questions about the generalizability of their findings.

Wang et al. 2004 [10] and Rastogi et al. 2021 [1] have also explored this topic by focusing on the segmented relationship between electricity load and weather features in different U.S. regions. They confirmed the critical role of humidity alongside temperature in predicting energy demand, yet their analyses were still predominantly limited to these two variables, overlooking the potential contributions of other weather factors.

Ihara et al. 2008 [11] provided additional insights into the sensitivity of electricity consumption to temperature and humidity, using data from Tokyo and various U.S. regions, respectively. Their methodologies highlighted the need for high-precision estimates but did not fully explore the combined effects of multiple weather variables.

These studies collectively advance our understanding of the intricate relationship between weather and energy demand, but they also highlight significant gaps. Many focused on limited geographical areas or specific datasets, lacked a comprehensive comparison of diverse weather features, and often used coarse temporal resolutions. These limitations underscore the need for a more detailed and expansive approach.

Our research addresses these gaps by incorporating a wider range of weather features and analyzing data at a finer temporal (hourly) and spatial (city-level) resolution utilizing real-world data. We aim to quantify the importance of each weather parameter, providing a more nuanced and comprehensive understanding of their roles in energy load forecasting. By building on the foundational work of these previous studies, we seek to enhance the robustness and accuracy of predictive models, ultimately contributing to more resilient and efficient grid operations.

III. METHODOLOGY

In this study, we explored how various weather parameters influence energy load prediction, inspecting whether these effects differ across locations. Temporal variables such as load history, weather patterns, and the calendar effect were also analyzed for their impact on model robustness. Data from the weathersource.com and Alberta Electric System Operator (AESO), and calendar sources were preprocessed to align timestamps and prepare for predictive modeling. To identify the most relevant features, we employed four feature selection methods: Mutual Information, Principal Component Analysis (PCA), Lasso Regression, and Correlation Heatmaps, evaluating their effectiveness across multiple cities in Alberta. A hybrid

JESTSA JOURNAL
JOURNAL OF ENGINEERING ADVANCES AND TECHNOLOGIES FOR SUSTAINABLE APPLICATIONS
Volume 1, Issue 1 (January, 2025)
Print ISSN: 3062-5629 Online ISSN: 3062-5637
Open Access
https://jeats.journals.ekb.eg/

model (time-series temporal matrix in an ANN structure) was utilized to run several scenarios and benchmark the preliminary evaluation from the aforementioned feature selection methods. These steps provided a comprehensive analysis of key weather features for having an optimized and reliable model for accurate load prediction.

III.I. Data and Pre-Processing:

Using real-world load and weather data, we conducted an in-depth analysis of how different weather parameters impact energy load prediction and if different locations show different behavior .

Unlike most existing models that predominantly rely on temperature, our research investigated the effect of inclusion of other weather features and if it can improve load prediction accuracy. The research also studied the potential variation of this behaviour across different cities in the area of study.

Furthermore, to optimize the robustness of the developed ML model, we investigated the effective use of temporal variables such as load and weather, as well as a modified variable associated with hour of the day.

To practice our theories, the following data sources were utilized:

Load Data: The load data was obtained from the Alberta Electric System Operator (AESO) for 42 major cities in Alberta, spanning from January 2011 to October 2023. The data, recorded hourly and measured in megawatt-hours (MWh), is comprehensive and contains no missing values [12]. The geographical distribution of these cities is illustrated in **Error! Reference source not found.**.

Weather Data: Weather data corresponding to the same period and frequency as the AESO load data was sourced from weathersource.com [13]. The available features based on the local timestamp include:

- Cloud Coverage: The fraction of the sky covered by clouds.
- Dew Point: The temperature at which air becomes saturated with moisture.
- Feels Like: The apparent temperature considering humidity and wind.
- Freezing Rain Flag: Indicates the occurrence of freezing rain (boolean).
- Heat Index: The apparent temperature considering humidity.
- Ice Pellets Flag: Indicates the occurrence of ice pellets (boolean).
- Air Pressure: The force exerted by the atmosphere at a given point.
- Precipitates: Any form of water, liquid or solid, falling from the sky.
- Pressure Tend: The change in air pressure over time.
- Provisional Flag: Marks provisional data subject to revision (boolean).
- Solar Radiation: The intensity of sunlight reaching the ground.
- Rain Flag: Indicates the occurrence of rain (boolean).
- Relative Humidity: The amount of moisture in the air relative to what the air can hold at that temperature.
- Snow Flag: Indicates the occurrence of snow (boolean).
- Snow Fall: The amount of snow that has fallen.
- Specific Humidity: The mass of water vapor per unit mass of air.
- Temperature: The degree of heat present in the atmosphere.
- Visibility: The distance one can clearly see.
- Wet Bulb: The lowest temperature air can reach by evaporative cooling.
- Wind Chill: The perceived decrease in air temperature felt by the body due to wind.
- Wind Direction: The direction from which the wind is blowing.
- Wind Speed: The rate at which air is moving horizontally.

Calendar Data: Python's calendar library was employed to identify working and non-working dates, contributing to the temporal analysis.

Python was used to clean and integrate these data sources into a unified format, either as CSV files or data frames. The integration process ensured that the timestamps of the load and weather data were aligned, facilitating seamless analysis.

III.II. Feature Selection and Analysis

Feature selection is a crucial step in machine learning that involves identifying and choosing the most relevant features from a dataset for building a predictive model. The process serves multiple purposes, including enhancing model performance, reducing overfitting, improving model interpretability, and decreasing computational costs [14]. By selecting the most impactful variables, feature selection ensures that the machine learning model focuses on the key information while discarding irrelevant or redundant data that may otherwise introduce noise or confusion [15].

In predictive modeling, especially with large datasets, it's common to encounter many input variables, some of which may not significantly contribute to the target prediction. For instance, in time series forecasting, like energy load prediction, where weather conditions and calendar variables are used as inputs, not all weather parameters will have a meaningful influence on the model. Feature selection techniques are employed to filter out the less relevant features, thereby reducing the dimensionality of the dataset and improving the overall efficiency of the learning process [5].

One widely used approach to feature selection is statistical-based methods, which involve evaluating the relationship between each input variable and the target variable using statistical measures. These techniques assess the strength of each feature's contribution to the model by analyzing the correlation or association with the output variable (in our case, energy load). This allows practitioners to rank the features based on their importance and select those with the strongest relationship to the target [16].

However, choosing the appropriate statistical measures for feature selection depends heavily on the type of data. For example, if both the input and output variables are continuous, correlation coefficients such as Pearson's correlation can be applied. For categorical variables, measures like mutual information or Chi-square may be more suitable. The challenge lies in selecting the right statistical technique for the dataset in question, ensuring that the best set of features is identified to build an optimized model [17].

Feature selection offers several key advantages [18]:

Enhanced Model Performance: By focusing on the most relevant features, the model becomes more accurate and efficient, as it can better learn from the data.

Reduced Overfitting: Including too many irrelevant features can lead to overfitting, where the model performs well on training data but fails to generalize to unseen data. Feature selection helps prevent this by removing noise.

Improved Interpretability: A model with fewer features is not only easier to interpret but also more transparent. This is crucial when explaining the model's behavior to stakeholders or for regulatory purposes.

Faster Training and Inference: With a smaller set of features, the time required for both training the model and making predictions is significantly reduced, making the process more computationally efficient.

In this research, the goal was to develop a robust model for energy load prediction by selecting the most impactful weather features. While temperature has traditionally been the main focus in load prediction models, our work aimed to explore the inclusion of additional weather parameters such as relative humidity, wind speed, outside pressure, and dew point. These features, although less commonly used, were hypothesized to have a significant effect on energy consumption depending on the geographic location.

From the comprehensive set of weather features, temperature, relative humidity, dew point, wind speed, and air pressure were selected for detailed analysis. These five weather parameters were chosen based on their established physical relationships with energy consumption as justified below. They represent a comprehensive set of factors that can significantly influence energy demand, making them suitable for detailed analysis in this research. While other weather parameters may also play a role, these five are considered to have the most direct and significant impact on energy load, providing a solid foundation for feature selection in this domain.

Reasons that temperature, relative humidity, dew point, wind speed, and air pressure were chosen for feature analysis:

Temperature:

- Direct correlation: Temperature is a primary driver of energy consumption, especially for heating and cooling systems. It is the most important feature, if not the only one, in almost all load prediction models. Higher temperatures typically lead to increased cooling loads, while lower temperatures necessitate increased heating loads.
- Physical basis: Temperature directly influences the rate of heat transfer between buildings and their surroundings.

Relative Humidity:

- Impact on HVAC systems: High humidity can increase the cooling load by making it more difficult for air conditioners to dehumidify the indoor air.
- Comfort levels: Humidity affects human comfort levels, influencing heating and cooling demands.

Dew Point:

- Temperature-humidity relationship: Dew point is a measure of the amount of water vapor in the air. It's closely related to temperature and humidity, making it a valuable parameter for understanding the overall atmospheric conditions.

- Condensation and energy consumption: High dew points can lead to increased condensation and moisture-related issues in buildings, affecting energy consumption.

Wind Speed:

- Heat transfer: Wind can influence heat transfer between buildings and the environment. Higher wind speeds can increase heat loss in colder weather and reduce cooling loads in warmer weather.
- Ventilation: Wind can affect the effectiveness of ventilation systems, which can impact energy consumption.

Air Pressure:

- Weather patterns: Air pressure is associated with various weather patterns, including storms and temperature changes. It can indirectly influence energy consumption by affecting heating and cooling needs.
- Ventilation: Changes in air pressure can affect the efficiency of ventilation systems.

Having chosen the features of interest, the next step is the method to analyze their importance and benchmark the result. In this study, four feature selection methods were utilized to determine the most important weather and temporal features for energy load prediction: Mutual Information, Principal Component Analysis (PCA), Lasso Regression, and Correlation Heatmaps. Each method offers distinct advantages and challenges for handling feature selection in machine learning models.

III.II.I. Mutual Information (MI) Method

Mutual Information (MI) was first proposed by Shannon 1963 [19] in the context of information theory and further developed by Cover and Thomas 2006 [15]. MI measures the amount of shared information between two variables, quantifying their dependency. In the context of feature selection, mutual information is used to measure the dependency between input features and the target variable.

MI has the advantage of capturing non-linear dependencies between features and the target, making it suitable for complex datasets where linear correlation may not be sufficient. In our application of energy load prediction, mutual information helps us understand how each weather feature contributes to the variability in energy load.

Pros:

- Captures non-linear relationships between features and the target.
- Does not assume any specific model structure.

Cons:

- Sensitive to the number of samples in the dataset.
- Requires discretization or kernel density estimation, which can introduce bias.

III.II.II. Principal Component Analysis (PCA)

PCA was first introduced by Pearson in 1901 [20] and later generalized by Hotelling 1933 [21]. PCA is a dimensionality reduction technique that identifies linear combinations of features that capture the maximum variance in the data. It does this by transforming the input features into new orthogonal components called principal components, which are ranked by the amount of variance they explain.

In load prediction, PCA allows us to reduce the dimensionality of weather and temporal features, retaining only the components that contribute the most to load variability. This aids in simplifying the model without sacrificing predictive power.

Pros:

- Reduces dimensionality, simplifying the model and mitigating overfitting.
- Identifies the most significant directions of variance in the dataset.

Cons:

- Assumes linear relationships between features.
- Loses interpretability, as principal components are combinations of original features.

III.II.III. Lasso Regression (L1 Regularization)

Lasso (Least Absolute Shrinkage and Selection Operator) was introduced by Tibshirani in 1996 [22]. Lasso is a linear regression technique that introduces L1 regularization, which enforces sparsity in the model by shrinking some coefficients to zero. This effectively performs both feature selection and regularization, helping to mitigate overfitting by reducing the complexity of the model.

In the context of energy load prediction, Lasso selects the most relevant weather and temporal features by shrinking the less important feature coefficients to zero, providing a sparse solution that simplifies interpretation and reduces the risk of overfitting.

Pros:

- Simultaneously performs feature selection and regularization.
- Provides a sparse solution, enhancing model interpretability.

Cons:

- Sensitive to the choice of the regularization parameter λ.
- Can be unstable if features are highly correlated.

III.II.IV. Correlation Heatmap

The concept of correlation, specifically Pearson correlation, was first introduced by Pearson in 1895 [23]. The correlation matrix computes pairwise correlation coefficients between features, quantifying the linear relationships between them. This method is useful for identifying highly correlated features and understanding their direct relationships with the target variable.

For our study, we used correlation heatmaps to visualize the relationships between weather features and energy load. This method provides an intuitive way to identify both positive and negative correlations, enabling us to select features that are strongly related to energy load.

Pros:

- Simple and intuitive to understand.
- Visual representation of feature relationships.

Cons:

- Only captures linear relationships.
- Sensitive to noise in the data.

Each of these methods was leveraged to highlight key weather features that influence energy load prediction, and in combination, they provided a comprehensive understanding of the feature space.

IV. RESULTS

To evaluate the performance of different feature selection methods for energy load prediction, we randomly selected seven cities across Alberta: Calgary, Fort McMurray, Cold Lake, Red Deer, and Provost. The four feature selection techniques – Mutual Information (MI), Principal Component Analysis (PCA), Lasso Regression (L1), and Heatmap correlation—were applied to the dataset for each city to rank the most important weather features. The outcome of these experiments is summarized in Table 1.

IV.I. Observations and Performance Comparison:

Mutual Information (MI):

Surprisingly, the MI method returned exactly the same feature ranking for all cities. The top five features were consistently: temperature, dew point, relative humidity, mean sea-level pressure, and wind speed, in that order. While MI is known to capture non-linear relationships between features and the target variable, this consistency across all cities is somewhat unexpected. A location-based behavior in feature importance would be more in line with physical expectations, as energy load should be influenced by unique local weather conditions. This lack of variation is also inconsistent with findings in other studies, such as the work by Maia-Silva et al. (2020), which reported location-dependent variations in weather impact on load prediction.

Principal Component Analysis (PCA):

PCA, while aiming to capture variance in the data, exhibited similar behavior to MI. The top features identified were mostly the same across cities, except for a swap between relative humidity and air pressure. This suggests that PCA, despite its dimensionality reduction capabilities, may be too generalized to capture city-specific variations in weather's influence on load. The transformed feature space created by PCA also reduces the direct interpretability of the selected features, which further complicates its utility in understanding local weather impacts.

Lasso Regression (L1):

The Lasso method provided more variability in feature importance across cities but sometimes produced rankings that seemed counterintuitive. For instance, in Calgary, Lasso ranked wind speed as the most important feature, while for Edmonton, dew point was ranked highest, with temperature only in third place. These results conflict with physical expectations and prior research, where temperature typically plays a dominant role in energy load predictions. The method's focus on penalizing feature coefficients might sometimes over-penalize features, leading to unexpected rankings. This inconsistency questions the reliability of Lasso for this particular application.

Heatmap (Correlation-Based Method):

In contrast to the other methods, the Heatmap correlation approach provided a more varied and location-specific feature ranking. The top feature for each city varied between temperature, dew point, and relative humidity, which aligns well with physical expectations. For example, temperature and relative humidity were identified as top contributors for load prediction in most cities, with dew point emerging as a secondary feature, which is in line with the role these weather parameters play in energy demand variations. The Heatmap method's output reflects a more logical and legitimate prediction of feature importance, grounded in the actual physical relationships between weather variables and energy load.

IV.II. Preliminary Evaluation

As represented above, MI and PCA don't exhibit the sensitivity to location-specific variations, as proposed by Maia-Silva et. al (2020). Lasso, though flexible, sometimes produces rankings that contradict physical intuition, further reducing confidence in its applicability for this task. For instance, temperature which is, undoubtedly a vital parameter to predict load, is at the bottom of the list for the city of Cold Lake. To provide a more robust and practical approach while minimizing potential bias, we implemented various combinations of features (limited to three at most to better visualize the impact of each feature) for a selection of randomly chosen cities. We then applied a machine learning (ML) model to benchmark the precision and performance of these feature combinations. This benchmarking process is designed to identify the feature selection method that most accurately supports energy load prediction across diverse geographic regions.

Before discussing the results, we first detail the ML setup used to benchmark the following combinations of weather features (along with additional input features, as described in the next section):

- Temperature only
- Combination 1: ['temp', 'dewPt', 'relHum']
- Combination 2: ['temp', 'dewPt', 'mslPres']
- Combination 3: ['temp', 'windSpd', 'mslPres']
- Combination 4: ['temp', 'windSpd', 'relHum']
- Combination 5: ['temp', 'dewPt']
- Combination 6: ['temp', 'relHum']

IV.III. Development of the Predictive Model

To evaluate our hypothesis that incorporating additional weather features and temporal adjustments leads to more robust energy load predictions, we developed a temporal-based Artificial Neural Network (ANN) model. This model leverages multiple input features, capturing both historical and recent data, to predict energy load for the lead time(s) ahead. The input features for the model include:

1. Load History: Temporal data with a lagged time of a selected number of hours, capturing the recent trends and patterns in energy consumption. The depth of number of hours selected is part of the model optimization which will be presented in our next publication.
2. Weather Features: A case dependent input – as discussed above – to practice and benchmark 7 scenarios of different weather features combinations. This feature is also a temporal input and a history of weather features with chosen depth was fed into the model. The depth of this feature was also a model optimization parameter that will be explained in our next publication.
3. Calendar Effect, including:
   a. Hour of the Day: Sine and cosine transformed representations of the hour of the day (elaborated below), ensuring the model accurately captures the cyclical nature of daily energy consumption patterns.
   b. Working/Non-working Days: Binary indicator capturing whether a given day is a working day or a non-working day, reflecting the variation in energy load based on human activities and calendar events.

The Artificial Neural Network (ANN) model employed in this study follows a deep feed-forward architecture, optimized for energy load prediction. The model consists of eight fully connected layers, with ReLU (Rectified Linear Unit) activation functions applied to each layer to introduce non-linearity. The input layer accepts features corresponding to the weather parameters, while the final output layer consists of a single neuron, designed for the regression task of load prediction. [24]

The network architecture begins with a large number of neurons in the first hidden layer (2048), followed by progressively smaller layers with 1024, 512, 256, 128, 64, 32, and 8 neurons, respectively. This design allows the network to capture complex patterns in the input data by progressively reducing the dimensionality of the feature space.

To optimize training, the model uses the Adam optimizer, starting with a learning rate of 0.01, and employs a mean squared error (MSE) loss function, which is standard for regression tasks. A learning rate scheduler adjusts the learning rate dynamically during training. The model was trained for 100 epochs with a batch size equal to the lag size, and validated using a separate validation dataset to monitor performance [25].

The model is designed to predict the energy load for one hour ahead, as the only output feature, which could be repeated in a loop to predict a period of lead time, e.g. 72 hours ahead, providing a short to mid-term forecast aiding grid management and energy planning. The model was used to predict only one step for 1, 6, 12, and 24 hours ahead. It enables us to investigate the impact and effectiveness of weather feature selections on different lead times.

Of the whole dataset, 2011 to 2017 was assigned to training, 2018 to validation, and 2019 to 2020 to test dataset, to accomplish the data splitting step. Afterwards, standardization and normalization were applied to the train, validation, and test datasets based on the training dataset.

To complete the input data, the calendar effect was incorporated by distinguishing between working and non-

JESTSA JOURNAL
JOURNAL OF ENGINEERING ADVANCES AND TECHNOLOGIES FOR SUSTAINABLE APPLICATIONS
Volume 1, Issue 1 (January, 2025)
Print ISSN: 3062-5629 Online ISSN: 3062-5637
Open Access
https://jeats.journals.ekb.eg/

working days, which reflects the varying energy consumption patterns associated with different types of days. On working days, energy demand is typically higher due to increased industrial, commercial, and office activities, while non-working days, such as weekends and holidays, often see reduced consumption as businesses close and residential use becomes more prominent. By integrating the calendar date into the model, we can capture these behavioral shifts in energy usage, which are essential for improving the accuracy of load predictions. Understanding the influence of calendar variations allows the model to adapt to the cyclical nature of energy demand, providing more precise forecasting, especially when predicting loads during holiday seasons or over weekends. Time of day is another critical factor that significantly impacts energy consumption. Typically, energy usage patterns follow daily cycles, with peak and off-peak hours corresponding to human activities such as work, leisure, and sleep. To capture these cyclical patterns, it is essential to represent the 24-hour format in a way that reflects the continuous and cyclic nature of time.

Using the 24-hour format directly in a machine learning model can be problematic. For instance, while 23:00 and 1:00 are just two hours apart, their numerical representations (23 and 1) are far apart, which misrepresents their true relationship. To address this issue and optimize the use of hour of the day, we employed sine and cosine transformations to encode the hour of the day, thereby maintaining the cyclical nature of time .For this purpose, sine and cosine transformations were utilized. The sine and cosine transformations convert each hour into two components that capture the cyclical relationship. The formulae for these transformations are:

$$hour\_sin = \sin(2\pi.hour/24)$$
$$hour\_cos = cos(2\pi.hour/24)$$

By transforming the hour of the day using these sine and cosine functions, we create two continuous features that maintain the cyclical nature of time.

This approach ensures that the model correctly understands the proximity of 23:00 to 1:00 and captures the cyclical daily patterns in energy consumption.

Incorporating these transformed time features into our model enhances its ability to predict energy load by accurately reflecting the underlying temporal dynamics. This transformation aligns the hour-based data with the natural periodicity of daily human activities, leading to more logical and robust forecasting results.

IV.IV. ANN Model Setup and Training

The ANN was trained to predict the load one hour ahead using the following inputs:

- Load history (temporal with lagged time of n hours – the depth of this selection, n, was optimized which will be presented in our next publication)
- Weather history (temporal with lagged time of m hours – the depth of this selection, m, was optimized which will be presented in our next publication)
- Hour of the day (sine and cosine transformed)
- Working/non-working days

We incorporated the time-series nature of load and weather data to be utilized as input for an ANN structure. Simply put, the whole input matrix can be described by the following formula:

$$\hat{y}_{t+1} = \sigma\left(\sum_{i=1}^{n} \omega_i . x_{t-i} + \sum_{k=1}^{p}\sum_{j=1}^{m} weather\_feature_k + \omega_{hs}.hour\_sin + \omega_{hc}.hour\_cos + \omega_c.workingStatus + b\right)$$

Where:

$\hat{y}_{t+1}$ is the predicted load at time $t+1$, n is depth of load history selected as input, p is total number of weather features selected, m is depth of weather history selected as input, x_{t-1} is the load at time $t-i$ for $i = 0, 1, 2, ...$, $hour\_sin$ and $hour\_cos$ are the hour (time) of the day (24-hour format) transformed by sine and cosine transformers, respectively, $workingStatus$ illustrates if the date is a working or non-working calendar date (1 for working and 0 for non-working including weekends and stat holidays), $\omega_{<index>}$ are the weights, b is the bias, and σ is the activation function - ReLu in our case: $ReLu(z) = \max(0, z)$.

This is technically like a hybrid model that incorporates the temporal (time-series) nature of the data and ANN structure.

The learning process involved minimizing the RMSE (Root Mease Square Error) loss function to adjust the weights and biases.

To predict the load for a certain number of lead hours ahead, we can use the forecasts recursively:

1. Predict the load for the next hour.
2. Use the predicted load as the actual load at time *t*.
3. Repeat steps 1 and 2 for 72 iterations.

This iterative process allows us to extend the prediction horizon up to the desired number of hours by leveraging the model trained for 1-hour ahead predictions.

IV.V. Scenario Analysis

Using the machine learning (ML) setup described earlier, we conducted a comprehensive analysis by running 112 scenarios. These scenarios spanned across four randomly chosen cities (Calgary, Fort McMurray, Cold Lake, and Red Deer), seven different weather feature combinations, and four lead times (1, 6, 12, and 24 hours ahead). For each scenario, the model was tasked with predicting the load for 50 randomly selected data points from the test dataset. The precision of these predictions was calculated, averaged, and then benchmarked for performance analysis

Focusing on the one-hour-ahead prediction in a clear pattern emerges. With the exception of Provost, one of the weakest predictions consistently occurs when using only a single weather feature (temperature). Similarly, with the same exception, the most accurate predictions are obtained when temperature and dew point are selected, along with either relative humidity or air pressure. This pattern, however, does not hold uniformly for predictions at other lead times.

Looking at all lead time predictions, it can be realized that while Calgary – a major city – maintains its dependency on multiple weather features, other cities such as Cold Lake and Red Deer show a reduced sensitivity to multiple features, suggesting that in some cases, temperature alone becomes more relevant. Nevertheless, with the exception of Red Deer, the load predictions for the other cities generally improve when additional weather features, aside from temperature, are included in the model.

One key observation across all lead times is that adding more features is not always beneficial. For instance, in many cases, the inclusion of wind speed in the weather feature set tends to destabilize the predictions, making them more fragile. This indicates that more features do not always enhance model performance, and in some cases, might introduce noise, detracting from predictive accuracy.

An important takeaway is that the model's sensitivity to feature combinations varies not only across locations but also across different lead times. This suggests that a more practical approach could involve developing individually trained ML models for different lead times or at least for distinct lead time ranges. This concept warrants further investigation and is currently being explored by our research group for future publication.

However, if simplicity is preferred, we can average the precision across the four lead times to obtain a more general performance benchmark. This aggregate analysis aligns most closely with the one-hour-ahead lead time results, and thus can be qualitatively used to rank feature combinations.

To address the key question of this research—determining the most effective method for weather feature selection—we compared the performance of the ML model with the feature selection methods previously discussed. It is essential to note that the correlation methods we explored primarily assess row-level data (i.e., the dependency of the output on a single row, which corresponds to a one-hour-ahead prediction). Therefore, a fair comparison involves benchmarking the feature selection methods against the one-hour lead time results produced by the ML model.

From a physical perspective, we have already discounted the reliability of the Lasso method due to its inconsistent feature rankings. Analyzing the relative error plot for one-hour-ahead predictions, we observe minimal differences between combinations 1 and 2 (the selection of temperature and dew point, along with either relative humidity or air pressure). Therefore, our focus shifts to the two-feature combinations—combinations 5 (temperature and dew point) and 6 (temperature and relative humidity)—to compare them with the top two prioritized features selected by the feature selection methods.

For Calgary, combination 5 (temperature and dew point) yields a lower error than combination 6 (temperature and relative humidity), aligning well with the feature rankings suggested by the Heatmap method. Conversely, for Cold Lake, combination 6 outperforms combination 5, which also agrees with the Heatmap results. In contrast, MI and PCA fail to demonstrate the expected location-specific sensitivity, offering identical feature importance across cities, which further undermines their suitability for this task.

In conclusion, based on the analysis of our dataset, the Heatmap method offers the most reliable dependency calculation for feature selection, particularly in its ability to account for geographic and temporal variations in weather-load relationships.

IV.VI. Mapping and Visualization

, with a clearer understanding of the importance of weather features and the identification of the optimal feature selection method, we extended our analysis to visualize these patterns across the entire dataset by mapping them onto a geographic representation of Alberta. To achieve this, we selected the top five weather features—temperature, relative humidity, dew point, air pressure, and wind speed—and applied the Heatmap method to prioritize these features for each city in the dataset.

A distinct pattern emerges from the map. While temperature consistently plays the most significant role in load prediction across Alberta, relative humidity becomes notably more important in the southern part of the province and parts of the western region compared to other areas. This trend is less pronounced in the northern regions, but still observable. Ignoring humidity in the southern and western parts of Alberta may introduce significant errors in energy load prediction, as these areas show a greater dependency on this feature.

This visualization highlights the regional variations in feature importance, emphasizing the need for tailored models that account for these geographic differences when predicting energy load. The implications of this finding are clear: incorporating localized weather patterns into energy load models is crucial for improving prediction accuracy, especially in regions where factors such as humidity exhibit a stronger influence.

In other words, some key observations emerged from this analysis could be summarized as:

- Location-Based Variability: The impact of weather features on energy load is highly location-dependent, with different cities displaying distinct patterns. This variability underscores the importance of localized analysis in energy load forecasting.
- Behavioral Trends: Despite the location-based differences, certain behavioral trends are evident across the province:
  - Temperature: Temperature consistently has the most significant impact on energy load across all cities. However, this impact is notably larger in the eastern part of Alberta. This region experiences colder winters, likely increasing the demand for heating and thus amplifying the influence of temperature on energy consumption.
  - Humidity and Dew Point: As we move west and south within Alberta, the importance of humidity and dew point increases. These regions typically have milder winters and higher humidity levels, making these factors more relevant in determining energy demand, particularly for cooling during the warmer months.

V. SUMMARY AND CONCLUSION

The primary objective of this study was to identify the critical weather features required to construct a robust model for energy load prediction. Our findings revealed that relying solely on temperature is inadequate for accurate load forecasting. Instead, the inclusion of additional weather features significantly improves prediction accuracy, with the specific features required varying by geographical location. Using data from Alberta, provided by AESO, and weather parameters from Weathersource.com, we observed clear geographic dependencies in the impact of weather features on energy load, highlighting the need for location-specific feature engineering.

To address this, we applied a hybrid Artificial Neural Network (ANN) and temporal model to predict hourly energy loads at various lead times, utilizing data from 42 major cities across Alberta. Our method involved selecting input features, including temperature, relative humidity, dew point, air pressure, and wind speed, using four feature selection methods: Mutual Information (MI), Principal Component Analysis (PCA), Lasso regression, and Heatmap correlation. We then benchmarked these methods with a machine learning model that incorporated not only the temporal history of these weather features but also the temporal history of load data, the sine-cosine transformation of the hour of the day, and calendar day (working or non-working).

Through a comprehensive analysis, we identified the most influential weather features for a set of randomly selected cities and determined the most effective feature selection methodology. The use of the optimal feature selection method enabled us to study and visualize the importance of weather features for all cities, resulting in a clear behavioral pattern mapped geographically across Alberta.

An additional insight from this study was that grouping lead times and developing individual machine learning models for each group could potentially provide more accurate and reliable load predictions. This approach is currently under investigation and will be the subject of future publications.

This study demonstrates that the combination of careful feature selection, appropriate temporal transformations, and a deep understanding of the local climate's influence on energy consumption can significantly enhance the accuracy of load prediction models. These insights contribute to the development of more precise and reliable forecasting tools, which are essential for efficient energy management and planning.

To demonstrate the robustness of the model when all five weather features are used to predict the 72-hour lead time load for the city of Calgary, four randomly selected predictions are presented in **Error! Reference source not found.**. These results highlight the model's reliability, which will be further explored in our future publications.

VI. ACKNOWLEDGEMENT

We would like to extend our sincere gratitude to weathersource.com for granting us access to the weather data of various locations across Alberta. This valuable data was instrumental in the development and validation of our load prediction models.

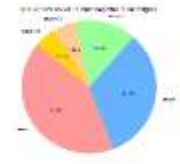
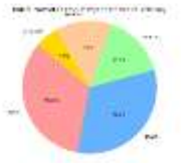
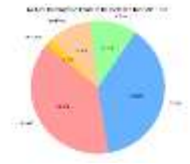
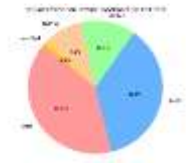
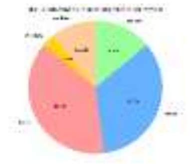
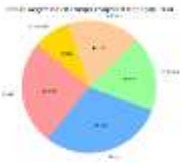
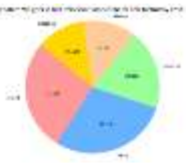
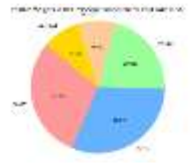
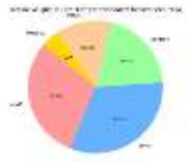
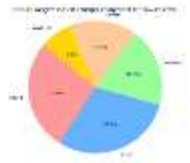
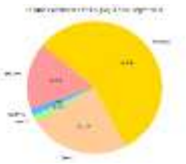
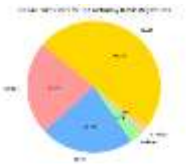
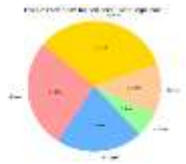
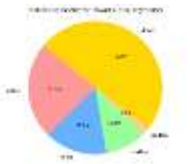
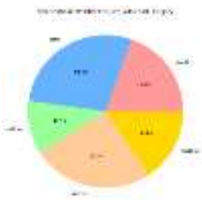
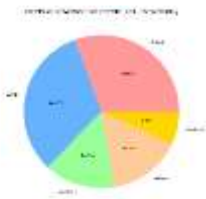
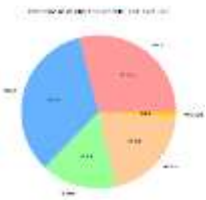
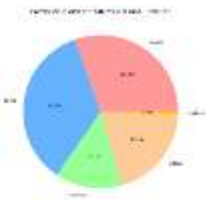
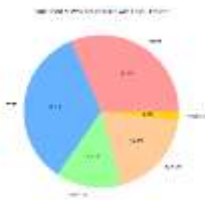
VII. REFERENCES

1. Rastogi, Deeksha; Lehner, Flavio; Kuruganti, Teja; Evans, Katherine J; Kurte, Kuldeep; Sanyal,

Jibonananda “The Role of Humidity in Determining Future Electricity Demand in the Southeastern United States.” Environmental Research Letters, vol. 16, no. 11, 2021, pp. 114017-, https://doi.org/10.1088/1748-9326/ac2fdf.
2. Auffhammer, Maximilian; Baylis, Patrick; Hausman, Catherine H. “Climate Change Is Projected to Have Severe Impacts on the Frequency and Intensity of Peak Electricity Demand across the United States.” Proceedings of the National Academy of Sciences - PNAS, vol. 114, no. 8, 2017, pp. 1886–91, https://doi.org/10.1073/pnas.1613193114.
3. Ralston Fonseca, Francisco; Jaramillo, Paulina; Bergés, Mario; Severnini, Edson. “Seasonal Effects of Climate Change on Intra-Day Electricity Demand Patterns.” Climatic Change, vol. 154, no. 3–4, 2019, pp. 435–51, https://doi.org/10.1007/s10584-019-02413-w.
4. Rastogi, Deeksha; Holladay, James Scott; Evans, Katherine J; Preston, Ben L; Ashfaq, Moetasim. “Shift in Seasonal Climate Patterns Likely to Impact Residential Energy Consumption in the United States.” Environmental Research Letters, vol. 14, no. 7, 2019, pp. 74006-, https://doi.org/10.1088/1748-9326/ab22d2.
5. Psiloglou, B.E.; Giannakopoulos, C.; Majithia, S.; Petrakis, M. “Factors Affecting Electricity Demand in Athens, Greece and London, UK: A Comparative Assessment.” Energy (Oxford), vol. 34, no. 11, 2009, pp. 1855–63, https://doi.org/10.1016/j.energy.2009.07.033.
6. Maia-Silva, D., Kumar, R., Nateghi, R. (2020). “The Critical Role of Humidity in Modeling Summer Electricity Demand across the United States” Nature Communications, vol. 11, no. 1, 2020, pp. 1686–88, https://doi.org/10.1038/s41467-020-15393-8.
7. Beccali, M.; Cellura, M.; Lo Brano, V.; Marvuglia, A. “Short-Term Prediction of Household Electricity Consumption: Assessing Weather Sensitivity in a Mediterranean Area.” Renewable & Sustainable Energy Reviews, vol. 12, no. 8, 2008, pp. 2040–65, https://doi.org/10.1016/j.rser.2007.04.010.
8. Mirasgedis, S.; Sarafidis, Y.; Georgopoulou, E.; Lalas, D.P.; Moschovits, M.; Karagiannis, F.; Papakonstantinou, D. “Models for Mid-Term Electricity Demand Forecasting Incorporating Weather Influences.” Energy (Oxford), vol. 31, no. 2, 2006, pp. 208–27, https://doi.org/10.1016/j.energy.2005.02.016.
9. Friedrich, Luiz; Armstrong, Peter; Afshari, Afshin. “Mid-Term Forecasting of Urban Electricity Load to Isolate Air-Conditioning Impact.” Energy and Buildings, vol. 80, 2014, pp. 72–80, https://doi.org/10.1016/j.enbuild.2014.05.011.
10. Wang, Yaoping, and Jeffrey M. Bielicki. “Acclimation and the Response of Hourly Electricity Loads to Meteorological Variables.” Energy, vol. 142, 2018, pp. 473–85, https://doi.org/10.1016/j.energy.2017.10.037.
11. Ihara, T.; Genchi, Y.; Sato, T.; Yamaguchi, K.; Endo, Y. “City-Block-Scale Sensitivity of Electricity Consumption to Air Temperature and Air Humidity in Business Districts of Tokyo, Japan.” Energy (Oxford), vol. 33, no. 11, 2008, pp. 1634–45, https://doi.org/10.1016/j.energy.2008.06.005.
12. https://www.aeso.ca/market/market-and-system-reporting/data-requests/hourly-load-by-area-and-region/
13. https://weathersource.com/products/onpoint-weather/
14. Abedinia, Oveis; Amjady, Nima; Zareipour, Hamidreza. “A New Feature Selection Technique for Load and Price Forecast of Electrical Power Systems.” IEEE Transactions on Power Systems, vol. 32, no. 1, 2017, pp. 62–74, https://doi.org/10.1109/TPWRS.2016.2556620.
15. Cover, T.M., and Thomas, J.A. (2006). Elements of Information Theory (2nd ed.). Wiley.
16. Witten, Ian H., and Eibe Frank. Data Mining: Practical Machine Learning Tools and Techniques, Second Edition. 2nd ed., Elsevier Science, 2005.
17. Introduction to Statistical and Machine Learning Methods for Data Science. SAS Institute, 2021.
18. Bishop, Christopher M. Pattern Recognition and Machine Learning. Springer, 2006.
19. Shannon, Claude Elwood, and Warren Weaver. The Mathematical Theory of Communication. University of Illinois Press, 1963.
20. Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. Philosophical Magazine Series 6, 2(11), 559–572.
21. Hotelling, H. (1933). "Analysis of a complex of statistical variables into principal components" Journal of Educational Psychology, 24(6), 417-441, https://doi.org/10.1037/h0071325.
22. Tibshirani, R. (1996). "Regression Shrinkage and Selection via the Lasso" Journal of the Royal Statistical Society: Series B (Statistical Methodology), 58(1), 267–288, https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.
23. Pearson, Karl. “Note on Regression and Inheritance in the Case of Two Parents.” Proceedings of the Royal Society of London, vol. 58, 1895, pp. 240–42.
24. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
25. Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. International Conference on Learning Representations (ICLR).

JEATSA JOURNAL
JOURNAL OF ENGINEERING ADVANCES AND TECHNOLOGIES FOR SUSTAINABLE APPLICATIONS
Volume 1, Issue 1 (January, 2025)
Print ISSN: 3062-5629 Online ISSN: 3062-5637
Open Access
https://jeats.journals.ekb.eg/

Table 1: Features prioritization by selected feature selection methods

Method/City	Calgary	Fort McMurray	Cold Lake	Red Deer	Provost
MI					
PCA					
Lasso					
Heatmap					
Logo (color codes)					

JEATSA JOURNAL
JOURNAL OF ENGINEERING ADVANCES AND TECHNOLOGIES FOR SUSTAINABLE APPLICATIONS
Volume 1, Issue 1 (January, 2025)
Print ISSN: 3062-5629 Online ISSN: 3062-5637
Open Access
https://jeats.journals.ekb.eg/

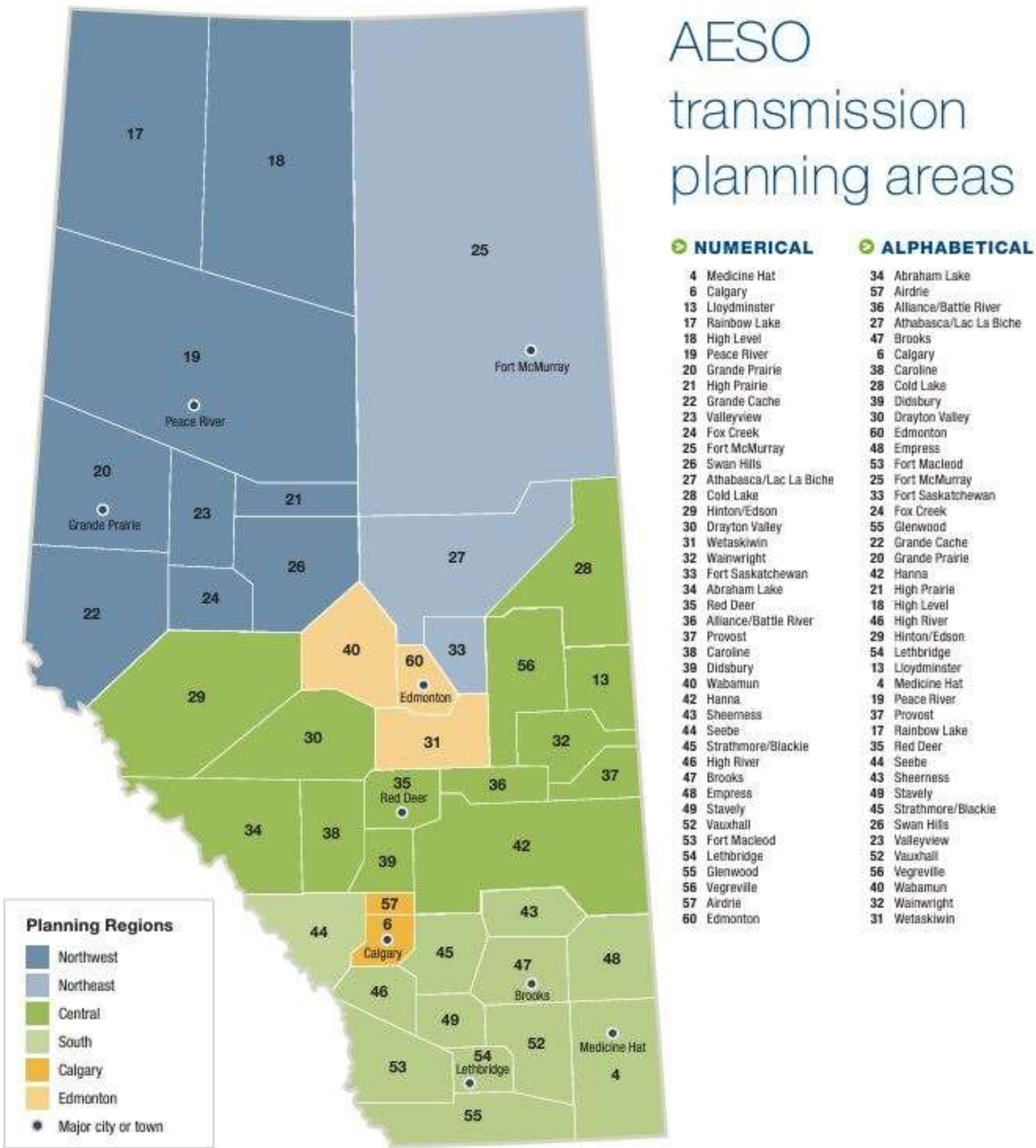


Figure 1 – Alberta transmission area based on AESO's plan [12]

JEATSA JOURNAL
JOURNAL OF ENGINEERING ADVANCES AND TECHNOLOGIES FOR SUSTAINABLE APPLICATIONS
Volume 1, Issue 1 (January, 2025)
Print ISSN: 3062-5629 Online ISSN: 3062-5637
Open Access
https://jeats.journals.ekb.eg/

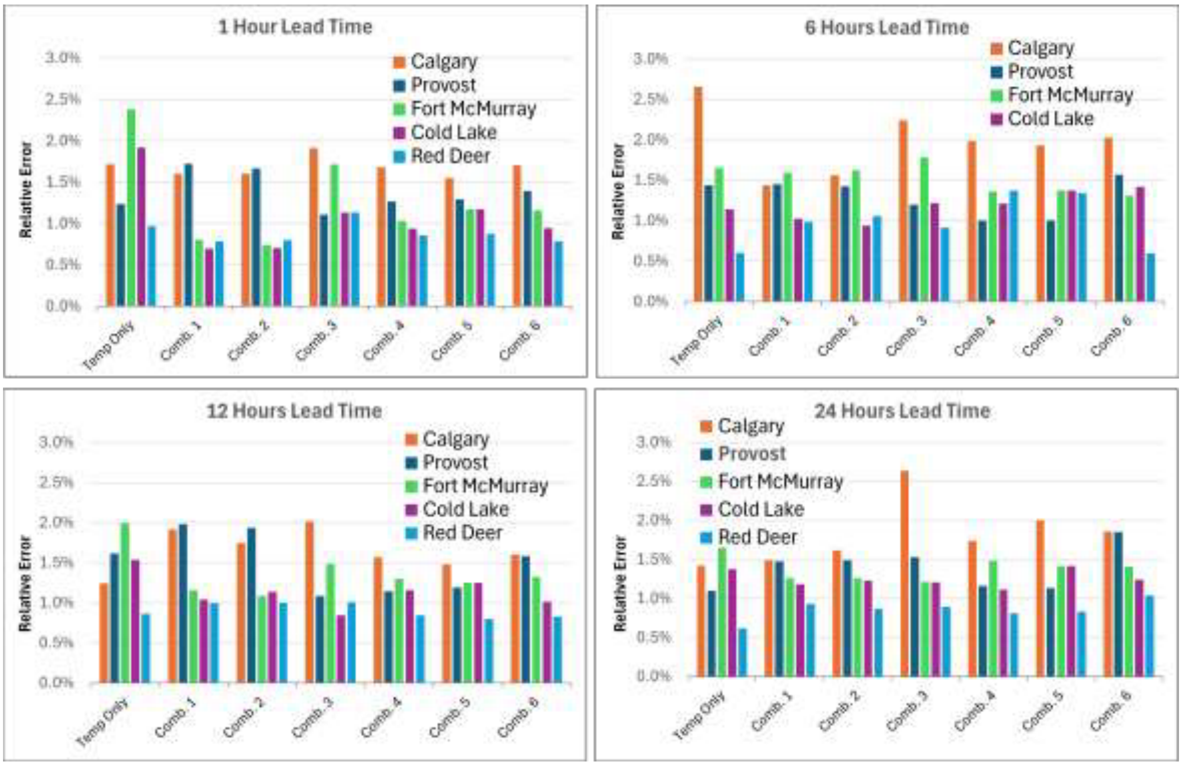


Figure 2 - Precision of different lead times by different combinations of weather features

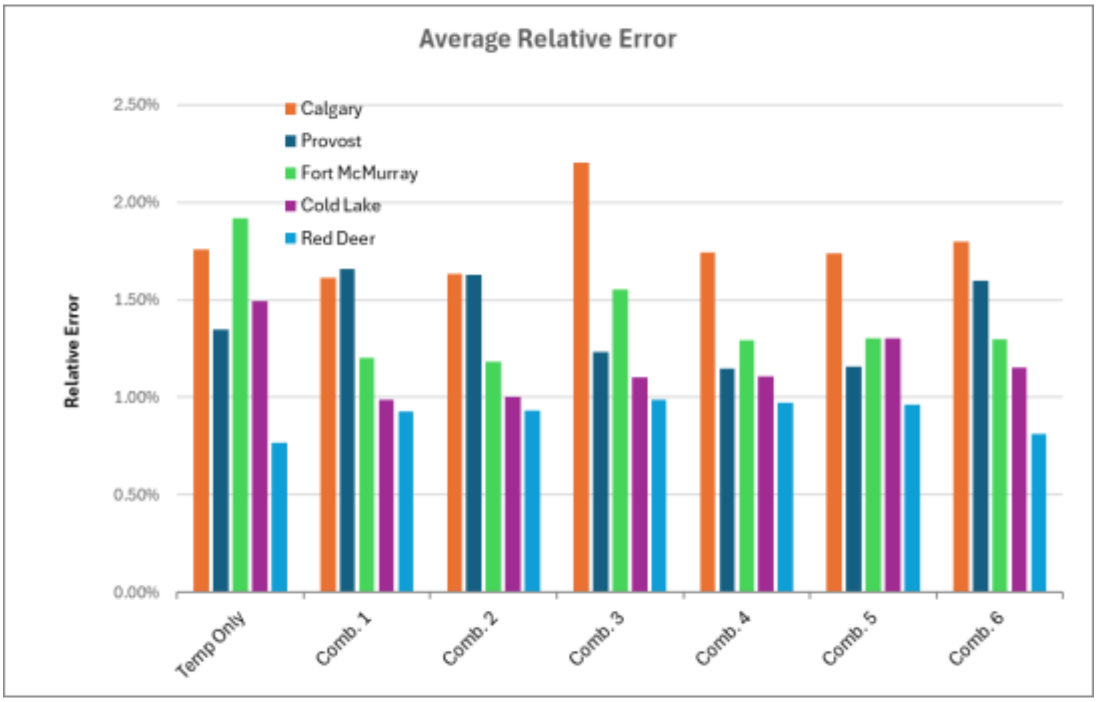


Figure 3 – Averaged precision of different feature selection for the selected cities

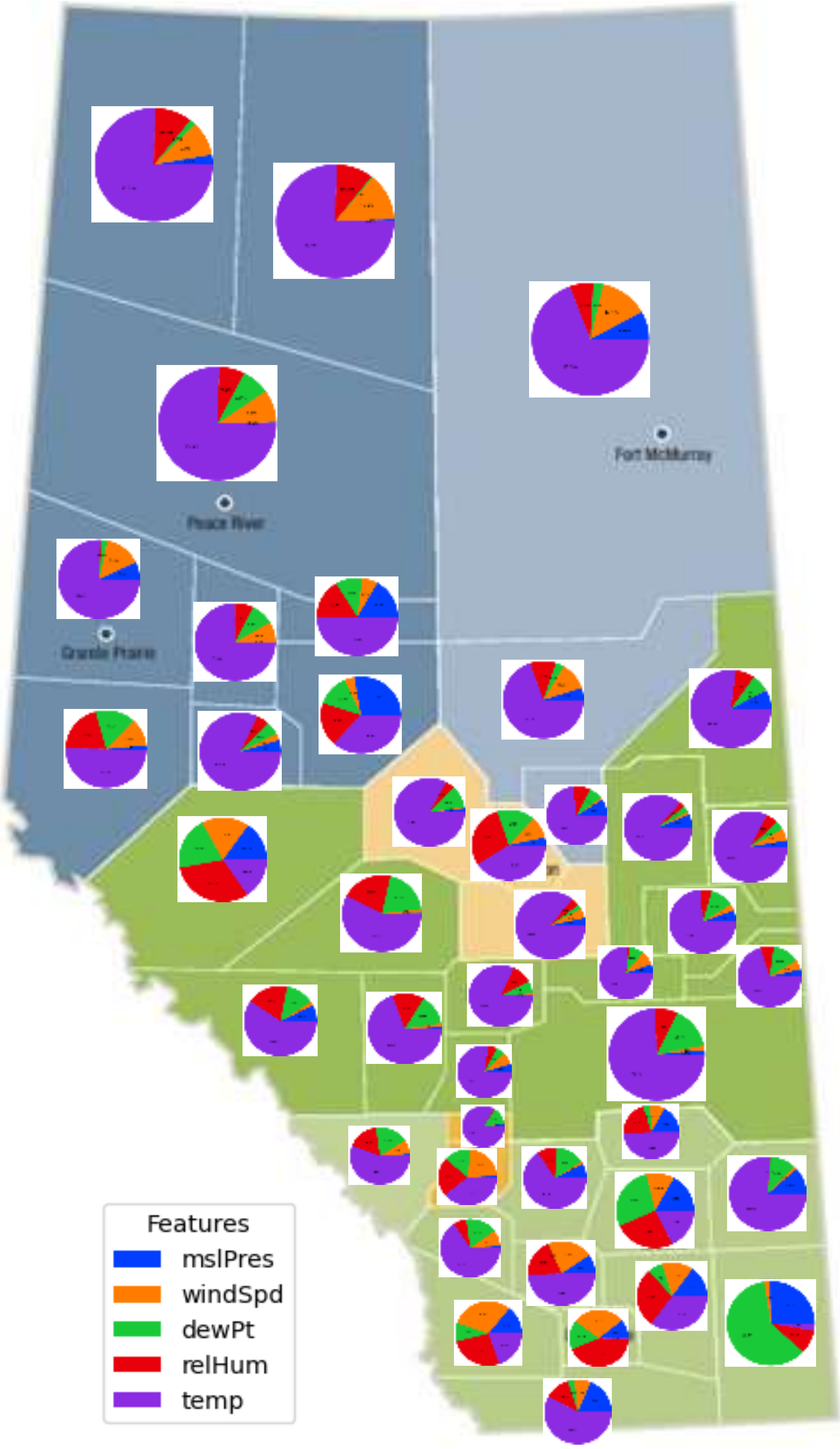


Figure 4 – Areal analysis of weather features importance (after detrending)

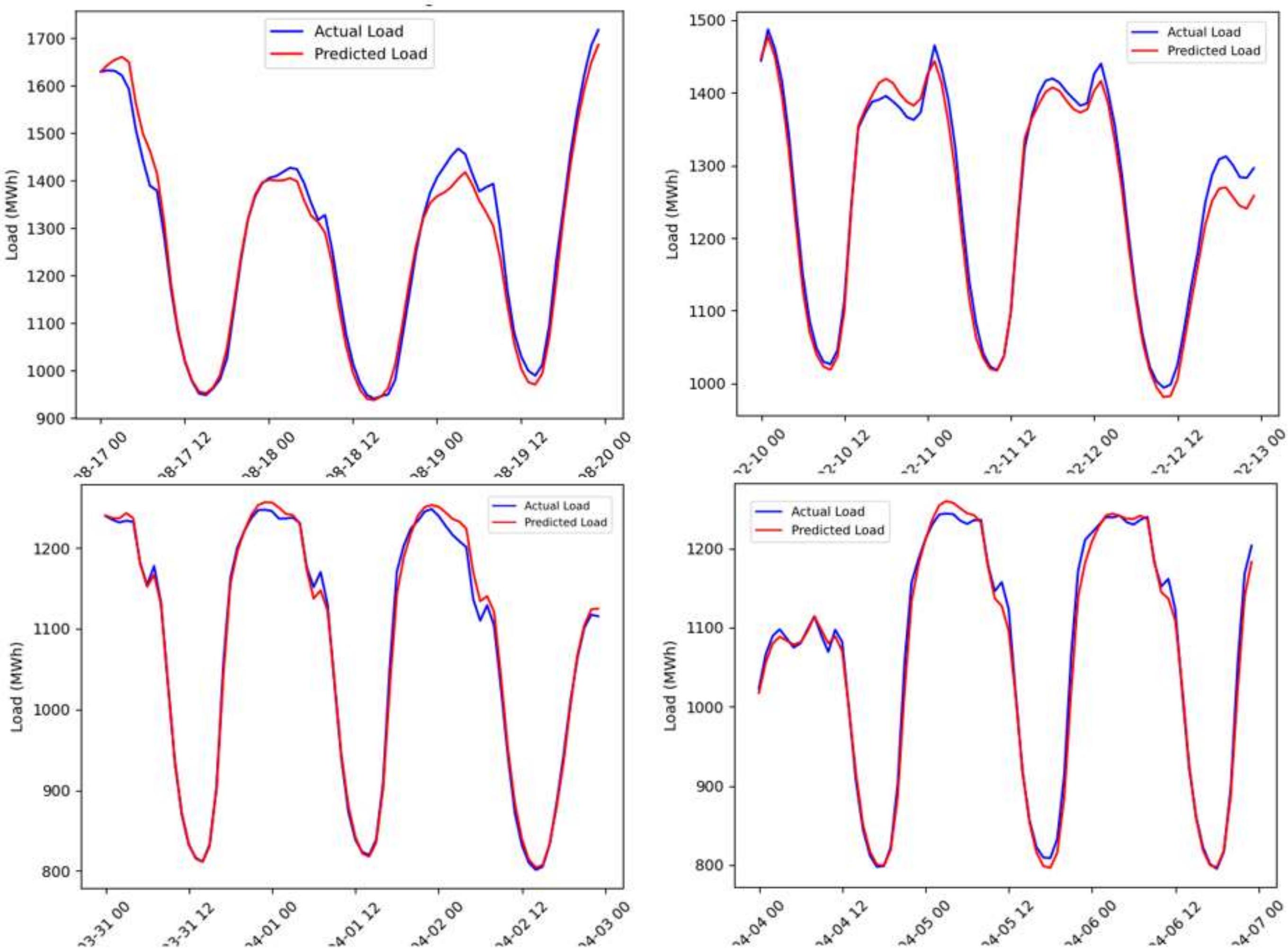


Figure 5 – Some prediction (red) vs. actual (blue) loads for the City of Calgary