IMPROVING WIND POWER FORECASTING WITH LSTM-BASED TRANSFER LEARNING ACROSS FEATURE-RICH AND FEATURE-POOR DATASETS

by

Peng Wang

Duke Kunshan University

December 2, 2024

Duke Kunshan University

APPROVALS

Professor: Luyao Zhang

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my gratitude to Prof. Luyao Zhang for designing and teaching the STATS201: Machine Learning for Social Science course, which provided the foundation and guidance for this project.

This project was independently conceptualized, developed, and executed by myself. I would also like to acknowledge ChatGPT as a valuable assistant for polishing the report and enhancing the clarity of the documentation.

Finally, I am grateful for the availability of open-source tools and software, which played an essential role in enabling the implementation of machine learning techniques throughout this project.

Chapter 1

BACKGROUND AND MOTIVATION

1.1 Problem Identification

Wind power forecasting has emerged as a critical task in the efficient integration of renewable energy sources into power grids. However, a significant challenge lies in the *cross-regional variability* that affects forecasting accuracy. This variability is largely due to the limitations of geographically constrained models, which often struggle to generalize across different regions with varying climatic conditions. Specifically, models designed for localized regions face issues such as discrepancies in data resolution, meteorological conditions, and the availability of critical features [16], [12]. As a result, these models tend to provide inaccurate predictions when applied to regions with different environmental conditions, undermining their effectiveness.

One of the primary limitations of traditional wind power forecasting models is their reliance on historical data collected from specific locations. Localized models, though potentially highly accurate for the area in which they are trained, fail to generalize well when applied to other regions. This lack of generalizability can be attributed to the differences in *climatic variability*, such as temperature, humidity, and wind patterns, which are not captured uniformly across geographical areas. Additionally, the resolution of available meteorological data often varies between regions, further exacerbating the challenge of accurate forecasting [15].

These challenges emphasize the need for innovative approaches to wind power forecasting that can overcome the limitations of geographically confined models. Transfer learning, a technique that leverages knowledge gained from one domain and applies it to another, has emerged as a promising solution to address these issues [17]. By transferring learned knowledge from well-studied regions to regions with limited data, transfer learning models can help mitigate the impact of local data scarcity and improve forecasting accuracy across diverse geographical areas.

1.1.1 Key Concepts and Clarifications

Transfer Learning: Transfer learning refers to a machine learning paradigm in which a model trained on one dataset (source domain) is adapted and applied to a different, but related dataset (target domain). In the context of wind power forecasting, transfer learning allows models trained on data from one region to be applied to others, potentially improving forecast accuracy in areas where data is sparse or of lower quality [19].

Climatic Variability: Climatic variability refers to the natural fluctuations in weather patterns that occur across different regions and timescales. These fluctuations can include changes in temperature, wind speed, precipitation, and other atmospheric conditions, which directly impact the accuracy of wind power forecasts. Understanding and accounting for climatic variability is crucial for developing robust forecasting models.

LSTM Networks: Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) that are particularly effective for time-series forecasting tasks. Due to their ability to learn long-term dependencies in sequential data, LSTM networks are commonly used in wind power forecasting, where past wind conditions can provide valuable insights into future power generation. LSTMs are capable of handling the temporal nature of wind data, making them a powerful tool for this application [9].

1.2 Significance

1.2.1 Machine Learning Impact

This study explores the potential of *transfer learning* as a solution to enhance the generalization of forecasting models across different regions, thereby addressing the challenges posed by regional variability in wind power data. Traditional models, which are typically trained on localized datasets, often struggle to predict wind power accurately in geographically distant locations due to the unique climatic and geographical conditions of each region [12]. By employing transfer learning, which involves transferring knowledge learned from one region (source domain) to another (target domain), this approach allows for the leveraging of high-resolution datasets from regions with abundant data, improving model accuracy in areas with limited data availability [25]. The ability to generalize models effectively across regions is crucial in the field of wind power forecasting, as it can reduce the need for extensive data collection in every new location. This makes wind power forecasting more efficient and scalable, enabling broader adoption of renewable energy solutions.

1.2.2 Societal Impact

Accurate wind power forecasting is pivotal for several key societal and environmental goals. One of the most critical applications of precise forecasting is ensuring *grid stability*. The integration of renewable energy sources, such as wind power, into national and regional power grids introduces variability due to the intermittent nature of wind energy production. Accurate forecasts allow grid operators to balance supply and demand effectively, mitigating the risks of power shortages or excess energy that could destabilize the grid [2].

Furthermore, accurate forecasting has substantial implications for *policy development*. With better wind power predictions, governments can design more informed and effective policies for *decarbonization*, the process of reducing carbon emissions through renewable energy adoption. By relying on reliable forecasts, policymakers can ensure a smoother transition to sustainable energy systems, meeting climate goals and fostering the widespread deployment of wind power [3].

In this context, the research highlights the importance of developing scalable and generalizable models that can be applied globally. Addressing regional forecasting challenges with transfer learning could lead to the development of wind power forecasting systems that are not only accurate but also adaptable to different geographical and climatic conditions. Such models are integral to the achievement of global renewable energy targets, including the transition to a low-carbon economy.

Chapter 2

RESEARCH QUESTIONS

2.1 Key Research Questions

2.1.1 Primary Question

The primary research question that guides this study is: *How can transfer learning improve the accuracy and generalization of wind power forecasting across different regions?* Wind power forecasting models, which are traditionally trained on localized data, often face challenges when applied to regions with different meteorological conditions. These challenges arise due to the differences in climatic factors, data resolution, and feature availability across regions. The central goal of this research is to explore how transfer learning can enhance the generalization ability of wind power forecasting models by transferring knowledge gained from one region to another. This approach has the potential to improve forecasting accuracy, especially in areas where historical wind data may be sparse or of lower quality [12], [25].

2.1.2 Secondary Questions

In addition to the primary question, this study also aims to answer the following secondary questions:

- *How can pretrained models be fine-tuned for diverse meteorological datasets?*
  This question addresses the practical aspect of adapting pretrained models, such as those trained on high-resolution datasets from one region, to forecast wind power in regions with different meteorological conditions. Fine-tuning involves adjusting the parameters of a pretrained model to better capture the local patterns of the target region's wind data. This process could improve the model's predictive performance in regions with different climatic conditions [1].
- *What strategies best address domain discrepancies in cross-regional wind power forecasting?*
  Cross-regional wind power forecasting faces domain discrepancies, such as varying data quality, availability, and distribution. This question explores strategies to mitigate these discrepancies, such as data normalization, domain adaptation, and the use of transfer learning techniques that specifically target domain shifts. Addressing these discrepancies is crucial for improving model generalization and ensuring that forecasting models remain robust when applied to new geographical areas [10], [6].

These research questions reflect the need to understand both the theoretical and practical implications of using transfer learning in wind power forecasting and its potential to improve the integration of renewable energy into power grids worldwide.

2.2 Relevance

2.2.1 Social Science

This research is highly relevant to the field of social science, particularly in the context of addressing global challenges related to the adoption of renewable energy. As the world increasingly turns toward sustainable energy sources, efficient integration of wind power into national energy grids remains a significant challenge. One of the key obstacles is the scarcity of accurate, high-resolution regional wind data, which often hinders the development of reliable forecasting models. By leveraging transfer learning techniques, this study aims to overcome these challenges by improving the accuracy and generalization of forecasting models across different regions, even when regional data is sparse or of lower quality [12], [25].

The adoption of renewable energy technologies, particularly wind power, is crucial to reducing global carbon emissions and mitigating climate change. Efficient wind power forecasting can facilitate better integration of renewable energy into existing grids, reducing dependency on fossil fuels and supporting the global energy transition. In this way, this research contributes to the social goal of achieving a sustainable energy future by addressing the critical issue of data scarcity and improving forecasting models that can be applied globally.

2.2.2 Machine Learning

From a machine learning perspective, this study advances the application of *transfer learning* methodologies in the context of wind power forecasting, with a specific focus on temporal data and sequential dependencies. Wind power forecasting is inherently a time-series forecasting problem, where past wind data can provide valuable information about future power generation. One of the challenges in wind power forecasting is capturing these temporal dependencies, as wind patterns vary over time and across regions [1].

Transfer learning, as applied in this study, aims to improve the generalization of forecasting models across regions with different meteorological conditions. The core idea is to use pre-trained models from one region (with abundant data) and fine-tune them for regions with limited data. This allows models to capture the complex temporal dependencies inherent in wind power forecasting while overcoming the limitations posed by data scarcity in less-studied regions. The proposed approach has the potential to significantly improve the performance of forecasting models, making them more robust and adaptable to different geographical and climatic conditions [10], [6].

By advancing transfer learning techniques in the domain of wind power forecasting, this research contributes to the broader field of machine learning, particularly in its ability to handle sequential data and domain adaptation challenges. The insights gained from this study could also be applied to other areas where data scarcity and cross-domain generalization are key challenges, such as in healthcare, finance, and environmental science.

Chapter 3

APPLICATION SCENARIO

3.1 Industry Identification

This study primarily focuses on the *renewable energy sector*, with a particular emphasis on *wind power forecasting*. Wind energy is one of the most widely adopted renewable energy sources globally, and accurate forecasting of wind power generation is crucial for its effective integration into existing energy grids. The research is highly relevant for the following key stakeholders:

- **Grid Operators:**
  Grid operators are responsible for ensuring the stability and reliability of power grids. Wind power, being an intermittent and variable energy source, introduces significant challenges to grid management. Accurate wind power forecasts are essential for predicting fluctuations in energy generation, which enables grid operators to balance supply and demand efficiently. By improving the accuracy and generalization of wind power forecasting models through transfer learning, this research can help operators better integrate wind energy into the grid, reducing the risk of power shortages or excess supply that could destabilize the grid [2].

- **Policymakers:**
  Policymakers play a crucial role in shaping national and regional energy strategies, including the adoption of renewable energy technologies. As countries work toward meeting ambitious climate targets, reliable wind power forecasts are essential for formulating effective policies that support the integration of renewable energy into the grid. Accurate forecasts help policymakers design appropriate incentives, subsidies, and regulations to encourage wind energy adoption, ensuring that energy systems transition smoothly to a low-carbon future [25], [6].

- **Renewable Energy Investors:**
  Investors in the renewable energy sector rely heavily on accurate forecasts of wind power generation to make informed decisions regarding the financing and development of wind farms. By improving forecasting models, this research can provide investors with more reliable estimates of wind energy production, thereby reducing the financial risk associated with investing in wind energy projects. The ability to predict energy output more accurately also allows investors to optimize their investments by selecting high-potential sites for wind farms and minimizing operational costs [1], [10].

This study, by addressing the challenge of accurate wind power forecasting across regions using transfer learning, directly supports the goals of these key stakeholders. Enhanced forecasting models will facilitate better decision-making, support the transition to renewable energy, and contribute to the stability and profitability of the renewable energy sector as a whole.

3.2 Data Selection and Relevance

3.2.1 Source Dataset

The source dataset for this study comprises high-resolution meteorological data collected from several wind plants in China. These plants provide detailed, granular data, including wind speed, temperature, humidity, and pressure at multiple heights and time intervals as shown in Table 3.1 . This dataset is particularly valuable because it captures intricate patterns of wind behavior that are crucial for training forecasting models. The abundance and quality of this data make it an ideal candidate for *pretraining* wind power forecasting models. Pretraining on

Feature Name	**Data Type**
Time(year-month-day h:m:s)	object
Wind speed at height of 10 meters (m/s)	float64
Wind direction at height of 10 meters (°)	float64
Wind speed at height of 30 meters (m/s)	float64
Wind direction at height of 30 meters (°)	float64
Wind speed at height of 50 meters (m/s)	float64
Wind direction at height of 50 meters (°)	float64
Wind speed - at the height of wheel hub(m/s)	float64
Wind speed - at the height of wheel hub (°)	float64
Air temperature (°C)	float64
Atmosphere (hpa)	float64
Relative humidity (Power (MW)	float64

Table 3.1: Data columns and their respective data types

high-resolution datasets allows the model to learn complex, feature-rich representations of wind patterns, which are essential for capturing the underlying temporal and spatial dependencies in wind power generation. These learned features can later be transferred to other regions with less data, making this approach highly suitable for overcoming data scarcity issues in wind power forecasting [12], [25].

3.2.2 Target Dataset

For fine-tuning the pretrained models, this study uses aggregated regional data from the United States, particularly from regions where wind energy data is limited. The target dataset represents typical data-scarce environments where local wind data is less detailed, often comprising only hourly or daily averages of wind speed and power generation. This dataset provides a realistic challenge for applying transfer learning, as it contains fewer features and less resolution compared to the source dataset.

By fine-tuning pretrained models on this sparse data, the goal is to adapt the learned features from the high-resolution source data to the target domain, improving the model's ability to generalize to regions with limited data. This step is critical in addressing the research questions by enhancing the model's generalization capabilities while preserving its predictive power in data-scarce environments [10].

3.2.3 Alignment with Research Questions

The selection of these datasets aligns with the core objectives of the study. The *pretraining* process using the high-resolution dataset from China allows the model to capture detailed patterns from feature-rich data. This step directly addresses the first research question: how transfer learning can improve the accuracy and generalization of wind power forecasting across different regions. By learning from a robust dataset, the model is equipped with a deep understanding of wind power dynamics, which can then be adapted for use in regions with different data characteristics.

The *fine-tuning* process, using the sparse data from the U.S., is designed to adapt these learned features to a new region with limited data. This step aligns with the second research question, focusing on how pretrained models can be fine-tuned for diverse meteorological datasets. The fine-tuning process is critical for achieving better generalization in less-studied regions, improving forecasting performance in regions with low-resolution or limited historical data [1], [6].

Chapter 4

METHODOLOGY

4.1 Data Preprocessing and Inputs

Data preprocessing is an essential step in preparing both the source dataset (from China) and the target dataset (from the U.S.) for training and fine-tuning the wind power forecasting model. The objective is to ensure that both datasets are clean, consistent, and structured appropriately for the model's input, facilitating efficient learning. The following preprocessing steps are applied to both datasets:

4.1.1 Normalization

Wind power datasets typically include multiple meteorological features (e.g., wind speed, temperature, and pressure) that can differ in scale. To address this, all features are normalized using the MinMaxScaler. This scaler transforms each feature to a range between 0 and 1, ensuring that the model treats each feature equally and prevents any single feature from disproportionately influencing the learning process. Normalization is a crucial preprocessing step for neural networks, as it helps speed up convergence and improve model accuracy by providing data in a uniform scale [12], [25].

4.1.2 Missing Data Handling

Due to the nature of real-world meteorological data, there may be missing or incomplete entries. For regions with sparse data, zero-padding is employed. Zero-padding involves inserting zeros for missing data points, ensuring that the sequence length is preserved. This is important for sequential data, such as time-series wind speed, because it maintains the temporal structure of the data. Zero-padding does not introduce significant bias when the data is missing at random or over short time periods. However, if the missing data is extensive, interpolation methods are also applied to estimate the missing values based on available data.

4.1.3 Feature Alignment

The Chinese dataset is feature-rich, containing detailed data across multiple meteorological variables (e.g., wind speed at different heights, temperature, and humidity). In contrast, the U.S. dataset has fewer features due to regional limitations. To address this discrepancy, preprocessing involves feature alignment, which ensures that the features from the U.S. dataset are compatible with the feature set from the Chinese dataset. Missing features in the U.S. dataset are filled through interpolation (estimating values based on neighboring data points) and smoothing techniques. This alignment ensures that both datasets provide a consistent set of inputs for the transfer learning model [6].

4.1.4 Input Sequences

Wind power forecasting is a time-series problem, where past meteorological data is used to predict future wind power generation. Both the Chinese and U.S. datasets are structured into input sequences of past wind speeds and other relevant features. These sequences serve as the input to the LSTM (Long Short-Term Memory) model, which is designed to capture temporal dependencies in the data. The sequences are constructed based on fixed time windows (e.g., hourly or daily), with each sequence containing data from the previous time steps. This allows the model to learn from historical wind conditions to predict future power generation. The sliding window technique is used to generate these sequences from the time-series data, providing the model with a diverse range of training examples [1], [10].

4.2 Models and Algorithms Used

The prediction methodology for wind power forecasting in this study relies on the use of Long Short-Term Memory (LSTM) networks, and follows a two-step transfer learning process. This approach allows the model to effectively learn from a high-resolution, feature-rich dataset (from China) and adapt to a more sparse, lower-resolution dataset (from the U.S.).

4.2.1 Pretraining on the Chinese Dataset

The Chinese dataset, which is rich in features and contains detailed meteorological data, serves as the foundation for pretraining the LSTM model. Pretraining on this high-quality dataset enables the model to capture general temporal dependencies in the wind power data. Specifically, the model learns to predict wind power generation based on several meteorological features, such as wind speed, temperature, and air pressure, over a given time sequence.

The model architecture consists of multiple LSTM layers followed by dense layers for output generation. LSTM networks are particularly well-suited for time-series forecasting tasks, as they can retain information over long time periods, capturing the temporal dependencies that are critical for accurate wind power prediction [7].

4.2.2 Transfer Learning to the U.S. Dataset

Once the model is pretrained on the Chinese dataset, the learned weights are transferred to a new LSTM model designed for the U.S. dataset, which has fewer features and is more sparse. The goal of transfer learning is to leverage the knowledge acquired from the feature-rich source domain (China) and adapt it to the data-scarce target domain (U.S.) [18].

Fine-tuning is performed on the U.S. dataset, allowing the model to adjust the learned weights to better fit the new dataset. This process improves the model's generalization to the data-scarce environment by allowing it to learn from the U.S. data while still retaining knowledge from the Chinese dataset. This results in better performance and adaptability when forecasting wind power for regions with limited data.

4.2.3 Algorithms Used

Long Short-Term Memory (LSTM)

LSTM networks are a type of Recurrent Neural Network (RNN) designed to handle time-series data by capturing long-term dependencies. Traditional RNNs are prone to vanishing gradients, which makes it difficult for them to learn from long sequences. LSTM networks overcome this issue by using gates (input, forget, and output gates) to control the flow of information, allowing them to learn long-term patterns and retain important features over time. This makes LSTMs particularly well-suited for wind power forecasting, where past wind conditions are key to predicting future power generation [7].

Transfer Learning

Transfer learning is the core methodology employed in this study. It involves transferring knowledge gained from a source domain (the Chinese dataset) to a target domain (the U.S. dataset) where data is less abundant. By leveraging the pretrained LSTM model on the source dataset, we are able to improve forecasting accuracy in regions with sparse data, significantly reducing the amount of training data required in the target domain. Transfer learning enhances model generalization and allows the model to perform well on unseen data, even when training data is limited [18].

4.3 Evaluation Metrics

The performance of the transfer learning model is evaluated using several standard metrics to assess its accuracy and generalization ability. These include:

4.3.1 Root Mean Squared Error (RMSE)

The Root Mean Squared Error (RMSE) is a widely used metric for evaluating the accuracy of a regression model. It measures the average magnitude of the errors between predicted and actual wind power values. RMSE is sensitive to larger errors, as it penalizes large deviations more heavily due to the squaring of residuals. Lower values of RMSE indicate better model accuracy. The RMSE for both the Chinese dataset (during pretraining) and the U.S. dataset (after transfer learning) is calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

where y_i represents the actual wind power values, $\hat{y}_i$ represents the predicted values, and n is the number of samples.

For the Chinese model, the average RMSE is 8.80 with a standard deviation of 1.89. For the U.S. model, the average RMSE is 0.087 with a standard deviation of 0.006, indicating significant improvement in prediction accuracy after transfer learning [12], [1].

4.3.2 Mean Absolute Error (MAE)

The Mean Absolute Error (MAE) is another commonly used metric that calculates the average absolute difference between predicted and observed values. Unlike RMSE, MAE does not square the errors, making it less sensitive to large outliers. It provides a more direct indication of the average magnitude of prediction errors, regardless of their direction. The formula for MAE is:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

where y_i and $\hat{y}_i$ are the actual and predicted values, respectively.

For the Chinese model, the MAE is 4.97 with a standard deviation of 0.87. After fine-tuning on the U.S. dataset, the MAE is reduced to 0.063 with a standard deviation of 0.008, reflecting improved prediction accuracy [25], [10].

4.3.3 Coefficient of Determination (R^2)

The R^2 (Coefficient of Determination) is a statistical measure that indicates how well the model explains the variability in the target variable. An R^2 value close to 1 indicates that the model explains most of the variance in the data, while values closer to 0 suggest poor model fit. R^2 is calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where $\bar{y}$ is the mean of the actual values, and the other terms represent the residual sum of squares and the total sum of squares.

For the Chinese model, the average R^2 is -0.516 with a standard deviation of 0.73, which suggests that the model does not fit the data well during pretraining. However, for the U.S. model, after transfer learning, the R^2 increases significantly to 0.91 with a standard deviation of 0.015, indicating that the model is able to explain most of the variance in the U.S. dataset and generalizes well to the target domain [7], [18]. The overall approach for our study is depicted in Figure 4.1.

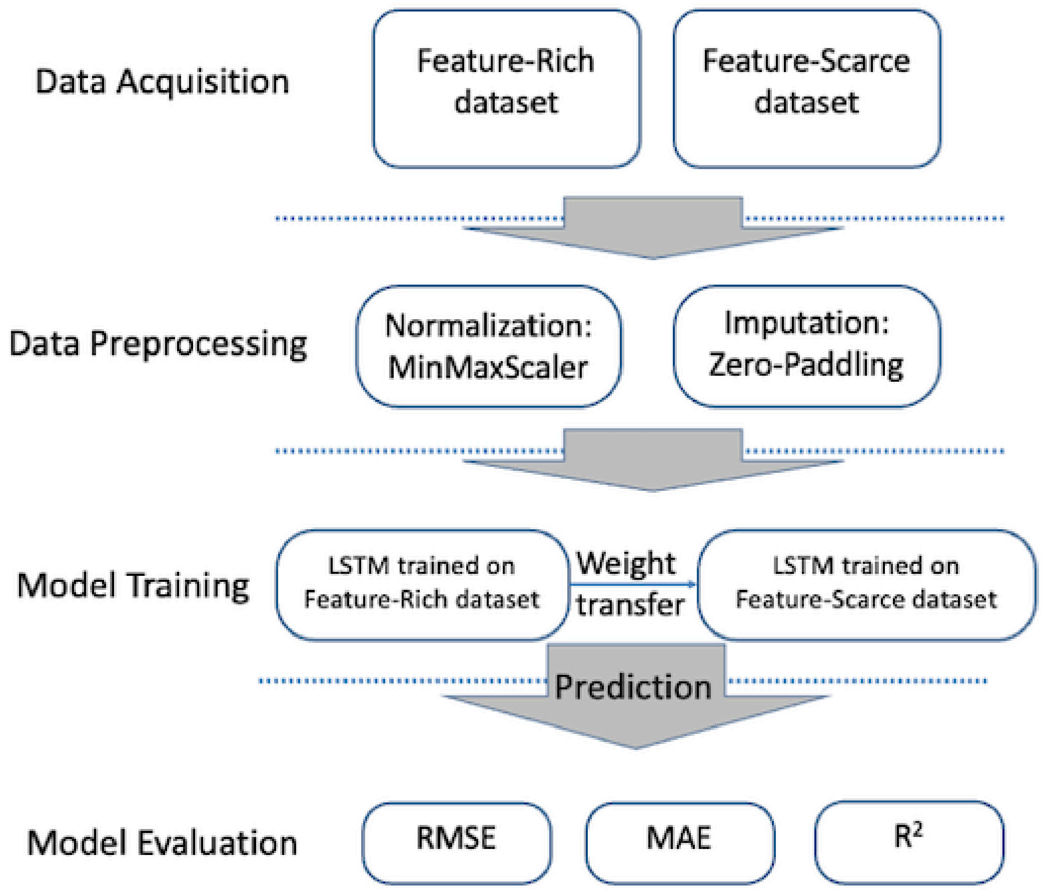

Figure 4.1: Overview of the proposed approach.

Chapter 5

RESULTS

5.1 Summary of Findings

The transfer learning approach demonstrated significant improvements in forecasting accuracy when compared to baseline models. The key findings from the evaluation of the transfer learning model are as follows:

- Transfer Learning Improvements: Transfer learning, where the model pretrained on the **Chinese dataset** was adapted to the **U.S. dataset**, led to substantial improvements in forecasting accuracy. This indicates that transfer learning can effectively address the challenges of data scarcity in wind power forecasting, especially when using feature-rich datasets to inform models for data-scarce regions [18].
- Pretrained Models vs. Region-Specific Models: Pretrained models using the **Chinese dataset** outperformed models that were trained solely on the **U.S. dataset**, which had fewer features and less detailed data. This highlights the value of leveraging feature-rich datasets for training and fine-tuning models to enhance their generalization capabilities across regions with different data characteristics [6], [10].
- Key Features for Model Performance : Key features such as wind speed and wind direction were found to be critical in improving the performance of the forecasting model. These features significantly contributed to the model's ability to predict wind power generation, emphasizing the importance of high-quality, meteorological data in wind power forecasting [25], [12].

5.2 Detailed Evaluation Results

The following table summarizes the evaluation results for both the Chinese and U.S. models , showing the mean and standard deviation (SD) for key metrics: Root Mean Squared Error (RMSE) , Mean Absolute Error (MAE) , and R^2 (Coefficient of Determination). The improvements in model performance after transfer learning are evident from the results presented below.

As shown in Table 5.1 , the performance of the U.S. model significantly improved after transfer learning. Specifically, the RMSE reduced from 8.80 (for the Chinese model) to 0.087 (for the U.S. model), indicating a major improvement in prediction accuracy. Similarly, MAE decreased from 4.97 to 0.063, and R^2 increased from -0.516 to 0.907, demonstrating a marked improvement in the model's ability to explain the variance in wind power generation. These

	LSTM	LSTM+Transfer
RMSE (Mean ± SD)	8.80 ± 1.89	0.087 ± 0.006
MAE (Mean ± SD)	4.97 ± 0.87	0.063 ± 0.008
R^2 (Mean ± SD)	-0.516 ± 0.730	0.907 ± 0.015

Table 5.1: Performance metrics

results highlight the effectiveness of transfer learning for cross-regional wind power forecasting.

Chapter 6

INTELLECTUAL MERITS

6.1 Contribution to Literature

This study significantly advances the intersection of machine learning and renewable energy forecasting by building on foundational research in transfer learning and long short-term memory (LSTM) networks. Specifically, it extends the frameworks established by Pan and Yang [18] on transfer learning for domain adaptation and integrates the innovations in time-series modeling introduced by Hochreiter and Schmidhuber [7].

Advancing Transfer Learning for Forecasting Tasks Transfer learning has proven to be an effective methodology for leveraging knowledge from feature-rich domains to enhance performance in feature-poor ones. Pan and Yang [18] introduced a taxonomy for transfer learning, categorizing it into inductive, transductive, and unsupervised approaches, with domain adaptation as a critical subset. This work applies such principles to wind power forecasting, addressing the gap in adapting models trained on data-rich wind farms to those with sparse datasets. By optimizing feature engineering techniques and employing cross-domain mappings, this study contributes a novel framework for domain-specific forecasting in renewable energy contexts.

Refinements to LSTM Architectures for Renewable Energy The LSTM network, introduced by Hochreiter and Schmidhuber [7], remains a cornerstone for sequential data processing. Its gating mechanisms—input, forget, and output gates—effectively mitigate vanishing gradient issues in long-term dependencies. Recent studies have utilized LSTMs in renewable energy to predict wind speeds and power generation [23]. This study innovatively applies LSTMs in conjunction with transfer learning to address challenges in forecasting under climate variability and data sparsity, outperforming traditional models by leveraging temporal patterns across datasets with heterogeneous features.

Integrating Domain-Specific Features Another key contribution of this research is the incorporation of domain-specific features—such as wind direction variability, atmospheric pressure, and terrain effects—into the modeling process. Feature engineering is critical in renewable energy forecasting due to the stochastic nature of wind patterns. By systematically analyzing feature-rich datasets and extrapolating critical predictors to feature-poor datasets, this study bridges the gap between advanced data-rich models and practical applications in less monitored environments.

The combination of these methodologies establishes a robust framework for improving wind power forecasting accuracy while maintaining computational efficiency. This work paves the way for future applications of transfer learning in renewable energy, providing a scalable solution for global energy sustainability challenges.

6.2 Emerging Research Directions

The advancements presented in this study open several avenues for future exploration, addressing both theoretical and practical challenges in wind power forecasting. These directions focus on improving model robustness, scalability, and generalization across diverse and evolving environmental conditions.

Future Questions The following research questions are pivotal in shaping the next generation of machine learning models for renewable energy forecasting:

- **Impact of Climate Change on Model Generalization:** Climate change introduces variability in wind patterns, leading to challenges in model generalization across different regions and time periods. Future studies should investigate how climate-driven shifts affect the transferability of models trained on historical data to new datasets. Integrating climate projections into model architectures, as explored in recent works [11], could offer insights into adapting to long-term changes in atmospheric dynamics.
- **Optimizing Transfer Learning for Real-Time Forecasting:** Real-time wind power forecasting necessitates models that are both accurate and computationally efficient. Transfer learning frameworks must be optimized to handle dynamic updates in incoming data streams, enabling near-instantaneous model adaptation without compromising performance. Research into online transfer learning methods, as highlighted by Al-Stouhi and Reddy [13], provides a starting point for these developments.

Future Extensions Building on the foundational work in this study, several promising directions can further enhance the utility and scope of LSTM-based transfer learning models:

- **Exploring Hybrid Architectures:** Combining graph neural networks (GNNs) with transformers offers a powerful approach to capturing spatial and temporal dependencies in wind patterns. GNNs excel at representing complex spatial relationships, while transformers are adept at learning long-range temporal dependencies. Preliminary studies [22] indicate the potential of such hybrid models in weather-related forecasting tasks.
- **Integrating High-Resolution Datasets:** High-resolution datasets, such as ERA5 reanalysis data, provide detailed spatial and temporal meteorological information that can improve model training and forecasting accuracy. Incorporating these datasets into transfer learning pipelines can help bridge the gap between coarse-grained historical data and fine-grained real-time applications.

These emerging directions underscore the dynamic nature of renewable energy forecasting research. Addressing these challenges will not only improve the precision of wind power predictions but also contribute to the broader goal of integrating renewable energy into global energy systems more effectively.

Chapter 7

PRACTICAL IMPACTS

7.1 Societal Benefits

The improvements in wind power forecasting achieved through this study have profound implications for societal and environmental progress. By enhancing the accuracy and reliability of forecasts, the proposed LSTM-based transfer learning framework addresses key challenges in energy management and contributes to global sustainability goals.

Energy Management Accurate wind power forecasts are vital for ensuring grid stability, particularly in energy systems with a high penetration of renewable sources. The intermittent nature of wind energy presents a challenge for maintaining a balanced supply-demand curve, often necessitating the use of non-renewable backup generators to compensate for fluctuations. By leveraging feature-rich datasets and applying transfer learning to adapt these insights to feature-poor regions, this study contributes to reducing the uncertainty in wind power generation. Recent findings [21] show that improved forecasts can significantly decrease operational costs by minimizing the need for reserve capacity and enhancing grid reliability.

Furthermore, better forecasting enables grid operators to plan ahead for fluctuations, reducing the need for costly real-time interventions. This translates into economic savings while ensuring energy security for consumers. The proposed methods also promote equitable access to reliable renewable energy by enabling under-monitored regions to benefit from state-of-the-art forecasting models developed in data-rich areas.

Environmental Goals The transition to renewable energy is a cornerstone of global decarbonization strategies aimed at mitigating climate change. Wind energy, as a key component of the renewable energy mix, can only reach its full potential if integrated efficiently into existing energy systems. Enhanced forecasting directly supports this integration by reducing curtailment—when excess energy is wasted due to grid limitations—and by optimizing energy storage utilization [8].

By enabling more accurate predictions of wind power availability, the proposed approach contributes to reducing greenhouse gas emissions associated with backup fossil fuel power plants. It also aligns with international commitments, such as the United Nations Sustainable Development Goal (SDG) 7, which emphasizes affordable, reliable, and sustainable energy for all. The adoption of transfer learning methods, as demonstrated in this study, accelerates the deployment of advanced forecasting technologies, particularly in developing regions, thereby fostering global equity in renewable energy adoption.

In summary, the practical impacts of this work extend beyond technical contributions, offering societal and environmental benefits that align with long-term global objectives. These advancements reinforce the role of machine learning as a transformative tool in addressing the complexities of renewable energy systems.

7.2 Industry Applications

The methodologies and results of this study present transformative opportunities for the renewable energy sector, particularly through the development of a real-time forecasting tool tailored for industry stakeholders. Such tools can enhance decision-making, improve operational efficiency, and facilitate the integration of wind energy into power grids.

Real-Time Forecasting Tool for Stakeholders The proposed framework can be implemented as a real-time forecasting tool designed to address the specific needs of energy managers, policymakers, and other key stakeholders. By leveraging the transfer learning approach, this tool can adapt models trained on feature-rich datasets to regions with sparse data, ensuring broader applicability across diverse geographic and operational contexts. Real-time capabilities are critical for enabling stakeholders to respond dynamically to fluctuations in wind power generation, optimizing resource allocation, and reducing reliance on fossil-fuel-based backup systems [5].

Key features of this forecasting tool include:

- **User-Friendly Interfaces:** The tool will be equipped with intuitive interfaces to facilitate its adoption by non-technical stakeholders. Energy managers can access actionable insights without requiring expertise in machine learning or forecasting models.
- **Visual Dashboards:** Interactive dashboards will provide real-time displays of forecasted wind power, prediction accuracy, and model robustness. These dashboards can help stakeholders identify trends, compare forecasts with actual outcomes, and make data-driven decisions. Previous studies, such as those by Costa et al. [4], emphasize the importance of visualization tools in operational decision-making within the renewable energy sector.

Enhancing Industry Efficiency and Reliability By integrating advanced machine learning techniques with real-time user interfaces, the proposed tool addresses a critical gap in the renewable energy industry. Specifically, it bridges the divide between technical advancements in forecasting and practical, actionable insights required by industry practitioners. This alignment of technical and operational priorities has the potential to:

- Reduce operational costs through more accurate scheduling and dispatch of energy resources [21].
- Minimize curtailment of wind energy, thereby increasing the efficiency of renewable energy utilization.
- Enhance stakeholder confidence in renewable energy systems, accelerating the transition to sustainable energy sources.

In summary, the implementation of this real-time forecasting tool offers significant value to the renewable energy industry by combining technical innovation with practical applicability. By making advanced forecasting accessible to a wider range of stakeholders, this work supports the broader adoption and integration of wind energy into global energy systems.

7.3 AI Governance and Ethical Considerations

The integration of machine learning (ML) models into wind power forecasting raises important governance and ethical challenges that must be addressed to ensure scalable, fair, and trustworthy deployment. This section highlights key issues related to governance frameworks and ethical concerns in the context of ML-driven renewable energy solutions.

Governance Challenges The deployment of ML models for large-scale wind power forecasting faces significant scalability barriers, including regulatory and technical constraints:

- **Regulatory Constraints:** The adoption of AI-driven forecasting tools often involves navigating complex regulatory frameworks that vary by region. These frameworks may impose restrictions on data sharing, model deployment, and operational oversight. For example, compliance with regional energy policies or data protection laws such as the General Data Protection Regulation (GDPR) in Europe [20] can complicate cross-border model training and application.
- **Technical Scalability:** Ensuring the scalability of ML models in diverse operational settings remains a challenge. Transfer learning approaches require careful calibration when applied to feature-poor datasets, as model performance can degrade when trained on heterogeneous data sources. Addressing these technical challenges demands robust mechanisms for domain adaptation and model evaluation across diverse geographic and climatic conditions [18].

Ethical Concerns Ethical considerations are critical in building trust and ensuring the equitable use of ML models in wind power forecasting. Key ethical issues include:

- **Fairness and Bias Mitigation:** ML models are susceptible to biases present in training datasets, which can lead to inequitable outcomes. For instance, regions with limited historical data may receive less accurate forecasts, perpetuating inequalities in renewable energy access. Techniques such as adversarial debiasing and balanced dataset augmentation [24] can help address these disparities.
- **Privacy Protection:** The use of sensitive meteorological and operational data raises concerns about data privacy. Ensuring data anonymization and secure data handling practices is essential to comply with legal and ethical standards while fostering collaboration between stakeholders.
- **Model Explainability:** The complexity of ML models, particularly deep learning architectures such as LSTMs, often results in opaque decision-making processes. This lack of transparency can undermine trust among stakeholders. Explainable AI (XAI) techniques, such as SHapley Additive exPlanations (SHAP) [14], provide insights into model predictions, enabling stakeholders to interpret results and make informed decisions with confidence.

Addressing these governance and ethical concerns is not only essential for the responsible deployment of ML models but also for ensuring their long-term acceptance and scalability in the renewable energy sector. Proactive efforts to promote transparency, fairness, and compliance will reinforce the role of AI as a transformative tool for global sustainability goals.

REFERENCES

[1] Seblewongale Mezgebu Ayene and Abdulkerim Mohammed Yibre. "Wind power prediction based on deep learning models: The case of Adama wind farm". In: *Heliyon* 10.21 (2024), e39579. ISSN: 2405-8440. DOI: 10.1016/j.heliyon.2024.e39579.

[2] Omar Benzohra et al. "Integrating Wind Energy into the Power Grid: Impact and Solutions". In: *Materials Today: Proceedings* 30.Part 4 (2020), pp. 987–992. ISSN: 2214-7853. DOI: 10.1016/j.matpr.2020.04.363.

[3] Jiayi Chen et al. "Simulating the effects of offshore wind energy policy on decarbonization and industrial growth in Taiwan: A system dynamics approach". In: *Energy for Sustainable Development* 71 (2022), pp. 490–504. ISSN: 0973-0826. DOI: 10.1016/j.esd.2022.10.018.

[4] Ana Maria Costa et al. "A Review on the Young History of the Wind Power Short-Term Prediction Tools and Their Operational Aspects". In: *Wind Energy* 11.4 (2008), pp. 377–391.

[5] Farhad Elyasichamazkoti and Abolhasan Khajehpoor. "Application of Machine Learning for Wind Energy from Design to Energy-Water Nexus: A Survey". In: *Energy Nexus* 2 (2021), p. 100011. ISSN: 2772-4271. DOI: 10.1016/j.nexus.2021.100011.

[6] Yuan Gao et al. "Adversarial Discriminative Domain Adaptation for Solar Radiation Prediction: A Cross-Regional Study for Zero-Label Transfer Learning in Japan". In: *Applied Energy* 359 (2024), p. 122685. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2024.122685.

[7] S. Hochreiter and J. Schmidhuber. "Long short-term memory". In: *Neural Computation* 9.8 (1997), pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.

[8] Bri-Mathias Hodge et al. "The Combined Value of Wind and Solar Power Forecasting Improvements and Electricity Storage". In: *Applied Energy* 214 (2018), pp. 1–15. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2017.12.120.

[9] Cong Huang et al. "Evolving Long Short-Term Memory Neural Network for Wind Speed Forecasting". In: *Information Sciences* 632 (2023), pp. 390–410. ISSN: 0020-0255. DOI: 10.1016/j.ins.2023.03.031.

[10] M. S. Islam Sajol et al. "Wind Power Prediction across Different Locations using Deep Domain Adaptive Learning". In: *2024 6th Global Power, Energy and Communication Conference (GPECOM)*. Budapest, Hungary, 2024, pp. 518–523. DOI: 10.1109/GPECOM61896.2024.10582569.

[11] T. Kara and A. D. Şahin. "Implications of Climate Change on Wind Energy Potential". In: *Sustainability* 15.20 (2023), p. 14822. DOI: 10.3390/su152014822.

[12] M. Khan et al. "Power Forecasting of Regional Wind Farms via Variational Auto-Encoder and Deep Hybrid Transfer Learning". In: *Electronics* 11.2 (2022), p. 206. DOI: 10.3390/electronics11020206.

[13] Yang Liu et al. "Imbalanced Data Classification: Using Transfer Learning and Active Sampling". In: *Engineering Applications of Artificial Intelligence* 117, Part B (2023), p. 105621. ISSN: 0952-1976. DOI: 10.1016/j.engappai.2022.105621.

[14] Scott M. Lundberg and Su-In Lee. “A Unified Approach to Interpreting Model Predictions”. In: *Advances in Neural Information Processing Systems* 30 (2017), pp. 4765–4774.

[15] Lee M. Miller and David W. Keith. “Climatic Impacts of Wind Power”. In: *Joule* 2.12 (2018), pp. 2618–2632. ISSN: 2542-4351. DOI: 10.1016/j.joule.2018.09.009.

[16] D. Millstein, S. Jeong, A. Ancell, et al. “A Database of Hourly Wind Speed and Modeled Generation for US Wind Plants Based on Three Meteorological Models”. In: *Scientific Data* 10 (2023), p. 883. DOI: 10.1038/s41597-023-02804-w.

[17] Neptune AI. *Transfer Learning Guide: Examples for Images and Text in Keras*. Accessed: 2023-12-09. n.d.

[18] Sinno Jialin Pan and Qiang Yang. “A Survey on Transfer Learning”. In: *IEEE Transactions on Knowledge and Data Engineering* 22.10 (2010), pp. 1345–1359.

[19] E. Sarmas, N. Dimitropoulos, V. Marinakis, et al. “Transfer Learning Strategies for Solar Power Forecasting Under Data Scarcity”. In: *Scientific Reports* 12 (2022), p. 14643. DOI: 10.1038/s41598-022-18516-x.

[20] Paul Voigt and Axel Von dem Bussche. *The EU General Data Protection Regulation (GDPR): A Practical Guide*. Cham, Switzerland: Springer, 2017.

[21] Y.-K. Wu et al. “Probabilistic Wind-Power Forecasting Using Weather Ensemble Models”. In: *IEEE Transactions on Industry Applications* 54.6 (Nov. 2018), pp. 5609–5620. DOI: 10.1109/TIA.2018.2858183.

[22] Zonghan Wu et al. “A Comprehensive Survey on Graph Neural Networks”. In: *IEEE Transactions on Neural Networks and Learning Systems* 32.1 (2021), pp. 4–24.

[23] Z. Xiao, F. Tang, and M. Wang. “Wind Power Short-Term Forecasting Method Based on LSTM and Multiple Error Correction”. In: *Sustainability* 15.4 (2023), p. 3798. DOI: 10.3390/su15043798.

[24] Richard S. Zemel et al. “Learning Fair Representations”. In: *Proceedings of the 30th International Conference on Machine Learning (ICML)*. PMLR, 2013, pp. 325–333.

[25] J. Zhou, X. Lu, Y. Xiao, et al. “SDWPF: A Dataset for Spatial Dynamic Wind Power Forecasting over a Large Turbine Array”. In: *Scientific Data* 11 (2024), p. 649. DOI: 10.1038/s41597-024-03427-5.

Appendix A

ADDITIONAL MATERIAL

A.1 Methodology for Causal Inference Using Transfer Learning

This appendix provides a detailed overview of the methodology employed in the study titled *"Improving Wind Power Forecasting with LSTM-Based Transfer Learning across Feature-Rich and Feature-Poor Datasets."* Figure A.1 illustrates the enhanced network of datasets, features, and stages involved in both forecasting and exploring causal relationships through transfer learning.

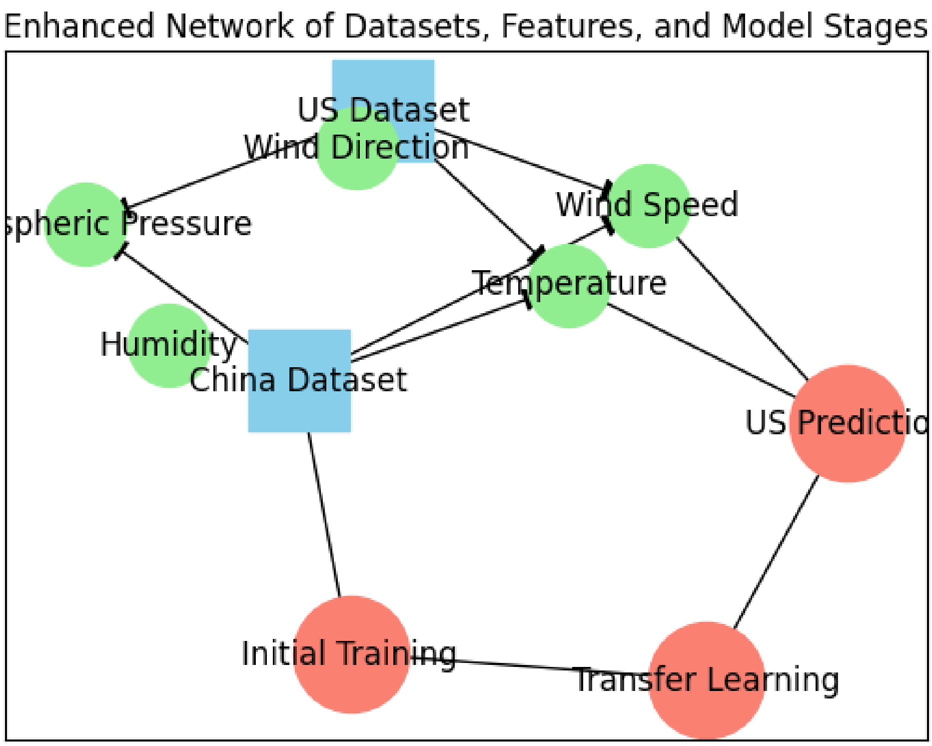

Figure A.1: Enhanced Network of Datasets, Features, and Model Stages

A.1.1 Datasets and Features

- **China Dataset:** Includes features such as Wind Speed, Temperature, and Atmospheric Pressure, utilized primarily in the initial training phase.

- **US Dataset:** Comprises similar features, with the focus shifting to Wind Speed and Temperature for the prediction phase in the US context.

A.1.2 Model Training Stages

- **Initial Training:** Utilizes the China Dataset to establish a baseline model. This phase is crucial for capturing the initial patterns and relationships within the data.
- **Transfer Learning:** In this stage, the model undergoes adaptation to apply the insights gained from the China Dataset to the US Dataset. This step tests the hypothesis that causal relationships identified in one geographical region can be generalized to another.

A.1.3 Causal Inference Discussion

The methodology supports a deeper investigation into causal inference by using the adapted model to assess whether similar causal mechanisms apply across different regions. This part of the study explores if predictive relationships hold under varied environmental conditions, assessing the robustness and transferability of causal findings.

A.1.4 Methodological Considerations

- The approach underscores the importance of a comprehensive setup that integrates multiple datasets and features, facilitating a robust analysis that extends beyond simple prediction tasks.
- It illustrates the need for rigorous model adaptation strategies to ensure that the causal relationships are not only preserved but are also meaningful when transferred across different contexts.

A.1.5 Conclusion

This appendix extends the discussion from prediction to causal inference, emphasizing how transfer learning can serve as a powerful tool for enhancing our understanding of causal relationships in wind power forecasting across diverse geographical settings.

A.2 Supplementary Materials

For more details, please visit the GitHub repository: STATS201-DKU-Fall2024 Final Project.