Enhancing Sea Level Pressure Simulation through Mutual Information Analysis of Meteorological Indicators and Regression Modeling

Kaifang Mu, *Member, IAENG*, Qiang Ai†, *Member, IAENG*, Xize Lu, Rui Zhang

***Abstract*—This paper focuses on the meteorological data of Qinghai Province in China and studies the simulation effects of regression models on sea level pressure (SLP). To validate the performance of the models, meteorological data from 2022 is used, and the data is divided into seasons: spring, summer, autumn, and winter. The models are trained using SLP data from different seasons to evaluate their simulation effects. The results indicate significant differences in the performance of the models across different seasons, with lower predictive accuracy in spring and winter, while autumn and summer yield relatively higher prediction performance. The model evaluation metrics include the correlation coefficient R, RMSE, MAE, and MAPE, which provide a comprehensive assessment of the model performance.**

***Index Terms*—sea level pressure, mutual information, regression model, day-by-day data, NOAA**

I. INTRODUCTION

CLIMATE change and human activities have profound impacts on ecosystems, making their scientific evaluation and research critically important [1]. The relationships among various climate variables are a key component of climate change research. Compared to individual climate variables, dependencies between climate variables may lead to earlier and more substantial deviations from natural variability [2]. Modeling variable combinations that contribute to extreme climate events is a complex interdisciplinary task. To this end, some studies propose using influence diagrams to define, map, analyze, model, and communicate the risks associated with such combinations [3]. Among numerous climate variables, the mean sea level pressure (MSLP), its spatiotemporal variability, and the occurrence of extreme events are essential for understanding the complexity of the climate system [4]. Sea level pressure has significant practical applications in climate system characterization, weather forecasting, and meteorological disaster warning. It is a key parameter reflecting atmospheric pressure and is influenced by factors such as altitude, latitude, temperature, and humidity, capturing the circulation patterns and variations in the lower troposphere [5]. Its changes have notable impacts on precipitation and temperature [6]. Studies have shown that SLP in the Western Pacific and East Asia significantly affects sea level changes, precipitation, and temperature in East Asia [7]. Therefore, continued research on the relationship between surface temperature and sea level pressure is necessary [8]. Sea level changes are closely related to global warming, which in turn affects sea level pressure in various regions. Additionally, regional sea level pressure is closely associated with air quality. Air quality reflects the chemical state of the atmosphere at specific times and locations. Like weather, air quality impacts everyone. Air pollutants include gases and particulates that may cause non-carcinogenic and/or carcinogenic adverse health effects [9]. Currently, studies have been conducted on sea level pressure, such as research in Istanbul and Turkey, which revealed the spatiotemporal variations of MSLP and identified relationships between MSLP and oceanic data such as sea surface temperature (SST) and sea level anomalies (SLA) [8]. However, there are still limited studies on the simulation and prediction of sea level pressure, requiring further exploration. In response, we conducted research on the simulation and prediction of sea level pressure.

Our study area is located in the Tibetan Plateau of China, where the high-altitude features are more sensitive to climate change [10]. This region has unique topography with non-zonal plateau climate characteristics, significantly impacting global climate [11]. The degradation of permafrost in this area is dramatically altering the hydrological conditions on a regional and even continental scale, attracting the attention of hydrologists, climatologists, ecologists, engineers, and policymakers [12]. Even small fluctuations in the climate of the Tibetan Plateau strongly affect the terrestrial ecosystem responses [13]. Therefore, it is essential to study climate indicators of the Tibetan Plateau. Sea level pressure in this region shows distinct regional and seasonal characteristics. Currently, limited studies exist on the interactions and responses between sea level pressure and other climate indicators in this area, necessitating further research. This study thus selects climate indicators from meteorological stations in Xining and surrounding areas to analyze them over time.

Climate simulation research primarily uses statistical methods and numerical models. For example, Nan Lyu et al. applied the maximum entropy model to assess the potential impact of climate change on Chinese grouse, predicting outcomes across three time periods under two greenhouse

Manuscript received October 19, 2024; revised July 4, 2025. This work was supported in part by the College Student Innovation Training Program of Qinghai Normal University in 2025 (Project Nos. qhnucxcy2025022, qhnucxcy2025026).

Kaifang Mu is a Postgraduate Student at the School of Geographical Sciences, Qinghai Normal University, Xining 810008, China (e-mail: kaifang.mu@aliyun.com).

Qiang Ai is a Postgraduate Student at the College of Computer, Qinghai Normal University, Xining 810008, China (Corresponding author to provide phone: +8617813140425; e-mail: qiang.ai@outlook.com).

Xize Lu is a Postgraduate Student at the School of Environmental Science and Engineering, Shandong University, Qingdao 266237, China (e-mail: 202212875@mail.sdu.edu.cn).

Rui Zhang is a Computer Scientist at the Aisess (Dalian) Computer Services Co., Ltd., Dalian, China (e-mail: zhangrui@essays.ltd).

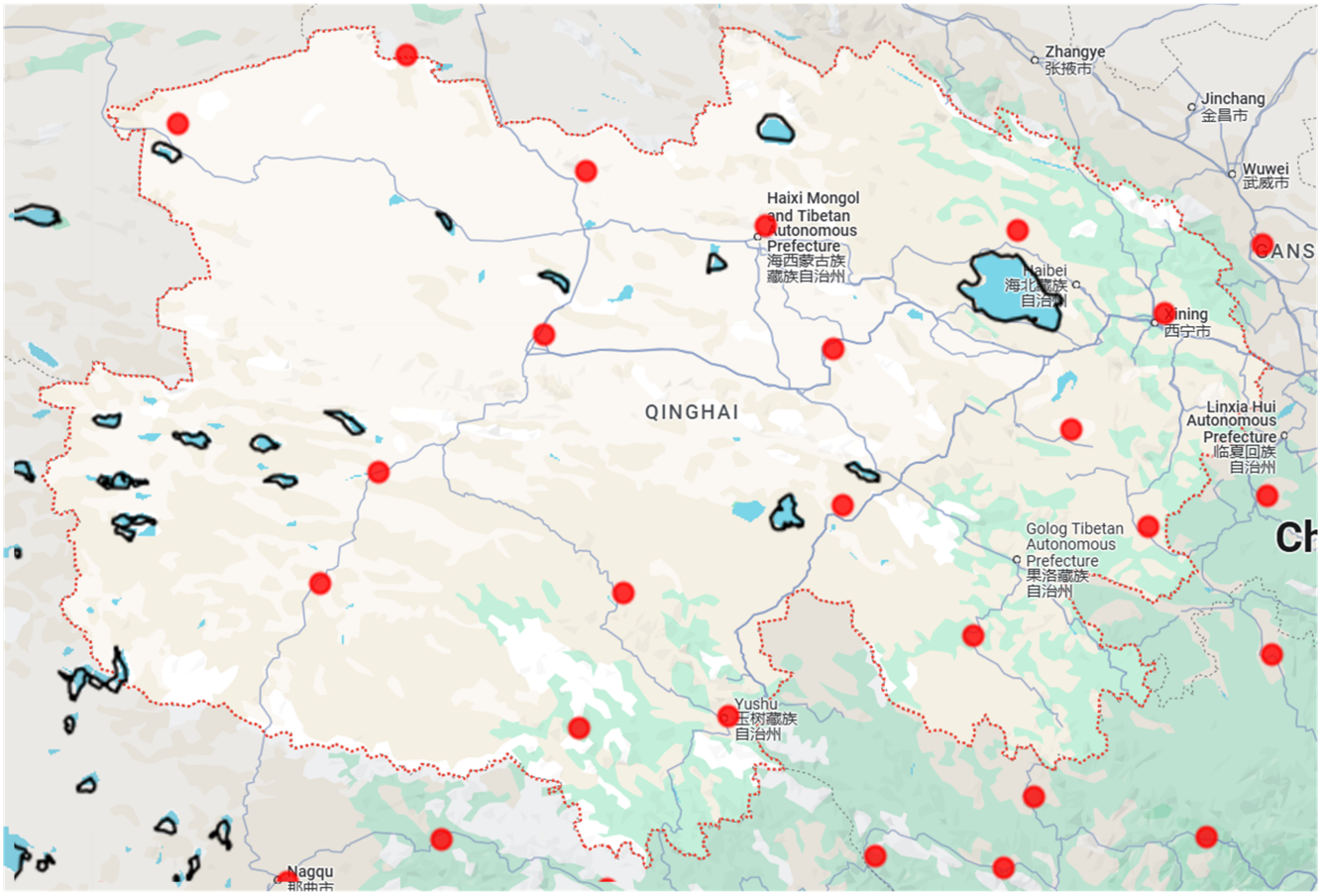

Fig. 1. Meteorological station location schematic.

gas emission scenarios [14]. Taylor et al. used the time series forecasting algorithm Prophet to examine temporal data characteristics and trends, projecting future dynamics [15]. Joanna et al. used a random forest-based ensemble data mining approach to identify features linearly related to air quality concentration, establishing a regression model for effective classification and prediction of air pollutants [16]. Other studies include machine learning-based tropical cyclone generation prediction models [17] and deep neural network models that automatically extract data features, achieving high predictive skill [18].Feature selection methods also include the correlation coefficient method [19] and principal component analysis (PCA) [20], but the former requires feature variables to be independent, and PCA is mainly applied to linear problems. SLP (Sea Level Pressure) serves as a core indicator of the atmospheric dynamic system, yet traditional physical models for its simulation exhibit high complexity, rendering real-time prediction challenging. Given that non-linear relationships may exist between climate indicators[21], this study uses mutual information analysis to select optimal features and applies an improved CatBoost [22], model to construct a regression model for predicting mean sea level pressure.Enhancing SLP simulation accuracy and achieving intelligent meteorological modeling are essential for advancing meteorological prediction capabilities.

II. METHODOLOGY

A. Overview of Study Area

This study focuses on Qinghai Province, China, as the research area. Fig. 1 illustrates the study area, highlighting its geographical features and significance.

Qinghai Province is located in the northeastern part of the Tibetan Plateau, characterized by its diverse topography and climatic conditions. The province is known for its high altitude, vast grasslands, and numerous lakes, including the famous Qinghai Lake. The unique geographical and meteorological characteristics of Qinghai make it an ideal location for studying the impacts of climate variability on environmental and ecological systems.

The region experiences a complex climate influenced by its altitude and geographical location, with significant variations in temperature and precipitation across different areas. This variability presents an opportunity to analyze meteorological data and its correlation with various environmental indicators.

In this study, we will utilize meteorological data collected from various stations across Qinghai Province to explore the relationships between sea level pressure and other meteorological indicators, contributing to a better understanding of climatic patterns in this unique region.

B. Data Details and Pre-processing

The meteorological station data used in this study is sourced from the National Centers for Environmental Information (NCEI) under the National Oceanic and Atmospheric

Administration (NOAA). The data used are the daily meteorological data from all meteorological stations in Qinghai Province in 2022, which includes various meteorological features summarized in the Table I

These features are essential for analyzing the relationships between sea level pressure and other meteorological indicators in the context of this study.

C. Mutual Information-Based Selection of Features

Feature extraction is a crucial step in the data preprocessing pipeline, particularly in the context of supervised learning. The goal is to select the most informative features that contribute significantly to the prediction of the target variable. One effective method for measuring the relevance of features is Mutual Information (MI).

1) Introduction to Mutual Information: Mutual Information quantifies the amount of information obtained about one random variable through another random variable. It measures the dependency between variables and is particularly useful when dealing with non-linear relationships. For two random variables X and Y, the mutual information $I(X;Y)$ can be defined as:

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} P(x,y) \log \left(\frac{P(x,y)}{P(x)P(y)} \right) \quad (1)$$

Where:

- $P(x,y)$ is the joint probability distribution of X and Y,
- $P(x)$ and $P(y)$ are the marginal probabilities of X and Y.

A higher MI value indicates a greater amount of shared information, suggesting a stronger relationship between the variables.

2) Feature Selection Process: In this study, we employed mutual information to evaluate the relevance of meteorological features in predicting sea level pressure (SLP). The following steps outline our feature selection process:

1) **Data Preparation**: We began by collecting a dataset containing various meteorological indicators along with SLP. The features selected for analysis included 'ELEVATION', 'TEMP', 'DEWP', 'STP', 'VISIB', 'WDSP', 'MXSPD', 'GUST', 'MAX', 'MIN', 'PRCP', 'SNDP', and 'FRSHTT'.
2) **Calculating Mutual Information**: We computed the mutual information between each feature and the target variable (SLP) using the `mutual_info_regression` function from the `sklearn` library. This function evaluates the mutual information score for each feature relative to SLP.
3) **Feature Ranking**: The features were ranked based on their mutual information scores. Features with higher scores were considered more informative and relevant for predicting SLP.
4) **Visualization**: To provide a clear understanding of the relationship between features and SLP, a mutual information matrix was generated and visualized using a heatmap. This graphical representation highlighted the strength of associations between each feature and the target variable.

3) Results: The results show in Fig. 2, indicated that the heatmap presents mutual information scores as color-coded tiles in a multi-column horizontal layout, making the mutual information distribution of each feature readily apparent. Darker regions (e.g., columns TEMP, MAX, MIN) correspond to the highest scores, while lighter areas (e.g., SNDP, MXSPD) indicate lower information content. This figure intuitively demonstrates the differences in mutual information strength between features, assisting readers in understanding why features with low mutual information values were excluded from modeling, thus simplifying the model structure and mitigating the curse of dimensionality. Fig. 3, a bar plot, displays the mutual information scores between candidate meteorological features (TEMP, MAX, MIN, DEWP, STP, FRSHTT, etc.) and sea level pressure (SLP), ranked from highest to lowest. Temperature (TEMP) exhibits the highest mutual information score, indicating that it contains the richest information for SLP prediction. This is followed by maximum temperature (MAX) and minimum temperature (MIN), reflecting that temperature fluctuations significantly influence air pressure. Dew point temperature (DEWP), sea level standard pressure (STP), and frost indication (FRSHTT) also show secondary but still significant information content. This result validates the rationality of focusing on these six factors during the feature selection phase, providing a solid data foundation for subsequent model training.

By employing mutual information as a feature selection technique, we ensured that our predictive model was built on the most relevant meteorological indicators, thereby enhancing the accuracy and interpretability of the results.

D. Regression Models-Based Numerical Simulation of SLP

CatBoost (Categorical Boosting) is an open-source machine learning library released in 2017, belonging to the family of Boosting algorithms. Like XGBoost and LightGBM, CatBoost is regarded as one of the significant improvements in gradient boosting decision tree methods. However, unlike the other two Boosting algorithms, CatBoost specifically handles categorical features and uses oblivious trees as its base model, supporting categorical variables. This algorithm mitigates the risk of overfitting by balancing the structure of the trees. In oblivious trees, the index of each leaf node can be encoded using a binary vector of the same length as the tree's depth. The CatBoost algorithm first performs binary transformation on all floating-point features, statistical information, and one-hot encoded features, and then uses these binary features for model predictions. Additionally, CatBoost effectively addresses gradient shift and prediction issues, demonstrating superior model performance compared to LightGBM and XGBoost. The complete tree-building process of the CatBoost algorithm is illustrated in Table II.

The CatBoost algorithm features a powerful gradient boosting framework that automatically handles categorical features in a special way. This capability makes CatBoost particularly convenient and efficient when dealing with datasets that contain many categorical features. It employs a sorting-based approach, which reduces the need for hyperparameter tuning and mitigates the risk of overfitting, thereby enhancing the model's performance. Additionally, CatBoost

TABLE I
METEOROLOGICAL FEATURES COLLECTED

Feature	Description
STATION	Station identifier
DATE	Date of the observation
LATITUDE	Latitude of the station
LONGITUDE	Longitude of the station
ELEVATION	Elevation of the station (meters)
NAME	Name of the weather station
TEMP	Air temperature (°C)
TEMP_ATTRIBUTES	Attributes of the temperature measurement
DEWP	Dew point temperature (°C)
DEWP_ATTRIBUTES	Attributes of the dew point measurement
SLP	Sea level pressure (hPa)
SLP_ATTRIBUTES	Attributes of the sea level pressure measurement
STP	Station pressure (hPa)
STP_ATTRIBUTES	Attributes of the station pressure measurement
VISIB	Visibility (kilometers)
VISIB_ATTRIBUTES	Attributes of the visibility measurement
WDSP	Wind speed (m/s)
WDSP_ATTRIBUTES	Attributes of the wind speed measurement
MXSPD	Maximum sustained wind speed (m/s)
GUST	Wind gust (m/s)
MAX	Maximum temperature (°C)
MAX_ATTRIBUTES	Attributes of the maximum temperature measurement
MIN	Minimum temperature (°C)
MIN_ATTRIBUTES	Attributes of the minimum temperature measurement
PRCP	Precipitation (mm)
PRCP_ATTRIBUTES	Attributes of the precipitation measurement
SNDP	Snow depth (cm)
FRSHTT	Frost condition

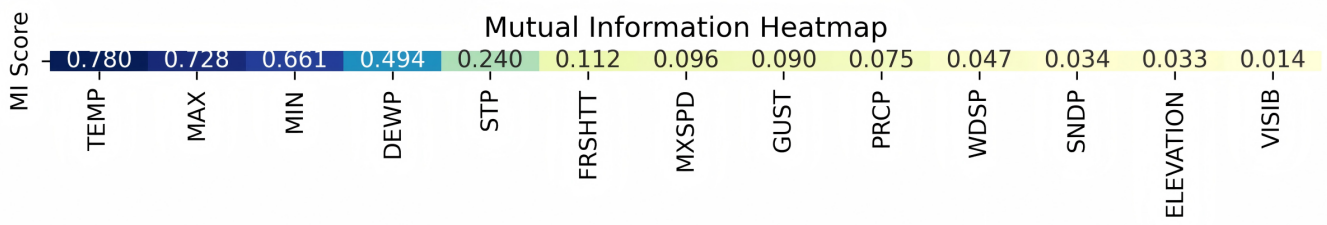

Fig. 2. Mutual Information Heatmap.

TABLE II
CATBOOST ALGORITHM PROCESS

Algorithm:	CatBoost
Input:	x, y, I, L, s, $Mode$
Initialization:	n, i, i, i, $\{(x_i, y_i), ...\}$, α
1:	$\sigma \leftarrow$ any permutation of $[1, n]$, $r = 0, \ldots, s$;
2:	$M(i) \leftarrow 0$ for $i = 1, \ldots, n$;
3:	If $Mode =$ Plain, then
	$M(i) \leftarrow 0$ for $r = 1, \ldots, s$;
4:	If $Mode =$ Ordered, then
	for $j \leftarrow 1$ to $[\log n]$ do
	$M(i) \leftarrow 0$ for $r = 1, \ldots, s$;
5:	for $t \leftarrow 1$ to I
	$T_t \leftarrow$ build tree $(M, \{(x, y)\}, L, \sigma, \text{Mode})$;
	for $i \leftarrow 1, \ldots, n$ do
	$\text{leaf}_i \leftarrow \text{GetLeaf}(x_i)$;
	$(0, 0)$ grad $\leftarrow \text{CalcGradient}(L, M, y)$;
	foreach leaf j in T_t do
	$b_{\text{avg}}(\text{grad}_j) \leftarrow \frac{1}{\lvert i:\text{leaf}_i = j \rvert} \sum_{i:\text{leaf}_i = j} \text{grad}_i$;
	$M_i \leftarrow M_i + a \cdot b_{\text{avg}}(\text{grad}_j)$ for $i = 1, \ldots, n$;
Return:	$\{(x, y, \text{leaf})\}_1^n$;
	$\text{GetLeaf}(x, T)$ ApplyMode j;

utilizes a computation method that arranges multiple learning datasets, allowing it to calculate statistical data for categorical features during the training process.

1) Improving CatBoost Algorithm with Stacking Model Ensemble Method: Ensemble learning accomplishes learning tasks by constructing and combining multiple learners. Ensemble classification methods process different individual classifiers similarly to enhance classification efficiency. This scenario is regarded as the precise assignment of objects, and this combination is achieved by aggregating the classification results obtained from various classifiers, resulting in a final classifier with optimal predictive capability. To improve the model's accuracy in predicting the meteorological dataset, this study employs the Stacking model fusion method to enhance the CatBoost algorithm. The first layer uses Random Forest, XGBoost, and CatBoost as base learners for prediction. The second layer takes the predictions from the base learners as input for a multiple linear regression model, ultimately generating the prediction results for the meteorological data. The Stacking algorithm flow is illustrated in Table III.

In the Stacking method, there are two stages of models. The first-stage model consists of base models that take the original training set as input, allowing for multiple base models to be selected for training. In the second-stage model, the training set for the meta-model is derived from the predictions of the base models on the original training set, while the test set consists of the predictions of the base models on the original test set. To avoid the risk of

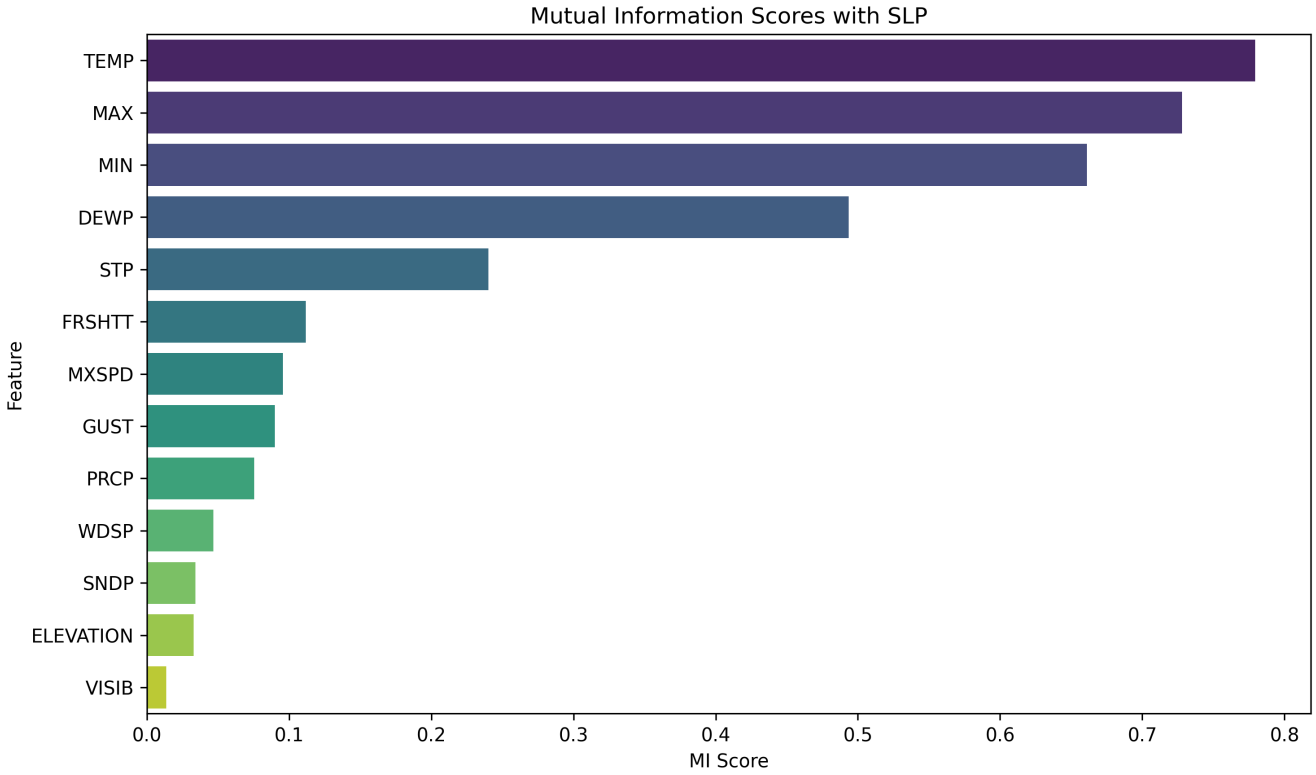

Fig. 3. Mutual Information Score Bar Plot.

TABLE III
ALGORITHM FLOW OF THE STACKING MODEL FUSION METHOD

Algorithm: Stacking
Input: Training data $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$
Process: Given the ensemble classifier H
1: **Step 1: Learn the first-level classifiers**
2: for $t = 1, 2, \ldots, T$
3: Learn a base classifier h_t based on D
4: End of Step 1
5: **Step 2: Construct a new dataset** D'
6: for $i = 1, 2, \ldots, m$
7: Construct a new dataset containing $\{(\mathrm{x}'_i, y_i)\}$, where $x'_i = \{h_1(x_i), h_2(x_i), \ldots, h_T(x_i)\}$
8: End of Step 2
9: **Step 3: Learn the second-level classifier**
10: Learn a new classifier based on the newly constructed dataset
Output: $H(x) = (h_1(x), h_2(x), \ldots, h_T(x))$

overfitting during training, the dataset for the meta-learner is generated from the base learners. Generally, to generate training samples for the meta-learner, unused samples from the base learners can be utilized through methods such as cross-validation. In this study, k-fold cross-validation is one of the sampling reuse methods adopted. The initial training set D is randomly divided into k subsets of similar size $D_1, D_2, \ldots, D_k$. In each round of validation, the test set D_j and training set D'_j are determined. To achieve the training of the meta-learner, T base learning algorithms are selected, and in each round, the base learner h_t is trained on D_j using the T-th learning algorithm and generates a model. Specifically, for each sample x_i in D_j, its training example part is $z_i = (z_1, z_2, \ldots, z_T)$, and the label part is y_i. Thus, at the end of the entire cross-validation process, the training set for the meta-learner generated from these T base learners is $\{(z_i, y_i)\}_{i=1}^{m}$, and then D' will be used to train the meta-learner.

2) Improving the Loss Function of the CatBoost Algorithm: Focal loss (FL) was originally proposed by Kaiming He [23] and is primarily used in the field of image processing to address issues related to dense object detection tasks and model performance problems caused by data imbalance. This section starts with the cross-entropy loss function used in CatBoost, analyzes the data imbalance problem, provides the definition and formula of the focal loss function, and describes the process of improving the CatBoost algorithm's loss function using focal loss. The formula for the cross-entropy loss function is:

$$\text{Loss} = L(y, \hat{p}) = -y \log(\hat{p}) - (1 - y) \log(1 - \hat{p}) \quad (2)$$

Where $\hat{p}$ is the predicted class probability and y is the true label, corresponding to 0 and 1 in binary classification. The formula for the cross-entropy loss function can be derived as follows:

$$L(y,\hat{p}) = \begin{cases} -\log(\hat{p}) & \text{if } y = 1 \\ -\log(1-\hat{p}) & \text{if } y = 0 \end{cases} \tag{3}$$

By defining

$$\hat{p}_i = \begin{cases} \hat{p} & \text{if } y = 1 \\ 1-\hat{p} & \text{if } y = 0 \end{cases} \tag{4}$$

we can express the cross-entropy loss function uniformly as:

$$L_{\text{ce}} = -\log(\hat{p}) \tag{5}$$

First, we make an initial improvement to the cross-entropy loss function by introducing a parameter adjustment factor $\gamma(1-\hat{p})$. This factor means that if the probability of correct classification is higher, the adjustment factor $\gamma(1-\hat{p})$ becomes smaller, indicating that the loss weight for correctly classified samples is reduced. Conversely, if the classification is incorrect, the loss weight for those samples increases. The formula for this first improvement is as follows:

$$L_{\text{fl}}(y,\hat{p}) = \begin{cases} -\gamma(1-\hat{p})\log(\hat{p}) & \text{if } y = 1 \\ -\gamma(1-\hat{p})\log(1-\hat{p}) & \text{if } y = 0 \end{cases} \tag{6}$$

Next, we proceed with a second improvement to the cross-entropy loss function by introducing a weight factor α_i. The purpose of adding the weight factor α_i is to reduce the weight of certain samples, primarily to address the issue of sample imbalance. The formula for this second improvement is as follows:

$$L_{\text{fl}}(y,\hat{p}) = \begin{cases} -\gamma_i\alpha_i\hat{p}\log(\hat{p}) & \text{if } y = 1 \\ -\gamma_i\alpha_i(1-\hat{p})\log(1-\hat{p}) & \text{if } y = 0 \end{cases} \tag{7}$$

Similarly, we can express the Focal Loss function uniformly as:

$$L_{\text{fl}} = (1-\hat{p})\log(\hat{p}) - \gamma_i\alpha_i\hat{p} \tag{8}$$

E. Mode results Construction: Stacking Regression

To harness the complementary strengths of multiple algorithms, we employ a two-stage stacking ensemble:

First-Stage Base Learners: The first layer consists of three distinct regression models, each trained on the six features selected by mutual information: {TEMP, MAX, MIN, DEWP, STP, FRSHTT}.

leftmargin=1.5em

- **Random Forest (RF)**
  An ensemble of decision trees using bootstrap aggregation. RF mitigates overfitting and captures complex nonlinear relationships.
- **XGBoost**
  An optimized gradient-boosting framework with built-in regularization and column sampling. It offers rapid convergence and robust generalization.
- **Gradient Boosting Decision Trees (GBDT)**
  A classical boosting tree model that excels at fitting complicated patterns but requires careful tuning to prevent overfitting.

Each base learner is configured with 100 trees (or boosting rounds) and fitted on the training set. Their out-of-fold predictions on validation splits become inputs to the meta-learner.

Second-Stage Meta Learner: The second layer stacks the predictions from the base learners, $\hat{y}_{\text{RF}}$, $\hat{y}_{\text{XGB}}$, $\hat{y}_{\text{GBDT}}$, and applies a multiple linear regression:

$$\hat{y} = \beta_1\,\hat{y}_{\text{RF}} + \beta_2\,\hat{y}_{\text{XGB}} + \beta_3\,\hat{y}_{\text{GBDT}} + \beta_0.$$

Using 2022 data, the fitted coefficients are:

$$\hat{y} = 1.673\,\hat{y}_{\text{RF}} + 1.215\,\hat{y}_{\text{XGB}} - 3.198\,\hat{y}_{\text{GBDT}} + 1357.312.$$

Interpretation: The positive weights on RF and XGBoost indicate their strong predictive contributions, whereas the negative weight on GBDT serves to correct its systematic overestimation. The intercept 1357.312 represents the baseline SLP when the base-learner outputs are zero. This stacking framework combines the robustness of RF, the efficiency of XGBoost, and the flexibility of GBDT, yielding improved forecasting accuracy overall.

III. EXPERIMENTS

A. Evaluation Metrics

In this study, the correlation coefficient R, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) are selected as evaluation standards for the model simulation results.

The correlation coefficient R is defined as:

$$R = \frac{\text{cov}(y_i,\hat{y}_i)}{\sqrt{\text{var}(y_i)\cdot\text{var}(\hat{y}_i)}} \tag{9}$$

The formula for Root Mean Square Error (RMSE) is:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2} \tag{10}$$

The formula for Mean Absolute Error (MAE) is:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i-\hat{y}_i| \tag{11}$$

The formula for Mean Absolute Percentage Error (MAPE) is:

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \times 100 \tag{12}$$

Where y_i is the observed value; $\hat{y}_i$ is the simulated value; N is the total number of samples; $\text{cov}(y_i,\hat{y}_i)$ is the covariance between y_i and $\hat{y}_i$; $\text{var}(y_i)$ denotes the variance of y_i; and $\text{var}(\hat{y}_i)$ denotes the variance of $\hat{y}_i$.

B. Experimental Analysis of Model Validation Results

1) Model Evaluation: Leave-One-Season-Out (LOSO) Method: To thoroughly investigate the cross-seasonal generalization capability of the model, this study employs the "Leave-One-Season-Out (LOSO)" cross-validation method. The specific approach is as follows: in each iteration, the data from one season (spring, summer, autumn, or winter) is designated as the test set, while the data from the remaining three seasons serve as the training set. This process is repeated four times, and finally, the results of the four sets of tests are aggregated. This strategy simulates the real-world scenario where only meteorological data from the past three seasons are available for predicting the SLP of the next season, effectively evaluating the model's stability and adaptability under different seasonal meteorological conditions.

Based on LOSO validation, the R^2, RMSE, MAE, and MAPE for each season were calculated, with results shown in Table IV.For summer and autumn, the R^2 values both exceed 0.7, RMSE is below 5 hPa, and MAPE is lower than 0.4%, indicating the model exhibits excellent fitting performance for SLP in warm seasons. In contrast, winter and spring show R^2 values below 0.04, RMSE exceeding 10 hPa, and MAPE ranging between 1% and 1.5%. This suggests the model has a weak response under extreme and cold climatic conditions, struggling to capture complex atmospheric pressure variations.The model's superior performance in summer and autumn may be attributed to the larger variability and more pronounced trendiness of SLP during this period, which provides the model with stronger signal training capability. Additionally, the significant amplitude of changes in some factors during this period (such as temperature and dew point) significantly influences the model, making their impacts more pronounced.This seasonal discrepancy indicates that SLP exhibits greater regularity in warm seasons, with feature variables demonstrating stronger explanatory power for SLP. In cold seasons, however, atmospheric pressure changes are more complex, influenced by sudden weather events and boundary layer processes.

TABLE IV
MODEL EVALUATION METRICS BY SEASON

Season	R^2	RMSE	MAE	MAPE
1 Winter	0.0002	326.68	24.77	1.38
2 Spring	0.0340	9.13	7.38	0.73
3 Summer	**0.7276**	3.13	2.47	0.25
4 Autumn	**0.7572**	4.10	3.11	0.31

2) Model Error and Stability Analysis: The residual distribution of the test set is shown in Fig. 4 .The histogram displays the frequency distribution of prediction errors (predicted value minus true value) for all test samples (across all four seasons), overlaid with a kernel density estimation curve. This figure is used to examine whether the residuals follow a normal approximation and whether systematic bias exists. The error distribution is generally bell-shaped and symmetric, with a slight positive skew, indicating that the model slightly overestimates SLP in most cases. However, the overall errors are concentrated within the range of ± 10 hPa, and extremely large errors ($|error| > 50$ hPa) occur with extremely low frequency, demonstrating good model stability. This result suggests that although errors are relatively large in individual seasons, the overall residual structure of the model remains stable, and no significant overfitting or underfitting is observed.

3) Explanation of Seasonal Differences: As shown in Fig. 5, the box plots of predicted values across different seasons reveal the seasonal differences in the model output. This box plot compares the predicted SLP distributions of the stacking model in four seasons: winter (1), spring (2), summer (3), and autumn (4). The predicted medians in summer and autumn are close to the observed medians, with smaller box heights and shorter whiskers, indicating lower volatility and more stable predictions in these two seasons. In contrast, the boxes are taller and the whiskers are longer in winter and spring, suggesting a wider error distribution and occasional large deviations in these two seasons. Thus, there are obvious differences in the model's seasonal adaptability, with the best performance in summer and autumn. Considering the climatic characteristics of Qinghai Province, the atmospheric stratification is relatively stable in summer and autumn, and the pressure field is more continuously influenced by temperature and humidity. In winter and spring, however, the significant influence of the southward movement of cold air and the topographic lifting effect leads to rapid and drastic changes in air pressure, increasing the prediction difficulty.

Additionally, the overall performance of the model simulation follows a pattern of gradually decreasing effectiveness from cold to warm seasons. When combined with the study's climatic indicators, mutual information results show that temperature-related metrics and seasonal characteristic indices (such as meteorological condition indicators) play a dominant role. Further considering the impact of seasonal activities like sand-dust events—under the influence of Qinghai Province's four-season meteorological characteristics—on sea level pressure (SLP), the discrepancies in the SLP simulation model may also be attributed to such seasonal activity patterns.

4) Fitting Performance of the Model in Common Pressure Ranges: As shown in Fig. 6, to rigorously assess the fit of the model within the typical SLP range, we applied an automatic zoom on the 'Observed vs. Predicted' scatter plot: the 1st and 99th percentiles of both observed and predicted values define the core domain, within which points are plotted along with the reference line $y = x$.

The resulting point cloud concentrates between approximately 990 and 1045 hPa, with most points lying just above or below the line, indicating minimal overall bias. Predictions mejor when observed SLP < 1030 hPa and slightly overestimate or dampen when observed > 1030 hPa.

A horizontal band near 1010 hPa reveals a 'flat top' effect, likely due to the convergence of certain base or meta-learner outputs at that pressure. This approach cleanly exposes residual structure and bias patterns within the most relevant SLP interval.

IV. DISCUSSION

In this study, we propose a sea-level pressure (SLP) prediction framework based on mutual information feature selection and Stacking ensemble regression, and comprehensively validate it using daily observational data from Qinghai Province. The specific discussions are as follows:

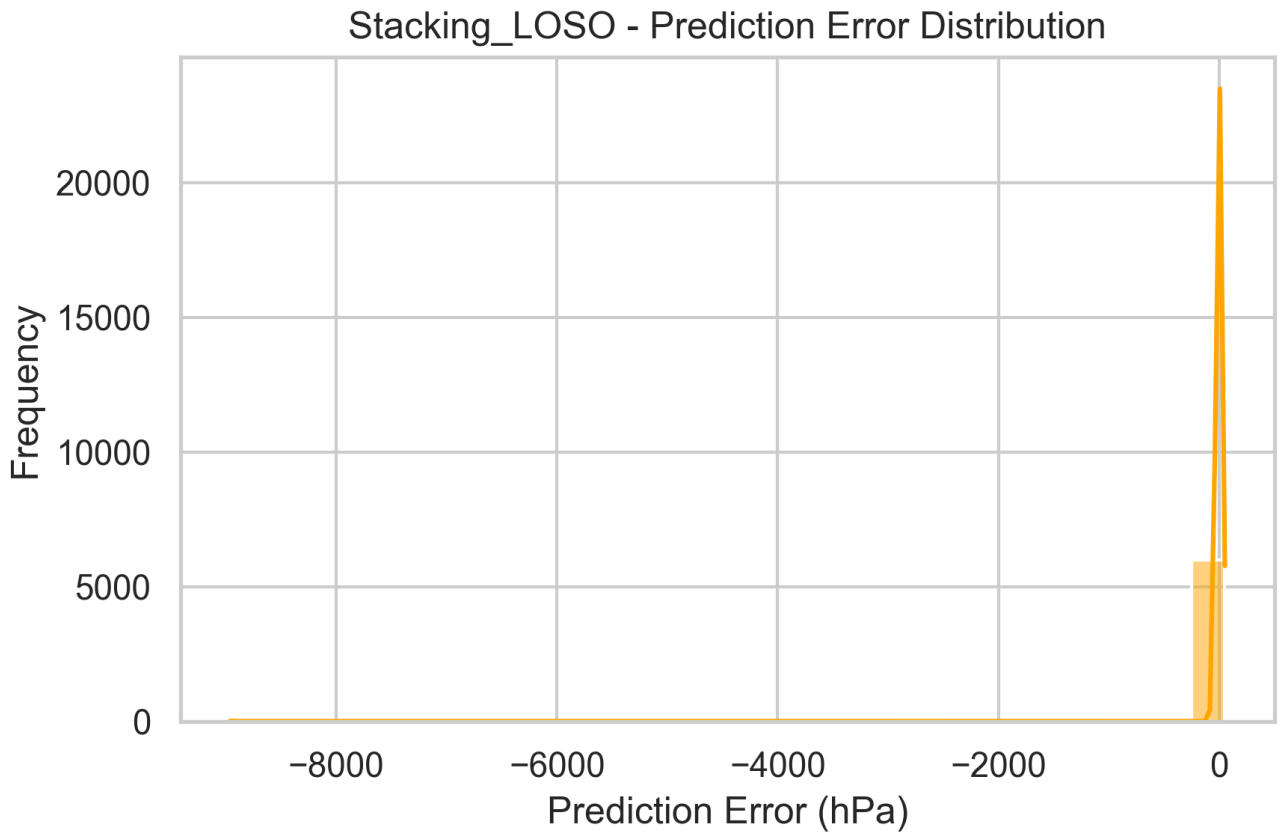

Fig. 4. Stacking LOSO error dist.

1) Feature Selection and Model Interpretability: Mutual information analysis effectively reduced dimensionality by selecting six key predictors of SLP (TEMP, MAX, MIN, DEWP, STP, FRSHTT), balancing model simplicity and performance. This step ensures that the subsequent stacking framework operates on high-information variables, enhancing interpretability.

2) Ensemble Modeling Performance: The two-stage stacking ensemble—combining RF, XGBoost, and GBDT followed by a linear meta-learner—demonstrated robust performance in summer and autumn, while yielding limited accuracy in spring and winter. These seasonal disparities underscore the model's ability to capture non-linear dependencies under stable warm-season conditions, but also its challenges in highly volatile cold seasons.

3) Seasonal Variability and Climatic Interpretation: Summer and autumn in Qinghai feature relatively stable synoptic conditions and gradual diurnal temperature changes, providing strong predictive signals. In contrast, winter and spring are dominated by sudden cold-air outbreaks and orographic lifting, leading to rapid pressure fluctuations that exceed the model's current explanatory capacity.

4) Error Structure and Model Stability: Residual distributions are approximately symmetric and centered around zero, indicating minimal systematic bias. The scarcity of extreme residuals further attests to the model's stability, though elevated cold-season errors suggest the need for additional predictors such as teleconnection indices or lagged features.

5) Limitations and Future Work: The present framework does not incorporate large-scale climate drivers (e.g., ENSO, NAO) or advanced time-series architectures (e.g., LSTM). Future research should explore these factors, along with spatial correlation models, to enhance performance across all seasons.

6) Practical Implications: The proposed approach provides an interpretable and efficient tool for regional SLP forecasting, particularly valuable for operational meteorological services during high-impact warm-season events.

V. Conclusion

In this study, we developed a two-stage stacking regression framework combined with mutual information–based feature selection to simulate daily sea level pressure (SLP) in Qinghai Province. Key findings and contributions are as follows:

1) **Feature Importance:** Mutual information analysis identified air temperature (TEMP), daily maximum temperature (MAX), daily minimum temperature (MIN), and dew point temperature (DEWP) as the most informative predictors of SLP. These variables capture both thermal and moisture dynamics critical to pressure variations.
2) **Modeling Strategy:** The stacking ensemble, comprising Random Forest, XGBoost, and GBDT as first-stage base learners, and a multiple linear regression as the second-stage meta learner, effectively integrates complementary strengths of each algorithm.

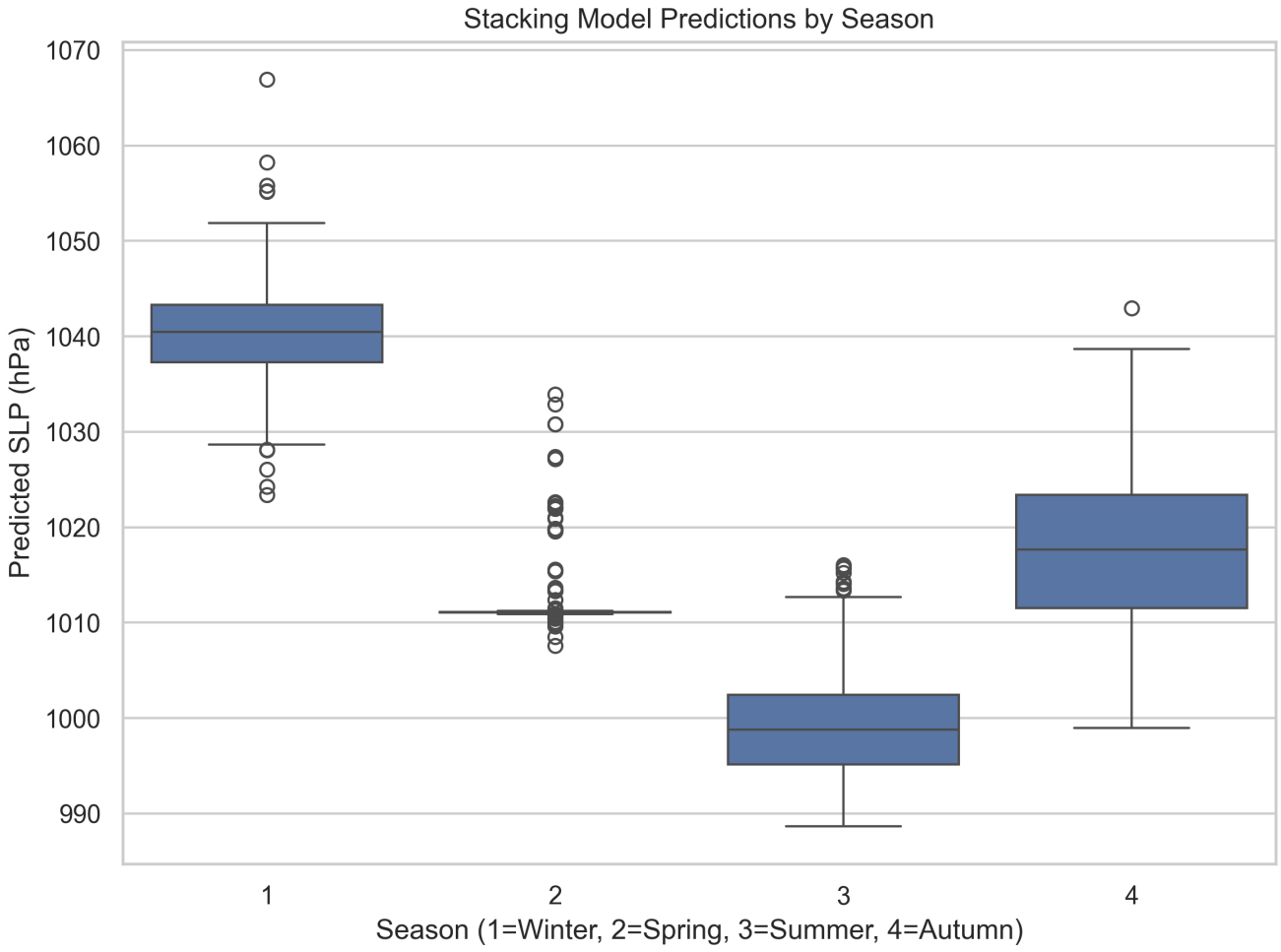

Fig. 5. seasonal predictions boxplot.

3) **Validation Scheme:** Leave-One-Season-Out (LOSO) cross-validation demonstrated strong predictive performance in summer and autumn, while revealing limitations in spring and winter. This emphasizes the need for season-specific tuning in operational forecasting.
4) **Practical Implications:** The proposed approach offers a robust and interpretable tool for regional SLP prediction, aiding meteorological services in the Tibetan Plateau region. Its demonstrated accuracy in warm seasons suggests potential for real-time application during high-impact weather periods.

REFERENCES

[1] Y. Liu, Z. Zhang, L. Tong *et al.*, "Assessing the effects of climate variation and human activities on grassland degradation and restoration across the globe," *Ecological Indicators*, vol. 106, p.105504, 2019.
[2] C. R. Mahony and A. J. Cannon, "Wetter summers can intensify departures from natural variability in a warming climate," *Nature communications*, vol. 9, no. 1, p.783, 2018.
[3] M. Leonard, S. Westra, A. Phatak, M. Lambert, B. van den Hurk, K. McInnes, J. Risbey, S. Schuster, D. Jakob, and M. Stafford-Smith, "A compound event framework for understanding extreme impacts," *Wiley Interdisciplinary Reviews: Climate Change*, vol. 5, no. 1, pp113–128, 2014.
[4] S. Patil, H. Singh, S. Bansod, and N. Singh, "Trends in extreme mean sea level pressure and their characteristics during the summer monsoon season over the indian region," *International journal of remote sensing*, vol. 32, no. 3, pp701–715, 2011.
[5] S. L. Durden, "Trends in intense typhoon minimum sea level pressure," *Atmosphere*, vol. 3, no. 1, pp124–131, 2012.
[6] C. G. Piecuch, P. R. Thompson, and K. A. Donohue, "Air pressure effects on sea level changes during the twentieth century," *Journal of Geophysical Research: Oceans*, vol. 121, no. 10, pp7917–7930, 2016.
[7] Y. Wang, D. Jiang, and W. Hua, "Projection of typhoon-related environmental fields in the western north pacific," *Chinese Journal of Atmospheric Sciences*, vol. 44, no. 3, pp552–564, 2020.
[8] F. Islek and Y. Yuksel, "Spatio-temporal long-term evaluations of the mean sea level pressure, sea level change, and sea surface temperature over two enclosed seas," *Regional Studies in Marine Science*, vol. 66, p.103130, 2023.
[9] Y. Zhang, M. Bocquet, V. Mallet, C. Seigneur, and A. Baklanov, "Real-time air quality forecasting, part i: History, techniques, and current status," *Atmospheric environment*, vol. 60, pp632–655, 2012.
[10] J. Wang, X. Zhang, and Y. Gao, "Research status and prospects of the relationship between vegetation dynamics and environmental factors on the qinghai-tibet plateau," *Earth Science Frontiers*, vol. 28, no. 4, pp70–82, 2021.
[11] H.-y. YU and J.-c. XU, "Effects of climate change on vegetations on qinghai-tibet plateau: A review." *Chinese Journal of Ecology*, vol. 28, no. 04, p.747, 2009.
[12] H. Gao, J. Wang, Y. Yang, X. Pan, Y. Ding, and Z. Duan, "Permafrost hydrology of the qinghai-tibet plateau: A review of processes and modeling," *Frontiers in Earth Science*, vol. 8, p.576838, 2021.
[13] J. A. Klein, J. Harte, and X.-Q. Zhao, "Experimental warming causes large and rapid species loss, dampened by simulated grazing, on the tibetan plateau," *Ecology Letters*, vol. 7, no. 12, pp1170–1179, 2004.
[14] N. Lyu and Y.-H. Sun, "Predicting threat of climate change to the chinese grouse on the qinghai—tibet plateau," *Wildlife biology*, vol. 20, no. 2, pp73–82, 2014.
[15] S. J. Taylor and B. Letham, "Forecasting at scale," *The American Statistician*, vol. 72, no. 1, pp37–45, 2018.
[16] J. A. Kamińska, "The use of random forests in modelling short-term air pollution effects based on traffic and meteorological conditions: A

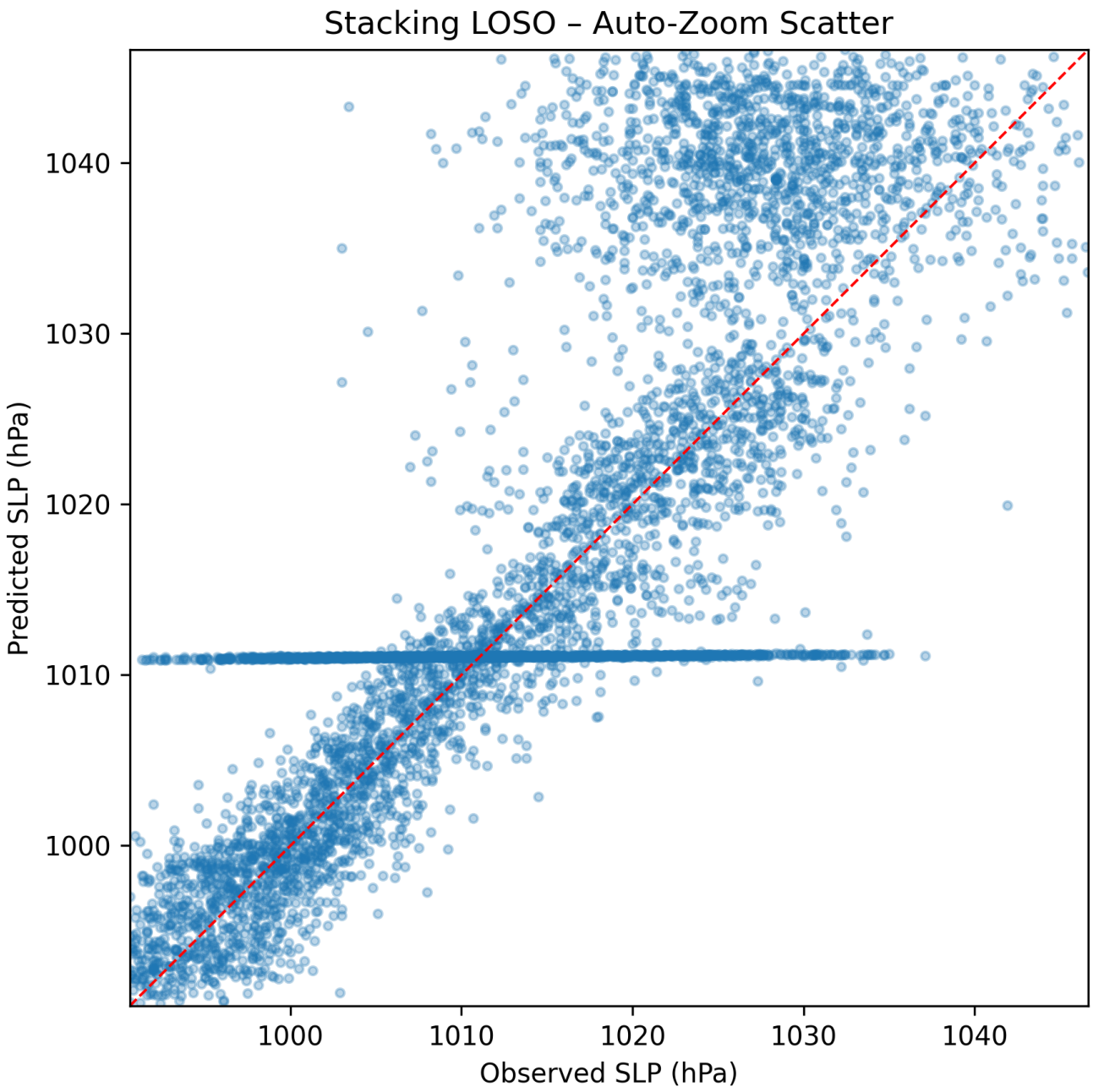

Fig. 6. Stacking LOSO auto zoom.

case study in wrocław," *Journal of environmental management*, vol. 217, pp164–174, 2018.

[17] M.-S. Park, M. Kim, M.-I. Lee, J. Im, and S. Park, "Detection of tropical cyclone genesis via quantitative satellite ocean surface wind pattern and intensity analyses using decision trees," *Remote sensing of environment*, vol. 183, pp205–214, 2016.

[18] D. Matsuoka, M. Nakano, D. Sugiyama, and S. Uchida, "Detecting precursors of tropical cyclone using deep neural networks," in *The 7th International Workshop on Climate Informatics, CI*, 2017.

[19] L. Wang, X. Wu, T. Zhao, G. Cheng, X. Zhang, L. Tang, M. Jia, and Y. Chen, "A scheme for rolling statistical forecasting of pm2. 5 concentrations based on distance correlation coefficient and support vector regression," *Acta Sci. Circumst*, vol. 37, no. 04, pp1268–1276, 2017.

[20] I. Juhos, L. Makra, and B. Tóth, "Forecasting of traffic origin no and no2 concentrations by support vector machines and neural networks using principal component analysis," *Simulation Modelling Practice and Theory*, vol. 16, no. 9, pp1488–1502, 2008.

[21] W. Zhanshan, L. I. Yunting, C. Tian, Z. Dawei, S. Feng, and P. Libo, "Spatial-temporal characteristics of pm2.5 in beijing in 2013," *Acta Geographica Sinica*, vol. 35, no. 5, pp1529–1536, 2015.

[22] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "Catboost: unbiased boosting with categorical features," *Advances in neural information processing systems*, vol. 31, 2018.

[23] T. Lin, "Focal loss for dense object detection," *arXiv preprint arXiv:1708.02002*, 2017.

Kaifang Mu(M'2025) is currently a postgraduate student at the School of Geographical Sciences, Qinghai Normal University. He received a Bachelor of Engineering degree from Heilongjiang University of Science and Technology and a Bachelor of Arts degree from Nankai University. He became a Member (M) of IAENG in 2025. He has worked as an assistant engineer in The Second Construction Co. Ltd. of CSCEC 7th Division, one of the world's top 500 enterprises. He has participated in national projects in civil engineering, geology and geographical sciences. Now he is a cooperative scholar of Urban Computing Laboratory, Aisess (Dalian) Computer Services Co., Ltd.

His research interests include: soil erosion, climate and environment.

Qing Ai(M'2024) is currently a postgraduate student at the College of Computer, Qinghai Normal University. He received his bachelor's degree from Software College, Liaoning Technical University. He became a Member (M) of IAENG in 2024. He has been worked at the Institute of Information Engineering, Chinese Academy of Sciences, and now he is a cooperative scholar of Urban Computing Laboratory, Aisess (Dalian) Computer Services Co., Ltd.

His research interests include: deep learning, combinatorial networks, urban computing and intelligent transport.

Xize Lu is currently a postgraduate student at the School of Environmental Science and Engineering, Shandong University. He received his bachelor's degree from Northeastern University. His work as the first author was published in *Advanced Energy Materials* (IF = 24.4), a top journal in the field of energy materials.

His research interests include: water pollution control chemistry and energy environmental catalysis.

Rui Zhang is currently the general manager of Aisess (Dalian) Computer Services Co., Ltd. He received his bachelor's degree from Software College, Liaoning Technical University. He has worked as a full-stack engineer for many years, accumulating extensive software development experience, particularly excelling in the field of mobile development. His mastery of various technology stacks enables him to independently manage the entire project lifecycle, from requirements analysis and design to development, testing, and deployment. His in-depth research in mobile application development not only enhances user experience but also drives successful project delivery. His contributions are reflected not only in technical expertise but also in his attention to detail and commitment to teamwork.

Now his research interests include: deep learning, combinatorial networks, urban computing and intelligent transport.