# SELECTION OF BEST METHOD AMONG DIFFERENT METHODS FOR DETERMINING MISSING RAINFALL DATA: CRITICAL REVIEW

[1]Monalika Malaviya, [2]Dr.Vilin Parekh
[1]Student, [2]Principal
[1]Civil Engineering Department,
[1]Parul Institute of Technology, Vadodara, India

***Abstract :*** Here, an attempt has been made to critically review research papers from 2006 onwards, to determine missing rainfall data. These papers discuss selected zones, numbers of rain gauge stations, duration and methods used for estimation of missing precipitation records as well as algorithms used to select the best one. Around 10 algorithms and 20 methods were discussed in selected research papers. Some methods need only one input variable to generate one output like Artificial Neural Network Method (ANN) & Closest Station Method (CSM) and some methods must need more than one input variables to generate output like Arithmetic Average Method, Inverse Distance Method & Normal Ratio Method. Here, coefficient of determination ($R^2$) and Willmott Agreement Index (d) algorithm are the best from all algorithms, which were used in these research papers. But for more accurate result, Authors should have used A Modified Index of Agreement ($d_1$) and A Refined Index of Agreement ($d_1$'). Critically considering all the merits and demerits of the algorithms and methods discussed, the best algorithm and method to get the best optimum result were determined.

***IndexTerms*** **- Rain gauge stations, Rainfall Data, Missing Data, Algorithms, Methods.**

## I. INTRODUCTION

Missing rainfall data may vary in length from one or two days or months to several years. It is necessary to estimate the missing data in order to utilize partial records, especially in data-sparse areas. For filling in missing records, the commonly available methods are arithmetic average method, inverse square distance (ISD) or national weather service method, normal ratio method, linear or multiple regression methods, the kriging method, etc.

The scientist/hydrologists come across the problem of missing rainfall data due to a variety of reasons. Measurements of hydrologic variables like rainfall, stream flows, etc. are prone to various instrumental/systematic, manual and random errors. Instrumental errors in rain gage measurements can be of various types: (1) raindrop splash from outside the rain gauge (2) water loss during measurement (3) adhesion loss on the surface of the gauge (4) raindrop splash from the collector. Sometimes, complete data may be lost if rain gauge malfunctions for a specific time period. Errors in recording of rainfall data were possible due to tree growth, instrumentation problems or techniques used in measuring the rainfall amounts. These errors were critical as they affect the continuity of rainfall data and ultimately influence the results of hydrologic models that use rainfall as input. Therefore, evaluation of missing data was important tasks for formulation of hydrological models.
To tackle missing rainfall data is a part of nearly all research. There are a number of different models available for dealing with missing rainfall data.

In the tentative specification phase namely model identification, the goal is to employ computationally simple methods to narrow down the range of parsimonious models. The Box Jenkins method is the only one, which is suitable for stationary time series data. Here one has to observe the time series graph for any anomaly and transform the data appropriately (Cryer and Chan, 2008). If the variance grows with time, it must be necessary to stabilize the variance. Then check for additive and lasting level shifts unaccounted for by the model using the outlier statement. Determine the regression variables after removing the outliers. Then augment the original dataset with these regression variables. The data are identified by preliminary values of autoregressive order p, the order of differencing d, the moving average order q and their corresponding seasonal parameters P, D and Q. Stoffer and Dhumway (2010) examined the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF). The Autocorrelation Function (ACF) methods amount of linear dependence between explanations in a time series that are separated by a lag q. The order of difference incidence from non-stationary time series to stationary series is called parameter d. Furthermore, a time series plot and ACF of data will typically suggest whether any differencing is needed and if differencing is called for, the time plots will illustration some kind of linear trend.

## II. RESEARCH PAPERS

This section comprises review of research papers, different methods used in them for computing missing precipitation records and their limitations.

**A. Reviews of Research Papers:**

The objectives of the study by Arumugam and Saranya (2018) were to examine records of rainfall data from different rain gauge stations and determine the missing rainfall data. In this study, monthly rainfall data from rain gauge stations of 11 years considered. The model was denoted as Seasonal ARIMA i.e. SARIMA (1, 1, 1) (0, 1, 1). A serious problem in analyzing rainfall data is what to do when missing or extreme values occur perhaps as a result of a breakdown in automatic counting equipment. Outlier detection and missing rainfall value estimation can be determined by using SARIMA procedure. The model fitted the data well and the stochastic seasonal variation was successfully modeled. It can be used for forecasting and modeling the time series of monthly rainfall data. The model can be improved by replacing missing rainfall values by mean values and identifying Seasonal Additive Outlier (SAO) and Innovative Outlier (IO) outliers. The fitted SARIMA (1, 1, 1) (0, 1, 1) gives improved forecasting results than observed data. [1]

Records of rainfall data from different rain gauge stations were examined by Miro et al. (2017) to determine the missing rainfall data. The study was carried out with the aim of performing a comparative validation of ten methods capable to infill multiple series at the same time in dense station networks for a climate with high rainfall irregularity. In this study, daily rainfall data of 60 years were considered. For finding missing rainfall data, Multiple Imputation Methods like 6 Linear, 2 Non-linear and 2 Hybrid Methods were used. These methods were chosen for two considerations: (1) The availability of codes in the MATLAB language, so that these can be applied under identical validation criteria in the same programming environment and (2) The applicability to dense network, allowing a multiple imputation in a large number of stations at a time. The imputation methods tested are: k-Nearest Neighbor (KNN), Iterated Local Least Squares (ILLS), Regularized Expectation Maximization (RegEM), Principal Component Analysis Approaches (PPCA, VBPCA), Data Interpolating Empirical Orthogonal Functions (DINEOF), Self-Organizing Map (SOM), Non-Linear Principal Component Analysis (NLPCA) and Hybrid Approaches (SOM+EOF, NLPCA+EOF). The evaluation and selection of the best algorithm is carried out by ranking the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Bias Percentage (BIASP), Root Mean Square Percentage Error (RMSPE), Mean Absolute Percentage Error (MAPE) and Coefficient of Determination ($R^2$). The best results were achieved by applying a pre-processing and post-processing of the data. The pre-processing was consisted in a division of the data into three groups according to a classification of three synoptic situations that generalize the type of precipitation according to its basic origin in the region and the post-processing has consisted in a bias correction using the Quantile Matching (QM) technique, so that the estimated data preserves the statistical properties of the observed data and adjusts well the trigger level between wet and dry events. [6]

Saeed et al. (2016) introduced several single imputation algorithms, believed to be more competent in treating missing daily rainfall data without depending on the rainfall records of adjoining rain gauge stations. The study developed an efficient single imputation algorithm in treating missing rainfall data without depending on the rainfall records and homogeneity of adjoining stations. In the study, hourly rainfall data from 6 rain gauge stations of 1 year duration were considered. For finding missing rainfall data, Column Arithmetic Means, Column Geometric Means, Column Harmonic Means, Column Medians, Column Midranges and Row Arithmetic Means methods were used. To evaluate imputation algorithms and to select the best algorithm, they were ranked by the average of Bray-Curtis dissimilarity, Mean Square Error (MSE) and Normalized Root Mean Square Error (RMSE). Based on the analysis, it was proved that the proposed algorithms replacing, missing rainfall data with Column Geometric Means, Column Harmonic Means and Column Medians were more superior than the others. [14]

Sungmin e**t** al. (2016) evaluated WegenerNet daily rainfall data of 8 years through careful comparison with data from 150 Austrian national weather stations. The Bias, relative Bias (rBias), Mean Absolute Error (MAE), relative Mean Absolute Error (rMAE) and Root Mean Square Error (RMSE) were used to evaluate imputation algorithms and to select the best algorithm. Since rain gauge measurements were considered close to true rainfall, there are increasing needs for WegenerNet data for the validation of rainfall data products such as remote sensing based estimations or model outputs. The analysis of the study improved WegenerNet data for user applications and a new version of the data, v5.0, is now available at the WegenerNet data portal (www.wegenernet.org). [15]

To examine records of rainfall data from different rain gauge stations and determine the missing rainfall data Ghuge and Regulwar (2013) used Artificial Neural Network (ANN) Method. In the study, monthly rainfall data of 10 years from 6 rain gauge stations were considered. The performances of the method were evaluated Root Mean Square Error (RMSE), Mean Relative Error (MRE), Mean Absolute Error (MAE) and Coefficient of Determination ($R^2$). Values of RMSE were found towards higher side in ANN. The variation in rainfall was increased year by year as increase in the error. The performance of proposed ANN model was studied and observed that the values predicted by the program were reliable to use. [4]

Nkuna and Odiyo (2011) examined records of 1 year rainfall data from 5 rain gauge stations and determined the missing rainfall data. For finding missing rainfall data, Artificial Neural Network (ANN) method was used. To evaluate imputation algorithms and to select the best algorithm the Root Mean Square Error (RMSE) and Coefficient of Determination ($R^2$) were used. The Shuffled Complex Evolution (SCE) was used to find optimal parameters of the ANNs. Evolutionary process of SCE tried to balance between a wide-scan of a large solution space and deep search of promising locations. It depends mainly on partitioning the solution space into local communities and perform local search within these communities. Then, it shuffled these local communities to perform global search. The major drawback was the fact that these approaches are data intensive and work as black box, thus no process insight was provided. This drawback was rectified by Modified Shuffled Frog Leaping Algorithm (MSFLA). The study showed that ANNs were suitable for estimating missing rainfall data and produced reliable rainfall data. [10]

Mair and Fares (2010) examined records from long-term rain gauges in Makaha Valley, Hawaii, for data homogeneity and to compare four different methods of estimating missing rainfall data such as direct substitution, inverse distance, multilinear regression and normal ratio Method. [8]

Sorjamaa et al. (2010) presented an improved methodology for the determination of missing rainfall values in a spatiotemporal database. The methodology performed denoising projection in order to accurately fill the missing rainfall values in the database. The improved methodology was empirical orthogonal functions (EOF) pruning and it was based on an original linear projection method called empirical orthogonal functions (EOF). [13]

Villazón and Willems (2010) applied Linear regression and multiple linear regression techniques for the estimation of monthly precipitation data. [16]

Ng et al. (2009) evaluated the performance of different estimation techniques for the filling of missing rainfall observations in extreme daily hydrologic series. Generalized regression neural networks were proposed for the estimation of missing rainfall observations with their input configuration determined through an optimization approach of genetic algorithm. [9]

Partal and Cigizoglu (2009) aimed to predict the daily precipitation data from meteorological data from Turkey using the wavelet—neural network method, which combines two methods: (1) Discrete Wavelet Transforms (DWT) and (2) Artificial Neural Networks (ANN). [11]

Patel et al. (2008) concluded the effectiveness of the artificial neural network method for Mehsana district, Gujarat, India, compared to the arithmetic average method, inverse square distance (ISD) (National Weather Service method), normal ratio method, linear and multiple regression methods. [12]

Suhaila et al. (2008) examined 30 years rainfall data from 20 rain gauge stations and determine the missing rainfall data. For finding missing rainfall data, Inverse Distance Weighting Method, Normal Ratio Method, Coefficient of Correlation Weighting Method, Modified Coefficient of Correlation, Modified Coefficient of Correlation with Inverse Distance Weighting Method, Modified Normal Ratio with Inverse Distance Method and Modified Old Normal Ratio with Inverse Distance Methods were used. The existing methods which include the inverse distance, normal ratio and coefficient of correlation weighting methods had been explored and revised. Some modifications or revisions had been made to the existing methods and these new modified methods had been tested for estimation of missing rainfall values. Evaluation and selection of the best algorithm ranked by Mean Absolute Error (MAE) and Coefficient of Determination ($R^2$). The result indicated that the performance of these modified methods improved the estimation of missing rainfall records at the target station based on the Similarity Index (S-index), Mean Absolute Error (MAE) and Coefficient of Determination ($R^2$).

De Silva et al. (2007) developed and introduced a new method (Aerial Precipitation Ratio Method) for missing rainfall data estimation. In the study, monthly rainfall data of 20-30 years from 7 principle stations with 3-4 surrounding stations were considered and different frequency was chosen for different stations. The performance of the method was commonly used error measures Mean, Standard Deviation, Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Correlation Coefficient ($R^2$). Results showed that Inverse Distance Method was the most suitable for Long country stations, Normal Ratio Method was the most suitable method for Mid country and Up country intermediate zone stations, Arithmetic Mean Method was more suitable for Up country wet zone and Aerial Precipitation Ratio Method was more suitable for Mid country wet zone. [3]

Lucio et al. (2007) examined 45 years records of monthly rainfall data from 15 rain gauge stations and determine the missing rainfall data by using Artificial Neural Network (ANN) Method. Here different neural network architectures and learning rates were tested, aimed at establishing a network that resulted in the best possible reconstruction of missing rainfall data. The performance of the method was tested by commonly used error measure Coefficient of Determination ($R^2$). [7]

Coulibaly and Evora (2007) investigated six different types of artificial neural networks namely the multilayer perceptron (MLP) network and its variations (the time-lagged feedforward network (TLFN)), the generalized radial basis function (RBF) network, the recurrent neural network (RNN) & its variations (the time delay recurrent neural network (TDRNN)), and the counter propagation fuzzy-neural network (CFNN) along with different optimization methods for infilling missing daily rainfall records.

Considering the closest station method and characteristics of cluster analysis Wu et al. (2007), aimed to calibrate existing models and developed new models for estimating missing global solar radiation data using commonly measured meteorological data and proposed a strategy for selecting the optimal models under different situations of available meteorological data. [17]

Garcia et al. (2006) examined 31 years records of rainfall data from 106 different rain gauge stations and determined the missing rainfall data. For finding missing rainfall data, Closest Station Method was used by Cluster Analysis. To evaluate imputation algorithms and to select the best algorithm the Root Mean Square Error (RMSE), Willmott Agreement Index (d), Mean Error (ME), Mean Absolute Error (MAE) and Coefficient of Correlation (r) were used. Daily, weekly, Bi-weekly and monthly observed rain and no rain values were estimated. Values of r for daily, weekly, bi-weekly and monthly time scale were 0.345, 0.531, 0.556 and 0.695 respectively. Values of d for daily, weekly, bi-weekly and monthly time scale were 0.565, 0.723, 0.741 and 0.830 respectively. It shows clearly that d is the better parameter to evaluate the performance of the model. Here it shows for the Closest Station Method better results are obtained when the monthly time series is used, as when d approaches unity then there is complete agreement with the observed and predicted values. [2]

**B. Limitations/Demerits:**

Limitations of methods used by various researchers for finding missing rainfall precipitation records are discussed here:

- **Artificial Neural Network Method (ANN):**
  ANN develops a probing solution but it does not give a clue as to why and how it happens.
  There is no specific rule for determining the Artificial Neural Network Structure.
  Before being introduced to ANN, problems have to be translated into numerical values.
  Values do not give the optimum results.
- **Cluster Analysis:**
  There are different methods of clustering. Every method usually gives very different results.
  The results will be affected by the way in which the variables are ordered with the exception of simple linkage.
  The analysis is not stable when cases are dropped because selection of a case depends on similarity of one case to the cluster.
- **Non Renewable Empower time series:**
  Once sources of non renewable energies are gone, they cannot be replaced.
- **Seasonal ARIMA Model:**
  It is appropriate only for a time series that is stationary.
- **Multiple Imputation Methods:**
  To generate multiple imputations, more work is needed.

Need more space to store the data.
Required more work to analyze the data.

- **Arithmetic Mean Method:**
  Highly affected by extreme values.
  It cannot average the percentages and ratios properly and also has highly skewed distributions.
  If any data is missing, it cannot be computed accurately.
  Sometimes the mean does not coincide with any of the observed value.
- **Regression Methods:**
  The relationship between the variables remains unchanged.
  If more data are taken into consideration, the functional relationship that is established between any two or more variables on the basis of some limited data may not hold good.
  It involves very complicated and lengthy procedures of calculations and analysis.
  In case of qualitative phenomenon, it cannot be used.
- **Simple and Modified correlation Methods:**
  The correlation analysis can be used when the variables are two measurable on a scale.
  Cause and effect cannot be established in correlation analysis as it is not certain that one variable caused another to happen, it could be one or the other or it could even be an unknown variable that causes the correlation.
- **Spatiotemporal database:**
  Keep information about some part of a real or artificial domain. These databases allow capturing some essence of time, which generally consist in a snapshot view of the world limited to the last update regarding the temporal aspects of facts that occur in the real word. It becomes critical when there is the need to capture the evolution of facts over the time.
- **Time series:**
  One rarely knows the true shape of the distribution with which she can work.
  The observations within each series are not independent of each other.
  It is rarely reasonable to assume that the time sequence of the causal patterns matches the time periods.
  Analyzing causal patterns is the familiar problem that correlation does not imply causation.
- **Fuzzy Logic:**
  Fuzzy systems lack the capability of neural network type pattern as well as machine learning recognition.
  Validation and verification of a fuzzy knowledge based system require extensive testing with hardware.
  Evaluation of exact fuzzy rules and membership functions is a hard task.
  Important concern for fuzzy control is stability.

Limitations of algorithms, which are used by researchers to select the best one among the selected algorithms:

- **Mean:**
  It is highly affected by the presence of a few abnormally low scores.
  In absence of a single item, it's value become inaccurate.
  It cannot be determined by inspection.
- **Standard deviation:**
  It doesn't give you the full range of the data.
  It can be hard to calculate.
  It is only used with data where an independent variable is plotted against the frequency of it.
  It assumes a normal distribution pattern.
- **Variance:**
  There may be a time gap, which may affect the remedial action taking ability to a certain extent. All sources of variance may not be available in accounting data, which makes acting upon variances difficult.
  It gives added weight to numbers far from the mean (outliers), since squaring these numbers can skew interpretations of the data.
  It is not easily interpreted, and the square root of its value is usually taken to get the standard deviation of the data set in question.
- **Root Mean Square Error (RMSE):**
  It does not increase with the variance of the errors.
- **Mean Relative Error (MRE) and Mean Absolute Error (MAE):**
  The relative size of the error is not always obvious.

- **Mean Square Error (MSE):**
  It has heavily weighting outliers and the result of the squaring of each term, which effectively weights large errors more heavily than small ones.
- **Coefficient of Determination ($R^2$):**
  Values of any axis are multiplied by a constant but $R^2$ stay the same.
  Sometimes it is hard to tell a big error from a small error.
- **Coefficient of correlation (r):**
  It only measures linear relationships between two axis.
- **Willmott Agreement Index (d):**
  A Refined Index of Agreement is modified of Willmott Agreement Index.
- **A Modified Index of Agreement ($d_1$):**
  It is overly sensitive to extreme values due to the squared differences.
- **A Refined Index of Agreement ($d_1$'):**
  It hasn't more predictive ability than the observed mean.

**C. Merits:**

Merits of the algorithms, which are used for evaluate the optimum results.

- **Mean:**
  It is simple to understand and easy to calculate.
  It is rigidly defined.
  It is least affected function of sampling.
  It is suitable for further algebraic treatment.
  It takes into account all the values in the series.
- **Standard deviation:**
  It shows how much data is clustered around a mean value.
  It gives a more accurate idea of how the data is distributed.
  Not as affected by extreme values.
- **Variance:**
  It treats all deviations from the mean the same regardless of direction; as a result, the squared deviations cannot sum to zero and give the appearance of no variability at all in the data.
- **Root Mean Square Error (RMSE):**
  It is more useful when large errors are particularly undesirable.
- **Mean Absolute Error (MAE):**
  It requires linear programming to compute the gradient.
- **Mean Square Error (MSE):**
  Almost always strictly positive and not zero is because of randomness or the estimator does not account for information that could produce a more accurate estimate.
- **Coefficient of Determination ($R^2$):**
  It provides a measure of the strength of the correlation.
- **Coefficient of correlation (r):**
  It provides the positive or negative direction of the correlation.
- **Willmott Agreement Index (d):**
  The relatively high values may be obtained even for a poor model fit.
- **A Modified Index of Agreement ($d_1$):**
  The values obtained by the model are more accurate than Willmott Agreement Index.
- **A Refined Index of Agreement ($d_1$'):**
  It undermines interpretations of index values associated with poorly performing models.

Considering the errors magnitude adopted in the Monte Carlo Simulations as well as the case of study, the findings indicate that the original version of the Willmott index may lead the user to erroneously select a predicting model that generates poor estimates. This statement is consistent with previous studies. The results also indicate that the two newer versions of this index (modified and refined) overcome such problem, leading to more rigorous evaluations of the predicting models. Therefore, they should be preferred over the original version. [5]

Around 10 algorithms and 20 methods were discussed in selected research papers. Some methods need only one input variable to generate one output like Artificial Neural Network Method (ANN) & Closest Station Method (CSM) and some methods must need more than one input variables to generate output like Arithmetic Average Method, Inverse Distance Method & Normal Ratio Method. From the study, it is concluded that ANN and CSM are the best method to get optimum result and Willmott Agreement Index (d) algorithm is the best from all algorithms, which were used in these research papers.

As discussed above, a refined index of agreement ($d_1$') can be used for evaluating the algorithms for the better results.

## III. CONCLUSION

From the review of different research papers, it is concluded that Artificial Neural Network Method (ANN) and Closest Station Method are the best methods for determining the missing rainfall data in which only one input variable is considered to generate output instead of using other methods, for which more than one input variable is needed to generate the output. To evaluate the different algorithms Willmott Agreement Index (d) is the best tool. But for more accurate result, Refined Index of Agreement ($d_1$') can be used.

## REFERENCES

[1] Arumugam. P, and Saranya. R, 2018, "Outlier Detection and Missing Value in Seasonal ARIMA Model Using Rainfall Data," *Materials Today: Proceedings* 5, 1791–1799.

[2] B. Rosmina, 2007, "Artificial neural network for precipitation and water level predictions of bedup river," *IAENG, International Journal of Computer Science*, 34(2).

[3] De Silva R. P., 2007, "A Comparison of Methods used in Estimating Missing Rainfall Data," *The Journal of Agricultural Sciences*, vol.3, no.2, 101-108.

[4] Ghuge H. K., and Regulwar D.G., "Artificial Neural Network Method for Estimation of Missing Data," *International Journal of Advanced Technology in Civil Engineering*, ISSN: 2231 –5721, Volume-2, Issue-1.

[5] H.R.Pereira et al., 2018, "On the performance of three indices of agreement: an easy-to-use r-code for calculating the Willmott indices," Bragantia vol.77 no.2.

[6] J. J. Miró, 2017, "Multiple Imputation of Rainfall Missing Data in the Iberian Mediterranean Context," *Atmospheric Research*, S0169-8095, 30125-4.
[7] Lucio P. S., 2007, "Spatiotemporal monthly rainfall reconstruction via artificial neural network – case study: south of Brazil," *Adv. Geosci*., 10, 67–76.
[8] M. Alan and F. Ali, 2010, "Assessing rainfall data homogeneity and estimating missing records in Mâkaha Valley, O'ahu, Hawai'I," *J. of Hydrol. Engg*. 15(1).
[9] Ng W. W., 2009, "Comparative studies in problems of missing extreme daily streamflow records," *J. of Hydrol. Engg*. 14(1).
[10] Nkuna T.R., and Odiyo J.O., 2011, "Filling of missing rainfall data in Luvuvhu River Catchment using artificial neural networks," *Physics and Chemistry of the Earth*, 36, 830-835.
[11] P. Turgay and C. H.Kerem, 2009, "Prediction of daily precipitation using wavelet—neural networks," *Hydrol. Sci. J*. 54(2), 234 – 246.
[12] Patel N.R, Suryanarayana T.M.V. and Shete D.T., 2008, "Comparison of ANN and conventional methods for predicting missing climate data," Proc. *of International conference on "Operations Research for a Growing Nation in conjunction with 41st Annual convention of Operation Research Society of India*, Tirupati.
[13] Sorjamaa A., 2010, "An improved methodology for filling missing values in spatiotemporal climate data set," *J. of Computational Geosciences* 14(1).
[14] Saeed G. A. A., 2016, "Determination of the Best Single Imputation Algorithm for Missing Rainfall Data Treatment," *Journal of Quality Measurement and Analysis* 12(1-2), 79-87.
[15] Sungmin O., 2016, "Validation and correction of rainfall data from the WegenerNet high density network in southeast Austria," *Journal of Hydrology*, S0022-1694, 30764-8.
[16] Villazon M. F. and Willems P., 2010, "Filling gaps and daily disaccumulation of precipitation data for rainfall-runoff model," *BALWOIS 2010 - Ohrid*, Republic of Macedonia – 25.
[17] Wu Guofeng, 2007, "Methods and strategy for modeling daily global solar radiation with measured meteorological data – A case study in Nanchang station, China," *Energy Conversion and Management* 48, 2447–2452.