Colored Noise Minimisation using Machine Learning

Under the guidance and supervision of
Prof. Shekha Rai

Department Of Electrical Engineering
National Institute Of Technology
Rourkela - 769008

Akula Hemanth Kalyan Kumar	120EE0452
Kandimalla Vishwas Reddy	120EE0695
Annadanam Lakshmi Haritha	120EE1086

Contents

Introduction

- Power System stability is an important factor for the interconnection among large scale power grids [1].
- Low frequency oscillations, caused due to system interconnections or heavy load conditions affects the power system stability. [1].
- Wide Area Measurement System (WAMS) provides on-line estimation of the network by installation of Phasor Measurement Units (PMU)[2].
- The power signal from the PMU is retrieved as a minor component of additive colored noise along with additive white noise because the PMU employs an anti-aliasing filter prior to sampling[2].
- Mitigation of these noises using some identification algorithms is important[2],[3].

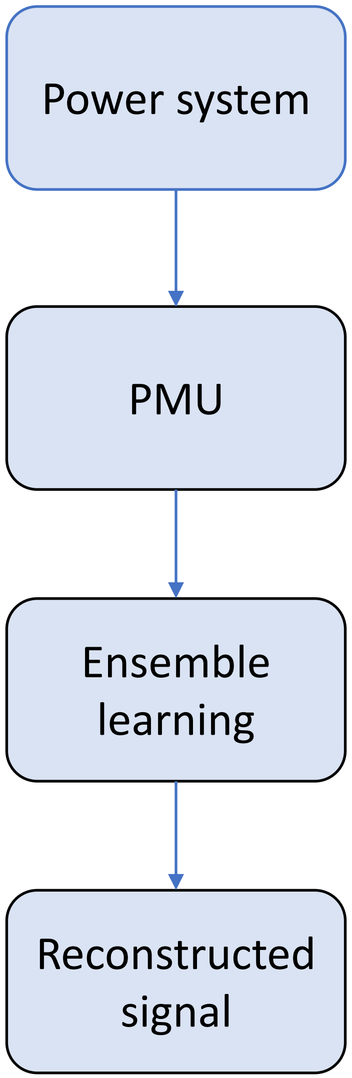

Figure 1: Block diagram of proposed method for coloured noise minimisation

Literature Review

Title of Paper	Name of the Journal/Publisher/Conference and Publication year	Review
A Modified TLS-ESPRIT Based Method for Low Frequency Mode Identification in Power Systems Utilizing Synchro phasor Measurement	IEEE Transactions on Power Systems, vol. 26, no. 2, pp. 719-727, May 2011	This report provides a review on how colored noise and white noise are being added to a clean signal and different algorithms to filter out high frequency components at the PMUs output.
Synchronized measurement based estimation of inter-area electromechanical modes using Ibrahim time domain method.	IEEE Transactions on Power Systems Res 118:85-95	This report provides how the Weiner filter plays a significant role in noise suppression and filtering of the clean signal.
An efficient K-SVD based Algorithm for detection of Oscillatory mode from ambient data for synchro phasor application	2021 IEEE 18th India Council International Conference (INDICON), 2021	Got a review on the effective application of K-SVD for the identifying and removing the oscillatory modes from ambient data and how ambient data is obtained from PMU.

Literature Review

Title of Paper	Year of Publication	Review
Using sparsity to estimate oscillatory mode from ambient data	Indian Academy of Sciences,2018	Got a review report on how it creates an effective dictionary from an input signal and then taking the representation coefficients that most accurately represent the noise-free signal from it.
Online Transfer Function Estimation and Control Design Using Ambient Synchro phasor Measurements	IEEE Transactions on Power Systems,2022	In this report, a methodology for identifying and designing adaptive controls for low-frequency oscillations in power systems is given and the design is based upon ambient synchro phasor measurements.
An XgBoost based method for identifying electromechanical oscillations from ambient measurements using WAMS	IEEE Conference, ICSTNSN, 2023	This report provides a review on how we are minimizing colored noise using XgBoost algorithm which enables more reliable and precise estimation.

Motivation

- On-site PMUs often face varying degrees of data quality difficulties that need to be handled. These issues might be caused as PMU uses anti-aliasing filter before sampling.

- Colored noise causes an estimation bias that greatly lowers the accuracy of the predicted data when utilizing the current techniques.

- Proposing few other methods with XgBoost algorithm to implement a better solution for colored noise reduction.

Objectives

- To pass the PMU data through Ensemble learning to remove outliers and recover the data loss due to faulty PMU, communication loss etc.

- To minimise the error of the reconstructed signal using Stacking and comparing reconstructed signal with clear signal.

Machines Learning Algorithms

Independent component analysis

- ICA assumes that the observed data is a linear combination of independent, non-Gaussian signals. The goal of ICA is to find a linear transformation of the data that results in a set of independent components.

- Independent component analysis (ICA) is a statistical technique to decompose multivariate data into statistically independent components. [5]

- Ability to separate mixed signals: ICA is a powerful tool for separating mixed signals into their independent components. This is useful in a variety of applications, such as signal processing, image analysis, and data compression.

- We can use the FastICA algorithm. FastICA is an iterative algorithm that uses a non-linear optimization technique to find the independent components.

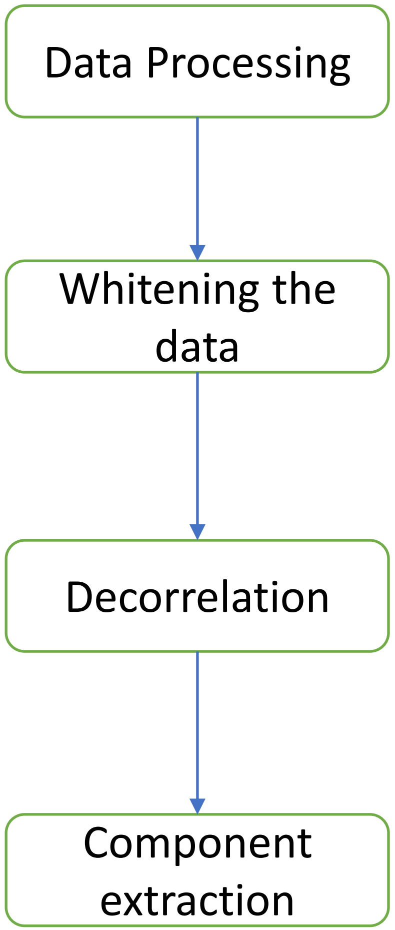

Figure 2 : Block diagram of working of ICA

Sparse Principal component analysis

- Sparse Component Analysis (SCA) is a dimensionality reduction technique aiming to extract meaningful components from high-dimensional data while promoting sparsity in the representation.

- $X \approx W \cdot Z$, where X is the original data matrix,
  W is the sparse loading matrix, and
  Z is the sparse representation matrix.

- Minimizing $\|X - W \cdot Z\|^2 + \lambda \|Z\|$ where λ controls sparsity through L1 regularization, encouraging many coefficients in W to be zero and $\lambda||Z||$ is L1 regularization term.

Non-Negative Matrix factorization

- Non-Negative matrix factorization is commonly used for dimensionality reduction and feature extraction. For a signal if applying NMF will separate a signal into additive components.[6]

- NMF decomposes a matrix into the product of two lower-rank matrices with non-negative elements. So a signal matrix containing noise can be represented as the product of two matrices representing the signal and the noise. [6]

$$X \approx W \bullet H$$

where X = signal matrix with noise

W = appropriate signal matrix

H = Noise matrix

- We need to choose our rank of the matrix such that there is no problem of overfitting and simplification of data representation.

Ensemble stacking

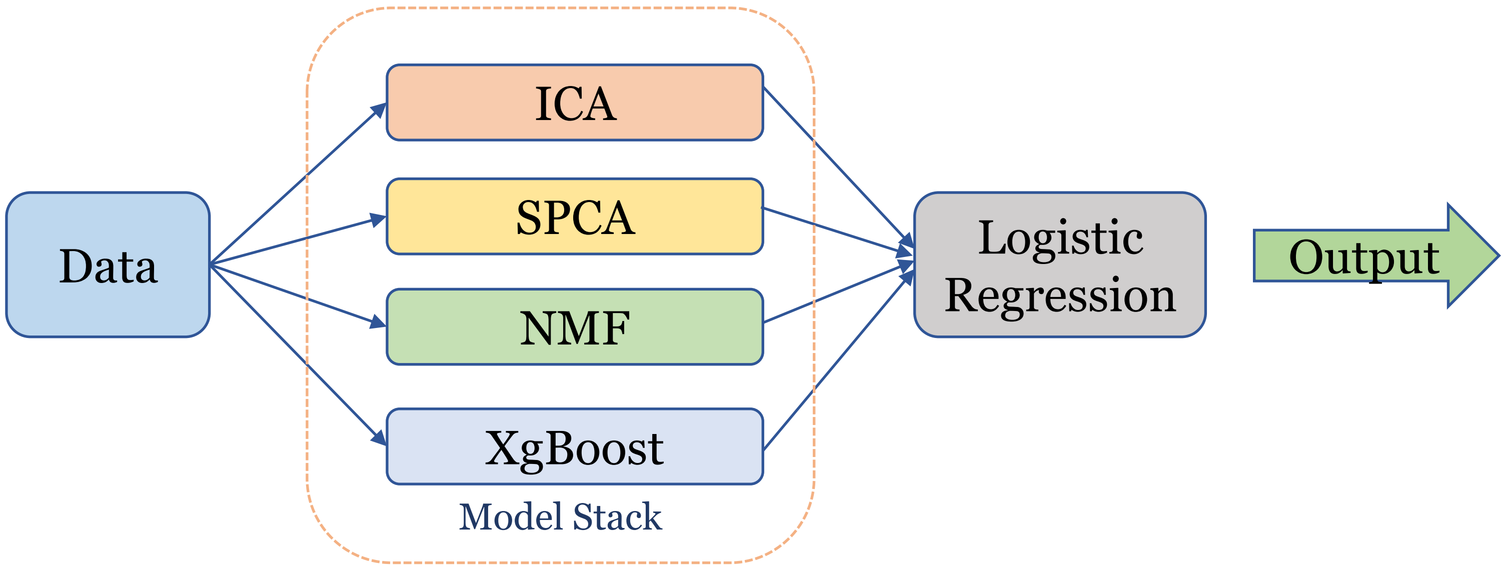

Figure 3 - : Stacked model of Machine learning algorithms

- XgBoost is an efficient open source implementation of the gradient boosting algorithm.
- In stacking an algorithm takes the outputs of sub-models as input and attempts to learn how to best combine the input predictions to make a better output prediction.
- XgBoost doesn’t always give appropriate output with all the input data types, hence stacking is performed with ICA, SPCA, NMF and XgBoost for a better inclusion.

Results and Discussion

➢ This test signal corresponding to Ensemble stacking was simulated with a sampling frequency of 10 Hz.

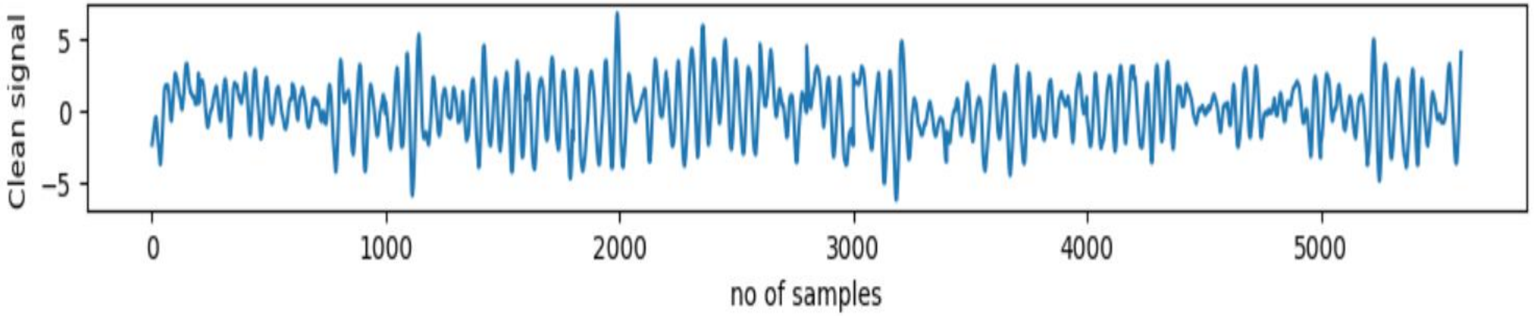

Figure 4 : Clean signal

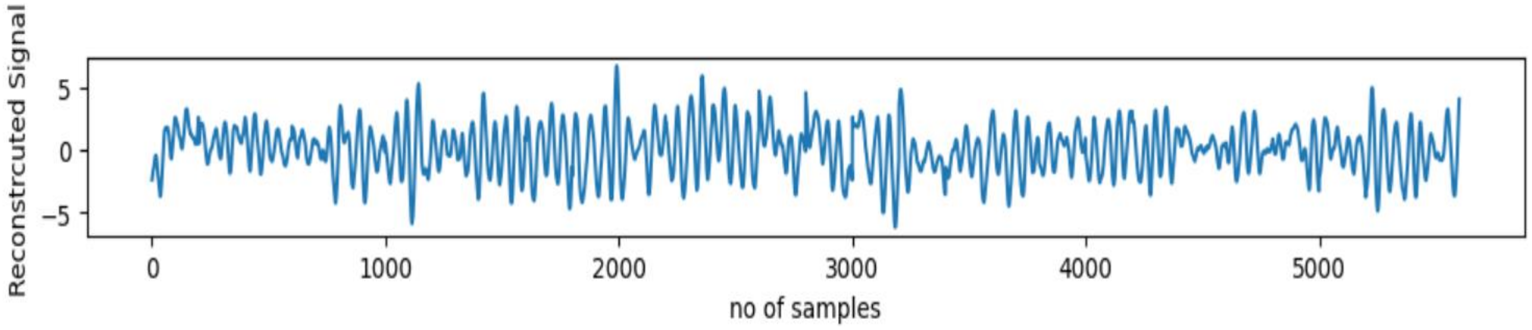

Figure 5 : Reconstructed signal with proposed method

Method	MSE
ICA	0.999
SPCA	0.2161
NMF	0.7115
XgBoost	0.8457
Stacking	0.0017

Table 1: Errors of proposed models

Conclusion

- In this paper we used different machine learning algorithms so that the proposed method is compatible with different input data types.
- The presence of colored noise is less of an issue with this method, which enables a more reliable and precise estimation of the modes from ambient data.
- The results obtained from different algorithms like ICA, SPCA, NMF and XgBoost are compared.
- For an optimized solution we used ensemble stacking which makes a better output prediction based on all the outputs of sub-models.
- Hence, it is reasonable to draw the conclusion that suggested ensemble stacking scheme which preserves the signal characteristics and gives precise results.

Future Work

- Testing of 2 area networks in practical scenarios.
- Testing of time signals and necessary modifications.
- Exploration of other machine learning algorithms for better predictions.

References

- [1] Kundur P 1994 Power system stability and control. New York, McGraw-Hill
- [2] P. Tripathy, S. C. Srivastava and S. N. Singh, "A Modified TLS-ESPRIT-Based Method for Low-Frequency Mode Identification in Power Systems Utilizing Synchro phasor Measurements," in *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 719-727, May 2011, doi: 10.1109/TPWRS.2010.2055901.
- [3] S. Rai, P. Tripathy, and S. Nayak, "A robust TLS-ESPIRIT method using covariance approach for identification of low-frequency oscillatory mode in power systems," in Power Systems Conference (NPSC), 2014 Eighteenth National, Dec 2014, pp. 1–6.
- [4] Zheng, Shiqiang & Shuangyi, Zhang & Song, Youyi & Lin, Zhizhe & Wang, Fei & Zhou, Teng. (2020). A Noise-eliminated Gradient Boosting Model for Short-term Traffic Flow Forecasting. 10.1109/ICDH51081.2020.00036.
- [5] S. Makeig, A.J. Bell, T.-P. Jung, and T.-J. Sejnowski. Independent component analysis of electro-encephalographic data. In Advances in Neural Information Processing Systems 8, pp. 145-151. MIT Press, 1996.
- [6] Yu-Xiong Wang ,Yu-Jin Zhang, "Nonnegative Matrix Factorization: A Comprehensive Review" in *IEEE Transactions on Knowledge and data Engineering (Volume : 25 , Issue : 6, June 2013).* doi :10.1109/TKDE.2012.51
- [7] D. L. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

THANK YOU