IEEE WCCI 2020
IEEE World Congress on Computational Intelligence
Virtual Conference – July 19-24, 2020

Solar Power Forecasting Based on Ensemble Learning Methods

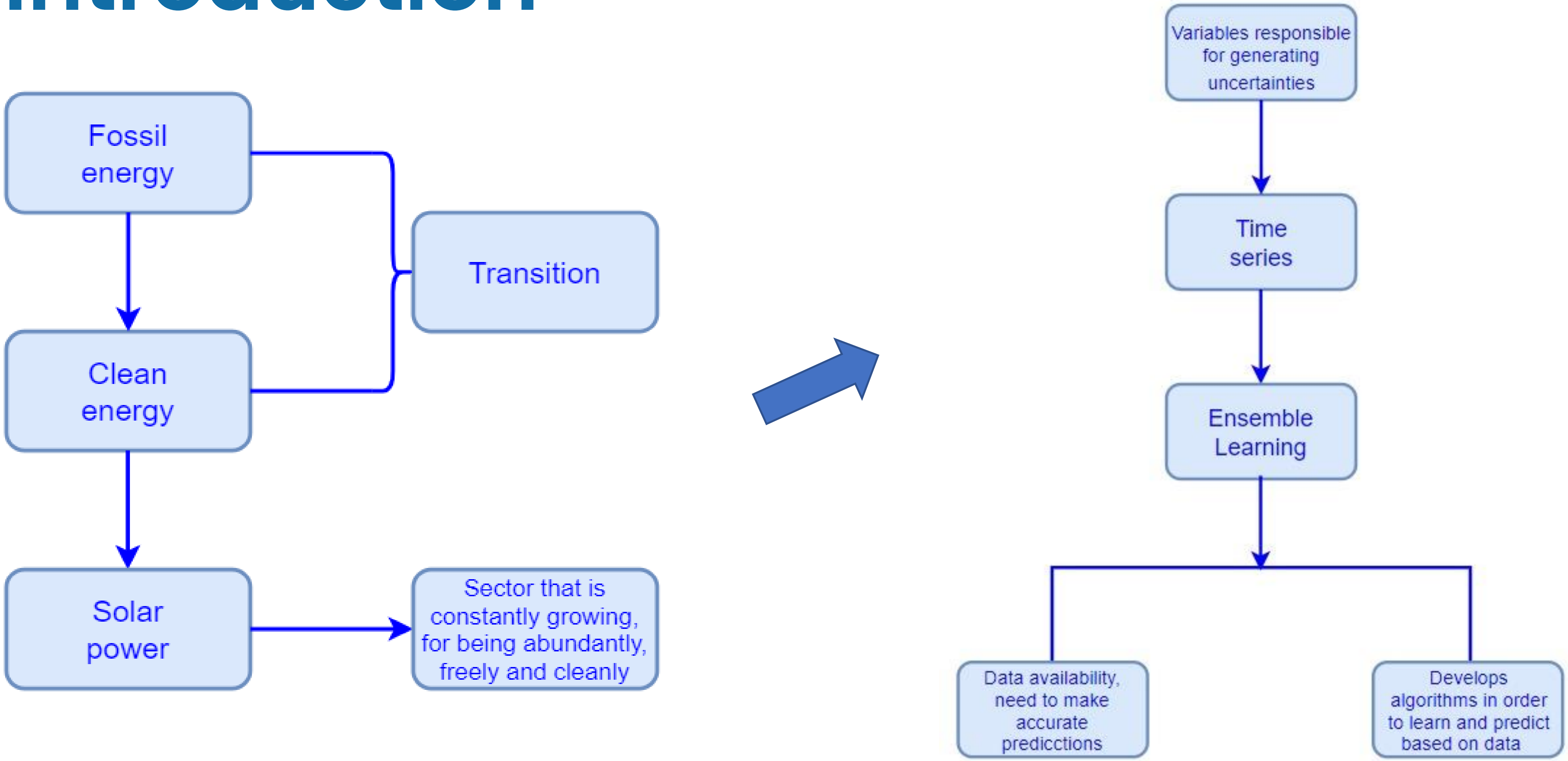
N. Fraccanabbia, M.H.D.M. Ribeiro, R.G. da Silva, S.R. Moreno, L.S.Coelho, V.C. Mariani

Mechanical Engineering Graduate Program (PPGEM). Pontifical Catholic University of Parana (PUCPR), Parana, Brazil

Agenda

- Introduction
- Objectives
- Material
- Methods
- Methodology
- Results
- Conclusion
- Acknowledgments

Introduction



Objectives

MAIN OBJECTIVE:

- Forecasting solar power generation.

SPECIFIC OBJECTIVES:

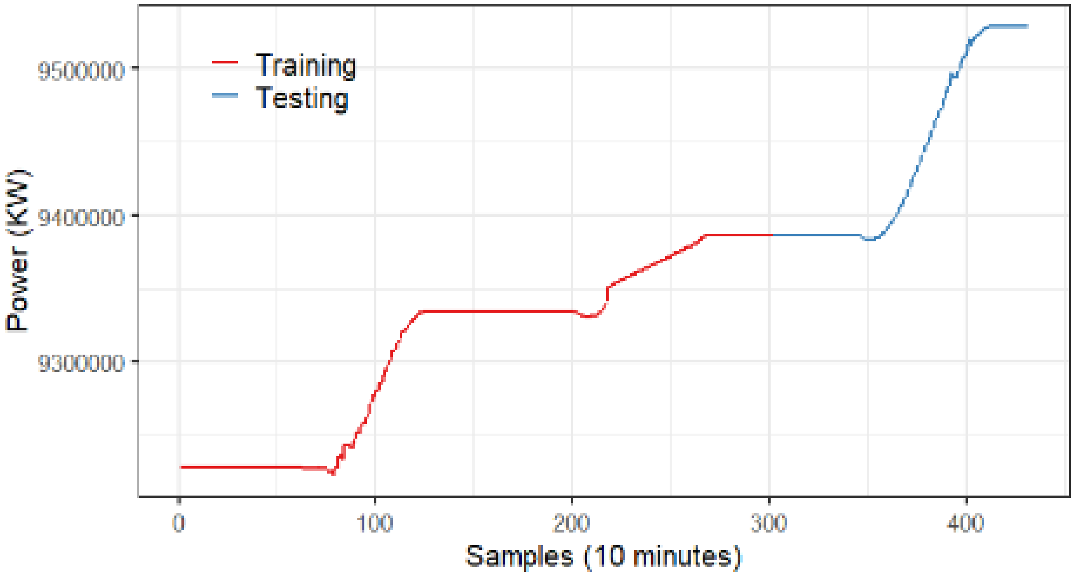
- Applying regression approaches;
- Developing a model to forecast solar power generation;
- Identifying suitable machine learning techniques;
- Improving the forecasting accuracy of solar power generation.

Material

- Photovoltaic plant of 26.35 MWp;
- Location: Artigas, Uruguay;
- No. of samples: 432;
- Period: 04/14/2018, 00:00h to 04/16/2018, 23h50;
- Collection: every 10 minutes;

Material

- The dataset was split into training (70%) and test (30%) sets:



- The system inputs and output (3 variables) are described in the following table:

INPUTS AND OUTPUT OF THE SYSTEM

Type	Description	Unit Measure
Input	Temperature	C°
Input	Radiation	W/m^2
Output	Power	kW

Methods

- Stacking-Ensemble Learnig (STACK);

Models used in the STACK method:

- CUBIST;
- Multivariate Adaptive Regression Splines (MARS);
- Linear Model (LM);
- Bayesian Regularized Neural Network (BRNN);
- Support Vector Regression with Linear, and Radial Basis Function kernel (SVRL and SVRR).

Methods

Pre-processing techniques:

- Principal Component Analysis (PCA);

(DU; ZHU, 2019)

- Correlation Matrix (CORR)

(KUHN; JOHNSON, 2013)

Methods

Performance Measures:

Relative Root Mean Square Error (RRMSE)

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

Determination Coefficient (R^2)

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

Symmetric Mean Absolute Percentage Error (sMAPE)

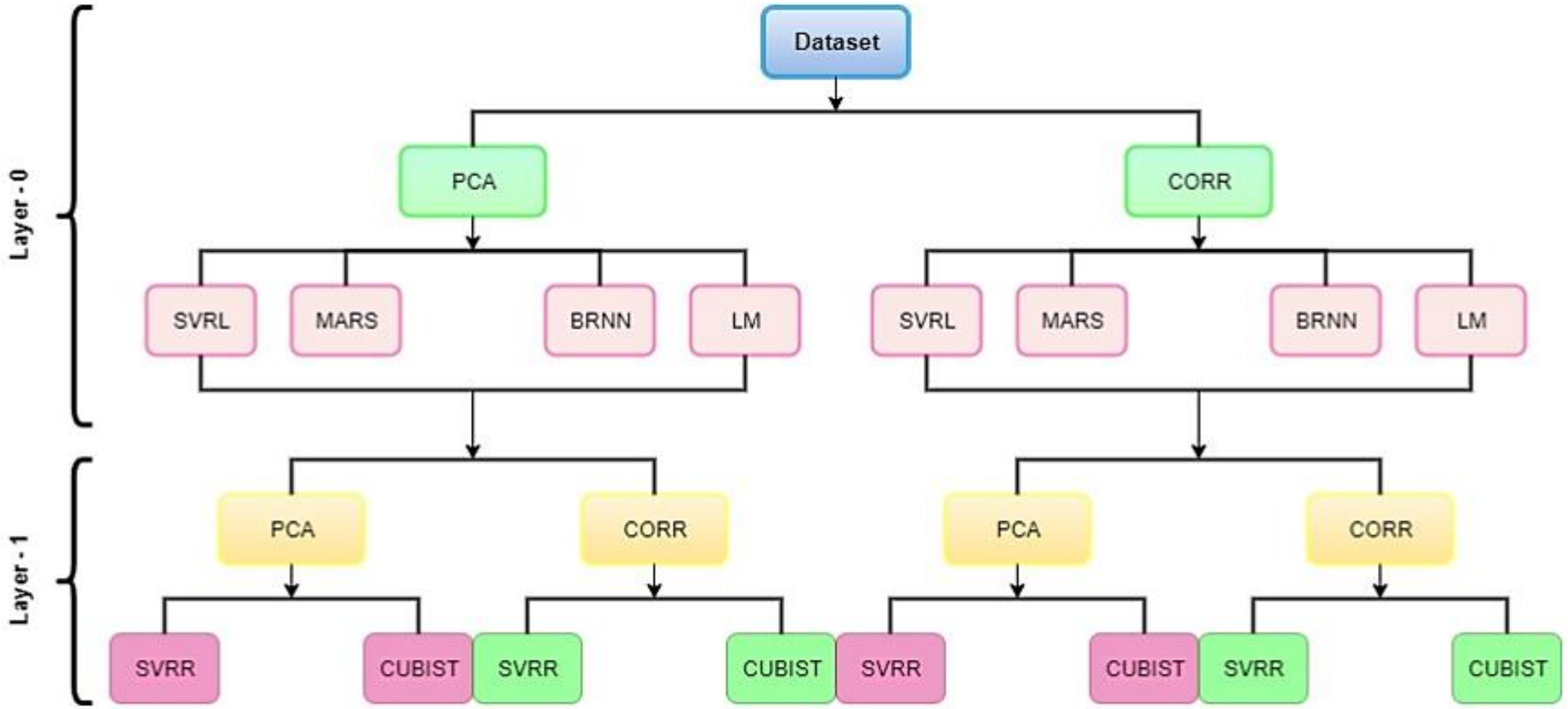
- $sMAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{|y_i| + |\hat{y}_i|/2}\right|$

Statistical test:
Diebold-Mariano test

$$\mathrm{DM} = \frac{\frac{\sum_{n}^{i=1}[d_i]}{n}}{\sqrt{\frac{\mathrm{var}(d_i)}{n-1}}}$$

Methodology

Steps for developing the proposed model:



Results

- CORR models presented better results than PCA models due to the fact that CORR removed the high correlated features;
- The combination CORR-LM obtained best results of the single models;
- The CORR-CORR-CUBIST-STACK obtained best results of the STACK models;
- Comparing best single and STACK models, the CORR-LM presented better performance.

PERFORMANCE MEASURES RESULTS OF THE MODELS

METRIC METHODS \Layer-0				METRIC METHODS STACK \Layer-1			
	sMAPE	RRMSE	R^2		sMAPE	RRMSE	R^2
PCA-SVRL	9.449549e-04	9.779133e-04	0.9997	PCA-PCA-SVRR-STACK	0.0108	0.0145	0.7025
PCA-MARS	9.473985e-04	1.018069e-03	0.9984	PCA-PCA-CUBIST-STACK	0.0056	0.0083	0.3338
PCA-BRNN	1.742184e-03	2.488815e-03	0.9946	PCA-CORR-SVRR-STACK	0.0103	0.0131	0.5773
PCA-LM	1.851099e-04	2.111932e-04	0.9998	PCA-CORR-CUBIST-STACK	0.0057	0.0084	0.4026
CORR-SVRL	1.074919e-03	1.171680e-03	0.9999	CORR-PCA-SVRR-STACK	0.0117	0.0147	0.6002
CORR-MARS	2.798447e-05	6.844400e-05	0.9999	CORR-PCA-CUBIST-STACK	0.0039	0.0059	0.9489
CORR-BRNN	1.969842e-03	3.000576e-03	0.9959	CORR-CORR-SVRR-STACK	0.0101	0.0132	0.4232
CORR-LM	**1.100414e-05**	**2.372946e-05**	**0.9999**	CORR-CORR-CUBIST-STACK	0.0014	0.0024	0.9874

Results

DIEBOLD-MARIANO TEST RESULTS

Model	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)	(N)	(O)	(P)
(A) PCA-PCA-SVRR-STACK	-	8.8299*	8.9003*	8.7990*	1.2679	8.6965*	8.8302*	8.6453*	8.6432*	8.6432*	8.6451*	8.6432*	8.6432*	8.6432*	8.6467*	8.6432*
(B) PCA-PCA-CUBIST-STACK	-	-	-8.7708*	-1.757726	-8.9679*	7.2834*	-8.8186*	7.2895*	7.2863*	7.2860*	7.2895*	7.2866*	7.2861*	7.2866*	7.2870*	7.2866*
(C) PCA-CORR-SVRR-STACK	-	-	-	8.7187*	-9.2565*	8.5678*	-9.9921*	8.4939*	8.4909*	8.4908*	8.4937*	8.4909*	8.4909*	8.4909*	8.4958*	8.4909*
(D) PCA-CORR-CUBIST-STACK	-	-	-	-	-8.9366*	7.8333*	-8.7683*	7.6795*	7.6726*	7.6724*	7.6789*	7.6729*	7.6725*	7.6729*	7.6813*	7.6729*
(E) CORR-PCA-SVRR-STACK	-	-	-	-	-	8.8166*	9.2023*	8.7595*	8.7572*	8.7572*	8.7593*	8.7572*	8.7572*	8.7572*	8.7611*	8.7572*
(F) CORR-PCA-CUBIST-STACK	-	-	-	-	-	-	-8.6131*	7.2270*	7.2159*	7.2150*	7.2258*	7.2168*	7.2150*	7.2168*	7.2181*	7.2168*
(G) CORR-CORR-SVRR-STACK	-	-	-	-	-	-	-	8.5384*	8.5353*	8.5353*	8.5382*	8.5354*	8.5353*	8.5354*	8.5403*	8.5354*
(H) CORR-CORR-CUBIST-STACK	-	-	-	-	-	-	-	-	6.8960*	6.8658*	4.8139*	6.9226*	6.8642*	6.9228*	-7.4720*	6.9229*
(I) PCA-SVRL	-	-	-	-	-	-	-	-	-	-2.1311**	-6.8965*	10.2051*	-8.3503*	10.2451*	-7.1830*	10.2507*
(J) PCA-MARS	-	-	-	-	-	-	-	-	-	-	-6.8604*	6.6480*	-4.4925*	6.6627*	-7.1694*	6.6631*
(K) PCA-BRNN	-	-	-	-	-	-	-	-	-	-	-	6.9237*	6.8613*	6.9240*	-7.4023*	6.9240*
(L) PCA-LM	-	-	-	-	-	-	-	-	-	-	-	-	-9.5986*	4.4538*	-7.1964*	8.5665*
(M) CORR-SVRL	-	-	-	-	-	-	-	-	-	-	-	-	-	9.6156*	-7.1692*	9.6163*
(N) CORR-MARS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-7.1965*	1.3622
(O) CORR-BRNN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7.1965*
(P) CORR-LM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: *1% significance level; **5% significance level.

Conclusion

- In this case all single models presented better results when compared to STACK models, obtaining satisfactory values in all performance measures.
- And the single model CORR-LM obtained the best results of them all, in the three performance measures.
- As proposed for future research:

(i) Using different algorithms in both STACK layers;

(ii) Increasing the number of steps ahead for forecasting;

(iii) Using other renewable energy sources - such as wind and biomass - for comparative studies.

Acknowledgments

CAPES

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico

FUNDAÇÃO ARAUCÁRIA
Apoio ao Desenvolvimento Científico e Tecnológico do Paraná

PUCPR

Thank you!