An Overview of Machine Learning Applications on the Energy Sector

Marcelo Guarido
Daniel Trad

CREWES

HOSTS

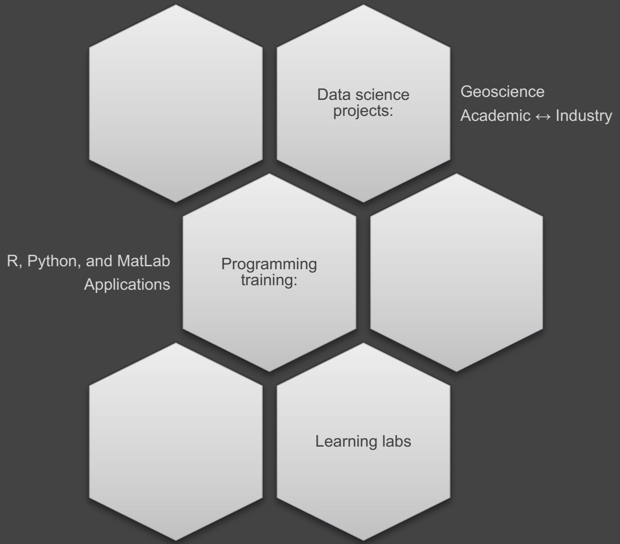
Marcelo Guarido

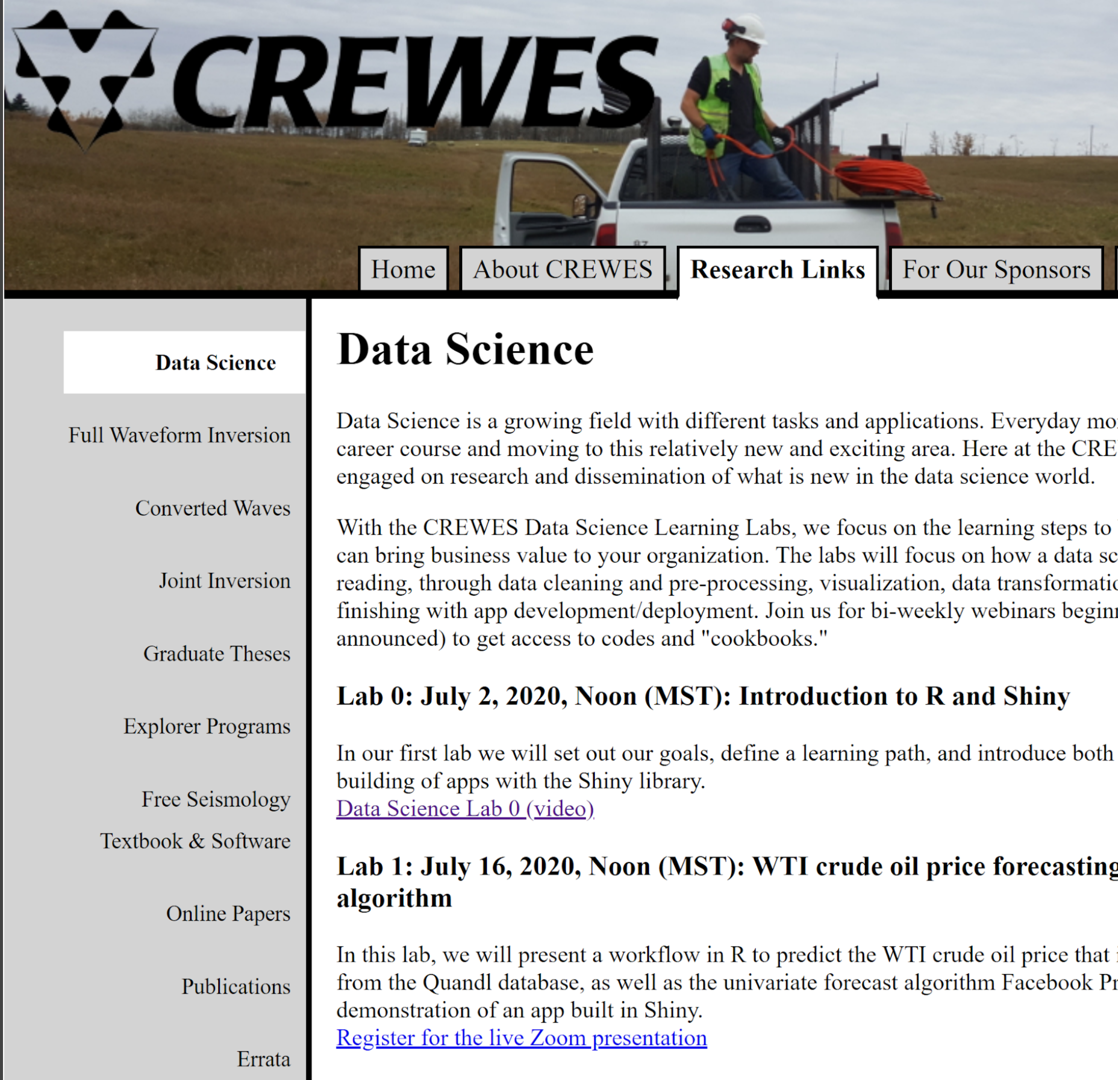
Data Scientist with PhD in Geophysics Head of the CREWES Data Science Initiative.

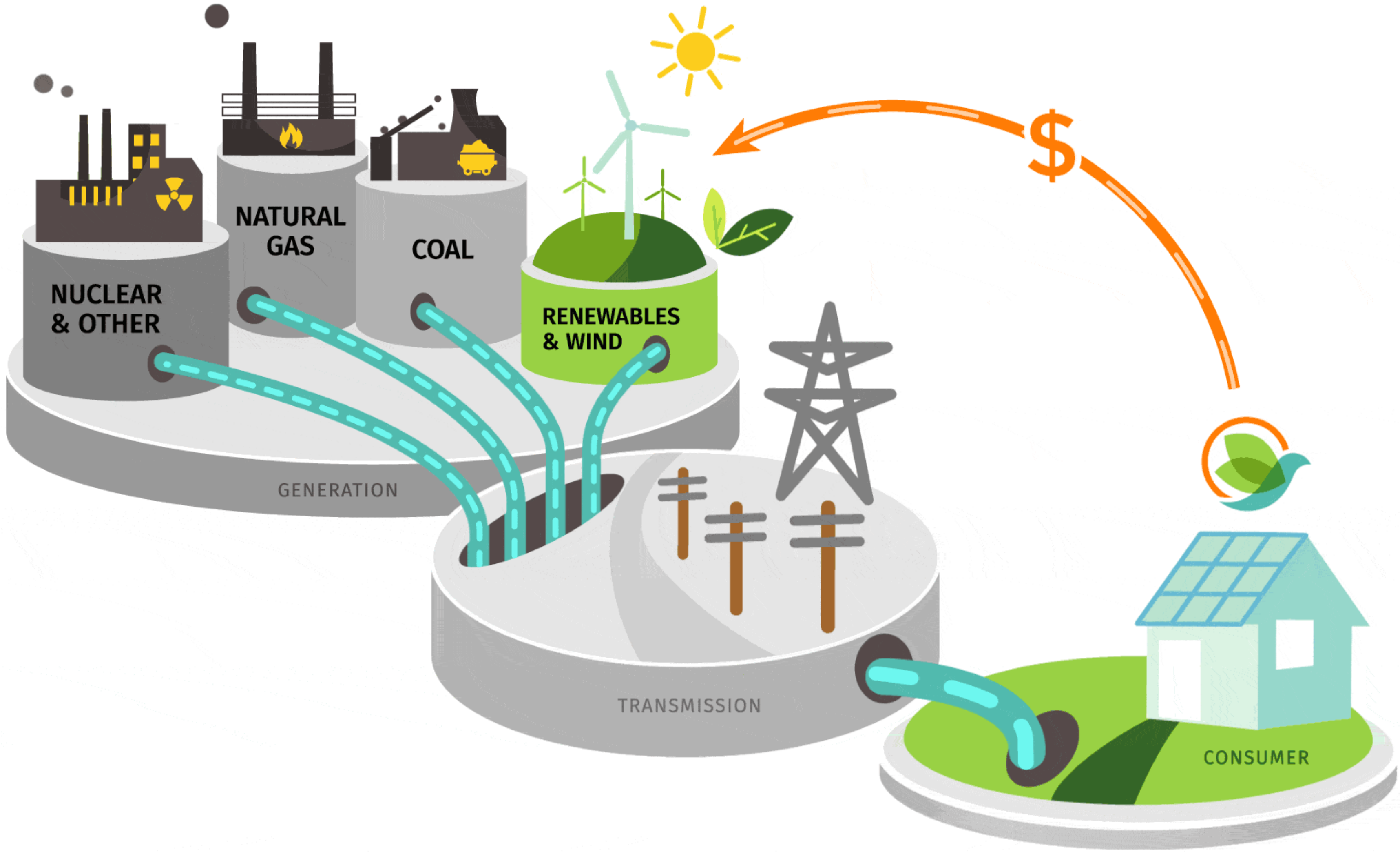
Daniel Trad

Associate Professor and Chair in Exploration Geophysics at University of Calgary.

CREWES Data Science Initiative






NUCLEAR & OTHER
NATURAL GAS
COAL
RENEWABLES & WIND
$
GENERATION
TRANSMISSION
CONSUMER

TOTAL

SIEMENS

BERKSHIRE HATHAWAY ENERGY

Ørsted

Vestas

EDF ENERGY

AVANGRID RENEWABLES

eni

seas-nve

NEXTera ENERGY

Chevron

ecoplexus

AMERICAN ELECTRIC POWER

GE Energy

nationalgrid renewables

Xcel Energy

mCloud

Georgia Power

bp

Invenergy

SUZLON
POWERING A GREENER TOMORROW

Exxon

DUKE ENERGY

BORREGO SOLAR

CYPRESS CREEK RENEWABLES

ENERGY CONSUMPTION AND PRICE

01

Machine learning techniques to analyze and forecast energy consumption and price

OIL & GAS

02

Machine learning as a tool to help interpretation and decision making

RENEWABLE ENERGY

03

How machine learning is used to optimize the use of renewable resources

LEARNING SESSIONS 2021

04

What expect for this year

"Just as **electricity transformed almost everything 100 years ago**, today I actually have a hard time thinking of an industry that I don't think **artificial intelligence will transform in the next several years**."

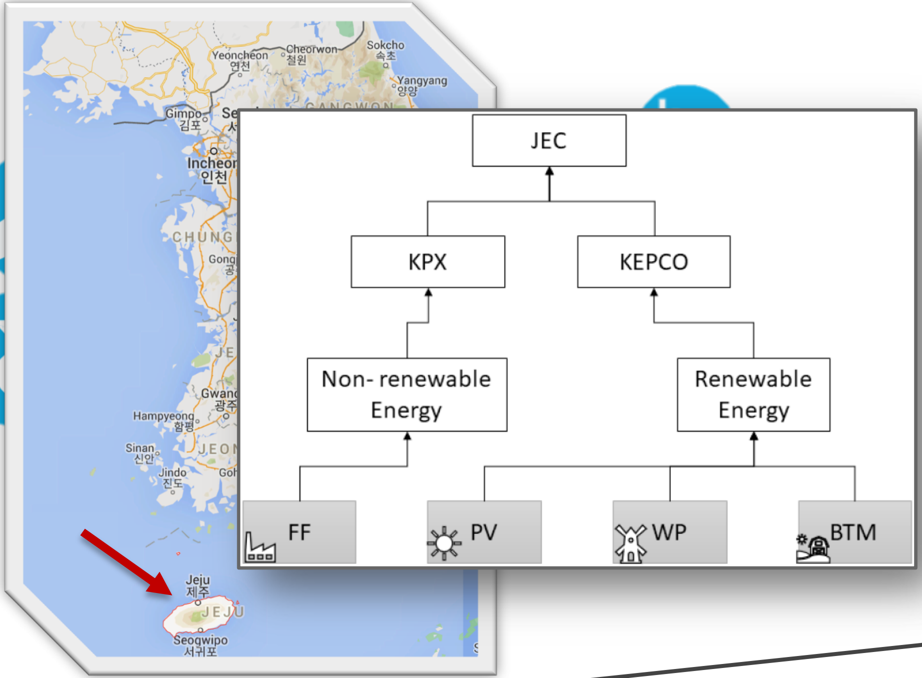
— Andrew NG

01

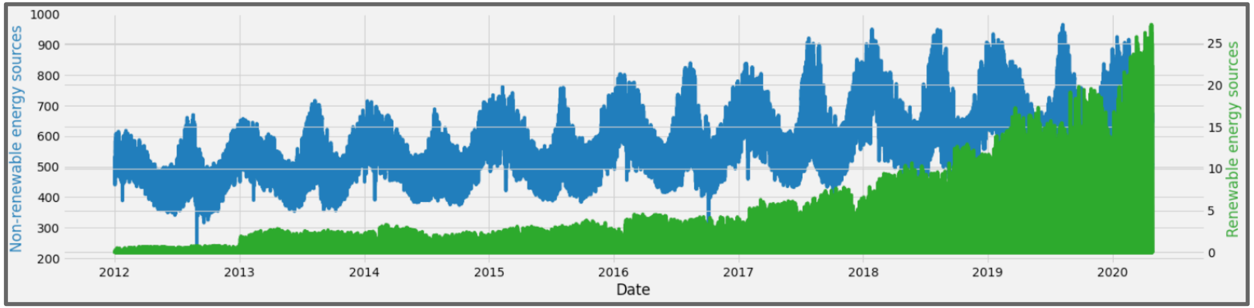
ENERGY CONSUMPTION AND PRICE

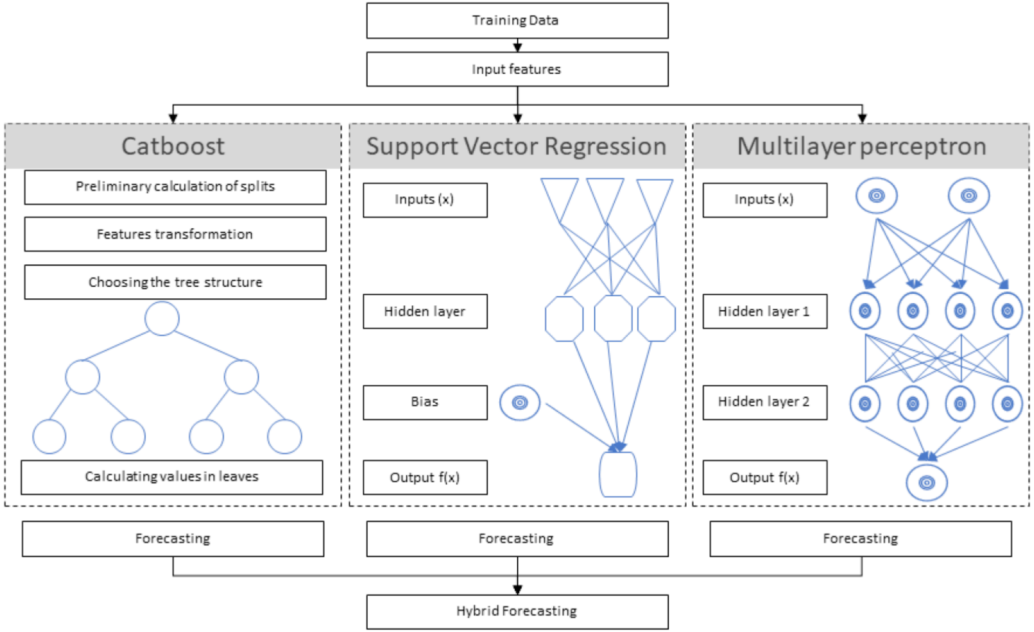
Examples of energy consumption and price forecasting (plus other things)

MACHINE LEARNING-BASED APPROACH TO PREDICT ENERGY CONSUMPTION OF RENEWABLE AND NONRENEWABLE POWER SOURCES by Khan, P.W.; Byun, Y.-C.; Lee, S.-J.; Kang, D.-H.; Kang, J.-Y.; Park, H.-S, 2020



- **JEJU ISLAND** is used as a test lab
- **REPLACE** non-renewable energy by renewable energy by 2030
- **FORECAST** energy consumption (MW) from all sources





ENERGY MW

Total consumption (renewable + non-renewable)

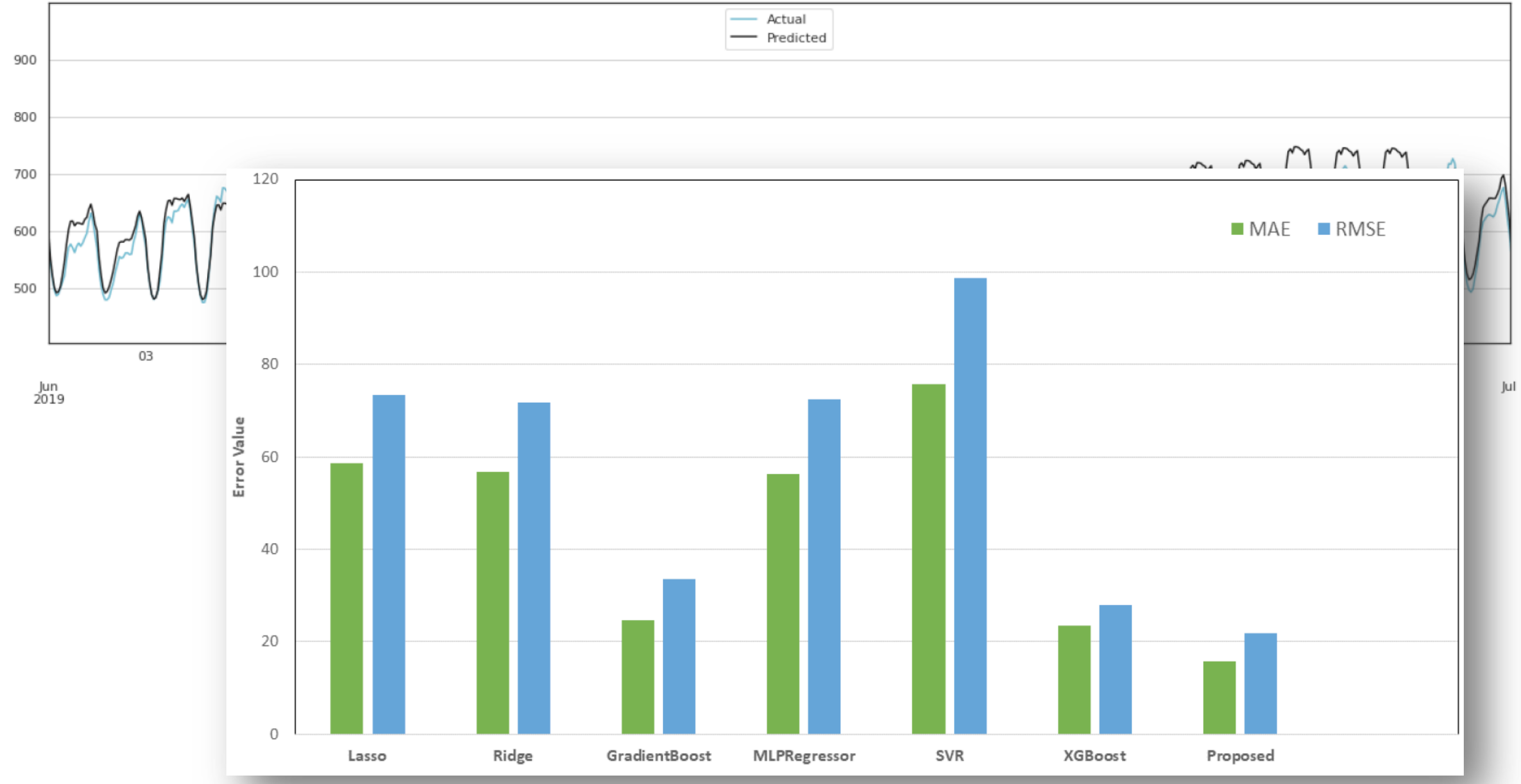
MODEL

Hybrid model

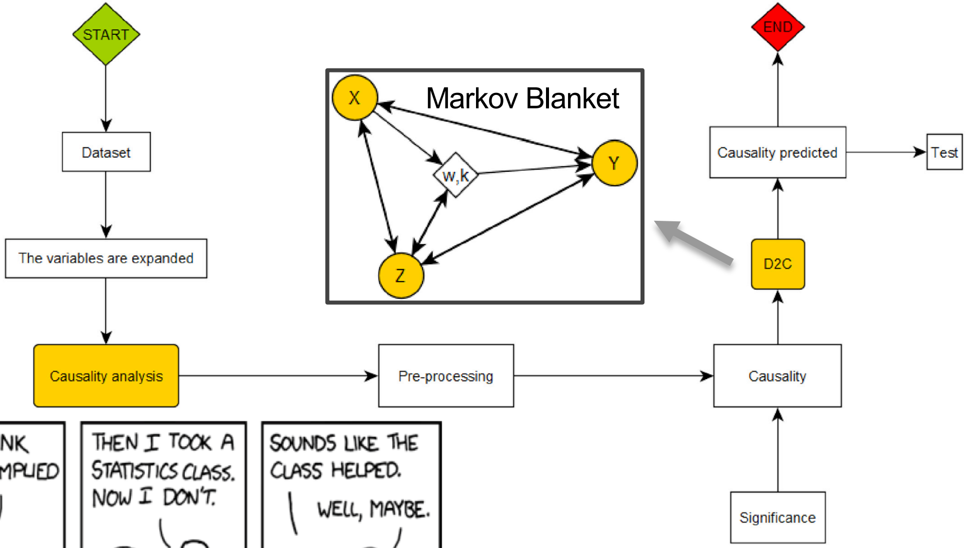
COMPARE

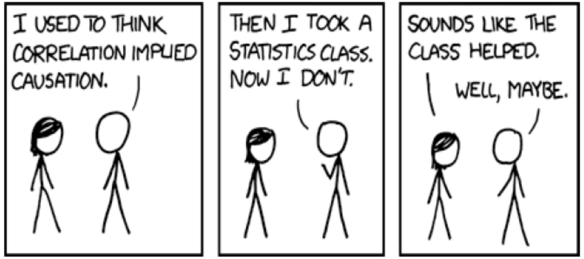
Against other models

FORECASTING



A MACHINE LEARNING APPROACH ON THE RELATIONSHIP AMONG SOLAR AND WIND ENERGY PRODUCTION, COAL CONSUMPTION, GDP, AND CO_2 EMISSIONS by Magazzino, C.; Mele, M.; Schneider, N., 2021





CAUSAL RELATIONSHIP among solar and wind energy production, coal consumption, economic growth, and CO_2 emissions

USA, CHINA, AND INDIA are studied in this paper

COUNTRIES with the larger investments on renewable energy, but also the highest CO_2 emissions.

DISCUSSION

Relationships found on the paper:

1. There is apparent unilateral causality between **GDP** and **CO_2 emissions**. Higher GDP causes higher CO_2 emissions, but the opposite is not true (do you agree with that information?).

2. Bi-directional causality between **GDP** and **coal consumption**.

3. No apparent cause effect of **wind and solar energies** with **CO_2 emissions**. Only India showed some correlation, but due to equipment construction and drying peat bogs.

4. Strong causal effect of **coal consumption** with **CO_2 emissions**.

5. Possibility to forecast **CO_2 emissions**.

TEXT-BASED CRUDE OIL PRICE FORECASTING: A DEEP LEARNING APPROACH

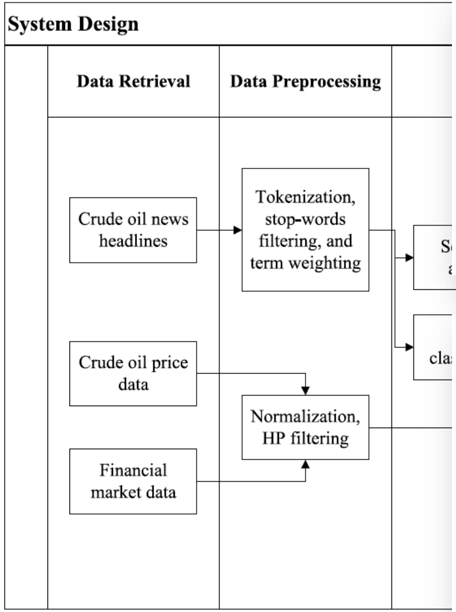
by Li, X.; Shang, W.; Wang, S., 2019

CRUDE OIL FORECAST is a complex task that involves supply and demand variables, as well as unpredictable factors

CRUDE OIL NEWS column from Investing.com

FORECAST brent crude oil price

METHODOLOGY



Topic words in each of the topics.

	Topic proportion	Top 20 topic words with the largest weights
Topic 1	20.50%	**natural**, **gas**, futures, **weather**, U.S, supply, forecasts, data, low, prices, high, gains, rally, focus, report, storage, mild, outlook, **warm**, off
Topic 2	68.80%	**crude**, **oil**, U.S, stocks, data, futures, low, **supply**, **dollar**, Asia, gains, up, ahead, **euro**, NYNEX, high, China, fall, rise, fed
Topic 3	3.92%	oil, white, house, **Iran**, **Russia**, state, energy, export, **China**, inflation, new, bank, ban, rates, exclusive, U.S, manufacturing, index, fuel, **Mexico**
Topic 4	6.76%	**stocks**, **trade**, close, down, low, up, high, **composite**, FTSE, **Dax**, **S&P**, **Dow**, **Jones**, **industrial**, China, U.S, Malaysia, mixed, Russia, **shares**

Note: Topic words with the largest weights in each of the topics evaluated by the LDA model. The words in bold are those with specific meanings.

Forecasting performances of the random forest, SVR and linear regression models.

	Text features (1)	Financial features (2)	Combination: (1) + (2)	Percentage improvement from (2) to (1) + (2)
MAE				
Random forest	0.0785	0.0082	0.0073	**12.32%**
SVR	0.0252	0.0032	0.0030	**6.67%**
Linear regression	0.0854	0.0035	0.0045	−22.22%
RMSE				
Random forest	0.0883	0.0092	0.0088	**4.55%**
SVR	0.0325	0.0041	0.0040	**2.50%**
Linear regression	0.0953	0.0044	0.0056	−21.42%

RESOURCES – PART 01

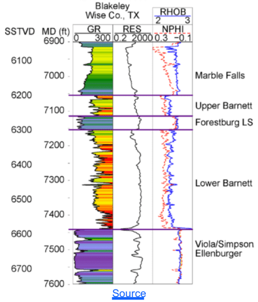
- Khan, P.W.; Byun, Y.-C.; Lee, S.-J.; Kang, D.-H.; Kang, J.-Y.; Park, H.-S. Machine Learning-based Approach To Predict Energy Consumption Of Renewable And Nonrenewable Power Sources. *Energies,* **2020**, *13*, 4870. https://doi.org/10.3390/en13184870
- Magazzino, C.; Mele, M.; Schneider, N. A Machine Learning Approach On The Relationship Among Solar And Wind Energy Production, Coal Consumption, GDP, And CO_2 Emissions. *Renewable Energy,* **2021**, *167*, 99-167. https://doi.org/10.1016/j.renene.2020.11.050
- Li, X.; Shang, W.; Wang, S. Text-based Crude Oil Price Forecasting: A Deep Learning Approach. *International Journal of Forecasting,* **2019**, *4*, 1548-1560. https://doi.org/10.1016/j.ijforecast.2018.07.006

02

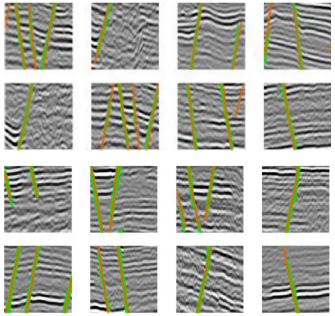
OIL & GAS

Machine learning as a tool to help interpretation and decision making

MACHINE LEARNING METHODS IN GEOSCIENCE by Maniar, H.; Ryali, S.; Kulkarni, M. S.; Abubakar, A., 2018



Source

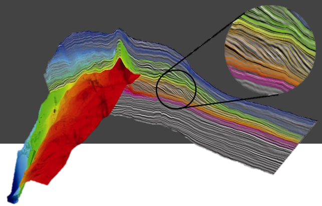


WELL LOG CORRELATION

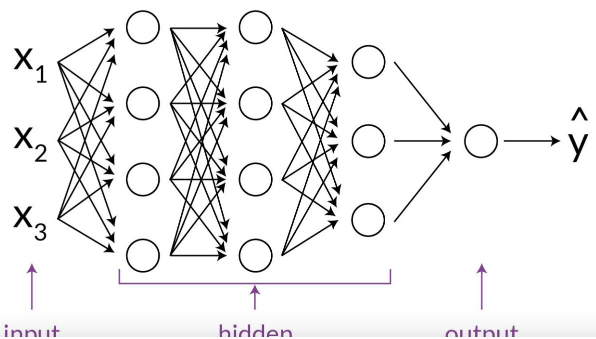
Using GR and resistivity (triple combo)

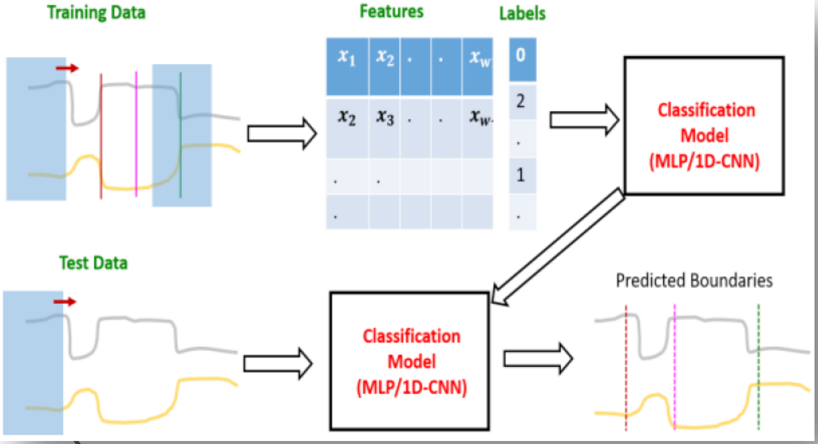
FAULT DETECTION

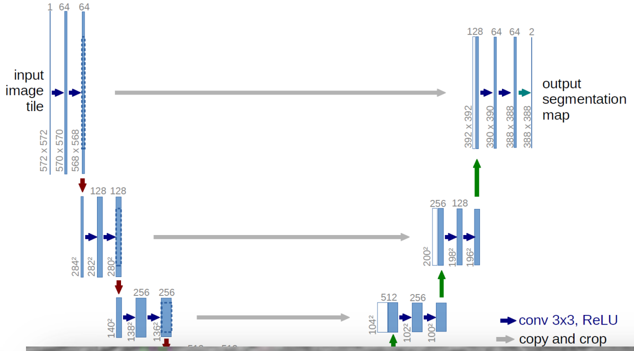
On 2D seismic data

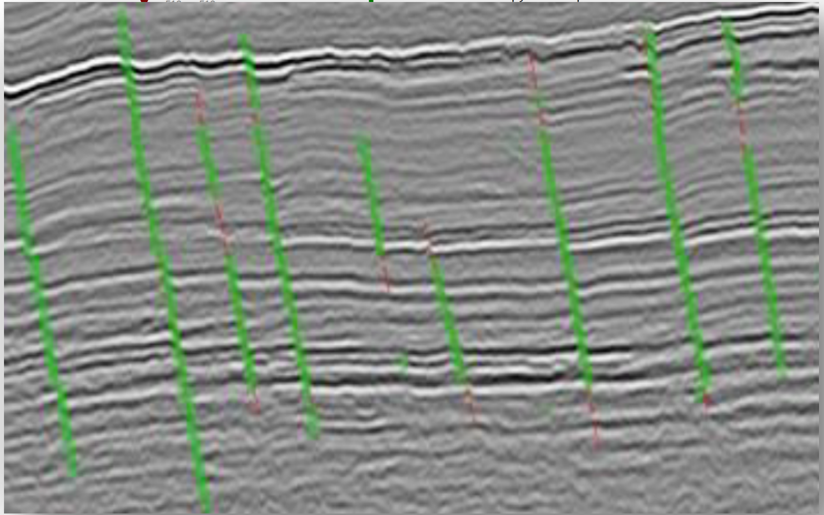


Deep Neural Network

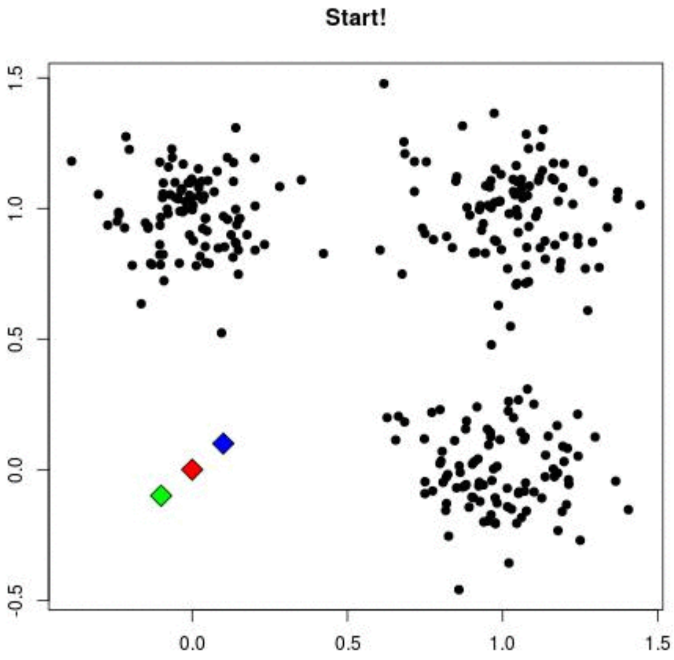








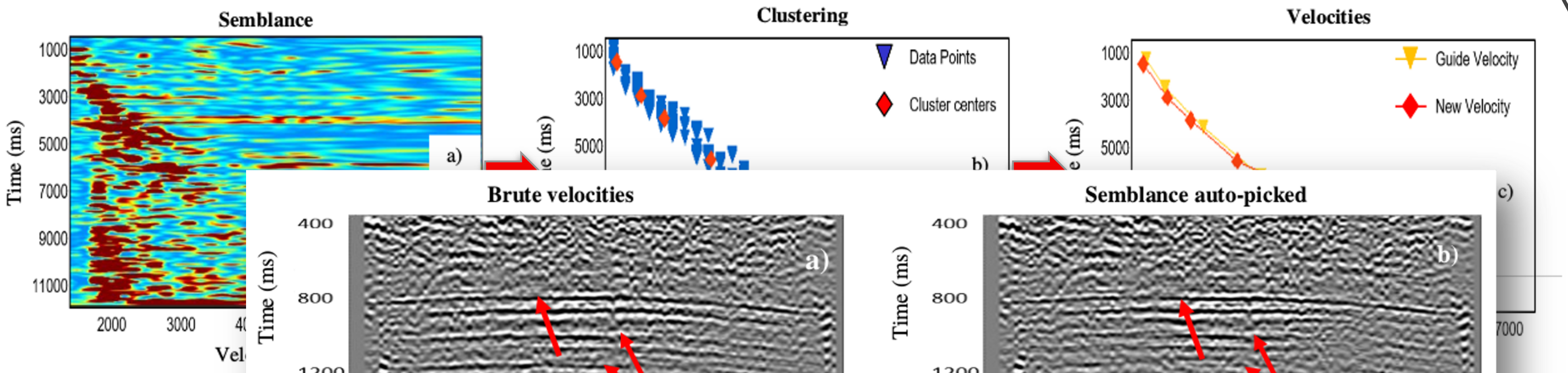
MACHINE LEARNING ASSISTED VELOCITY AUTO-PICKING by Kenneth Smith, 2017

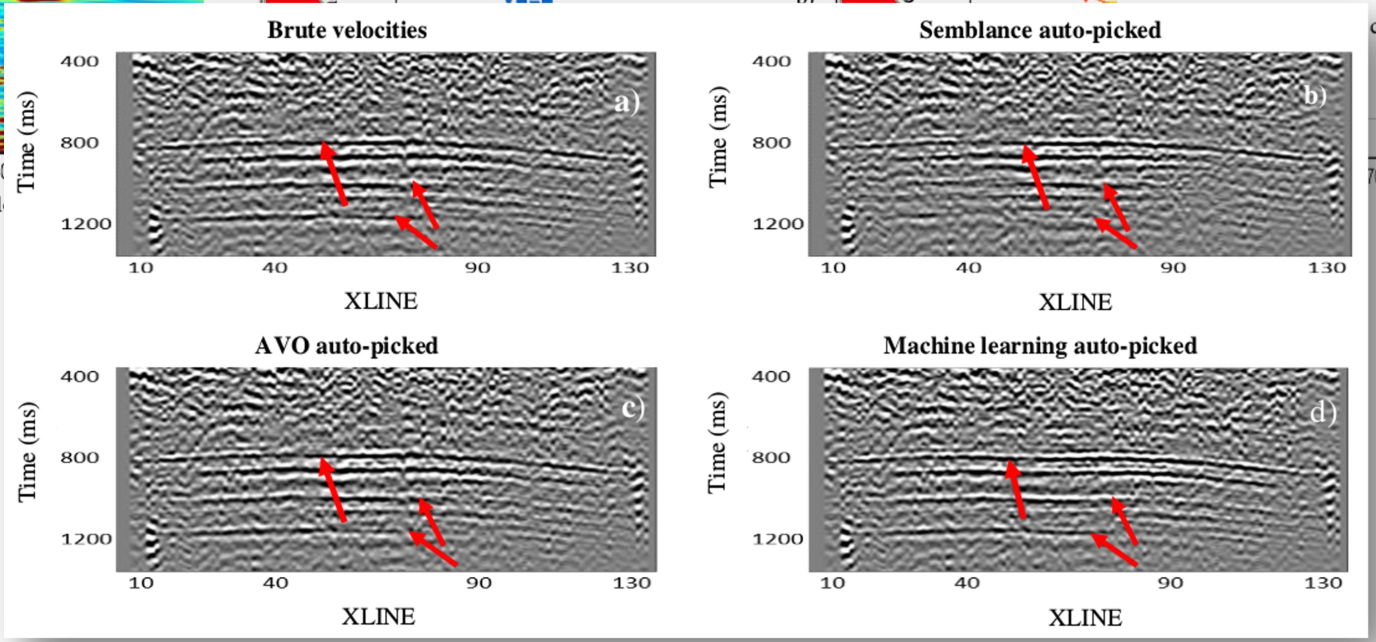


SEMBLANCE auto-picking

UNSUPERVISED learning using k-nearest neighbors

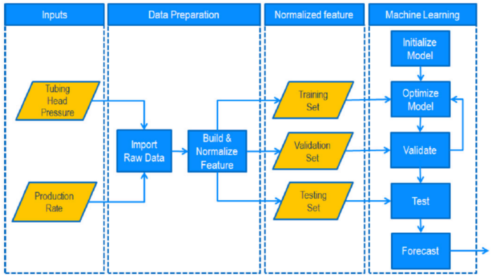
CLUSTER centroid is used as the picking point

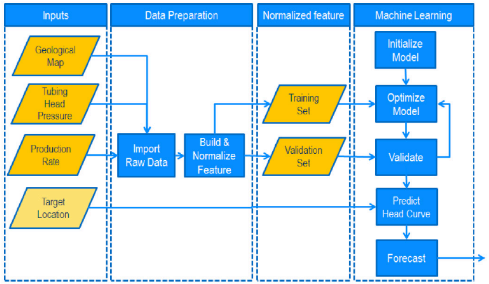

Semblance
Time (ms)
1000
3000
5000
7000
9000
11000
2000
3000
a)
Clustering
1000
3000
5000
Data Points
Cluster centers
b)
Velocities
1000
3000
5000
Guide Velocity
New Velocity
c)
7000


Brute velocities
Time (ms)
400
800
1200
a)
10
40
90
130
XLINE
Semblance auto-picked
Time (ms)
400
800
1200
b)
10
40
90
130
XLINE
AVO auto-picked
Time (ms)
400
800
1200
c)
10
40
90
130
XLINE
Machine learning auto-picked
Time (ms)
400
800
1200
d)
10
40
90
130
XLINE

DATA DRIVEN PRODUCTION FORECASTING USING MACHINE LEARNING

by Q.; Banerjee, R.; Gupta, S.; Li, J.; Zhou, W.; Jeyachandra, B., 2018



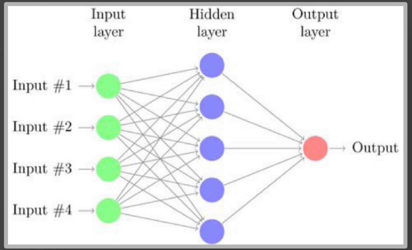


EXISTING WELL

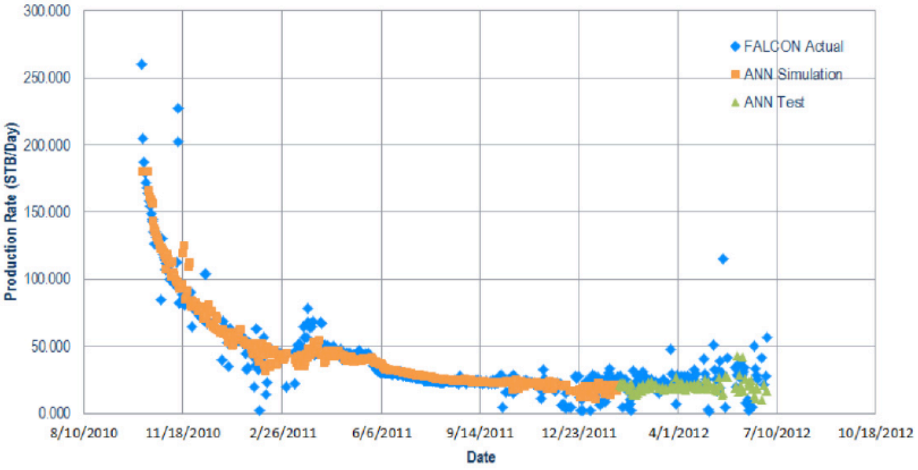
Using existing production history

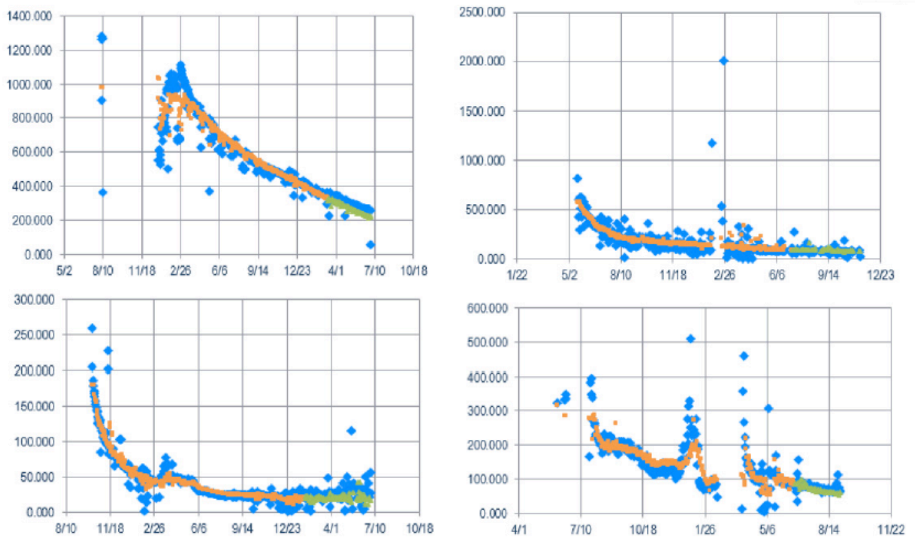
NEW WELL

Using historical production from nearby wells with similar geology

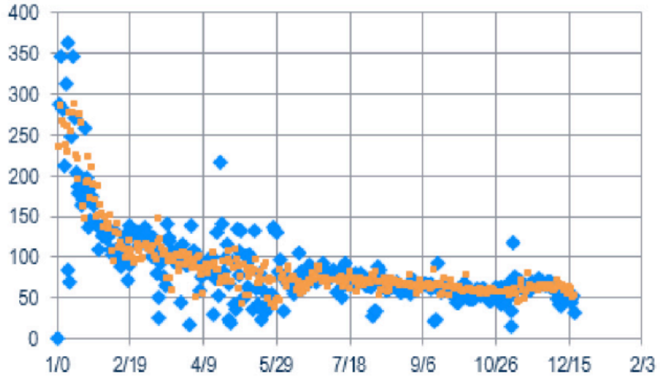


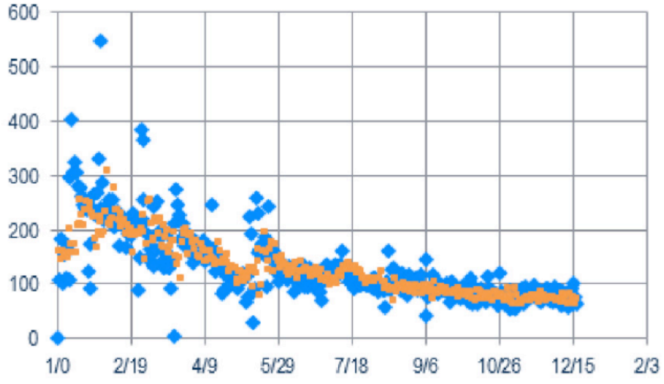
FORECASTING FOR EXISTING WELLS

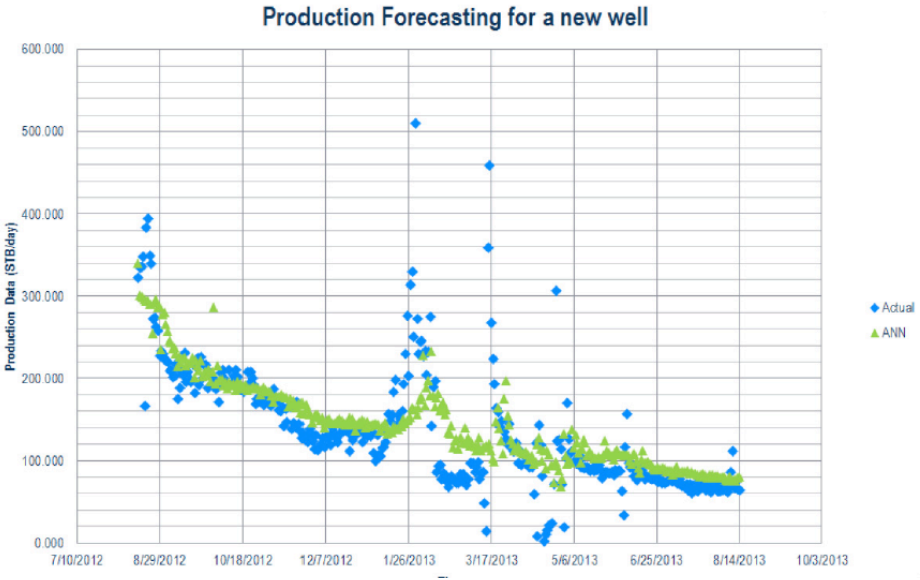


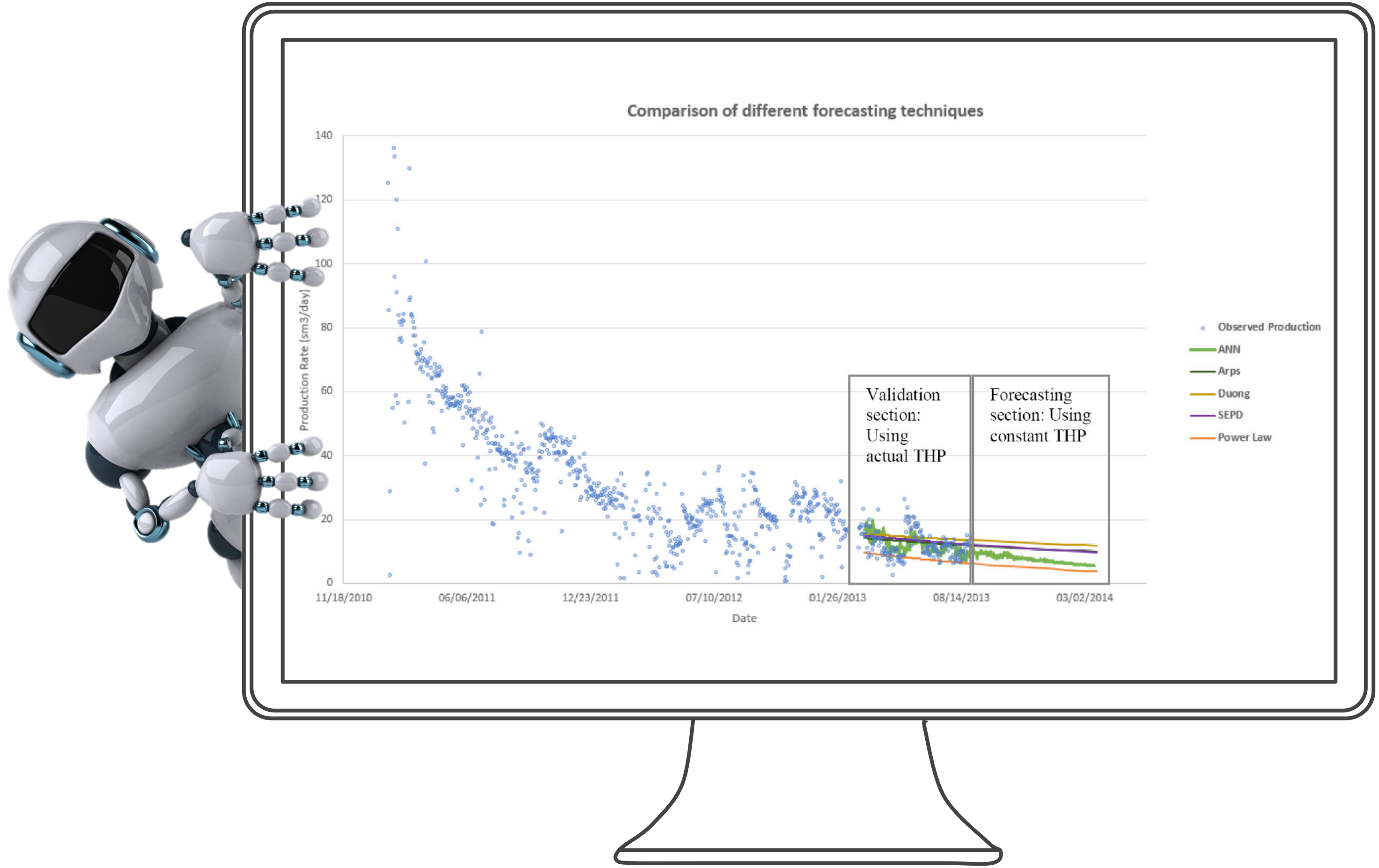


FORECASTING FOR NEW WELLS



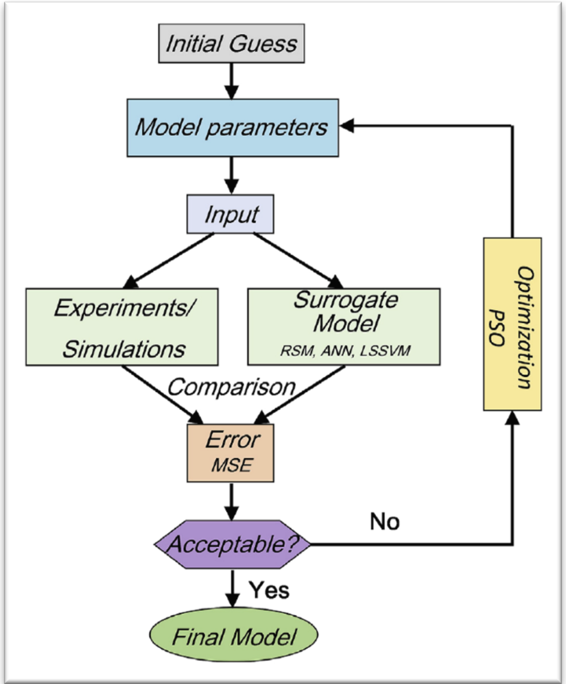






Comparison of different forecasting techniques
Production Rate (sm3/day)
140
120
100
80
60
40
20
0
11/18/2010
06/06/2011
12/23/2011
07/10/2012
01/26/2013
08/14/2013
03/02/2014
Date
Validation section: Using actual THP
Forecasting section: Using constant THP
Observed Production
ANN
Arps
Duong
SEPD
Power Law

APPLICATION OF ARTIFICIAL INTELLIGENCE TO FORECAST HYDROCARBON PRODUCTION FROM SHALES by Panja, P; Velasco, R.; Pathak, M.; Deo, M., 2018

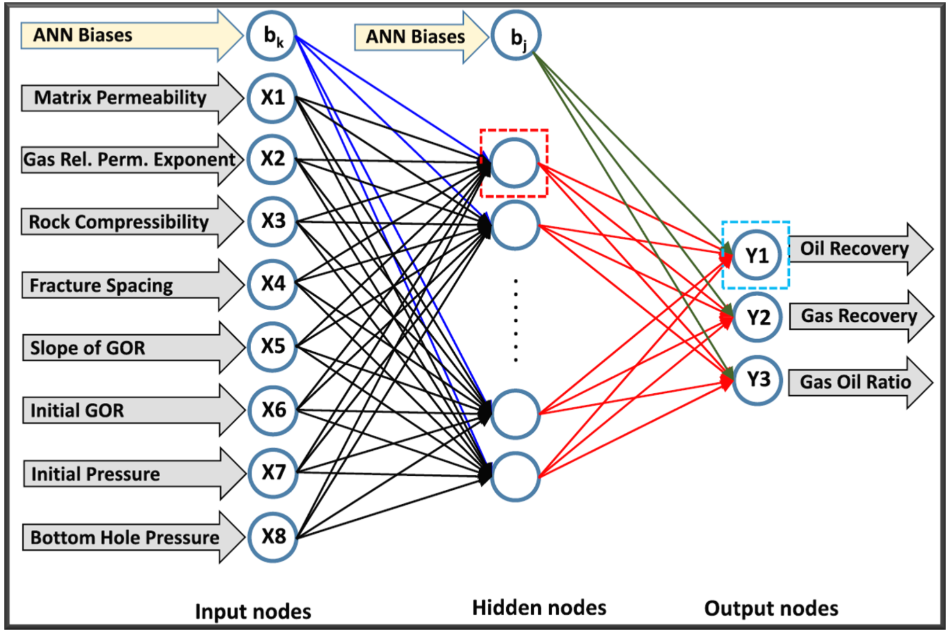


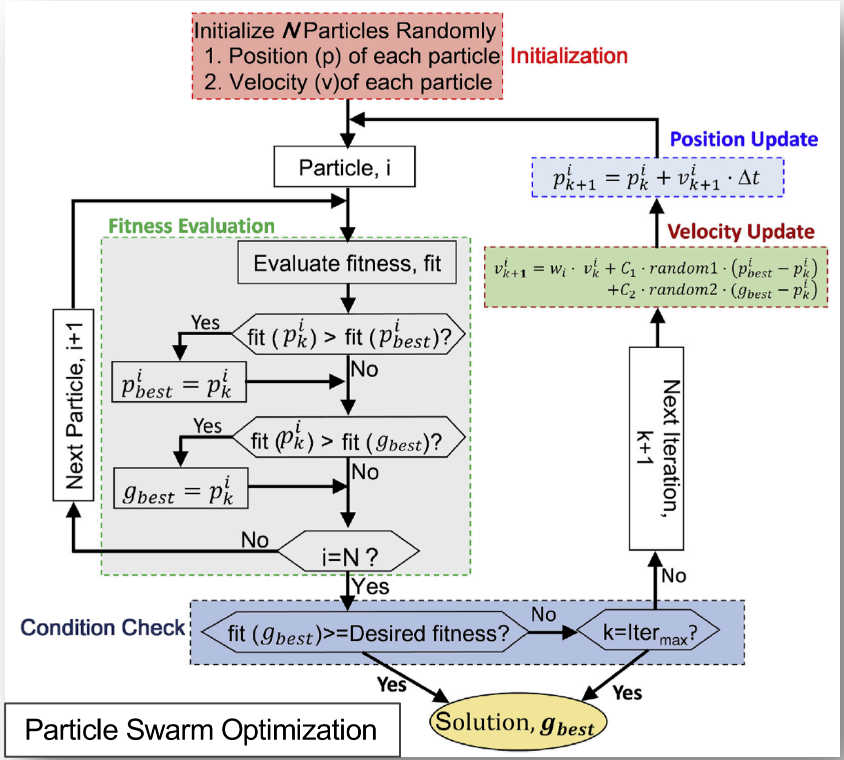
RESPONSE SURFACE MODEL (RSM) is a quadratic fit of the data

LEAST SQUARE SUPPORT VECTOR MACHINE (LSSVM) solves a linear system

ARTIFICIAL NEURAL NETWORKS (ANN) is used with 14 hidden layers

METHODOLOGY





Surrogate Model		Particle Swarm Optimization				
Name	Parameters Optimized	Number of Particles for a single parameter	C_1	C_2	w	Maximum Iteration
RSM	1 intercept 44 coefficients	100	2	2	0.6	1000
LSSVM	1 Regularization parameter (γ) 1 Kernel parameter (σ)	100	2	2	0.6	1000
ANN	15 Biases 126 wt	100	2	2	0.6	1000

FORECASTING

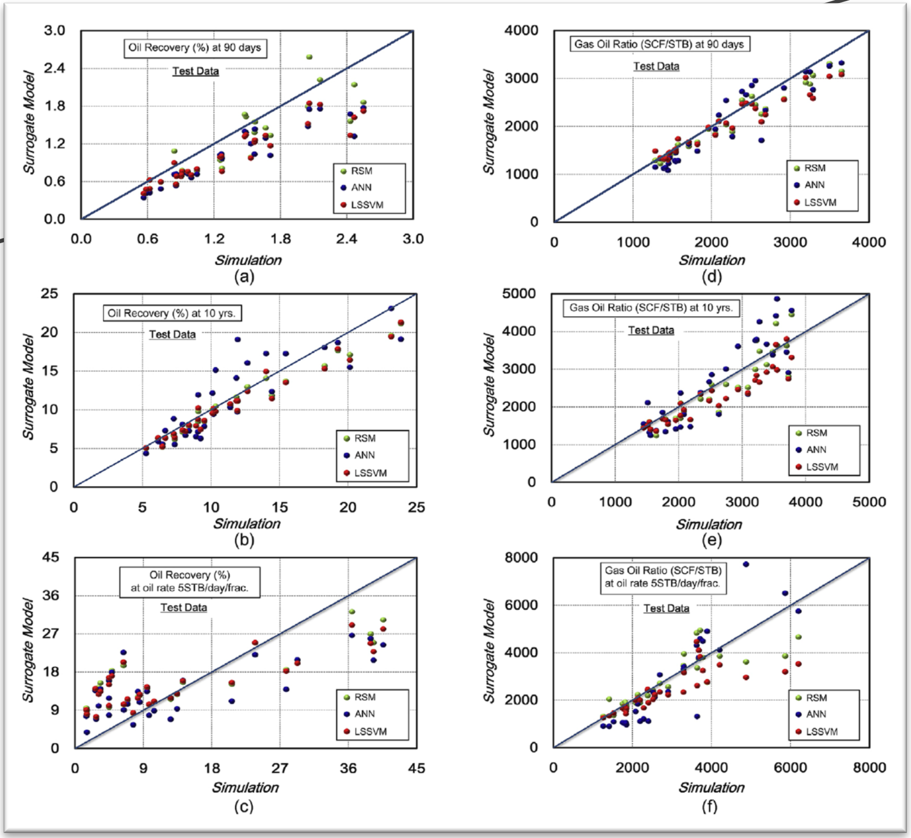
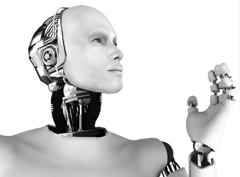


Table A1
Coefficient of determination (R^2) of RSM, LSSVM and ANN for all models.

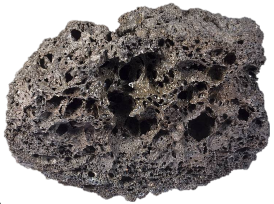
Output	Model	Training Data			Test Data		
		RSM	LSSVM	ANN	RSM	LSSVM	ANN
Oil Recovery	90 days	0.99	0.99	0.96	0.69	0.52	0.51
	1 year	0.98	0.99	0.98	0.78	0.69	0.53
	5 years	0.99	0.99	0.99	0.63	0.81	0.60
	10 years	0.99	0.99	0.98	0.91	0.9	0.72
	15 years	0.99	0.99	0.99	0.97	0.93	0.84
	Rate Based	0.98	0.98	0.99	0.57	0.54	0.48
Gas Oil Ratio	90 days	0.98	0.99	0.95	0.92	0.84	0.80
	1 year	0.98	0.98	0.96	0.93	0.91	0.90
	5 years	0.98	0.98	0.97	0.41	0.73	0.30
	10 years	0.88	0.92	0.93	0.76	0.73	0.46
	15 years	0.83	0.77	0.84	0.79	0.75	0.32
	Rate Based	0.84	0.88	0.92	0.68	0.45	0.43

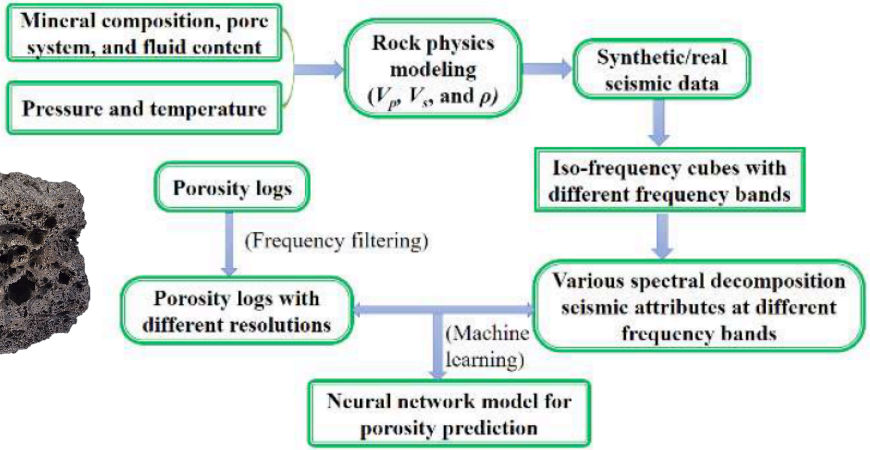
Table A2
Normalized Root Mean Square Error (NRMSE) of RSM, LSSVM and ANN for all models.

Output	Model	Training Data			Test Data		
		RSM	LSSVM	ANN	RSM	LSSVM	ANN
Oil Recovery	90 days	1.9	1.9	3.5	16.5	20.3	20.7
	1 year	2.4	2.3	2.5	12.4	14.7	18.1
	5 years	2.0	1.9	2.1	16.1	11.5	16.7
	10 years	1.9	1.7	2.6	7.9	8.5	14
	15 years	2.7	2.4	2.1	4.9	7.3	10.8
	Rate Based	3.5	3.3	2.4	20.7	21.2	22.6
Gas Oil Ratio	90 days	2.6	2.0	4.6	8.7	11.8	13.3
	1 year	3.3	3.3	4.2	7.9	9.3	9.7
	5 years	3.0	3.1	3.7	24.0	16.1	26.2
	10 years	5.7	4.6	4.3	16.1	17.2	24.3
	15 years	5.8	6.8	5.6	14.4	15.5	25.7
	Rate Based	5.2	4.5	3.8	14.1	18.4	18.8



POROSITY PREDICTION USING MACHINE LEARNING by Jiang, L.; Castagna, J. P.; Russell, B., 2020

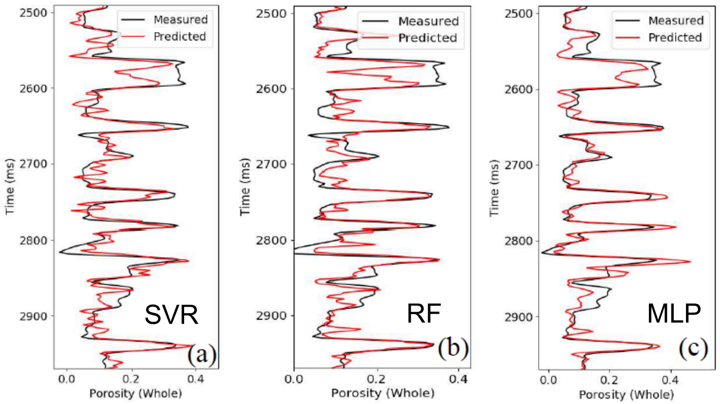




POROSITY predicted from seismic data

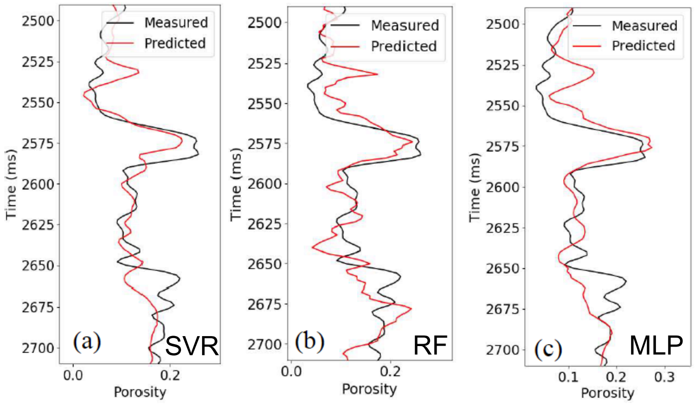
SPECTRAL decomposition

HIGHER correlation then conventional methods



Methods	SVR	RF	MLP	PNN
Conventional seismic attributes	0.88	0.90	0.89	0.74
Conventional seismic attributes and SD attributes	0.91	0.89	0.91	0.82
SD seismic attributes	0.84	0.83	0.89	0.70
SD seismic attributes (narrow frequency bands)	0.86	0.80	0.94	N/A

SYNTHETIC



Methods	SVR	RF	MLP	PNN
Conventional seismic attributes	0.56	0.48	0.66	0.41
Conventional seismic attributes and SD attributes	0.74	0.56	0.64	0.78
SD seismic attributes	0.58	0.50	0.55	0.72
SD seismic attributes (narrow frequency band)	0.81	0.72	0.71	N/A

FIELD DATA

RESOURCES – PART 02

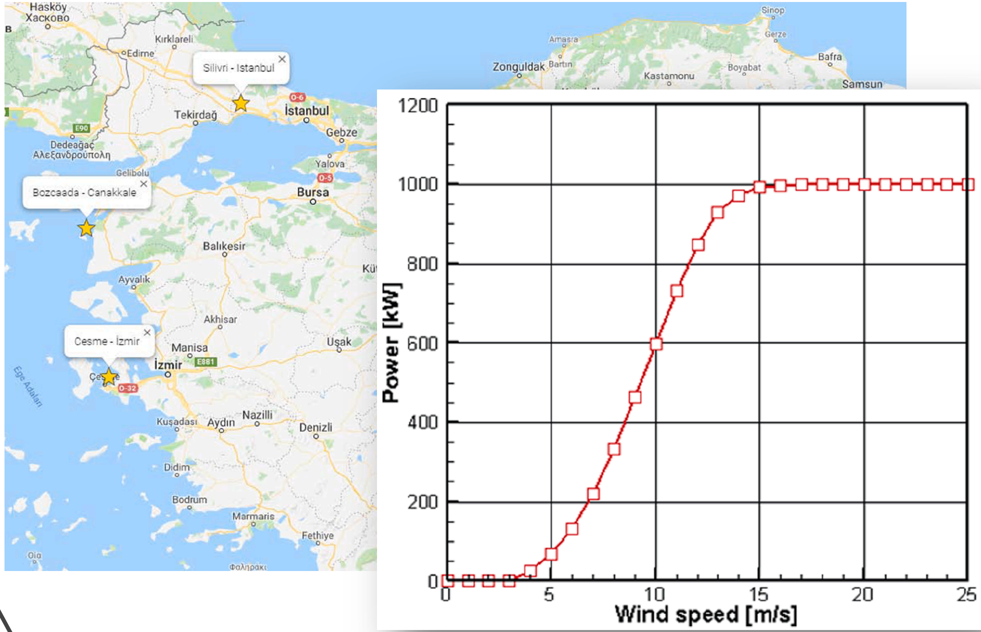
- Maniar, H.; Ryali, S.; Kulkarni, M. S.; Abubakar, A. Machine Learning Methods In Geoscience. *SEG Technical Program Expanded Abstracts,* **2018**, 4638-4642. 10.1190/segam2018-2997218.1
- Smith, K. Machine Learning Assisted Velocity Autopicking. *SEG Technical Program Expanded Abstracts,* **2017**, 5686-5690. 10.1190/segam2017-17684719.1
- Cao, Q.; Banerjee, R.; Gupta, S.; Li, J.; Zhou, W.; Jeyachandra, B. Data Driven Production Forecasting Using Machine Learning. *SPE Argentina Exploration and Production of Unconventional Resources Symposium*, Buenos Aires, Argentina, June 2016. https://doi.org/10.2118/180984-MS
- Panja, P; Velasco, R.; Pathak, M.; Deo, M. Application Of Artificial Intelligence To Forecast Hydrocarbon Production From Shales. *Petroleum,* **2018**, *4*, 75-89. https://doi.org/10.1016/j.petlm.2017.11.003
- Jiang, L.; Castagna, J. P.; Russell, B. Porosity Prediction Using Machine Learning. *SEG Technical Program Expanded Abstracts,* **2020**, 3862-3866. 10.1190/segam2020-w13-04.1

03

RENEWABLE ENERGY

How machine learning is used to optimize the use of renewable resources

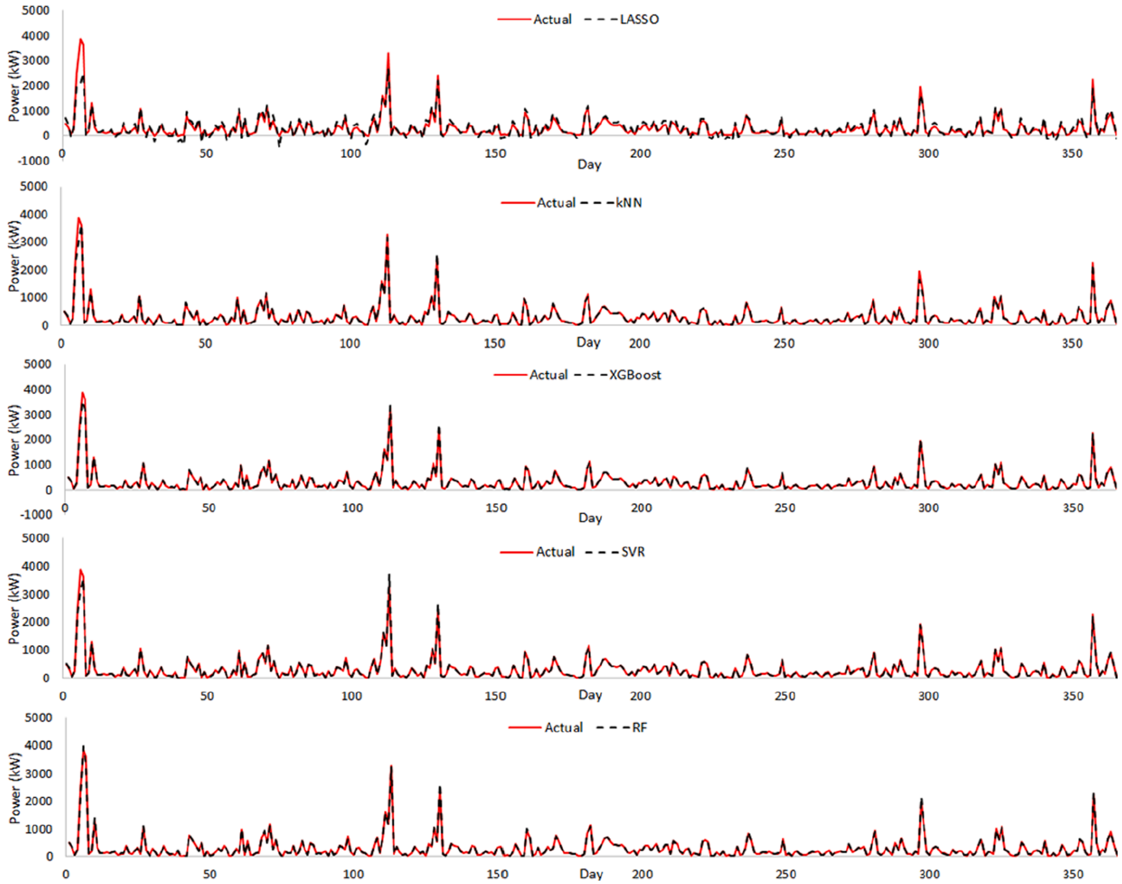
WIND POWER FORECASTING BASED ON DAILY WIND SPEED DATA USING MACHINE LEARNING ALGORITHMS by Demolli, H.; Dokuz, A. S.; Ecemis, A.; Gokcek, M., 2019



WIND energy needs efficiency study prior installation

FORECAST wind power from historical data

MODELS showed to be independent of location



Forecasting accuracy of the algorithms for daily wind speed and standard deviation.

Algorithm/Metric	R^2	RMSE	MAE
LASSO	0.8619	164.61	88.85
kNN	0.9852	53.82	7.197
xGBoost	0.9939	34.40	6.528
SVR	0.992	38.52	**5.430**
RF	**0.995**	**30.224**	7.048

Forecasting accuracy of the algorithms using only daily wind speed.

Algorithm/Metric	R^2	RMSE	MAE
LASSO	0.823	186.10	94.67
kNN	0.943	105.76	42.47
xGBoost	0.932	115.47	41.58
SVR	**0.955**	**93.13**	**32.63**
RF	0.921	123.88	42.76

QUANTIFYING ROOFTOP PHOTOVOLTAIC SOLAR ENERGY POTENTIAL: A MACHINE LEARNING APPROACH by Assouline, D.; Mohajeri, N.; Scartezzini, J. L., 2017

PHOTOVOLTAIC (PV) potential on rooftops in Switzerland using SVM and GIS

RESULTS show that 81% of rooftops (328 km^2) could be used for solar energy

TOTAL energy generated could be of around 17.86 TW h, or 28% of energy consumption in 2015

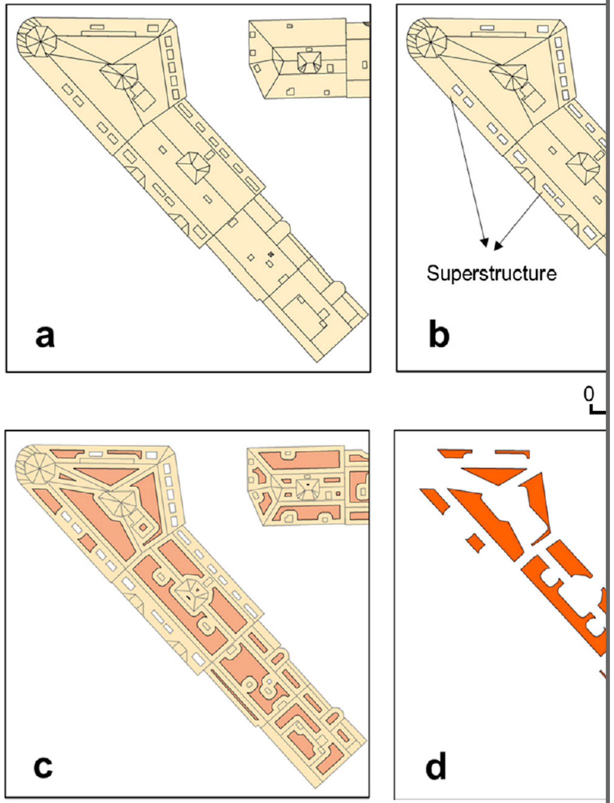


Fig. 3. Schematic presentation of different steps to estimate available roof area using ArcGIS. (a) Building polygon
(e.g. chimney, dormers, staircase), (b) removing the superstructures from roof surfaces, (c) creating 1 m^2 buffer ar
roof area for PV installation after removing the areas smaller than 28 m^2.

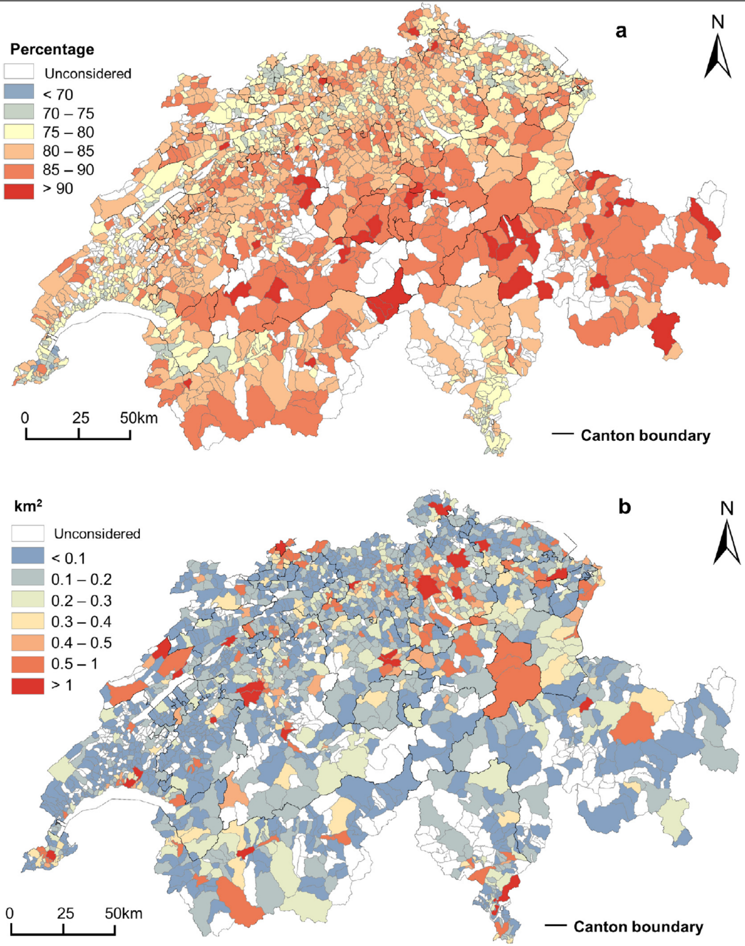


Fig. 4. (a) The ratio of average available roof area for PV (A_R) to the average ground floor area (A_f) for each commune, percentage (b) total available roof area in each commune, km^2.

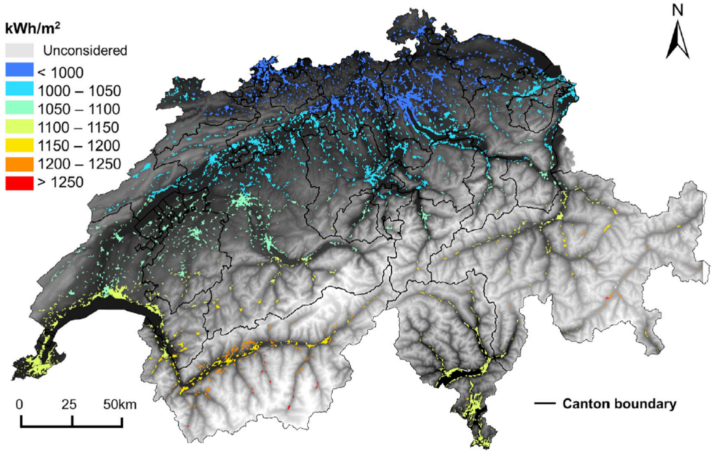


Fig. 6. Annual tilted solar radiation for one specific slope and azimuth configuration (slope = 35° and azimuth = 10°) out 63 configuration.

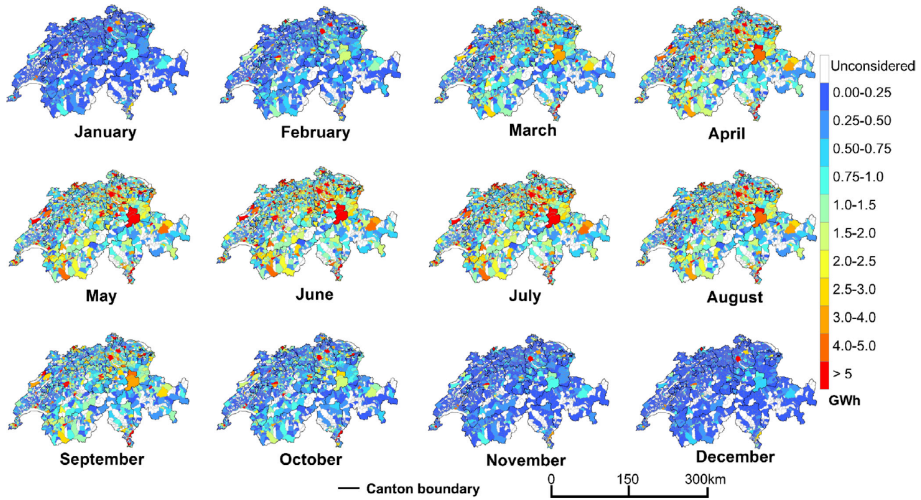
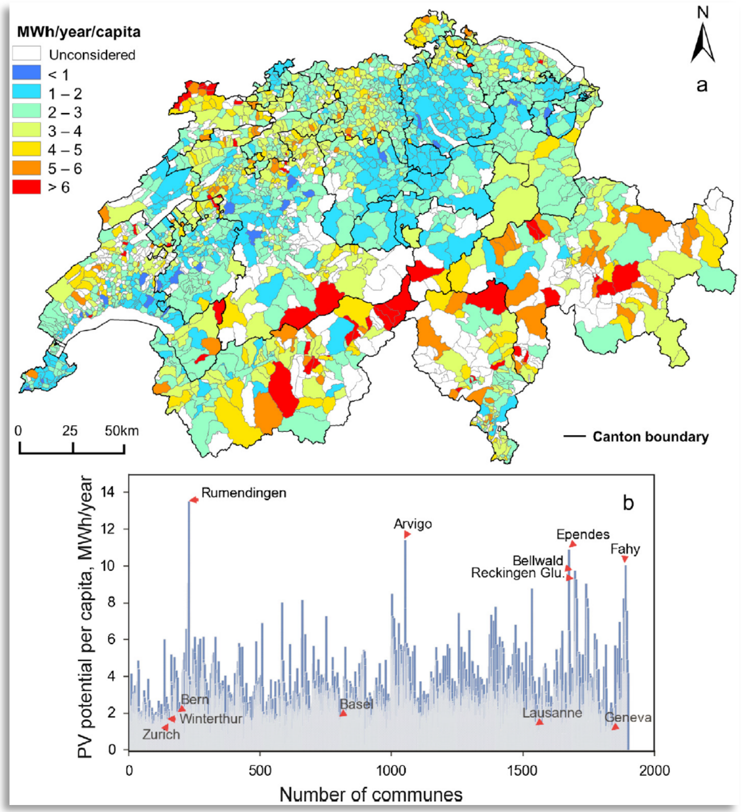
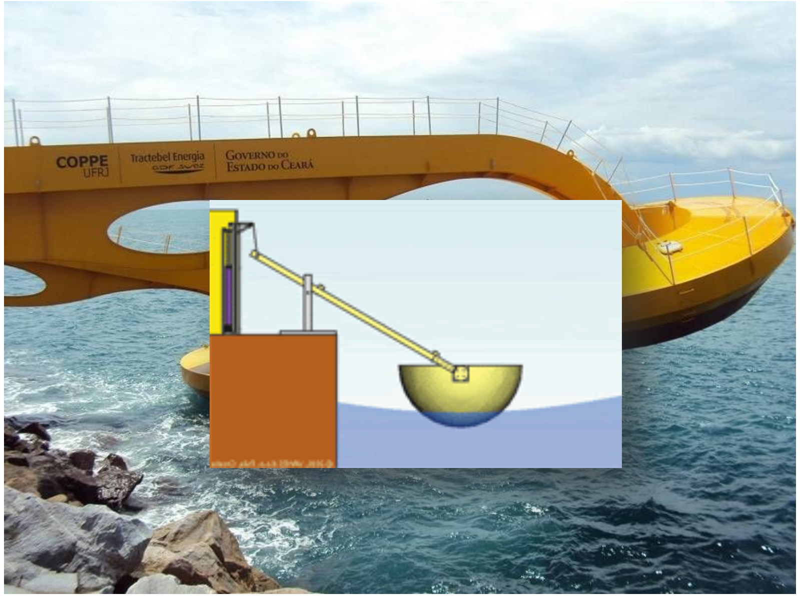


Fig. 7. Technical potential of rooftop PV electricity production for each commune in Switzerland, GW h/month.



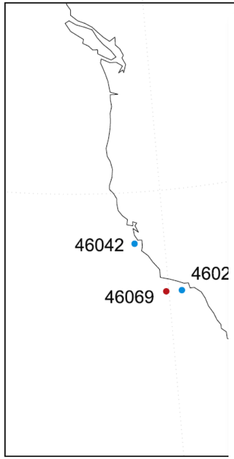
SIGNIFICANT WAVE HEIGHT AND ENERGY FLUX PREDICTION FOR MARINE ENERGY APPLICATIONS: A GROUPING GENETIC ALGORITHM – EXTREME LEARNING MACHINE APPROACH by Cornejo-Bueno, L.; Nieto-Borge, J. C.; García-Díaz, P.; Rodríguez, G.; Salcedo-Sanz, S., 2016

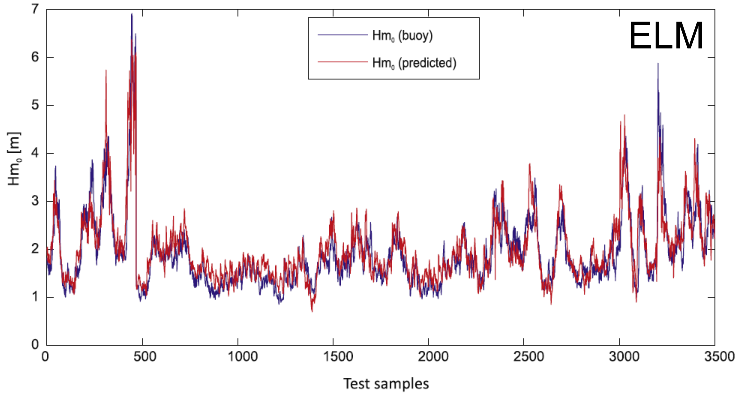


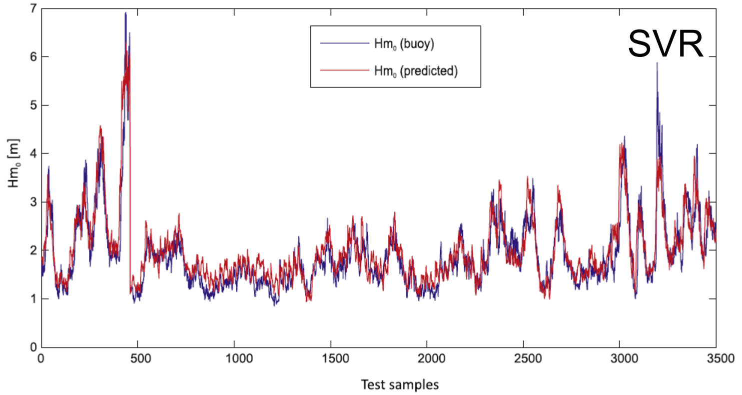
KINNECT ENERGY is converted to electric energy from wind-generated waves with the Wave Energy Converter (WEC)

MEASURING 2 important parameters: wave height (H_{m_0}) and wave energy flux (P)

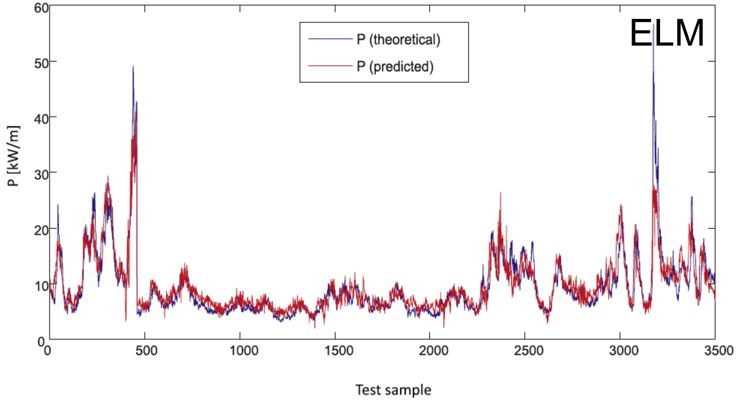
FEATURE SELECTION with GGA and predictions with ELM



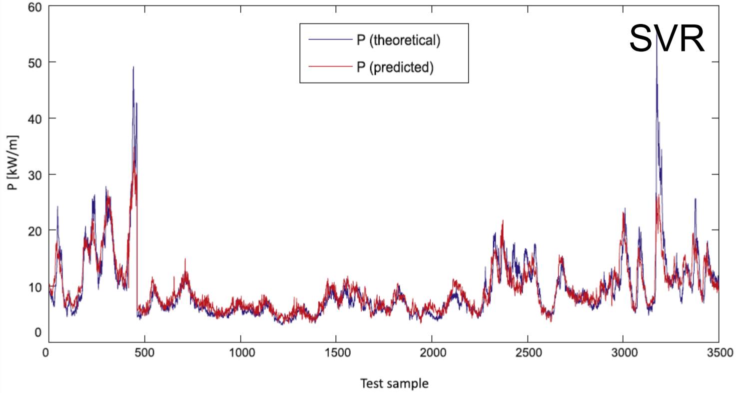




Experiments	RMSE	MAE	r^2
All features-ELM	0.4653 m	0.3582 m	0.6624
All features-SVR	0.6519 m	0.4986 m	0.3949
GGA-ELM	0.4351 m	0.3118 m	0.6742
GGA-SVR	0.3923 m	0.2889 m	0.7472



(a)



Experiments	RMSE	MAE	r^2
All features-ELM	3.4183 kW/m	2.4265 kW/m	0.6243
All features-SVR	4.4419 kW/m	2.8993 kW/m	0.3129
GGA-ELM	2.8739 kW/m	1.8715 kW/m	0.7101
GGA-SVR	2.6626 kW/m	1.6941 kW/m	0.7548

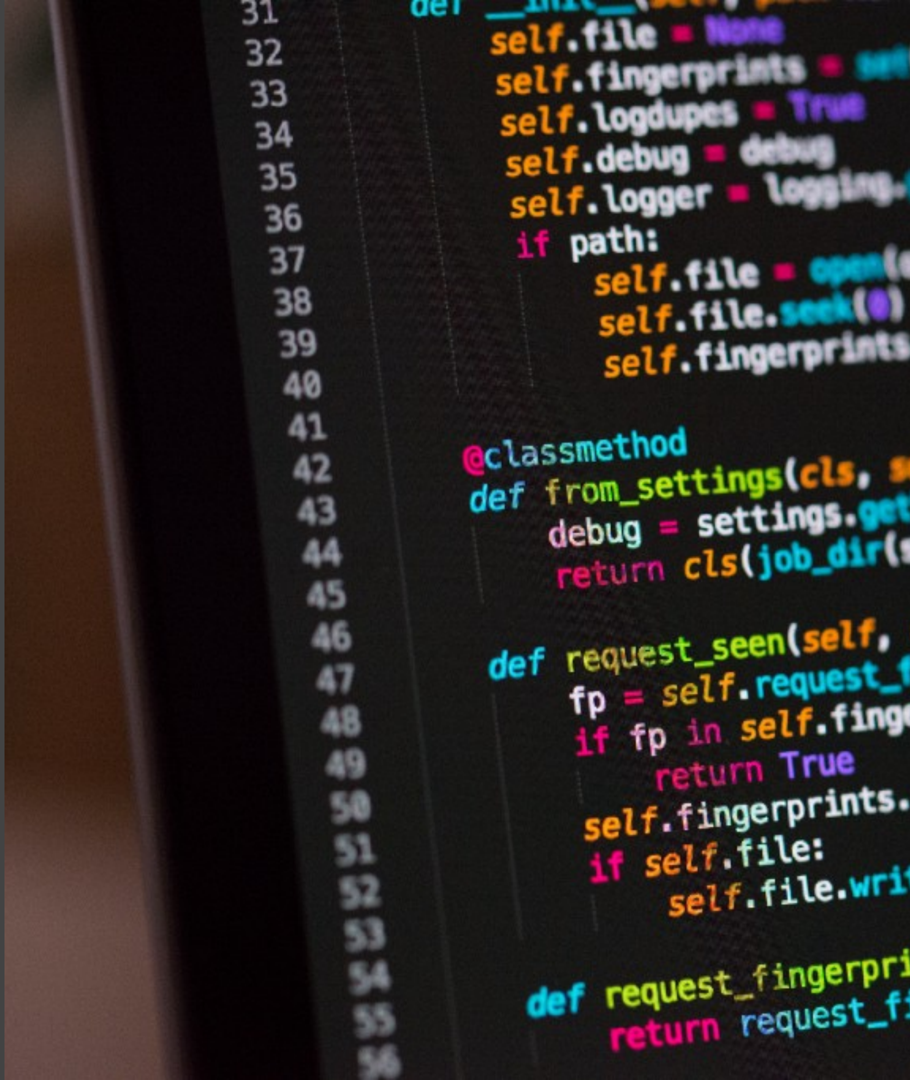
RESOURCES – PART 03

- Demolli, H.; Dokuz, A. S.; Ecemis, A.; Gokcek, M. Wind Power Forecasting Based On Daily Wind Speed Data Using Machine Learning Algorithms. *Energy Conversion and Management,* **2019**, *198*. https://doi.org/10.1016/j.enconman.2019.111823
- Assouline, D.; Mohajeri, N.; Scartezzini, J. L. Quantifying Rooftop Photovoltaic Solar Energy Potential: A Machine Learning Approach. *Solar Energy,* **2017**, *141*, 278-296. https://doi.org/10.1016/j.solener.2016.11.045
- Cornejo-Bueno, L.; Nieto-Borge, J. C.; García-Díaz, P.; Rodríguez, G.; Salcedo-Sanz, S. Significant Wave Height And Energy Flux Prediction For Marine Energy Applications: A Grouping Genetic Algorithm – Extreme Learning Machine Approach. *Renewable Energy*, **2016**, *97*, 380-389. https://doi.org/10.1016/j.renene.2016.05.094

04

LEARNING SESSIONS 2021

What expect for this year



PLANS FOR 2021

Still bi-weekly

“Reproduce” the methodology of a paper (hands-on)

Focus in Python

CODES and all the material will be available at the CREWES website (1-2 days after the lab)


Thank you!