

Dynamic Time Scan Forecasting

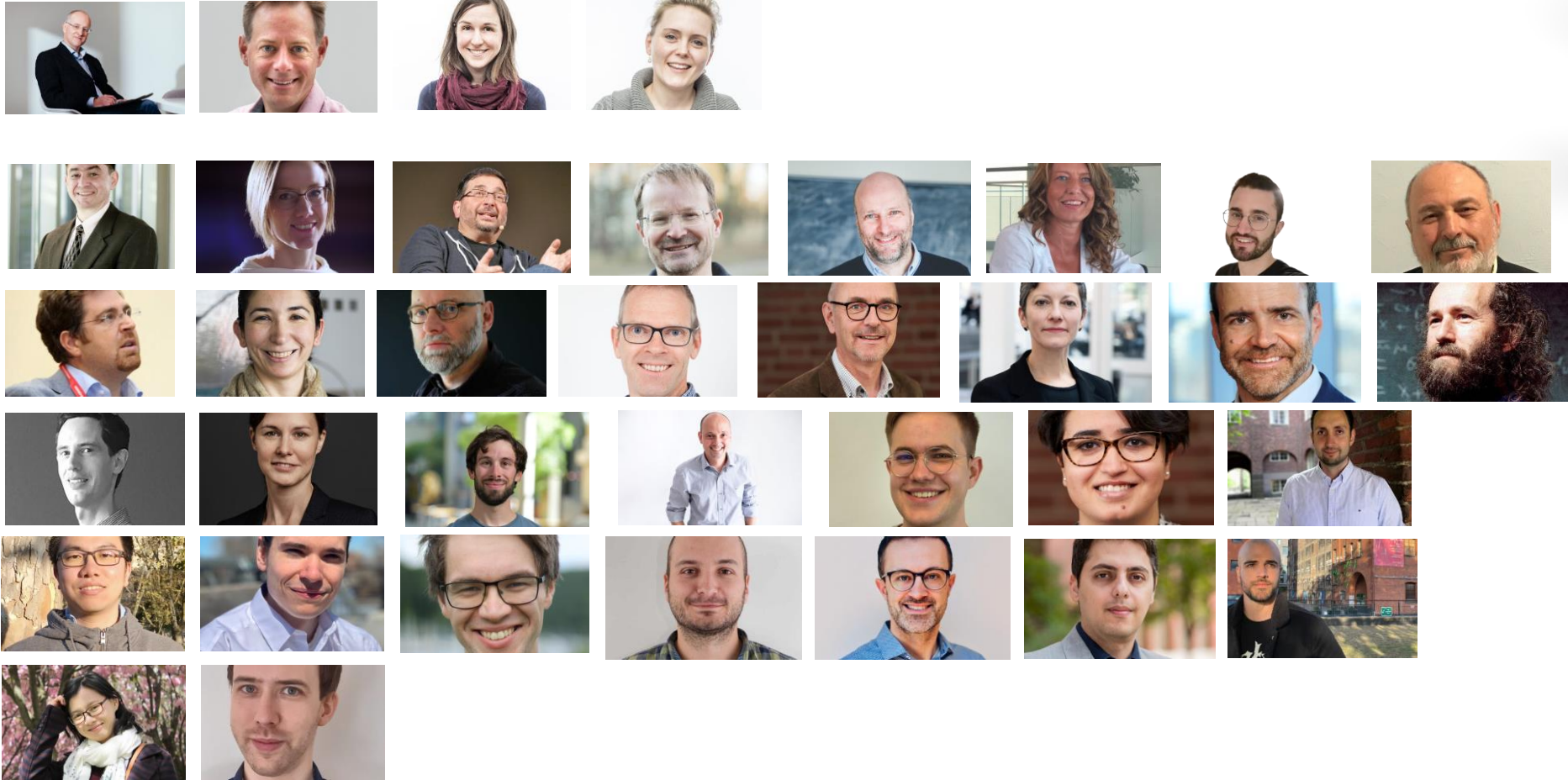
Marcelo Azevedo Costa - Production Engineer/UFMG
(macosta@ufmg.br)

Ramiro Ruiz-Cárdenas - Consultant

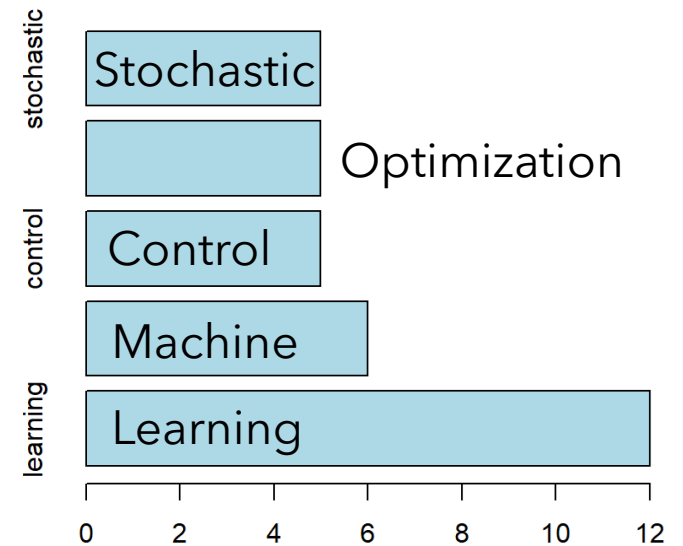
Leandro Brioschi Mineti - Falconi Consultants for Results

Marcos Oliveira Prates - Statistics/UFMG

Just Another Time Series Model ?



Wordcloud analysis using the presentation titles:



Learning (list of the titles)

- Merging optimization and **machine learning**
- Language-based representation learning for acting and planning
- **Machine learning** to accelerate solving of constraint programs
- Primal heuristics for mixed-integer programming through a **machine learning** lens
- A hybrid approach to safe learning in automatic control
- Learning beam search: utilizing **machine learning** to guide beam search for solving combinatorial optimization problems
- A fast matheuristic for two-stage stochastic programs through supervised learning
- Seeking transparency in **machine learning** through optimized explanations
- **Machine learning**-supported decomposition algorithms for a large scale hub location problem
- Learning-based model predictive control with applications to autonomous racing and multi-agent coverage control
- Learning probabilistic circuits using stochastic computation graphs
- Reinforcement learning for guiding metaheuristics
- Learning stationary nash equilibrium policies in n-player stochastic games with independent chains

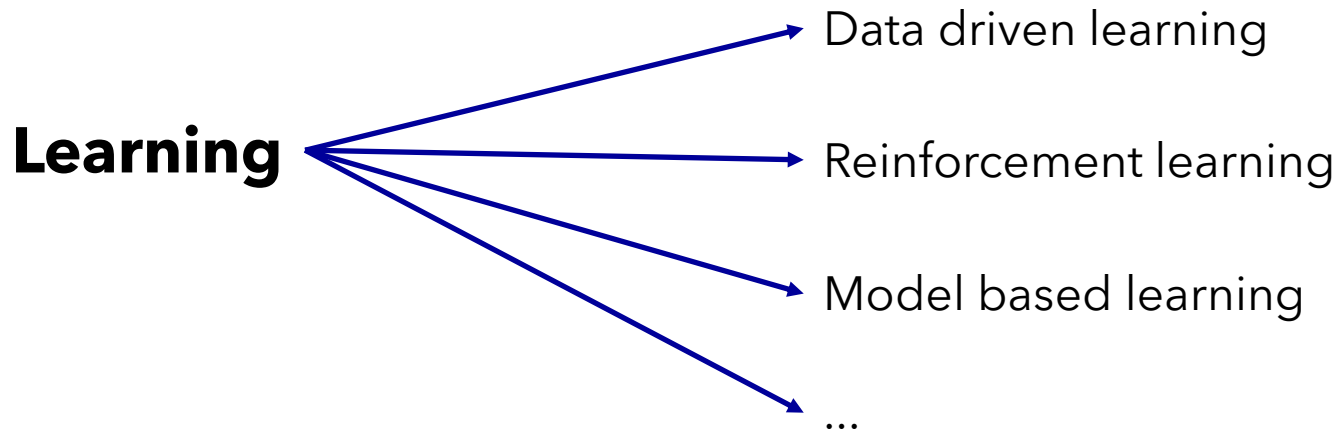
Learning (list of the titles)

- 1) Language-based representation **learning** for acting and planning
- 2) A hybrid approach to safe **learning** in automatic control
- 3) **Learning** beam search: utilizing **machine learning** to guide beam search for solving combinatorial optimization problems
- 4) A fast matheuristic for two-stage stochastic programs through supervised **learning**
- 5) **Learning**-based model predictive control with applications to autonomous racing and multi-agent coverage control
- 6) **Learning** probabilistic circuits using stochastic computation graphs
- 7) Reinforcement **learning** for guiding metaheuristics
- 8) **Learning** stationary nash equilibrium policies in n-player stochastic games with independent chains

What is Learning?



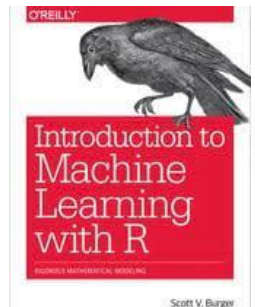
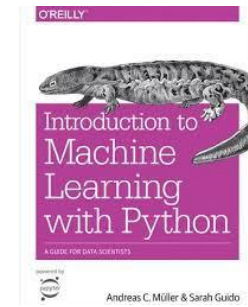
- The acquisition of knowledge or skills through study, experience, or being taught (Oxford Dictionary)



Hybrid AI - Where **data-driven and model-based** methods meet

Hybrid AI - Where **LEARNING** methods meet

What about **Machine Learning**? Data-driven methods?



How many different **Machine Learning** methods exist?

Regarding Machine Learning Models:

✓ What is your **favorite** Machine Learning model?

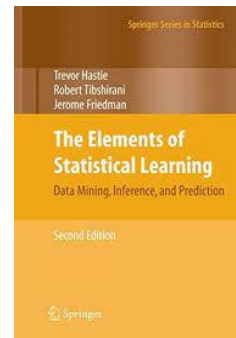
✓ What is your **second** favorite Machine Learning model?

1. Linear (neural networks)

227. xyf (Self-Organizing Maps)

Appendix A: Encyclopedia of Machine Learning Models in **caret**

Statistical Elements of **Machine Learning**



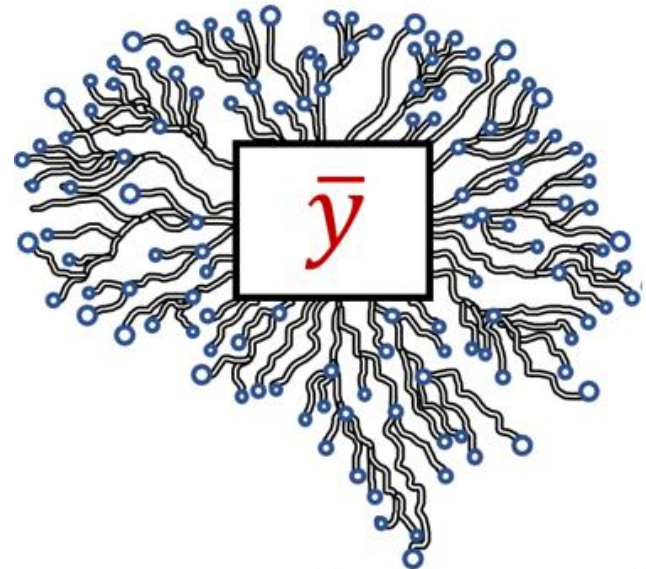
Statistical Decision Theory

$$f(x) = E(Y|X = x)$$

The best prediction of Y at any point $X=x$ is the conditional mean (pg. 18)

Since there is typically at most one observation at any point x , we settle for:

$$\hat{f}(x) = \frac{1}{N_{k(x)}} \sum_{N_{k(x)}} y_i | x$$

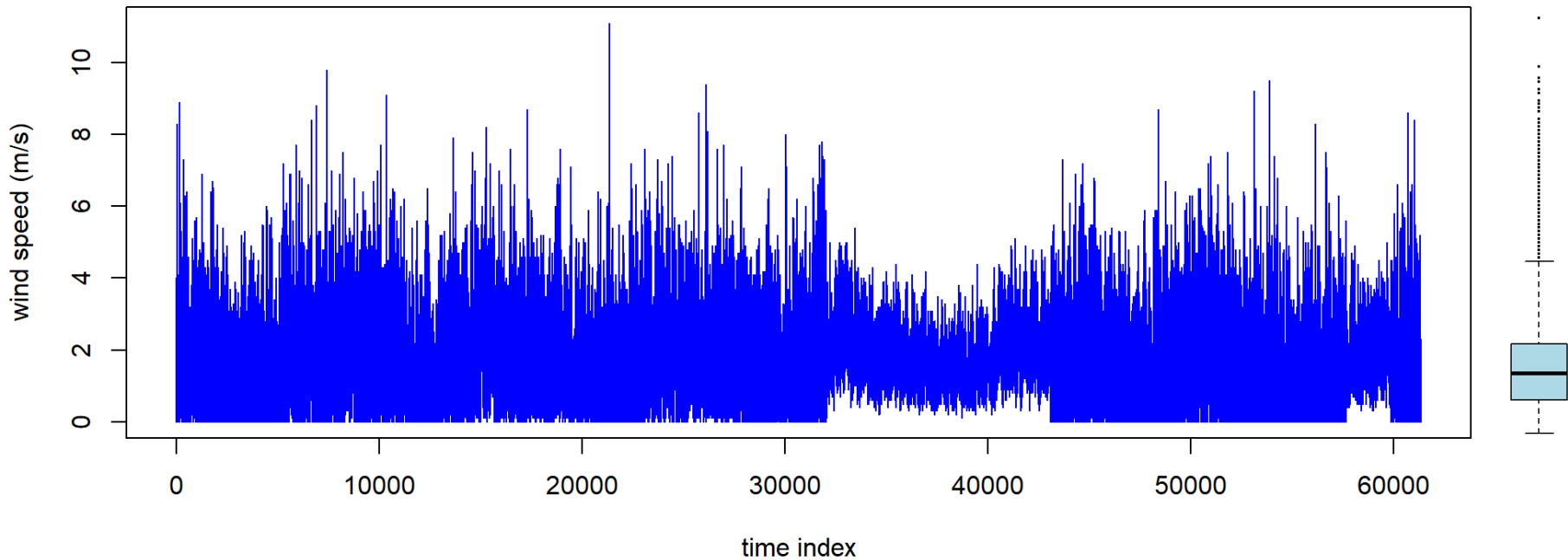


The Wind Speed Time Series Case Study



Wind speed data from January 1, 2009 to December 31, 2015 at every 30 minutes (**61.341** observations).

Final goal: one-day-ahead prediction, i.e., 48 Steps ahead



Time Series Forecasting Literature Review

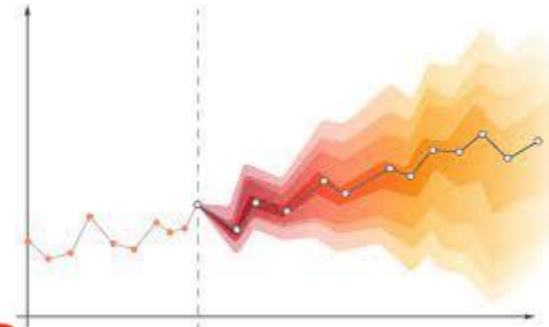
Makridakis Competitions (M-competition)

by Spyros Makridakis

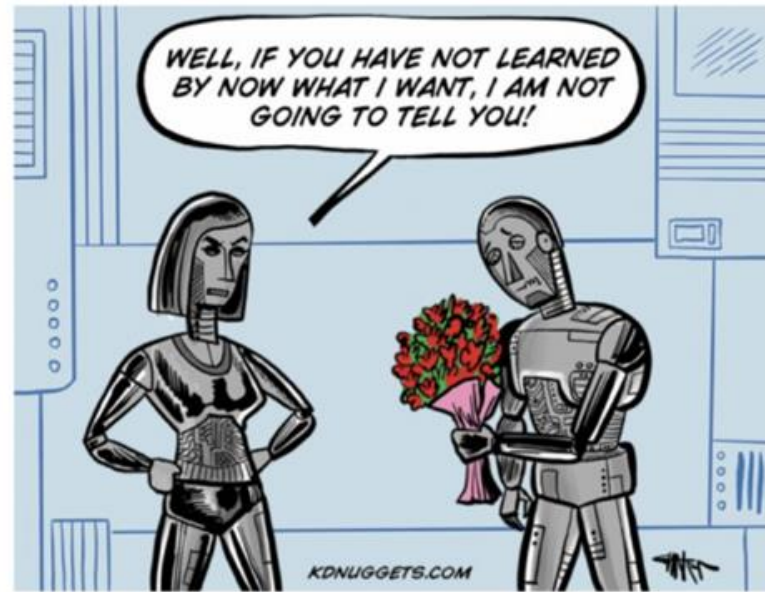
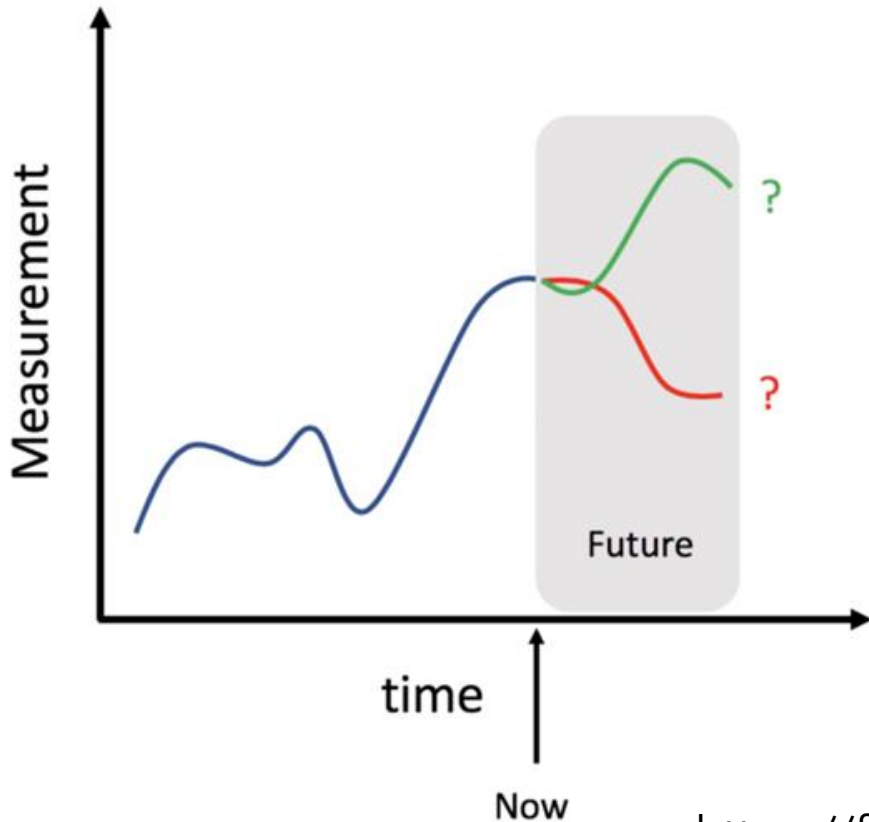
- **M-Competition (1982)**
- **M2-Competition (1993)**
 - The M2-Competition - A real-time judgmentally based forecasting study (International Journal of Forecasting)
- **M3-Competition (2000)**
 - **The M3-Competition: results, conclusions and implications** (International Journal of Forecasting)
- **M4-Competition (2020)**
 - The M4 Competition: 100,000 time series and 61 forecasting methods (International Journal of Forecasting)
- **M5-Competition (2021)**
 - **M5 accuracy competition: Results, findings, and conclusions** (International Journal of Forecasting)
- **M6-Competition (2022-2024)**

<https://forecasters.org/resources/time-series-data/>

Main findings



Machine Learning Time Series Problem



Main findings

- M5-Competition (2021)
 - The M5 “Accuracy” competition clearly showed that **ML methods have entered the mainstream of forecasting applications**, at least in the area of retail sales forecasting.
 - **From a practical perspective, it is necessary to determine the extra costs incurred to run ML methods** versus the standard statistical methods, and whether their accuracy improvements would justify higher costs.



Some Machine Learning results

300

Multi-Layer-Perceptron
(Neural Network)

Do larger data sets require more complex methods?

Not necessarily!

Selected methods for Wind Speed (Time Series) Forecasting based on Literature Review

- **Näive method**

- replicate the observed wind speed in the previous day, i.e., the last 48 observations, as the forecast values.

- **Time series based approaches**

- TBATS (Exponential smoothing state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components) model
- ARIMA (Autoregressive Integrated Moving Average) model

- **Hybrid approaches**

- **Machine learning approaches**

- **Analog-based approaches**

- PSF (Pattern Sequence-Based Forecasting) algorithm
- AnEn (weather analogs ensemble) method
- **Dynamic Time Scan Forecasting (Renewable Energy, 2021)**

Scan Statistics



- ✓ Clustering of random points in two dimensions. **Biometrika** 52 (1965), 263-267.
- ✓ Kulldorff M. A spatial scan statistic. **Communications in Statistics: Theory and Methods**, 1997; 26:1481-1496.

Clarivate English ▾ Products

Web of Science™ Search Sign In ▾ Register

Search > Results for scan statistics (All Fields)

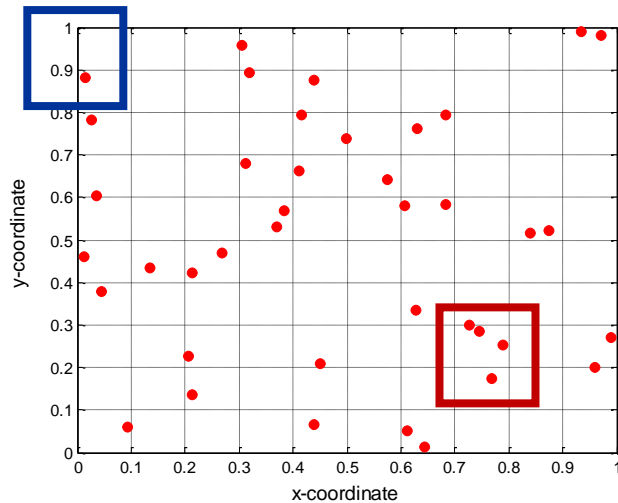
16,390 results from Web of Science Core Collection for:

scan statistics (All Fields) Analyze Results Citation Report Create Alert

November 9, 2022

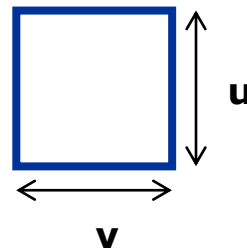
Clustering of random points in two dimensions

by J. I. Naus



Objective: (anomaly detection) to obtain the upper and lower bounds of the probability of finding at least one cluster of dimensions \mathbf{v} and \mathbf{u} containing at least \mathbf{n} points,

$$P(n | N, u, v).$$

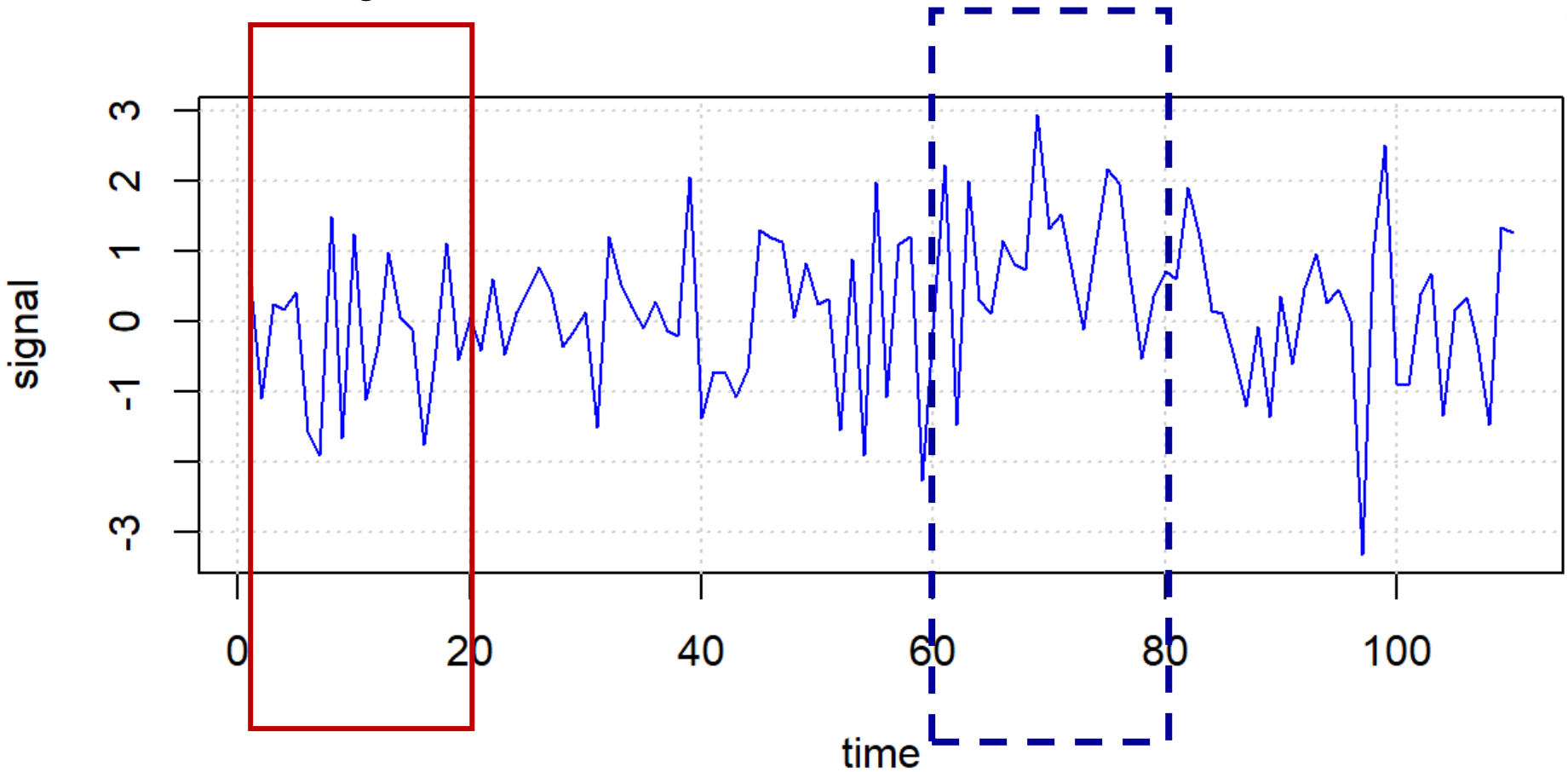


Clustering of random points in two dimensions. *Biometrika* **52** (1965), 263-267.

Scan Statistics in Time Series



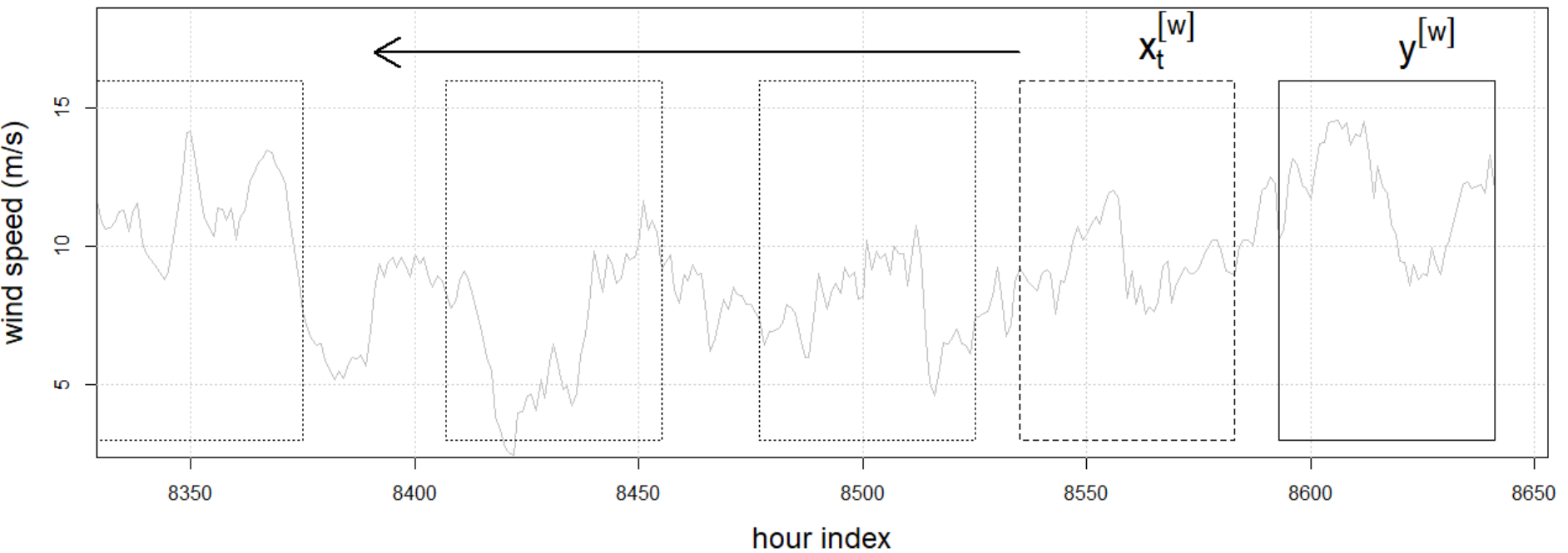
Scanning window



$$\hat{\mu}_{inside} \neq \hat{\mu}_{outside}$$

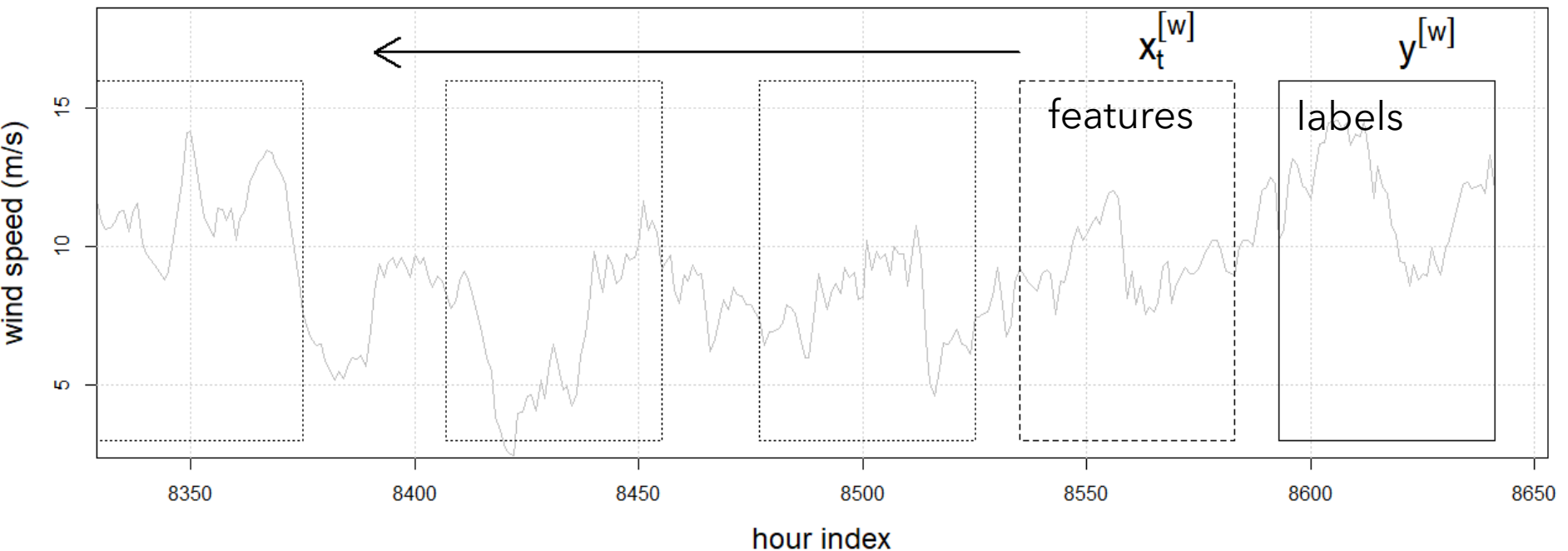
Dynamic Time Scanning Process

Scanning process



Dynamic Time Scanning Process

Scanning process

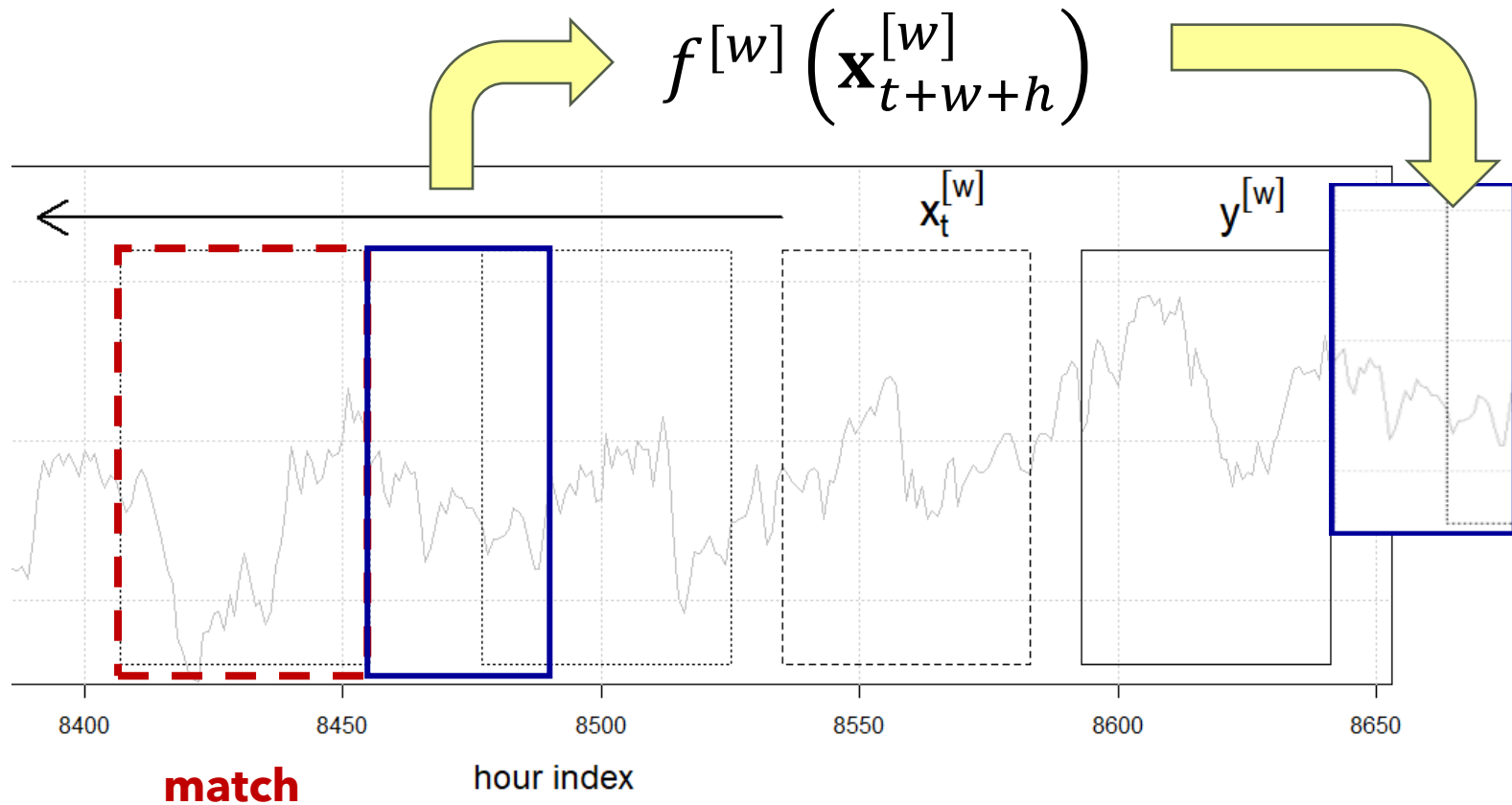


Similarity function:

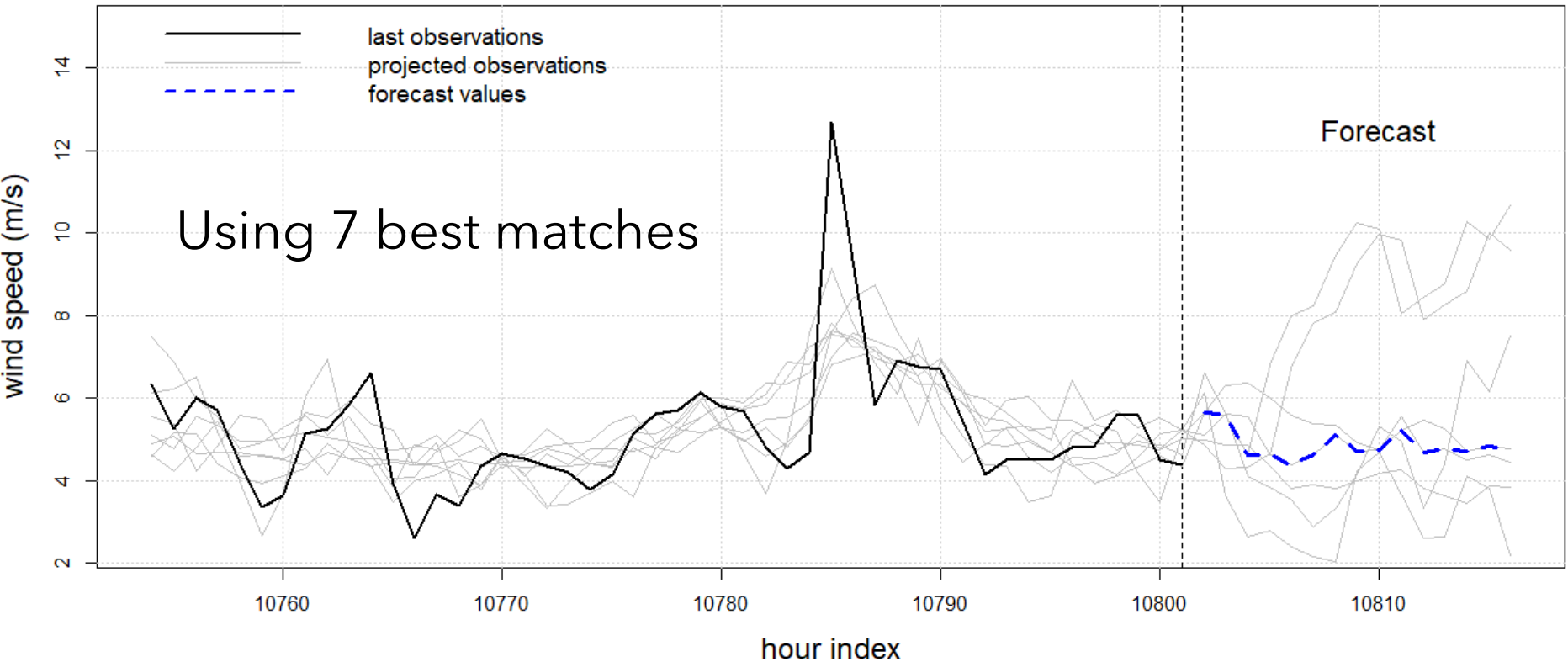
$$y^{[w]} = \beta_0^{[w]} + \beta_1^{[w]} \mathbf{x}_t^{[w]}$$

$$y^{[w]} = f(\mathbf{x}_t^{[w]})$$

Dynamic Time Scan Forecasting

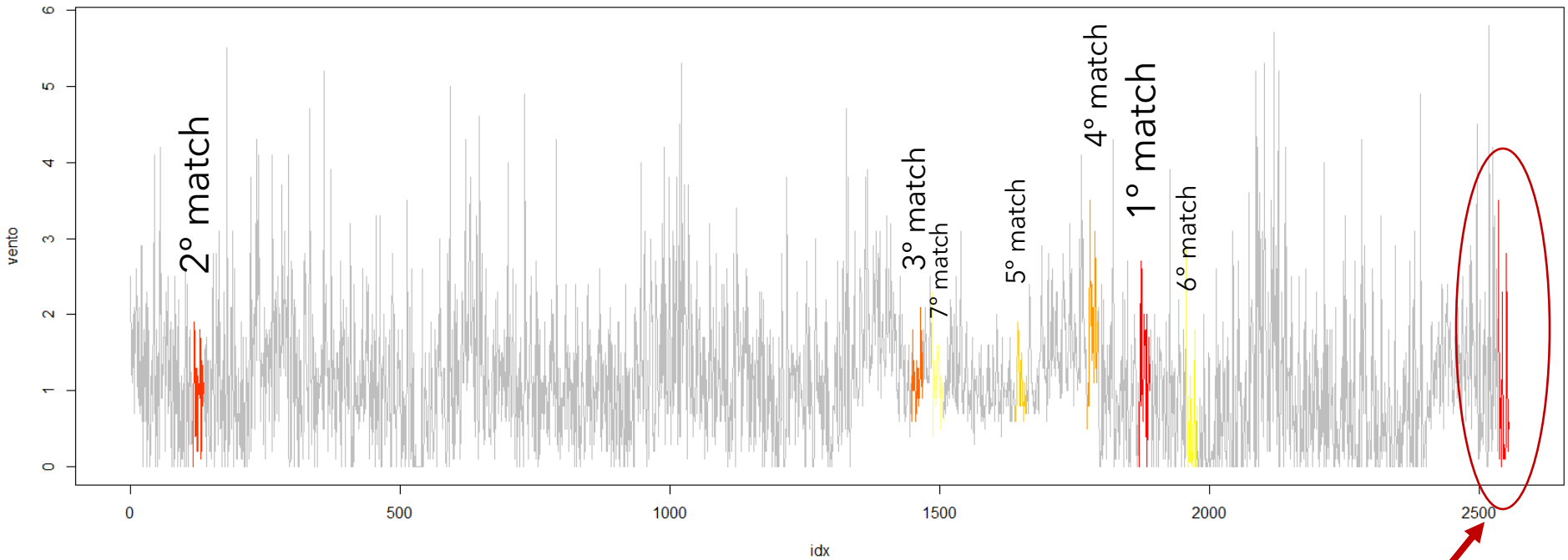


Dynamic Time Scan Forecasting



The **median** function is used to create the final point forecasts to minimize extreme values

Case study



Required parameters:

window size = 20

best matches = 7

k.prediction = 48

(forecast steps)

Last
observations

Parameter tuning

1) Window size:

- 18, 24 e 36 days

2) $f(\mathbf{x})$: Similarity function:

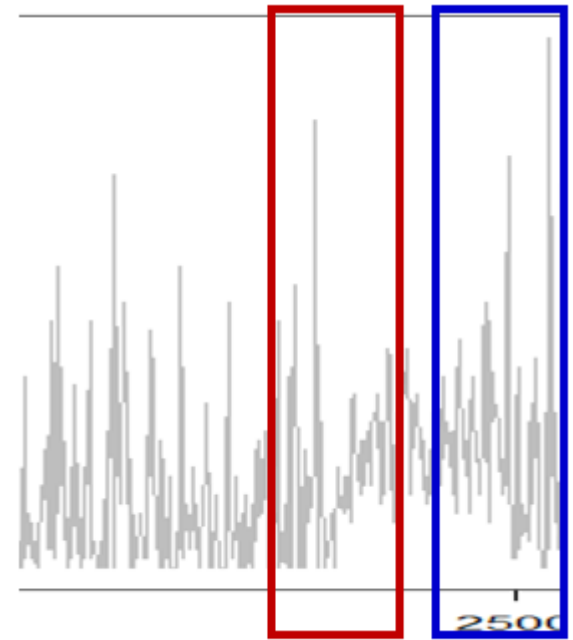
$$y = \beta_0 + \beta_1 x \quad (\text{linear})$$

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$$

3) Best matches

- 5, 12, 24



Ensemble version: combining the n elements of a grid of parameters with best forecast performance in the previous day.

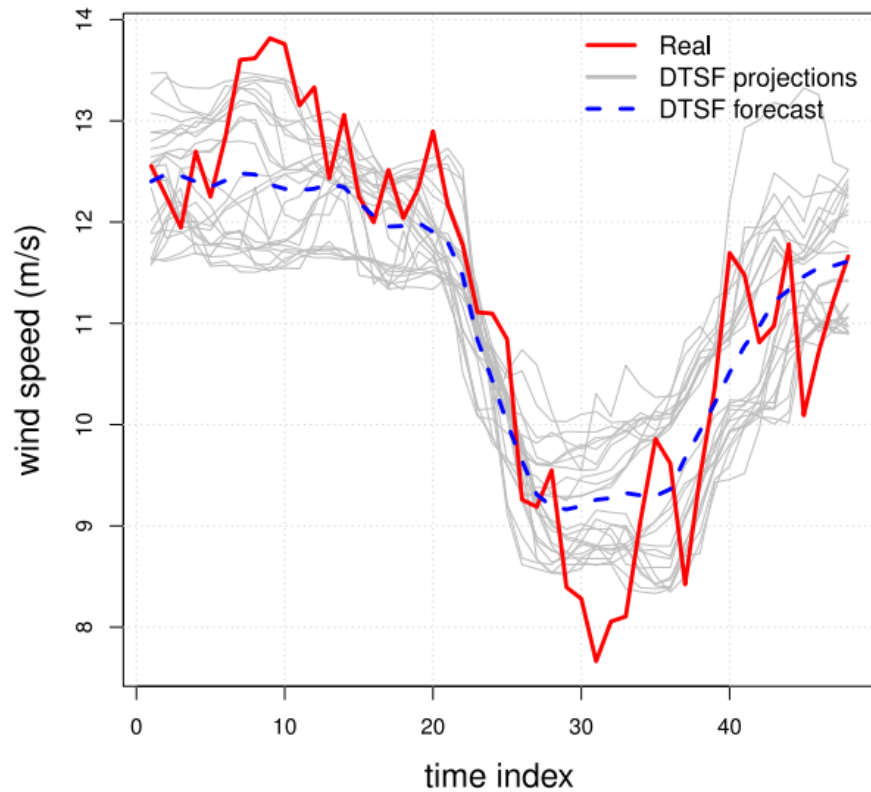
scanning window: 96, 192, 288, 384 and 480

best matches : 5, 10, 30, 50, 70 and 90

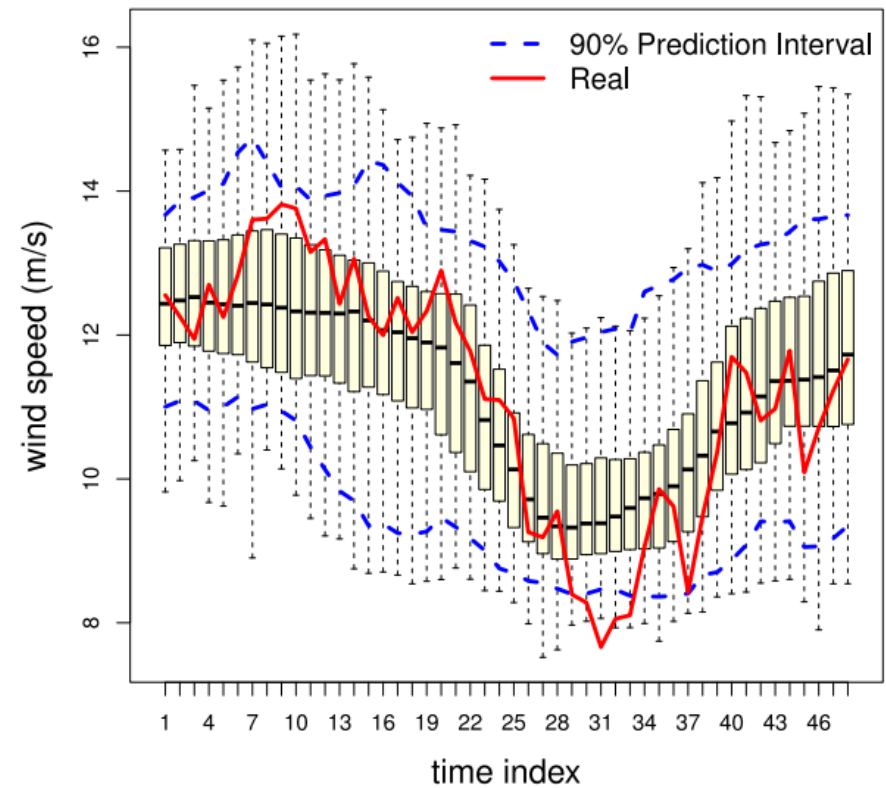
Similarity functions of degrees: 1, 2 and 3

} 120 different combinations of the three parameters.

Ensemble Dynamic Time Scan Forecasting



(a) Projected values.



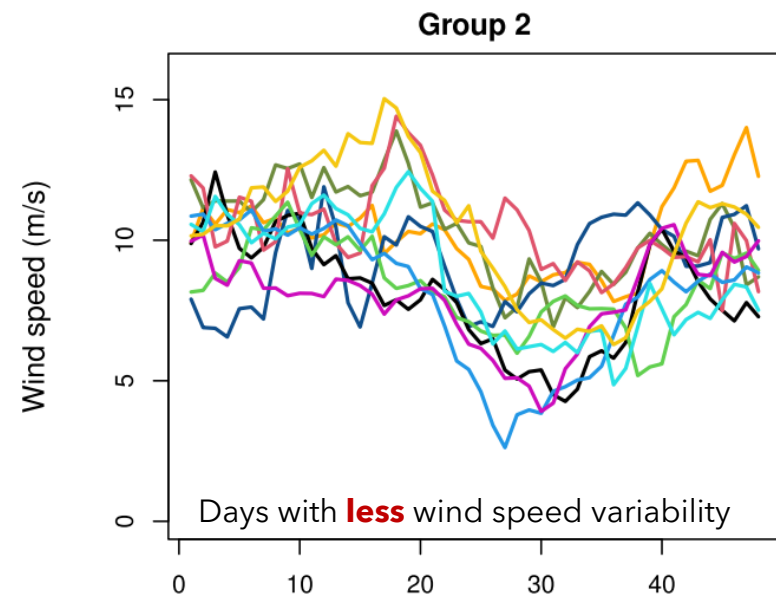
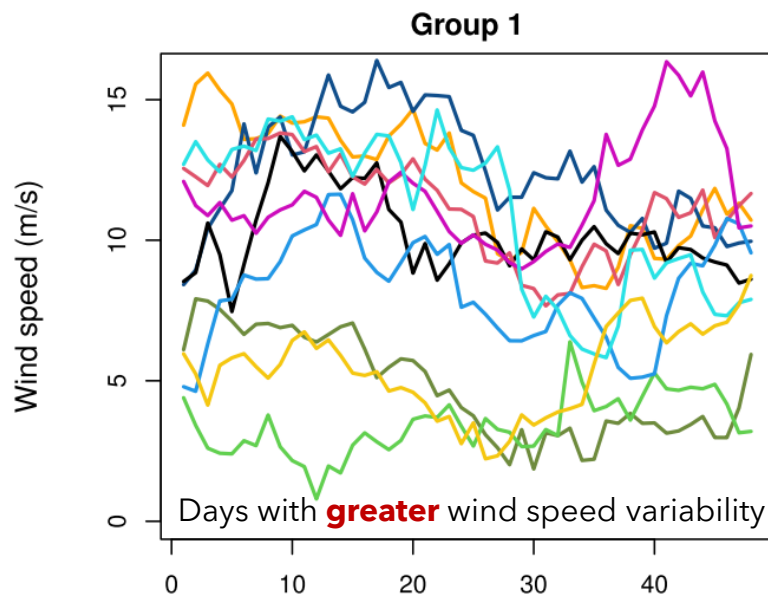
(b) Empirical prediction interval.

Data Validation

November 21, 2011 to June 22, 2016 (241,200 observations)

Winter	Spring	Summer	Autumn
2015-06-28	2015-09-30	2016-01-31	2016-04-03
2015-07-04	2015-10-01	2016-02-05	2016-04-05
2015-08-08	2015-10-26	2016-02-24	2016-04-12
2015-08-11	2015-12-02	2016-02-25	2016-04-13
2015-09-18	2015-12-06	2016-03-13	2016-05-18

Divided into two groups based on the variability of the wind speed through its diurnal cycle, as influenced by the prevalent turbulence intensity

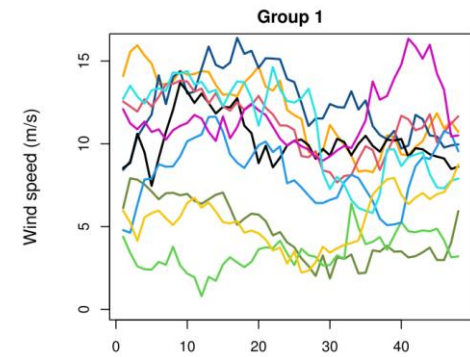


Forecast objective: **48 steps ahead - next 24 hours.**

Results

(using different error statistics)

Days with **greater** wind speed variability

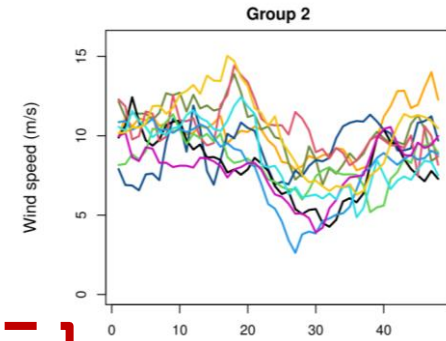


Method	MAE	RMSE	sMAPE	MAPE	MF	AvgRelMAE
naïve	2.36	2.79	0.31	31.45	0.0032	1.000
ARIMA	2.27	2.60	0.31	34.00	0.0033	0.989
TBATS	2.04	2.39	0.28	29.47	0.0025	0.935
NNET.1(*)	1.96	2.34	0.27	28.25	0.0028	0.903
NNET.2	2.04	2.39	0.28	26.59	0.0023	0.900
STL+ETS	2.10	2.41	0.29	26.52	0.0022	0.913
hybrid.1	2.16	2.52	0.27	30.32	0.0021	1.007
hybrid.2(*)	1.89	2.22	0.25	25.80	0.0018	0.893
PSF	2.87	3.26	0.38	44.70	0.0047	1.149
AnEn	3.00	3.35	0.38	52.58	0.0096	1.145
forecAn	1.91	2.28	0.27	26.37	0.0021	0.869
DTSF	1.72	2.07	0.23	26.84	0.0021	0.871
eDTSF	1.89	2.27	0.25	26.98	0.0020	0.891

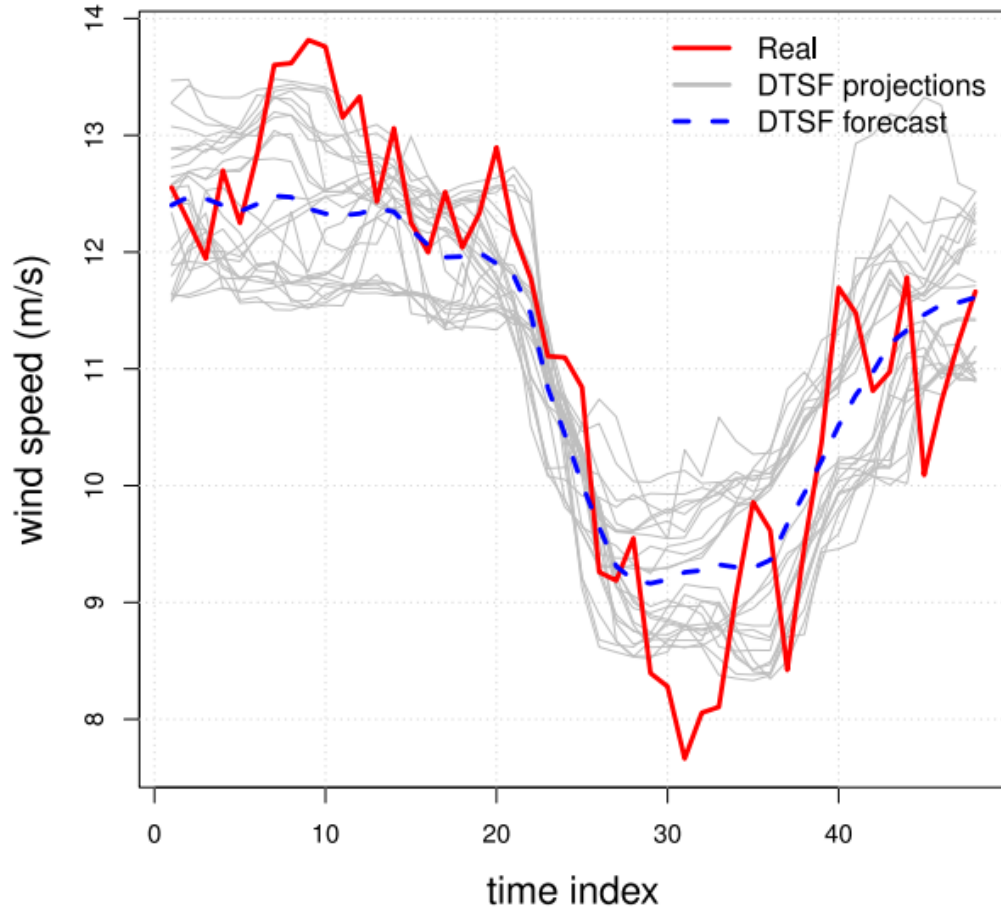
Results

(using different error statistics)

Days with **less** wind speed variability



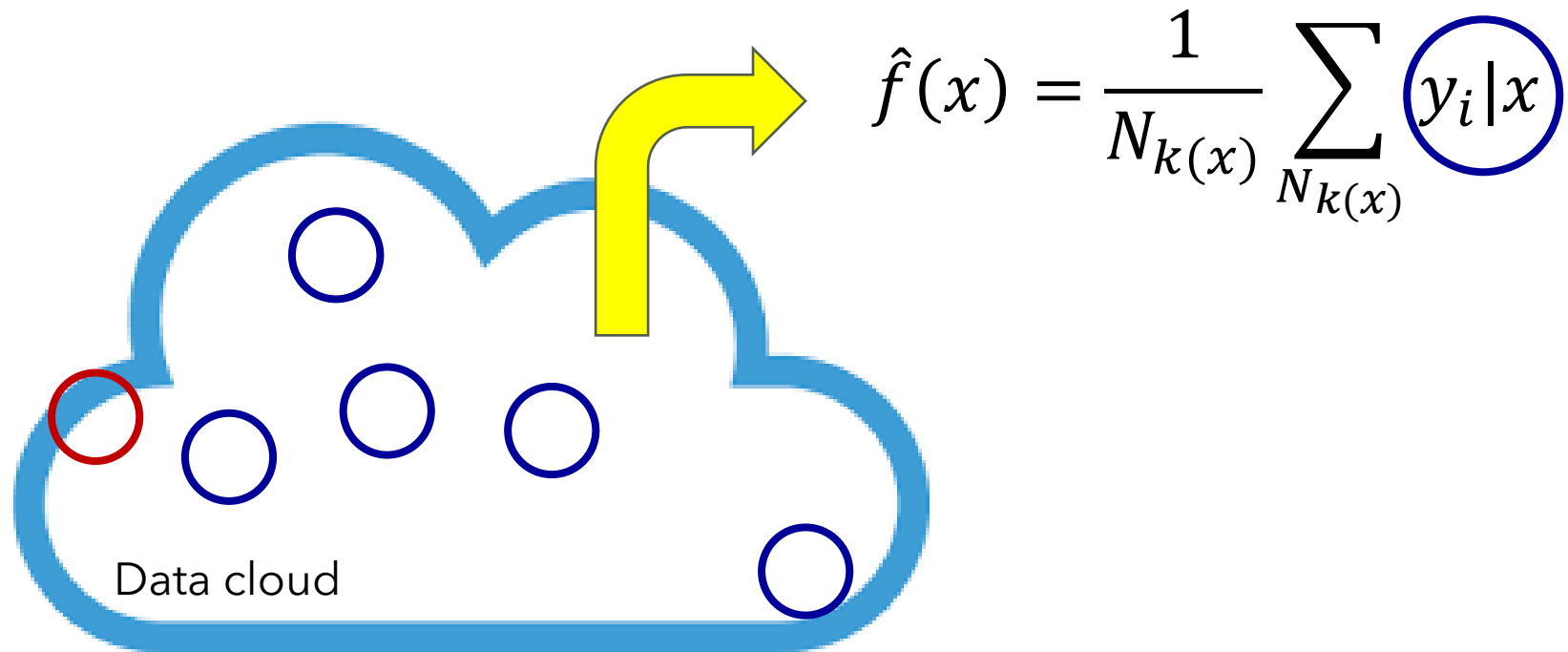
Method
naïve
ARIMA
TBATS
NNET.1(*)
NNET.2
STL+ETS
hybrid.1
hybrid.2(*)
PSF
AnEn
forecAn
DTSF
eDTSF



vgRelMAE
1.000
0.878
0.821
0.909
0.875
0.890
1.016
0.875
0.928
0.908
0.866
0.931
0.791

(a) Projected values.

Visual conclusion



Scanning data may provide a simpler and effective **Machine Learning** solution!

<https://github.com/leandrominetti/DTScanF>

By the way...

M5-competition



“Before presenting the five winning methods, we note that most of the methods utilized **LightGBM**, which is a ML algorithm for performing nonlinear regression using gradient boosted trees (Ke et al., 2017)”.

“The **winner** used an equal weighted combination (**arithmetic mean**) of various **LightGBM** models”

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions. International Journal of Forecasting.