International Workshop on Machine Learning for Space Weather: Fundamentals, Tools and Future Prospects

The Abdus Salam
International Centre
for Theoretical Physics

Buenos Aires, Argentina, 7th November, 2022 – 11th November, 2022

On performance assessment of machine learning-based GNSS ionospheric delay correction model based on space weather predictors in immediate positioning environment

Renato Filjar,
Laboratory for Spatial Intelligence,
University of Applied Sciences Hrvatsko Zagorje Krapina,
Krapina, CROATIA

- Content

- Machine learning-based model development

- Model performance assessment

- Case study of GNSS ionospheric delay correction model performance assessment

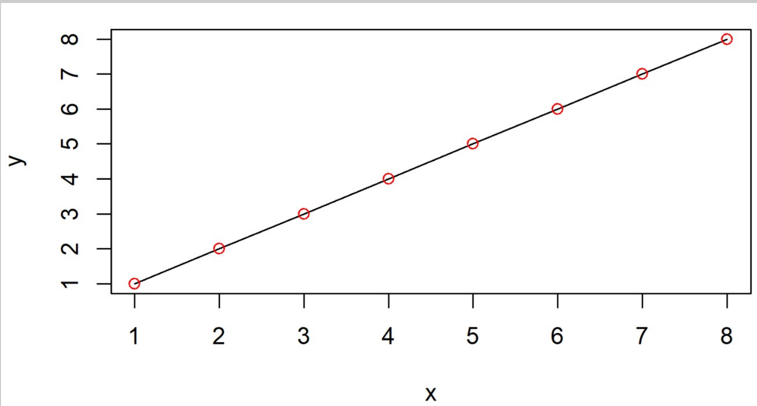
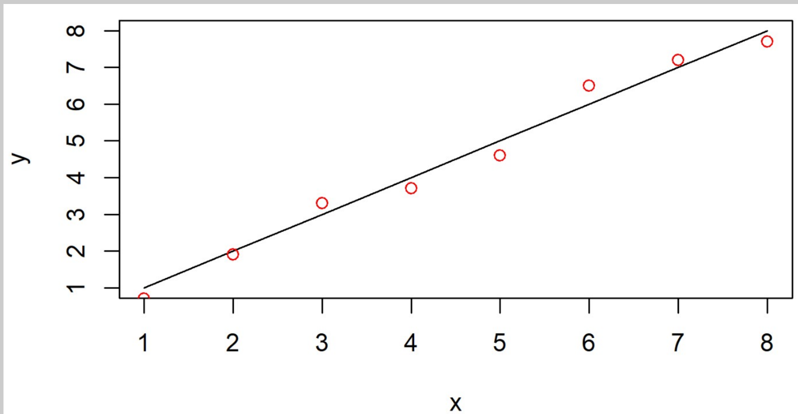
- Discussion

All models are wrong,
but some are useful.

G E P Box

Box, G E P. (1979). Robustness in the strategy of scientific model building (report).

- <u>Rationale and nature of statistical modelling</u>
- A *model* of a *phenomenon* or a system is a main outcome of every research. → description, or prediction

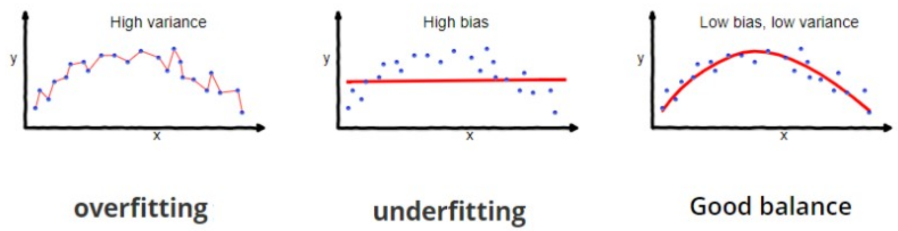
Ideal world (Mathematics)	Real world (Statistics)
$y = ax + b$	$y = ax + b + \epsilon$ $\hat{y} = \hat{a}x + \hat{b}$
	

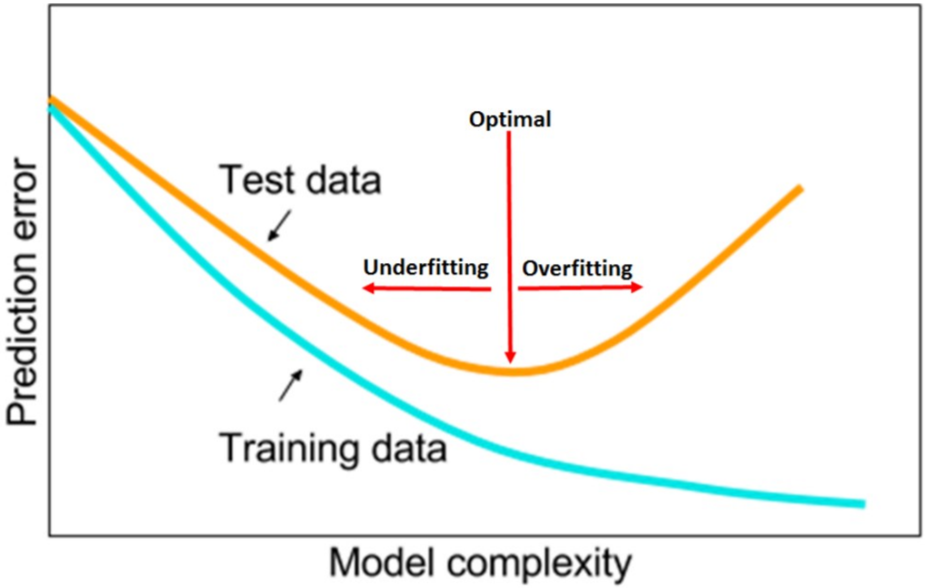
- Machine learning
- A set of statistics-based methods for predictive or descriptive model development based on observations (a sample of population) of a statistical variable, or variables
- Original observations set split randomly into: training, and testing sub-sets (*Pareto principle*, or other)
- Supervised, unsupervised, and reinforcement learning
- Scientific methodology (a common-sense methodology): problem statement, data selection, statistical analysis of data (properties), model structure and model development method definition (hypothesis), model development, model performance assessment → decision on the model to solve the problem in the most efficient manner (optimisation, inference)

International Workshop on Machine Learning for Space Weather: Fundamentals, Tools and Future Prospects
Buenos Aires, Argentina, 7th November, 2022 – 11th November, 2022
On performance assessment of machine learning-based GNSS ionospheric delay correction model based on space weather predictors in immediate positioning environment (R Filjar, *Croatia*)

- <u>Underfitting and overfitting</u>
- Statistics as a tool for detection, identification, and extraction of hidden knowledge in data
- Statistics as the means for '*finding the signal in the noise*'.
- Silver, N. (2015). The Signal and the Noise: Why So ManyPredictions Fail – but Some Don't. Penguin Books. New York, NY.

Source:
https://medium.com/almabetter/bias-variance-tradeoff-c8c2a0fb643e

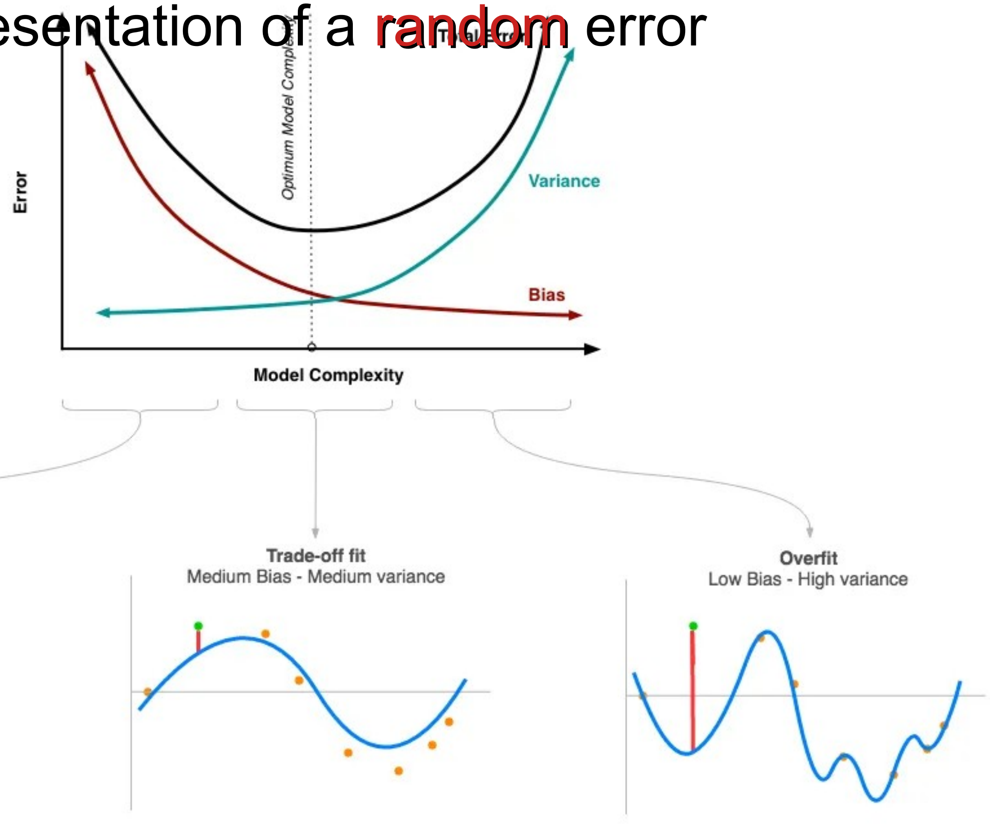




International Workshop on Machine Learning for Space Weather: Fundamentals, Tools and Future Prospects
Buenos Aires, Argentina, 7th November, 2022 – 11th November, 2022
On performance assessment of machine learning-based GNSS ionospheric delay correction model based on space weather predictors in immediate positioning environment (R Filjar, *Croatia*)

- <u>Bias vs Variance – the need for a balance</u>
- Bias as a representation of systematic error
- Variance as a representation of a random error

Source: https://medium.com/almabetter/bias-variance-tradeoff-c8c2a0fb643e

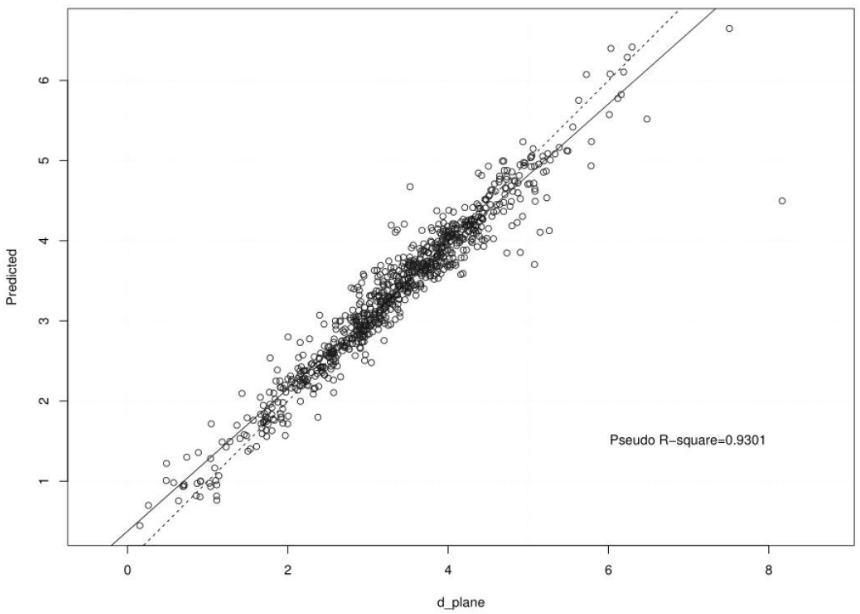


International Workshop on Machine Learning for Space Weather: Fundamentals, Tools and Future Prospects
Buenos Aires, Argentina, 7th November, 2022 – 11th November, 2022
On performance assessment of machine learning-based GNSS ionospheric delay correction model based on space weather predictors in immediate positioning environment (R Filjar, *Croatia*)

- <u>Model performance assessment - Goodness of fit</u>
- Predicted vs observed target values of the testing sub-set

<u>Predicted vs Observed (P-O) diagram</u>

Source: Filić, M, Filjar, R. (2018). ISBN 978-613-9-90118-0.



<u>Statistical tests</u>

Pearson χ^2 test

$$\chi^2 = \frac{\sum_{i=1}^{N} (O_i - E_i)^2}{E_i}$$

O_i denotes observed count of bin *I*

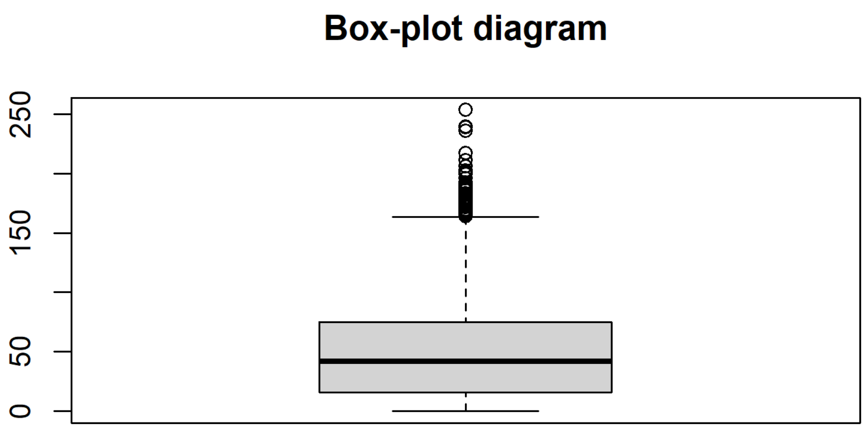
E_i denotes en expected count of bin I, defined as:

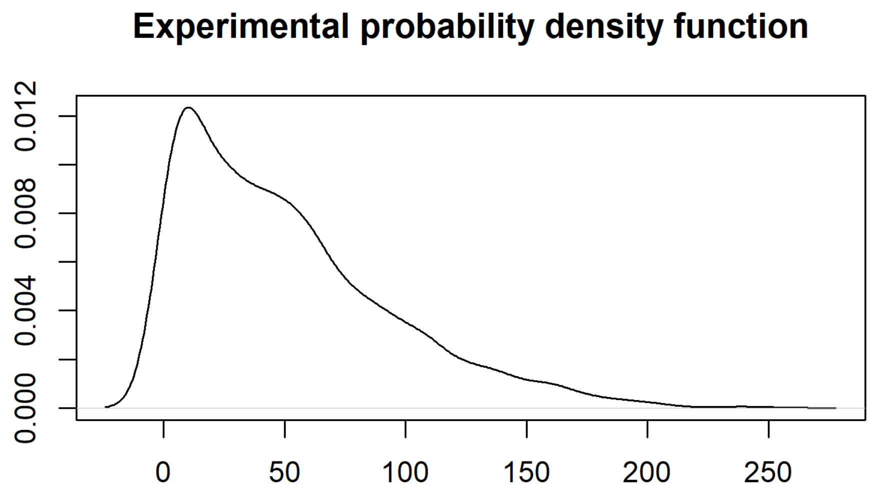
$$E = [F(Y_u) - F(Y_l)] N$$

Determined χ^2 value is then examined for conformance with $\chi 2$ statistical distribution

- Model performance assessment - Statistical distribution of residuals
- *residual = predicted value – observed value*

  (for the same set of predictors values)
- A new statistical variable





- <u>Model performance assessment - Is my model good (enough)?</u>

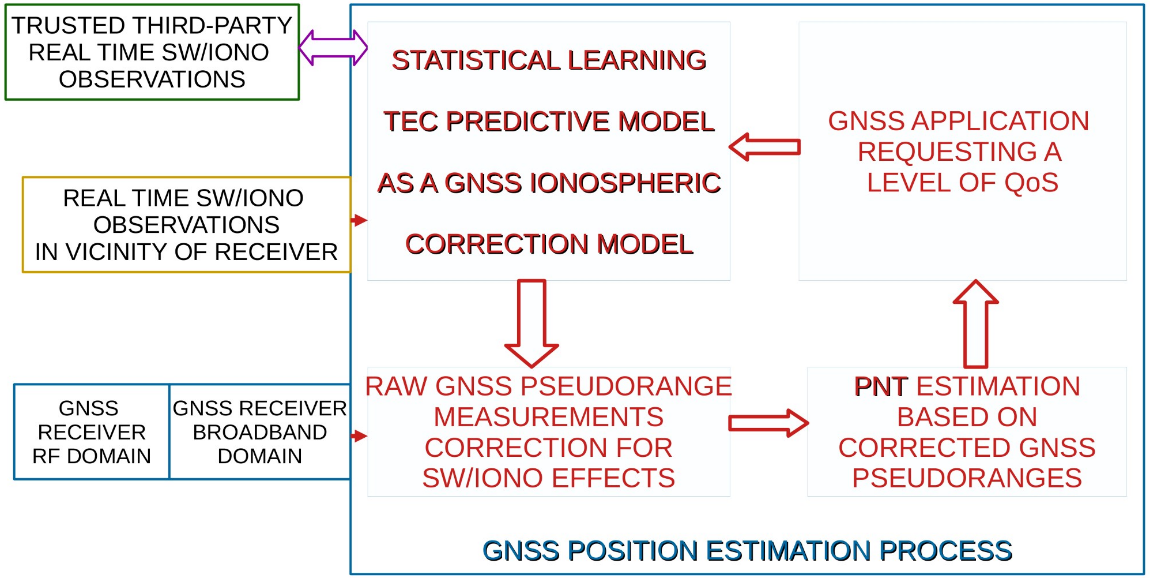
```
Call:
lm(formula = y ~ x2 + x3)

Residuals:
    Min      1Q  Median      3Q     Max 
-4818.1 -2745.4  -639.2  2035.2  7487.4 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -7507.67     695.51 -10.795  < 2e-16 ***
x2             28.04      40.23   0.697    0.487    
x3            223.70      32.73   6.835 7.27e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3243 on 97 degrees of freedom
Multiple R-squared:  0.9426,	Adjusted R-squared:  0.9414 
F-statistic: 796.4 on 2 and 97 DF,  p-value: < 2.2e-16
```

- Model performance assessment - Essential model performance indicators
- *P-O diagram*
- *RMSE*
- *adjR2*
- *statistical significance of predictors*
- *comparison with bechmark/reference model preformance in the same scenario/case study*

International Workshop on Machine Learning for Space Weather: Fundamentals, Tools and Future Prospects
Buenos Aires, Argentina, 7th November, 2022 – 11th November, 2022
On performance assessment of machine learning-based GNSS ionospheric delay correction model based on space weather predictors in immediate positioning environment (R Filjar, *Croatia*)

Case study of GNSS ionospheric delay correction model performance assessment

- Self-adaptive environment-aware Software-Defined Radio (SDR) GNSS position estimation algorithm
- Adaptive GNSS ionospheric delay correction model for short-term rapidly developing geomagnetic storm in sub-equatorial region



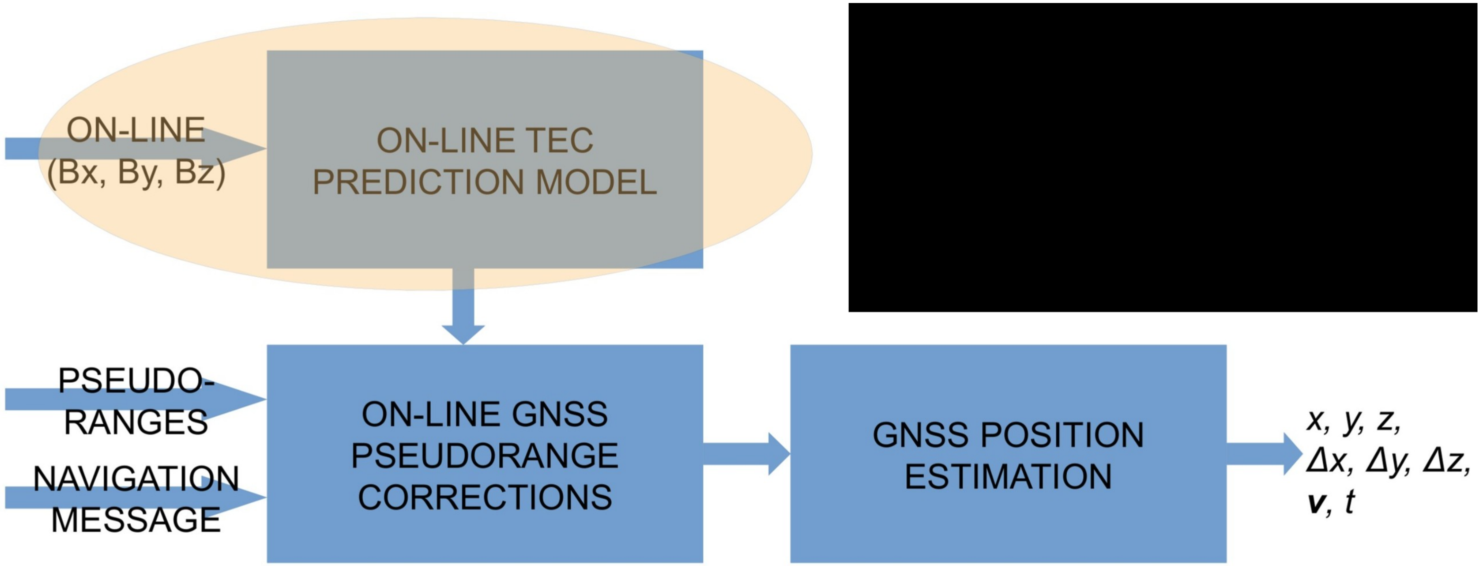
Reference:
Filjar, R. (2022). Statistical learning TEC predictive model for GNSS ionospheric delay mitigation in self-adaptive environment - aware SDR GNSS position estimation algorithm. *The United Nations/Azerbaijan Workshop on the International Space Weather Initiative: The Sun, Space Weather and Geosphere.* Baku, Azerbaijan.
Available at: https://www.unoosa.org/documents/pdf/psa/activities/2022/ISWI2022/s6_01.pdf

International Workshop on Machine Learning for Space Weather: Fundamentals, Tools and Future Prospects
Buenos Aires, Argentina, 7th November, 2022 – 11th November, 2022

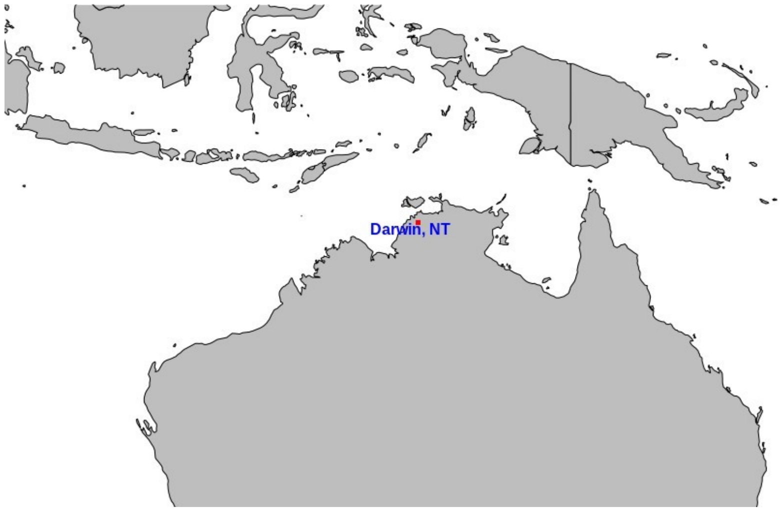
On performance assessment of machine learning-based GNSS ionospheric delay correction model based on space weather predictors in immediate positioning environment (R Filjar, *Croatia*)

Case study reference:
Filjar, R, Weintrit, A, Iliev, T, Malčić, G, Jukić, O, Sikirica, N. (2020). doi: https://doi.org/10.23919/FUSION45008.2020.9190264

- Case study of GNSS ionospheric delay correction model performance assessment
- Positioning environment situation awareness → geomagnetic field observation in the time and the place of positioning
- Self-adaptivness → S/M L-based TEC prediction model



- <u>Case study of GNSS ionospheric delay correction model performance assessment</u>
- Geomagnetic field observations taken at the Intermagnet Kakadu, NT, Australia reference station
- GNSS pseudorange observations taken at the IGS Darwin, NT, Australia reference station
- A short-term rapidly developing geomagnetic storm, identified by its Dst patterns

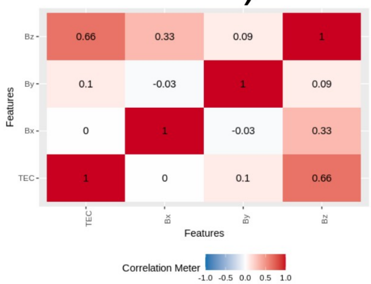


International Workshop on Machine Learning for Space Weather: Fundamentals, Tools and Future Prospects

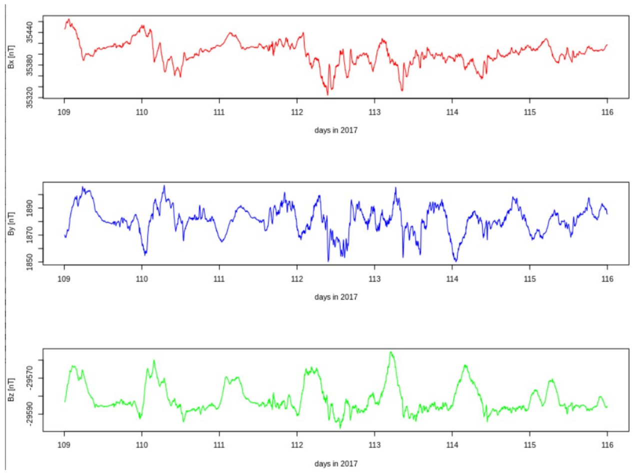
Buenos Aires, Argentina, 7th November, 2022 – 11th November, 2022

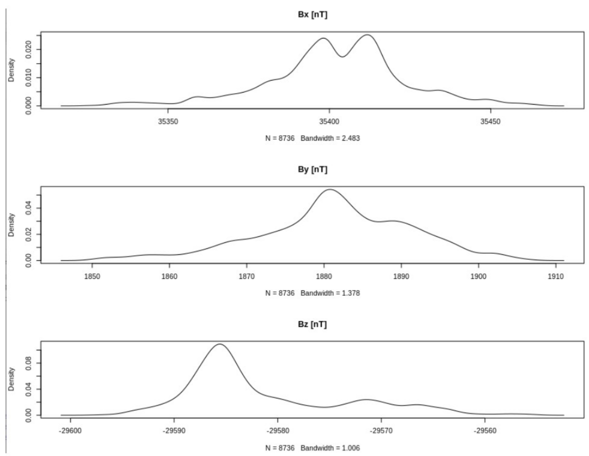
On performance assessment of machine learning-based GNSS ionospheric delay correction model based on space weather predictors in immediate positioning environment (R Filjar, *Croatia*)

- **Data**
- Time span:
- Day 109 – Day 116 in 2017
- Source:
  - IGS (Darwin, NT, Australia)
  - INTERMAGNET (Kakadu, NT, Australia)

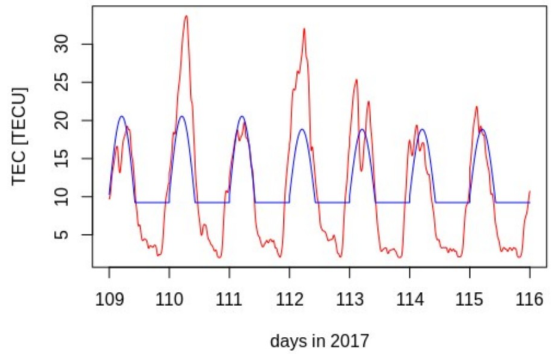


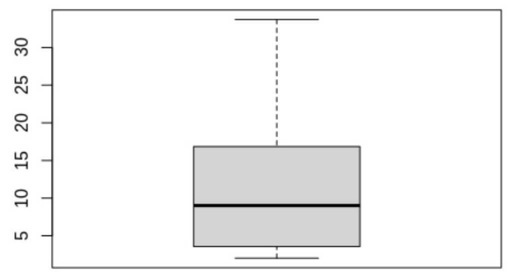
predictors: Bx, By, Bz

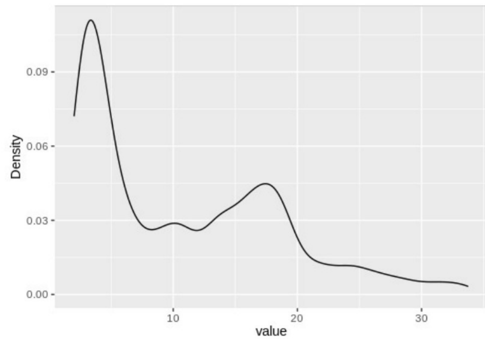




target: TEC







International Workshop on Machine Learning for Space Weather: Fundamentals, Tools and Future Prospects

Buenos Aires, Argentina, 7th November, 2022 – 11th November, 2022

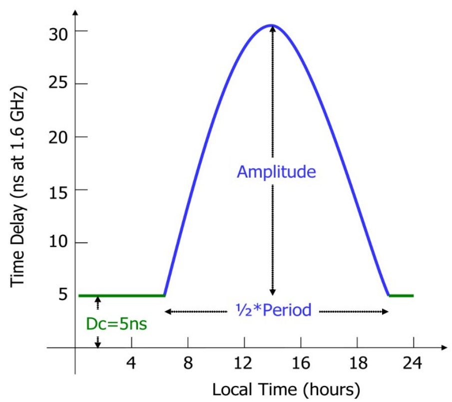
On performance assessment of machine learning-based GNSS ionospheric delay correction model based on space weather predictors in immediate positioning environment (R Filjar, *Croatia*)

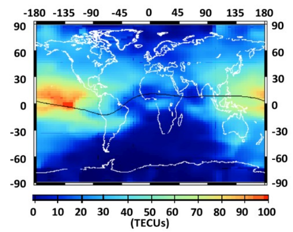
- <u>Case study of GNSS ionospheric delay correction model performance assessment</u>

- Geomagnetic field observations-based candidate models:
- LRM … Linear Regression Model
- MMLPNN … Monotone Multi-layer Perceptron Neural Network Model
- RFM … Random Forest Model,
- Klobuchar … standard GPS Klobuchar Model

Source:
Sanz Subirana, J, Juan Zornoza, J M, Hernández-Pajares, M. (2013). GNSS Data Processing, Volume I: Fundamentals and Algorithms. ESA. Noordwijk, The Netherlands. ISBN 978-92-9221-886-7.
Available at: https://gssc.esa.int/navipedia/GNSS_Book/ESA_GNSS-Book_TM-23_Vol_I.pdf

Klobuchar model





$$Ion_{VERT} = \begin{cases} DC + A\cos\left[\frac{2\pi(t-\Phi)}{P}\right] & (day) \\ DC; \quad if \left[\frac{2\pi(t-\Phi)}{P}\right] \geq \frac{\pi}{2} & (night) \end{cases}$$

Being :

$$A = \sum_{n=0}^{3} \alpha_n \varphi^n \quad ; \quad P = \sum_{n=0}^{3} \beta_n \varphi^n$$

$\alpha_n, \beta_n =$ Klobuchar coefficients

$\varphi =$ Geomagnetic latitude

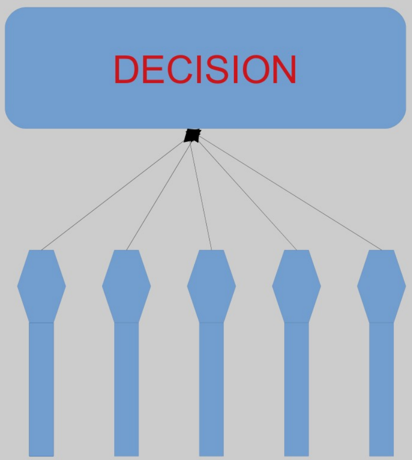
$$Ion_{SLANT} = Ion_{VERT} \; m(elev)$$

$$m(elev) = \left[1 - \left(\frac{R_E}{R_E + h}\cos(elev)\right)^2\right]^{-1/2}$$

Where:
DC= 5ns
Φ= 14h (constant phase offset)
t = Local time

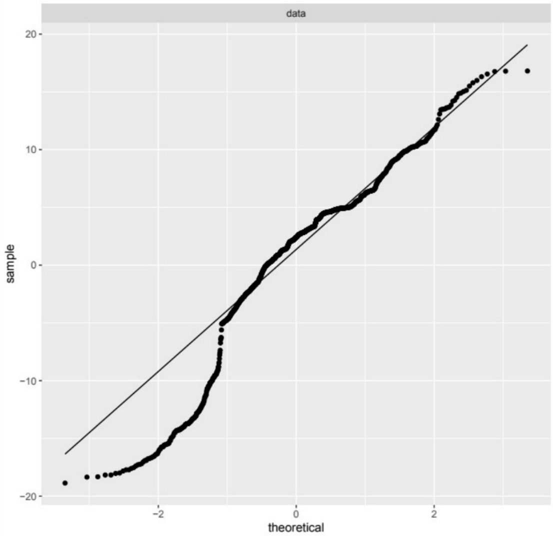
International Workshop on Machine Learning for Space Weather: Fundamentals, Tools and Future Prospects
Buenos Aires, Argentina, 7th November, 2022 – 11th November, 2022
On performance assessment of machine learning-based GNSS ionospheric delay correction model based on space weather predictors in immediate positioning environment (R Filjar, *Croatia*)

- <u>Case study of GNSS ionospheric delay correction model performance assessment</u>
- Machine learning-based methods for model development

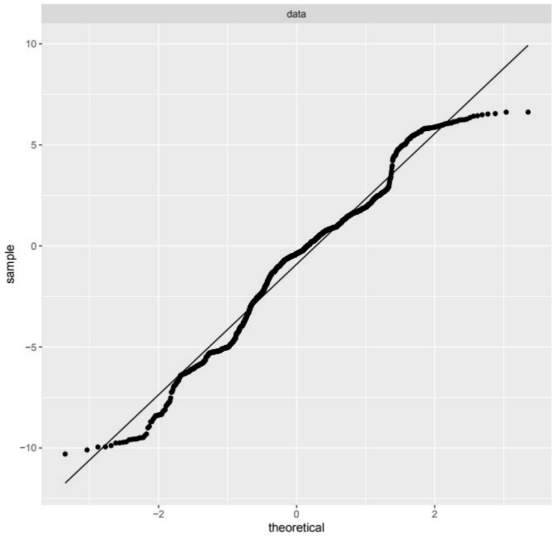
LINEAR REGRESSION (LR)	MONOTONE MULTI-LAYER PERCEPTRON NEURAL NETWORK (MMLPNN)	RANDOM FOREST (RF)
$\hat{y} = \beta_0 + \sum_{i=1}^{3} \beta_i x_i$ $\boldsymbol{x} = (x_1 = B_x, x_2 = B_y, x_3 = B_z)$	$\hat{y}(x) = w_b + \sum_{l=1}^{L} w_l \, tanh \cdot \left(w_{b,l} + \sum_{h=1}^{H} w_{l,h} \, tanh\left(w_{b,h} + \sum_{i=1}^{I} w_{h,i} \, x_i\right)\right)$ $\frac{\partial \hat{y}}{\partial x_j} = \sum_{l=1}^{L} w_l \cdot (1-\Theta_1^2) \cdot \sum_{l=1}^{H} w_{l,h} \cdot (1-\Theta_2^2) \cdot w_{h,j} \geq 0$ $\Theta_1 = tanh\left(w_{b,l} + \sum_{h=1}^{H} w_{l,h} \, tanh\left(w_{b,h} + \sum_{i=1}^{I} w_{h,i} \, x_i\right)\right)$ $\Theta_2 = tanh\left(w_{b,h} + \sum_{i=1}^{I} w_{h,i} \, x_i\right)$	

- Case study of GNSS ionospheric delay correction model performance assessment
- Model performance assessment – q-q diagrams (conformity of residuals to normal distribution)

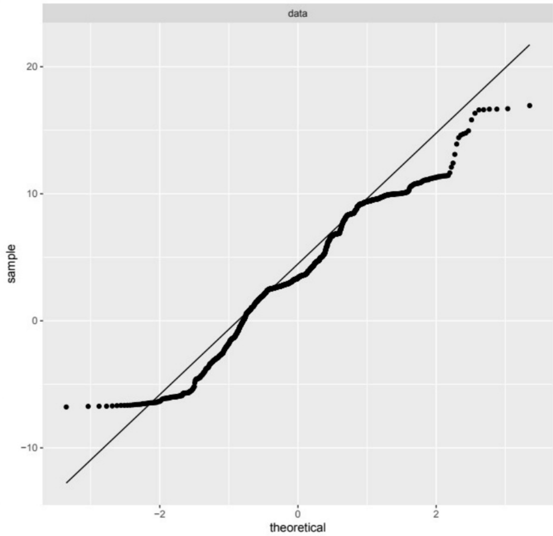
L R



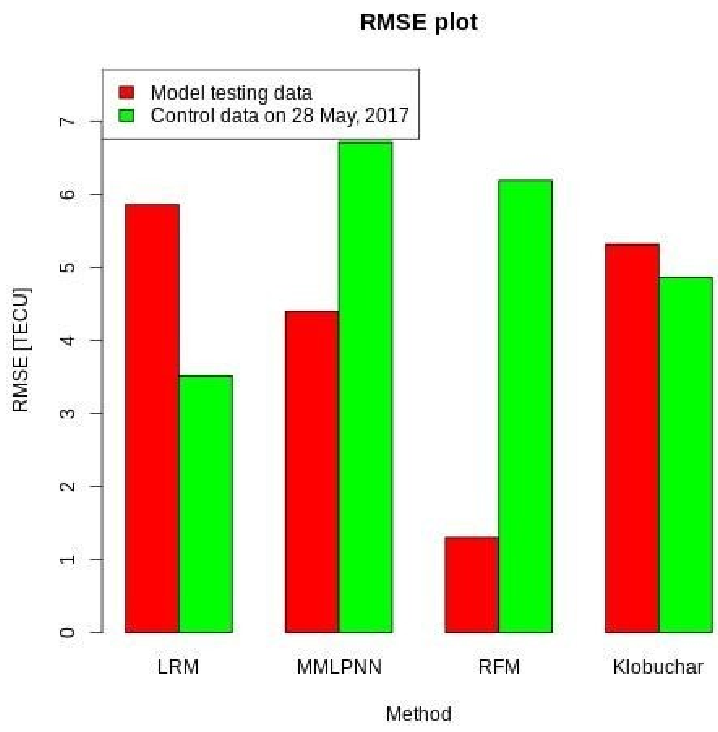
M M L P N N



R F



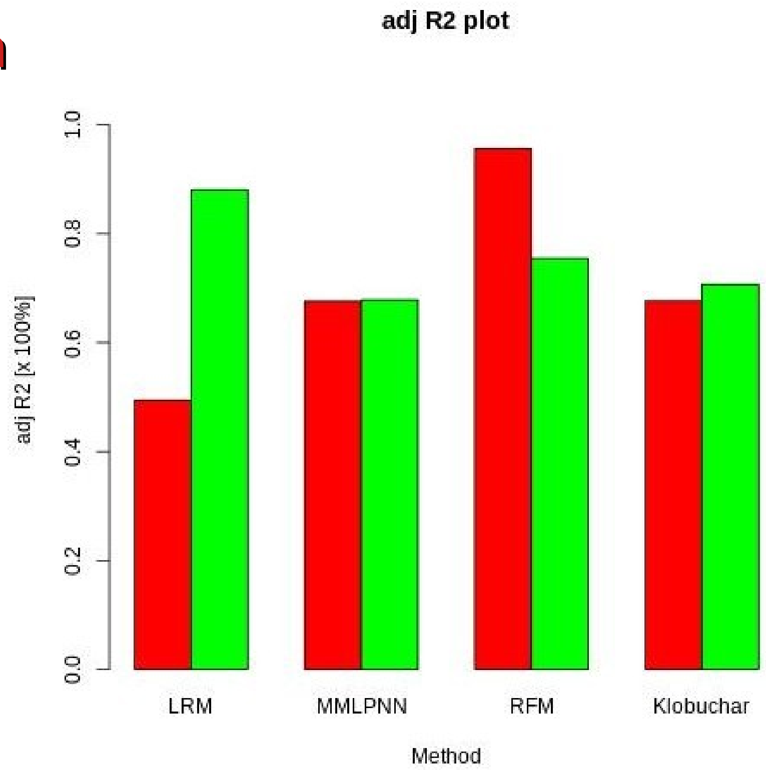
- <u>Case study of GNSS ionospheric delay correction model performance assessment</u>
- Model performance assessment – bias
- Root Means Square Error (RMSE)
- LRM … Linear Regression Model
- MMLPNN … Monotone Multi-layer Perceptron Neural Network Model
- RFM … Random Forest Model,
- Klobuchar … standard GPS Klobuchar Model



- <u>Case study of GNSS ionospheric delay correction model performance assessment</u>
- Model performance assessment – variance
- Adjusted coefficient of determination
- *adjR2* parameter → for objective comparison of described original

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

$$adjR2 = 1 - (1 - R^2) \cdot \frac{s_N - 1}{s_N - p}$$



Discussion

- Essential practical methodology for a machine learning-, and observations-based predictive model performance assessment is outlined, for the purpose of GNSS-related ionospheric delay correction model development
- A case study of a machine learning self-adaptive positioning (i. e. geomagnetic) environment-aware GNSS ionospheric delay model development, and its performance assessment is presented in the scenario of short-term raoidly developing geomagnetic storm in sub-equatorial region
- Future research: model development for classes of short-term rapidly developing storms, battery of statistical tests to be selected for usage in model performance assessment, computational requirements to be defined in distributed positioning environment

International Workshop on Machine Learning for Space Weather: Fundamentals, Tools and Future Prospects

Buenos Aires, Argentina, 7th November, 2022 – 11th November, 2022

On performance assessment of machine learning-based GNSS ionospheric delay correction model based on space weather predictors in immediate positioning environment (R Filjar, *Croatia*)

Reference list

- Filjar, R. (2022). An application-centred resilient GNSS position estimation algorithm based on positioning environment conditions awareness. Proc ION ITM 2022, 1123 - 1136. Long Beach, CA. doi: 10.33012/2022.18247
- Filjar, R, Weintrit, A, Iliev, T, Malčić, G, Jukić, O, Sikirica, N. (2020). Predictive Model of Total Electron Content during Moderately Disturbed Geomagnetic Conditions for GNSS Positioning Performance Improvement. Proc FUSION2020, 256-262. Sun City, South Africa. doi: https://doi.org/10.23919/FUSION45008.2020.9190264
- Filić, M, Filjar, R. (2019). On correlation between SID monitor and GPS-derived TEC observations during a massive ionospheric storm development. URSI AP-RASC 2019 Meeting. New Delhi, India. doi: 10.23919/URSIAP-RASC.2019.8738664
- Filić, M, and Filjar, R. (2018). Forecasting model of space weather-driven GNSS positioning performance (monograph). Lambert Academic Publishing. Riga, Latvia. ISBN 978-613-9-90118-0.
- Filić, M, Filjar, R. (2018). Modelling the Relation between GNSS Positioning Performance Degradation, and Space Weather and Ionospheric Conditions using RReliefF Features Selection. Proc of 31st International Technical Meeting ION GNSS+ 2018, 1999-2006. Miami, FL. doi: https://doi.org/10.33012/2018.16016
- Filić, M, Grubišić, L, Filjar, R. (2018). Improvement of standard GPS position estimation algorithm through utilization of Weighted Least-Square approach. Proc of 11th Annual Baška GNSS Conference, 7-19. Baška, Krk Island, Croatia. Available at: https://www.pfri.uniri.hr/web/hr/dokumenti/zbornici-gnss/2018-GNSS-11.pdf

Reference list

- Maindonald, J, John Brown, W. (2010). Data Analysis and Graphics Using R: An Example-Based Approach (3rd). Cambridge University Press. Cambridge, UK. ISBN 9781139194648
- Lindholm, A, Wahlström, N, Lindsten, F, and Schön, T B. (2022). Machine Learning – A First Course for Engineers and Scientists. Cambridge University Press. Cambridge, UK. ISBN: 9781108843607. Available at: http://smlbook.org/
- James, G, Witten, D, Hastie, T, Tibshirani, R. (2021). An Introduction to Statistical Learning with Applications in R. Springer Verlag. ISBN 978-1071614174. Available at: https://www.statlearning.com/
- Efron, B, Hastie, T. (2016). Computer Age Statistical Inference: Algorithms, Evidence and Data Science. Cambridge University Press. Cambridge, UK. ISBN 9781107149892. Available at: https://hastie.su.domains/CASI/order.html
- Biecek, P, Burzykowski, T. (2020). Explanatory Model Analysis: Explore, Explain, and Examine Predictive Models. With examples in R and Python. CRC Press. Boca Raton, FL. ISBN 9780367135591. Available at: https://ema.drwhy.ai/


International Workshop on Machine Learning for Space Weather:
Fundamentals, Tools and Future Prospects
Buenos Aires, Argentina, 7th November, 2022 – 11th November, 2022
On performance assessment of machine learning-based GNSS ionospheric delay correction model based on space weather predictors in immediate positioning environment (R Filjar, Croatia)
In appreciation of your attention, and
with invitation to
Baška SIF (Spatial Information Fusion) Meetings,
every October in Baška, Krk Island, Croatia
Dr Renato Filjar
Laboratory for Spatial Intelligence, Krapina University of Applied Sciences, Krapina, Croatia
E-mail: renato.filjar@gmail.com