Katholische Universität Eichstätt-Ingolstadt

MIDS Mathematical Institute for Machine Learning and Data Science

PARAMETER AND ITS UNCERTAINTY ESTIMATION

Data Lab Project Presentation, Summer Semester 2024

Authors : Faustino Vazquez | Elshan Dashtiyev

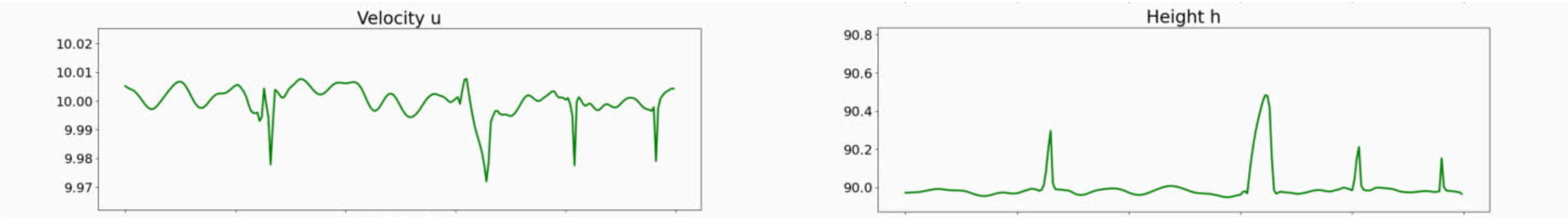
Supervisors : Prof. Tijana Janjic | Dr. Ramezani Ziarani, Maryam

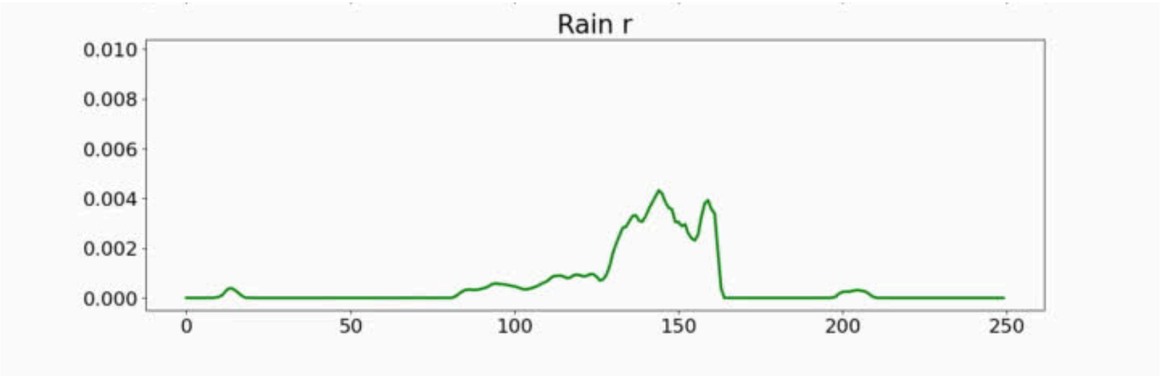
STRUCTURE

1. Introduction
2. Database
3. Exploratory Data Analysis
4. Models implementation
5. Analysis
6. Conclusions

INTRODUCTION

Features





INTRODUCTION

Target Variables

α - **Half-life of influence of rain of roughly 1 hour**

ϕ - **Geopotential constant value above first threshold to allow for unstable convection.**

h - height threshold for rain.

TABLE 1 Lower and upper bounds for the uniform distributions of the model parameters

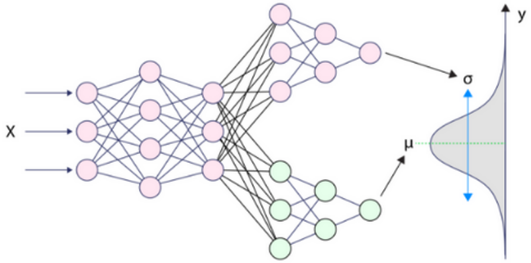
Parameter	Lower bound	Upper bound
α	0.0003	0.001
ϕ_c	899.7	899.9
h_r	90.15	90.25

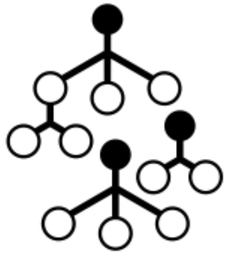
INTRODUCTION

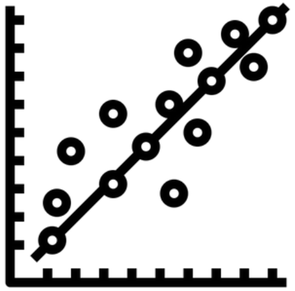
Materials and methods

Python - Machine Learning

Neural Network / Random Forest / Linear Regression







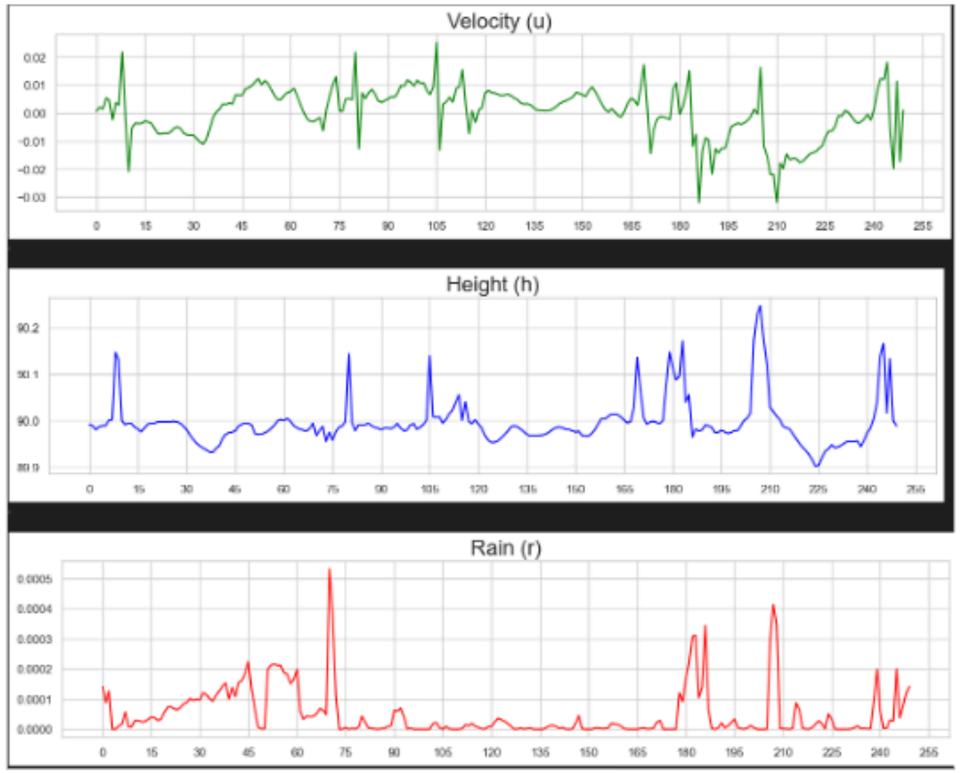
INTRODUCTION

Background Objective and Research question

Explore the dynamics between the elements in Shallow Water and their impact. Analyse the behaviour of elements like height, rain and velocity.

Is it possible possible to predict the target variables depending on the behaviour of the features (‘velocity’, ‘rain’ and ‘height’))?

DATABASE



16,035 Instances x 3 Features x 3 Target Variables

FEATURES :

- Velocity (250 gridpoints)
- Height (250 gridpoints)
- Rain (250 gridpoints)

TARGET VALUES (UNIFORMLY DISTRIBUTED):

- alpha
- phic
- h_rain

KATHOLISCHE UNIVERSITÄT EICHSTÄTT-INGOLSTADT

MIDS Mathematical Institute for Machine Learning and Data Science

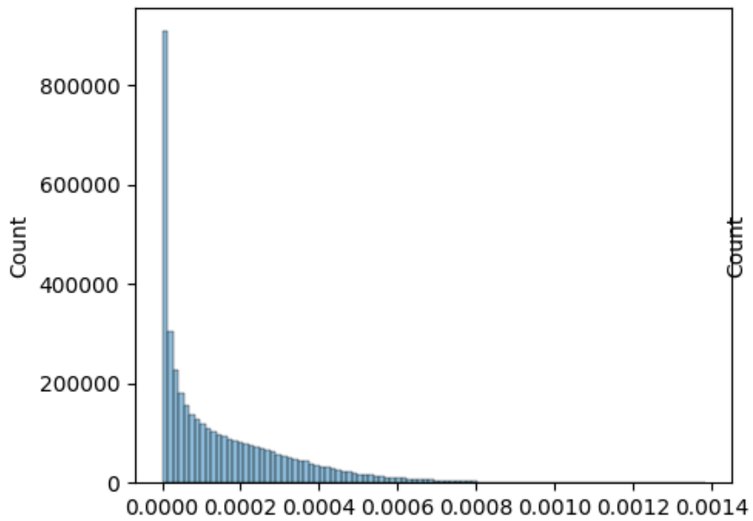
EXPLORATORY DATA ANALYSIS

```
Run Number,phic,h_rain,alpha,Velocity,Height,Rain
1,899.7067188312412,90.1907397606142,0.0008921197793287174,"[[0.010799882501991228], [0.009855476258547132], [0.007824513066283515], [0.004944743411602115], [0.001616
2,899.7999178528956,90.18041924484294,0.0004002048744743250 7,"[[0.010351237367773532], [0.008246628456653601], [0.004290072955496145 6], [0.001421996639231445 8], [-0.000
3,899.8278607932398,90.24856183970041,0.0009550807291113917,"[[0.009881864105294548], [0.008554994957547545], [0.00613092415599859], [0.002256930549810080 6], [0.000426
4,899.8879854916166,90.21210563149941,0.0009174910870069957,"[[-0.002534188269910481 3], [-0.000264394990128503 26], [-0.000633245006754056 4], [-0.002115299679991800 36],
5,899.8712398317421,90.2423514571513 4,0.0003991504336066022 6,"[[0.015061786674556123], [0.014376011505152105], [0.011034052251985942], [0.007257528238807642], [0.0035
6,899.8688415546837,90.24866733120018,0.0005416647974170053,"[[0.011005684652196 4], [0.009738672672746194], [0.007549075616643731 4], [0.004459727804616771], [-2.549785
7,899.8109553179236,90.2110135030212,0.0004989073320662292,"[[0.009500079032884648], [0.008850836004543916], [0.006900603875919598], [0.004425422284853345], [0.001943
8,899.8645250440704,90.2124242305189 3,0.0005469528321476138,"[[0.013600628685508 62], [0.012776393867767867], [0.010058943331846057], [0.008843537969649014], [0.0058046
9,899.7345092388506,90.22463208989497,0.0003763218613716044 6,"[[0.014510502039047333], [0.012655957764622498], [0.010217423177265823], [0.007197032845064476], [0.0039
10,899.8320855292686,90.1562195468977 7,0.0006788319416691394,"[[0.006554707630059 65], [0.002690547717377908], [0.002033174033607925], [-0.001369528381814924 8], [-0.00
11,899.7751139415824,90.21753020410496,0.0004009835427435319 7,"[[0.008831297265674392], [0.008441239235073091], [0.005900988549173243], [0.001909193773034648 4], [-0.00
12,899.7894462079158,90.18047999852911,0.0003143252607497747,"[[0.008756263784609855], [0.009122887177551898], [0.006184496822997534], [0.003162658863487146 4], [-0.000
13,899.8987359093056,90.24393498770435,0.0005136604402169411,"[[0.001007793934851854], [-0.001042913924747851 1], [-0.005483288244378856], [-0.008094133433171303], [-0.
14,899.7045279274354,90.19586099988302,0.0003066525698869307,"[[0.010236144769270532], [0.009177689440975327], [0.006967957975673 1], [0.003852871501442573], [-0.000270
15,899.8085032982411,90.2003591798047,0.0008180677987231635,"[[-0.001895793909980323 3], [-0.002370349758562832], [-0.004698004351347151 4], [-0.007531979892984049], [-0
16,899.7869610393045,90.23483472886168,0.0008215277487002598,"[[0.006608384885208 66], [0.005326678869910525], [0.003160320237517614], [0.003399705823824744], [-0.00
17,899.7419469664047,90.2129832026862 3,0.0007046007180999155,"[[0.011541697930515 1], [0.008929113092046493], [0.005767598205591606], [0.002113469075586873 8], [-0.002
18,899.8680370094883,90.2309007595638 2,0.0003908955426438928,"[[0.018827511266122244], [0.016777263534307 9], [0.013840283638156375], [0.010250043179759269], [0.005462
19,899.7693155940057,90.24877534883534,0.0009288635606196006,"[[0.005562185238868745], [0.003061815608523966 7], [0.000534628104832448 8], [-0.001831038392103716 8], [-0.
20,899.7073868780951,90.23066766905936,0.0008229131149037518,"[[0.006072686893686015], [0.004893360325858259], [0.004517071072738831], [0.005162332846309971], [0.00759
21,899.7828382153513,90.1579798406691,0.0003540815488820076,"[[0.012124389156296626], [0.010275238264669341], [0.009503921818875223], [0.006593514520406656], [0.00383
22,899.7568390103763,90.24144315642462,0.0006546664113928884,"[[0.007318113281749933 5], [0.007305416586037751], [0.006135708291382744 5], [0.003394192780314563], [0.001
23,899.7165329429637,90.18713301293599,0.0005817231458642 14,"[[0.004994802959887812], [0.003086137756522829], [0.008844088052840871 4], [4.222553142854228e-05], [-0.002
24,899.8456773516169,90.17758309616958,0.0005246716704372262,"[[0.025944833704816 99], [0.025989442236256 2], [0.023913200404863263], [0.021296654981648677], [0.0197040
25,899.7741638923962,90.19589106385502,0.0003728072174382247 7,"[[0.013288713181027 2], [0.012290513575942682], [0.010440382863768265], [0.007897361915138 64], [0.004770
26,899.871069786052,90.20318149373315,0.0003318594335049553 4,"[[0.008668701331267401], [0.007855609905391369], [0.006099849006861955], [0.003625283088419081 2], [0.000
27,899.7010488274345,90.20428499993989,0.0008983233672363665,"[[0.007169501232975401], [0.006858539025473366], [0.008490319752772282], [0.004753549704709454], [0.0044
28,899.8853222095892,90.20479275489151,0.0004476327481751923 7,"[[0.007657881718938723], [0.007088988422016743], [0.005312822793524738], [0.002629366431329286 3], [-0.00
29,899.7548779368137,90.15452759485053,0.0003248941960944507 3,"[[0.014391181816972335], [0.013464561779590 35], [0.012028870909254242], [0.010381165908207141], [0.0084
30,899.8709064628924,90.22483450619653,0.0003662145638098984,"[[0.011833959386422347], [0.011427024110893412], [0.009500132278264944], [0.007444692666298269], [0.0051
31,899.726684384547,90.21394801069712,0.0008101220909145747,"[[0.015781394910166882], [0.014357333390579513], [0.012230410813566829], [0.009597528802555263], [0.006763
32,899.7891136960388,90.24106105680829,0.0005185944725832 15,"[[0.012071646257758832], [0.010554243784574603], [0.008226212374418536], [0.005060926824765204], [0.0017
33,899.8542560048664,90.17796505069425,0.0003794886110744956,"[[0.008356452651012085], [0.005706545933701146 2], [0.002302168156249854 5], [-0.001287850789591815], [-0.00
34,899.8459573091498,90.16681549540868,0.0006791793415298745,"[[0.002281166138698073 3], [0.001570777079410543 9], [0.001443092561676201], [0.006773520401651165], [0.00
35,899.7782195596844,90.23765907396138,0.0007298108371021803,"[[0.009329283939737222 3], [0.008147414560732463], [0.006075995815783286], [0.003309360058661 3], [0.000134
36,899.8237065321881,90.16881481441489,0.0005118468836645366,"[[0.010585739531144 54], [0.008252943326085727], [0.005026475718527291], [0.001400296327093399 1], [-0.0019
37,899.8325013094407,90.23655317218434,0.0006846576069052119,"[[0.008569988511015826], [0.007987464509666545], [0.006777247086379465], [0.004997367653641758], [0.002
38,899.8759674540511,90.19023876599886,0.0004392756474980228 5,"[[0.004290778668115319 4], [0.003618840002243611 5], [0.002774019871238299], [0.001774812332998128 4], [0.0
39,899.7353545300091,90.22504733556649,0.0003663747543284894,"[[0.006832823397565527], [0.005967516426168741], [0.004127970674117654], [0.001548653869653928 4], [-0.00
40,899.8278743859249,90.24412173506228,0.0005346630154005231,"[[0.006926761459778286], [0.005592400674705447], [0.003137685544957872 5], [-0.000220406584592153 3], [-0.0
41,899.8522239357325,90.23651369351094,0.0007682862173203921,"[[0.013573845536485 12], [0.012602935727245686], [0.010168583184029132], [0.008674513995564586], [0.00627
42,899.7746910042664,90.19557690281006,0.0009681026090434866,"[[0.005450959925100973], [0.004560192353246 25], [0.002758771128361742], [0.000252422845308745 2], [-0.0036
43,899.7904076911364,90.20803889563095,0.0009782910694895776,"[[-0.000959466799560740 5], [-0.003377372686630304], [-0.004781755561374765], [-0.007327055842591774], [-0
44,899.8712316613678,90.21220826683493,0.0009127998047990 41,"[[0.005194614085933514], [0.005330998614205603], [0.004503423158021121], [0.001944775187822329], [0.00071
```

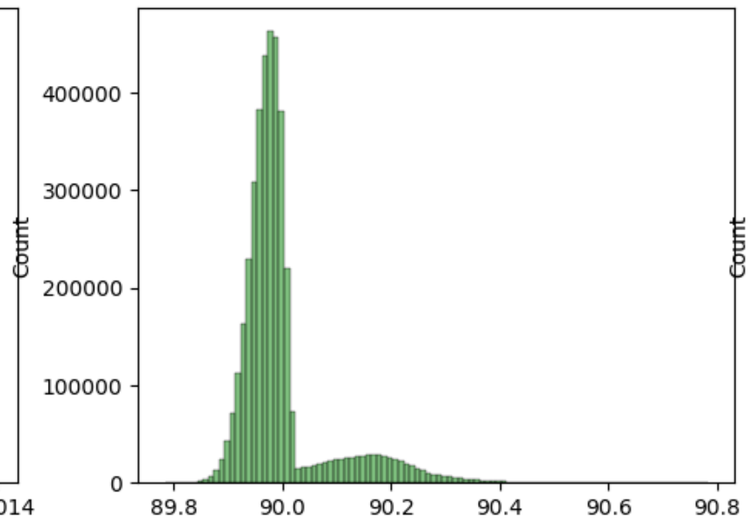
```
Information about the data:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 16035 entries, 0 to 16034
Data columns (total 6 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   phic      16035 non-null  float64
 1   h_rain    16035 non-null  float64
 2   alpha     16035 non-null  float64
 3   Velocity  16035 non-null  object
 4   Height    16035 non-null  object
 5   Rain      16035 non-null  object
dtypes: float64(3), object(3)
memory usage: 751.8+ KB
```

EXPLORATORY DATA ANALYSIS

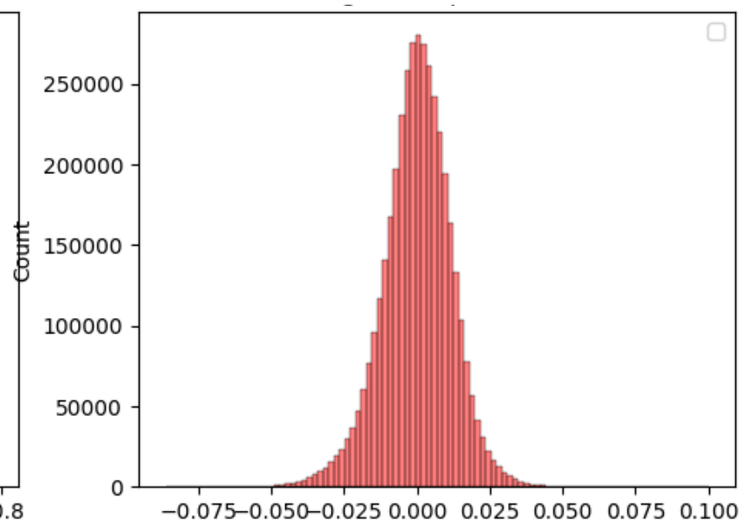
Histograms Features



rain



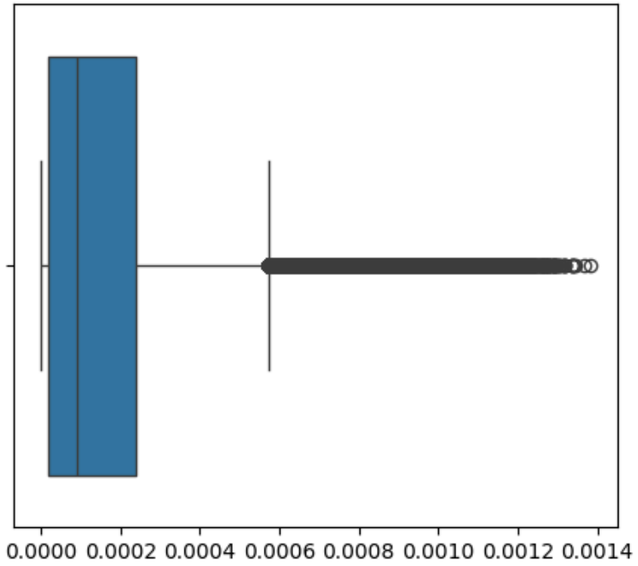
height



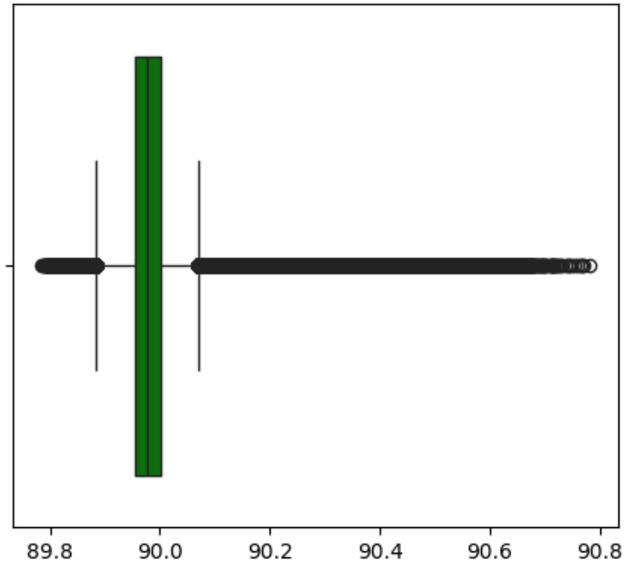
velocity

EXPLORATORY DATA ANALYSIS

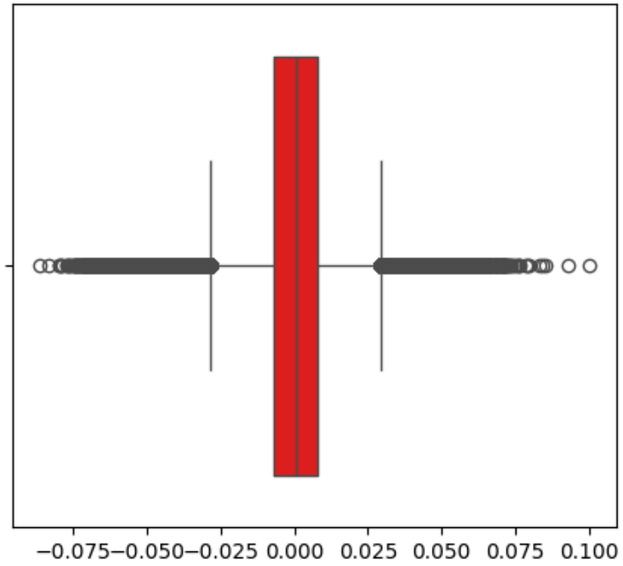
Boxplots Features



rain



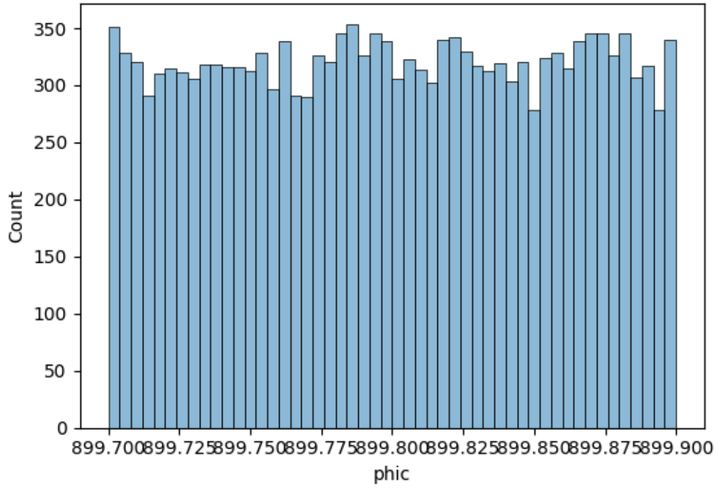
height



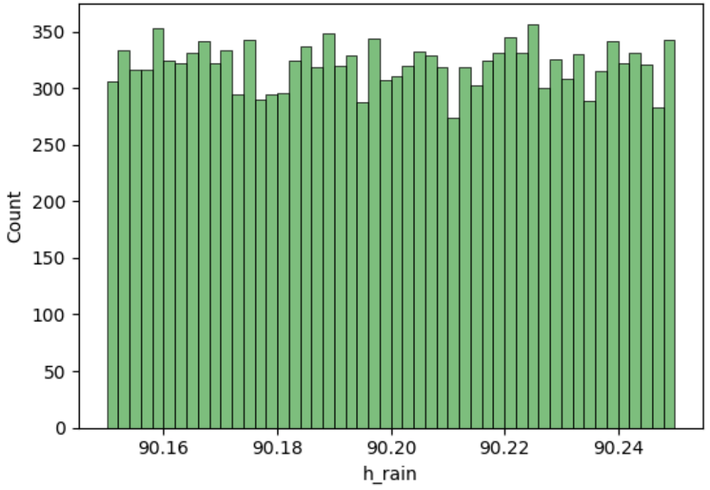
velocity

EXPLORATORY DATA ANALYSIS

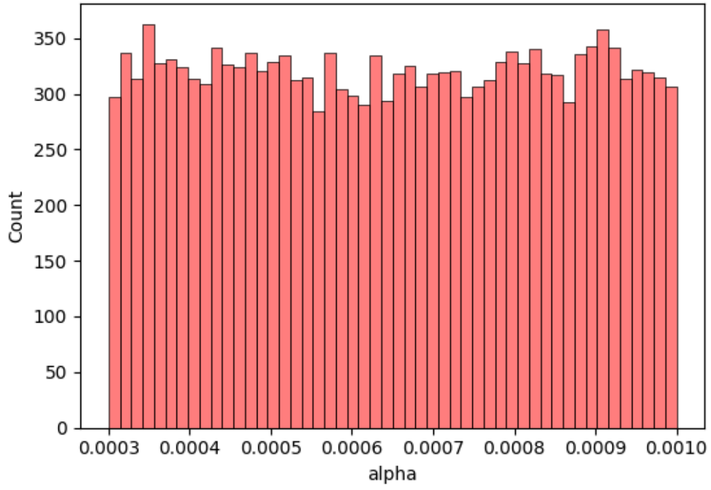
Histograms Target class



phic



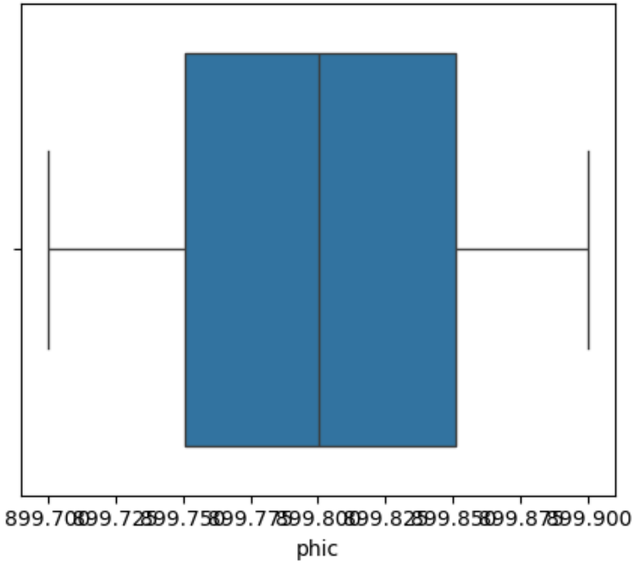
h_rain



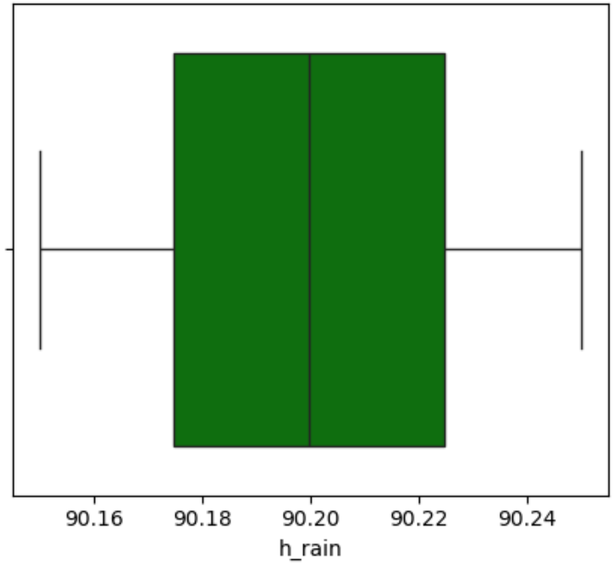
alpha

EXPLORATORY DATA ANALYSIS

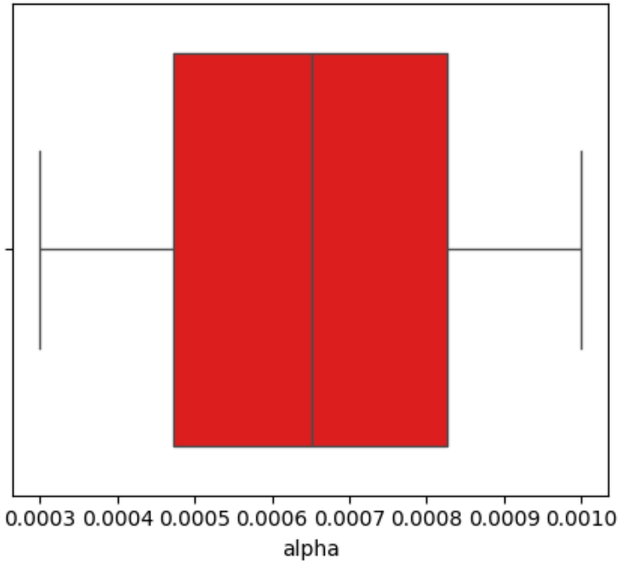
Boxplots Target class



phic



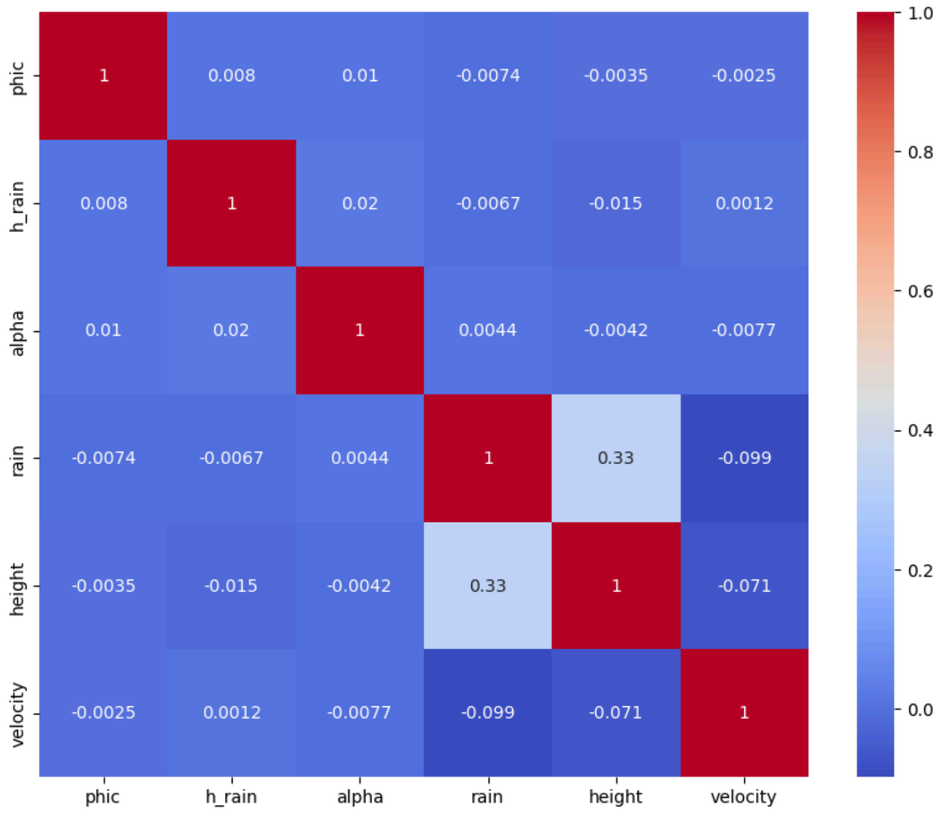
h_rain



alpha

EXPLORATORY DATA ANALYSIS

Correlation Heatmap



MODELS

Error Metrics

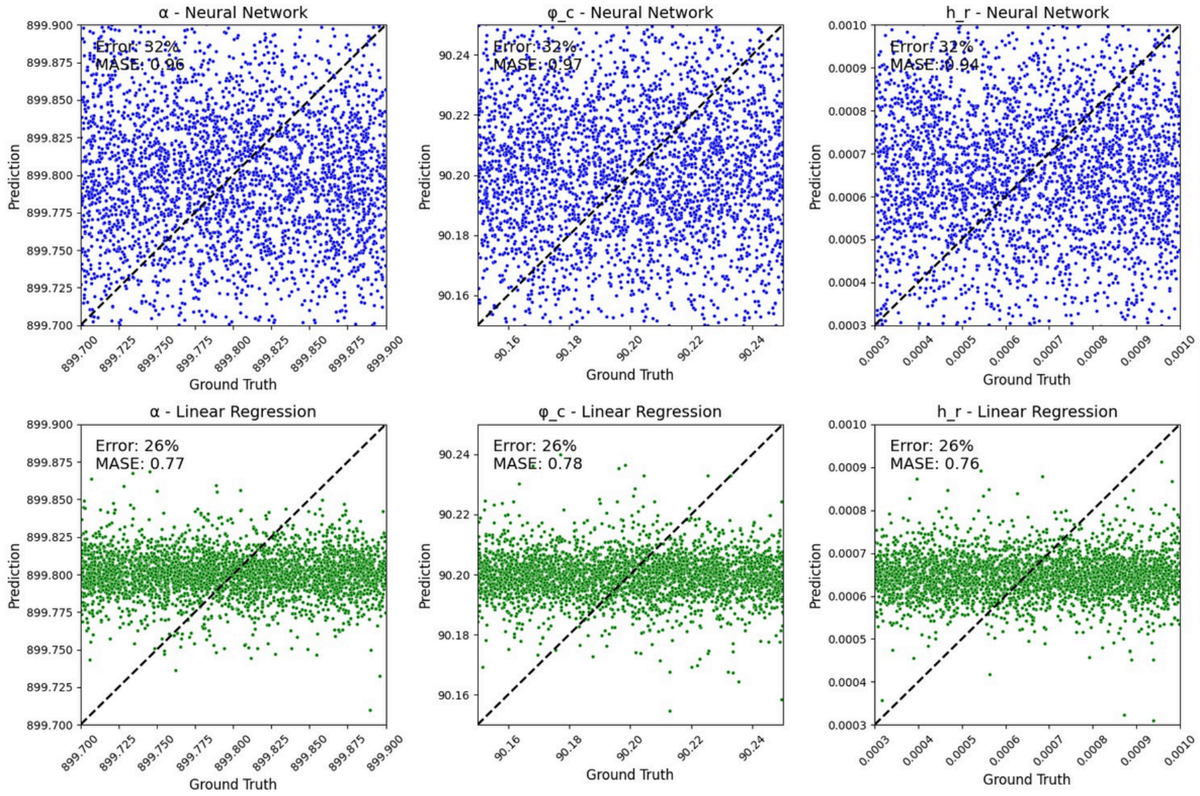
$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right|$$

$$MASE = \frac{MAE}{(1/n-1)\sum_{i=2}^{n} |a_i - a_{i-1}|}$$

MODELS

Linear Regression | Neural Networks - Original Data

DATA SPLITTED: TRAIN SIZE = 80%TEST SIZE = 20%



Mean Absolute Percentage Error (MAPE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	21.748950374904872	27.315075066383475	24.86813741632042	25.083049396716596
h_rain	21.637021350786885	27.10094817946425	25.052237672976585	25.092820309525504
alpha	22.239522216984412	27.46224386688945	25.180741695915444	25.23936194672509

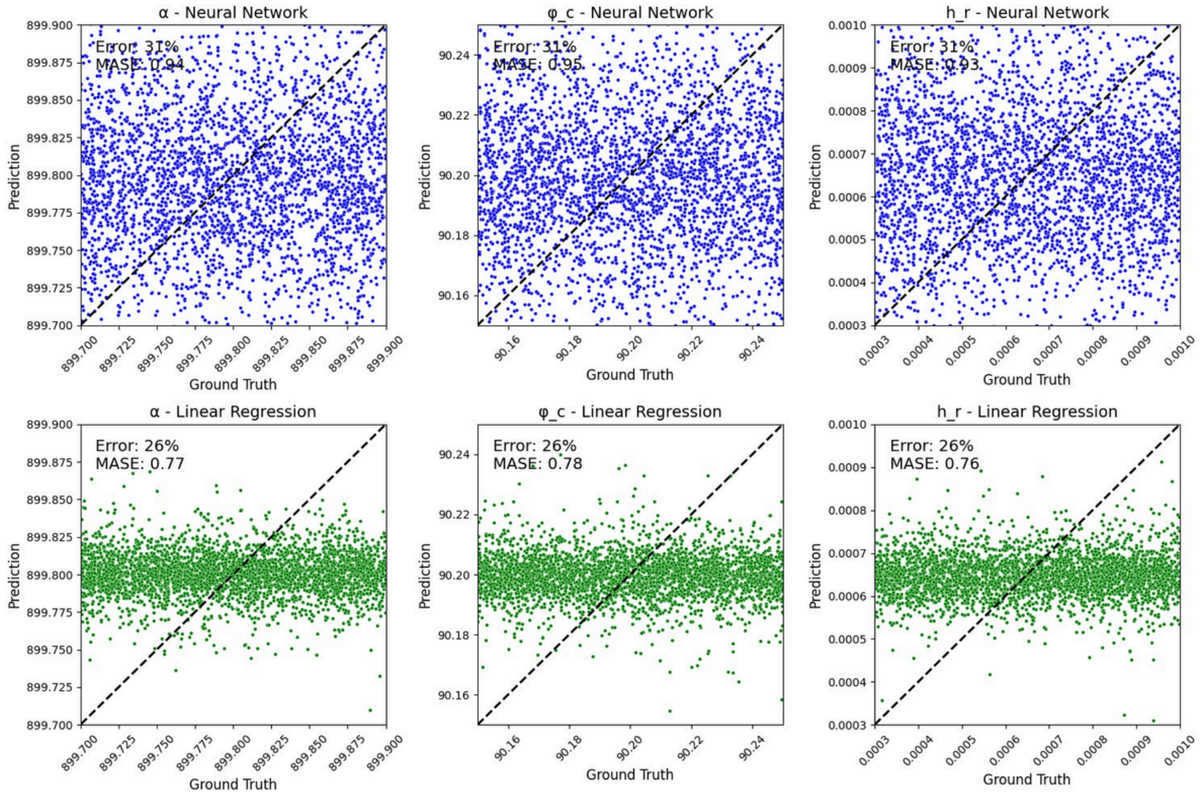
Mean Absolute Scaled Error (MASE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	0.649299232824065	0.8206454105239567	0.7424203130653794	0.7535871426066171
h_rain	0.6454064816943108	0.8244807178182957	0.7472783019876008	0.7633882904716909
alpha	0.6653643215571101	0.8149204999119286	0.7533600295608973	0.7489582262388157

MODELS

Linear Regression | Neural Networks - Standart Scaler

DATA SPLITTED: TRAIN SIZE = 80%TEST SIZE = 20%



Mean Absolute Percentage Error (MAPE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	11.81220474218926	31.4195055714978	23.917158413268265	25.649218596490662
h_rain	11.660333584316156	31.323282336401938	24.0778589158509	25.746360006992163
alpha	11.37657797950784	31.331917532867614	24.180259572647238	25.589224866097414

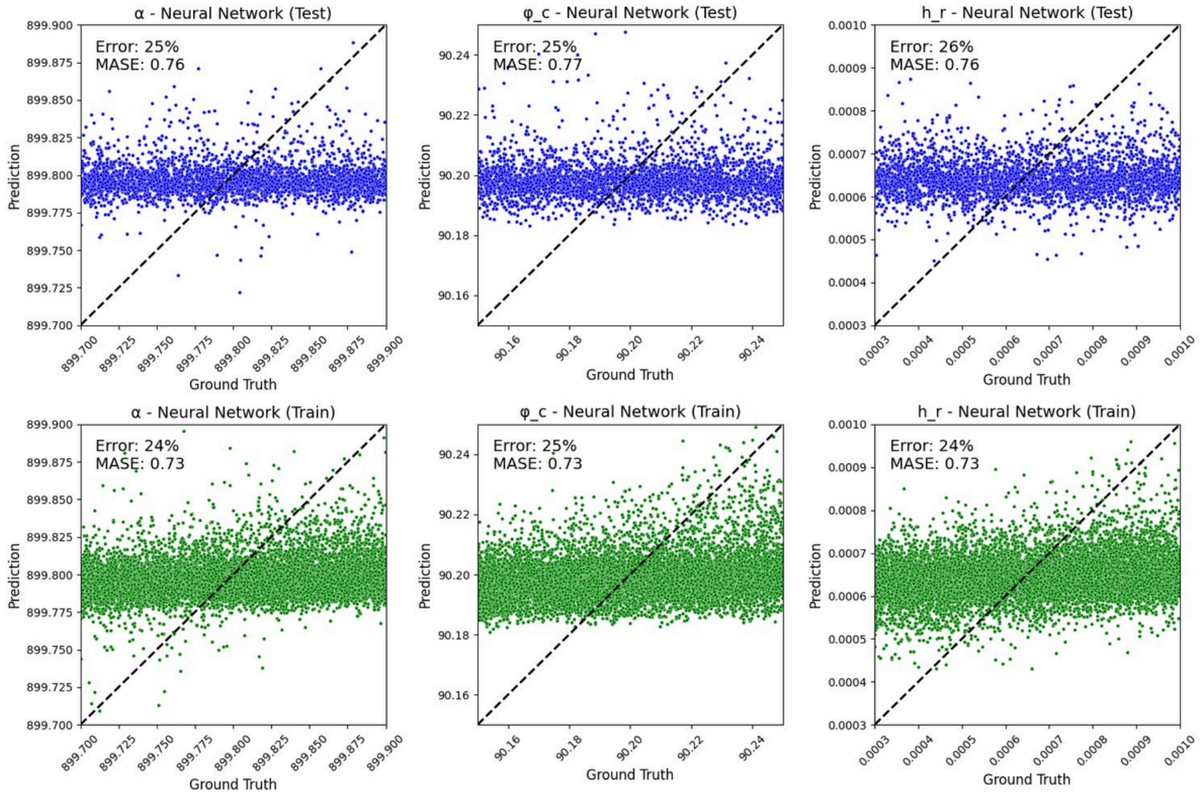
Mean Absolute Scaled Error (MASE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	0.3526448561818384	0.9439576126193457	0.7140295205687434	0.7705969496177801
h_rain	0.34781381189337385	0.9529350092891233	0.7182137484486133	0.7832706530857918
alpha	0.3403656344377705	0.929750024172946	0.7234275021134335	0.7593401333596426

MODELS

Linear Regression | Neural Networks - MinMaxScaler

DATA SPLITTED: TRAIN SIZE = 80%TEST SIZE = 20%



Mean Absolute Percentage Error (MAPE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	24.346112038009178	25.338765043165008	23.917158413267988	25.64921859649079
h_rain	24.504316948472464	25.3316915674618	24.077858915850932	25.746360006991935
alpha	24.442620204451345	25.53891629589814	24.180259572647063	25.58922486609759

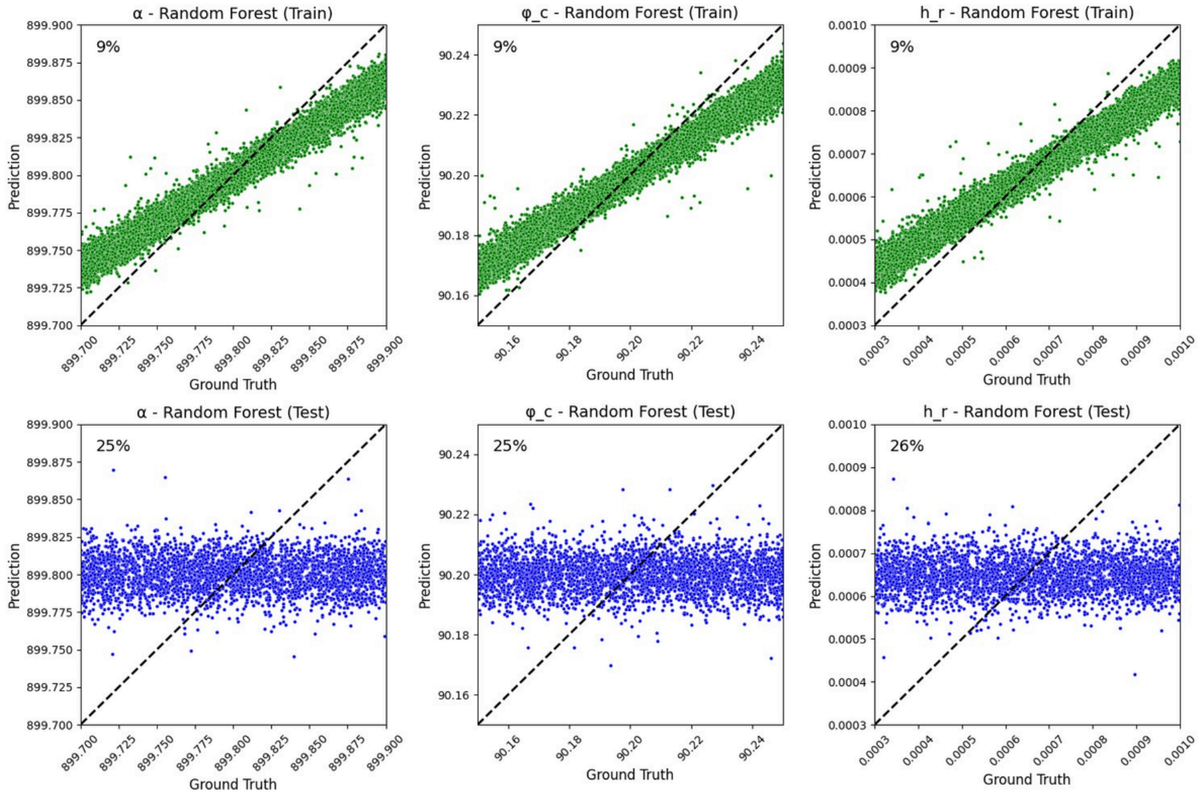
Mean Absolute Scaled Error (MASE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	0.7268356217672061	0.761269781997029	0.714029520568735	0.7705969496177839
h_rain	0.7309344817677838	0.7706553700183288	0.7182137484486141	0.7832706530857847
alpha	0.7312768345802221	0.7578472660842894	0.7234275021134285	0.759340133359648

MODELS

Random Forest - StandarScaler

DATA SPLITTED: TRAIN SIZE = 80%TEST SIZE = 20%



Mean Absolute Percentage Error (MAPE)

Metric	Random Forest (Train)	Random Forest (Test)
phic	9.364441743620073	25.395726537989766
h_rain	9.377166628219616	25.41092632771062
alpha	9.494532551318882	25.6092293818387

Mean Absolute Scaled Error (MASE)

Metric	Random Forest (Train)	Random Forest (Test)
φ_c	0.28	0.77
h_r	0.28	0.77
α	0.28	0.77

ANALYSIS

COMPARASION

Linear Regression | Neural Networks - Original Data

Mean Absolute Percentage Error (MAPE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	21.748950374904872	27.315075066383475	24.86813741632042	25.083049396716596
h_rain	21.637021350786885	27.10094817946425	25.052237672976585	25.092820309525504
alpha	22.239522216984412	27.46224386688945	25.180741695915444	25.23936194672509

Mean Absolute Scaled Error (MASE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	0.649299232824065	0.8206454105239567	0.7424203130653794	0.7535871426066171
h_rain	0.6454064816943108	0.8244807178182957	0.7472783019876008	0.7633882904716909
alpha	0.6653643215571101	0.8149204999119286	0.7533600295608973	0.7489582262388157

Linear Regression | Neural Networks - Standart Scaler

Mean Absolute Percentage Error (MAPE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	11.81220474218926	31.4195055714978	23.917158413268265	25.649218596490662
h_rain	11.660333584316156	31.323282336401938	24.0778589158509	25.746360006992163
alpha	11.37657797950784	31.331917532867614	24.180259572647238	25.589224866097414

Mean Absolute Scaled Error (MASE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	0.3526448561818384	0.9439576126193457	0.7140295205687434	0.7705969496177801
h_rain	0.3478381189337385	0.9529350092891233	0.7182137484486133	0.7832706530857918
alpha	0.3403656344377705	0.929750024172946	0.7234275021134335	0.7593401333596426

Linear Regression | Neural Networks - MimMaxScaler

Mean Absolute Percentage Error (MAPE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	24.346112038009178	25.338765043165008	23.917158413267988	25.64921859649079
h_rain	24.504316948472464	25.3316915674618	24.077858915850932	25.746360006991935
alpha	24.442620204451345	25.53891629589814	24.180259572647063	25.58922486609759

Mean Absolute Scaled Error (MASE)

Metric	Neural Network (Train)	Neural Network (Test)	Linear Regression (Train)	Linear Regression (Test)
phic	0.7268356217672061	0.761269781997029	0.714029520568735	0.7705969496177839
h_rain	0.7309344817677838	0.7706553700183288	0.7182137484486141	0.7832706530857847
alpha	0.7312768345802221	0.7578472660842894	0.7234275021134285	0.759340133359648

Random Forest - StandarScaler

Mean Absolute Percentage Error (MAPE)

Metric	Random Forest (Train)	Random Forest (Test)
phic	9.364441743620073	25.395726537989766
h_rain	9.377166628219616	25.41092632771062
alpha	9.494532551318882	25.6092293818387

Mean Absolute Scaled Error (MASE)

Metric	Random Forest (Train)	Random Forest (Test)
φ_c	0.28	0.77
h_r	0.28	0.77
α	0.28	0.77

CONCLUSIONS

The database may not be large enough

Correlations between variables and target variables are not high.

The numbers of features is not really big

REFERENCES

https://github.com/wavestoweather/SWM_largeensemble

https://keras.io/examples/keras_recipes/bayesian_neural_networks/

https://stackoverflow.com/questions/51237635/difference-between-standard-scaler-and-minmaxscaler

https://www.pmorgan.com.au/tutorials/mae,-mape,-mase-and-the-scaled-rmse/

THANK YOU FOR YOUR ATTENTION!