Table 5.1: Results of the ablation study of the Machine learning architecture

	Table 5.1:	Results of	of the al	olation s	tudy of t	he Με	achine lear	ning arc	hitecture.	
ID	Config.	MSE	RMSE	MAE	MAPE	R^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
1	Linear Regression:	141776.29	376.53	252.09	0.87	1	1.26	0.88	-60.543908	0.86
2	Ridge Regression: alpha=0.01	141877.73	376.67	252.12	0.87	1	1.26	0.19	-60.413898	0.86
3	Ridge Regression: alpha=0.1	143002.98	378.16	252.65	0.88	1	1.26	0.2	-59.291826	0.86
4	Ridge Regression: alpha=1	158254.53	397.81	263.42	0.91	1	1.31	0.21	-53.272749	0.88
5	Ridge Regression: alpha=10	265867.72	515.62	346.11	1.2	0.99	1.66	0.18	-45.975812	0.92
6	Lasso Regression: alpha=0.01	4796216.90	2190.03	1859.82	7.23	0.87	7.17	3.27	15.625081	1.01
7	Lasso Regression: alpha=0.1	51073484.	1 2 146.57	6004.68	23.43	-0.34	24.64	0.19	-20.620023	1
8	Lasso Regression: alpha=1	51073484.	172146.57	6004.68	23.43	-0.34	24.64	0.13	-20.620023	1
9	Random Forest: max_depth=5, min_samples_spli n_estimators=50	_t <u>363,</u> 156.9	602.63	435.79	1.48	0.99	1.94	155.66	-36.752922	0.94
10	Random Forest: max_depth=5, min_samples_spli n_estimators=100	_t <u>366</u> 169.09	605.12	435.92	1.48	0.99	1.95	289.49	-33.181601	0.95
11	Random Forest: max_depth=5, min_samples_spli n_estimators=200	_t <u>35</u> 8,232.25	598.53	431.07	1.46	0.99	1.93	597.8	-38.577704	0.94
12	Random Forest: max_depth=5, min_samples_spli n_estimators=50	_t <u>36</u> 1,047.26	600.87	432.52	1.47	0.99	1.94	149.53	-35.486951	0.94
13	Random Forest: max_depth=5, min_samples_spli n_estimators=100	_t <u>36</u> 1,284.02	601.07	433.51	1.47	0.99	1.94	293.46	-36.394789	0.94
14	Random Forest: max_depth=5, min_samples_spli n_estimators=200	_t <u>35</u> 6002.1	596.66	429.94	1.46	0.99	1.93	598.08	-40.202596	0.94
15	Random Forest: max_depth=5, min_samples_spli n_estimators=50	_t <u>35</u> ₁₇ 3,28.77	597.77	428.27	1.45	0.99	1.91	147.14	-36.857328	0.94
16	Random Forest: max_depth=5, min_samples_spli n_estimators=100	_t <u>36</u> 7,999.61	606.63	436.9	1.48	0.99	1.95	290.59	-32.860473	0.95
17	Random Forest: max_depth=5, min_samples_spli n_estimators=200	±3576,72.33	598.06	429.28	1.45	0.99	1.92	627.38	-37.410473	0.94
18	Random Forest: max_depth=10, min_samples_spli n_estimators=50	_t <u>150,</u> 481.85	387.92	267.59	0.93	1	1.29	269.58	99.428901	1.34
19	Random Forest: max_depth=10, min_samples_spli n_estimators=100	_t 150,345.15	387.74	267.28	0.93	1	1.28	529.02	101.082790	1.35

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1		01 0110 00	91001011 5	taa, or c	7110 1110		Training		
ID		MOD	DAGE	3.645	MADE	D.2	CLIDA (C	O	Overfit	Loss
ID	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	E Time	Gap	Ratio
								(s)		
20	Random Forest: max_depth=10, min_samples_sp n_estimators=2	olit ¹⁴ 8,797.62	385.74	265.66	0.92	1	1.29	1066.13	99.690336	1.35
21	Random Forest: max_depth=10, min_samples_sp n_estimators=5	olit <u>15</u> 1,491.27	389.22	267.65	0.93	1	1.3	261.08	96.190329	1.33
22	Random Forest: max_depth=10, min_samples_sp n_estimators=10	olit=3,222.53	386.29	266.1	0.92	1	1.29	527.68	94.224497	1.32
23	Random Forest: max_depth=10, min_samples_sp n_estimators=20	149435.28	386.57	266.92	0.92	1	1.29	1058.31	95.272409	1.33
24	Random Forest: max_depth=10, min_samples_sp n_estimators=5	$_{0}^{\mathrm{blit}\stackrel{1500}{=}10,48.05}$	387.36	266.87	0.92	1	1.29	274.54	91.887739	1.31
25	Random Forest: max_depth=10, min_samples_sp n_estimators=10	149296.06 oo	386.39	265.57	0.92	1	1.29	548.87	92.885181	1.32
26	Random Forest: max_depth=10, min_samples_sp n_estimators=20	148903.82 oo	385.88	265.81	0.92	1	1.29	1080.86	91.662446	1.31
27	Random Forest: max_depth=15, min_samples_sp n_estimators=5	olit <u>142,</u> 872.21	377.98	260.28	0.9	1	1.27	369.28	200.962790	2.14
28	Random Forest: max_depth=15, min_samples_sp n_estimators=10	olit <u>143,</u> 612.18	378.96	260.68	0.9	1	1.27	759.03	203.683264	2.16
29	Random Forest: max_depth=15, min_samples_sp n_estimators=20	olit=2,150.81	377.03	258.96	0.9	1	1.26	1498.74	202.137978	2.16
30	Random Forest: max_depth=15, min_samples_sp n_estimators=5	olit <u>144</u> 728.8	380.43	262.72	0.91	1	1.28	367.5	188.572266	1.98
31	Random Forest: max_depth=15, min_samples_sp n_estimators=10	olit <u>143,</u> 400.99	378.68	261.35	0.91	1	1.27	711.89	189.268092	2
32	Random Forest: max_depth=15, min_samples_sp n_estimators=20	olit <u>142,</u> 892.33	378.01	259.45	0.9	1	1.26	1378.22	187.875811	1.99
33	Random Forest: max_depth=15, min_samples_spacestimators=50	$_{ m blit}_{ m 0}^{146923.82}$	383.31	263.56	0.91	1	1.28	335.28	176.322837	1.85
34	Random Forest: max_depth=15, min_samples_sp n_estimators=1	$_{ m blit}$ $\frac{1445}{10}$ $\frac{92.12}{00}$	380.25	260.97	0.9	1	1.27	673.28	175.653349	1.86
35	Random Forest: max_depth=15, min_samples_sp n_estimators=2	$_{ m blit}_{ m 50}^{ m 143289.95}_{ m 500}$	378.54	260.08	0.9	1	1.26	1340.55	172.847759	1.84
36	Gradient Boosting: learning_rate=0 max_depth=3, n_estimators=50 subsample=0.7		05810.29	5719.46	22.32	-0.22	23.5	98.95	-18.215639	1
37	Gradient Boosting: learning_rate=0 max_depth=3, n_estimators=5 subsample=0.8		16810.03	5719.32	22.32	-0.22	23.5	114.19	-18.472051	1
38	Gradient Boosting: learning_rate=0 max_depth=3, n_estimators=5 subsample=1.0		25 809.97	5719.24	22.32	-0.22	23.5	142.6	-18.558061	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

								Training		
ID	~ ^	MCE	DMCE	MAE	MADE	ъ2	CVDMC	J	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	E Time	Gap	Ratio
								(s)		
	Gradient Boosting:									
39	learning_rate=	0.0042129249	.48490.7	5448.18	21.27	-0.11	22.42	198.37	-15.783533	1
00	max_depth=3, n_estimators=:		.1010011	0110110		0.11		100.0.	1000000	-
	subsample=0.7 Gradient									
	Boosting:									
40	learning_rate= max_depth=3,	0.0042128905	.96490.68	5448.14	21.27	-0.11	22.42	226.88	-15.840599	1
	n_estimators= subsample=0.8	100,								
	Gradient									
41	Boosting: learning_rate= max_depth=3,	0.00/191506	: ngunn en	5448.18	21.27	0.11	22.42	201 16	15 620222	1
41	max_depth=3, n_estimators=		.U#19U.09	3440.10	21.27	-0.11	22.42	284.46	-15.639223	1
	subsample=1.0	100,								
	Gradient Boosting:									
42	learning_rate= max_depth=3,	0.0034816123	3. 75 9900.52	4946.47	19.32	0.09	20.42	395.49	-8.176807	1
	n_estimators=									
	subsample=0.7 Gradient									
40	Boosting:	0.000 400 500.0	0F1001 F0	40.45.01	10.00	0.00	00.49	450.00	E 050410	4
43	learning_rate= max_depth=3,		.6%1901.52	4947.01	19.32	0.08	20.43	453.63	-7.272416	1
	n_estimators=: subsample=0.8	200,								
	Gradient Boosting:									
44	learning_rate=	0.0034824996	5.050901.27	4946.88	19.32	0.08	20.43	567.59	-7.442802	1
	n_estimators=									
	subsample=1.0 Gradient									
	Boosting:	0.001								
45	learning_rate= max_depth=5,		5.0 6 800.49	5715.25	22.3	-0.22	23.44	155.66	-20.350767	1
	n_estimators= subsample=0.7	50,								
	Gradient									
46	Boosting: learning_rate= max_depth=5.	0.0046242079	.281800.15	5714.98	22.3	-0.22	23.44	179.49	-20.683194	1
10	max_depth=5, n_estimators=			0.11.00		0.22	20.11	1.0.10	20.000101	-
	subsample=0.8									
	Gradient Boosting:									
47	learning_rate= max_depth=5,	0.00 <mark>4</mark> 6241654	.55800.12	5714.9	22.3	-0.22	23.44	225.12	-20.736567	1
	n_estimators= subsample=1.0	50,								
	Gradient									
48	Boosting: learning_rate= max_depth=5.	0.00/1280011	949/171 55	5440.06	21.22	-0.1	22.3	312.85	-19.965498	1
40	max_depth=5, n_estimators=:			5440.00	21.22	-0.1	22.5	312.00	-19.905496	1
	subsample=0.7	,								
	Gradient Boosting:									
49	learning_rate= max_depth=5,	$^{0.00}41882237$	7. 62 471.65	5440.21	21.23	-0.1	22.3	357.9	-19.874906	1
	n_estimators=									
	subsample=0.8 Gradient									
F0	Boosting: learning_rate= max_depth=5.	0.004, 05,4550	00471 05	F 490 F	01.00	0.1	00.0	450.1	00 440707	4
50	max_depth=5, n_estimators=		.00%4/1.07	5439.7	21.22	-0.1	22.3	450.1	-20.446725	1
	n_estimators= subsample=1.0	100,								
	Gradient Boosting:									
51	learning_rate=	0.0034364793	3.75862.15	4930.2	19.23	0.1	20.19	622.69	-18.375008	1
	n_estimators=									
	subsample=0.7 Gradient									
	Boosting:	0.004	omo o o : :	1000 ==	10.55		00.1-	H 1 -	10.051	_
52	learning_rate= max_depth=5,		.8862.18	4930.23	19.23	0.1	20.19	715.86	-18.351626	1
	n_estimators=: subsample=0.8	200,								

Table 5.1: Results of the ablation study of the Machine learning architecture.

					- 0		Training	Overfit	Loss
ID Config.	\mathbf{MSE}	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMSE	E Time	Gap	Ratio
							(s)	Сир	
Gradient									
Boosting: learning_rate	=0.00 1 4363942	2.65862.08	4930.21	19.23	0.1	20.19	900.78	-18.405702	1
max_depth= n_estimators			1000.21	10.20	0.1	20.10	000.10	10.100102	1
subsample=1. Gradient	0								
Boosting:	-0.001								
		2.562800.13	5714.21	22.3	-0.22	23.44	209.96	-19.057124	1
n_estimators subsample=0.									
Gradient Boosting:									
55 learning_rate	=0.00 1 6242396	.86800.18	5714.27	22.3	-0.22	23.44	240.27	-18.987459	1
max_depth= n_estimators			0,11,2,		0.22	20.11	_10	10.001.100	-
subsample=0. $Gradient$	8								
Boosting:	-0.001								
		.26800.2	5714.28	22.3	-0.22	23.45	301.72	-18.959817	1
$n_{estimators}$ $subsample=1$									
Gradient Boosting:									
57 learning_rate max_depth=	=0.0041871123	3.4 6 3470.79	5438.09	21.22	-0.1	22.31	418.51	-17.392053	1
n_estimators	=100,								
subsample=0. Gradient									
Boosting: learning rate	=0.00\ddrag{1870746}	900470.76	E 420 07	01 00	0.1	00.21	485.57	17 416664	1
58 max_depth= n_estimators		1.300±10.10	5438.07	21.22	-0.1	22.31	400.07	-17.416664	1
subsample=0.									
Gradient Boosting:									
59 learning_rate max_depth=	=0.00 <u>4</u> 1873405	.662470.97	5438.15	21.22	-0.1	22.31	603.58	-17.190756	1
n_estimators subsample=1.	=100,								
Gradient									
Boosting: 60 learning_rate	=0.00 1 4335939	.65859.69	4925.5	19.22	0.1	20.19	837.88	-14.267054	1
max_depth= n_estimators			1020.0	10.22	0.1	20.10	001.00	11.20.001	-
subsample=0. $Gradient$	7								
Boosting:	-0.001								
		3.8 58 59.79	4925.6	19.22	0.1	20.19	961.4	-14.147892	1
n_estimators subsample=0.									
Gradient Boosting:									
62 learning_rate max_depth=	=0.00 3 4340677	7.25860.09	4925.62	19.22	0.1	20.19	1210.31	-13.827036	1
n_estimators	=200,								
subsample=1. Gradient	0								
Boosting: learning_rate	=0.01 3, 19666813	9/0/19/1/79	2609 70	14.47	0.49	15 46	00.07	0.592665	1
63 max_depth= n_estimators	_{3,} 19000013 =50.		3698.79	14.47	0.48	15.46	99.07	9.583665	1
subsample=0.									
Gradient Boosting:	0.01								
		5.8 4 32.59	3696.46	14.46	0.48	15.46	113.95	7.589609	1
$n_{estimators}$ subsample=0									
Gradient Boosting:									
65 learning_rate	=0.01 19660227	.24 7 433.99	3697.76	14.47	0.48	15.46	143.21	8.856403	1
n_estimators	=50,				-	-			
subsample=1. Gradient	U								
Boosting:	=0.01 7 707468.4	49077C 00	2201 61	0.00	0.0	0.0	107 75	7 00 4000	1
66 max_depth= n_estimators		45∠110.23	2291.61	8.98	0.8	9.8	197.75	7.084020	1
subsample=0.									

Table 5.1: Results of the ablation study of the Machine learning architecture.

								Training		
ID	G 8	MCE	DMCE	MAE	MADE	$\mathbf{R^2}$	CVDM	J	Overfit	\mathbf{Loss}
ID	Config.	MSE	RMSE	MAE	MAPE	R-	CVRIVI	SE Time	Gap	Ratio
								(s)		
	Gradient Boosting:									
67	learning_rate=0.	⁰¹ 7710375.	342776.76	2291.6	8.98	0.8	9.8	227.32	7.227094	1
	n_estimators=10									
	subsample=0.8 Gradient									
	Boosting:	01								
68	learning_rate=0. max_depth=3,		312775.24	2290.29	8.98	0.8	9.8	285.47	6.065712	1
	n_estimators=10 subsample=1.0	10,								
	Gradient Boosting:									
69	learning_rate=0. max_depth=3,	⁰¹ 1416068.	851189.99	957.21	3.73	0.96	4.17	396.34	-15.960600	0.99
	n_estimators=20									
	subsample=0.7 Gradient									
70	Boosting:	01, 41,5000	F01100 00	050.05	0.70	0.00	4.15	155 15	10 500000	0.00
70	learning_rate=0. max_depth=3,		591189.66	956.65	3.72	0.96	4.17	455.45	-16.590333	0.99
	n_estimators=20 subsample=0.8	10,								
	Gradient Boosting:									
71	learning_rate=0. max_depth=3,	⁰¹ 1417310.	311190.51	956.73	3.73	0.96	4.18	572.3	-15.502716	0.99
	n_estimators=20 subsample=1.0	0,								
	Gradient									
72	Boosting: learning_rate=0. max_depth=5.	04803245/	1 //8251 1/	3664.02	14.29	0.5	14.97	156.58	-15.380127	1
12	max_depth=5, n estimators=50		1.16001.11	0004.02	14.20	0.0	14.01	100.00	-10.000121	1
	subsample=0.7	•								
	Gradient Boosting:	0.1								
73	learning_rate=0. max_depth=5,		6.545351.17	3664.95	14.3	0.5	14.97	179.36	-15.036305	1
	n_estimators=50 subsample=0.8),								
	Gradient Boosting:									
74	learning_rate=0. max_depth=5,	⁰¹ 18931482	2.84351.03	3663.45	14.29	0.5	14.97	226.24	-15.375018	1
	n_estimators=50									
	subsample=1.0 Gradient									
75	Boosting: learning_rate=0. max_depth=5.	017000075	692664 27	2245.27	9.76	0.91	0.15	222.01	19 910649	0.00
75	max_depth=5, n_estimators=10		0&004.57	2243.21	8.76	0.81	9.15	323.91	-13.810648	0.99
	subsample=0.7	•,								
	Gradient Boosting:									
76	learning_rate=0. max_depth=5,	⁰¹ 7103358.	272665.21	2245.33	8.76	0.81	9.15	360.31	-12.869325	1
	n_estimators=10 subsample=0.8	00,								
	Gradient Boosting:									
77	learning_rate=0.	⁰¹ 7099672.	2 2664.52	2245.52	8.76	0.81	9.15	453.55	-13.556401	0.99
	n_estimators=10									
	subsample=1.0 Gradient									
70	Boosting: learning_rate=0. max_depth=5.	01,1141==	7/1055 54	967 79	9 90	0.07	9.50	607 50	15 000412	0.00
78	max_depth=5, n_estimators=20		(41U05.54	867.72	3.39	0.97	3.56	627.53	-15.288416	0.99
	subsample=0.7	~,								
	Gradient Boosting:	0.1								
79	learning_rate=0. max_depth=5,		741054.8	865.86	3.38	0.97	3.56	721.64	-15.642747	0.99
	n_estimators=20 subsample=0.8	00,								
	Gradient									
80	Boosting: learning_rate=0. max_depth=5	⁰¹ 113624.	471055.28	865.72	3.38	0.97	3.56	906.73	-15.216075	0.99
	n_estimators=20									
	subsample=1.0									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:	10000100	01 0110 0	51441011 8	tuaj or (,110 1,11		Training		
						- 2		Ü	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSI	E Time	Gap	Ratio
								(s)	P	
	Gradient									
81	Boosting: learning_rate=0.0 max_depth=7,	⁰¹ 18889500	.48346.21	3656.32	14.27	0.5	14.96	210.14	-5.388431	1
01	max_depth=7, n_estimators=50,		7.16010.21	0000.02	11.21	0.0	11.00	210.11	0.000101	-
	subsample=0.7 Gradient									
	Boosting.	31								
82	learning_rate=0.0 max_depth=7,		1.64/346.17	3656.49	14.27	0.5	14.96	241.35	-5.466385	1
	n_estimators=50, subsample=0.8	,								
	Gradient Boosting:									
83	learning_rate=0.0	01 18890316	5.24346.3	3655.96	14.27	0.5	14.96	304.13	-5.247794	1
	n_estimators=50,									
	subsample=1.0 Gradient									
0.4	Boosting: learning_rate=0.0 max_depth=7.	017070499	7 0050 4	0000 47	0.75	0.01	0.14	400.00	c 00000c	1
84	max_depth=7, n_estimators=100		7 2659.4	2239.47	8.75	0.81	9.14	420.28	6.839926	1
	subsample=0.7	~,								
	Gradient Boosting:									
85	learning_rate=0.0 max_depth=7,		132659.1	2239.95	8.75	0.81	9.13	482.7	6.615332	1
	n_estimators=100 subsample=0.8	0,								
	Gradient Boosting:									
86	learning_rate=0.0	⁰¹ 7069960.	872658.94	2239.64	8.75	0.81	9.13	607.95	6.355840	1
	n_estimators=100									
	subsample=1.0 Gradient									
97	Boosting: learning_rate=0.0	014112010	21105405	990.15	9.49	0.07	2.50	944.9	22 002062	1.09
87	max_depth=7, n_estimators=200		эшоэ4.9э	880.15	3.43	0.97	3.59	844.2	32.092962	1.03
	subsample=0.7	~,								
	Gradient Boosting:									
88	learning_rate=0.0 max_depth=7,		991056.86	882.4	3.44	0.97	3.6	968.28	33.778811	1.03
	n_estimators=200 subsample=0.8	0,								
	Gradient Boosting:									
89	learning_rate=0.0 max_depth=7,	⁰¹ 1120141.	121058.37	883.43	3.44	0.97	3.6	1215.36	35.006641	1.03
	n_estimators=200									
	subsample=1.0 Gradient									
90	Boosting: learning_rate=0.1 max_depth=3.	1,045600 0	4 405.6	355.47	1.27	0.99	1 50	99.95	23.516872	1.05
90	max_depth=3, n_estimators=50,		4 495.0	355.47	1.21	0.99	1.58	99.90	23.310672	1.05
	subsample=0.7 Gradient									
	Boosting:	1								
91	learning_rate=0.1 max_depth=3,		9 480.24	340.7	1.21	0.99	1.56	114.35	11.762571	1.03
	n_estimators=50, subsample=0.8	,								
	Gradient Boosting:									
92	learning_rate=0.1 max_depth=3,	¹ ,242489.4	3 492.43	352.91	1.26	0.99	1.59	143.83	21.636576	1.05
	n_estimators=50, subsample=1.0									
	Gradient									
93	Boosting: learning_rate=0.1 max_depth=3	1, ₂₁₉₉₁₂₋₃	6 468 95	338.32	1.19	0.99	1.46	203.49	55.273062	1.13
00	n_estimators=100		. 100.00	555.62	1.10	0.00	1.10	200.40	30.210002	1.10
	subsample=0.7 Gradient									
6.1	Boosting:	l 000 = 01 =	a 4 5 0.00	0.46.01	1 22	0.00		0.45	00 10 150-	1.10
94			б 478.26	349.31	1.23	0.99	1.47	247.45	66.134598	1.16
	n_estimators=100 subsample=0.8	σ,								
94	Gradient Boosting: learning_rate=0.1 max_depth=3,		6 478.26	349.31	1.23	0.99	1.47	247.45	66.134598	1.16

Table 5.1: Results of the ablation study of the Machine learning architecture.

		1: Results			<u> </u>			Training		
ID		MOD	DAGE	24.5	MADE	D.2	CLIDAG	J	Overfit	Loss
ID	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	SE Time	Gap	Ratio
								(s)		
	Gradient Boosting:									
95	learning_rate=0 max_depth=3,	0.1,222713.04	471.92	344.19	1.22	0.99	1.48	290.58	61.225109	1.15
	n_estimators=1									
	subsample=1.0 Gradient									
96	Boosting: learning_rate=(max_depth=3.	0.1,200065 06	150 11	330.7	1 16	0.99	1.43	411.6	95 471697	1.23
90	max_depth=3, n_estimators=2		406.11	550.7	1.16	0.99	1.45	411.6	85.471627	1.23
	subsample=0.7	,								
	Gradient Boosting:	. 1								
97	learning_rate=0 max_depth=3,		456.08	327.13	1.14	0.99	1.41	468.07	79.509514	1.21
	n_estimators=2 subsample=0.8	00,								
	Gradient Boosting:									
98	learning_rate=0 max_depth=3,	0.1,204400.01	452.11	328.95	1.16	0.99	1.43	583.41	81.585657	1.22
	n_estimators=2									
	subsample=1.0 Gradient									
99	Boosting: learning_rate=(max_depth=5.	0.1,161760 79	409.91	285.06	1	1	1.31	158.63	49 916511	1.12
99	max_depth=5, n_estimators=5		402.21	200.00	1	1	1.51	100.00	42.316511	1.12
	subsample=0.7 Gradient									
	Boosting:	. 1								
100	learning_rate=0 max_depth=5,		404.33	288.67	1.01	1	1.33	181.83	45.607891	1.13
	n_estimators=5 subsample=0.8	0,								
	Gradient Boosting:									
101	learning_rate=0 max_depth=5,).1,160477.82	400.6	286.34	1	1	1.3	227.92	45.264500	1.13
	n_estimators=5 subsample=1.0	0,								
	Gradient									
102	Boosting: learning_rate=(max_depth=5).1, ₁₅₈₂₂₁₋₇₀	397 77	283.76	0.99	1	1.29	324.98	72.869016	1.22
102	max_depth=5, n_estimators=1		001.11	200.10	0.55	1	1.20	024.00	12.005010	1.22
	subsample=0.7 Gradient									
	Boosting:) 1								
103	learning_rate=0 max_depth=5,		388.03	275.59	0.95	1	1.26	372.19	66.049842	1.21
	n_estimators=1 subsample=0.8	00,								
	Gradient Boosting:									
104	learning_rate=0 max_depth=5,	^{).1,} 153345.71	391.59	277.08	0.95	1	1.26	467.95	70.826350	1.22
	n_estimators=1 subsample=1.0	00,								
	Gradient									
105	Boosting: learning_rate=(max_depth=5.).1, ₁₆₂₃₄₃ 78	402.92	285.58	0.99	1	1.32	664.75	105.238828	1.35
100	n_estimators=2		102.02	200.00	0.00		1.02	001.10	100.200020	1.00
	subsample=0.7 Gradient									
100	Boosting: learning_rate=0	0.1.150005 4	801.22	200.00	0.05	4	1.00	F04 55	00 410044	1.04
106	max_depth=5, n estimators=2		391.26	280.08	0.97	1	1.28	764.57	99.419344	1.34
	$\overline{\text{subsample}}=0.8$,								
	Gradient Boosting:									
107	learning_rate=0 max_depth=5,		387.74	273.53	0.94	1	1.27	959.99	93.267642	1.32
	n_estimators=2 subsample=1.0	00,								
	Gradient Boosting:									
108	learning_rate=0).1, _{131377.87}	362.46	255.67	0.9	1	1.21	212.7	78.442396	1.28
	n_estimators=5									
	subsample=0.7									

Table 5.1: Results of the ablation study of the Machine learning architecture.

								Training		_
ID	Config.	MSE	RMSE	MAE	MAPE	$\mathbf{R^2}$	CVRMS	E Time	Overfit	Loss
	J							(s)	Gap	Ratio
	Gradient Boosting:									
109	learning_rate=0.1 max_depth=7,	,136219.14	369.08	259.16	0.91	1	1.23	242.84	87.598024	1.31
	n_estimators=50, subsample=0.8									
	Gradient									
110	Boosting: learning_rate=0.1 max_depth=7.	,133422.15	365.27	256.21	0.9	1	1.22	306.42	85.556633	1.31
	n_estimators=50,									
	subsample=1.0 Gradient									
111	Boosting: learning_rate=0.1 max_depth=7,	··130675-73	361 49	253.63	0.88	1	1.21	442.19	107.118349	1.42
111	n_estimators=100		001.40	200.00	0.00	1	1.21	442.10	101.110040	1.42
	subsample=0.7 Gradient									
112	Boosting: learning_rate=0.1 max_depth=7.	1.41 E €0. 99	276 25	262 75	0.02	1	1.29	E00 E0	109 400410	1.40
112	max_depth=7, n_estimators=100		310.20	263.75	0.93	1	1.29	508.52	123.492412	1.49
	subsample=0.8 Gradient									
119	Boosting: learning_rate=0.1 max_depth=7.	1.99600.00	205.05	055.00	0.0	1	1.04	C2C 20	117 000004	1 47
113	max_depth=7, n_estimators=100		305.05	255.23	0.9	1	1.24	636.38	117.666924	1.47
	subsample=1.0 Gradient									
114	Boosting: learning_rate=0.1	197905 0	270 55	200.02	0.01	1	1.07	015 49	1.49.75.0001	1.69
114	max_depth=7, n_estimators=200		370.55	260.63	0.91	1	1.27	915.43	143.758091	1.63
	subsample=0.7 Gradient	,								
	Boosting:	100000 10		25.4.24	0.00			1010	101000010	
115	learning_rate=0.1 max_depth=7, n_estimators=200		359.57	254.01	0.89	1	1.21	1049.77	134.280218	1.6
	subsample=0.8	·,								
	Gradient Boosting:									
116	learning_rate=0.1 max_depth=7, n_estimators=200		365.93	255.71	0.9	1	1.25	1328.25	139.773314	1.62
	subsample=1.0	,,								
	XGBoost: gamma=0,	001								
117	learning_rate=0.0 max_depth=3, n_estimators=50,		2 8 810.1	5719.34	22.32	-0.22	23.5	0.92	-18.572252	1
	subsample=0.7									
	XGBoost: gamma=0, learning_rate=0.0	001								
118	learning_rate=0.0 max_depth=3, n_estimators=50,		76 810.1	5719.31	22.32	-0.22	23.5	0.42	-18.549472	1
	subsample=0.8									
	XGBoost: gamma=0, learning_rate=0.0	001								
119	learning_rate=0.0 max_depth=3, n_estimators=50,		26 810.01	5719.26	22.32	-0.22	23.5	0.4	-18.625938	1
	subsample=1.0									
100	XGBoost: gamma=0, learning_rate=0.0 max_depth=3,	012122162	00100.00	F 4 40 0 5	01.05	0.11	22.42	0.00	15 050110	
120	max_depth=3, n_estimators=100		09 490.93	5448.27	21.27	-0.11	22.42	0.63	-15.856118	1
	subsample=0.7	•								
101	XGBoost; gamma=0, learning_rate=0.0 max_depth=3,	040105014	4014 O 1 1 7	F 4 4 0 B 0	01.07	0.11	00.40	0.6	15 550005	-1
121	max_depth=3, n_estimators=100		464491.17	5448.39	21.27	-0.11	22.42	0.6	-15.578925	1
	subsample=0.8									
100	XGBoost: gamma=0, learning_rate=0.0 max_depth=3,	00/195100	96401 16	5448.3	91.97	0.11	22.42	0.66	15 554470	1
122	max_depth=3, n_estimators=100	42133198.),	ω491.10	5448.3	21.27	-0.11	22.42	0.66	-15.554479	1
	subsample=1.0									

101

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:	rtcsurts	or the a	Diation 5	tudy of t	JIC IVI	acimic ica		intecture.	
								Training	Overfit	$_{ m Loss}$
ID	Config.	MSE	\mathbf{RMSE}	MAE	MAPE	$\mathbf{R^2}$	CVRMS	E Time	Gap	Ratio
								(s)	Gap	itatio
	XGBoost:									
102	gamma=0, learning_rate=0.0 max_depth=3,	044022201	70001.04	40 <i>46</i> 79	10.22	0.00	20.42	1.00	0.045025	1
123	max_depth=3, n_estimators=200		1.73901.04	4946.72	19.32	0.08	20.43	1.02	-8.245035	1
	$\overline{\text{subsample}}=0.7$,								
	XGBoost: gamma=0,									
124	learning_rate=0.0 max_depth=3,	34825344	4.0 9 901.3	4946.8	19.32	0.08	20.43	1.02	-7.878553	1
	n_estimators=200 subsample=0.8									
	XGBoost: gamma=0,									
125	learning_rate=0.0 max_depth=3,	001 34827754	1.745901.5	4946.81	19.32	0.08	20.43	1	-7.587614	1
120	max_depth=3, n_estimators=200),	1.14001.0	4040.01	10.02	0.00	20.40	1	-1.001014	1
	subsample=1.0 XGBoost;									
		001								
126	learning_rate=0.0 max_depth=5,	46251683	3.66800.86	5715.44	22.3	-0.22	23.44	0.63	-20.471576	1
	n_estimators=50, subsample=0.7									
	XGBoost: gamma=0,									
127	learning_rate=0.0 max_depth=5,	46250699	0.65800.79	5715.42	22.3	-0.22	23.44	0.61	-20.502796	1
	n_estimators=50,									
	subsample=0.8 XGBoost;									
100	gamma=0, learning_rate=0.0 max_depth=5,	01401722	000000 = 4	E71E 10	20.2	0.22	22.44	0.66	20 607707	1
128	max_depth=5, n_estimators=50,		3.99600.34	5715.18	22.3	-0.22	23.44	0.66	-20.697797	1
	subsample=1.0									
	XGBoost: gamma=0,	201								
129	learning_rate=0.0 max_depth=5,	41892493	3.66472.44	5440.74	21.23	-0.1	22.31	1.07	-20.007426	1
	n_estimators=100 subsample=0.7),								
	XGBoost:									
130	learning_rate=0.0	$^{001}_{41890915}$	6472.32	5440.68	21.23	-0.1	22.31	1.01	-20.043158	1
	n_estimators=100),				-		-		
	subsample=0.8 XGBoost:									
101	XGBoost: gamma=0, learning_rate=0.0 max_depth=5,	001,	100450 10	F 4 40 F0	01.00	0.1	00.01	1.00	20.102220	4
131	max_depth=5, n_estimators=100		3.169472.16	5440.59	21.23	-0.1	22.31	1.02	-20.102339	1
	subsample=1.0	,,								
	XGBoost: gamma=0,									
132	learning_rate=0.0 max_depth=5,	⁰⁰ 34376909	0.656863.18	4931.06	19.24	0.1	20.2	2.71	-19.093021	1
	n_estimators=200 subsample=0.7),								
	XGBoost: gamma=0,									
133	learning_rate=0.0	$\frac{001}{3}4374207$	7.15/862.95	4930.89	19.24	0.1	20.2	1.94	-19.141245	1
-55	n_estimators=200									
	subsample=0.8 XGBoost:									
104	XGBoost: gamma=0, learning_rate=0.0	001.05.4000	. #F0.60.0F	4001 1	10.04	0.1	20.10	1.00	10.000004	4
134	max_depth=5, n_estimators=200		8.43862.95	4931.1	19.24	0.1	20.19	1.82	-18.896324	1
	$\overline{\text{subsample}}=1.0$,,								
	XGBoost: gamma=0,									
135	learning_rate=0.0 max_depth=7,	46239824	1.5%799.99	5714.01	22.3	-0.22	23.44	1.06	-20.422145	1
	n_estimators=50, subsample=0.7									
	XGBoost: gamma=0,									
136	learning_rate=0.0	001 46238207	7.367799-87	5713.93	22.29	-0.22	23.44	1.06	-20.442522	1
100	n_estimators=50,			3110.00	22.20	0.22	20.11	1.00	20.112022	_
	subsample=0.8									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:							Training		
ID	G S	Mee	DMen	MAE	MADE	\mathbb{R}^2	CVDMS	_	Overfit	Loss
ID	Config.	MSE	RMSE	WAL	MAPE	n-	CVRMS		Gap	Ratio
								(s)		
	XGBoost: gamma=0,									
137	learning_rate=0.0	$^{001}_{46242483}$	3.764800.18	5713.99	22.3	-0.22	23.44	1.11	-19.981506	1
	n_estimators=50		.,,,	0,-0.00						_
	subsample=1.0									
	XGBoost: gamma=0, learning rate=0.0	001								
138	learning_rate=0.0 max_depth=7,	~41864503	3.767470.28	5437.59	21.22	-0.1	22.3	1.89	-20.280690	1
	n_estimators=10 subsample=0.7	0,								
	XGBoost: gamma=0,									
139	learning_rate=0.0 max_depth=7,	$^{001}_{41862803}$	3.83470.15	5437.46	21.22	-0.1	22.3	1.88	-20.231068	1
	n_estimators=10	0,								
	subsample=0.8 XGBoost:									
1.40	XGBoost: gamma=0, learning_rate=0.0 max_depth=7,	001	00150 10	F 40 F 01	01.00	0.1	22.2	1.00	10.004490	
140	max_depth=7, n_estimators=10		2.09470.48	5437.61	21.22	-0.1	22.3	1.92	-19.624430	1
	subsample=1.0	0,								
	XGBoost; gamma=0,									
141	learning_rate=0.0 max_depth=7,	0034321408	35858.45	4924.45	19.22	0.1	20.19	3.58	-20.043108	1
	n_estimators=20 subsample=0.7	0,								
	XGBoost:									
142	gamma=0, learning_rate=0.0 max_depth=7,	00 1 4320044	85858 33	4924.31	19.22	0.1	20.19	3.54	-19.785785	1
142	max_depth=7, n_estimators=20	0,		4324.51	13.22	0.1	20.13	5.54	-13.703703	1
	subsample=0.8	,								
	XGBoost: gamma=0,	0.01								
143	learning_rate=0.0 max_depth=7,		.05858.56	4923.96	19.22	0.1	20.19	3.7	-18.995445	1
	n_estimators=20 subsample=1.0	0,								
	XGBoost: gamma=0,									
144	learning_rate=0.0	⁰¹ 19666296	5.248434.67	3698.43	14.47	0.48	15.46	0.39	8.451086	1
	n_estimators=50	,								
	subsample=0.7 XGBoost:									
1.45	XGBoost: gamma=0, learning_rate=0.0 max_depth=3,	01,004,7401	14490 55	9.607	1.4.40	0.40	15 45	0.4	6.400050	1
145	max_depth=3, n_estimators=50		.144432.55	3697	14.46	0.48	15.45	0.4	6.480059	1
	subsample=0.8	,								
	XGBoost: gamma=0,									
146	learning_rate=0.0 max_depth=3,	⁰¹ 19681157	7.544436.35	3698.55	14.47	0.48	15.47	0.41	10.309046	1
	n_estimators=50 subsample=1.0									
	XGBoost:									
147	gamma=0, learning_rate=0.0 max_depth=3,	01 ₇₇₀₅₉₁₄	112775 95	2290.37	8.98	0.8	9.8	0.61	5.042698	1
	n_estimators=10		112		0.00	0.0	0.0	0.01	0.012000	-
	subsample=0.7									
	XGBoost: gamma=0, learning rate=0.	01		220: =:		0 -	0 -	0 -	F 00=	
148	learning_rate=0.0		952776.81	2291.52	8.98	0.8	9.8	0.6	5.807418	1
	n_estimators=10 subsample=0.8	υ,								
	XGBoost: gamma=0,									
149	learning_rate=0.0 max_depth=3,	⁰¹ 7708988.	562776.51	2290.6	8.98	0.8	9.8	0.61	5.759622	1
	n_estimators=10									
	subsample=1.0 XGBoost: gamma=0,									
150	learning_rate=0.0	04,17067	47I 100 74	056 72	9 79	0.06	110	1.06	19 697700	0.00
150	max_depth=3, n_estimators=20		411190.14	956.73	3.73	0.96	4.18	1.96	-18.627706	0.98
	subsample=0.7	,								

Table 5.1: Results of the ablation study of the Machine learning architecture.

				010001011 0	ready of e	110 1110	achine lea			
						- 0		Training	Overfit	Loss
$^{\mathrm{ID}}$	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	SE Time	Gap	Ratio
								(s)	Сир	100010
	XGBoost: gamma=0,									
151	learning_rate=0 max_depth=3,	0.01 1417431.	861190.56	956.58	3.72	0.96	4.18	1.06	-18.837539	0.98
101	n_estimators=2	00,	001100.00	000.00	52	0.00	1.10	1.00	10.00.000	0.00
	subsample=0.8									
450	XGBoost: gamma=0, learning rate=0	.01						4.04	40 4000=0	0.00
152	learning_rate=0 max_depth=3, n_estimators=2	1418742.	661191.11	957	3.73	0.96	4.18	1.01	-18.100379	0.99
	subsample=1.0	00,								
	XGBoost: gamma=0,									
153	max_depth=5,	18935015 1 ^{10.0}	5.9 9 351.44	3664.04	14.29	0.5	14.96	0.6	-18.463796	1
	n_estimators=5 subsample=0.7	0,								
	XGBoost:									
154	learning_rate=0	0.01 18933980	0.74351.32	3664.08	14.29	0.5	14.96	0.63	-18.254487	1
	n_estimators=5									
	subsample=0.8 XGBoost;									
155	gamma=0, learning_rate=0 max_depth=5,	0.01,0024024	6#9#1 49	3664.44	14.29	0.5	14.96	0.62	17 757541	1
100	max_depth=5, n_estimators=5		1.04001.42	3004.44	14.29	0.5	14.90	0.02	-17.757541	1
	subsample=1.0	,								
	XGBoost: gamma=0,	0.01								
156	learning_rate=0 max_depth=5,		492664.38	2245.73	8.76	0.81	9.14	1.02	-18.742297	0.99
	n_estimators=1 subsample=0.7	00,								
	XGBoost: gamma=0,									
157	learning_rate=0 max_depth=5,	^{0.01} 7094694.	922663.59	2244.21	8.75	0.81	9.13	1.02	-19.230075	0.99
	n_estimators=1 subsample=0.8									
	XGBoost:									
158	learning_rate=0	0.01 <mark>7</mark> 098347.	792664.27	2245.8	8.76	0.81	9.14	1.01	-17.975113	0.99
	n_estimators=1	00,			0	0.02	9.22			0.00
	subsample=1.0 XGBoost: gamma=0,									
150	gamma=0, learning_rate=0 max_depth=5,	0.01,116174	251056 40	960 96	2 20	0.07	2 57	1.0	01 074717	0.00
159	max_depth=5, n_estimators=2		551050.49	869.86	3.39	0.97	3.57	1.8	-21.074717	0.98
	subsample=0.7	,								
	XGBoost: gamma=0,	0.01								
160	learning_rate=0 max_depth=5,	7.011114230.	851055.57	867.34	3.39	0.97	3.56	1.82	-21.714502	0.98
	n_estimators=2 subsample=0.8	00,								
	XGBoost: gamma=0,									
161	learning_rate=0 max_depth=5,	^{0.01} 1113514.	591055.23	866.61	3.38	0.97	3.56	1.77	-21.579264	0.98
	n_estimators=2 subsample=1.0									
	XGBoost: gamma=0,									
162	learning_rate=0	0.01 18871005	5.54344.08	3653.82	14.27	0.5	14.96	1.08	-16.624806	1
102	n_estimators=5			0000.02	11.2.	0.0	11.00	1.00	10.02 1000	-
	subsample=0.7 XGBoost:									
1.00	XGBoost: gamma=0, learning_rate=0 max_depth=7,	0.01,0070100	0.000.44.03	9659.0	14.07	0.5	1400	1.07	15 050510	4
163	max_depth=7, n_estimators=5		0.242344.21	3653.8	14.27	0.5	14.96	1.07	-15.672513	1
	subsample=0.8	~,								
	XGBoost: gamma=0,	0.01								
164	learning_rate=0 max_depth=7,		3.7 4 343.3	3652.17	14.26	0.5	14.96	1.07	-15.514100	1
	n_estimators=5 subsample=1.0	υ,								

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:				, , , , , , , , , , , , , , , , , , ,			Training		
						- 0		J	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	E Time	Gap	Ratio
								(s)	Gap	Itatio
	XGBoost:									
1.05	gamma=0, learning_rate=0.0 max_depth=7,	01,000,400	0.00050.00	0007.64	0.74	0.01	0.14	1.00	6 506605	4
165			962658.66	2237.64	8.74	0.81	9.14	1.98	-6.706627	1
	n_estimators=100 subsample=0.7	υ,								
	XGBoost: gamma=0,									
166	learning_rate=0.0	⁰¹ 7070326.	912659.01	2237.57	8.74	0.81	9.14	1.99	-5.293910	1
	max_depth=7, n_estimators=100				0	0.0-			0.2000	
	subsample=0.8 XGBoost;									
	gamma=0	11								
167	learning_rate=0.0 max_depth=7,		112661.19	2239.3	8.75	0.81	9.15	2.09	-1.618217	1
	n_estimators=100 subsample=1.0	0,								
	XGBoost:									
168	learning_rate=0.0	⁰¹ 1133146.	451064.49	884.38	3.45	0.97	3.64	3.78	25.978215	1.03
100	max_depth=7, n_estimators=200		101001.10	001.00	0.10	0.0.	0.01	00	20.0.0210	1.00
	subsample=0.7									
	XGBoost: gamma=0,	11								
169	learning_rate=0.0 max_depth=7,		571065.85	885.8	3.46	0.97	3.65	4.82	28.505839	1.03
	n_estimators=200 subsample=0.8	0,								
	XGBoost:									
170	learning_rate=0.0 max_depth=7,	⁰¹ 135954.	271065.81	885.89	3.46	0.97	3.65	3.8	29.931985	1.03
	max_deptn=7, n_estimators=200		_,		0.20	0.0.	0.00	0.0		
	subsample=1.0 XGBoost:									
	gamma=0	1								
171	learning_rate=0.1 max_depth=3,		3 496.41	353.91	1.26	0.99	1.6	0.4	19.467531	1.04
	n_estimators=50, subsample=0.7	,								
	XGBoost: gamma=0,									
172	learning_rate=0.1 max_depth=3,	¹ ,239886.8	4 489.78	345.4	1.21	0.99	1.56	0.39	11.818338	1.02
	n_estimators=50,									
	subsample=0.8 XGBoost:									
170	XGBoost: gamma=0, learning_rate=0.1 max_depth=3,	1,004010 =	404.05	0.40.61	1.00	0.00		0.00	0.05.1000	1.01
173	max_depth=3, n_estimators=50,		4 484.37	343.61	1.22	0.99	1.57	0.38	6.974202	1.01
	subsample=1.0	,								
	XGBoost: gamma=0,									
174	learning_rate=0.1 max_depth=3,	¹ ,231046.4	1 480.67	346.92	1.22	0.99	1.51	0.55	60.112151	1.14
	n_estimators=100 subsample=0.7	0,								
	XGBoost:									
175	gamma=0, learning_rate=0.1 max_depth=3,	1,007051 0	0 177 91	949 71	1.10	0.00	1 /10	0.50	E6 6E99E9	1 19
175	max_depth=3, n_estimators=100		8 411.34	342.71	1.19	0.99	1.48	0.59	56.652358	1.13
	subsample=0.8	~,								
	XGBoost: gamma=0,									
176	learning_rate=0.1 max_depth=3,	¹ ,217122.2	5 465.96	341.46	1.21	0.99	1.48	0.57	48.219801	1.12
	n_estimators=100 subsample=1.0									
	XGBoost:									
177	gamma=0, learning_rate=0.1 max_depth=3,	1, ₂₀₄₇₇ 2 0	3 459 59	327.99	1.16	0.99	1.43	0.88	69.495350	1.18
111	max_depth=3, n_estimators=200		0 404.04	941.93	1.10	0.99	1.40	0.00	00.400000	1.10
	subsample=0.7									
	XGBoost: gamma=0,	1								
178	learning_rate=0.1 max_depth=3,		3 454.4	329.81	1.15	0.99	1.42	0.89	71.982471	1.19
	n_estimators=200 subsample=0.8	U,								
	-									

Table 5.1: Results of the ablation study of the Machine learning architecture.

							achine lear	Training		
ID		MOD	DAGE	2645	MADE	D2		J	Overfit	Loss
$^{\mathrm{ID}}$	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Сцр	100013
	XGBoost:									
179	gamma=0, learning_rate=0.1, max_depth=3,	108571 98	445.61	325.09	1.15	0.99	1.42	0.88	63.930992	1.17
113	max_depth=3, n_estimators=200,		440.01	323.03	1.10	0.33	1.42	0.00	03.930992	1.11
	subsample=1.0									
	XGBoost; gamma=0,									
180	learning_rate=0.1, max_depth=5,	161150.53	401.44	286.42	1	1	1.33	0.58	30.081291	1.08
	n_estimators=50, subsample=0.7									
	XGBoost: gamma=0,									
101	gamma=0, learning_rate=0.1, max_depth=5,	164100.04	40F 10	000.74	1.00	-1	1.94	0.0	22 000700	1.00
181	max_depth=5, n_estimators=50,	164122.04	405.12	289.74	1.02	1	1.34	0.6	33.222582	1.09
	subsample=0.8									
	XGBoost: gamma=0,									
182	learning_rate=0.1, max_depth=5,	159129.15	398.91	285.51	1	1	1.31	0.58	28.342860	1.08
	n_estimators=50,									
	subsample=1.0 XGBoost;									
	gamma=0.									
183	learning_rate=0.1, max_depth=5,		392.22	280.41	0.97	1	1.29	0.89	57.534747	1.17
	n_estimators=100, subsample=0.7									
	XGBoost:									
184	learning_rate=0.1,	155902.15	394.84	282	0.98	1	1.3	0.87	59.001192	1.18
101	max_depth=5, n_estimators=100,		001.01	202	0.00	-	1.0	0.01	00.001102	1.10
	subsample=0.8									
	XGBoost: gamma=0,									
185	learning_rate=0.1, max_depth=5,		392.32	280.85	0.97	1	1.28	0.86	59.825262	1.18
	n_estimators=100, subsample=1.0									
	XGBoost: gamma=0,									
186	learning_rate=0.1,	1/0/58 13	386 6	275.95	0.96	1	1.28	1.45	82.204145	1.27
100	max_depth=5, n_estimators=200,		300.0	210.30	0.90	1	1.20	1.40	02.204145	1.21
	subsample=0.7									
	XGBoost: gamma=0,									
187	learning_rate=0.1, max_depth=5,	154769.02	393.41	281.08	0.98	1	1.3	1.43	90.743298	1.3
	n_estimators=200,									
	subsample=0.8 XGBoost;									
100	gamma=0, learning_rate=0.1, max_depth=5,	1 40000 01	906.01	070.00	0.05	-1	1.00	1 41	09.004990	1.00
188	max_depth=5, n_estimators=200,		380.91	273.98	0.95	1	1.28	1.41	83.804228	1.28
	subsample=1.0									
	XGBoost: gamma=0,									
189	learning_rate=0.1, max_depth=7,	142512.46	377.51	264.71	0.94	1	1.28	1.09	74.577352	1.25
	n_estimators=50,									
	subsample=0.7 XGBoost;									
100	XGBoost; gamma=0, learning_rate=0.1, max_depth=7,	1 1000 1 5 =	850 85	000.00	0.05	_	1.01	1.60	F O 401046	1.00
190		142924.98	378.05	266.93	0.95	1	1.31	1.08	78.481240	1.26
	n_estimators=50, subsample=0.8									
	XGBoost: gamma=0,									
191	learning_rate=0.1,	139218.4	373.12	260.17	0.92	1	1.27	1.08	76.405197	1.26
-	n_estimators=50,			-	-		-	-	/	-
	subsample=1.0									
	XGBoost: gamma=0, learning_rate=0.1									
192	learning_rate=0.1, max_depth=7,	135909.04	368.66	255.65	0.9	1	1.26	1.74	102.778209	1.39
	n_estimators=100, subsample=0.7									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:			-	v -			Training		
		3. f G F	DIGE	2545	16455	D 2	CT ID 1 10	· ·	Overfit	\mathbf{Loss}
ID	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	SE Time	Gap	Ratio
								(s)		
	XGBoost: gamma=0,									
193	learning_rate=0.1	1,137366.73	370.63	260.02	0.92	1	1.28	2.68	107.827440	1.41
	n_estimators=100				0.0_					
	subsample=0.8 XGBoost:									
104	XGBoost: gamma=0, learning_rate=0.1 max_depth=7,		005 44	051.0	0.00		1.00		104 604406	
194	max_depth=7, n_estimators=100		365.44	251.9	0.88	1	1.26	1.77	104.634426	1.4
	subsample=1.0	,,								
	XGBoost: gamma=0,									
195	learning_rate=0.1 max_depth=7,	136673.97	369.69	257.28	0.9	1	1.26	2.88	136.850338	1.59
	n_estimators=200 subsample=0.7),								
	XGBoost:									
196	learning_rate=0.1	1,137564.29	370.9	260.64	0.92	1	1.28	2.87	141.399707	1.62
	n_estimators=200				0.0_					
	subsample=0.8 XGBoost: gamma=0,									
107	gamma=0, learning_rate=0.1 max_depth=7,	1194405 1	1 200 01	050.00	0.00	1	1.00	0.77	196 970090	1.50
197	max_depth=7, n estimators=200		1 300.01	252.98	0.88	1	1.26	2.77	136.372939	1.59
	$\overline{\text{subsample}}=1.0$	-,								
	XGBoost: gamma=0.1,	001								
198	learning_rate=0.0 max_depth=3,		.2 8 810.1	5719.34	22.32	-0.22	23.5	0.4	-18.572252	1
	n_estimators=50, subsample=0.7									
	XGBoost: gamma=0.1,									
199	learning_rate=0.0 max_depth=3,	$^{001}_{46377438}$.7 6 810.1	5719.31	22.32	-0.22	23.5	0.41	-18.549472	1
	n_estimators=50, subsample=0.8									
	XGBoost: gamma=0.1,									
200	learning_rate=0.0	001 46376202	269810.01	5719.26	22.32	-0.22	23.5	0.39	-18.625938	1
200	max_depth=3, n_estimators=50,		.20010.01	0113.20	22.02	-0.22	20.0	0.00	-10.020300	1
	subsample=1.0									
	XGBoost: gamma=0.1, learning_rate=0.0	001								
201	learning_rate=0.0 max_depth=3,		.09490.93	5448.27	21.27	-0.11	22.42	0.6	-15.856118	1
	n_estimators=100 subsample=0.7	,								
	XGBoost: gamma=0.1,									
202	learning_rate=0.0 max_depth=3,	$^{00}42135314$.464491.17	5448.39	21.27	-0.11	22.42	0.6	-15.578925	1
	n_estimators=100 subsample=0.8),								
	XGBoost: gamma=0.1,									
203	learning_rate=0.0	42135198	.86491.16	5448.3	21.27	-0.11	22.42	0.63	-15.554479	1
	n_estimators=100									
	subsample=1.0 XGBoost: gamma=0.1,									
204	gamma=0.1, learning_rate=0.0 max_depth=3,	04899901	780001.04	4046 79	10 22	0 00	20.43	1 00	8 345035	1
204	max_depth=3, n_estimators=200	040220U1),	. 100001.04	4946.72	19.32	0.08	20.45	1.02	-8.245035	1
	subsample=0.7	,								
	XGBoost: gamma=0.1,	001								
205	learning_rate=0.0 max_depth=3,	′′34825344	.0 9 901.3	4946.8	19.32	0.08	20.43	1.01	-7.878553	1
	n_estimators=200 subsample=0.8	J,								
	XGBoost: gamma=0.1,									
206	learning_rate=0.0 max_depth=3,	34827754	.74901.5	4946.81	19.32	0.08	20.43	1.04	-7.587614	1
	n_estimators=200 subsample=1.0									
	sausampie=1.0									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:				-			Training		
ID	G S	Men	DMen	MAE	MADE	$\mathbf{R^2}$	CVRMSI	J	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	R-	CVKMSI	Lime	Gap	Ratio
								(s)	•	
	XGBoost: gamma=0.1,									
207	learning_rate=0.0	001 46251646	.461800.86	5715.43	22.3	-0.22	23.44	0.61	-20.477036	1
201	max_depth=5, n_estimators=50,		. 101000.00	0110.10	22.0	0.22	20.11	0.01	20.111000	-
	subsample=0.7									
	XGBoost: gamma=0.1,	0.1								
208	learning_rate=0.0 max_depth=5,		.961800.79	5715.42	22.3	-0.22	23.44	0.61	-20.503898	1
	n_estimators=50, subsample=0.8									
	XGBoost:									
209	learning_rate=0.0	0016247328	98/800 54	5715.18	22.3	-0.22	23.44	0.63	-20.697797	1
203	max_depth=5, n_estimators=50,	10211020		0110.10	22.0	-0.22	20.44	0.00	-20.031131	1
	subsample=1.0									
	XGBoost: gamma=0.1,	001								
210	learning_rate=0.0 max_depth=5,	^{'0} 41892282	2.364472.42	5440.71	21.23	-0.1	22.31	1.14	-20.035027	1
	n_estimators=100 subsample=0.7),								
	XGBoost:									
211	learning_rate=0.0	00 1 1891000	.35472.33	5440.69	21.23	-0.1	22.31	1.11	-20.037673	1
	n_estimators=100			0 0 - 0						_
	subsample=0.8									
	XGBoost: gamma=0.1, learning_rate=0.0	001								
212	learning_rate=0.0 max_depth=5,		3.1 6 9472.16	5440.59	21.23	-0.1	22.31	1.07	-20.102339	1
	n_estimators=100 subsample=1.0),								
	XGBoost: gamma=0.1,									
213	learning_rate=0.0 max_depth=5,	34376138	3.55863.12	4930.92	19.24	0.1	20.2	1.87	-19.195708	1
	n_estimators=200),								
	subsample=0.7 XGBoost:									
01.4	XGBoost: gamma=0.1, learning_rate=0.0 max_depth=5,	001.07.4000	~~~~~	1000.00	10.04	0.1	20.2	0.75	10 155051	
214	max_depth=5, n_estimators=200	34374083	5.85862.94	4930.86	19.24	0.1	20.2	2.77	-19.155851	1
	subsample=0.8	,								
	XGBoost: gamma=0.1,									
215	learning_rate=0.0 max_depth=5,	34374238	3.45862.95	4931.1	19.24	0.1	20.19	1.89	-18.896324	1
	n_estimators=200									
	subsample=1.0 XGBoost;									
910	gamma=0.1, learning_rate=0.0 max_depth=7,	0010010101	econo e n	E71F 17	90.9	0.00	09.45	0.07	00 050470	-1
216	max_depth=7, n_estimators=50,		.wouu.b7	5715.17	22.3	-0.22	23.45	0.97	-20.250470	1
	subsample=0.7									
	XGBoost: gamma=0.1,	201								
217	learning_rate=0.0 max_depth=7,	46248769	.16800.64	5715.14	22.3	-0.22	23.44	0.93	-20.128805	1
	n_estimators=50, subsample=0.8									
	XGBoost:									
218	learning_rate=0.0	0016242150	969800 16	5714.64	22.3	-0.22	23.44	0.94	-20.401812	1
210	max_depth=7, n_estimators=50,			0114.04	44.0	-0.22	20.44	0.34	-20.401012	1
	subsample=1.0									
	XGBoost: gamma=0.1,	001								
219	learning_rate=0.0 max_depth=7,		.467471.91	5440.12	21.23	-0.1	22.31	1.6	-19.739786	1
	n_estimators=100 subsample=0.7									
	XGBoost: gamma=0.1,									
220	learning_rate=0.0	0011889961	76471 65	5439 74	21.22	-0.1	22.31	1.6	-19.711504	1
220	n_estimators=100			0400.14	41.44	-0.1	22.01	1.0	10.111004	1
	subsample=0.8									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:							Training		
ID	G 0	MCE	DMCE	MAE	MADE	\mathbb{R}^2	CVDMC	· ·	Overfit	Loss
ID	Config.	\mathbf{MSE}	RMSE	MAL	MAPE	n-	CVRMS		Gap	Ratio
								(s)		
221	XGBoost: gamma=0.1, learning_rate=0.0 max_depth=7, n_estimators=100	⁰ 41870080	0.33470.71	5438.83	21.22	-0.1	22.3	1.64	-20.238807	1
222	subsample=1.0 XGBoost: gamma=0.1, learning_rate=0.0 max_depth=7, n_estimators=200 subsample=0.7		0.25861.68	4929.57	19.23	0.1	20.2	2.9	-19.142773	1
223	XGBoost; gamma=0.1, learning_rate=0.0 max_depth=7, n_estimators=200 subsample=0.8		3.75861.39	4929.13	19.23	0.1	20.19	2.97	-18.876707	1
224	XGBoost: gamma=0.1, learning_rate=0.0 max_depth=7, n_estimators=200 subsample=1.0	⁰¹ 34343081 ,	03860.3	4927.97	19.23	0.1	20.19	3.03	-19.115755	1
225	XGBoost: gamma=0.1, learning_rate=0.0 max_depth=3, n_estimators=50, subsample=0.7	¹ 19666296	5.28434.67	3698.43	14.47	0.48	15.46	0.39	8.451086	1
226	XGBoost: gamma=0.1, learning_rate=0.0 max_depth=3, n_estimators=50, subsample=0.8	¹ 19647481	.144432.55	3697	14.46	0.48	15.45	0.39	6.480059	1
227	XGBoost: gamma=0.1, learning_rate=0.0 max_depth=3, n_estimators=50, subsample=1.0	¹ 19681157	7.54436.35	3698.55	14.47	0.48	15.47	0.39	10.309046	1
228	XGBoost: gamma=0.1, learning_rate=0.0 max_depth=3, n_estimators=100 subsample=0.7		112775.95	2290.37	8.98	0.8	9.8	0.63	5.042698	1
229	XGBoost gamma=0.1, learning_rate=0.0 max_depth=3, n_estimators=100 subsample=0.8	¹ 7710686.9	952776.81	2291.52	8.98	0.8	9.8	0.62	5.807418	1
230	XGBoost gamma=0.1, learning_rate=0.0 max_depth=3, n_estimators=100 subsample=1.0	¹ 7708988.	562776.51	2290.6	8.98	0.8	9.8	0.62	5.759622	1
231	XGBoost gamma=0.1, learning_rate=0.0 max_depth=3, n_estimators=200 subsample=0.7	¹ 1417867.4	471190.74	956.73	3.73	0.96	4.18	1.06	-18.627706	0.98
232	XGBoost gamma=0.1, learning_rate=0.0 max_depth=3, n_estimators=200 subsample=0.8	¹ 1417431.8	861190.56	956.58	3.72	0.96	4.18	1.02	-18.837539	0.98
233	XGBoost: gamma=0.1, learning_rate=0.0 max_depth=3, n_estimators=200 subsample=1.0	¹ 1418742.0	661191.11	957	3.73	0.96	4.18	1.02	-18.100379	0.99
234	XGBoost: gamma=0.1, learning_rate=0.0 max_depth=5, n_estimators=50, subsample=0.7	¹ 18926746	5.2 4 850.49	3661.66	14.29	0.5	14.97	0.62	-20.192375	1

Table 5.1: Results of the ablation study of the Machine learning architecture.

								Training		
ID	G	MCE	DMCE	MAE	MAPE	\mathbb{R}^2	CVDMS	· ·	Overfit	Loss
ID	Config.	MSE	RMSE	MAL	MAPE	n-	CVRMS	E lime	Gap	Ratio
								(s)		
	XGBoost: gamma=0.1,									
235	learning_rate=0.0 max_depth=5,	¹ 18930165	5.94350.88	3662.98	14.29	0.5	14.97	0.62	-19.037527	1
	n_estimators=50,									
	subsample=0.8 XGBoost: gamma=0.1,									
236	gamma=0.1, learning_rate=0.0 max_depth=5,	4,0025012	0 60051 54	2664 46	14.29	0.5	14.97	0.66	17 697020	1
230	max_depth=5, n_estimators=50,	10933912	2.04331.34	3664.46	14.29	0.5	14.97	0.66	-17.687920	1
	subsample=1.0									
	XGBoost: gamma=0.1,	.1								
237	learning_rate=0.0 max_depth=5,		2 2668.15	2244.56	8.76	0.81	9.17	1.06	-19.659681	0.99
	n_estimators=100 subsample=0.7	١,								
	XGBoost: gamma=0.1,									
238	learning_rate=0.0 max_depth=5,	¹ 7107441.:	292665.98	2242.88	8.75	0.81	9.16	1.08	-20.338987	0.99
	n_estimators=100									
	subsample=0.8 XGBoost: gamma=0.1,									
239	gamma=0.1, learning_rate=0.0 max_depth=5,	1 ₇₁ 02200	070665-01	2244.16	8.75	0.81	9.15	1.03	10 690194	0.99
239	max_depth=5, n_estimators=100		912005.01	2244.10	0.10	0.61	9.10	1.05	-19.620134	0.99
	subsample=1.0	,								
	XGBoost: gamma=0.1,	1								
240	learning_rate=0.0 max_depth=5,	1163395.	551078.61	884.9	3.45	0.97	3.64	1.76	-25.685012	0.98
	n_estimators=200 subsample=0.7	',								
	XGBoost: gamma=0.1,									
241	learning_rate=0.0 max_depth=5,	¹ 1157426.	321075.84	883.04	3.44	0.97	3.63	1.82	-24.081997	0.98
	n_estimators=200 subsample=0.8	,								
	XGBoost:									
242	learning_rate=0.0 max_depth=5,	¹ 1145450	241070-26	878.63	3.43	0.97	3.6	1.74	-23.345004	0.98
	n_estimators=200		_ 11010120	0.0.00	0.10	0.0.	0.0	211 2	20.010001	0.00
	subsample=1.0 XGBoost:									
0.49	XGBoost: gamma=0.1, learning_rate=0.0 max_depth=7,	4,000,000	04946 91	9050.00	14.07	0.5	1405	0.00	01 407167	-1
243	max_depth=7, n_estimators=50,	18890390).84346.31	3656.68	14.27	0.5	14.95	0.83	-21.407167	1
	subsample=0.7									
	XGBoost: gamma=0.1,									
244	learning_rate=0.0 max_depth=7,	18897938	3.049347.18	3658.32	14.28	0.5	14.96	0.91	-19.228515	1
	n_estimators=50, subsample=0.8									
	XGBoost: gamma=0.1,									
245	learning_rate=0.0 max_depth=7,	¹ 18902314	1.52347.68	3660.3	14.28	0.5	14.96	0.92	-16.925616	1
	n_estimators=50,									
	subsample=1.0 XGBoost: gamma=0.1,									
246	gamma=0.1, learning_rate=0.0 max_depth=7,	¹ 7093068	542663 28	2237.31	8.74	0.81	9.14	1.37	-21.832961	0.99
240	max_depth=7, n_estimators=100	, 099000. ,	o-±000.∠0	4401.01	0.14	0.01	9.14	1.01	-21.002701	0.33
	subsample=0.7 XGBoost:									
	gamma-() 1	1								
247	learning_rate=0.0 max_depth=7,	7078789.	412660.6	2235.03	8.73	0.81	9.13	1.43	-21.988012	0.99
	n_estimators=100 subsample=0.8	',								
	XGBoost: gamma=0.1,									
248	learning_rate=0.0 max_depth=7,	¹ 7052200.	292655.6	2230.97	8.72	0.81	9.11	1.49	-23.066939	0.99
	n_estimators=100 subsample=1.0									
	sausampie=1.0									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:				<u> </u>			 Training		
			DAGE	2547		D 2		_	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Gup	164010
	XGBoost:									
249	gamma=0.1, learning_rate=0.01 max_depth=7,	L 160018 6	3/1077 /6	883.18	3.44	0.97	3.63	2.62	-25.467600	0.98
243	max_depth=7, n_estimators=200,		941077.40	000.10	0.44	0.91	5.05	2.02	-25.407000	0.90
	subsample=0.7	,								
	XGBoost; gamma=0.1,									
250	learning_rate=0.01 max_depth=7,	¹ 1156044.4	1075.2	882.78	3.44	0.97	3.62	2.38	-22.908478	0.98
	n_estimators=200, subsample=0.8									
	XGBoost:									
251	gamma=0.1, learning_rate=0.01 max_depth=7,	L ₁₁₀₁₇₆ 3	211067 70	875.9	3.42	0.97	3.59	2.54	-22.784651	0.98
201	max_depth=7, n_estimators=200,	1140170.0	эшоот.тэ	010.9	5.42	0.91	5.55	2.04	-22.764031	0.90
	subsample=1.0									
	XGBoost: gamma=0.1,									
252	learning_rate=0.1, max_depth=3,	249810.15	499.81	352.66	1.25	0.99	1.61	0.43	-5.344989	0.99
	n_estimators=50, subsample=0.7									
	XGBoost: gamma=0.1,									
253	learning_rate=0.1, max_depth=3,	248620.79	498.62	353.21	1.25	0.99	1.6	0.42	-0.078202	1
	n_estimators=50,			000		0.00		V <u>-</u>	0.00.000	_
	subsample=0.8									
	XGBoost: gamma=0.1,									
254	learning_rate=0.1, max_depth=3,	246236.51	496.22	350.49	1.24	0.99	1.6	0.43	3.338862	1.01
	n_estimators=50, subsample=1.0									
	XGBoost: gamma=0.1,									
255	learning_rate=0.1,	247266.8	497.26	349.76	1.24	0.99	1.6	0.49	-7.003853	0.99
	n_estimators=100,									
	subsample=0.7 XGBoost:									
050	XGBoost; gamma=0.1, learning_rate=0.1, max_depth=3,	0.40504.50	100 50	055 50	1.00	0.00	1.0	0 =	2 400000	1.01
256	max_depth=3, n_estimators=100,		498.79	355.52	1.26	0.99	1.6	0.5	2.690003	1.01
	subsample=0.8	,								
	XGBoost: gamma=0.1,									
257	learning_rate=0.1, max_depth=3,	246671.24	496.66	352.24	1.25	0.99	1.59	0.52	5.227642	1.01
	n_estimators=100,	,								
	subsample=1.0 XGBoost: gamma=0.1,									
258	gamma=0.1, learning_rate=0.1, max_depth=3,	0.47949.01	407.24	240.97	1.24	0.00	1.6	0.72	6 026160	0.00
200	max_depth=3, n_estimators=200,	447343.91	491.34	349.87	1.24	0.99	1.6	0.73	-6.926169	0.99
	subsample=0.7	,								
	XGBoost: gamma=0.1,									
259	learning_rate=0.1, max_depth=3,		498.68	355.35	1.26	0.99	1.6	0.71	2.572762	1.01
	n_estimators=200, subsample=0.8	,								
	XGBoost:									
260	learning_rate=0.1,	246674 71	496 66	352.25	1.25	0.99	1.59	0.76	5.231110	1.01
200	n_estimators=200,		100.00	552.25	1.20	0.00	1.00	0.10	0.201110	1.01
	subsample=1.0									
	XGBoost: gamma=0.1,									
261	learning_rate=0.1, max_depth=5,	246961.9	496.95	361.38	1.28	0.99	1.57	0.58	23.018749	1.05
	n_estimators=50, subsample=0.7									
	XGBoost:									
262	learning_rate=0.1,	234893.06	8 484.66	350.44	1.23	0.99	1.54	0.51	17.307546	1.04
	n_estimators=50,									~ -
	subsample=0.8									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:	nesuits	or the ar	mation s	study of t	пе ма			mieciure.	
							1	Training	Overfit	Loss
ID	Config.	MSE	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time		
								(s)	Gap	Ratio
	XGBoost:									
000	gamma=0.1, learning_rate=0.1 max_depth=5,	202050 05	, 40° 0°	255.05	1.04	0.00	1 50	0.59	00 607066	1.00
263	max_depth=5, n_estimators=50,	230059.27	485.86	355.05	1.24	0.99	1.53	0.53	28.627266	1.06
	subsample=1.0									
	XGBoost; gamma=0.1,									
264	learning_rate=0.1 max_depth=5,	246704.71	496.69	361.02	1.27	0.99	1.57	0.64	22.759798	1.05
	n_estimators=100 subsample=0.7	,								
	XGBoost:									
265	gamma=0.1, learning_rate=0.1 max_depth=5,	232565 76	482.25	347.77	1.22	0.99	1.53	0.63	15.771287	1.03
200	max_depth=5, n_estimators=100	, ,	102.20	011.11	1.22	0.00	1.00	0.00	10.111201	1.00
	subsample=0.8 XGBoost;									
	gamma-() 1									
266	learning_rate=0.1 max_depth=5,	'236066.07	485.87	355.06	1.24	0.99	1.53	0.6	28.634245	1.06
	n_estimators=100 subsample=1.0	',								
	XGBoost: gamma=0.1,									
267	learning_rate=0.1 max_depth=5,	244366	494.33	358.56	1.26	0.99	1.56	0.84	21.353447	1.05
	n_estimators=200 subsample=0.7	,								
	XGBoost: gamma=0.1,									
268	learning_rate=0.1	232440.12	482.12	347.59	1.22	0.99	1.53	0.81	15.640821	1.03
200	max_depth=5, n_estimators=200		102.12	011.00	1.22	0.00	1.00	0.01	10.010021	1.00
	subsample=0.8 XGBoost:									
	gamma=0 1									
269	learning_rate=0.1 max_depth=5,		485.87	355.07	1.24	0.99	1.53	0.84	28.637919	1.06
	n_estimators=200 subsample=1.0	',								
	XGBoost: gamma=0.1,									
270	learning_rate=0.1 max_depth=7,	'239311.66	489.19	353.71	1.24	0.99	1.56	0.56	17.114508	1.04
	n_estimators=50, subsample=0.7									
	XGBoost: gamma=0.1,									
271	learning_rate=0.1	'242235.15	492.17	358.13	1.26	0.99	1.56	0.57	23.351680	1.05
	n_estimators=50,			000.20						
	subsample=0.8 XGBoost:									
070	XGBoost: gamma=0.1, learning_rate=0.1 max_depth=7,	2007011.05	407.15	95415	1.04	0.00	1.54	0.50	90 144776	1.07
272	max_depth=7, n_estimators=50,	237311.97	487.15	354.15	1.24	0.99	1.54	0.58	30.144776	1.07
	subsample=1.0									
	XGBoost: gamma=0.1,									
273	learning_rate=0.1 max_depth=7,	'236613.64	486.43	350.69	1.23	0.99	1.55	0.68	15.291527	1.03
	n_estimators=100 subsample=0.7	,								
	XGBoost:									
274	learning_rate=0.1 max_depth=7,	'242101.95	492.04	357.95	1.26	0.99	1.56	0.7	23.216435	1.05
	n_estimators=100									
	subsample=0.8 XGBoost: gamma=0.1,									
975	gamma=0.1, learning_rate=0.1 max_depth=7,	1027200 10	10710	25/10	1.04	0.00	1 54	0.75	20 155071	1.07
275	max_depth=7, n_estimators=100		487.16	354.16	1.24	0.99	1.54	0.75	30.155271	1.07
	subsample=1.0	•								
	XGBoost: gamma=0.1,									
276	learning_rate=0.1 max_depth=7,		486.45	350.72	1.23	0.99	1.55	0.89	15.312016	1.03
	n_estimators=200 subsample=0.7	',								

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1	· recours	Of the a	DIGUIOII E	dudy of t	JIC IVIC			intecture.	
								Training	Overfit	Loss
ID	Config.	MSE	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMS	E Time	Gap	Ratio
								(s)	Сар	itatio
	XGBoost: gamma=0.1,									
277	learning_rate=0	.1,241058 4	5 401 80	357.75	1.26	0.99	1.56	0.9	23.070425	1.05
211	max_depth=7, n_estimators=20		0 491.09	551.15	1.20	0.33	1.50	0.9	25.070425	1.00
	subsample=0.8	,								
	XGBoost: gamma=0.1,	_								
278	learning_rate=0 max_depth=7,	·1,237326.5	7487.16	354.17	1.24	0.99	1.54	0.94	30.159722	1.07
	n_estimators=20 subsample=1.0	00,								
	XGBoost;									
279	gamma=0.5, learning_rate=0 max_depth=3,	.0016977499)	5710.94	20.20	0.22	22.5	0.49	19 579959	1
219	max_depth=3, n_estimators=50	- 40377423).	0.20010.1	5719.34	22.32	-0.22	23.5	0.42	-18.572252	1
	subsample=0.7	,								
	XGBoost; gamma=0.5,									
280	learning_rate=0 max_depth=3,	$^{.001}6377438$	3. 76 810.1	5719.31	22.32	-0.22	23.5	0.42	-18.549472	1
	n_estimators=50 subsample=0.8),								
	XGBoost: gamma=0.5,									
281	learning_rate=0 max_depth=3,	.0016376202	260810.01	5719.26	22.32	-0.22	23.5	0.4	-18.625938	1
201	max_depth=3, n_estimators=50		20010.01	0110.20	22.02	-0.22	20.0	0.4	-10.020300	1
	subsample=1.0									
	XGBoost: gamma=0.5,	001								
282	learning_rate=0 max_depth=3,		3.0 9 490.93	5448.27	21.27	-0.11	22.42	0.6	-15.856118	1
	n_estimators=10 subsample=0.7	00,								
	XGBoost:									
283	learning_rate=0	$^{.001}_{-42135314}$	1.464491.17	5448.39	21.27	-0.11	22.42	0.63	-15.578925	1
	n_estimators=10					-				
	subsample=0.8 XGBoost:									
20.4	gamma=0.5, learning_rate=0 max_depth=3,	.001		* 440.0	04 0 -		22.42	0.00		
284	max_depth=3,	42135198	3.86491.16	5448.3	21.27	-0.11	22.42	0.63	-15.554479	1
	n_estimators=10 subsample=1.0	,								
	XGBoost: gamma=0.5,									
285	learning_rate=0 max_depth=3,	34822301	.759901.04	4946.72	19.32	0.08	20.43	1.05	-8.245035	1
	n_estimators=20 subsample=0.7	00,								
	XGBoost: gamma=0.5,									
286	gamma=0.5, learning_rate=0 max_depth=3,	.0014825344	L 0 6 001 3	4946.8	19.32	0.08	20.43	1.04	-7.878553	1
200	max_depth=3, n estimators=20		.00001.0	4340.0	13.32	0.08	20.40	1.04	-1.010000	1
	subsample=0.8									
	XGBoost: gamma=0.5,	001								
287	learning_rate=0 max_depth=3,	.0034827754	1.7 4 901.5	4946.81	19.32	0.08	20.43	1.15	-7.587614	1
	n_estimators=20 subsample=1.0	00,								
	XGBoost:									
288	learning_rate=0	$\frac{.001}{46271192}$	2.08802.29	5714.77	22.3	-0.22	23.46	0.61	-20.492637	1
	n_estimators=50					-				
	subsample=0.7 XGBoost:									
	gamma-() 5	001								
289	learning_rate=0 max_depth=5,		.05802.38	5714.87	22.3	-0.22	23.46	0.63	-20.327532	1
	n_estimators=50 subsample=0.8	ν,								
	XGBoost: gamma=0.5,									
290	learning_rate=0 max_depth=5,	0.0016259746	6.86801.45	5714.18	22.3	-0.22	23.45	0.64	-20.804088	1
	n_estimators=50									
	subsample=1.0									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:	1 (CSUITGS	or the a	oration s	day or i	JIC IVI	_		inicciure.	
								lraining	Overfit	$_{ m Loss}$
ID	Config.	MSE	RMSE	MAE	MAPE	${f R^2}$	CVRMSE	\mathbf{Time}	Gan	Ratio
								(s)	Gap	Italio
	XGBoost;									
201	gamma=0.5, learning_rate=0.0 max_depth=5,	011000165	: CC171 79	E 420 02	91 99	0.1	20.22	1.06	20 520021	1
291	max_depth=5, n_estimators=100		0.000474.73	5438.92	21.22	-0.1	22.33	1.06	-20.530021	1
	subsample=0.7	,								
	XGBoost: gamma=0.5,									
292	learning_rate=0.0 max_depth=5,	$^{04}_{41926609}$	0.034475.08	5439.23	21.22	-0.1	22.33	1.96	-20.000141	1
	n_estimators=100	,								
	subsample=0.8 XGBoost;									
200	gamma-() 5	01		- 10=	24.22		22.22	4.00	04 05 4000	
293	learning_rate=0.0 max_depth=5,	41902769	0.76473.23	5437.75	21.22	-0.1	22.32	1.08	-21.074220	1
	n_estimators=100 subsample=1.0	,								
	XGBoost: gamma=0.5,									
294	learning_rate=0.0 max_depth=5,	34426459	0.654867.41	4927.36	19.23	0.1	20.25	1.93	-20.159756	1
	n_estimators=200	,								
	subsample=0.7 XGBoost;									
		01		400=0	40.00		00.05		40 4000=4	
295	learning_rate=0.0 max_depth=5,	34433817	7.656868.03	4927.9	19.23	0.1	20.25	1.88	-19.196271	1
	n_estimators=200 subsample=0.8	,								
	XGBoost: gamma=0.5,									
296	learning_rate=0.0 max_depth=5,	⁰¹ 34409036	6.054865.92	4926.34	19.22	0.1	20.24	1.83	-20.063315	1
	n_estimators=200									
	subsample=1.0 XGBoost:									
	gamma-0.5	01								
297	learning_rate=0.0 max_depth=7,	46271192	2.068802.29	5714.77	22.3	-0.22	23.46	0.64	-20.492637	1
	n_estimators=50, subsample=0.7									
	XGBoost: gamma=0.5,									
298	learning_rate=0.0 max_depth=7,	$^{01}_{46272321}$.05802.38	5714.87	22.3	-0.22	23.46	0.65	-20.327532	1
	n_estimators=50,									
	subsample=0.8									
	XGBoost: gamma=0.5,	01								
299	learning_rate=0.0 max_depth=7,	46259746	5.86801.45	5714.18	22.3	-0.22	23.45	0.75	-20.804088	1
	n_estimators=50, subsample=1.0									
	XGBoost: gamma=0.5,									
300	learning_rate=0.0	$^{01}_{41922165}$	5.65474.73	5438.92	21.22	-0.1	22.33	1.05	-20.530021	1
	n_estimators=100									
	subsample=0.7 XGBoost;									
		01								
301	learning_rate=0.0 max_depth=7,		0.064475.08	5439.23	21.22	-0.1	22.33	1.08	-20.000141	1
	n_estimators=100 subsample=0.8	,								
	XGBoost: gamma=0.5,									
302	learning_rate=0.0 max_depth=7,	011902769	0.761473.23	5437.75	21.22	-0.1	22.32	1.14	-21.074220	1
	n_estimators=100									
	subsample=1.0 XGBoost:									
	XGBoost: gamma=0.5, learning_rate=0.0	01				_				
303	learning_rate=0.0 max_depth=7,		0.654867.41	4927.36	19.23	0.1	20.25	1.99	-20.159756	1
	n_estimators=200 subsample=0.7	,								
	XGBoost: gamma=0.5,									
304	learning_rate=0.0	$\frac{01}{3}4433817$	7.656868.03	4927.9	19.23	0.1	20.25	1.98	-19.196271	1
- '	n_estimators=200				-		-	-		
	subsample=0.8									

Table 5.1: Results of the ablation study of the Machine learning architecture.

								Training		
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	E Time	Overfit	Loss
	g.							(s)	Gap	Ratio
	XGBoost;							(3)		
005		001	0540.05.00	1000 01	10.00	0.1	20.24	0.05	00.000015	
305	learning_rate=0.0 max_depth=7,		6.054865.92	4926.34	19.22	0.1	20.24	2.05	-20.063315	1
	n_estimators=200 subsample=1.0	,								
	XGBoost: gamma=0.5,									
306	learning_rate=0.0 max_depth=3,	¹ 19666296	5.248434.67	3698.43	14.47	0.48	15.46	0.42	8.451086	1
	n_estimators=50, subsample=0.7									
	XGBoost: gamma=0.5,									
307	learning_rate=0.0 max_depth=3,)1 19647481	.144432.55	3697	14.46	0.48	15.45	0.42	6.480059	1
	n_estimators=50,									
	subsample=0.8 XGBoost: gamma=0.5,									
308	learning_rate=0.0 max_depth=3,)1 ₁₉₆₈₁₁₅₇	5443635	3698.55	14.47	0.48	15.47	0.41	10.309046	1
000	n_estimators=50,		.01100.00	0000.00	11.11	0.40	10.41	0.41	10.000040	
	subsample=1.0 XGBoost;									
200	gamma=0.5, learning_rate=0.0 max_depth=3,	17705014	1 10775 05	2000 27	0.00	0.0	0.0	0.61	F 0.40000	1
309	max_depth=3, n_estimators=100		112775.95	2290.37	8.98	0.8	9.8	0.61	5.042698	1
	subsample=0.7	,								
	XGBoost: gamma=0.5,	11								
310	learning_rate=0.0 max_depth=3,		952776.81	2291.52	8.98	0.8	9.8	0.63	5.807418	1
	n_estimators=100 subsample=0.8	J,								
	XGBoost: gamma=0.5,									
311	learning_rate=0.0 max_depth=3,	⁰¹ 7708988.	562776.51	2290.6	8.98	0.8	9.8	0.63	5.759622	1
	n_estimators=100 subsample=1.0	0,								
	XGBoost: gamma=0.5,									
312	learning_rate=0.0 max_depth=3,	⁰¹ 1420569.	291191.88	954.26	3.72	0.96	4.2	1.04	-21.385434	0.98
	n_estimators=200 subsample=0.7	Ο,								
	XGBoost: gamma=0.5,									
313	learning_rate=0.0 max_depth=3,	01 ₁ 416307.	711 1 90.09	955.18	3.72	0.96	4.18	1.05	-20.998093	0.98
010	n_estimators=200		. 11100.00	000.10	02	0.00	1.10	1.00	20.000000	0.00
	subsample=0.8 XGBoost;									
214	XGBoost: gamma=0.5, learning_rate=0.0 max_depth=3,	1419511	451101 O2	956.89	2 79	0.06	/ 10	1.04	19 617056	0 00
314	max_depth=3, n_estimators=200		401191.03	990.69	3.73	0.96	4.18	1.04	-18.617856	0.98
	$\overline{\text{subsample}}=1.0$	•								
	XGBoost: gamma=0.5, learning_rate=0.0	01								
315	learning_rate=0.0 max_depth=5,		.04366.33	3665.24	14.3	0.5	15.08	0.65	-17.164667	1
	n_estimators=50, subsample=0.7	1								
	XGBoost: gamma=0.5,	0.1								
316	learning_rate=0.0 max_depth=5,		7.442365.65	3664.07	14.3	0.5	15.07	1.57	-16.566855	1
	n_estimators=50, subsample=0.8	1								
	XGBoost: gamma=0.5,									
317	learning_rate=0.0 max_depth=5,	⁰¹ 19033740	.149362.77	3662.28	14.29	0.5	15.06	0.66	-17.892872	1
	n_estimators=50, subsample=1.0									
	XGBoost:									
318	learning_rate=0.0	01 <mark>7275643.</mark>	782697.34	2254.95	8.8	0.81	9.35	1.02	-23.590356	0.99
320	max_depth=5, n_estimators=100				0.0	0.01	0.00	1.02	_0.00000	0.00

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:	resures	or the a	oration s	tudy of t	iic wi	aciiiic ica		inicciurc.	
								Training	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	E Time		Ratio
								(s)	Gap	natio
	XGBoost: gamma=0.5,									
910	gamma=0.5, learning_rate=0.03 max_depth=5.	15064100	0007.0	0070.00	0.0	0.01	0.00	1.01	00 104500	0.00
319	max_depth=5, n_estimators=100.		2695.2	2256.63	8.8	0.81	9.32	1.01	-22.184563	0.99
	subsample=0.8	,								
	XGBoost: gamma=0.5,									
320	learning_rate=0.03 max_depth=5,	¹ 7236700.	672690.11	2255.55	8.8	0.81	9.28	1.03	-21.098389	0.99
	n_estimators=100.									
	subsample=1.0 XGBoost;									
901	gamma=0.5, learning_rate=0.03 max_depth=5,	4000000	1 41 1 40 70	004.05	0.0	0.07	4.01	1.50	99 061000	0.07
321	max_depth=5, n_estimators=200,		141149.79	924.35	3.6	0.97	4.01	1.59	-33.861982	0.97
	subsample=0.7	,								
	$\begin{array}{c} \text{XGBoost:} \\ \text{gamma=0.5}, \end{array}$									
322	learning_rate=0.03 max_depth=5,	1306245.0	041142.91	920.9	3.58	0.97	3.97	1.57	-33.386408	0.97
	n_estimators=200, subsample=0.8	,								
	XGBoost:									
323	learning_rate=0.03 max_depth=5,	1 1278248.:	371130.6	916.1	3.56	0.97	3.9	1.54	-33.185714	0.97
0_0	$n_estimators=200$.			0 - 0	0.00	0.0.			331233123	0.0.
	subsample=1.0 XGBoost:									
224	gamma=0.5	المممد	0.40.00.00	000504	440		4 7 00	0.00		
324	learning_rate=0.03 max_depth=7, n estimators=50,	19064801	.04366.33	3665.24	14.3	0.5	15.08	0.62	-17.164667	1
	subsample=0.7									
	XGBoost: gamma=0.5,									
325	learning_rate=0.03 max_depth=7,	¹ 19058907	7.4 42 365.65	3664.07	14.3	0.5	15.07	0.63	-16.566855	1
	n_estimators=50, subsample=0.8									
	XGBoost: gamma=0.5,									
326	learning_rate=0.0	¹ †9033740	149362.77	3662.28	14.29	0.5	15.06	0.67	-17.892872	1
020	$ max_depth=7, \\ n_estimators=50, $	10000110		5002.20	11.20	0.0	10.00	0.01	11.002012	1
	subsample=1.0									
	XGBoost: gamma=0.5, learning_rate=0.0	L								
327	learning_rate=0.03 max_depth=7,		782697.34	2254.95	8.8	0.81	9.35	1	-23.590356	0.99
	n_estimators=100. subsample=0.7	,								
	XGBoost: gamma=0.5,									
328	learning_rate=0.03 max_depth=7,	$^{1}7264129$	2695.2	2256.63	8.8	0.81	9.32	1.02	-22.184563	0.99
	n_estimators=100. subsample=0.8	,								
	XGBoost: gamma=0.5,									
329	learning_rate=0.03 max_depth=7,	¹ 7236700 (6 7 2690 11	2255.55	8.8	0.81	9.28	1.05	-21.098389	0.99
020	max_depth=7, n_estimators=100,		012000.11	2200.00	0.0	0.01	0.20	1.00	21.000000	0.00
	subsample=1.0									
	XGBoost: gamma=0.5, learning_rate=0.0	1								
330	learning_rate=0.03 max_depth=7,		141149.79	924.35	3.6	0.97	4.01	1.59	-33.861982	0.97
	n_estimators=200. subsample=0.7	,								
	XGBoost: gamma=0.5,									
331	learning_rate=0.03 max_depth=7,	1306245.0	041142.91	920.9	3.58	0.97	3.97	1.63	-33.386408	0.97
	n_estimators=200. subsample=0.8									
	XGBoost: gamma=0.5,									
332	learning_rate=0.0	1278248	371130.6	916.1	3.56	0.97	3.9	1.68	-33.185714	0.97
002	n_estimators=200.		0.11100.0	010.1	5.50	0.01	9.9	1.00	50.100114	0.31
	subsample=1.0									

Table 5.1: Results of the ablation study of the Machine learning architecture.

							r	Training		
TD		MOD	DATCE	2645	MADE	D2		_	Overfit	Loss
$^{\mathrm{ID}}$	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Gup	100010
	XGBoost:									
333	gamma=0.5, learning_rate=0.1 max_depth=3,	995659 14	570.25	411.75	1 51	0.00	1.07	0.20	27 425675	0.04
ააა	max_depth=3, n_estimators=50,	555052.14	579.55	411.75	1.51	0.99	1.97	0.39	-37.425675	0.94
	subsample=0.7									
	XGBoost: gamma=0.5,									
334	learning_rate=0.1 max_depth=3,	315679.17	561.85	395.77	1.44	0.99	1.9	0.39	-44.366390	0.93
	n_estimators=50,									
	subsample=0.8									
	XGBoost: gamma=0.5, learning_rate=0.1 max_depth=3.									
335		'305129.8	552.39	389.2	1.41	0.99	1.87	0.36	-30.170200	0.95
	n_estimators=50, subsample=1.0									
	XGBoost: gamma=0.5,									
336	learning_rate=0.1	329307.44	573.85	406.35	1.49	0.99	1.95	0.47	-39.147266	0.94
330	max_depth=3, n_estimators=100		0.0.00	100.00	1.10	0.00	1.00	0.1.	00.11.200	0.01
	subsample=0.7									
	XGBoost: gamma=0.5,									
337	learning_rate=0.1 max_depth=3,		559.3	393.02	1.43	0.99	1.89	0.46	-43.543486	0.93
	n_estimators=100 subsample=0.8),								
	XGBoost:									
338	learning_rate=0.1	305134 78	552 39	389.21	1.41	0.99	1.87	0.45	-30.165519	0.95
990	max_depth=3, n_estimators=100		002.00	000.21	1.41	0.55	1.01	0.40	-50.100013	0.50
	subsample=1.0									
	XGBoost: gamma=0.5,									
339	learning_rate=0.1 max_depth=3,	'329460.73	573.99	406.54	1.49	0.99	1.95	0.63	-39.013616	0.94
	n_estimators=200 subsample=0.7),								
	XGBoost: gamma=0.5,									
340	learning_rate=0.1	°319681 97	559 18	392.83	1.43	0.99	1.89	0.66	-43.667619	0.93
040	max_depth=3, n_estimators=200		000.10	002.00	1.40	0.55	1.00	0.00	-40.001013	0.55
	subsample=0.8									
	XGBoost: gamma=0.5,									
341	learning_rate=0.1 max_depth=3,	305137.22	552.39	389.22	1.41	0.99	1.87	0.67	-30.163214	0.95
	n_estimators=200 subsample=1.0),								
	XGBoost: gamma=0.5,									
342	learning_rate=0.1	334169 18	578.07	415.87	1.52	0.99	1.95	0.43	-30.391276	0.95
942	max_depth=5, n_estimators=50,	554102.10	310.01	410.01	1.02	0.00	1.00	0.40	50.031210	0.30
	subsample=0.7									
	XGBoost: gamma=0.5,									
343	learning_rate=0.1 max_depth=5,	'317987.52	563.9	403.77	1.46	0.99	1.89	0.42	-31.245274	0.95
	n_estimators=50, subsample=0.8									
	XGBoost:									
344	learning_rate=0.1	308073 25	555.85	398.91	1.43	0.99	1.82	0.42	-19.325209	0.97
944	max_depth=5, n estimators=50,	000010.20	555.55	990.91	1.40	0.99	1.04	0.44	-13.020203	0.31
	subsample=1.0									
	XGBoost: gamma=0.5,									
345	learning_rate=0.1 max_depth=5,	'327601.17	572.36	413.97	1.51	0.99	1.91	0.5	-32.688035	0.95
	n_estimators=100 subsample=0.7),								
	XGBoost: gamma=0.5,									
346	learning_rate=0.1	"317930 ∩ହ	563.86	403.71	1.46	0.99	1.89	1.5	-31.288223	0.95
940	max_depth=5, n_estimators=100		505.00	400.11	1.40	0.99	1.00	1.0	-01.400440	0.30
	subsample=0.8									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:				-			Training		
ID	G 6	MSE	DMCE	MAE	MAPE	$\mathbf{R^2}$	CVRMS	•	Overfit	Loss
ID	Config.	MSE	RMSE	MAL	MAPE	n-	CVRIVIS	E Iline	Gap	Ratio
								(s)		
	XGBoost: gamma=0.5,									
347	learning_rate=0.1 max_depth=5,	¹ ,308979.62	555.86	398.91	1.43	0.99	1.82	0.55	-19.319544	0.97
	n_estimators=100									
	subsample=1.0 XGBoost:									
9.40	XGBoost: gamma=0.5, learning_rate=0.1 max_depth=5,	1,005000 00	F 50 10	411.01	1 -	0.00	1.0	0.70	91 110747	0.05
348	max_depth=5, n_estimators=200	~323082.30 n	0 070.10	411.21	1.5	0.99	1.9	0.72	-31.118747	0.95
	subsample=0.7	~,								
	XGBoost: gamma=0.5,	1								
349	learning_rate=0.1 max_depth=5,		563.69	403.47	1.46	0.99	1.89	0.75	-31.460312	0.95
	n_estimators=200 subsample=0.8	0,								
	XGBoost:									
350	learning_rate=0.	¹ ,308983.13	555.86	398.92	1.43	0.99	1.82	0.73	-19.316470	0.97
	n_estimators=200									
	subsample=1.0 XGBoost;									
251	gamma=0.5, learning_rate=0.1 max_depth=7,	1,224160 16	170.07	415 07	1 50	0.00	1.05	0.45	20 201076	0.05
351	max_depth=7, n_estimators=50.		5 578.07	415.87	1.52	0.99	1.95	0.45	-30.391276	0.95
	subsample=0.7	,								
	XGBoost: gamma=0.5,									
352	learning_rate=0.7 max_depth=7,		563.9	403.77	1.46	0.99	1.89	0.43	-31.245274	0.95
	n_estimators=50, subsample=0.8	,								
	XGBoost:									
353	learning_rate=0.	¹ ,308973.25	555.85	398.91	1.43	0.99	1.82	0.45	-19.325209	0.97
	n_estimators=50									
	subsample=1.0 XGBoost: gamma=0.5,									
354	gamma=0.5, learning_rate=0.1 max_depth=7,	1,227601 17	579.26	413.97	1.51	0.99	1.91	0.57	-32.688035	0.95
354	max_depth=7, n_estimators=10	- 3 27001.17 0,	372.30	413.97	1.51	0.99	1.91	0.57	-52.000055	0.95
	subsample=0.7	,								
	XGBoost: gamma=0.5,	1								
355	learning_rate=0.: max_depth=7,		563.86	403.71	1.46	0.99	1.89	0.55	-31.288223	0.95
	n_estimators=100 subsample=0.8	0,								
	XGBoost: gamma=0.5,									
356	learning_rate=0.7 max_depth=7,	¹ ,308979.62	555.86	398.91	1.43	0.99	1.82	0.57	-19.319544	0.97
	n_estimators=100									
	subsample=1.0 XGBoost: gamma=0.5,									
357	gamma=0.5, learning_rate=0.1 max_depth=7,	1,325082 36	570.16	411.21	1.5	0.99	1.9	0.74	-31.118747	0.95
551	max_depth=7, n_estimators=20		010.10	411.21	1.0	0.55	1.5	0.14	-31.110747	0.33
	subsample=0.7									
	XGBoost: gamma=0.5, learning rate=0.	1.04==		105 :-		0				0
358	learning_rate=0.1 max_depth=7, n estimators=200		563.69	403.47	1.46	0.99	1.89	0.75	-31.460312	0.95
	subsample=0.8	υ,								
	XGBoost: gamma=0.5,									
359	learning_rate=0.1 max_depth=7,	¹ ,308983.13	555.86	398.92	1.43	0.99	1.82	0.77	-19.316470	0.97
	n_estimators=200 subsample=1.0	0,								
	XGBoost:									
360	learning_rate=0.0	001 46377423	28810.1	5719.34	22.32	-0.22	23.5	0.41	-18.572252	1
500	n_estimators=50.			3,10.01	0_	J.22	_5.0	U.11	10.012202	1
	subsample=0.7									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:	recours	Of the a	DIGUIOII 5	dudy of t	IIC IVIC		<u>~</u>	intecture.	
								Training	Overfit	$_{ m Loss}$
ID	Config.	MSE	\mathbf{RMSE}	MAE	MAPE	$\mathbf{R^2}$	CVRMSI	E Time	Gap	Ratio
								(s)	Gap	Itatio
	XGBoost;									
361	gamma=1, learning_rate=0.0 max_depth=3,	0016377438	7.6810.1	5719.31	22.32	-0.22	23.5	0.4	-18.549472	1
301	max_depth=3, n_estimators=50,			0113.31	22.02	-0.22	25.5	0.4	-10.043412	1
	subsample=0.8									
	XGBoost: gamma=1,	001								
362	learning_rate=0.0 max_depth=3,		2. 26 810.01	5719.26	22.32	-0.22	23.5	0.39	-18.625938	1
	n_estimators=50, subsample=1.0	,								
	XGBoost:									
363	learning_rate=0.0	$^{001}_{42132163}$.0 % 490.93	5448.27	21.27	-0.11	22.42	0.6	-15.856118	1
000	max_depth=3, n_estimators=100	0,		0110.21	21.21	0.11	22.12	0.0	10.000110	1
	subsample=0.7 XGBoost;									
		001								
364	learning_rate=0.0	~42135314	.464491.17	5448.39	21.27	-0.11	22.42	0.62	-15.578925	1
	n_estimators=100 subsample=0.8	J,								
	XGBoost: gamma=1,									
365	learning_rate=0.0 max_depth=3,	$^{004}2135198$	3.86491.16	5448.3	21.27	-0.11	22.42	0.62	-15.554479	1
	n_estimators=100 subsample=1.0									
	XGBoost: gamma=1,									
366	learning_rate=0.0	001/1822301	750001.04	4946.72	19.32	0.08	20.43	1.03	-8.245035	1
500	max_depth=3, n_estimators=200			1010.12	15.52	0.00	20.40	1.00	-0.240000	1
	subsample=0.7 XGBoost:									
	$a_{2}m_{2}-1$	001								
367	learning_rate=0.0 max_depth=3,	34825344	.0 9 901.3	4946.8	19.32	0.08	20.43	1	-7.878553	1
	n_estimators=200 subsample=0.8	υ,								
	XGBoost: gamma=1,									
368	learning_rate=0.0 max_depth=3,	34827754	.74901.5	4946.81	19.32	0.08	20.43	1.03	-7.587614	1
	n_estimators=200	0,								
	subsample=1.0 XGBoost;									
369	gamma=1, learning_rate=0.0 max_depth=5,	0016260080	16802.2	5714.99	22.3	-0.22	23.46	0.62	-20.793185	1
309	max_depth=5, n_estimators=50,	40203360	.10002.2	3114.33	22.3	-0.22	25.40	0.02	-20.795165	1
	subsample=0.7									
	XGBoost: gamma=1,	001								
370	learning_rate=0.0 max_depth=5,		3.36801.82	5714.46	22.3	-0.22	23.46	0.62	-21.046813	1
	n_estimators=50, subsample=0.8	,								
	XGBoost: gamma=1,									
371	learning_rate=0.0 max_depth=5,	0016266891	6801.98	5714.6	22.3	-0.22	23.46	0.61	-20.854686	1
	n_estimators=50,	,								
	subsample=1.0 XGBoost;									
372	gamma=1, learning_rate=0.0 max_depth=5,	00/1022710	067/7/95	5439.77	21.22	-0.1	22.33	1.01	-21.061268	1
312	max_depth=5, n_estimators=100		.w414.00	0403.11	21.22	-0.1	22.33	1.01	-21.001208	1
	subsample=0.7									
	XGBoost: gamma=1,	001								
373	learning_rate=0.0 max_depth=5,		6.68474.26	5438.89	21.22	-0.1	22.33	1.02	-21.332800	1
	n_estimators=100 subsample=0.8	U,								
	XGBoost:									
374	learning_rate=0.0 max_depth=5,	$^{001}_{41916612}$	2.46474.3	5438.52	21.22	-0.1	22.33	1.03	-21.111167	1
	n_estimators=100									
	subsample=1.0									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1	· recourse	or the a	DIGUIOII E	tuay or t	7110 1110				
						- 0		lraining	Overfit	Loss
ID	Config.	MSE	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Gup	164010
	XGBoost;									
975	gamma=1, learning_rate=0. max_depth=5,	0014440206	: 016969 E0	4020.15	10.99	0.00	20.25	1 70	20.844060	1
375	max_depth=5, n_estimators=20	04440360	.90000.39	4929.15	19.23	0.09	20.25	1.78	-20.844060	1
	subsample=0.7	ν,								
	XGBoost: gamma=1,									
376	learning_rate=0. max_depth=5,	$001 \over 34429255$	5.05867.64	4928.4	19.23	0.1	20.25	1.81	-21.088303	1
	n_estimators=20	0,								
	subsample=0.8 XGBoost;									
077	gamma=1, learning_rate=0. max_depth=5,	001,4001,70	- 	4007.0	10.00	0.1	90.95	0.70	01 097104	1
377	max_depth=5,	34422179	9.750867.04	4927.2	19.23	0.1	20.25	2.72	-21.037104	1
	n_estimators=20 subsample=1.0	υ,								
	XGBoost: gamma=1,									
378	learning_rate=0. max_depth=7,	0016269980	0.16802.2	5714.99	22.3	-0.22	23.46	0.63	-20.793185	1
	n_estimators=50									
	subsample=0.7 XGBoost;									
	gamma=1	001								
379	learning_rate=0. max_depth=7,		3.36801.82	5714.46	22.3	-0.22	23.46	0.62	-21.046813	1
	n_estimators=50 subsample=0.8	,								
	XGBoost: gamma=1,									
380	learning_rate=0. max_depth=7,	$^{001}_{46266891}$	6801.98	5714.6	22.3	-0.22	23.46	0.62	-20.854686	1
	n_estimators=50									
	subsample=1.0 XGBoost:									
	aamma-1	001								
381	learning_rate=0. max_depth=7,		0.067474.85	5439.77	21.22	-0.1	22.33	1.01	-21.061268	1
	n_estimators=10 subsample=0.7	0,								
	XGBoost: gamma=1,									
382	learning_rate=0.	⁰⁰ 41916075	5.68474.26	5438.89	21.22	-0.1	22.33	1.02	-21.332800	1
	max_depth=7, n_estimators=10	0,		0 -00.00		0.2				
	subsample=0.8									
	XGBoost: gamma=1,	001								
383	learning_rate=0. max_depth=7,	41916612	2.46474.3	5438.52	21.22	-0.1	22.33	1.02	-21.111167	1
	n_estimators=10 subsample=1.0	0,								
	XGBoost: gamma=1,									
384	learning_rate=0.	⁰⁰¹ 34440386	6.96868.59	4929.15	19.23	0.09	20.25	1.78	-20.844060	1
	n_estimators=20					0.00				_
	subsample=0.7 XGBoost;									
	gamma=1.	001								
385	learning_rate=0. max_depth=7,		6.05867.64	4928.4	19.23	0.1	20.25	1.82	-21.088303	1
	n_estimators=20 subsample=0.8	υ,								
	XGBoost;									
386	learning_rate=0.	$^{001}_{34422179}$	0.756867.04	4927.2	19.23	0.1	20.25	1.83	-21.037104	1
200	n_estimators=20									-
	subsample=1.0									
	XGBoost: gamma=1,	01								
387	learning_rate=0. max_depth=3,		5.2 4 34.67	3698.43	14.47	0.48	15.46	0.41	8.451086	1
	n_estimators=50 subsample=0.7	,								
	XGBoost:									
388	learning_rate=0.	⁰¹ 19647481	.14432.55	3697	14.46	0.48	15.45	0.39	6.480059	1
300	n_estimators=50			3301	11.10	0.10	10.10	0.00	0.10000	-
-	subsample=0.8									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:							Training		
ID		MOD	DAGE	24.5	MADE	D2		Ü	Overfit	Loss
$^{\mathrm{ID}}$	Config.	\mathbf{MSE}	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Сар	100010
	XGBoost:									
389	gamma=1, learning_rate=0.0 max_depth=3,	¹ 19681157	5443635	3698.55	14.47	0.48	15.47	0.4	10.309046	1
000	max_depth=3, n_estimators=50,	10001101	.01100.00	0000.00	11.11	0.40	10.41	0.4	10.000040	1
	subsample=1.0									
	XGBoost: gamma=1,	1								
390	learning_rate=0.0 max_depth=3,		112775.95	2290.37	8.98	0.8	9.8	0.61	5.042698	1
	n_estimators=100 subsample=0.7	,								
	XGBoost: gamma=1,									
391	learning_rate=0.0	¹ 7710686.	952776.81	2291.52	8.98	0.8	9.8	0.61	5.807418	1
	n_estimators=100	,								
	subsample=0.8 XGBoost;									
202	gamma=1, learning_rate=0.0 max_depth=3,	1		2222				0.00	F = F0000	
392	max_depth=3, n_estimators=100	7708988.	562776.51	2290.6	8.98	0.8	9.8	0.62	5.759622	1
	subsample=1.0	,								
	XGBoost: gamma=1,									
393	learning_rate=0.0 max_depth=3,	¹ 1501368.	391225.3	969.16	3.79	0.96	4.42	1.01	-18.830883	0.98
	n_estimators=200 subsample=0.7	,								
	XGBoost:									
394	learning_rate=0.0 max_depth=3,	¹t470544.	991212.66	964.25	3.76	0.96	4.34	1.01	-24.836076	0.98
	n_estimators=200						-			
	subsample=0.8 XGBoost:									
20.	gamma-1	1					4.00	4.00		
395	learning_rate=0.0 max_depth=3, n_estimators=200	1442083.	531200.87	956.27	3.73	0.96	4.28	1.02	-26.018782	0.98
	subsample=1.0	,								
	XGBoost; gamma=1,	_								
396	learning_rate=0.0 max_depth=5,	19134864	.946374.34	3667.73	14.31	0.5	15.13	0.56	-17.693049	1
	n_estimators=50, subsample=0.7									
	XGBoost:									
397	learning_rate=0.0	¹ 19100721	.82370.44	3665.18	14.31	0.5	15.12	0.6	-19.164637	1
	n_estimators=50,									
	subsample=0.8 XGBoost;									
200	gamma=1, learning_rate=0.0 max_depth=5,	40075041	00005 50	0000 45	140	0.5	15.1	0.6	10.010814	1
398	max_depth=5, n_estimators=50,	19075841	943307.59	3662.45	14.3	0.5	15.1	0.6	-18.810314	1
	subsample=1.0									
	XGBoost: gamma=1,									
399	learning_rate=0.0 max_depth=5,		542716.03	2262.21	8.83	0.81	9.47	0.89	-22.911903	0.99
	n_estimators=100 subsample=0.7	,								
	XGBoost:									
400	learning_rate=0.0 max_depth=5,	¹ 7364929.	562713.84	2262.01	8.83	0.81	9.45	0.91	-22.188066	0.99
	n_estimators=100									
	subsample=0.8 XGBoost;									
401	gamma=1, learning_rate=0.0 max_depth=5,	17227140	790706 97	2250 26	Q 01	0.01	0.41	0.02	99 Q101 <i>6</i> 9	0.00
401	max_depth=5, n_estimators=100		102100.81	2258.26	8.81	0.81	9.41	0.92	-23.810163	0.99
	subsample=1.0	•								
	XGBoost: gamma=1,	1								
402	learning_rate=0.0 max_depth=5,	1466449.	831210.97	960.45	3.75	0.96	4.35	1.39	-24.868947	0.98
	n_estimators=200 subsample=0.7	,								

Table 5.1: Results of the ablation study of the Machine learning architecture.

							,	Training		
								Iraining	Overfit	$_{ m Loss}$
$^{\mathrm{ID}}$	Config.	MSE	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Сар	Itatio
	XGBoost;									
403	gamma=1, learning_rate=0.0 max_depth=5,	4,24508	001107 71	953.86	3.71	0.96	4.28	1.41	-29.698369	0.98
405	max_depth=5, n_estimators=200	,	331131.11	300.00	5.71	0.90	4.20	1.41	-29.096309	0.96
	subsample=0.8									
	XGBoost: gamma=1,	1								
404	learning_rate=0.0 max_depth=5,	1396281.	211181.64	944.24	3.67	0.96	4.19	2.36	-32.025608	0.97
	n_estimators=200 subsample=1.0	,								
	XGBoost: gamma=1,									
405	learning_rate=0.0 max_depth=7,	¹ 19134864	.946374.34	3667.73	14.31	0.5	15.13	0.58	-17.693049	1
	n_estimators=50, subsample=0.7									
	XGBoost:									
406	gamma=1, learning_rate=0.0 max_depth=7,	¹ 19100721	842370 44	3665.18	14.31	0.5	15.12	0.6	-19.164637	1
100	max_depth=7, n_estimators=50,	10100121		5000.10	11.01	0.0	10.12	0.0	10.101001	-
	subsample=0.8 XGBoost;									
405	gamma=1, learning_rate=0.0 max_depth=7,	40055041	0.000.07 50	0000 45	140	0 =	15.1	0.0	10.010814	
407	max_depth=7, n_estimators=50,	19075841	94367.59	3662.45	14.3	0.5	15.1	0.6	-18.810314	1
	subsample=1.0									
	XGBoost: gamma=1,	1								
408	learning_rate=0.0 max_depth=7,		542716.03	2262.21	8.83	0.81	9.47	0.89	-22.911903	0.99
	n_estimators=100 subsample=0.7	,								
	XGBoost: gamma=1,									
409	learning_rate=0.0 max_depth=7,	¹ 7364929.	562713.84	2262.01	8.83	0.81	9.45	0.9	-22.188066	0.99
	n_estimators=100 subsample=0.8									
	XGBoost: gamma=1,									
410	learning_rate=0.0	¹ 7327140	782706 87	2258.26	8.81	0.81	9.41	0.94	-23.810163	0.99
110	max_depth=7, n_estimators=100	,	100.01	2200.20	0.01	0.01	0.11	0.01	20.010100	0.00
	subsample=1.0									
411	XGBoost: gamma=1, learning_rate=0.0 max_depth=7,	4,466,440	091010.05	000 45	0.75	0.00	4.05	1.4	04.000045	0.00
411	max_depth=7, n_estimators=200	1466449.	831210.97	960.45	3.75	0.96	4.35	1.4	-24.868947	0.98
	subsample=0.7	,								
	XGBoost: gamma=1,									
412	learning_rate=0.0 max_depth=7,		991197.71	953.86	3.71	0.96	4.28	1.42	-29.698369	0.98
	n_estimators=200 subsample=0.8	,								
	XGBoost: gamma=1,									
413	learning_rate=0.0 max_depth=7,	¹ 1396281.	211181.64	944.24	3.67	0.96	4.19	1.47	-32.025608	0.97
	n_estimators=200									
414	learning_rate=0.1	,429225 Q	7 655 15	461 84	1.75	0.99	2.37	0.38	-38.065057	0.95
117	n_estimators=50,	120220.0	, 000.10	101.04	1.10	0.00	2.01	0.30	30.000001	0.00
	subsample=0.7									
4	gamma=1, learning rate=0 1		4 400 07	445.00	4 a=	0.00	0.63	0.1	40.10007	0.01
415		400752.1	4 633.05	445.33	1.67	0.99	2.26	0.4	-40.126254	0.94
	subsample=0.8									
	XGBoost: gamma=1,									
416	learning_rate=0.1 max_depth=3,	381728.4	3 617.84	438.78	1.64	0.99	2.18	0.39	-32.908594	0.95
	n_estimators=50, subsample=1.0									
414 415	subsample=1.0 XGBoost; gamma=1, learning_rate=0.1 max_depth=3, n_estimators=50, subsample=0.7 XGBoost; gamma=1, learning_rate=0.1 max_depth=3, n_estimators=50, subsample=0.8	'429225.9'		461.84 445.33	1.75 1.67	0.99	2.37 2.26	0.38	-38.065057 -40.126254	

Table 5.1: Results of the ablation study of the Machine learning architecture.

							,	Training		
TD		MOD	DAGE	3.645	NAADE	D2		_	Overfit	Loss
$^{\mathrm{ID}}$	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	-	
	XGBoost: gamma=1,									
417	learning_rate=0.1	·,428802.79	654 83	461.42	1.75	0.99	2.37	0.49	-38.388234	0.94
411	max_depth=3, n_estimators=100	420002.1 <i>0</i>),	004.00	401.42	1.10	0.55	2.01	0.40	-50.500254	0.54
	subsample=0.7									
	XGBoost: gamma=1,									
418	learning_rate=0.1 max_depth=3,		632.98	445.24	1.67	0.99	2.26	0.44	-40.194807	0.94
	n_estimators=100 subsample=0.8),								
	XGBoost:									
419	gamma=1, learning_rate=0.1 max_depth=3,	·381733 37	617.85	438.79	1.64	0.99	2.18	0.49	-32.904624	0.95
413	max_depth=3, n_estimators=100		017.00	400.10	1.04	0.33	2.10	0.43	-52.304024	0.55
	subsample=1.0	,								
	XGBoost: gamma=1,									
420	learning_rate=0.1 max_depth=3,	'428092.45	654.29	460.7	1.75	0.99	2.37	0.62	-32.646724	0.95
	n_estimators=200 subsample=0.7),								
	XGBoost: gamma=1,									
421	learning_rate=0.1	399364.65	631.95	443.84	1.67	0.99	2.26	0.65	-35.352635	0.95
	n_estimators=200		002.00			0.00		0.00	331332	0.00
	subsample=0.8 XGBoost;									
	gamma=1.									
422	learning_rate=0.1 max_depth=3,		617.85	438.79	1.64	0.99	2.18	0.65	-32.902545	0.95
	n_estimators=200 subsample=1.0),								
	XGBoost: gamma=1,									
423	learning_rate=0.1	408550.62	639.18	445.51	1.69	0.99	2.31	0.42	-47.392631	0.93
	n_estimators=50,									
	subsample=0.7 XGBoost;									
10.1	gamma=1.		000 14	4.45 40	1.00	0.00	2.22	0.00	94 90009	0.05
424	learning_rate=0.1 max_depth=5,	'404672.21	636.14	447.49	1.68	0.99	2.28	0.39	-36.399037	0.95
	n_estimators=50, subsample=0.8									
	XGBoost: gamma=1,									
425	learning_rate=0.1 max_depth=5,	381499.18	617.66	440.38	1.63	0.99	2.15	0.41	-32.377064	0.95
	n_estimators=50,									
	subsample=1.0 XGBoost: gamma=1,									
196	gamma=1, learning_rate=0.1 max_depth=5	, 1000c0 c	620 0	444.00	1 60	0.00	0.21	0.49	47 760070	0.02
426	max_depth=5, n_estimators=100		638.8	444.99	1.68	0.99	2.31	0.48	-47.769072	0.93
	subsample=0.7	.,								
	XGBoost: gamma=1,									
427	learning_rate=0.1 max_depth=5,	'404652.21	636.12	447.47	1.68	0.99	2.28	0.5	-36.414736	0.95
	n_estimators=100 subsample=0.8),								
	XGBoost: gamma=1,									
428	learning_rate=0.1	381505 04°	617 66	440.39	1.63	0.99	2.15	0.49	-32.372436	0.95
£20	max_depth=5, n_estimators=100),	011.00	110.00	1.00	0.00	2.10	0.10	32.012400	0.00
	subsample=1.0									
	XGBoost: gamma=1,									
429	learning_rate=0.1 max_depth=5,		639.02	445.29	1.69	0.99	2.31	0.67	-47.552450	0.93
	n_estimators=200 subsample=0.7),								
	XGBoost:									
430	learning_rate=0.1 max_depth=5,	,404402.78	635.93	447.2	1.68	0.99	2.28	0.7	-36.610916	0.95
	n_estimators=200						-	•		
	subsample=0.8									

Table 5.1: Results of the ablation study of the Machine learning architecture.

	Table 5.1:	Kesults	of the a	biation s	tudy of t	ne Ma	achine leari	ning arc	nitecture.	
							7	Training	Overfit	Loss
ID	Config.	MSE	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time		
								(s)	Gap	Ratio
	XGBoost:									
401	gamma=1, learning_rate=0. max_depth=5,	1,001505.00	617 66	440.4	1.60	0.00	0.15	0.7	90 950105	0.05
431	max_depth=5, n_estimators=200		617.66	440.4	1.63	0.99	2.15	0.7	-32.370187	0.95
	subsample=1.0	ς,								
	XGBoost: gamma=1,									
432	learning_rate=0.7 max_depth=7,	$^{1,408550.62}$	639.18	445.51	1.69	0.99	2.31	0.41	-47.392631	0.93
	n_estimators=50 subsample=0.7									
	XGBoost:									
433	gamma=1, learning_rate=0. max_depth=7,	1, ₄₀₄₆ 72 21	636 14	447.49	1.68	0.99	2.28	0.39	-36.399037	0.95
400	max_depth=7, n_estimators=50.		050.14	441.43	1.00	0.33	2.20	0.55	-30.333031	0.30
	subsample=0.8									
	XGBoost: gamma=1,	1								
434	learning_rate=0.: max_depth=7,		617.66	440.38	1.63	0.99	2.15	0.4	-32.377064	0.95
	n_estimators=50 subsample=1.0	,								
	XGBoost: gamma=1,									
435	learning_rate=0.1 max_depth=7,	¹ ,408069.6	638.8	444.99	1.68	0.99	2.31	0.49	-47.769072	0.93
	n_estimators=100 subsample=0.7	0,								
	XGBoost: gamma=1,									
436	learning_rate=0.	¹ ,404652.21	636.12	447.47	1.68	0.99	2.28	0.49	-36.414736	0.95
100	n_estimators=10		000.12	111111	1.00	0.00	2.20	0.10	30.111.00	0.00
	subsample=0.8 XGBoost:									
40=	gamma=1, learning_rate=0.1 max_depth=7,	1	04= 00	440.00	4.00		2.45			
437	max_depth=7, n_estimators=10		617.66	440.39	1.63	0.99	2.15	0.51	-32.372436	0.95
	subsample=1.0	J,								
	XGBoost; gamma=1,									
438	learning_rate=0.1 max_depth=7,		639.02	445.29	1.69	0.99	2.31	0.67	-47.552450	0.93
	n_estimators=200 subsample=0.7	0,								
	XGBoost:									
439	learning_rate=0.1 max_depth=7,	¹ ,404402.78	635.93	447.2	1.68	0.99	2.28	0.7	-36.610916	0.95
	n_estimators=20									
	subsample=0.8 XGBoost;									
440	gamma=1, learning_rate=0.1 max_depth=7,	1,901507 00	617 66	440.4	1.63	0.99	2.15	0.7	-32.370187	0.95
440	max_depth=7, n_estimators=200		017.00	440.4	1.05	0.99	2.10	0.7	-32.370107	0.95
	subsample=1.0									
441	SVR: C=0.1,	148216.7	384.99	258.51	0.9	1	1.28	921.78	-56.072572	0.87
	epsilon=0.01, kernel=linear									
449	SVR: C=0.1,	1000007 6	01 410 64	1050 50	4.02	0.05	E 01	E07.47	756 000001	0.15
442	epsilon=0.01, kernel=rbf	1989897.6	21410.04	1050.52	4.03	0.95	5.21	507.47	756.028301	2.15
	SVR:									
443	C=0.1, epsilon=0.1,	3512111.5	61874.06	1581.55	5.76	0.91	4.8	1.66	452.917492	1.32
	kernel=linear SVR:									
444	C=0.1, epsilon=0.1,	6557051.8	42560.67	2081.01	7.97	0.83	8.17	6.28	582.779524	1.29
	kernel=rbf									
445	SVR: C=0.1,	11347187.	937368.56	2769.17	10.61	0.7	10.78	0.69	401.987520	1.14
110	epsilon=0.2, kernel=linear	_1011101.	23000.00	2.00.11	10.01	٠.,	10.10	0.00	-51.001020	2.2.1
4.40	SVR: C=0.1,	05705005	07070.00	4017.00	10 55	0.90	14.05	1.00	1150 000011	1.0
446	epsilon=0.2, kernel=rbf	25705097.	S⊗U7U.U2	4217.66	16.57	0.32	14.87	1.28	1158.003311	1.3
	1.011101—1.01									

Table 5.1: Results of the ablation study of the Machine learning architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R ²	CVRMS	Training E Time (s)	Overfit Gap	Loss Ratio
447	SVR: C=1.0, epsilon=0.01, kernel=linear	143044.3	1 378.21	254.57	0.89	1	1.27	9932.86	-60.848095	0.86
448	$\begin{array}{c} \mathrm{SVR:} \\ \mathrm{C=}1.0, \\ \mathrm{epsilon=}0.01, \\ \mathrm{kernel=rbf} \end{array}$	1688480.0	051299.42	1013.09	3.93	0.96	4.36	1309.78	906.864992	3.31
449	SVR: C=1.0, epsilon=0.1, kernel=linear	2946243.9	961716.46	1442.69	5.14	0.92	4.13	2.67	504.870654	1.42
450	$\begin{array}{c} \mathrm{SVR:} \\ \mathrm{C=}1.0, \\ \mathrm{epsilon=}0.1, \\ \mathrm{kernel=rbf} \end{array}$	6914955.9	992629.63	2167.62	8.26	0.82	7.98	3.86	786.395894	1.43
451	SVR: C=1.0, epsilon=0.2, kernel=linear	5834281.0	082415.43	1983.37	7.23	0.85	8.31	0.87	- 224.681832	0.91
452	SVR: C=1.0, epsilon=0.2, kernel=rbf	24779642	.845977.92	4095.46	16.1	0.35	14.88	1.09	1112.513789	1.29

 ${\bf Table~5.2:}~{\bf Results~of~the~ablation~study~of~the~Stacked~LSTM~architecture.}$

								Training	Overfit	Loss
ID	Config.	\mathbf{MSE}	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMS	E Time		Ratio
								(s)	Gap	natio
1	num_layers=1, dropout=True, epochs=10, batch_size=16	384695.55	620.24	443.36	1.45	0.99	2.01	332.54	-0.000123	0.68
2	num_layers=1, dropout=True, epochs=10, batch_size=32	476581.56	690.35	477.68	1.53	0.99	2.12	156.32	-0.000149	0.68
3	num_layers=1, dropout=True, epochs=10, batch_size=64	566230.43	752.48	549.87	1.72	0.99	2.08	79.88	-0.000150	0.72
4	num_layers=1, dropout=True, epochs=20, batch_size=16	330588.21	574.97	405.62	1.29	0.99	1.79	663.88	-0.000081	0.73
5	num_layers=1, dropout=True, epochs=20, batch_size=32	483506.39	695.35	508.58	1.69	0.99	2.27	312.9	-0.000039	0.89
6	num_layers=1, dropout=True, epochs=20, batch_size=64	441020.31	664.09	466.95	1.47	0.99	2.11	159.94	-0.000087	0.78
7	num_layers=1, dropout=True, epochs=50, batch_size=16	317355.3	563.34	412	1.36	0.99	1.83	1652.25	-0.000036	0.86
8	num_layers=1, dropout=True, epochs=50, batch_size=32	262201.1	512.06	349.15	1.12	0.99	1.66	780.65	-0.000059	0.75
9	num_layers=1, dropout=True, epochs=50, batch_size=64	3064671.9	51750.62	1438.97	4.06	0.94	3.23	393.7	0.001847	8.76
10	num_layers=1, dropout=False, epochs=10, batch_size=16	313059.99	559.52	384.83	1.24	0.99	1.83	322.14	-0.000050	0.81
11	num_layers=1, dropout=False, epochs=10, batch_size=32	385994.43	621.28	442.6	1.46	0.99	1.99	153.75	-0.000002	0.99

Table 5.2: Results of the ablation study of the Stacked LSTM architecture.

	Table 9.	u. 1050116	or the	rota (1011	study 01	one D	vacacu Li	STM archi	ioccourt.	
								Training	Overfit	$_{ m Loss}$
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	SE Time	Gap	Ratio
								(s)	Сар	Itatio
12	num_layers=1, dropout=False, epochs=10, batch_size=64	440168.68	663.45	449.03	1.44	0.99	2.12	78.27	-0.000026	0.92
13	num_layers=1, dropout=False, epochs=20, batch_size=16	293412.63	541.68	387.6	1.23	0.99	1.64	644.8	-0.000003	0.98
14	num_layers=1, dropout=False, epochs=20, batch_size=32	282089.41	531.12	363.25	1.16	0.99	1.71	310.56	-0.000019	0.91
15	num_layers=1, dropout=False, epochs=20, batch_size=64	410460.27	640.67	471.52	1.49	0.99	1.88	156.35	0.000041	1.17
16	num_layers=1, dropout=False, epochs=50, batch_size=16	235991.06	485.79	341.26	1.08	1	1.46	1611.98	0.000003	1.02
17	num_layers=1, dropout=False, epochs=50, batch_size=32	268392.88	518.07	370.08	1.17	0.99	1.54	762.08	0.000021	1.13
18	num_layers=1, dropout=False, epochs=50, batch_size=64	219325.22	468.32	315.39	1.01	1	1.52	389.03	-0.000018	0.89
19	num_layers=2, dropout=True, epochs=10, batch_size=16	378733.66	615.41	435.85	1.4	0.99	2.02	513.79	-0.000176	0.59
20	num_layers=2, dropout=True, epochs=10, batch_size=32	551913.8	742.91	550.4	1.85	0.99	2.48	246.34	-0.000065	0.85
21	num_layers=2, dropout=True, epochs=10, batch_size=64	590858.76	768.67	563.12	1.79	0.99	2.49	117.83	-0.000189	0.68
22	num_layers=2, dropout=True, epochs=20, batch_size=16	1117961.4	51057.34	864.66	2.82	0.98	3.35	1024.14	0.000470	2.61
23	num_layers=2, dropout=True, epochs=20, batch_size=32	445651.73	667.57	500.23	1.64	0.99	1.99	498.69	-0.000054	0.85
24	num_layers=2, dropout=True, epochs=20, batch_size=64	1927714.7	91388.42	1093.35	3.12	0.96	2.91	237.61	0.000965	3.76
25	num_layers=2, dropout=True, epochs=50, batch_size=16	305414.88	552.64	388.44	1.19	0.99	1.65	2571.78	-0.000018	0.92
26	num_layers=2, dropout=True, epochs=50, batch_size=32	3558243.3	371886.33	1525.8	4.52	0.93	5.03	1180.77	0.002209	11.24
27	num_layers=2, dropout=True, epochs=50, batch_size=64	4796429.2	282190.08	1808.18	5.04	0.91	3.1	591.34	0.003030	14.01
28	num_layers=2, dropout=False, epochs=10, batch_size=16	299566.83	547.33	376.22	1.2	0.99	1.77	495.24	-0.000056	0.79
29	num_layers=2, dropout=False, epochs=10, batch_size=32	313767.54	560.15	385.4	1.24	0.99	1.8	233.64	-0.000058	0.79
30	num_layers=2, dropout=False, epochs=10, batch_size=64	358836.9	599.03	425.95	1.37	0.99	1.9	116.63	-0.000046	0.84
31	num_layers=2, dropout=False, epochs=20, batch_size=16	254474.49	504.45	347.46	1.1	1	1.62	991.5	-0.000041	0.81

Table 5.2: Results of the ablation study of the Stacked LSTM architecture.

								Training	O	T a
ID	Config.	MSE	RMSE	MAE	MAPE	$\mathbf{R^2}$	CVRMS	SE Time	Overfit	Loss
	J							(s)	Gap	Ratio
32	num_layers=2, dropout=False, epochs=20, batch_size=32	254250.27	504.23	344.46	1.1	1	1.64	450.79	-0.000044	0.8
33	num_layers=2, dropout=False, epochs=20, batch_size=64	367098.58	605.89	450.5	1.4	0.99	1.76	227.73	0.000019	1.08
34	num_layers=2, dropout=False, epochs=50, batch_size=16 num_layers=2,	328401.43	3 573.06	421.69	1.28	0.99	1.58	2442.54	0.000058	1.35
35	num_layers=2, dropout=False, epochs=50, batch_size=32 num_layers=2,	247511.72	497.51	348.96	1.1	1	1.51	1232.29	0.000006	1.04
36	dropout=False, epochs=50, batch_size=64 num_layers=3,	224894.37	474.23	319.61	1.02	1	1.53	589.22	-0.000022	0.87
37	dropout=True, epochs=10, batch_size=16	424040.92	651.18	490.17	1.64	0.99	2.06	656.65	-0.000188	0.61
38	num_layers=3, dropout=True, epochs=10, batch_size=32	2170386.6	311473.22	1149.8	3.23	0.96	2.7	318.26	0.001021	3.23
39	num_layers=3, dropout=True, epochs=10, batch_size=64	889141.09	942.94	700.74	2.49	0.98	3.33	161.07	-0.000021	0.97
40	num_layers=3, dropout=True, epochs=20, batch_size=16	4191812.1	82047.39	1591.24	4.36	0.92	3.37	1306.13	0.002546	9.2
41	num_layers=3, dropout=True, epochs=20, batch_size=32	7168050.5	662677.32	2264.15	6.37	0.86	3.43	633.74	0.004584	16.25
42	num_layers=3, dropout=True, epochs=20, batch_size=64 num_layers=3,	2207328.8	81485.71	1205.74	3.44	0.96	2.39	321.83	0.001150	4.22
43	dropout=True, epochs=50, batch_size=16	3395723.6	91842.75	1417.6	3.87	0.93	3.27	3355.51	0.002101	10.86
44	num_layers=3, dropout=True, epochs=50, batch_size=32 num_layers=3,	2816321.4	11678.19	1307.26	3.69	0.94	3.73	1853.82	0.001682	8.09
45	dropout=True, epochs=50, batch_size=64	1005851.3	81002.92	798.37	2.93	0.98	3.04	1183.04	0.000436	2.74
46	num_layers=3, dropout=False, epochs=10, batch_size=16	514056.33	716.98	560.35	1.78	0.99	1.8	1082.58	0.000084	1.31
47	num_layers=3, dropout=False, epochs=10, batch_size=32	400410.01	632.78	455.39	1.42	0.99	1.87	338.07	-0.000015	0.95
48	num_layers=3, dropout=False, epochs=10, batch_size=64 num_layers=3,	579008.27	760.93	565.08	1.82	0.99	2.2	194.01	0.000053	1.16
49	dropout=False, epochs=20, batch_size=16	263879.45	5 513.69	361.26	1.15	0.99	1.62	1650.79	-0.000030	0.86
50	num_layers=3, dropout=False, epochs=20, batch_size=32	311075.68	557.74	403.95	1.29	0.99	1.7	837.36	-0.000015	0.93
51	num_layers=3, dropout=False, epochs=20, batch_size=64	283680.9	532.62	367.09	1.18	0.99	1.7	411.23	-0.000051	0.79

Table 5.2: Results of the ablation study of the Stacked LSTM architecture.

			<u> </u>					Training		
ID	C	MSE	RMSE	MAE	MAPE	$\mathbf{R^2}$	CVRM	SE Time	Overfit	Loss
עו	Config.	MSE	RWISE	MAL	MAPE	n-	CVRIVI		Gap	Ratio
								(s)		
52	num_layers=3, dropout=False, epochs=50, batch_size=16	240325.15	5 490.23	337.09	1.07	1	1.55	4232.35	0.000006	1.04
53	num_layers=3, dropout=False, epochs=50, batch_size=32	378951.56	615.59	472.76	1.45	0.99	1.55	2135.71	0.000093	1.56
54	num_layers=3, dropout=False, epochs=50, batch_size=64	245321.24	495.3	340.88	1.09	1	1.58	1026.73	-0.000001	0.99
55	num_layers=4, dropout=True, epochs=10, batch_size=16	1520894.2	211233.25	938.39	2.8	0.97	3.45	1240.47	0.000626	2.53
56	num_layers=4, dropout=True, epochs=10, batch_size=32	3680606.1	191918.49	1460.02	4.02	0.93	3.54	595.7	0.002031	5.25
57	num_layers=4, dropout=True, epochs=10, batch_size=64	2602226.3	3 1613.14	1182.73	3.34	0.95	3.19	287.33	0.001092	2.6
58	num_layers=4, dropout=True, epochs=20, batch_size=16	2225399.4	121491.78	1283.48	4.54	0.96	4.4	1862.99	0.001208	4.91
59	num_layers=4, dropout=True, epochs=20, batch_size=32	3151510.1	171775.25	1415.61	3.96	0.94	3.57	806.86	0.001816	6.48
60	num_layers=4, dropout=True, epochs=20, batch_size=64	6735232.4	142595.23	2226.73	6.35	0.87	3.02	401.49	0.004215	12.15
61	num_layers=4, dropout=True, epochs=50, batch_size=16	5699837.3	372387.43	1920.03	5.29	0.89	3.49	5165.87	0.003646	16.3
62	num_layers=4, dropout=True, epochs=50, batch_size=32	4787794.7	7 2188.1	1809.74	6.34	0.91	7.08	3033.25	0.003008	12.8
63	num_layers=4, dropout=True, epochs=50, batch_size=64	3343631.9	941828.56	1419.34	3.9	0.93	3.15	1456.14	0.002033	9.15
64	num_layers=4, dropout=False, epochs=10, batch_size=16	329796.62	2 574.28	405.4	1.3	0.99	1.8	1021.68	-0.000034	0.87
65	num_layers=4, dropout=False, epochs=10, batch_size=32	696324.13	3 834.46	676.24	2.21	0.99	2.04	476.84	0.000186	1.65
66	num_layers=4, dropout=False, epochs=10, batch_size=64	422199.29	9 649.77	474.31	1.52	0.99	1.99	263.19	-0.000053	0.84
67	num_layers=4, dropout=False, epochs=20, batch_size=16	250323.2	500.32	338.36	1.07	1	1.6	2156.1	-0.000040	0.81
68	num_layers=4, dropout=False, epochs=20, batch_size=32	299322.6	547.1	389.53	1.25	0.99	1.68	1013.48	-0.000020	0.91
69	num_layers=4, dropout=False, epochs=20, batch_size=64	496319.53	3 704.5	554.41	1.74	0.99	1.73	482.08	0.000096	1.39
70	num_layers=4, dropout=False, epochs=50, batch_size=16	225944.6	475.34	330.75	1.06	1	1.54	5066.28	0.000001	1
71	num_layers=4, dropout=False, epochs=50, batch_size=32	221558.31	1 470.7	319.77	1.02	1	1.52	2366.15	-0.000012	0.92

Table 5.2: Results of the ablation study of the Stacked LSTM architecture.

			s or the a		oracj or			Training		
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVDMC	Ü	Overfit	Loss
ID							CVRMS		Gap	Ratio
								(s)		
72	num_layers=4, dropout=False, epochs=50, batch_size=64	224493.78	8 473.81	319.47	1.02	1	1.52	1197.92	-0.000030	0.84
73	num_layers=5, dropout=True, epochs=10, batch_size=16	3761567.	161939.48	1606.07	4.6	0.93	4.17	1261.27	0.002182	6.72
74	num_layers=5, dropout=True, epochs=10, batch_size=32	1428606.0	011195.24	888.97	2.72	0.97	3.53	627.59	0.000484	1.99
75	num_layers=5, dropout=True, epochs=10, batch_size=64	2882271.	751697.73	1300.61	3.75	0.94	4.2	295.42	0.001285	2.89
76	num_layers=5, dropout=True, epochs=20, batch_size=16	1666747.	721291.03	993.54	3.34	0.97	4.49	2490.76	0.000808	3.47
77	num_layers=5, dropout=True, epochs=20, batch_size=32	1820279.0	061349.18	1074.46	3.23	0.96	3.86	1251.47	0.000902	3.66
78	num_layers=5, dropout=True, epochs=20, batch_size=64	3355491.	371831.8	1540.28	4.66	0.93	5.25	601.12	0.001909	6.08
79	num_layers=5, dropout=True, epochs=50, batch_size=16	2127101.4	481458.46	1163.71	4.21	0.96	5	6354.87	0.001207	5.97
80	num_layers=5, dropout=True, epochs=50, batch_size=32	2488392.0	621577.46	1234.49	3.58	0.95	4.01	3134.16	0.001454	7.03
81	num_layers=5, dropout=True, epochs=50, batch_size=64	3720370.	521928.83	1640.26	5.77	0.93	6.21	1536.97	0.002297	10.64
82	num_layers=5, dropout=False, epochs=10, batch_size=16	51070843	5.52146.39	5946.3	19.69	-0	24.06	1211.33	0.000225	1.01
83	num_layers=5, dropout=False, epochs=10, batch_size=32	313147.58	8 559.6	397.06	1.26	0.99	1.75	633.6	-0.000059	0.78
84	num_layers=5, dropout=False, epochs=10, batch_size=64	486057.6	5 697.18	522.86	1.65	0.99	2.06	298.31	0.000011	1.03
85	num_layers=5, dropout=False, epochs=20, batch_size=16	249225.3	499.22	347.99	1.11	1	1.58	2531.31	-0.000038	0.82
86	num_layers=5, dropout=False, epochs=20, batch_size=32	271780.9	1 521.33	365.49	1.15	0.99	1.64	1253.22	-0.000041	0.82
87	num_layers=5, dropout=False, epochs=20, batch_size=64	311775.92	2 558.37	394.17	1.28	0.99	1.86	604.83	-0.000038	0.85
88	num_layers=5, dropout=False, epochs=50, batch_size=16	267145.9	4 516.86	365.51	1.15	0.99	1.5	6359.32	0.000030	1.2
89	num_layers=5, dropout=False, epochs=50, batch_size=32	208583.58	8 456.71	313.34	1	1	1.46	3158.46	-0.000012	0.92
90	num_layers=5, dropout=False, epochs=50, batch_size=64	241447.7	4 491.37	346.76	1.11	1	1.53	1540.72	-0.000001	1

Table 5.3: Results of the ablation study of the LSTMCNN architecture.

		5.3: Resul						Training		
ID	G 0	MCE	DMCE	MAE	MADE	$\mathbf{R^2}$	CVDMC	J	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	R²	CVRMS		Gap	Ratio
								(s)		
1	num_layers=1, dropout=True, epochs=10,	338327.22	581.66	413.66	1.32	0.99	1.88	476.64	-0.000100	0.7
2	batch_size=16 num_layers=1, dropout=True, epochs=10,	256300.33	506.26	355.71	1.15	0.99	1.66	254.99	-0.000065	0.73
	batch_size=32 num_layers=1, dropout=True,									
3	epochs=10, batch_size=64 num_layers=1,	256364.57	506.32	358.68	1.16	0.99	1.61	130.68	-0.000016	0.92
4	dropout=True, epochs=20, batch_size=16	264072.72	513.88	365.47	1.18	0.99	1.68	996.61	-0.000022	0.89
5	num_layers=1, dropout=True, epochs=20, batch_size=32	229742.24	479.31	335.93	1.06	1	1.51	403.85	0.000001	1.01
6	num_layers=1, dropout=True, epochs=20, batch_size=64	227285.07	476.74	332.7	1.05	1	1.5	180.3	0.000024	1.18
7	num_layers=1, dropout=True, epochs=50, batch_size=16	240245.8	490.15	352.5	1.13	1	1.53	1853.75	0.000009	1.06
8	num_layers=1, dropout=True, epochs=50, batch_size=32	212389.7	460.86	327.96	1.03	1	1.45	889.15	0.000047	1.48
9	num_layers=1, dropout=True, epochs=50, batch_size=64	200016.09	447.23	314.3	1	1	1.43	450.19	0.000059	1.77
10	num_layers=1, dropout=False, epochs=10, batch_size=16	290735.93	539.2	377.8	1.21	0.99	1.75	365.02	-0.000079	0.71
11	num_layers=1, dropout=False, epochs=10, batch_size=32	260507.96	510.4	349.33	1.1	0.99	1.57	215.42	-0.000028	0.86
12	num_layers=1, dropout=False, epochs=10, batch_size=64	292076.55	540.44	391.77	1.22	0.99	1.53	117.76	0.000023	1.13
13	num_layers=1, dropout=False, epochs=20, batch_size=16	240846.25	490.76	341.15	1.09	1	1.59	868.49	-0.000020	0.89
14	num_layers=1, dropout=False, epochs=20, batch_size=32	219656.56	468.68	320.95	1.02	1	1.49	354.26	0.000002	1.01
15	num_layers=1, dropout=False, epochs=20, batch_size=64	215572.44	464.3	318.97	1.01	1	1.47	179.73	0.000022	1.18
16	num_layers=1, dropout=False, epochs=50, batch_size=16	262795.61	512.64	348.66	1.1	0.99	1.6	1843.73	0.000054	1.43
17	num_layers=1, dropout=False, epochs=50, batch_size=32	222722.33	471.93	322.81	1.02	1	1.5	1132.94	0.000059	1.64
18	num_layers=1, dropout=False, epochs=50, batch_size=64	217581.76	466.46	323.12	1.03	1	1.5	591.31	0.000092	2.62
19	num_layers=2, dropout=True, epochs=10, batch_size=16	451867.08	672.21	501.24	1.62	0.99	1.95	793.88	-0.000061	0.83
20	num_layers=2, dropout=True, epochs=10, batch_size=32	281182.72	530.27	377.68	1.21	0.99	1.66	416.39	-0.000053	0.78

Table 5.3: Results of the ablation study of the LSTMCNN architecture.

								Training	Overet	
ID	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	${f R^2}$	CVRMS	E Time	Overfit	Loss
								(s)	Gap	Ratio
21	num_layers=2, dropout=True, epochs=10, batch_size=64	244823.43	494.8	345.26	1.09	1	1.55	226.39	-0.000037	0.82
22	num_layers=2, dropout=True, epochs=20, batch_size=16	313607.87	560.01	405.07	1.26	0.99	1.74	1675.74	-0.000002	0.99
23	num_layers=2, dropout=True, epochs=20, batch_size=32	219451.9	468.46	325.64	1.04	1	1.52	823.08	-0.000016	0.9
24	num_layers=2, dropout=True, epochs=20, batch_size=64	217581.77	466.46	325.63	1.04	1	1.51	388.48	0.000009	1.07
25	num_layers=2, dropout=True, epochs=50, batch_size=16	229610.34	479.18	337.44	1.08	1	1.52	4231.23	0.000007	1.05
26	num_layers=2, dropout=True, epochs=50, batch_size=32 num_layers=2,	216926.85	465.75	331.38	1.06	1	1.47	1625.05	0.000045	1.43
27	dropout=True, epochs=50, batch_size=64	200798.33	448.11	314.22	0.99	1	1.41	782.53	0.000059	1.75
28	num_layers=2, dropout=False, epochs=10, batch_size=16	614730.07	784.05	591.04	1.8	0.99	1.82	572.64	0.000147	1.54
29	num_layers=2, dropout=False, epochs=10, batch_size=32	263973.33	513.78	351.91	1.11	0.99	1.61	278.26	-0.000020	0.9
30	num_layers=2, dropout=False, epochs=10, batch_size=64	224528.81	473.84	322.74	1.02	1	1.52	140.78	-0.000018	0.89
31	num_layers=2, dropout=False, epochs=20, batch_size=16	287447.75	536.14	377.05	1.18	0.99	1.56	1131.2	0.000017	1.1
32	num_layers=2, dropout=False, epochs=20, batch_size=32	222159.1	471.34	322.69	1.02	1	1.51	545.19	0.000009	1.06
33	num_layers=2, dropout=False, epochs=20, batch_size=64	207125.69	455.11	310.19	0.99	1	1.46	273.02	0.000021	1.18
34	num_layers=2, dropout=False, epochs=50, batch_size=16	218627.86	467.58	324.21	1.04	1	1.52	2811.4	0.000029	1.25
35	num_layers=2, dropout=False, epochs=50, batch_size=32	218120.13	467.03	331.72	1.04	1	1.46	1349.76	0.000066	1.8
36	num_layers=2, dropout=False, epochs=50, batch_size=64	199339.95	446.48	313.25	1	1	1.44	684.64	0.000084	2.61
37	num_layers=3, dropout=True, epochs=10, batch_size=16	493530.74	702.52	515.99	1.56	0.99	1.95	399.16	0.000022	1.07
38	num_layers=3, dropout=True, epochs=10, batch_size=32	244475.61	494.44	343.41	1.09	1	1.59	191.18	-0.000058	0.74
39	num_layers=3, dropout=True, epochs=10, batch_size=64	228717.5	478.24	330.03	1.05	1	1.53	97.33	-0.000020	0.89
40	num_layers=3, dropout=True, epochs=20, batch_size=16	280393.82	529.52	385.04	1.22	0.99	1.61	790.28	-0.000000	1

Table 5.3: Results of the ablation study of the LSTMCNN architecture.

Traini ID Config. MSE RMSE MAE MAPE R ² CVRMSE Time (s) 41	Overfit Gap 3 0.000037	Loss Ratio
(s) 41	Gap 3 0.000037	
11 num_layers=3, dropout=True, epochs=20, batch_size=3, dropout=True, epochs=20, batch_size=32, num_layers=3, dropout=True, epochs=20, batch_size=64 225070.43 474.42 336.56 1.06 1 1.5 192.9	3 0.000037	1.25
41 dropout=True, epochs=20, batch_size=32 num_layers=3, dropout=True, epochs=20, batch_size=64 225070.43 474.42 336.56 1.06 1 1.5 192.9		1.25
41 epochs=20, 212079.91 522.19 502.15 1.21 0.99 1.35 581.0 batch_size=32 num_layers=3, dropout=True, epochs=20, batch_size=64 225070.43 474.42 336.56 1.06 1 1.5 192.9		1.25
num_layers=3, dropout=True, epochs=20, batch_size=64 225070.43 474.42 336.56 1.06 1 1.5 192.9	6 0.000044	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	6 0.000044	
$batch_size=64$		1.4
1		
num_layers=3, dropout=True, 405401.45.636.71 413.1 1.31 0.00 1.87 1076.5	76 0 000110	1.70
43 dropout=1rue, 405401.45 636.71 413.1 1.31 0.99 1.87 1976.7 batch size=16	6 0.000119	1.76
num_layers=3,		
44 dropout=True, 190439.42 436.39 299.22 0.95 1 1.41 955.3	8 0.000038	1.41
batch_size=32		
num_layers=3, 45	6 0.000067	2.04
epochs=50, 131403.40 401.00 001.00 0.50 1 1.00 402.0 batch_size=64	0.000001	2.04
num_layers=3,		
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	7 -0.000067	0.76
batch_size=16 num_layers=3,		
47 dropout=False, 305954 29 553 13 372 8 1 18 0 99 1 73 182 3	7 -0.000003	0.98
epochs=10, 500001:25 500:15 572:5 1:16 5:55 1:17 5:25.5 batch_size=32		
num_layers=3, 40 dropout=False, 249400 46 409 5 246 12 1 1 1 1 1 59 09 69	0.000006	0.07
48 dropout=raise, 248499.46 498.5 346.13 1.1 1 1.58 92.65 batch_size=64	-0.000006	0.97
num_layers=3,		
49 dropout=False, 281927.85 530.97 375.29 1.2 0.99 1.7 759.0	7 0.000002	1.01
batch_size=16		
num_layers=3, 50	7 0.000012	1.08
epochs=20, 240000:10 450:02 041.45 1:05 1 1:04 500:0 batch_size=32	. 0.000012	1.00
num_layers=3, 51 dropout=False, 220527.06 460.6 322.04 1.02 1 1.51 1.88	0.000004	4.40
epochs=20, 220027.00 403.0 522.54 1.02 1 1.01 100	0.000024	1.19
$batch_size=64$ $num_layers=3$,		
52 dropout=False, 298935.74 546.75 398.97 1.25 0.99 1.68 1916.8	0.000072	1.55
batch_size=16		
num_layers=3, 53	9 0.000068	1.76
epochs=50, batch_size=32	9 0.000000	1.70
num_layers=3,		
$54 \frac{\text{dropout=False,}}{\text{epochs=}50,} 222837.19 472.06 331.22 1.06 1 1.5 468.6$	4 0.000094	2.6
batch_size=64 num_layers=4,		
55 dropout=True, 428328 07 654 47 486 09 1 51 0 99 1 94 540 5	8 -0.000077	0.79
epochs=10, 125625:07 501:17 150:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1:07 5:50 1		
num_layers=4, 56	0.000040	0.04
56 dropout=1rue, 302299.65 549.82 396 1.25 0.99 1.69 278.5 batch_size=32	5 -0.000040	0.84
$num_{layers}=4,$		
57 dropout=True, 266604.76 516.34 364.99 1.18 0.99 1.61 133.6	4 -0.000017	0.92
batch_size=64 num_layers=4,		
58 dropout=True, 278298 84 527 54 372 29 1 16 0 99 1 62 1085 5	28 -0.000024	0.89
epochs=20, batch_size=16		2.00
num_layers=4, 50 dropout=True, 202212.5 540.57 380.84 1.2 0.00 1.72 524.4		
99 epochs=20, 292212.9 940.97 900.04 1.2 0.99 1.72 924.4	7 0.000009	1.05
batch_size=32 num_layers=4,		
60 dropout=True, 207524.23 455.55 312.61 0.99 1 1.47 274.3	1 0.000018	1.14
batch_size=64		

Table 5.3: Results of the ablation study of the LSTMCNN architecture.

								Training		
ID	Camba	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	Ü	Overfit	Loss
ш	Config.	MSE	RWISE	MAL	MAFE	n	CVRIVIS	(s)	Gap	Ratio
61	num_layers=4, dropout=True, epochs=50, batch_size=16	247569.56	497.56	354.23	1.16	1	1.62	2728.43	0.000015	1.1
62	num_layers=4, dropout=True, epochs=50, batch_size=32	205792.53	453.64	313.99	1	1	1.46	1349.3	0.000039	1.38
63	num_layers=4, dropout=True, epochs=50, batch_size=64	198022.2	445	310.49	0.98	1	1.41	685.59	0.000062	1.84
64	num_layers=4, dropout=False, epochs=10, batch_size=16	319114.99	564.9	398.82	1.27	0.99	1.73	529.26	-0.000074	0.75
65	num_layers=4, dropout=False, epochs=10, batch_size=32	357871.21	598.22	436.25	1.33	0.99	1.59	259.41	0.000030	1.14
66	num_layers=4, dropout=False, epochs=10, batch_size=64	274078.89	523.53	363.27	1.16	0.99	1.6	131.23	0.000012	1.07
67	num_layers=4, dropout=False, epochs=20, batch_size=16	277739.01	527.01	352.15	1.11	0.99	1.66	1055.26	-0.000003	0.98
68	num_layers=4, dropout=False, epochs=20, batch_size=32	232515.6	482.2	328.41	1.04	1	1.53	516.8	0.000003	1.02
69	num_layers=4, dropout=False, epochs=20, batch_size=64	210168.8	458.44	307.84	0.98	1	1.47	260.86	0.000016	1.12
70	num_layers=4, dropout=False, epochs=50, batch_size=16	237731.18	487.58	344.49	1.07	1	1.49	2638.88	0.000043	1.36
71	num_layers=4, dropout=False, epochs=50, batch_size=32	212653.23	461.14	324.18	1.03	1	1.43	1296.93	0.000072	1.99
72	num_layers=4, dropout=False, epochs=50, batch_size=64	204065.03	451.74	314.33	1	1	1.43	655.77	0.000090	2.85
73	num_layers=5, dropout=True, epochs=10, batch_size=16	509895.41	714.07	529.39	1.63	0.99	2.09	778.76	-0.000033	0.91
74	num_layers=5, dropout=True, epochs=10, batch_size=32	477905.9	691.31	514.08	1.54	0.99	1.92	380	0.000064	1.25
75	num_layers=5, dropout=True, epochs=10, batch_size=64	378562.52	615.27	453.63	1.42	0.99	1.95	193.3	0.000043	1.2
76	num_layers=5, dropout=True, epochs=20, batch_size=16	364388.54	603.65	455.87	1.5	0.99	1.87	1551.51	0.000017	1.07
77	num_layers=5, dropout=True, epochs=20, batch_size=32	396317.47	629.54	463.91	1.44	0.99	1.98	762.55	0.000080	1.42
78	num_layers=5, dropout=True, epochs=20, batch_size=64	359369.7	599.47	444.02	1.42	0.99	1.93	387.24	0.000080	1.48
79	num_layers=5, dropout=True, epochs=50, batch_size=16	456851.02	675.91	520.35	1.71	0.99	2.28	3883.5	0.000114	1.58
80	num_layers=5, dropout=True, epochs=50, batch_size=32	518313.86	719.94	560.56	1.79	0.99	2.33	1911.11	0.000237	3.04

APPENDIX 133

Table 5.3: Results of the ablation study of the LSTMCNN architecture.

					-			Training	Overfit	Loss
ID	Config.	MSE	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMS	E Time	Gap	Ratio
								(s)	Gap	natio
81	num_layers=5, dropout=True, epochs=50, batch_size=64	533333.41	730.3	582.93	1.94	0.99	2.43	965.25	0.000283	4.51
82	num_layers=5, dropout=False, epochs=10, batch_size=16	446857.43	668.47	496.5	1.54	0.99	1.87	729.8	0.000004	1.01
83	num_layers=5, dropout=False, epochs=10, batch_size=32	299745.76	547.49	384.03	1.26	0.99	1.85	355.97	-0.000011	0.95
84	num_layers=5, dropout=False, epochs=10, batch_size=64	272502.59	522.02	360.95	1.15	0.99	1.62	180.02	0.000012	1.07
85	num_layers=5, dropout=False, epochs=20, batch_size=16	286510.94	535.27	380.94	1.21	0.99	1.72	1459.52	-0.000001	0.99
86	num_layers=5, dropout=False, epochs=20, batch_size=32	218200.59	467.12	321.38	1.02	1	1.51	720.49	-0.000002	0.99
87	num_layers=5, dropout=False, epochs=20, batch_size=64	230579.73	480.19	334.35	1.08	1	1.53	364.1	0.000040	1.34
88	num_layers=5, dropout=False, epochs=50, batch_size=16	253615.19	503.6	362.85	1.18	1	1.6	3615.78	0.000043	1.33
89	num_layers=5, dropout=False, epochs=50, batch_size=32	195609.61	442.28	308.61	0.98	1	1.42	1833.34	0.000057	1.74
90	num_layers=5, dropout=False, epochs=50, batch_size=64	199303.86	446.43	312.51	1	1	1.44	952.49	0.000088	2.86

 ${\bf Table~5.4:}~{\bf Results~of~the~ablation~study~of~the~Stacked~LSTM~architecture.}$

ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	Training E Time (s)	Overfit Gap	Loss Ratio
1	num_layers=1, dropout=True, epochs=10, batch_size=16	508043.99	712.77	517.91	1.63	0.99	2.22	299.65	-0.000212	0.62
2	num_layers=1, dropout=True, epochs=10, batch_size=32	874665.43	935.24	719.97	2.23	0.98	2.4	148.18	-0.000022	0.96
3	num_layers=1, dropout=True, epochs=10, batch_size=64	1078683.2	1038.6	796.51	2.38	0.98	2.73	74.4	0.000028	1.04
4	num_layers=1, dropout=True, epochs=20, batch_size=16	512159.85	715.65	511.87	1.57	0.99	2.05	598.86	-0.000075	0.82
5	num_layers=1, dropout=True, epochs=20, batch_size=32	387530.64	622.52	434.01	1.4	0.99	2.07	287.17	-0.000200	0.57
6	num_layers=1, dropout=True, epochs=20, batch_size=64	399058.48	631.71	449.94	1.46	0.99	2.06	144.6	-0.000203	0.57
7	num_layers=1, dropout=True, epochs=50, batch_size=16	446699.59	668.36	481.75	1.45	0.99	1.96	1974.53	0.000024	1.09

Table 5.4: Results of the ablation study of the Stacked LSTM architecture.

								Training	Ower	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	${f R^2}$	CVRMS	SE Time	Overfit	
								(s)	Gap	Ratio
8	num_layers=1, dropout=True, epochs=50, batch_size=32	334627.18	578.47	398.04	1.24	0.99	1.81	853.2	-0.000078	0.75
9	num_layers=1, dropout=True, epochs=50, batch_size=64	345288.76	587.61	412.6	1.3	0.99	1.89	378.99	-0.000107	0.69
10	num_layers=1, dropout=False, epochs=10, batch_size=16	417110.59	645.84	486.64	1.56	0.99	1.85	284.43	-0.000005	0.98
11	num_layers=1, dropout=False, epochs=10, batch_size=32	471071.66	686.35	498.84	1.53	0.99	1.88	139.12	0.000030	1.1
12	num_layers=1, dropout=False, epochs=10, batch_size=64	360953.98	600.79	415.06	1.33	0.99	1.92	72.77	-0.000099	0.71
13	num_layers=1, dropout=False, epochs=20, batch_size=16	320375.36	566.02	395.92	1.26	0.99	1.84	575.22	-0.000017	0.93
14	num_layers=1, dropout=False, epochs=20, batch_size=32	373224.63	610.92	452.43	1.45	0.99	1.68	273.38	0.000020	1.09
15	num_layers=1, dropout=False, epochs=20, batch_size=64	388286.18	623.13	463.43	1.46	0.99	1.78	136.29	0.000020	1.08
16	num_layers=1, dropout=False, epochs=50, batch_size=16	245022	495	333.49	1.07	1	1.62	1409.65	-0.000011	0.94
17	num_layers=1, dropout=False, epochs=50, batch_size=32	295516.17	543.61	381.74	1.22	0.99	1.62	684.3	0.000020	1.11
18	num_layers=1, dropout=False, epochs=50, batch_size=64	291143.76	539.58	373.83	1.18	0.99	1.6	341.12	0.000021	1.12
19	num_layers=2, dropout=True, epochs=10, batch_size=16	628984.93	793.09	591.33	1.89	0.99	2.58	303.33	-0.000199	0.68
20	num_layers=2, dropout=True, epochs=10, batch_size=32	774379.41	879.99	646.67	2.17	0.98	3.01	151.41	-0.000158	0.77
21	num_layers=2, dropout=True, epochs=10, batch_size=64	845085.32	919.29	670.38	2.06	0.98	2.75	76.5	-0.000169	0.77
22	num_layers=2, dropout=True, epochs=20, batch_size=16	720776.27	848.99	644.95	1.95	0.99	2.48	609.55	0.000040	1.09
23	num_layers=2, dropout=True, epochs=20, batch_size=32	1512568.5	51229.87	902.35	2.54	0.97	2.82	310.63	0.000577	2.27
24	num_layers=2, dropout=True, epochs=20, batch_size=64	1261253.0	71123.06	871.86	2.55	0.97	2.77	153.13	0.000352	1.69
25	num_layers=2, dropout=True, epochs=50, batch_size=16	1160881.1	91077.44	828.96	2.46	0.98	3.06	1550.41	0.000478	2.52
26	num_layers=2, dropout=True, epochs=50, batch_size=32	1181495.9	31086.97	807.42	2.34	0.98	2.89	771.12	0.000485	2.52
27	num_layers=2, dropout=True, epochs=50, batch_size=64	2942853.2	41715.47	1254.91	3.45	0.94	3.53	384.95	0.001668	5.98

Table 5.4: Results of the ablation study of the Stacked LSTM architecture.

		4: nesums						Training		
ID		MOD	DAGE	3.645	MADE	D 2	CLIDAG	Ü	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	E Time (s)	Gap	Ratio
28	num_layers=2, dropout=False, epochs=10, batch_size=16	318744.52	564.57	388.81	1.25	0.99	1.86	276.75	-0.000085	0.72
29	num_layers=2, dropout=False, epochs=10, batch_size=32	386719.02	621.87	436.98	1.4	0.99	2.01	137.36	-0.000046	0.85
30	num_layers=2, dropout=False, epochs=10, batch_size=64	375478.17	612.76	435.01	1.38	0.99	1.94	69.63	-0.000052	0.83
31	num_layers=2, dropout=False, epochs=20, batch_size=16	300515.88	548.19	374.79	1.2	0.99	1.78	566.6	-0.000035	0.85
32	num_layers=2, dropout=False, epochs=20, batch_size=32	520345.28	721.35	561.23	1.72	0.99	1.68	278.31	0.000116	1.48
33	num_layers=2, dropout=False, epochs=20, batch_size=64	996910.24	998.45	860.17	2.72	0.98	2.03	138.86	0.000441	2.84
34	num_layers=2, dropout=False, epochs=50, batch_size=16	289117.65	537.7	374	1.17	0.99	1.61	1381.91	0.000022	1.12
35	num_layers=2, dropout=False, epochs=50, batch_size=32	240775.73	490.69	334.04	1.07	1	1.62	686.43	-0.000016	0.91
36	num_layers=2, dropout=False, epochs=50, batch_size=64	279473.69	528.65	373.09	1.16	0.99	1.63	341.91	0.000025	1.15
37	num_layers=3, dropout=True, epochs=10, batch_size=16	591300.23	768.96	546.49	1.7	0.99	2.26	413.76	-0.000203	0.66
38	num_layers=3, dropout=True, epochs=10, batch_size=32	613262.78	783.11	571.35	1.86	0.99	2.61	209.07	-0.000251	0.62
39	num_layers=3, dropout=True, epochs=10, batch_size=64	770258.49	877.64	652.48	2.05	0.98	2.78	105.07	-0.000303	0.63
40	num_layers=3, dropout=True, epochs=20, batch_size=16	433907.35	658.72	465.11	1.5	0.99	2.16	841.72	-0.000144	0.67
41	num_layers=3, dropout=True, epochs=20, batch_size=32	501137.65	707.91	519.97	1.64	0.99	2.26	405.66	-0.000103	0.77
42	num_layers=3, dropout=True, epochs=20, batch_size=64	729943.47	854.37	616.04	1.84	0.99	2.42	204.46	-0.000012	0.98
43	num_layers=3, dropout=True, epochs=50, batch_size=16	421552.89	649.27	470.45	1.49	0.99	2.05	2042.96	-0.000013	0.96
44	num_layers=3, dropout=True, epochs=50, batch_size=32	560162.63	748.44	564.21	1.83	0.99	2.08	1016.52	0.000070	1.22
45	num_layers=3, dropout=True, epochs=50, batch_size=64	1004884.3	81002.44	798.3	2.51	0.98	3.09	513.16	0.000357	2.09
46	num_layers=3, dropout=False, epochs=10, batch_size=16	321503.39	567.01	399.17	1.29	0.99	1.86	384.16	-0.000075	0.74
47	num_layers=3, dropout=False, epochs=10, batch_size=32	864406.88	929.73	762.46	2.35	0.98	1.93	188.14	0.000290	1.97

Table 5.4: Results of the ablation study of the Stacked LSTM architecture.

						_		Training	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	SE Time (s)	Gap	Ratio
48	num_layers=3, dropout=False, epochs=10, batch_size=64	341600.56	584.47	401.04	1.28	0.99	1.9	95.99	-0.000093	0.71
49	num_layers=3, dropout=False, epochs=20, batch_size=16 num_layers=3,	492356.58	701.68	536.69	1.69	0.99	1.75	769.35	0.000101	1.43
50	num_layers=3, dropout=False, epochs=20, batch_size=32 num_layers=3,	371477.06	609.49	436.53	1.35	0.99	1.73	421.95	0.000007	1.03
51	dropout=False, epochs=20, batch_size=64 num_layers=3,	418665.9	647.04	483.62	1.54	0.99	1.77	194.83	0.000033	1.13
52	dropout=False, epochs=50, batch_size=16 num_layers=3,	235469.59	485.25	324.19	1.03	1	1.56	1841.92	-0.000014	0.92
53	dropout=False, epochs=50, batch_size=32 num_layers=3,	474424.95	688.79	533.51	1.66	0.99	1.68	896.49	0.000147	1.84
54	dropout=False, epochs=50, batch_size=64 num_layers=4,	353874.01	594.87	444.04	1.38	0.99	1.62	460.71	0.000054	1.29
55	dropout=True, epochs=10, batch_size=16 num_layers=4,	1347342.0	41160.75	870	2.55	0.97	2.99	377.01	0.000237	1.35
56	dropout=True, epochs=10, batch_size=32 num_layers=4,	602899.24	776.47	564.23	1.77	0.99	2.46	192.76	-0.000310	0.57
57	dropout=True, epochs=10, batch_size=64 num_layers=4,	1178835.8	11085.74	825.5	2.52	0.98	3.19	97.72	-0.000061	0.93
58	dropout=True, epochs=20, batch_size=16 num_layers=4,	628965.31	793.07	560.72	1.73	0.99	2.35	752.54	-0.000024	0.95
59	dropout=True, epochs=20, batch_size=32 num_layers=4,	583591.15	763.93	581.14	1.82	0.99	2.41	378.44	-0.000097	0.8
60	dropout=True, epochs=20, batch_size=64 num_layers=4,	786549.74	886.88	643.3	1.94	0.98	2.48	192.63	-0.000031	0.95
61	dropout=True, epochs=50, batch_size=16 num_layers=4,	961362.69	980.49	778.4	2.36	0.98	2.89	1887.26	0.000346	2.12
62	dropout=True, epochs=50, batch_size=32 num_layers=4,	630574.71	794.09	589.73	1.74	0.99	2.18	946.61	0.000109	1.34
63	dropout=True, epochs=50, batch_size=64 num_layers=4,	2029474.2	91424.6	1111.83	3.27	0.96	3.92	522.59	0.001049	4.15
64	dropout=False, epochs=10, batch_size=16 num_layers=4,	388524.93	623.32	451.41	1.43	0.99	1.83	493.14	-0.000042	0.86
65	dropout=False, epochs=10, batch_size=32 num_layers=4,	1210524.9	71100.24	897.77	2.69	0.98	2.26	250.75	0.000493	2.49
66	dropout=False, epochs=10, batch_size=64 num_layers=4,	448902.39	670	481.41	1.51	0.99	2.03	132.2	-0.000034	0.9
67	dropout=False, epochs=20, batch_size=16	288569.86	537.19	374.29	1.2	0.99	1.74	1029.07	-0.000042	0.82

Table 5.4: Results of the ablation study of the Stacked LSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R ²	CVRMS	Training E Time (s)	Overfit Gap	Loss Ratio
68	num_layers=4, dropout=False, epochs=20, batch_size=32	308502.4	555.43	381.5	1.22	0.99	1.8	503.63	-0.000026	0.89
69	num_layers=4, dropout=False, epochs=20, batch_size=64	308070.97	555.04	389.37	1.25	0.99	1.76	255.14	-0.000041	0.84
70	num_layers=4, dropout=False, epochs=50, batch_size=16	243290.25	493.24	338.3	1.07	1	1.6	2059.35	-0.000003	0.98
71	num_layers=4, dropout=False, epochs=50, batch_size=32	247081.5	497.07	338.25	1.08	1	1.62	900.35	-0.000007	0.96
72	num_layers=4, dropout=False, epochs=50, batch_size=64	476455.75	690.26	519.64	1.55	0.99	1.78	451.39	0.000152	1.88

Table 5.5: Results of the ablation study of the BiLSTM architecture.

	Table	7.0. 1000	1105 01 01.	- ablati	on study	01 011	e BillSTN.	Training	- Caro.	
ID	Config.	MSE	RMSE	MAE	MAPE	$\mathbf{R^2}$	CVRMSI	Ü	Overfit	Loss
110	Comig.	WISL	TUISE	WIZE	1V1211 L	10	C V ICIVISI		Gap	Ratio
	2. 1. 1. 04							(s)		
1	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=Tru epochs=10, batch_size=32		640.76	430.01	1.35	0.99	1.94	365.3	-0.000074	0.79
2	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=Fals epochs=10, batch_size=32	589661.59	767.89	579.95	1.77	0.99	1.98	355.32	0.000058	1.17
3	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=Tru epochs=10, batch_size=32	477187.93 e,	690.79	492.32	1.6	0.99	2.23	346.99	0.000008	1.03
4	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=Fals epochs=10, batch_size=32	547839.7	740.16	535.04	1.71	0.99	2.39	339.48	-0.000042	0.9
5	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=Tru epochs=10, batch_size=64	1328547.6 e,	21152.63	957.53	3.07	0.97	2.53	185.19	0.000461	2.04
6	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=True, temporal_att=Fals epochs=10, batch_size=64	448395.44 se,	669.62	477.49	1.5	0.99	2.08	180.36	-0.000067	0.82
7	units_lstm1=64, units_lstm2=32, dropout_rate=0.0, feature_att=False, temporal_att=Tru epochs=10, batch_size=64	392220.73	626.28	443.74	1.44	0.99	1.98	177.28	-0.000088	0.75

Table 5.5: Results of the ablation study of the BiLSTM architecture.

							,	Training		
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	J	Overfit	Loss
מו	Comig.	WISE	IUMSE	WIAL	WIALE	11	CVIUNISE		Gap	Ratio
								(s)		
8	units_lstm1=64, units_lstm2=32, dropout_rate=0. feature_att=Fals temporal_att=Fals epochs=10,	e,657870.71	811.09	603.66	1.93	0.99	2.61	174.05	-0.000101	0.82
9	batch_size=64 units_lstm1=64, units_lstm2=32, dropout_rate=0. feature_att=Tructemporal_att=Tructemporal_att=Tructemporal_size=32	e, 313162.62	559.61	387.64	1.24	0.99	1.8	732.84	-0.000061	0.78
10	units_lstm1=64, units_lstm2=32, dropout_rate=0.feature_att=True temporal_att=Fa epochs=20, batch_size=32	e, 398778.35	631.49	465.23	1.47	0.99	1.81	710.16	0.000017	1.07
11	units_lstm1=64, units_lstm2=32, dropout_rate=0. feature_att=Fals temporal_att=Tr epochs=20, batch_size=32	0, e,323979.77 rue,	569.19	391.02	1.22	0.99	1.7	695.79	-0.000000	1
12	units_lstm1=64, units_lstm2=32, dropout_rate=0. feature_att=Fals temporal_att=Fa epochs=20, batch_size=32	0, e,301410.02 _{llse,}	549.01	389.43	1.24	0.99	1.69	675.19	-0.000027	0.88
13	units_lstm1=64, units_lstm2=32, dropout_rate=0. feature_att=True temporal_att=True epochs=20,	e, 365143.65	604.27	412.78	1.3	0.99	1.87	368.42	-0.000057	0.81
14	batch_size=64 units_lstm1=64, units_lstm2=32, dropout_rate=0. feature_att=Tru temporal_att=Fa epochs=20,	⁰ , 327926.67 _e , 327936.67	572.65	400.09	1.27	0.99	1.82	357.95	-0.000051	0.82
15	batch_size=64 units_lstm1=64, units_lstm2=32, dropout_rate=0. feature_att=Fals temporal_att=Trepochs=20,	e,350281.78	591.85	425.54	1.36	0.99	1.83	350.7	-0.000029	0.89
16	batch_size=64 units_lstm1=64, units_lstm2=32, dropout_rate=0. feature_att=Fals temporal_att=Fals epochs=20, batch_size=64	e,448543.69	669.73	490.61	1.53	0.99	1.99	352.03	-0.000041	0.88
17	units_lstm1=64, units_lstm2=32, dropout_rate=0. feature_att=Trutemporal_att=Tr epochs=50, batch_size=32	0, 262239.84 rue,	512.09	352.85	1.09	0.99	1.58	1842.44	-0.000008	0.96
18	units_lstm1=64, units_lstm2=32, dropout_rate=0. feature_att=True temporal_att=Fz epochs=50,	0, e, 228250.23 _{llse,}	477.76	326.87	1.04	1	1.52	1800.97	-0.000011	0.94
19	batch_size=32 units_lstm1=64, units_lstm2=32, dropout_rate=0. feature_att=Fals temporal_att=Tr epochs=50, batch_size=32	$^{0},_{ m e,220601.64}$ rue,	469.68	318.46	1.01	1	1.5	1765.21	-0.000019	0.89

 $\textbf{Table 5.5:} \ \operatorname{Results} \ \text{of the ablation study of the BiLSTM architecture}.$

								Training		
ID	a a Mc	יבוני	DMCE	MAT	MADE	D 2	CVDM	_	Overfit	Loss
ID	Config. MS	E	RMSE	MAE	MAPE	\mathbb{R}^2	CVKM	SE Time	Gap	Ratio
								(s)		
	units_lstm1=64, units_lstm2=32,									
20	dropout_rate=0.0, feature_att=False,25691	10.12	2 506.86	350.53	1.1	0.99	1.57	1719.88	0.000013	1.08
	temporal_att=False, epochs=50,									
	batch_size=32 units_lstm1=64,									
	units_lstm2=32,									
21	dropout_rate=0.0, feature_att=True, 24197	77.29	491.91	323.44	1.03	1	1.59	924.98	-0.000043	0.79
	temporal_att=True, epochs=50,									
	batch_size=64 units_lstm1=64,									
	units_lstm2=32,									
22	dropout_rate=0.0, feature_att=True, 28170 temporal_att=False,	01.17	530.76	381.52	1.22	0.99	1.56	921.71	0.000007	1.04
	epochs=50,									
	batch_size=64 units_lstm1=64,									
	units_lstm2=32, dropout_rate=0.0,	. .		0.40.00				20110	0.000044	0.04
23	feature_att=False,24435 temporal_att=True,	38.78	3 494.31	343.36	1.11	1	1.56	894.16	-0.000011	0.94
	epochs=50, batch_size=64									
	units_lstm1=64, units_lstm2=32,									
24	dropout_rate=0.0, feature_att=False,24581	17 9/	1 405 8	346.49	1.1	1	1.55	875.59	-0.000006	0.96
24	$temporal_att=False,$	17.54	495.6	340.49	1.1	1	1.55	610.09	-0.000000	0.90
	epochs=50, batch_size=64									
	units_lstm1=64, units_lstm2=32,									
25	dropout_rate=0.2, feature_att=True, 49488	83.47	703.48	489.56	1.56	0.99	2.29	374.65	-0.000130	0.72
	temporal_att=True, epochs=10,		, , , , , ,			0.00		0,1200	0.000	****
	batch_size=32									
	units_lstm1=64, units_lstm2=32,									
26	dropout_rate=0.2, feature_att=True, 65243	36.3	807.74	613.71	2.08	0.99	2.72	365.16	0.000023	1.05
	temporal_att=False, epochs=10,									
	batch_size=32 units_lstm1=64,									
	units_lstm2=32, dropout_rate=0.2,									
27	feature_att=False,57090 temporal_att=True,	04.92	755.58	552.37	1.7	0.99	1.99	359.4	-0.000001	1
	epochs=10,									
	batch_size=32 units_lstm1=64,									
20	units_lstm2=32, dropout_rate=0.2,			05001	2.00				0.000000	0.00
28	feature_att=False,77066 temporal_att=False,	56.91	877.88	659.01	2.06	0.98	2.35	355.57	-0.000038	0.93
	epochs=10, batch_size=32									
	units_lstm1=64, units_lstm2=32,									
29	dropout_rate=0.2, feature_att=True, 44988	RQ 51	670.74	467.18	1.48	0.99	2.15	192.6	-0.000159	0.66
23	temporal_att=True,	1	. 010.14	TO1.10	1.40	0.00	2.10	102.0	0.000103	0.00
	epochs=10, batch_size=64									
	units_lstm1=64, units_lstm2=32,									
30	dropout_rate=0.2, feature_att=True, 56664	49.29	752.76	524.97	1.64	0.99	2.33	189.09	-0.000067	0.85
	temporal_att=False, epochs=10,				-					
	batch_size=64 units_lstm1=64,									
	units_lstm2=32,									
31	dropout_rate=0.2, feature_att=False,42237	75.57	649.9	463.32	1.47	0.99	2.06	184.55	-0.000146	0.66
	temporal_att=True, epochs=10,									
	batch_size=64									

Table 5.5: Results of the ablation study of the BiLSTM architecture.

							e BiLSTM	Training		
						- 0		Ü	Overfit	$_{ m Loss}$
$^{\mathrm{ID}}$	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Cup	100010
	units_lstm1=64,									
	units_lstm2=32, dropout_rate=0.	2,								
32	feature_att=Fals temporal_att=Fa		806.83	584.55	1.87	0.99	2.52	180.87	-0.000237	0.65
	epochs=10,	,								
	batch_size=64 units_lstm1=64,									
	units_lstm2=32,	2								
33	dropout_rate=0. feature_att=True	e,628605.92	792.85	511.46	1.59	0.99	2.39	746.23	0.000133	1.45
	temporal_att=Trepochs=20,	rue,								
	batch_size=32									
	units_lstm1=64, units_lstm2=32,									
34	dropout_rate=0. feature_att=True	$_{ m e.}^{2,}$	715.17	555.61	1.72	0.99	1.78	739.72	0.000050	1.17
_	temporal_att=Fa									
	batch_size=32									
	units_lstm1=64, units_lstm2=32,									
35	dropout rate=0.	2,425628.00	660.02	513.57	1.61	0.99	1.64	725.5	0.000042	1.16
39	feature_att=Fals temporal_att=Tr	se,455026.95 rue,	000.02	515.57	1.01	0.99	1.04	120.0	0.000042	1.10
	epochs=20, batch_size=32									
	units_lstm1=64,									
0.0	units_lstm2=32, dropout_rate=0.	2,	- 04 4 -0				0.4		0.000040	0.04
36	feature_att=Fals temporal_att=Fa		614.76	447.22	1.47	0.99	2.1	697.51	-0.000048	0.84
	epochs=20, batch_size=32	,								
	units_lstm1=64,									
	units_lstm2=32, dropout_rate=0	2								
37	dropout_rate=0. feature_att=True		572.61	402.12	1.29	0.99	1.83	376.52	-0.000105	0.68
	temporal_att=Trepochs=20,	rue,								
	batch_size=64 units_lstm1=64,									
	units_lstm2=32,	0								
38	dropout_rate=0. feature_att=True	e, 557728.94	4 746.81	517.8	1.63	0.99	2.2	372.87	0.000075	1.24
	temporal_att=Fa epochs=20,	alse,								
	batch_size=64									
	units_lstm1=64, units_lstm2=32,	_								
39	dropout_rate=0. feature_att=Fals	2, se,353208.87	594.31	417.93	1.32	0.99	1.89	364.21	-0.000086	0.74
	temporal_att=Trepochs=20,									
	$batch_size=64$									
	units_lstm1=64, units_lstm2=32,									
40	dropout_rate=0. feature_att=Fals	2, se.493191.61	702.28	506.23	1.63	0.99	2.28	359.2	-0.000134	0.72
10	temporal_att=Fa	alse,		000.20	1.00	0.00	2.20	330.2	0.000101	0.12
	epochs=20, batch_size=64									
	units_lstm1=64, units_lstm2=32,									
41	dropout_rate=0. feature_att=True	2, 1080606 0	51043.84	730.88	2.03	0.98	2.21	1877.12	0.000533	3.54
41	temporal_att=Tr		01040.04	100.00	۵.05	0.30	4.41	1011.12	0.000000	0.04
	epochs=50, batch_size=32									
	units_lstm1=64, units_lstm2=32,									
40	dropout_rate=0. feature_att=True	2,000460 61	E90 41	976.16	1 15	0.00	1.01	1041.9	0.000000	4
42	feature_att=Tru- temporal_att=Fa	e, 283462.61 alse,	532.41	376.12	1.15	0.99	1.61	1841.3	0.000000	1
	epochs=50, batch_size=32	*								
	units_lstm1=64,									
	units_lstm2=32, dropout_rate=0.	2,								
43	feature_att=Fals temporal_att=Tr	se, 265332	515.1	364.17	1.12	0.99	1.55	1805.92	0.000004	1.02
	epochs=50,	,								
	batch_size=32									

 $\textbf{Table 5.5:} \ \operatorname{Results} \ \text{of the ablation study of the BiLSTM architecture}.$

	Table (5.0. 1005	d105 O1 01	ic ablat.	on study	01 011	C DILOT	M archited Training	ourc.	
			D. 600			5 2	GI ID I I	_	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	SE Time	Gap	Ratio
								(s)		
44	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=False epochs=50, batch_size=32 units_lstm1=64,	254523.96	504.5	350.94	1.1	1	1.58	1752.24	-0.000023	0.88
45	units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=Tru epochs=50, batch_size=64	285424.12 e,	534.25	373.3	1.2	0.99	1.73	943.62	-0.000032	0.86
46	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=True, temporal_att=Falsepochs=50, batch_size=64		499.85	340.5	1.08	1	1.6	903.81	-0.000052	0.77
47	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=Tru epochs=50, batch_size=64	,314084.29	560.43	415.37	1.26	0.99	1.54	868.3	0.000024	1.13
48	units_lstm1=64, units_lstm2=32, dropout_rate=0.2, feature_att=False, temporal_att=False, epochs=50, batch_size=64	241639.66	491.57	342.76	1.09	1	1.57	844.81	-0.000053	0.76
49	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=True epochs=10, batch_size=32	440429.67	663.65	452.19	1.44	0.99	2.11	369.47	-0.000085	0.78
50	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=Falsepochs=10, batch_size=32	398772.05 se,	631.48	458.06	1.5	0.99	2.06	375.73	-0.000069	0.8
51	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=Tru epochs=10, batch_size=32		540.48	372.14	1.2	0.99	1.78	358.22	-0.000095	0.68
52	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False temporal_att=False epochs=10, batch_size=32	,915443.88	956.79	761.24	2.4	0.98	2.43	356.94	0.000197	1.46
53	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=True epochs=10, batch_size=64	385519.87	620.9	434.54	1.38	0.99	2.01	188.72	-0.000051	0.84
54	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=True, temporal_att=Falsepochs=10, batch_size=64	502824.24	709.1	537.12	1.71	0.99	1.96	190.54	-0.000023	0.94
55	units_lstm1=128, units_lstm2=64, dropout_rate=0.0, feature_att=False, temporal_att=Tru epochs=10, batch_size=64	381167.23 e,	617.39	442.62	1.41	0.99	1.97	180.6	-0.000094	0.74

Table 5.5: Results of the ablation study of the BiLSTM architecture.

							ı	Training		
ID	Canta	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Overfit	Loss
ш	Config.	MISE	IUVISE	WIAL	WAIL	11	CVICIOSI		\mathbf{Gap}	Ratio
								(s)		
	units_lstm1=128 units_lstm2=64,	,								
56	dropout_rate=0. feature_att=Fals	$_{ m e.629979.99}^{ m 0,}$	793.71	592.38	1.9	0.99	2.45	182.56	-0.000110	0.8
	temporal_att=Fa	ilse,		00-100		0.00			0.0000	0.0
	epochs=10, batch_size=64									
	units_lstm1=128 units_lstm2=64,	,								
57	dropout_rate=0. feature_att=Tru	0, a 415894 7	644.9	473.1	1.55	0.99	1.87	737.43	0.000011	1.04
٠.	temporal_att=Tr	rue,	011.0	1.0.1	1.00	0.00	1.01	101110	0.000011	1.01
	epochs=20, batch_size=32									
	units_lstm1=128 units_lstm2=64,	,								
58	dropout_rate=0. feature_att=Tru	0, • 437778 37	661 65	503.64	1.55	0.99	1.72	741.08	0.000044	1.17
00	temporal_att=Fa		001.00	000.01	1.00	0.00		. 11.00	0.000011	1.1.
	epochs=20, batch_size=32									
	units_lstm1=128 units_lstm2=64,	,								
59	dropout_rate=0. feature_att=Fals	0, e,283300.68	532.26	376.34	1.18	0.99	1.62	705.78	-0.000028	0.87
	temporal_att=Trepochs=20,		00-1-0	0,0.0-		0.00			0.0000=0	
	$batch_size=32$									
	units_lstm1=128 units_lstm2=64,									
60	dropout_rate=0. feature_att=Fals	$_{ m e,288947.75}^{ m 0,}$	537.54	374.17	1.2	0.99	1.73	713.76	-0.000048	0.8
	temporal_att=Fa	ılse,								
	batch_size=32									
	units_lstm1=128 units_lstm2=64,									
61	dropout_rate=0. feature_att=Tru	$_{\rm e, 395405.08}^{0,}$	628.81	440.61	1.43	0.99	2.05	378.17	-0.000146	0.65
	temporal_att=Trepochs=20,	rue,								
	batch_size=64									
	units_lstm1=128 units_lstm2=64,									
62	dropout_rate=0. feature_att=Tru	$_{\rm e,}^{0,}332521.75$	576.65	392.37	1.25	0.99	1.85	381.17	-0.000043	0.84
	temporal_att=Fa epochs=20,	ılse,								
	batch_size=64 units_lstm1=128	:								
	units_lstm2=64,									
63	dropout_rate=0. feature_att=Fals	e,267120.38	516.84	350.1	1.12	0.99	1.68	359.66	-0.000047	0.79
	temporal_att=Trepochs=20,	rue,								
	batch_size=64 units_lstm1=128									
	units_lstm2=64, dropout_rate=0.									
64	feature_att=Fals temporal_att=Fa	e, 512357.42	715.79	532	1.73	0.99	2.29	365.24	-0.000016	0.96
	epochs=20,	iise,								
	batch_size=64 units_lstm1=128	,								
	units_lstm2=64, dropout_rate=0.	0,								
65	dropout_rate=0. feature_att=Tru temporal_att=Tr	e, 254926.97	504.9	345.37	1.11	1	1.6	1832.69	-0.000014	0.93
	epochs=50, batch_size=32	- /								
	units_lstm1=128	,								
0.0	units_lstm2=64, dropout_rate=0. feature_att=Tru	0,242545.	100.00	00= 0	1.00			1040 10	0.000001	4.04
66	feature_att=Tru temporal_att=Fa		492.28	337.8	1.08	1	1.57	1848.43	0.000001	1.01
	epochs=50, batch_size=32									
	units_lstm1=128	,								
67	units_lstm2=64, dropout_rate=0.	0, 142022 62	402 77	226 00	1.07	1	1 50	1764.24	0.000001	1.01
67	feature_att=Fals temporal_att=Ti	e,242823.08 rue,	492.11	336.88	1.07	1	1.58	1764.34	0.000001	1.01
	epochs=50, batch_size=32									

 $\textbf{Table 5.5:} \ \ \text{Results of the ablation study of the BiLSTM architecture}.$

								Training		
ID	G 0	MCE	DMCE	MAE	MADE	\mathbf{D}^2	CVDM	Ü	Overfit	$_{\rm Loss}$
ID	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	\mathbb{R}^2	CVRM	SE Time	Gap	Ratio
								(s)		
68	units_lstm1=128, units_lstm2=64, dropout_rate=0.0 feature_att=False temporal_att=Falsepochs=50, batch_size=32	,293252.49	9 541.53	374.28	1.19	0.99	1.73	1768.88	0.000037	1.22
69	units_lstm1=128, units_lstm2=64, dropout_rate=0.0 feature_att=True temporal_att=True pochs=50, batch_size=64	,'222351.09 ie,	9 471.54	310.01	0.99	1	1.53	936.92	-0.000036	0.81
70	units_lstm1=128, units_lstm2=64, dropout_rate=0.0 feature_att=True temporal_att=Fal epochs=50, batch_size=64	, ^{265852.18} ;se,	3 515.61	354.86	1.11	0.99	1.6	949.27	-0.000006	0.97
71	units_lstm1=128, units_lstm2=64, dropout_rate=0.0 feature_att=False temporal_att=Truepochs=50, batch_size=64	0.290072.02	2 538.58	395.67	1.24	0.99	1.54	893.97	0.000023	1.13
72	units_lstm1=128, units_lstm2=64, dropout_rate=0.0 feature_att=False temporal_att=Fal epochs=50, batch_size=64	,242020.15	5 491.96	338.95	1.07	1	1.57	911.82	-0.000017	0.91
73	units_lstm1=128, units_lstm2=64, dropout_rate=0.2 feature_att=True temporal_att=True pochs=10, batch_size=32	, 1745438.9	981321.15	1096.37	3.25	0.97	2.57	379.69	0.000756	2.75
74	units_lstm1=128, units_lstm2=64, dropout_rate=0.2 feature_att=True temporal_att=Fal epochs=10, batch_size=32	, 670897.76	819.08	642.98	2.1	0.99	2.2	379.13	0.000044	1.11
75	units_lstm1=128, units_lstm2=64, dropout_rate=0.2 feature_att=False temporal_att=Truepochs=10, batch_size=32	600971.81	775.22	586.44	1.78	0.99	1.91	374.71	0.000052	1.15
76	units_lstm1=128, units_lstm2=64, dropout_rate=0.2 feature_att=False temporal_att=Falsepochs=10, batch_size=32	,799989.59	9 894.42	693.8	2.2	0.98	2.41	361.43	-0.000007	0.99
77	units_lstm1=128, units_lstm2=64, dropout_rate=0.2 feature_att=True temporal_att=True pochs=10, batch_size=64	, 461364.08	3 679.24	499.44	1.59	0.99	2	192.12	-0.000136	0.7
78	units_lstm1=128, units_lstm2=64, dropout_rate=0.2 feature_att=True temporal_att=Fal epochs=10, batch_size=64	, 525544.04	4 724.94	524.68	1.68	0.99	2.37	195.78	-0.000115	0.76
79	units_lstm1=128, units_lstm2=64, dropout_rate=0.2 feature_att=False temporal_att=Truepochs=10, batch_size=64	,615099.24	4 784.28	587.6	1.83	0.99	2.04	183.81	0.000025	1.06

Table 5.5: Results of the ablation study of the BiLSTM architecture.

							e BiLSTM	Training		
		MOD	DMCE	24.5	MADE	D2		O	Overfit	$_{\mathrm{Loss}}$
ID	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)		
	units_lstm1=128, units_lstm2=64,									
80	dropout_rate=0.2		00706	E60 06	1 00	0.00	2.66	100 05	0.000109	0.7
80	feature_att=False temporal_att=Fal		001.00	569.06	1.82	0.99	2.66	186.85	-0.000192	0.7
	epochs=10, batch_size=64									
	units_lstm1=128,									
01	units_lstm2=64, dropout_rate=0.2	070100 50	0.07.00	701.00	0.20	0.00	2.07	010.0	0.000275	0.00
81	feature_att=True temporal_att=Tru		987.99	761.96	2.36	0.98	2.97	818.9	0.000375	2.29
	epochs=20, batch_size=32									
	units_lstm1=128,									
0.0	units_lstm2=64, dropout_rate=0.2	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	FOF 15	905.00	1.05	0.00	1.01	700 ==	0.000000	0.70
82	feature_att=True temporal_att=Fal		2 565.17	395.98	1.25	0.99	1.81	780.55	-0.000069	0.76
	epochs=20, batch_size=32									
	units_lstm1=128,									
09	units_lstm2=64, dropout_rate=0.2	, 910149 0	FF0.7	200 55	1.00	0.00	1 75	776 70	0.000020	0.00
83	feature_att=False temporal_att=Tru		558.7	390.55	1.22	0.99	1.75	776.72	-0.000030	0.88
	epochs=20, batch_size=32									
	units_lstm1=128, units_lstm2=64,									
0.4	dropout rate=0.2	, 420012.00	CF7 90	450.20	1 40	0.00	0.15	700 5	0.000017	0.04
84	feature_att=False temporal_att=Fal	e,432813.90 lse,	001.89	450.38	1.46	0.99	2.15	769.5	-0.000017	0.94
	epochs=20, batch_size=32									
	units_lstm1=128,									
95	units_lstm2=64, dropout_rate=0.2	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	, FCO 2	201 41	1.00	0.00	1.0	416.20	0.000000	0.00
85	feature_att=True temporal_att=Tru		2 508.5	391.41	1.23	0.99	1.8	416.39	-0.000098	0.69
	epochs=20, batch_size=64									
	units_lstm1=128,									
86	units_lstm2=64, dropout_rate=0.2	, , , , , , , , , , , , , , , , , , ,	L 661 60	481.61	1 55	0.99	2.14	417 55	0.000014	0.95
80	feature_att=True temporal_att=Fal	, 457654.74 lse,	1 001.09	461.01	1.55	0.99	2.14	417.55	-0.000014	0.95
	epochs=20, batch_size=64									
	units_lstm1=128, units_lstm2=64,									
87	dropout_rate=0.2 feature_att=False		574.50	398.09	1.26	0.00	1 0/	200 62	0.000079	0.76
01	temporal_att=Tri		. 574.59	396.09	1.20	0.99	1.84	398.63	-0.000072	0.76
	epochs=20, batch_size=64									
	units_lstm1=128, units_lstm2=64,									
00	dropout_rate=0.2	. 401479	701.05	511.1	1.64	0.99	2.26	397.47	-0.000076	0.82
88	feature_att=False temporal_att=Fal		701.05	511.1	1.64	0.99	4.40	551.41	-0.000070	0.02
	epochs=20, batch_size=64									
	units_lstm1=128, units_lstm2=64,									
89	dropout_rate=0.2 feature_att=True	9079071 0	131430 78	1078 28	2.96	0.96	2.95	2018.88	0.001216	7.2
00	temporal_att=Tri		.or 409.10	1010.20	2.30	0.90	2.30	2010.00	0.001210	1.4
	epochs=50, batch_size=32									
	units_lstm1=128, units_lstm2=64,									
90	dropout_rate=0.2 feature att=True	, 234363 1	484.11	333.47	1.05	1	1.53	2004.68	-0.000027	0.86
90	temporal_att=Fal		404.11	000.41	1.00	1	1.00	2004.00	0.000021	0.00
	epochs=50, batch_size=32									
	units_lstm1=128, units_lstm2=64,									
91	dropout_rate=0.2 feature_att=False	243623 <i>4</i> 3	8 493 58	349.94	1.12	1	1.59	1948.21	0.000005	1.03
01	temporal_att=Tri		. 100.00	010.04	1.12	1	1.00	1010.21	5.000000	1.00
	epochs=50, batch_size=32									

 $\textbf{Table 5.5:} \ \ \text{Results of the ablation study of the BiLSTM architecture}.$

	Table	2.3. 1600	G100 O1 0.		ion boardy	01 011		M archited Training		
TD		MOD	DMCE	MAD	MADE	D 2	CUDM	_	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	SE Time	Gap	Ratio
								(s)		
92	units_lstm1=128, units_lstm2=64, dropout_rate=0.2 feature_att=False temporal_att=Fal epochs=50, batch_size=32 units_lstm1=128,	,244835.37	494.81	340.89	1.09	1	1.58	1929.81	-0.000027	0.86
93	units_lstm2=64, dropout_rate=0.2 feature_att=True, temporal_att=True epochs=50, batch_size=64	, 275939.04	525.3	363.81	1.16	0.99	1.71	1034.81	-0.000019	0.91
94	units_lstm1=128, units_lstm2=64, dropout_rate=0.2 feature_att=True, temporal_att=Fal epochs=50, batch_size=64	256189.98	506.15	343.67	1.1	0.99	1.64	1041.89	-0.000029	0.86
95	units_lstm1=128, units_lstm2=64, dropout_rate=0.2 feature_att=False temporal_att=Truepochs=50, batch_size=64	,231014.92	480.64	330.13	1.05	1	1.56	987.04	-0.000024	0.87
96	units_lstm1=128, units_lstm2=64, dropout_rate=0.2 feature_att=False temporal_att=Fal epochs=50, batch_size=64	,276832.85	5 526.15	370.22	1.18	0.99	1.7	992.06	-0.000038	0.83
97	units_lstm1=256, units_lstm2=128, dropout_rate=0.0 feature_att=True, temporal_att=True epochs=10,		890.57	704.73	2.23	0.98	2.15	489.82	0.000160	1.42
98	batch_size=32 units_lstm1=256, units_lstm2=128, dropout_rate=0.0 feature_att=True, temporal_att=Fal epochs=10, batch_size=32	417720.26	6 646.31	450.97	1.41	0.99	1.94	466.23	-0.000060	0.83
99	units_lstm1=256, units_lstm2=128, dropout_rate=0.0 feature_att=False temporal_att=Tru epochs=10, batch_size=32	,389648.43	624.22	455.59	1.43	0.99	1.82	462.14	-0.000011	0.96
100	units_lstm1=256, units_lstm2=128, dropout_rate=0.0 feature_att=False temporal_att=Fal epochs=10, batch_size=32	,2049194.0	81431.5	1065.99	3.58	0.96	4.6	441.17	-0.000471	0.75
101	units_lstm1=256, units_lstm2=128, dropout_rate=0.0 feature_att=True, temporal_att=True epochs=10, batch_size=64	,478762.72	691.93	496.35	1.59	0.99	2.12	261.39	-0.000071	0.82
102	units_lstm1=256, units_lstm2=128, dropout_rate=0.0 feature_att=True, temporal_att=Fal epochs=10, batch_size=64	695249.75	833.82	653.77	2.1	0.99	2.06	250.38	0.000099	1.26
103	units_lstm1=256, units_lstm2=128, dropout_rate=0.0 feature_att=False temporal_att=Truepochs=10, batch_size=64	,421387.56	649.14	471.35	1.51	0.99	1.95	241.46	-0.000039	0.88

Table 5.5: Results of the ablation study of the BiLSTM architecture.

								Training		_
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSI	Ü	Overfit	Loss
110	Comig.	WISE	ICIVISE	WIZE	WIZE E	10	C V ICIVISI	(s)	Gap	Ratio
104	units_lstm1=256 units_lstm2=128 dropout_rate=0.0 feature_att=Fals temporal_att=Fa epochs=10, batch_size=64	, o, e,740653.72	860.61	625.96	1.95	0.99	2.64	230.22	-0.000066	0.88
105	units_lstm1=256 units_lstm2=128 dropout_rate=0.0 feature_att=Truetemporal_att=Truetemporal_att=Truetopochs=20, batch_size=32	, o, e, 333612.38 rue,	577.59	417.39	1.33	0.99	1.78	1009.85	-0.000045	0.83
106	units_lstm1=256 units_lstm2=128 dropout_rate=0.0 feature_att=Tructemporal_att=Fa epochs=20, batch_size=32	, o, e, 336834.98	580.37	415.36	1.29	0.99	1.73	931.63	-0.000025	0.9
107	units_lstm1=256 units_lstm2=128 dropout_rate=0.0 feature_att=Fals: temporal_att=Tr epochs=20, batch_size=32	, o, e,366972.5	605.78	455.73	1.43	0.99	1.64	921.71	0.000027	1.12
108	units_lstm1=256 units_lstm2=128 dropout_rate=0.0 feature_att=Fals temporal_att=Fa epochs=20, batch_size=32	, o, e,638056.96	798.78	581.92	1.84	0.99	2.53	880.46	-0.000169	0.72
109	units_lstm1=256 units_lstm2=128 dropout_rate=0.0 feature_att=True temporal_att=Tr epochs=20,	, o, e, 386917.22	622.03	454.07	1.49	0.99	1.9	520.7	-0.000031	0.89
110	batch_size=64 units_lstm1=256 units_lstm2=128 dropout_rate=0.6 feature_att=True temporal_att=Fa epochs=20,	, o, e, 574714.12	758.1	597.08	1.92	0.99	1.82	498.1	0.000129	1.49
111	batch_size=64 units_lstm1=256 units_lstm2=128 dropout_rate=0.0 feature_att=Fals temporal_att=Tr epochs=20, batch_size=64	, o, e,322726.56	568.09	404.31	1.27	0.99	1.79	477.29	-0.000007	0.97
112	units_lstm1=256 units_lstm2=128 dropout_rate=0.0 feature_att=Fals temporal_att=Fals epochs=20, batch_size=64	, o, e,4532053.7	32128.86	1573.66	4.85	0.91	6.04	458.23	-0.003883	0.44
113	units_lstm1=256 units_lstm2=128 dropout_rate=0.0 feature_att=True temporal_att=Tre epochs=50, batch_size=32	, e, 284370.51	533.26	372.48	1.21	0.99	1.72	2438.48	-0.000010	0.95
114	units_lstm1=256 units_lstm2=128 dropout_rate=0.0 feature_att=True temporal_att=Fa epochs=50, batch_size=32	; o, e, 215745.18	464.48	311.73	0.99	1	1.48	2327.55	-0.000016	0.9
115	units_lstm1=256 units_lstm2=128 dropout_rate=0.0 feature_att=Fals temporal_att=Tr epochs=50, batch_size=32	;), e,237715.12	487.56	339.75	1.07	1	1.55	2351.72	0.000002	1.01

Table 5.5: Results of the ablation study of the BiLSTM architecture.

	Table	5.5: Kes	uits of t	ne ablati	on study	oi th	e BiLST.	M architec	ture.	
_						_ ^		Training	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	SE Time	Gap	Ratio
								(s)	-	
	units_lstm1=256, units_lstm2=128,	1								
116	dropout_rate=0.0 feature_att=False temporal_att=False		478.6	323.49	1.03	1	1.53	2201.54	-0.000016	0.91
	epochs=50,	ise,								
	batch_size=32 units_lstm1=256,									
117	units_lstm2=128, dropout_rate=0.0), 262202 8	512.15	252 04	1 19	0.00	1.64	1299.58	0.000026	0.87
117	feature_att=True temporal_att=Tru	, 202292.8 ue,	312.13	353.84	1.12	0.99	1.64	1299.38	-0.000026	0.87
	epochs=50, batch_size=64									
	units_lstm1=256, units_lstm2=128,									
118	dropout_rate=0.0 feature_att=True	,256834.19	506.79	349.16	1.11	0.99	1.58	1261.94	-0.000024	0.88
	temporal_att=Fa epochs=50,	lse,								
	batch_size=64 units_lstm1=256,									
110	units_lstm2=128, dropout_rate=0.0),	FF0 9	405.04	1.00	0.00	1.50	1100.00	0.000000	1.10
119	feature_att=False temporal_att=Tr		550.3	405.24	1.26	0.99	1.56	1196.09	0.000032	1.19
	epochs=50, batch_size=64									
	units_lstm1=256, units_lstm2=128,									
120	dropout_rate=0.0 feature_att=False	4,478288.25	691.58	522.51	1.66	0.99	1.87	1177.12	0.000002	1.01
	temporal_att=Falepochs=50,	lse,								
	batch_size=64 units_lstm1=256,									
404	units_lstm2=128, dropout_rate=0.2	2,		4 40 7 04				40= 0=		
121	feature_att=True temporal_att=True		181925.68	1425.31	3.9	0.93	3.65	497.25	0.002085	5.71
	epochs=10, batch_size=32									
	units_lstm1=256, units_lstm2=128,									
122	dropout_rate=0.2 feature_att=True	,602692.68	776.33	576.69	1.76	0.99	2.1	479.78	-0.000008	0.98
	temporal_att=Fa epochs=10,	lse,								
	batch_size=32 units_lstm1=256,									
100	units_lstm2=128, dropout_rate=0.2	2,607409.44	770 41	5 00.60	1 77	0.00	1.00	475 11	0.000000	1.05
123	feature_att=False temporal_att=Tr		779.41	588.69	1.77	0.99	1.92	475.11	0.000082	1.25
	epochs=10, batch_size=32									
	units_lstm1=256, units_lstm2=128,									
124	dropout_rate=0.2 feature_att=False	9,8832993.6	62972.04	2304.06	6.88	0.83	8.18	435.62	-0.017074	0.26
	temporal_att=Fa. epochs=10,	lse,								
	batch_size=32 units_lstm1=256,									
105	units_lstm2=128, dropout_rate=0.2	2,	994.69	E00 41	1 00	0.00	0.45	257 40	0.000000	0.09
125	feature_att=True temporal_att=True		824.08	598.41	1.82	0.99	2.45	257.49	-0.000008	0.98
	epochs=10, batch_size=64									
	units_lstm1=256, units_lstm2=128,									
126	dropout_rate=0.2 feature_att=True	, 514347	717.18	513.56	1.62	0.99	2.21	246.42	-0.000083	0.81
	temporal_att=Falepochs=10,	ise,								
	batch_size=64 units_lstm1=256,									
107	units_lstm2=128, dropout_rate=0.2	2,	715 70	F16.04	1.00	0.00	0.0	007 40	0.000040	0.0
127	feature_att=False temporal_att=Tr		(15.72	516.84	1.66	0.99	2.3	237.49	-0.000040	0.9
	epochs=10, batch_size=64									

Table 5.5: Results of the ablation study of the BiLSTM architecture.

								Training		
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSI	Ü	Overfit	Loss
ш	Comig.	MSE	IUVISE	WIAL	WAIL	11	CVICIOISI		Gap	Ratio
								(s)		
	units_lstm1=256 units_lstm2=128 dropout_rate=0.	,								
128	dropout_rate=0. feature_att=Fals temporal_att=Fa	e,762581.39	9 873.26	639.09	2.08	0.99	2.78	225.5	-0.000205	0.72
	epochs=10,	use,								
	batch_size=64 units_lstm1=256	,								
	units_lstm2=128 dropout_rate=0.	,								
129	feature_att=Tru temporal att=Tr	e, 1720488.3	371311.67	971.17	2.69	0.97	2.72	950.54	0.000883	4.05
	epochs= $\overline{20}$,	. uc,								
	batch_size=32 units_lstm1=256									
	units_lstm2=128 dropout_rate=0.	2.								
130	feature_att=Tru temporal_att=Fa	e, 304694.05	5 551.99	383.38	1.23	0.99	1.79	907.39	-0.000077	0.73
	epochs=20, batch_size=32	,								
	units_lstm1=256									
101	units_lstm2=128 dropout_rate=0.	2	2 51/10	0FF 00	4 4 4	0.00	1.00	004.00	0.000000	0.55
131	feature_att=Fals temporal_att=T	e,264327.96 rue,	5 514.13	357.32	1.14	0.99	1.62	924.88	-0.000060	0.75
	epochs=20, batch_size=32									
	units_lstm1=256 units_lstm2=128									
199	dropout_rate=0. feature_att=Fals	2 , 2 , 2	5 079 91	700.88	9 99	0.00	2.08	875.95	0.000170	1 59
132	temporal_att=Fa		0 010.01	709.88	2.32	0.99	2.08	875.95	0.000179	1.53
	epochs=20, batch_size=32									
	units_lstm1=256 units_lstm2=128									
133	dropout_rate=0. feature_att=Tru	2,	531160.88	839.24	2.4	0.97	2.98	514.65	0.000599	2.88
100	temporal_att=T		001100.00	000.24	2.4	0.51	2.50	014.00	0.000000	2.00
	epochs= 20 , batch_size= 64									
	units_lstm1=256 units_lstm2=128									
134	dropout_rate=0. feature_att=Tru	$_{\rm e}^{2,}501128.56$	5 707.9	492.63	1.55	0.99	2.14	499.03	-0.000115	0.75
	temporal_att=Fa	ılse,								
	$batch_size=64$									
	units_lstm1=256 units_lstm2=128	,								
135	dropout_rate=0. feature_att=Fals	e,404774.37	7 636.22	430.95	1.36	0.99	1.97	481.34	-0.000019	0.94
	temporal_att=Trepochs=20,	rue,								
	batch_size=64 units_lstm1=256	i.								
	units_lstm2=128 dropout_rate=0.	,								
136	feature_att=Fals temporal_att=Fa	e,642272.07	7 801.42	613.5	1.93	0.99	2.15	456.42	-0.000029	0.94
	epochs=20,	use,								
	batch_size=64 units_lstm1=256									
	units_lstm2=128 dropout_rate=0. feature_att=Tru	,								
137	feature_att=Tru temporal_att=Tru	e, 285326.07 rue.	7 534.16	381.49	1.21	0.99	1.58	2429.79	-0.000005	0.97
	epochs=50, batch_size=32	•								
	units_lstm1=256									
190	units_lstm2=128 dropout_rate=0.	2,	7 450 50	202.25	0.00	1	1 45	9999 50	0.0000.40	0 77
138	feature_att=Tru temporal_att=Fa		450.79	303.37	0.96	1	1.45	2322.59	-0.000042	0.77
	epochs=50, batch_size=32									
	units_lstm1=256 units_lstm2=128									
139	dropout_rate=0. feature_att=Fals	2	1 476 59	323.57	1.03	1	1.53	2306.54	-0.000010	0.94
100	temporal_att=Tr		1 410.02	020.01	1.00	1	1.00	2000.04	-0.000010	0.34
	epochs=50, batch_size=32									

APPENDIX 149

Table 5.5: Results of the ablation study of the BiLSTM architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	R ²	CVRMS	Training E Time (s)	Overfit Gap	Loss Ratio
140	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=False epochs=50, batch_size=32	255076.78 e,	505.05	347.23	1.11	1	1.64	2193.36	-0.000028	0.86
141	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=True epochs=50, batch_size=64		81083.87	745.49	2.04	0.98	2.31	1291.21	0.000600	4
142	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=True, temporal_att=Fals epochs=50, batch_size=64		630.15	458.34	1.38	0.99	1.72	1249	0.000045	1.2
143	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=Tru epochs=50, batch_size=64		513.13	369.57	1.19	0.99	1.63	1469.69	0.000004	1.02
144	units_lstm1=256, units_lstm2=128, dropout_rate=0.2, feature_att=False, temporal_att=Fals epochs=50, batch_size=64	373514.86	6 611.16	434.6	1.43	0.99	2	1228.77	-0.000060	0.81

 ${\bf Table~5.6:}~{\bf Results~of~the~ablation~study~of~the~Transformers~architecture.}$

ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Training Time (s)	Overfit Gap	Loss Ratio
1	architecture=enco d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0 epochs=20, batch_size=64	414109.1	643.51	458.46	1.46	0.99	2.07	55.4	-0.000057	0.83
2	architecture=encod_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0 epochs=20, batch_size=128	503719.28	709.73	489.33	1.58	0.99	2.29	31.8	-0.000070	0.83
3	architecture=enco d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0 epochs=50, batch_size=64	349932.77	591.55	429.84	1.39	0.99	1.76	123.2	0.000003	1.01
4	architecture=enco d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0 epochs=50, batch_size=128	354360.79	595.28	420.04	1.35	0.99	1.91	67.19	-0.000031	0.89

Table 5.6: Results of the ablation study of the Transformers architecture.

								Training		
ID	G 6	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	_	Overfit	$_{\mathrm{Loss}}$
ш	Config.	MSE	RWISE	WAL	MAPL	n-	CVRIVIS	e ime	Gap	Ratio
								(s)		
	architecture=enc	oder,								
	$d_{model=32}$, $num_{heads=2}$,									
5	ff_dim=64, num_layers=1,	3441745.	581855.19	1460.6	4.08	0.93	3.18	59.36	0.001866	4.9
	dropout_rate=0.	2,								
	epochs=20, batch_size=64									
	architecture=enc d_model=32,	oder,								
	$num_heads=2$,									
6	ff_dim=64, num_layers=1,	6729769.	$3\ 2594.18$	2349.69	7	0.87	2.52	34.59	0.004005	7.89
	dropout_rate=0.	2,								
	epochs=20, batch_size=128									
	architecture=enc d_model=32,	oder,								
	num_heads=2,									
7	ff_dim=64, num_layers=1,	2412466.	971553.21	1363.16	4	0.95	2.29	129.21	0.001345	5.51
	dropout_rate=0. epochs=50,	2,								
	batch_size=64	,								
	architecture=enc d_model=32,	oder,								
0	$num_heads=2,$ $ff_dim=64,$	6084900	000 <i>166 6</i> 9	0191 9	<i>C</i> 1	0.00	2.04	76.4	0.002015	19 59
8	num_layers=1, dropout_rate=0.		992466.62	2131.3	6.1	0.88	2.94	76.4	0.003815	12.53
	epochs= 50 ,	<u>~</u> ,								
	batch_size=128 architecture=enc	oder.								
	d_model=32, num_heads=2,	,								
9	ff_dim=64,	525853.9	9 725.16	526.21	1.68	0.99	2.18	69.4	0.000035	1.11
	num_layers=2, dropout_rate=0.	0,								
	epochs=20, batch_size=64									
	architecture=enc	oder,								
	$d_{model=32}$, $num_{heads=2}$,									
10	ff_dim=64, num_layers=2,	473477.8	9 688.1	496.96	1.58	0.99	2.16	40.07	-0.000106	0.75
	dropout_rate=0. epochs=20,	0,								
	batch_size=128									
	architecture=enc d model=32,	oder,								
	num_heads=2, ff_dim=64,	o=00040		050 50			4.00			0.04
11	num_layers=2,	270964.2	1 520.54	358.53	1.15	0.99	1.69	151.81	-0.000036	0.84
	dropout_rate=0. epochs=50,	0,								
	batch_size=64 architecture=enc	oder								
	$d_{model=32}$	ouer,								
12	$num_heads=2, ff_dim=64,$	396799.8	7 629 92	467.88	1.5	0.99	1.81	84.2	0.000027	1.11
12	num_layers=2, dropout_rate=0.		1 025.52	401.00	1.0	0.55	1.01	04.2	0.000021	1.11
	epochs=50, batch_size=128	-,								
	architecture=enc	oder,								
	d_model=32, num_heads=2,									
13	ff_dim=64, num_layers=2,	4682441.	4 2163.89	1789.52	5.03	0.91	3.34	73.7	0.002792	8
	dropout_rate=0.	2,								
	epochs=20, batch_size=64									
	architecture=enc d_model=32,	oder,								
	num_heads=2,									
14	ff_dim=64, num_layers=2,		862677.99	2391.45	7.01	0.86	2.88	43.85	0.004279	8.03
	dropout_rate=0. epochs=20,	2,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

						-	<u> Transforme</u>	Training	-	
		3.600	D		2425	ъ 2		Ü	Overfit	Loss
$^{\mathrm{ID}}$	Config.	\mathbf{MSE}	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	F	
	architecture=enco	der,								
	d_model=32, num_heads=2,									
15	ff_dim=64, num_layers=2,	3313469.0	081820.29	1532.01	4.36	0.94	2.77	162.99	0.001951	7.37
	dropout_rate=0.2	,								
	epochs=50, batch_size=64									
	architecture=enco d_model=32,	der,								
	$num_heads=2$,									
16	ff_dim=64, num_layers=2,	5742742.0	032396.4	2022.02	5.71	0.89	3.37	93.97	0.003592	12.19
	dropout_rate=0.2 epochs=50,	,								
	$batch_size=128$	1								
	architecture=enco d_model=32,	der,								
17	num_heads=2, ff_dim=128,	458992.7	7 677 40	485.48	1.55	0.99	2.12	59.2	-0.000002	0.99
11	num_layers=1, dropout_rate=0.0		011.43	400.40	1.00	0.55	2.12	03.2	-0.000002	0.33
	epochs=20, batch_size=64	,								
	architecture=enco	der,								
	d_model=32, num_heads=2,									
18	ff_dim=128, num_layers=1,	448793.65	5 669.92	496.47	1.58	0.99	2.03	33.18	-0.000087	0.78
	dropout_rate=0.0 epochs=20,	,								
	$batch_size=128$									
	architecture=enco d_model=32,	der,								
10	num_heads=2, ff_dim=128,	205025 6	0 779 11	277 27	1.0	0.00	1 77	100.05	0.000000	0.0
19	num_layers=1, dropout_rate=0.0	305925.68	5 555.11	377.27	1.2	0.99	1.77	126.65	-0.000023	0.9
	epochs=50,	,								
	batch_size=64 architecture=enco	der,								
	d_model=32, num_heads=2,									
20	ff_dim=128, num_layers=1,	301926.8	5 549.48	382.5	1.22	0.99	1.78	68.83	-0.000051	0.8
	dropout_rate=0.0	,								
	epochs=50, batch_size=128									
	architecture=enco d_model=32,	der,								
	num_heads=2, ff_dim=128,									
21	num_layers=1,		612477.83	2127.52	6.09	0.88	3.06	62.23	0.003701	8.68
	dropout_rate=0.2 epochs=20,	,								
	batch_size=64 architecture=enco	der								
	d_model=32, num_heads=2,	der,								
22	ff_dim=128,	5008082.5	532237.87	1879.64	5.34	0.9	3.14	34.33	0.002937	7.18
	num_layers=1, dropout_rate=0.2			,,,,,	0.0-		0.22	0 -100		
	epochs=20, batch_size=128									
	architecture=enco	der,								
	d_model=32, num_heads=2,									
23	ff_dim=128, num_layers=1,	2859435.4	471690.99	1341.16	3.75	0.94	3.01	136.03	0.001663	6.82
	dropout_rate=0.2 epochs=50,	,								
	batch_size=64	J								
	architecture=enco d_model=32,	uer,								
24	num_heads=2, ff_dim=128,	4825924.0	049106 ₽	1815.41	5.1	0.91	3.31	75.59	0.002956	9.9
24	num_layers=1, dropout rate=0.2		U-1 2 1 3 U . U	1010.41	0.1	0.91	0.01	10.03	0.002300	<i>3.3</i>
	epochs=50, batch_size=128	,								
	Jaccii_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

Config. MSE RMSE MAE MAPE R2 CVRMSE Time Gap Ratio			2000 084					Iransform	Training		
architecture-encoder, d_model=32, num_beeds=22, num_byees=2, dropout_rate=0.0, epoches=20, architecture-encoder, d_model=32, num_byees=2, dropout_rate=0.0, epoches=0, d_model=32, num_byees=2, d_model=32, num_beads=2, d_model=32, num_byees=2, d_model=32, d_model	TD		MOD	DMCE	MAE	MADE	D 2		_	Overfit	Loss
Architecture=encoder,	Ю	Config.	MSE	RMSE	MAE	MAPE	R²	CVRMSI	E Time	Gap	Ratio
d_model=32, num_layers=2, 312672.47 559.17 395.32 1.26 0.99 1.77 69.01 -0.000080 0.73 num_layers=2, 312672.47 559.17 395.32 1.26 0.99 1.77 69.01 -0.000080 0.73 num_layers=2, dropout_rate=0.0, d_model=32, num_layers=2, 479944.64 692.78 523.28 1.71 0.99 2.07 39.64 -0.000059 0.85 num_layers=2, 479944.64 692.78 523.28 1.71 0.99 2.07 39.64 -0.000059 0.85 num_layers=2, 479944.64 692.78 523.28 1.71 0.99 1.67 159.97 -0.00018 0.91 num_layers=2, 47994.64 692.78 523.28 1.71 0.99 1.67 159.97 -0.00018 0.91 num_layers=2, 269971.33 519.59 372.35 1.23 0.99 1.67 159.97 -0.00018 0.91 num_layers=2, 269971.33 519.59 372.35 1.23 0.99 1.67 159.97 -0.00018 0.91 num_layers=2, 36673.65 621.91 466.04 1.54 0.99 1.88 84.9 0.00021 1.09 num_layers=2, 386773.65 621.91 466.04 1.54 0.99 1.88 84.9 0.00021 1.09 num_layers=2, 370.00000000000000000000000000000000000									(s)	- 1	
			oder,								
dropout_rate=0.0, epoches=0.0	25		312672.47	559.17	395.32	1.26	0.99	1.77	69.01	-0.000080	0.73
hatch_size=64 architecturesenceder, d_model=32, deponder d_model=32, d_mo		$dropout_rate=0.$	0,								
d_model=32, num_heads=2, and heads=2, and heads=2, num_heads=2, and heads=2, and heads=3, and heads=4, and heads=3, and heads=4, and heads=3, and heads=4, and heads=4, and heads=4, and heads=3, and heads=4, and heads=3, and heads=4, and heads=3, and h											
num_heads=2,		architecture=ence	oder,								
dropout_rate=0.0, epochs=20, latch_size=128 architecturesencoder, aum_heads=2, fl.meresenceder, aum_heads=2, fl.meresenceder, aum_heads=2, fl.meresenceder, d_model=32, num_layers=2, 386773.65 621.91 466.04 1.54 0.99 1.88 84.9 0.000021 1.09 dropout_rate=0.0, batch_size=64 architecturesencoder, d_model=32, num_layers=2, 386773.65 621.91 466.04 1.54 0.99 1.88 84.9 0.000021 1.09 dropout_rate=0.0, epochs=50, batch_size=128 architecturesencoder, tunn_heads=2, fl.meresenceder, d_model=32, num_layers=2, \$179565.15/275.87 1890.63 5.33 0.9 3.15 73.1 0.003119 8.61 dropout_rate=0.2, epochs=50, aum_layers=2, \$854808.852975.7 2659.63 7.8 0.83 2.84 43.82 0.005463 10.58 dropout_rate=0.2, epochs=20, epochs=30, epochs=64 architecturesencoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=64 architecturesencoder, d_model=32, num_layers=2, architecturesencoder, d_model=32, num_layers=2, architecturesencoder, d_model=32, num_layers=3, architecturesencoder, d_model=32, num_layers=1, dropout_rate=0.2, epochs=60, epochs=60, epochs=50, epochs=60, ep	26		479944.64	692.78	523.28	1.71	0.99	2.07	39.64	-0.000059	0.85
batch_size=128 architecturesencoder, d_model=32, num_lewise=2, depopul_rate=0.0, epochs=50, help the transported to the transpo		$dropout_rate=0.$	0,								
d_model=32, num_heads=2, 269971.33 519.59 372.35 1.23 0.99 1.67 159.97 -0.000018 0.91 nlm_heads=2, 269971.33 519.59 372.35 1.23 0.99 1.67 159.97 -0.000018 0.91 nlm_heads=2, num_heads=2, num_heads=3, num_heads=2, num_heads=4, n											
num_layers=2 269971.33 519.59 372.35 1.23 0.99 1.67 159.97 -0.000018 0.91 dim=128, num_layers=2, dropout_rate=0.0, batch_size=04 architecture=encoder, d_model=32, num_leads=2, num_leads=3, num_leads=3, num_leads=3, num_layers=2, 8854808.852975.7 2659.63 7.8 0.83 2.84 43.82 0.005463 10.58 dropout_rate=0.2, epochs=20, batch_size=128 architecture=encoder, d_model=32, num_layers=2, for num_layers=2, num_leads=3, num_layers=2, num_leads=3, num_layers=2, s314382.6 2305.29 1871.68 5.21 0.9 3.35 165.29 0.003336 12.68 dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, num_layers=2, num_leads=3, num_layers=2, s314382.6 3305.29 1871.68 5.21 0.9 3.35 165.29 0.003336 12.68 dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, num_layers=2, num_heads=4, num_layers=2, s319818.561794.39 1439.51 4.02 0.94 3.07 96.56 0.001908 7.65 dropout_rate=0.2, epochs=50, num_layers=2, num_heads=4, num_layers=2, num_heads=3, num_heads=3, num_heads=4, num_layers=2, num_heads=3, num_heads=3, num_heads=3, nu			oder,								
num_layers=2, dropout_rate=0.0, epochs=50, epochs=50, logout_rate=0.0, epochs=50, epochs=50, epochs=50, logout_rate=0.0, epochs=50, logout_rate=0.0, epochs=50, logout_rate=0.0, epochs=50, logout_rate=0.0, epochs=20, logout_rate=0.0, epochs=20, epochs=20, logout_rate=0.0, epochs=30, logout_rate=0.0, epochs=3		num_heads=2,									
dropout_rate=0.0, epochs=50, batch_size=64 architecture=encoder, d_model=32, mum_layers=2, dropout_rate=0.0, epochs=50, batch_size=128 architecture=encoder, dropout_rate=0.0, epochs=50, batch_size=128 architecture=encoder, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, dropout_rate=0.2, epochs=20, batch_size=128 architecture=encoder, dropout_rate=0.2, epochs=20, batch_size=128 architecture=encoder, dropout_rate=0.2, epochs=20, lam_layers=2, architecture=encoder, d_model=32, mum_layers=2, s14382.6 2305.29 1871.68 5.21 0.9 3.35 165.29 0.003336 12.68 architecture=encoder, d_model=32, mum_layers=2, s14382.6 2305.29 1871.68 5.21 0.9 3.35 165.29 0.003336 12.68 architecture=encoder, d_model=32, mum_layers=2, s14382.6 2305.29 1871.68 5.21 0.9 3.35 165.29 0.003336 12.68 architecture=encoder, d_model=32, mum_layers=2, s14382.6 2305.29 1439.51 4.02 0.94 3.07 96.56 0.001908 7.65 dropout_rate=0.2, epochs=50, size=64 architecture=encoder, d_model=32, mum_layers=1, s14382.6	27		269971.33	519.59	372.35	1.23	0.99	1.67	159.97	-0.000018	0.91
batch_size=64 architecture=encoder, d_model=32, num_heads=2, fl_dim=128, num_layers=2, d_encodel=32, num_heads=2, fl_dim=128, num_heads=2, fl_dim=128, num_heads=2, fl_dim=128, num_layers=2, num_heads=2, fl_dim=128, num_layers=2, num_heads=2, fl_dim=128, num_layers=2, num_heads=2, fl_dim=128, num_layers=2, num_heads=3, num_layers=2, num_heads=2, fl_dim=128, num_layers=2, num_heads=3, num_layers=3, num_layers=3, num_layers=3, num_layers=4, num_layers=6, num_lay		$dropout_rate=0.$	0,								
d_model=32, num_heads=2, ff_dim=128, s86773.65 621.91 466.04 1.54 0.99 1.88 84.9 0.000021 1.09 dropout_rate=0.0, batch_size=128 architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, s854808.852975.7 2659.63 7.8 0.83 2.84 43.82 0.005463 10.58 num_layers=2, dropout_rate=0.2, epochs=30, batch_size=128 architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, s854808.852975.7 2659.63 7.8 0.83 2.84 43.82 0.005463 10.58 num_layers=2, dropout_rate=0.2, epochs=30, batch_size=128 architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, s1 ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=1, 389512.46 624.11 437.53 1.4 0.99 1.98 60.99 -0.000060 0.81 dropout_rate=0.0, batch_size=64 architecture=encoder, d_model=32, num_layers=1, 389512.46 624.11 437.53 1.4 0.99 1.98 60.99 -0.000060 0.81 dropout_rate=0.0, batch_size=64 architecture=encoder, d_model=32, num_layers=1, 389512.46 624.11 437.53 1.4 0.99 1.98 60.99 -0.000060 0.81 dropout_rate=0.0, batch_size=64 architecture=encoder, d_model=32, num_leads=4, num_layers=2, dropout_rate=0.0, batch_size=64 architecture=encoder, d_model=32, num_heads=4, num_layers=1, dropout_rate=0.0, batch_size=64 architecture=encoder, d_model=32, num_heads=4, num_layers=1, dropout_rate=0.0, batch_size=64 architecture=encoder, d_model=32, num_heads=4, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, d_		batch_size=64									
### 1.54 1.54 1.54 1.55 1.88 1.88 1.89 1.09 ### 1.55 1.09 ### 1.56 1.55 1.55 1.09 ### 1.56 1.55 1.55 1.09 ### 1.56 1.55 1.55 1.09 ### 1.55 1.55 1.09 ### 1.55 1.09 ### 1.55 1.09 ### 1.55 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.00021 1.00021 1.09 ### 1.55 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.00021 1.			oder,								
Num_layers=2, Num_layers=1, Num_layers=2	20				10001			4.00	0.4.0		4.00
epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=2, fl_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128 architecture=encoder, d_model=32, num_heads=2, fl_dim=128, num_layers=2, sarchitecture=encoder, d_model=32, num_heads=2, fl_dim=128, num_layers=2, sarchitecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=18 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=14 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=14 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=14 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20,	28	num_layers=2,		621.91	466.04	1.54	0.99	1.88	84.9	0.000021	1.09
architecture=encoder, d_model=32, num_heads=2, flum=128,		epochs $=50$,	0,								
d_model=32, num_heads=2, ff_dim=128,			odor								
### 18		$d_{model=32}$	oder,								
num_layers=2, dropout_rate=0.2, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=50 batch_size=64 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20, epoch	20	ff_dim=128,	5170565 1	50075 87	1800 63	5 22	0.0	2.15	79.1	0.003110	8 61
epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=2, 8854808.852975.7 2659.63 7.8 0.83 2.84 43.82 0.005463 10.58 dropout_rate=0.2, epochs=20, batch_size=128 architecture=encoder, d_model=32, num_heads=2, fl_dim=128, objective batch_size=64 architecture=encoder, d_model=32, num_heads=2, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_heads=2, fl_dim=128, num_heads=2, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_heads=2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=4, specified batch_size=64 architecture=encoder, d_model=32, num_heads=4, specified batch_size=64, specified batch_size	23			.02210.01	1030.00	0.00	0.3	5.15	10.1	0.003113	0.01
architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=20, batch_size=128 architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, pochs=20,		epochs=20,	-,								
num_heads=2, fidim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=2, fidim=128, num_heads=2, fidim=128, num_heads=2, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_heads=2, fidim=128, num_heads=32, num_heads=4, fidim=64, num_heads=4, fidim=64, num_heads=4, fidim=64, architecture=encoder, d_model=32, num_heads=4, fidim=64, num_heads=4, fidim=64, num_heads=4, fidim=64, architecture=encoder, d_model=32, num_heads=4, fidim=64, num_heads=4, fidim=64, num_heads=4, fidim=64, architecture=encoder, d_model=32, num_heads=4, fidim=64, num_head			oder,								
30											
dropout_rate=0.2, epochs=20, batch_size=128 architecture=encoder, d_model=32, num_heads=2, ff_dim=128, size=64 architecture=encoder, d_model=32, num_heads=2, ff_dim=128, size=64 architecture=encoder, d_model=32, num_heads=2, ff_dim=128, size=64 architecture=encoder, d_model=32, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, dropou	30	ff_dim=128,	8854808.8	52975.7	2659.63	7.8	0.83	2.84	43.82	0.005463	10.58
batch_size=128 architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=2, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20,		dropout_rate=0.	2,								
architecture=encoder, d_model=32, num_heads=2, 31											
num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, architecture=encoder, d_model=32, num_heads=4, ff_dim=64, architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, epochs=20,		architecture=ence	oder,								
num_layers=2,		num_heads=2,									
dropout_rate=0.2, epochs=50, batch_size=64 architecture=encoder, d_model=32, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, from layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, from layers=1, dropout_rate=0.0, epochs=20, epochs=20, dropout_rate=0.0, epochs=20, epoc	31		5314382.6	2305.29	1871.68	5.21	0.9	3.35	165.29	0.003336	12.68
batch_size=64 architecture=encoder, d_model=32, num_heads=2, 32		dropout_rate=0.	2,								
d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_layers=1, dropout_rate=0.0, epochs=20,		batch_size=64									
32 ff_dim=128, 3219818.561794.39 1439.51 4.02 0.94 3.07 96.56 0.001908 7.65 dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20,			oder,								
num_layers=2, 3213616.501734.53 1435.51 4.02 0.34 3.07 90.50 0.001308 7.03 dropout_rate=0.2, epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, 389512.46 624.11 437.53 1.4 0.99 1.98 60.99 -0.000060 0.81 dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, 434194.56 658.93 467.79 1.49 0.99 2.11 35.51 -0.000117 0.72 dropout_rate=0.0, epochs=20,		num_heads=2,									
epochs=50, batch_size=128 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20 batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20,	32	num_layers=2,		61794.39	1439.51	4.02	0.94	3.07	96.56	0.001908	7.65
architecture=encoder, d_model=32, num_beads=4, 33		epochs= 50 ,	∠,								
d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81 1.4 0.99 1.98 60.99 -0.000060 0.81			nder								
33		$d_{model=32}$	J								
num_layers=1, dropout_rate=0.0, epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20,	33	ff_dim=64,	389512 46	624 11	437 53	1.4	0 99	1 98	60.99	-0.000060	0.81
epochs=20, batch_size=64 architecture=encoder, d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20,	55			, 024.11	101.00	1.4	0.00	1.00	00.00	0.000000	0.01
architecture=encoder, d_model=32, num_heads=4, 34 ff_dim=64, num_layers=1, dropout_rate=0.0, epochs=20,		epochs=20,	,								
num_heads=4, 34		architecture=ence	oder,								
34 $\underset{\substack{\text{fl} = \text{dim} = 64, \\ \text{num} = \text{layers} = 1, \\ \text{dropout} = \text{rate} = 0.0, \\ \text{epochs} = 20,}$ 434194.56 658.93 467.79 1.49 0.99 2.11 35.51 -0.000117 0.72 $\underset{\substack{\text{closed} = 0.0, \\ \text{epochs} = 20,}}{\text{flate}}$											
$ \frac{\text{dropout_rate=0.0}}{\text{epochs=20}} $	34	$ff_dim=64$,	434194.56	658.93	467.79	1.49	0.99	2.11	35.51	-0.000117	0.72
$epochs=20$, $batch_size=128$		dropout_rate=0.	0,								
		epochs=20, batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

						-		Training		
ID	G 6	MSE	DMCE	MAE	MAPE	\mathbb{R}^2	CVDM	_	Overfit	Loss
ID	Config.	MSE	RMSE	MAL	WAPE	n-	CVRMS		Gap	Ratio
								(s)		
	architecture=end d model=32,	coder,								
o=	num_heads=4, ff_dim=64,	2025.12.0		404.00				404.0	0.000000	
35	num_layers=1,	393542.9	627.33	464.83	1.47	0.99	1.72	131.2	0.000038	1.17
	dropout_rate=0 epochs=50,	.0,								
	batch_size=64 architecture=end	roder								
	$d_{model=32}$	Loder,								
36	num_heads=4, ff_dim=64,	353643.23	1 594 68	407.95	1.3	0.99	1.89	71.54	-0.000012	0.95
90	num_layers=1, dropout_rate=0		001.00	101.00	1.0	0.00	1.00	11.01	0.000012	0.00
	epochs=50, batch_size=128	*								
	architecture=end	coder,								
	$d_{model=32}$, $num_{heads=4}$,									
37	ff_dim=64, num_layers=1,	5399175.0	012323.61	1963.43	5.58	0.89	3.17	62.56	0.003225	8.11
	dropout_rate=0 epochs=20,	.2,								
	$batch_size=64$									
	architecture=end d_model=32,	coder,								
38	num_heads=4, ff_dim=64,	4500004	770100 22	1016 10	F 99	0.01	2.01	26 52	0.000527	F 69
30	num_layers=1, dropout_rate=0	4529804.7	1 12120.33	1816.12	5.23	0.91	2.91	36.53	0.002537	5.62
	epochs=20,	.2,								
	batch_size=128 architecture=end	coder,								
	$d_{model=32}$, $num_{heads=4}$,									
39	ff_dim=64, num_layers=1,	2728943.0	081651.95	1349.96	3.83	0.95	2.65	136.12	0.001536	5.75
	dropout_rate=0 epochs=50,	.2,								
	$batch_size=64$									
	architecture=end d_model=32,	coder,								
40	num_heads=4, ff_dim=64,	4649540	05015 <i>4.66</i>	1774 91	4.00	0.01	2.20	76 17	0.000015	0.00
40	num_layers=1, dropout_rate=0	4642549.9	952154.00	1774.31	4.98	0.91	3.29	76.17	0.002815	9.09
	epochs=50, batch_size=128	,								
	architecture=end	coder,								
	$d_{model=32}$, $num_{heads=4}$,									
41	ff_dim=64, num_layers=2,	420185.0	1 648.22	476.3	1.49	0.99	1.85	73.41	-0.000031	0.9
	dropout_rate=0 epochs=20,	.0,								
	batch_size=64									
	architecture=end d_model=32,	coder,								
42	num_heads=4, ff_dim=64,	45000e 0	670 59	407 21	1.62	0.99	2.05	42.43	0.000044	0.87
42	num_layers=2, dropout_rate=0	452296.8	672.53	497.31	1.02	0.99	2.05	42.43	-0.000044	0.87
	epochs=20, batch_size=128	,								
	architecture=end	coder,								
	$d_{model=32}$, $num_{heads=4}$,									
43	ff_dim=64, num_layers=2,	262604.53	3 512.45	355.05	1.12	0.99	1.6	156.75	-0.000036	0.83
	dropout_rate=0 epochs=50,	.0,								
	$batch_size=64$									
	architecture=end d_model=32,	coder,								
44	num_heads=4, ff_dim=64,	298825.93	2 5/6 65	384.99	1.22	0.99	1.76	86	0.000041	0.82
44	num_layers=2, dropout_rate=0		040.00	504.99	1.22	0.99	1.70	00	-0.000045	0.64
	epochs=50, batch_size=128	,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 9	. 0. 10.501	05 01 0110	ablation	i study O	1 0110	1ransforme		ccourc.	
								Training	Overfit	Loss
ID	Config.	\mathbf{MSE}	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Gap	Itatio
	architecture=ence	oder,								
	d_model=32, num_heads=4,									
45	ff_dim=64,	9133469.	933022.16	2621.68	7.49	0.82	3.38	73.15	0.005825	15.62
	num_layers=2, dropout_rate=0.					0.0_	0.00		0.0000	
	epochs=20, batch_size=64									
	architecture=ence	oder,								
	d_model=32, num_heads=4,									
46	$ff_dim=64$,	8917477.0	692986.21	2659.86	7.75	0.83	3	44.47	0.005546	11.45
	num_layers=2, dropout_rate=0.	2,								
	epochs=20, batch_size=128									
	architecture=ence	oder,								
	d_model=32, num_heads=4,									
47	ff_dim=64, num_layers=2,	3857032.9	9 1963.93	1600.66	4.49	0.92	2.86	173.92	0.002355	9.62
	dropout_rate=0. epochs=50,	2,								
	batch_size=64									
	architecture=ence d_model=32,	oder,								
	$num_heads=4,$									
48	ff_dim=64, num_layers=2,	4074733.5	282018.6	1644.62	4.61	0.92	3.04	97.11	0.002438	8.2
	dropout_rate=0. epochs=50,	2,								
	batch_size=128									
	architecture=enced_model=32,	oder,								
49	num_heads=4, ff_dim=128,	436749.5	660.87	493.99	1.52	0.99	1.85	59.32	-0.000015	0.95
43	num_layers=1, dropout_rate=0.		000.01	430.33	1.02	0.33	1.00	03.02	-0.000013	0.30
	epochs=20,	~,								
	batch_size=64 architecture=ence	oder,								
	d_model=32, num_heads=4,									
50	ff_dim=128, num_layers=1,	422404.9	4 649.93	488.13	1.61	0.99	2.03	32.88	-0.000047	0.86
	$dropout_rate=0.$	0,								
	epochs=20, batch_size=128									
	architecture=ence	oder,								
	d_model=32, num_heads=4,									
51	ff_dim=128, num_layers=1,	271292.1	520.86	358.92	1.15	0.99	1.67	128.73	-0.000033	0.85
	dropout_rate=0. epochs=50,	0,								
	batch_size=64									
	architecture=ence d_model=32,	oder,								
50	num_heads=4, ff_dim=128,	070051 7	7 500 01	900 11	1.10	0.00	1.77	70.1	0.000045	0.01
52	$num_layers=1$,	279851.7	7 529.01	366.11	1.16	0.99	1.7	70.1	-0.000045	0.81
	dropout_rate=0. epochs=50,	ο,								
	batch_size=128 architecture=ence	oder.								
	d_model=32, num_heads=4,	, ,								
53	ff_dim=128,	4417044.	732101.68	1825.97	5.27	0.91	2.66	59.36	0.002600	7.35
	num_layers=1, dropout_rate=0.			1020.0.	0.2.	0.01	2.00	00.00	0.002000	
	epochs=20, batch_size=64									
	architecture=ence	oder,								
	$d_{model=32}$, $num_{heads=4}$,									
54	ff_dim=128, num_layers=1,	6016407.	272452.84	2114.87	6.08	0.88	2.92	34.56	0.003600	8.19
	dropout_rate=0.	2,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 5.	6: Resul	ts of the	ablation	study o	t the	Transforme	rs archit	tecture.	
							7	Training	Overfit	Loss
ID	Config.	MSE	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Overni	
								(s)	Gap	Ratio
								(-)		
	architecture=enco d_model=32,	oder,								
	num_heads=4, ff_dim=128,	C057012	000501 44	0100.04	F 00	0.00	9.99	120 50	0.002020	15 10
55	$num_layers=1$,		262501.44	2102.94	5.92	0.88	3.33	132.52	0.003982	15.16
	dropout_rate=0.2 epochs=50,	2,								
	batch_size=64 architecture=enco	dor								
	$d_{model=32}$	der,								
56	num_heads=4, ff_dim=128,	3806633	261973.99	1661.53	4.72	0.92	3.32	74.42	0.002351	8.71
30	num_layers=1, dropout_rate=0.2		201913.99	1001.55	4.12	0.92	3.32	14.42	0.002331	0.71
	epochs=50,	-,								
	batch_size=128 architecture=enco	oder.								
	$d_{model=32}$, , ,								
57	num_heads=4, ff_dim=128,	403447.4	5 635 18	473.24	1.51	0.99	1.87	70.81	-0.000019	0.93
01	num_layers=2, dropout_rate=0.0		0 000.10	110.21	1.01	0.55	1.01	10.01	-0.000013	0.50
	epochs=20,	· ,								
	batch_size=64 architecture=enco	der,								
	d_model=32, num_heads=4,	,								
58	ff_dim=128,	463722.4	5 680.97	494.25	1.57	0.99	2.07	41.07	-0.000062	0.84
	num_layers=2, dropout_rate=0.0					0.00	,		0.00000	0.0-
	epochs=20, batch_size=128									
	architecture=enco	der,								
	$d_{model=32}$, $num_{heads=4}$,									
59	ff_dim=128,	253679.6	3 503.67	347.78	1.11	1	1.58	164.42	-0.000021	0.89
	num_layers=2, dropout_rate=0.0),								
	epochs=50, batch_size=64									
	architecture=enco	der,								
	$d_{model=32}$, $num_{heads=4}$,									
60	ff_dim=128, num_layers=2,	305792.8	8 552.99	395.69	1.26	0.99	1.73	87.53	-0.000009	0.96
	dropout_rate=0.0),								
	epochs=50, batch_size=128									
	architecture=enco d_model=32,	der,								
	$num_heads=4,$									
61	ff_dim=128, num_layers=2,	4020714.	172005.17	1644.58	4.64	0.92	3.64	74.93	0.002363	7.28
	dropout_rate=0.2 epochs=20,	2,								
	$batch_size=64$									
	architecture=enco d_model=32,	oder,								
	num_heads=4, ff_dim=128,									
62	$num_layers=2$,	8661933	2943.12	2577.88	7.43	0.83	3.3	44.96	0.005368	11.07
	dropout_rate=0.2 epochs=20,	2,								
	$batch_size=128$,								
	architecture=enco d_model=32,	der,								
CO.	num_heads=4, ff_dim=128,	F100000	0,000,00	1001 50	F 0	0.0	9.49	1.00.07	0.000000	19.10
63	num_layers=2,		062263.23	1861.52	5.2	0.9	3.43	169.97	0.003226	13.18
	dropout_rate=0.2 epochs=50,	·,								
	batch_size=64 architecture=enco	nder								
	$d_{model=32}$,								
64	num_heads=4, ff_dim=128,	4888774	312211.06	1885.92	5.37	0.9	3.04	95.78	0.003029	11
04	num_layers=2, dropout_rate=0.2		J42411.00	1000.32	0.01	0.0	0.04	55.16	5.000023	11
	epochs=50,	,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

							Iransform	Training		
TD		MOD	DMCE	MAE	MADE	D.2	CVDMC	_	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS	E Time	Gap	Ratio
								(s)	- 1	
	architecture=ence	oder,								
	d_model=64, num_heads=2,									
65	ff_dim=64, num_layers=1,	442077.23	664.89	486.7	1.58	0.99	1.98	64.61	-0.000019	0.94
	dropout_rate=0.0	0,								
	epochs=20, batch_size=64									
	architecture=ence	oder,								
	d_model=64, num_heads=2,									
66	ff_dim=64, num_layers=1,	601075.78	775.29	591.38	1.89	0.99	2.16	37.29	-0.000055	0.88
	dropout_rate=0.0	0,								
	epochs=20, batch_size=128									
	architecture=ence	oder,								
	$d_{model=64}$, $num_{heads=2}$,									
67	ff_dim=64, num_layers=1,	421154.77	648.96	504.18	1.61	0.99	1.66	130.69	0.000073	1.34
	dropout_rate=0.	0,								
	epochs=50, batch_size=64									
	architecture=ence d_model=64,	oder,								
	num_heads=2,									
68	ff_dim=64, num_layers=1,	359080.6	599.23	429.74	1.35	0.99	1.75	70.89	-0.000025	0.91
	dropout_rate=0.0 epochs=50,	0,								
	$batch_size=128$									
	architecture=ence d model=64,	oder,								
	num_heads=2,									
69	ff_dim=64, num_layers=1,	3330193.9	21824.88	1622.09	4.81	0.93	2.23	64.86	0.001804	4.88
	dropout_rate=0.2 epochs=20,	2,								
	batch_size=64	,								
	architecture=ence d_model=64,	oder,								
70	$num_heads=2,$ $ff_dim=64,$	1160000 0	41.000 OF	002 24	2.05	0.00	2.0	20.05	0.000020	1.04
70	num_layers=1, dropout_rate=0.5	1168239.9	41000.00	883.24	2.95	0.98	2.8	39.95	0.000030	1.04
	epochs=20,	<u>~</u> ,								
	batch_size=128 architecture=ence	nder								
	d_model=64, num_heads=2,	,								
71	ff_dim=64,	2031856.2	61425.43	1165.95	3.37	0.96	2.53	141.8	0.001030	3.91
• -	num_layers=1, dropout_rate=0.:		01120110	1100.00	0.01	0.00	2.00	111.0	0.001000	0.01
	epochs=50, batch_size=64									
	architecture=ence	oder,								
	$d_{model=64}$, $num_{heads=2}$,									
72	ff_dim=64, num_layers=1,	1879207.2	71370.84	1108.87	3.19	0.96	2.42	77.97	0.000912	3.48
	dropout_rate=0.	2,								
	epochs=50, batch_size=128									
	architecture=ence	oder,								
	$d_{model}=64, \\ num_{heads}=2,$									
73	ff_dim=64, num layers=2,	309881.51	556.67	387.97	1.25	0.99	1.82	73.22	-0.000098	0.68
	dropout_rate=0.0	0,								
	batch_size=64									
	architecture=ence d model=64.	oder,								
	num_heads=2,									
74	ff_dim=64, num_layers=2,	533014.51	730.08	556.93	1.74	0.99	2.09	40.25	0.000025	1.07
	dropout_rate=0.0 epochs=20,	0,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

		3. 2000df					Transforme	Training		
						- 0		J	Overfit	$_{ m Loss}$
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSI	E Time	Gap	Ratio
								(s)	Cap	100010
	architecture=enc	oder,								
	$d_{model}=64,$ $num_{heads}=2,$									
75	ff_dim=64, num_layers=2,	381409.23	3 617.58	480.08	1.57	0.99	1.66	156.04	0.000049	1.23
	dropout_rate=0.	0,								
	epochs=50, batch_size=64									
	architecture=enc d_model=64,	oder,								
	num_heads=2, ff_dim=64,									
76	num_layers=2,	607918.94	1 779.69	630.01	1.99	0.99	1.82	87.22	0.000141	1.52
	dropout_rate=0. epochs=50,	0,								
	batch_size=128 architecture=enc	oder								
	$d_{model}=64$,	oder,								
77	num_heads=2, ff_dim=64,	3197345.6	321788 11	1421.83	3.98	0.94	2.79	76.13	0.001763	5.24
• • •	num_layers=2, dropout_rate=0.		221100.11	1121.00	0.00	0.01	2.10	10.10	0.001100	0.21
	epochs=20, batch_size=64	,								
	architecture=enc	oder,								
	$d_{model=64}$, $num_{heads=2}$,									
78	ff_dim=64, num_layers=2,	2325625.5	59 1525	1287.67	3.77	0.95	2.27	47.09	0.001037	2.89
	dropout_rate=0.	2,								
	epochs=20, batch_size=128									
	architecture=enc d_model=64,	oder,								
	num_heads=2, ff_dim=64,									
79	num_layers=2,	1772781.4	441331.46	1026.08	2.89	0.97	2.9	162.09	0.000916	4.13
	dropout_rate=0. epochs=50,	2,								
	batch_size=64 architecture=enc	oder								
	$d_{model}=64$,	oder,								
80	$num_heads=2, ff_dim=64,$	4760496.2	292181.86	1834.18	5.2	0.91	2.81	91.47	0.002906	9.59
00	num_layers=2, dropout_rate=0.		202101.00	1001.10	0.2	0.01	2.01	01.11	0.002000	0.00
	epochs=50, batch_size=128									
	architecture=enc	oder,								
	$d_{model}=64,$ $num_{heads}=2,$									
81	ff_dim=128, num layers=1,	650258.86	806.39	615.19	1.87	0.99	2.04	60.43	0.000093	1.26
	dropout_rate=0. epochs=20,	0,								
	batch_size=64									
	architecture=enc d_model=64,	oder,								
	num_heads=2, ff_dim=128,	******		* 00 00			0.40	00.44		
82	num_layers=1,	552621.97	743.39	562.69	1.81	0.99	2.42	38.11	-0.000099	0.79
	dropout_rate=0. epochs=20,	0,								
	batch_size=128 architecture=enc	oder.								
	d_model=64, num_heads=2,	,								
83	ff_dim=128,	350935.1	592.4	436.25	1.37	0.99	1.65	124.98	0.000025	1.12
	num_layers=1, dropout_rate=0.			- 5 5						
	epochs=50, batch_size=64									
	architecture=enc	oder,								
	d_model=64, num_heads=2,									
84	ff_dim=128, num_layers=1,	326408.31	1 571.32	401.28	1.28	0.99	1.86	69.37	-0.000042	0.84
	dropout_rate=0. epochs=50,	0,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 5	.u. Itesui	ts of the	ablation	i study o	i the	<u>Transforme</u>		lecture.	
								Training	Overfit	Loss
$^{\mathrm{ID}}$	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Cap	100010
	architecture=enc	oder,								
	d_model=64, num_heads=2,									
85	ff_dim=128, num_layers=1,	3079735.	371754.92	1467.02	4.2	0.94	2.45	65.48	0.001684	5.07
	dropout_rate=0.	.2,								
	epochs=20, batch_size=64									
	architecture=enc d_model=64,	oder,								
	num_heads=2,									
86	ff_dim=128, num_layers=1,		5 2140.99	1709.54	4.78	0.91	3.44	37.2	0.002490	4.96
	dropout_rate=0. epochs=20,	.2,								
	batch_size=128	,								
	architecture=enc d_model=64,	oder,								
87	$num_heads=2, ff_dim=128,$	2452600	non og o 1 n	1503.96	4.21	0.93	2.76	135.48	0.002061	8.06
01	num_layers=1, dropout_rate=0.		021858.12	1505.90	4.21	0.95	2.70	133.46	0.002061	0.00
	epochs=50, batch_size=64	-,								
	architecture=enc	oder,								
	d_model=64, num_heads=2,									
88	ff_dim=128, num_layers=1,	1350120.	851161.95	903.79	2.61	0.97	2.68	99.62	0.000583	2.72
	dropout_rate=0.	.2,								
	epochs=50, batch_size=128									
	architecture=enc d_model=64,	oder,								
	num_heads=2, ff_dim=128,	000444		45.4.00	4.40					0.00
89	num_layers=2,	380441.4	6 616.8	454.69	1.49	0.99	1.87	76.95	-0.000056	0.82
	dropout_rate=0 epochs=20,	.0,								
	batch_size=64 architecture=enc	oder.								
	d_model=64, num_heads=2,	,								
90	ff_dim=128, num_layers=2,	484104.5	6 695.78	513.55	1.65	0.99	2.16	46.35	-0.000125	0.72
	dropout_rate=0.	.0,								
	epochs=20, batch_size=128									
	architecture=enc d_model=64,	oder,								
	num_heads=2, ff_dim=128,									
91	num_layers=2,	251032.9	3 501.03	347.43	1.12	1	1.62	152.94	-0.000029	0.86
	dropout_rate=0. epochs=50,	.0,								
	batch_size=64 architecture=enc	oder								
	d_model=64, num_heads=2,	oder,								
92	ff_dim=128,	427018	653.47	487.27	1.58	0.99	1.84	83.48	0.000034	1.13
	num_layers=2, dropout_rate=0.					0.00		001-0	0.00000	
	epochs=50, batch_size=128									
	architecture=enc	oder,								
	d_model=64, num_heads=2,									
93	ff_dim=128, num_layers=2,		531715.06	1411.08	4.01	0.94	3.08	79.58	0.001586	4.79
	dropout_rate=0.epochs=20,	.2,								
	batch_size=64									
	architecture=enc d_model=64,	odei,								
94	num_heads=2, ff_dim=128,	6302557	662510.49	2049.63	5.72	0.88	3.76	48.06	0.003727	7.53
01	num_layers=2, dropout_rate=0.			2010.00	J.12	0.00	5.10	10.00	5.005121	,
	epochs=20, batch size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

							Transforme	Training		
ID		MOD	DAGE	24.5	MADE	D.2		Ü	Overfit	$_{ m Loss}$
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)		
	architecture=enco	der,								
	d_model=64, num_heads=2,									
95	ff_dim=128, num_layers=2,	4505598.	552122.64	1768.96	4.96	0.91	2.97	166.29	0.002816	12.08
	dropout_rate=0.2 epochs=50,	!,								
	batch_size=64	1								
	architecture=enco d_model=64,	der,								
96	num_heads=2, ff_dim=128,	4747389	922178.85	1836.96	5.2	0.91	3.02	96.14	0.002910	10
00	num_layers=2, dropout_rate=0.2		J	1000.00	0.2	0.01	0.02	00.11	0.002010	10
	epochs=50, batch_size=128									
	architecture=enco	der,								
	d_model=64, num_heads=4,									
97	ff_dim=64, num_layers=1,	404028.0	7 635.63	457.11	1.47	0.99	2.1	65.35	-0.000074	0.79
	dropout_rate=0.0 epochs=20,),								
	batch_size=64 architecture=enco	dor								
	$d_{model}=64$	der,								
98	num_heads=4, ff_dim=64,	713448.7	8 844.66	681.07	2.27	0.99	2.22	38.42	0.000029	1.06
	num_layers=1, dropout_rate=0.0									
	epochs=20, batch_size=128									
	architecture=enco d_model=64,	der,								
	num_heads=4,									
99	ff_dim=64, num_layers=1,	373744.8	3 611.35	433.17	1.39	0.99	1.77	131.49	0.000015	1.06
	dropout_rate=0.0 epochs=50,	',								
	batch_size=64 architecture=enco	der.								
	d_model=64, num_heads=4,	,								
100	ff_dim=64, num_layers=1,	452850.83	2 672.94	522.34	1.64	0.99	1.73	71.5	0.000053	1.21
	dropout_rate=0.0),								
	epochs=50, batch_size=128									
	architecture=enco d_model=64,	der,								
101	num_heads=4, ff_dim=64,	1004010	7 1111 07	060.0	9.00	0.00	1.07	64.60	0.000001	1.00
101	num_layers=1, dropout_rate=0.2		5 1111.27	963.3	3.02	0.98	1.97	64.69	0.000381	1.83
	epochs=20,	',								
	batch_size=64 architecture=enco	der,								
	$d_{model=64}$, $num_{heads=4}$,									
102	ff_dim=64, num_layers=1,	1658617.	851287.87	1138.38	3.61	0.97	2.12	38.6	0.000520	1.85
	dropout_rate=0.2 epochs=20,	!,								
	$batch_size=128$									
	architecture=enco d_model=64,	der,								
103	num_heads=4, ff_dim=64,	1019950	1 1946 57	1062.19	3	0.96	2.7	140.35	0.000910	3.79
103	num_layers=1, dropout_rate=0.2		1 1346.57	1002.19	3	0.90	2.1	140.55	0.000910	3.19
	epochs=50, batch size=64	•								
	architecture=enco	der,								
	d_model=64, num_heads=4,									
104	ff_dim=64, num_layers=1,	2360241.	791536.31	1217.45	3.42	0.95	2.73	76.29	0.001285	4.97
	dropout_rate=0.2 epochs=50,	!,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

							Transforme	Training		
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Ü	Overfit Gap	Loss Ratio
105	architecture=enc d_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0. epochs=20, batch_size=64	391504.81 ₀ ,	625.7	442.33	1.42	0.99	1.97	69.29	-0.000054	0.83
106	architecture=encd_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0. epochs=20, batch_size=128	471533.28 0,	686.68	510.62	1.7	0.99	2.1	43.15	-0.000104	0.76
107	architecture=encd_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0. epochs=50, batch_size=64	305757.06 o,	5 552.95	389.1	1.24	0.99	1.72	162.21	0.000005	1.02
108	architecture=encd_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0. epochs=50, batch_size=128	353582.18	5 594.63	423.07	1.32	0.99	1.87	92.24	-0.000002	0.99
109	architecture=encd_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0. epochs=20, batch_size=64	1927915.6	11388.49	1054.5	2.96	0.96	2.76	74.88	0.000845	2.8
110	architecture=encd_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0. epochs=20, batch_size=128	4095135.7	92023.64	1829.9	5.45	0.92	2.47	49.14	0.002189	4.62
111	architecture=encd_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0. epochs=50, batch_size=64	5492807.8	22343.67	1987.72	5.63	0.89	2.95	174.48	0.003483	14.38
112	architecture=encd_model=64, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0. epochs=50, batch_size=128	4391923.8	92095.69	1727.24	4.84	0.91	3.03	104.95	0.002703	10.36
113	architecture=encd_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0. epochs=20, batch_size=64	708310.05	841.61	692.59	2.17	0.99	1.91	58.8	0.000158	1.49
114	architecture=enced_model=64, num_heads=4, ff_dim=128, num_layers=1, dropout_rate=0. epochs=20, batch_size=128	436693.74	660.83	496.87	1.56	0.99	1.9	32.92	-0.000048	0.86

Table 5.6: Results of the ablation study of the Transformers architecture.

								Training		
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRM	SE Time	Overfit	Loss
110	Comig.	MSE	IUMSE	WIAL	WAIL	16	C V ICIVIS		Gap	Ratio
								(s)		
	architecture=end d_model=64,	coder,								
115	num_heads=4, ff_dim=128,	265547 2	7 515 91	251 40	1 19	0.99	1.67	191 51	0 000026	0.83
115	num_layers=1, dropout_rate=0	265547.37	1 313.31	351.49	1.13	0.99	1.67	131.51	-0.000036	0.83
	epochs=50, batch_size=64	.0,								
	architecture=enc	oder,								
	$d_{model}=64,$ $num_{heads}=4,$									
116	ff_dim=128, num_layers=1,	289063.7	1 537.65	377.11	1.21	0.99	1.73	71.23	-0.000058	0.77
	dropout_rate=0 epochs=50,	.0,								
	batch_size=128	. 1.								
	architecture=end d_model=64,	coder,								
117	num_heads=4, ff_dim=128,	1805722.5	591343 77	1023.81	2.91	0.96	3.03	59.98	0.000753	2.58
111	num_layers=1, dropout_rate=0		001010.11	1020.01	2.01	0.00	0.00	00.00	0.000100	2.00
	epochs=20, batch_size=64									
	architecture=end d_model=64,	coder,								
	$num_heads=4$,									
118	ff_dim=128, num_layers=1,	1497040.9	921223.54	937.65	2.7	0.97	2.69	34.79	0.000415	1.69
	dropout_rate=0 epochs=20,	.2,								
	batch_size=128 architecture=end	oder,								
	d_model=64, num_heads=4,									
119	ff_dim=128, num_layers=1,	3683003.2	241919.12	1508.86	4.16	0.93	3.25	140.99	0.002248	9.58
	dropout_rate=0 epochs=50,	.2,								
	batch_size=64	_								
	architecture=end d_model=64,	coder,								
120	num_heads=4, ff_dim=128,	2451143.5	521565 61	1238.99	3.45	0.95	2.79	76.9	0.001377	5.65
120	num_layers=1, dropout_rate=0		21000.01	1230.33	0.40	0.55	2.13	10.5	0.001377	0.00
	epochs=50, batch_size=128									
	architecture=enc	coder,								
	$num_heads=4$,									
121	ff_dim=128, num_layers=2,	458065.26	6 676.81	507.15	1.75	0.99	2.19	69.82	-0.000000	1
	dropout_rate=0 epochs=20,	.0,								
	batch_size=64 architecture=end	oder.								
	d_model=64, num_heads=4,	,								
122	ff_dim=128, num_layers=2,	596449.62	2 772.3	584.5	1.79	0.99	1.99	43.85	0.000023	1.06
	dropout_rate=0	.0,								
	epochs=20, batch_size=128									
	architecture=end d_model=64,	coder,								
123	num_heads=4, ff_dim=128,	237236.6	487.07	332.33	1.06	1	1.57	155.75	-0.000037	0.82
123	num_layers=2, dropout_rate=0		407.07	<i>აა</i> ∠.აა	1.00	1	1.57	155.75	-0.000037	0.62
	epochs=50, batch_size=64	•								
	architecture=enc	coder,								
	d_model=64, num_heads=4,									
124	ff_dim=128, num_layers=2,	343066.2	585.72	414.58	1.32	0.99	1.9	94.25	0.000004	1.02
	dropout_rate=0 epochs=50,	.0,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 5	.u. icsui	.05 01 0110	ablation	i study O	1 0110	1 ransforme		ccourc.	
			D. 160	2645		D 2		Training	Overfit	Loss
ID	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)		
	architecture=enced_model=64,	oder,								
	num_heads=4,									
125	ff_dim=128, num_layers=2,		521471.83	1218.25	3.51	0.96	2.37	75.41	0.001052	3.48
	dropout_rate=0.2 epochs=20,	2,								
	batch_size=64 architecture=ence									
	$d_{model}=64$	oder,								
126	$num_heads=4, ff_dim=128,$	3212496	871792.34	1454.86	4.11	0.94	2.86	49.98	0.001692	4.4
120	num_layers=2, dropout_rate=0.:		0.11.02.01	1101.00	1.11	0.01	2.00	10.00	0.001002	1.1
	epochs=20, batch_size=128									
	architecture=ence	oder,								
	d_model=64, num_heads=4,									
127	ff_dim=128, num_layers=2,	4059712.	022014.87	1736.44	4.99	0.92	2.34	175.58	0.002501	10.44
	dropout_rate=0.5 epochs=50,	2,								
	batch_size=64 architecture=ence	oder								
	$d_{model}=64$,	oder,								
128	num_heads=4, ff_dim=128,	1903228.	771379.58	1088.72	3.06	0.96	2.72	105.12	0.000990	4.23
	num_layers=2, dropout_rate=0.:									
	epochs=50, batch_size=128									
	architecture=ence d_model=32,	dec,								
	num_heads=2,									
129	ff_dim=64, num_layers=1,	392867.9	626.79	449.61	1.45	0.99	2.03	60.83	-0.000085	0.76
	dropout_rate=0.0 epochs=20,	υ,								
	batch_size=64 architecture=ence	dec.								
	d_model=32, num_heads=2,	,								
130	ff_dim=64, num_layers=1,	486019.3	4 697.15	484.34	1.54	0.99	2.25	33.8	-0.000117	0.74
	dropout_rate=0.	0,								
	epochs=20, batch_size=128									
	architecture=ence d_model=32,	dec,								
191	num_heads=2, ff_dim=64,	207501 1	0 545 50	200 77	1.00	0.00	1 77	100.2	0.000000	0.97
131	num_layers=1, dropout_rate=0.0	297591.19	9 545.52	382.77	1.23	0.99	1.77	129.3	-0.000029	0.87
	epochs=50,	σ,								
	batch_size=64 architecture=ence	dec,								
	d_model=32, num_heads=2,									
132	ff_dim=64, num_layers=1,	353400.7	6 594.48	435.67	1.41	0.99	1.8	70.01	-0.000024	0.91
	dropout_rate=0.0 epochs=50,	0,								
	$batch_size=128$,								
	architecture=ence d_model=32,	iec,								
133	num_heads=2, ff_dim=64,	690408.4	5 820 01	615.07	1.84	0.99	2.2	60.66	0.000105	1.29
199	num_layers=1, dropout rate=0.5		5 650.91	015.07	1.04	0.99	2.2	00.00	0.000103	1.29
	epochs=20, batch size=64	•								
	architecture=ence	dec,								
	d_model=32, num_heads=2,									
134	ff_dim=64, num_layers=1,		061124.43	906.74	2.71	0.98	2.45	35.11	0.000396	1.85
	dropout_rate=0.5 epochs=20,	2,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 0.	iwsuit	01 1110	abia0101.	. Study O	. 0110	Transforme			
								Training	Overfit	Loss
$^{\mathrm{ID}}$	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Gap	Itatio
	architecture=enco	łec,								
	d_model=32, num_heads=2,									
135	ff_dim=64, num_layers=1,	612503.51	782.63	621.44	2.03	0.99	1.96	141.33	0.000169	1.68
	dropout_rate=0.2	2,								
	epochs=50, batch_size=64									
	architecture=enco d_model=32,	lec,								
	$num_heads=2$,									
136	ff_dim=64, num_layers=1,	566333.35	752.55	597.68	1.98	0.99	2.08	74.45	0.000094	1.32
	dropout_rate=0.2 epochs=50,	2,								
	batch_size=128 architecture=enco	loc								
	$d_{model=32}$	iec,								
137	num_heads=2, ff_dim=64,	361359.6	601.13	438.16	1.42	0.99	1.83	69.79	-0.000067	0.79
101	num_layers=2, dropout_rate=0.0		001.10	100.10	1.12	0.00	1.00	00.10	0.000001	0.10
	epochs=20, batch_size=64									
	architecture=enco	łec,								
	$d_{model=32}$, $num_{heads=2}$,									
138	ff_dim=64, num_layers=2,	405105.02	636.48	455.13	1.47	0.99	2.08	42.65	-0.000113	0.71
	dropout_rate=0.0 epochs=20,	Ο,								
	$batch_size=128$									
	architecture=enco d_model=32,	iec,								
139	$num_heads=2, ff_dim=64,$	283045.89	532.02	367.01	1.18	0.99	1.69	162.56	-0.000022	0.9
109	num_layers=2, dropout_rate=0.0		552.02	307.01	1.10	0.99	1.09	102.50	-0.000022	0.9
	epochs=50, batch_size=64	,								
	architecture=enco	lec,								
	$d_{model=32}$, $num_{heads=2}$,									
140	ff_dim=64, num_layers=2,	345214.97	587.55	427.01	1.36	0.99	1.76	86.51	-0.000010	0.96
	dropout_rate=0.0 epochs=50,	Ο,								
	$batch_size=128$									
	architecture=enco d_model=32,	lec,								
1.41	num_heads=2, ff_dim=64,	407147.00	COD 40	471.00	1 40	0.00	0.05	74.90	0.000000	0.01
141	num_layers=2, dropout_rate=0.2	467147.03	683.48	471.96	1.46	0.99	2.05	74.38	-0.000032	0.91
	epochs=20,	2,								
	batch_size=64 architecture=enco	lec,								
	d_model=32, num_heads=2,									
142	ff_dim=64,	1114508.3	91055.7	813.05	2.41	0.98	2.58	44.11	0.000292	1.63
	num_layers=2, dropout_rate=0.2	2,								
	epochs=20, batch_size=128									
	architecture=enco d_model=32,	lec,								
	$num_heads=2$,									
143	ff_dim=64, num_layers=2,	716232.41	846.31	706.55	2.21	0.99	1.71	172.8	0.000246	2.02
	dropout_rate=0.2 epochs=50,	2,								
	batch_size=64 architecture=enco	lec								
	d_model=32,	,								
144	num_heads=2, ff_dim=64,	372053.14	609.96	430.09	1.35	0.99	1.87	97.72	-0.000007	0.97
	num_layers=2, dropout_rate=0.2				50	2.00		-		
	epochs=50, batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 5.6: Results of the ablation study of the Transformers architectural Training							ccourt.		
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Ü	Overfit Gap	Loss Ratio
145	architecture=encd d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0 epochs=20, batch_size=64	465987.42	682.63	488.31	1.55	0.99	2.19	59.07	-0.000036	0.9
146	architecture=encod_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0 epochs=20, batch_size=128	627958.01	792.44	610.88	1.99	0.99	2.23	34.63	0.000024	1.06
147	architecture=encod_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0 epochs=50, batch_size=64	279716.71),	528.88	370.77	1.18	0.99	1.69	128.07	-0.000030	0.86
148	architecture=encod_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.0 epochs=50, batch_size=128	363894.19	603.24	449.15	1.49	0.99	1.84	69.36	-0.000006	0.98
149	architecture=encod_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2 epochs=20, batch_size=64	412106.83	641.96	464.77	1.47	0.99	1.94	60.03	-0.000085	0.77
150	architecture=encd d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2 epochs=20, batch_size=128	795719.8	892.03	705.28	2.25	0.98	2.31	34.39	0.000109	1.25
151	architecture=ence d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2 epochs=50, batch_size=64	334274.85	578.17	411.25	1.3	0.99	1.74	131.83	-0.000013	0.94
152	architecture=encd d_model=32, num_heads=2, ff_dim=128, num_layers=1, dropout_rate=0.2 epochs=50, batch_size=128	445856.13	667.72	511.96	1.64	0.99	1.81	72.09	0.000028	1.1
153	architecture=encd_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0 epochs=20, batch_size=64	439506.01	662.95	496.24	1.61	0.99	1.96	67.99	0.000017	1.06
154	architecture=encd d_model=32, num_heads=2, ff_dim=128, num_layers=2, dropout_rate=0.0 epochs=20, batch_size=128	802043.25	895.57	730.75	2.44	0.98	2.33	40.84	0.000136	1.33

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 9	. 0. 1005010	5 Of the	ablation	i study Oi	UIIC	Transforme		ccourc.	
							.1	Training	Overfit	Loss
ID	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time		
								(s)	Gap	Ratio
-	architecture=enc	dec.						. ,		
	d_model=32,	400,								
155	num_heads=2, ff_dim=128,	300999.89	548 63	401.62	1.27	0.99	1.6	161.07	-0.000022	0.9
100	num_layers=2, dropout_rate=0.		040.00	401.02	1.21	0.33	1.0	101.07	-0.000022	0.3
	epochs=50,	Ο,								
	batch_size=64 architecture=ence	dec,								
	d_model=32, num heads=2,									
156	ff_dim=128,	283445.52	532.4	377.94	1.22	0.99	1.74	95.91	-0.000051	0.79
	num_layers=2, dropout_rate=0.									
	epochs=50, batch_size=128									
	architecture=ence	dec,								
	d_model=32, num_heads=2,									
157	ff_dim=128, num_layers=2,	481134.66	693.64	483.55	1.48	0.99	1.99	77.86	-0.000013	0.96
	dropout_rate=0.	2,								
	epochs=20, batch_size=64									
	architecture=enced model=32,	dec,								
	num_heads=2,									
158	ff_dim=128, num_layers=2,	579708.15	761.39	556.52	1.73	0.99	2.35	44.83	-0.000038	0.91
	dropout_rate=0.2 epochs=20,	2,								
	$batch_size=128$	1								
	architecture=enced_model=32,	aec,								
159	num_heads=2, ff_dim=128,	326170.28	571 11	410.15	1.27	0.99	1.69	169.43	-0.000010	0.96
109	num_layers=2, dropout_rate=0.		371.11	410.15	1.21	0.99	1.09	109.40	-0.000010	0.90
	epochs=50, batch_size=64	-,								
	architecture=ence	dec,								
	d_model=32, num_heads=2,									
160	ff_dim=128, num_layers=2,	472349.18	687.28	486.81	1.46	0.99	1.88	95.3	0.000049	1.18
	dropout_rate=0.	2,								
	epochs=50, batch_size=128									
	architecture=enced_model=32,	dec,								
	$num_heads=4,$									
161	ff_dim=64, num_layers=1,	453348.52	673.31	488.23	1.56	0.99	2.07	61.69	-0.000087	0.78
	dropout_rate=0. epochs=20,	0,								
	batch_size=64									
	architecture=enced_model=32,	dec,								
162	$num_heads=4, ff_dim=64,$	40 <i>ee</i> 4 7 2 <i>e</i>	627 60	440.49	1 45	0.00	0.11	24.96	0.000004	0.75
102	num_layers=1, dropout_rate=0.	406647.36	057.09	449.42	1.45	0.99	2.11	34.26	-0.000094	0.75
	epochs=20,	0,								
	batch_size=128 architecture=ence	dec.								
	d_model=32, num_heads=4,	,								
163	ff_dim=64,	320377.07	566.02	397.69	1.28	0.99	1.76	130.67	-0.000013	0.94
	num_layers=1, dropout_rate=0.									
	epochs=50, batch_size=64									
	architecture=ence	dec,								
	d_model=32, num_heads=4,									
164	ff_dim=64, num_layers=1,	325984.47	570.95	395.58	1.27	0.99	1.85	70.42	-0.000044	0.83
	dropout_rate=0. epochs=50,	0,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 5.	o. nesun	S OI UIC	abiation	study of	one .	<u>Fransforme</u>		ccture.	
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE		Overfit Gap	Loss Ratio
								(s)		
165	architecture=encd d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2 epochs=20, batch_size=64	760850.34	872.27	624.88	1.86	0.99	2.3	61.15	0.000132	1.34
166	architecture=encd d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2 epochs=20, batch_size=128	653871.06	808.62	627.08	1.99	0.99	2.12	34.56	0.000030	1.07
167	architecture=encd d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2 epochs=50, batch_size=64 architecture=encd	351302.36	592.71	411.62	1.27	0.99	1.77	132.94	-0.000009	0.96
168	d_model=32, num_heads=4, ff_dim=64, num_layers=1, dropout_rate=0.2 epochs=50, batch_size=128	790292.56	888.98	736.6	2.34	0.98	1.89	73.27	0.000256	1.91
169	architecture=encd d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0 epochs=20, batch_size=64	680910.41	825.17	663.42	2.09	0.99	1.97	71.98	0.000163	1.54
170	architecture=encd d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0 epochs=20, batch_size=128	863263.99	929.12	781.59	2.6	0.98	2.25	47.46	0.000211	1.56
171	architecture=encd d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0 epochs=50, batch_size=64	275703.21	525.07	364.95	1.15	0.99	1.67	172.03	-0.000018	0.91
172	architecture=encd d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.0 epochs=50, batch_size=128	382130.28	618.17	440.28	1.39	0.99	1.76	90.34	0.000016	1.06
173	architecture=encd d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2 epochs=20, batch_size=64	727010.31	852.65	662.25	2.12	0.99	2.76	75.71	0.000187	1.61
174	architecture=encd d_model=32, num_heads=4, ff_dim=64, num_layers=2, dropout_rate=0.2 epochs=20, batch_size=128	2610860.6	11615.82	1462.24	4.58	0.95	2.38	44.57	0.001306	3.76

Table 5.6: Results of the ablation study of the Transformers architecture.

							<u> Transforme</u>	Training	-	
			D			D 2		Ü	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Сцр	100010
	architecture=enco	dec,								
	$d_{model=32}$, $num_{heads=4}$,									
175	ff_dim=64, num_layers=2,	285124.65	533.97	374.78	1.2	0.99	1.71	169.5	-0.000059	0.77
	dropout_rate=0.2	2,								
	epochs=50, batch_size=64									
	architecture=enco d_model=32,	dec,								
150	num_heads=4, ff_dim=64,	== 0000 00	050 50	= 4 = 0 =	0.00	0.00	1 50	00.54	0.000070	0.00
176	num_layers=2,	773926.29	879.73	747.37	2.38	0.98	1.76	93.54	0.000273	2.08
	dropout_rate=0.2 epochs=50,	2,								
	batch_size=128 architecture=enco	dec.								
	d_model=32, num_heads=4,	,								
177	ff_dim=128,	363578.02	602.97	434.69	1.4	0.99	1.94	57.54	-0.000041	0.86
	num_layers=1, dropout_rate=0.0									
	epochs=20, batch_size=64									
	architecture=enco	dec,								
	num_heads=4,									
178	ff_dim=128, num_layers=1,	867621.74	931.46	776.38	2.53	0.98	2.12	32.49	0.000242	1.7
	dropout_rate=0.0 epochs=20,	0,								
	$batch_size=128$	J								
	architecture=encod_model=32,	iec,								
179	num_heads=4, ff_dim=128,	295003.12	543 14	394.77	1.24	0.99	1.63	129.44	-0.000009	0.95
110	num_layers=1, dropout_rate=0.0		010.11	001.11	1.21	0.00	1.00	120.11	0.000000	0.00
	epochs=50, batch_size=64									
	architecture=enco	dec,								
	$d_{model=32}$, $num_{heads=4}$,									
180	ff_dim=128, num_layers=1,	346561.68	588.69	423.01	1.35	0.99	1.79	69.92	-0.000015	0.94
	dropout_rate=0.0 epochs=50,	0,								
	$batch_size=128$	1								
	architecture=encod_model=32,	iec,								
181	num_heads=4, ff_dim=128,	576241.29	750 11	562.22	1.72	0.99	1.99	58.42	0.000046	1.13
101	num_layers=1, dropout rate=0.2		103.11	302.22	1.12	0.33	1.33	30.42	0.000040	1.10
	epochs=20, batch_size=64	,								
	architecture=enco	dec,								
	$d_{model=32}$, $num_{heads=4}$,									
182	ff_dim=128, num_layers=1,	554178.98	744.43	550.86	1.72	0.99	2.25	33.32	-0.000029	0.93
	dropout_rate=0.2 epochs=20,	2,								
	$batch_size=128$	_								
	architecture=enco d_model=32,	dec,								
109	num_heads=4, ff_dim=128,	100700 77	600.94	E20.04	1.00	0.00	1 75	190.4	0.000007	1 41
183	num_layers=1, dropout_rate=0.2	489780.77	099.84	532.24	1.62	0.99	1.75	130.4	0.000097	1.41
	epochs=50,	- ,								
	batch_size=64 architecture=ence	dec,								
	d_model=32, num_heads=4,									
184	ff_dim=128, num layers=1,	628367.04	792.7	643.26	2.05	0.99	1.78	72.7	0.000164	1.62
	dropout_rate=0.2	2,								
	epochs=50, batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 5.	o. nesun	s or the	abiation	study of	the .	Iransforme		ecture.	
								Training	Overfit	Loss
ID	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Сар	100010
	architecture=ence	dec,								
	d_model=32, num_heads=4,									
185	ff_dim=128, num_layers=2,	376820.99	613.86	456.1	1.51	0.99	1.88	71.22	-0.000026	0.91
	dropout_rate=0.0	0,								
	epochs=20, batch_size=64									
	architecture=ence d_model=32,	dec,								
	num_heads=4, ff dim=128,									
186	num_layers=2,	397334.6	630.34	448.23	1.43	0.99	2.03	61.38	-0.000068	0.8
	dropout_rate=0.0 epochs=20,	0,								
	batch_size=128 architecture=ence	dec								
	d_model=32,	100,								
187	num_heads=4, ff_dim=128,	262107.77	511.96	358.17	1.13	0.99	1.63	181.14	-0.000022	0.89
	num_layers=2, dropout_rate=0.0									
	epochs=50, batch_size=64									
	architecture=ence d_model=32,	dec,								
	num_heads=4,									
188	ff_dim=128, num_layers=2,	527939.82	726.59	582.15	1.79	0.99	1.66	104	0.000116	1.47
	dropout_rate=0.0 epochs=50,	0,								
	batch_size=128 architecture=ence	dec								
	d_model=32,	100,								
189	num_heads=4, ff_dim=128,	1850771.6	1360.43	1185	3.69	0.96	2.16	91.16	0.000923	3.73
	num_layers=2, dropout_rate=0.:									
	epochs=20, batch_size=64									
	architecture=enced_model=32,	dec,								
	num_heads=4,									
190	ff_dim=128, num_layers=2,	436458.93	660.65	479.12	1.52	0.99	2.02	52.63	-0.000115	0.72
	dropout_rate=0.5 epochs=20,	2,								
	batch_size=128 architecture=ence	dec								
	$d_{model=32}$	100,								
191	num_heads=4, ff_dim=128,	306310.1	553.45	390.78	1.21	0.99	1.7	197.21	-0.000026	0.89
	num_layers=2, dropout_rate=0.:		0001-0			0.00			0.0000=0	0.00
	epochs=50, batch_size=64									
	architecture=ence	dec,								
	d_model=32, num_heads=4,									
192	ff_dim=128, num_layers=2,	353024.22	594.16	428.22	1.36	0.99	1.75	108.71	-0.000039	0.86
	dropout_rate=0.2 epochs=50,	2,								
	batch_size=128 architecture=ence	dec								
	d_model=64, num_heads=2,	,								
193	$ff_dim=64$,	618743.21	786.6	611.69	1.9	0.99	1.9	75.53	0.000118	1.39
	num_layers=1, dropout_rate=0.0		, , , , ,	0		0.00		, , , , ,	0.0000	
	epochs=20, batch_size=64									
	architecture=enced_model=64,	dec,								
	num_heads=2,									
194	ff_dim=64, num_layers=1,	695108.93	833.73	660.27	2.18	0.99	2.26	42.99	0.000003	1.01
	dropout_rate=0.0 epochs=20,	U,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

		3. 2000 dI						ners archit		
ID		MOD	DMCE	24.5	MADE	D 2	CLIDAG	J	Overfit	$_{ m Loss}$
ID	Config.	\mathbf{MSE}	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMS		Gap	Ratio
								(s)		
	architecture=enc d model=64,	dec,								
	num_heads=2,									
195	ff_dim=64, num_layers=1,	268710.07	518.37	356.59	1.13	0.99	1.65	153.95	-0.000046	0.8
	dropout_rate=0. epochs=50,	0,								
	batch_size=64 architecture=enc	dec								
	d_model=64, num_heads=2,	400,								
196	$ff_dim=64$,	326294.94	571.22	395.9	1.28	0.99	1.89	73.71	-0.000045	0.83
	num_layers=1, dropout_rate=0.	0,								
	epochs=50, batch_size=128									
	architecture=enc d_model=64,	dec,								
107	num_heads=2, ff_dim=64,	#10000 O		F91 09	1.60	0.00	0.00	00.54	0.000010	1.09
197	num_layers=1, dropout_rate=0.	512623.24	1 715.98	531.03	1.63	0.99	2.02	66.54	0.000010	1.03
	epochs=20,	2,								
	batch_size=64 architecture=enc	dec,								
	d_model=64, num_heads=2,									
198	ff_dim=64, num_layers=1,	1345593.5	3 1160	964.54	2.92	0.97	2.26	37.18	0.000465	2.03
	dropout_rate=0. epochs=20,	2,								
	$batch_size=128$	1								
	architecture=enc d_model=64,	dec,								
199	num_heads=2, ff_dim=64,	709754.96	842.47	672.36	2.08	0.99	1.82	139.25	0.000242	2
	num_layers=1, dropout_rate=0.			0,2.00		0.00			0.000=	
	epochs=50, batch_size=64									
	architecture=enc d_model=64,	dec,								
200	num_heads=2, ff_dim=64,	0400000			2.24		4.00	-0.04	0.000010	4 =0
200	num_layers=1,	813668.34	1 902.04	742.75	2.34	0.98	1.96	78.01	0.000243	1.78
	dropout_rate=0. epochs=50,	2,								
	batch_size=128 architecture=enc	dec,								
	d_model=64, num_heads=2,									
201	ff_dim=64, num_layers=2,	305143.61	552.4	383.5	1.24	0.99	1.82	74.48	-0.000085	0.71
	dropout_rate=0. epochs=20,	0,								
	$batch_size=64$,								
	architecture=enc d_model=64,	dec,								
202	$num_heads=2, ff_dim=64,$	1273157.9	111198 34	983.75	3.15	0.98	2.09	43.15	0.000465	2.15
202	num_layers=2, dropout_rate=0.		711120.01	000.10	0.10	0.00	2.00	10.10	0.000100	2.10
	epochs=20, batch_size=128									
	architecture=enc d_model=64,	dec,								
	num_heads=2, ff_dim=64,									
203	num_layers=2,	329278.12	2 573.83	425.03	1.39	0.99	1.68	162.8	0.000012	1.06
	dropout_rate=0. epochs=50,	υ,								
	batch_size=64 architecture=enc	dec,								
	d_model=64, num_heads=2,	•								
204	ff_dim=64, num_layers=2,	347821.85	589.76	430.28	1.35	0.99	1.66	90.73	-0.000014	0.95
	dropout_rate=0. epochs=50,	0,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 5	.u. nesun	ts of the	abiatioi	i study 0.	i the	<u> Transforme</u>		ecture.	
						- 2		Training	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)		
	architecture=enc d_model=64,	dec,								
	num_heads=2,									
205	ff_dim=64, num_layers=2,	1046856.8	361023.16	825.66	2.46	0.98	2.1	80.33	0.000378	2.12
	dropout_rate=0. epochs=20,	2,								
	batch_size=64 architecture=enc	dec								
	$d_{model}=64$,	acc,								
206	num_heads=2, ff_dim=64,	1123154.7	61059.79	761.22	2.24	0.98	2.76	47.83	0.000309	1.68
	num_layers=2, dropout_rate=0.									
	epochs=20, batch_size=128									
	architecture=enc d_model=64,	dec,								
	num_heads=2,									
207	ff_dim=64, num_layers=2,	678687.19	9 823.82	623.75	1.85	0.99	1.87	173.38	0.000223	1.93
	dropout_rate=0. epochs=50,	2,								
	batch_size=64 architecture=enc	dec,								
	$d_{model}=64,$ $num_{heads}=2,$									
208	ff_dim=64, num_layers=2,	662090.13	813.69	641.16	2.02	0.99	1.85	96.9	0.000174	1.63
	dropout_rate=0. epochs=50,	2,								
	$batch_size=128$	_								
	architecture=enc d_model=64,	dec,								
209	num_heads=2, ff_dim=128,	361820.23	2 601 51	427.96	1.38	0.99	1.98	67.41	-0.000083	0.75
203	num_layers=1, dropout_rate=0.		001.01	421.30	1.50	0.55	1.30	07.41	-0.000003	0.75
	epochs=20, batch_size=64	,								
	architecture=enc	dec,								
	$d_{model}=64, \\ num_{heads}=2,$									
210	ff_dim=128, num_layers=1,	451270.2	671.77	495.92	1.63	0.99	2.13	38.59	-0.000067	0.82
	dropout_rate=0. epochs=20,	0,								
	batch_size=128 architecture=enc	dec								
	d_model=64, num_heads=2,	400,								
211	ff_dim=128,	516965.32	2 719	567.3	1.79	0.99	1.77	138.21	0.000127	1.56
	num_layers=1, dropout_rate=0.									
	epochs=50, batch_size=64									
	architecture=enc d_model=64,	dec,								
010	num_heads=2, ff_dim=128,	200040 20		200 54	1.10	0.00		50.54	0.0000.40	0.00
212	num_layers=1, dropout_rate=0.	290640.26	539.11	369.54	1.19	0.99	1.77	76.54	-0.000040	0.83
	epochs=50,	0,								
	batch_size=128 architecture=enc	dec,								
	$d_{model}=64, \\ num_{heads}=2,$									
213	ff_dim=128, num_layers=1,	886459.27	941.52	758.33	2.26	0.98	2.14	65.07	0.000270	1.81
	dropout_rate=0. epochs=20,	2,								
	batch_size=64	doa								
	architecture=enc	uec,								
214	num_heads=2, ff_dim=128,	821741.93	3 906.5	714.03	2.19	0.98	2.19	37.76	0.000109	1.24
	num_layers=1, dropout_rate=0.			1.00		2.00		20		
	epochs=20, batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

Loss Ratio	Overfit	Training							
Ratio			~	- 2					
	Gap	SE Time	CVRMS	\mathbb{R}^2	MAPE	MAE	RMSE	onfig. MSE	ID ·
	P	(s)							
								chitecture=encdec, model=64,	
	0.000000	100.10			2.22			m_heads=2,	1
3 2.34	0.000328	138.16	1.9	0.98	2.23	739.9	917.58	m_layers=1, 041949.29	210
								opout $_{\text{rate}=0.2}$, ochs $=50$,	
								tch_size=64 chitecture=encdec,	
								_model=64, m_heads=2,	
2.18	0.000329	75.88	1.84	0.98	2.45	789.7	947.36	_dim=128, m_layers=1, 897492.52	216
								opout_rate= 0.2 , ochs= 50 ,	
								tch_size=128	1
								chitecture=encdec, _model=64,	
1.04	0.000014	76.58	1.92	0.99	1.7	560.17	746.36	m_heads=2, _dim=128, _557046.97	217
1.01	0.000011	10.00	1.02	0.00	1.1	000.11	110.00	m_layers=2, opout_rate=0.0,	1
								ochs=20, tch_size=64	
								chitecture=encdec, model=64,	
0.00	0.0001.40	40.00	1.05	0.00	1.00	101 10	200.00	m_heads=2,	1
2 0.63	-0.000142	46.03	1.97	0.99	1.36	421.16	603.03	_dim=128, m_layers=2, opout_rate=0.0,	210
								ochs=20, tch_size=128	•
								chitecture=encdec,	
								_model=64, .m_heads=2,	1
1.01	0.000002	168.06	1.61	0.99	1.22	387.58	541.84	$_{\text{m}}^{\text{dim}=128}, 293592.55$ $_{\text{m}}_{\text{layers}=2}, 293592.55$	110
								opout_rate=0.0, ochs=50,	
								tch_size=64 chitecture=encdec,	
								_model=64, .m_heads=2,	
1.35	0.000086	90.66	1.78	0.99	1.69	550.75	697.25	_dim=128, m_layers=2, 486155.32	220
								opout_rate=0.0,	
								ochs=50, tch_size=128	1
								chitecture=encdec, _model=64,	
3.15	0.000776	77 13	2.07	0.97	3 27	1005 68	91901 34	_dim=128, 1667551.2	001
3.10	0.000770	11.10	2.01	0.91	5.21	1099.00	21231.04	$m_{\text{layers}=2}$, 1007551.2 opout_rate=0.2,	
								ochs=20, tch size=64	
								chitecture=encdec,	
								m_heads=2,	1
3 1.51	0.000238	50.18	2.1	0.98	2.71	852.25	71013.72	m_layers=2, 1027016.1	144
								ochs=20,	
								chitecture=encdec,	
								_model=64, .m_heads=2,	1
3 1.2	0.000048	182.21	1.74	0.99	1.49	485.23	645.08	$_{\text{m}}^{\text{dim}=128}, 416133.29$	
								$\begin{array}{l} \text{opout_rate=0.2,} \\ \text{ochs=50,} \end{array}$	
								tch_size=64	1
								_model=64,	
1.13	0.000035	99.66	1.83	0.99	1.5	492.22	673.47	_dim=128, 453565.98	224
								$m_{\text{layers}=2}$, opout_rate=0.2,	
								ochs=50, tch_size=128	
3							71013.72	m_heads=2,	222

Table 5.6: Results of the ablation study of the Transformers architecture.

		2. 2505410	- 01 UIIO		- 2000, 0.	- 0110	ransioi me	Training		
		3.605	DMCT	3.6.5		ъ 2		· ·	Overfit	Loss
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)		
	architecture=enco	lec,								
	num_heads=4,									
225	ff_dim=64, num_layers=1,	383474.61	619.25	444.78	1.4	0.99	1.92	66.35	-0.000049	0.84
	dropout_rate=0.0 epochs=20,),								
	$batch_size=64$									
	architecture=ence d_model=64,	lec,								
224	num_heads=4, ff_dim=64,	* 00000	== 0.04	40 = F 0	1.00	0.00	2.04	00.00	0.000000	1.05
226	num_layers=1,	598833	773.84	607.58	1.93	0.99	2.04	38.62	0.000028	1.07
	dropout_rate=0.0 epochs=20,),								
	batch_size=128 architecture=enco	lec								
	$d_{model}=64$,	rec,								
227	$num_heads=4, ff_dim=64,$	267255.23	516.97	353.02	1.13	0.99	1.68	138.25	-0.000034	0.84
	num_layers=1, dropout_rate=0.0			000.02		0.00			0.00000	0.0-
	epochs=50, batch_size=64									
	architecture=enco	lec,								
	$d_{model}=64, \\ num_{heads}=4,$									
228	ff_dim=64, num_layers=1,	349280.53	591	425.4	1.4	0.99	1.92	74.4	-0.000028	0.89
	dropout_rate=0.0 epochs=50,),								
	$batch_size=128$	_								
	architecture=ence d_model=64,	iec,								
000	num_heads=4, ff_dim=64,	1070957 5	F1097 40	005 00	0.62	0.00	1.00	C4 FF	0.000200	0.00
229	num_layers=1, dropout_rate=0.2	1076357.5	51037.48	865.36	2.63	0.98	1.98	64.55	0.000382	2.09
	epochs=20,	-,								
	batch_size=64 architecture=enco	lec,								
	$d_{model}=64,$ $num_{heads}=4,$									
230	$ff_dim=64$,	1349117.4	51161.52	999.78	3.11	0.97	2.14	37.99	0.000499	2.18
	num_layers=1, dropout_rate=0.2	2,								
	epochs=20, batch_size=128									
	architecture=enco	lec,								
	num_heads=4,									
231	ff_dim=64, num_layers=1,	781237.04	883.88	677.06	2.03	0.98	1.95	140.63	0.000288	2.18
	dropout_rate=0.2 epochs=50,	2,								
	$batch_size=64$	loa								
	architecture=enco	iec,								
232	$num_heads=4, ff_dim=64,$	730305.57	854 58	715.71	2.35	0.99	1.87	82.51	0.000249	2
202	num_layers=1, dropout_rate=0.2		554.00	110.11	2.00	0.00	1.01	02.01	0.000240	2
	epochs=50, batch size=128	,								
	architecture=enco	lec,								
	$d_{model=64}$, $num_{heads=4}$,									
233	ff_dim=64, num_layers=2,	400385.81	632.76	451.73	1.43	0.99	1.89	76.84	-0.000012	0.96
	dropout_rate=0.0),								
	epochs=20, batch_size=64									
	architecture=enco d_model=64,	lec,								
a = .	num_heads=4, ff dim=64,	00065		100				10.55	0.00	0 ==
234	num_layers=2,	380961.65	617.22	438.02	1.41	0.99	1.93	46.62	-0.000072	0.78
	dropout_rate=0.0 epochs=20,	J,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 9	. 0. 10.501	05 01 0110	ablatioi	i study O.	1 0110	Transforme		occourc.	
								Training	Overfit	Loss
$^{\mathrm{ID}}$	Config.	MSE	\mathbf{RMSE}	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time		Ratio
								(s)	Gap	nano
	architecture=enc	dec,								
	$d_{model}=64,$ $num_{heads}=4,$	•								
235	ff_dim=64,	282194.61	1 531.22	368.2	1.18	0.99	1.69	168.68	-0.000010	0.95
	num_layers=2, dropout_rate=0.									
	epochs=50, batch_size=64									
	architecture=enc	dec,								
	$d_{model}=64,$ $num_{heads}=4,$									
236	ff_dim=64, num_layers=2,	277445.58	526.73	367.75	1.18	0.99	1.72	96.28	-0.000054	0.78
	dropout_rate=0. epochs=50,	0,								
	$batch_size=128$	_								
	architecture=enc d_model=64,	dec,								
237	num_heads=4, ff_dim=64,	1050054	27109E 11	778.54	2.32	0.98	2.24	78.21	0.000244	1.92
231	num_layers=2, dropout_rate=0.	1050854.3	011020.11	110.04	2.32	0.96	2.24	10.21	0.000344	1.92
	epochs=20,	2,								
	batch_size=64 architecture=enc	dec,								
	d_model=64, num_heads=4,									
238	ff_dim=64, num_layers=2,	2937457.0	011713.9	1487.15	4.43	0.94	2.39	48.05	0.001525	4.19
	dropout_rate=0.	2,								
	epochs=20, batch_size=128									
	architecture=enc d_model=64,	dec,								
	num_heads=4,									
239	ff_dim=64, num_layers=2,	352611.74	1 593.81	412.31	1.27	0.99	1.8	181.93	-0.000006	0.98
	dropout_rate=0. epochs=50,	2,								
	batch_size=64 architecture=enc	dec								
	d_model=64, num_heads=4,	acc,								
240	ff_dim=64,	852213.63	3 923.15	635.25	1.83	0.98	2.16	106.64	0.000298	2.05
	num_layers=2, dropout_rate=0.		, , , , , , , , , , , , , , , , , , , ,	000.20		0.00			0.000=00	
	epochs=50, batch_size=128									
	architecture=enc	dec,								
	d_model=64, num_heads=4,									
241	ff_dim=128, num_layers=1,	343403.56	5 586.01	414.94	1.32	0.99	1.83	62.88	-0.000091	0.72
	dropout_rate=0. epochs=20,	0,								
	$batch_size=64$	1								
	architecture=enc d_model=64,	aec,								
242	num_heads=4, ff_dim=128,	371587.66	609 58	440.76	1.44	0.99	1.92	35.68	-0.000088	0.74
242	num_layers=1, dropout_rate=0.		003.50	440.70	1.44	0.55	1.32	33.00	-0.000000	0.74
	epochs=20, batch_size=128	•								
	architecture=enc	dec,								
	$d_{model}=64,$ $num_{heads}=4,$									
243	ff_dim=128, num_layers=1,	300674.17	7 548.34	402.02	1.29	0.99	1.65	141.77	-0.000023	0.9
	dropout_rate=0. epochs=50,	0,								
	batch_size=64	_								
	architecture=enc d_model=64,	dec,								
0.4.4	num_heads=4, ff_dim=128,	206560.63	1 54450	200.00	1.00	0.00	1.70	77.00	0.000050	0.9
244	num_layers=1, dropout_rate=0.	296569.61	1 544.58	380.86	1.23	0.99	1.76	77.89	-0.000050	0.8
	epochs=50,	υ,								
	batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

	Table 5	.u. Hesul	ts of the	ablation	i study o	i the	1ransforme		lecture.	
								Training	Overfit	$_{ m Loss}$
ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSE	Time	Gap	Ratio
								(s)	Cap	100010
	architecture=enc	dec,								
	d_model=64, num_heads=4,									
245	ff_dim=128, num_layers=1,	526203.04	4 725.4	533.99	1.67	0.99	2.02	64.13	0.000007	1.02
	dropout_rate=0.	2,								
	$epochs=20,$ $batch_size=64$									
	architecture=enced_model=64,									
246	$num_heads=4$,	4320708.0								
	ff_dim=128, num_layers=1,	1924.34	5.92	0.92	2.12	35.95	0.002499	6.64		
	dropout_rate=0. epochs=20,									
	batch_size=128	J								
	architecture=enc d_model=64,	dec,								
247	$num_heads=4, ff_dim=128,$	539082.25	5 734 22	526.31	1.56	0.99	1.95	141.05	0.000134	1.58
211	num_layers=1, dropout_rate=0.		701.22	020.01	1.00	0.00	1.00	111.00	0.000101	1.00
	epochs=50, batch_size=64	•								
	architecture=ence	dec,								
	$d_{model}=64, \\ num_{heads}=4,$									
248	ff_dim=128, num_layers=1,	1114332.0	021055.62	919	2.98	0.98	1.96	77.35	0.000486	2.76
	dropout_rate=0. epochs=50,	2,								
	$batch_size=128$									
	architecture=ence d_model=64,	dec,								
0.40	num_heads=4, ff_dim=128,	240000 00	709.00	400.0	1.90	0.00	1.00	74.00	0.000000	0.70
249	num_layers=2, dropout_rate=0.	340888.08	5 363.60	422.8	1.38	0.99	1.88	74.29	-0.000060	0.79
	epochs=20,	0,								
	batch_size=64 architecture=ence	dec,								
	d_model=64, num_heads=4,									
250	ff_dim=128, num_layers=2,	1288543.5	561135.14	990.04	3.17	0.97	2.16	49.11	0.000519	2.44
	dropout_rate=0.	0,								
	$epochs=20, \\ batch_size=128$									
	architecture=enced_model=64,	dec,								
251	num_heads=4, ff_dim=128,									
	$num_layers=2$,	309046.68	8 555.92	389.33	1.24	0.99	1.62	170.99	0.000012	1.06
	$dropout_rate=0.$ $epochs=50,$	0,								
	batch_size=64 architecture=encdec,									
252	d_model=64, num_heads=4,	400,								
	ff_dim=128,	636628.53	3 797.89	656.13	2.13	0.99	1.89	106.41	0.000151	1.53
	num_layers=2, dropout_rate=0.									
	$epochs=50$, $batch_size=128$									
	architecture=enced_model=64,	dec,								
	$num_heads=4$,									
253	ff_dim=128, num_layers=2,	1021467.7	771010.68	809.35	2.42	0.98	2.52	82.26	0.000353	2.03
	dropout_rate=0. epochs=20,	2,								
	batch_size=64 architecture=encdec,									
	$d_{model}=64$	uec,								
254	num_heads=4, ff_dim=128,	661192.42	2 813 14	575.04	1.76	0.99	2.41	51.35	0.000014	1.03
404	num_layers=2, dropout_rate=0.		- 010.14	010.04	1.10	0.00	₩. 11	51.50	5.000014	1.00
	epochs=20, batch_size=128									

Table 5.6: Results of the ablation study of the Transformers architecture.

ID	Config.	MSE	RMSE	MAE	MAPE	\mathbb{R}^2	CVRMSI	Training E Time (s)	Overfit Gap	Loss Ratio
255	architecture=encd d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2 epochs=50, batch_size=64	341383.52	584.28	379.8	1.2	0.99	1.77	180.8	0.000010	1.05
256	architecture=encd d_model=64, num_heads=4, ff_dim=128, num_layers=2, dropout_rate=0.2 epochs=50, batch_size=128	673621.17	820.74	629.08	1.89	0.99	2.2	103.91	0.000190	1.71