Viscosity project summary_Section 2_ML 1275cP

Table of content

Standard 1275 cP

Machine learning segment

Set: Half

Observation of trends

Observation: GPR - scaling: Multiplication

The percentage error for this iteration is relatively consistent, ranging from -7 to -8%. A lot of repetitions in the sets of parameters generated are observed. The first 2 trials have the same sets of parameters and the rest of the trials share another set of parameters.

The most common set of parameters is: (AR, DR, DA, DD)
15.3, 7.0, 2.6, 4.5

Observation: GPR - scaling: Without

Repetitions of sets of parameters are also observed. The percentage error ranges from -3 to -7%, thus not as consistent as the trials with GPR - scaling: Multiplication.

The most common set of parameters is: (AR, DR, DA, DD)
11.1, 13.7, 4.1, 5.5

Observation: LIN - scaling: Multiplication

For this iteration, not as many repetitions are seen in the sets of parameters generated. Thus, the percentage errors are more distributed, ranging from -3 to -17%, with an average of -7%.

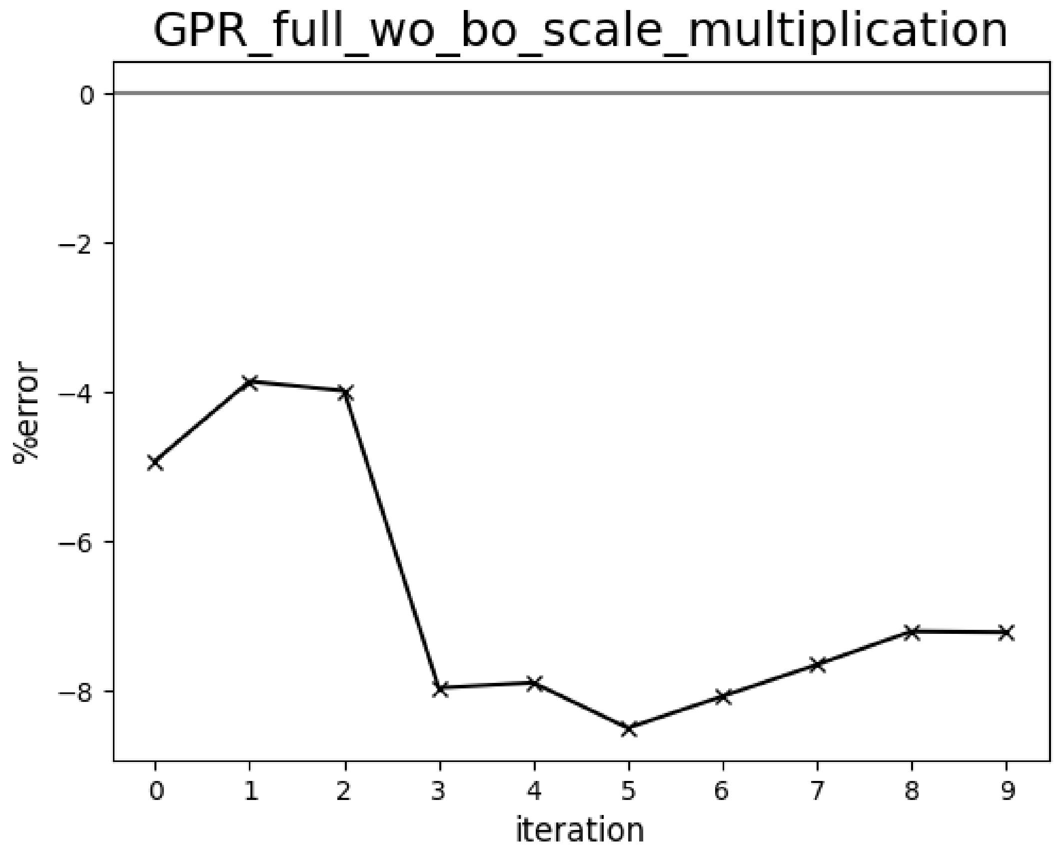
The most common set of parameters is: (AR, DR, DA, DD)
11.5, 7.3, 4.2, 3.6

Observation: LIN - scaling: Without

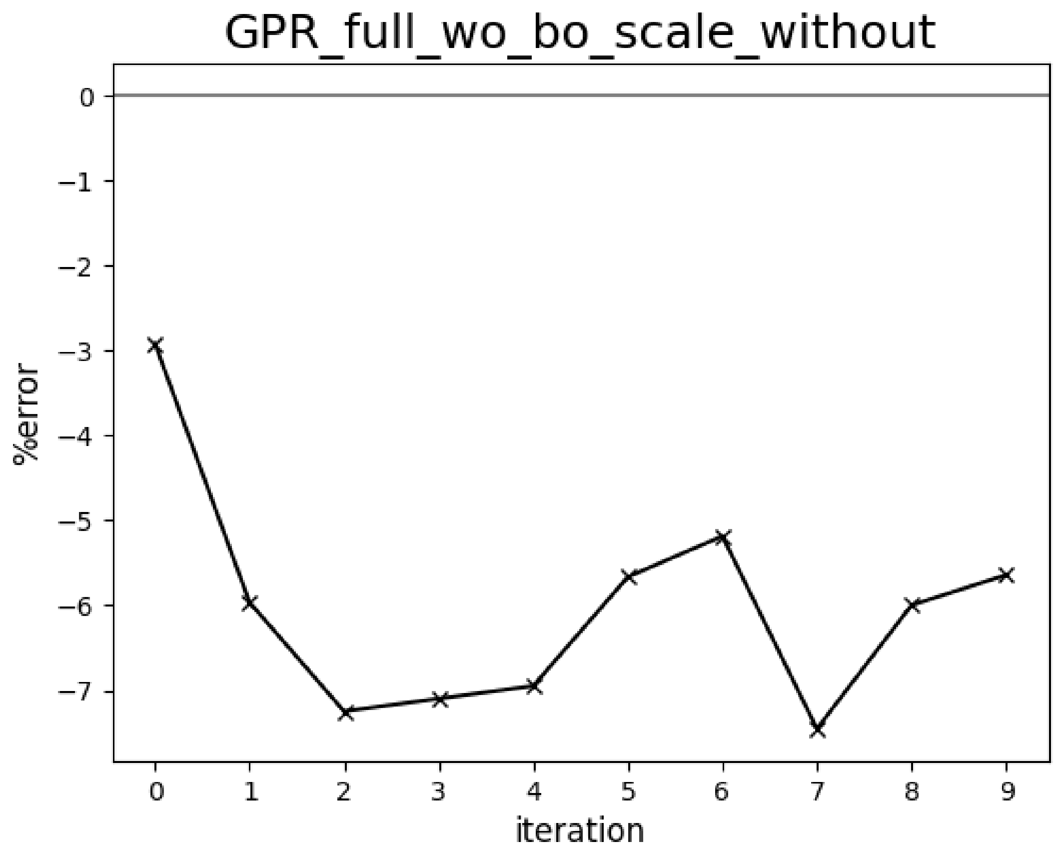
There are a lot of repetitions in the parameters generated, the parameters remained the same from the fourth trial onwards. The average is around -8%.

The most common set of parameters is: (AR, DR, DA, DD)
15.3, 7.0, 2.6, 4.5

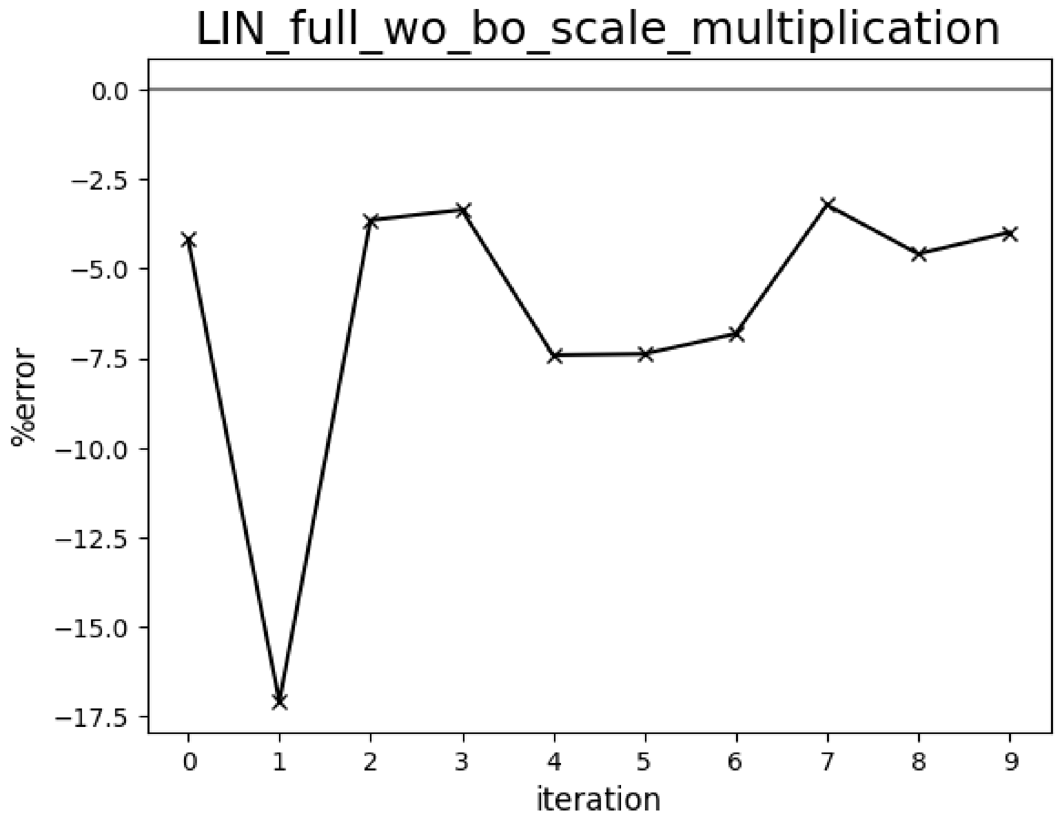
GPR scaling: multiplication



GPR scaling: without



LIN scaling: multiplication



LIN scaling: without

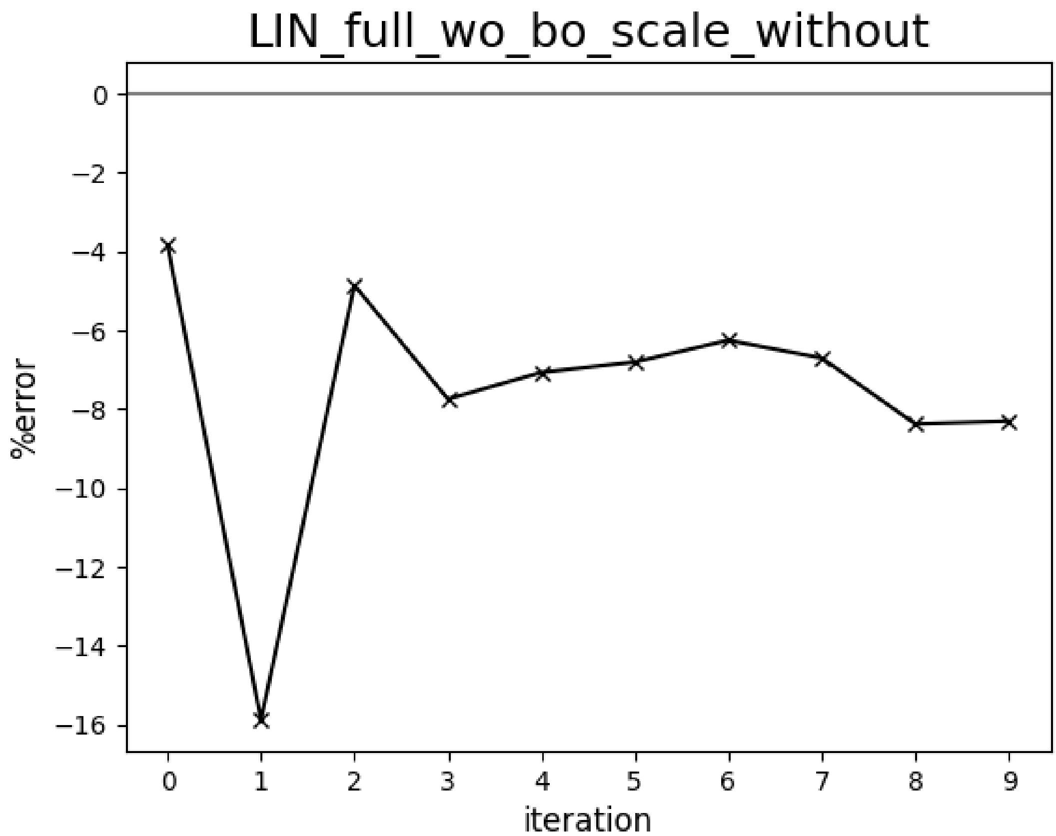
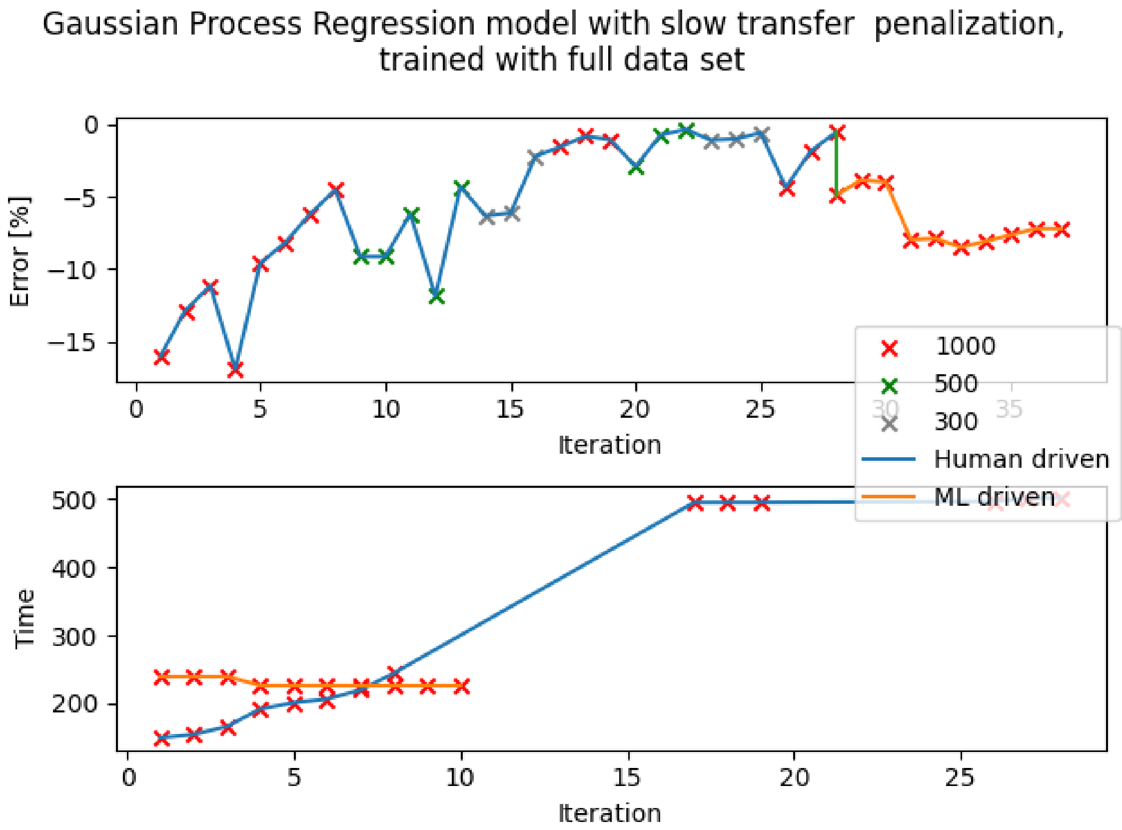
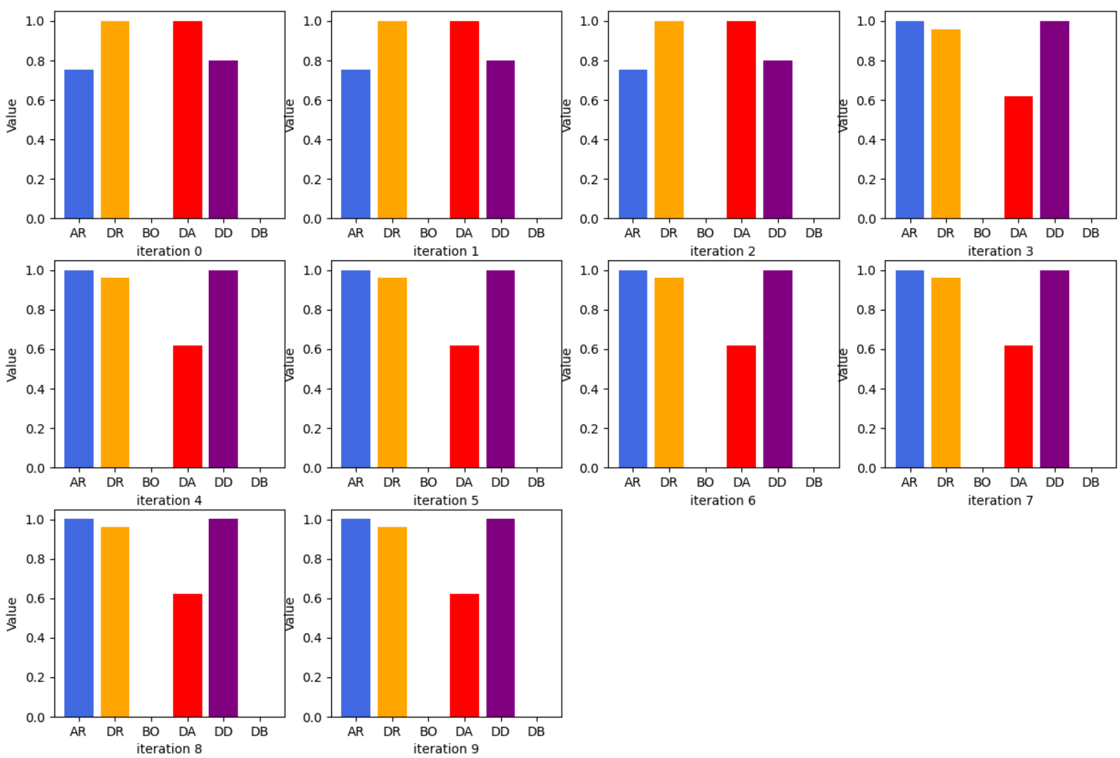
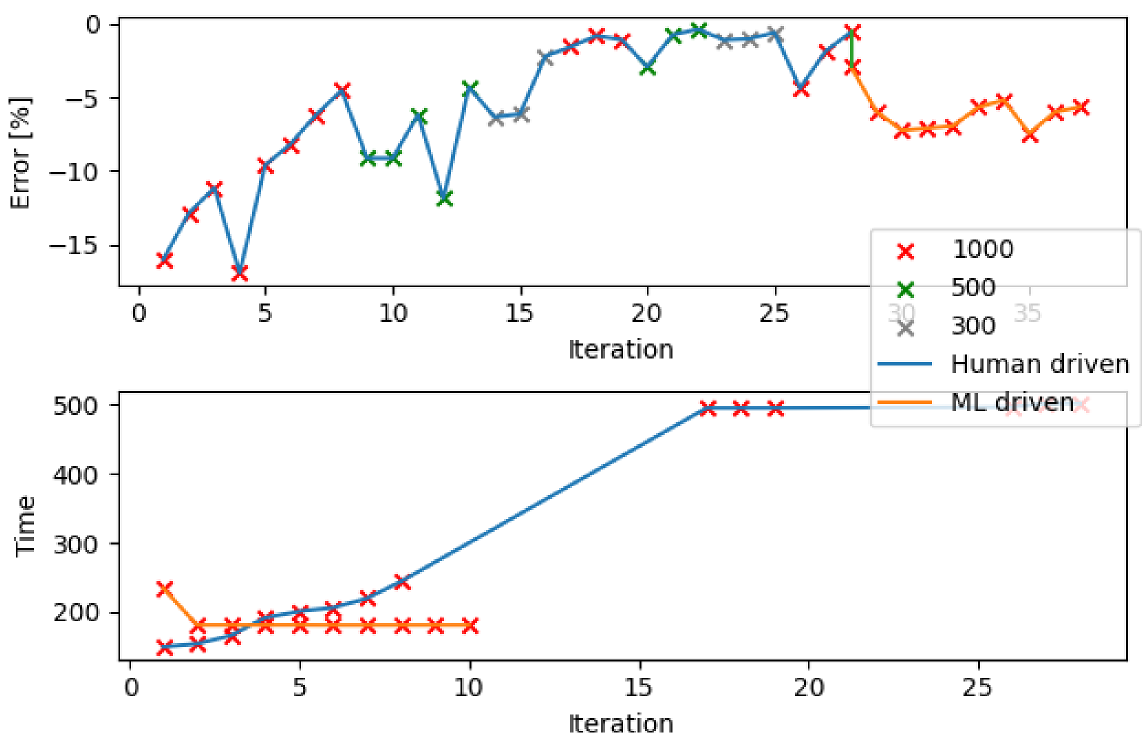


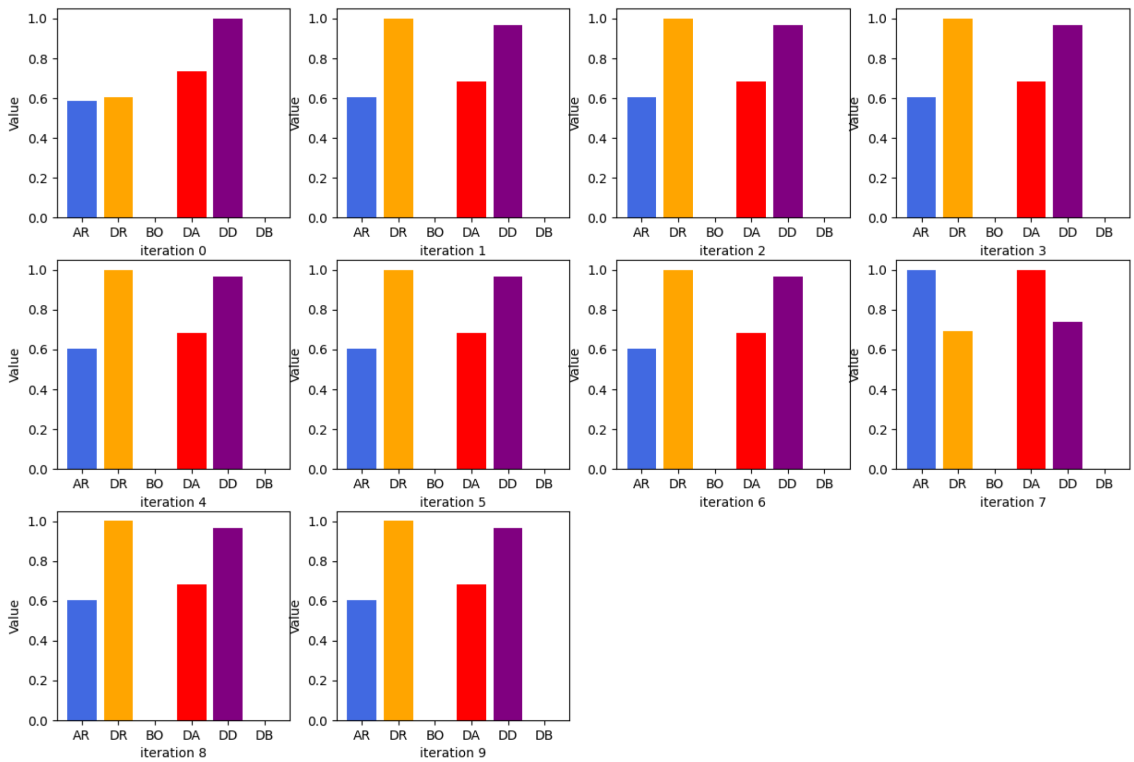
Diagram comparing human-driven and ML test trials (1275 cP, set:half)

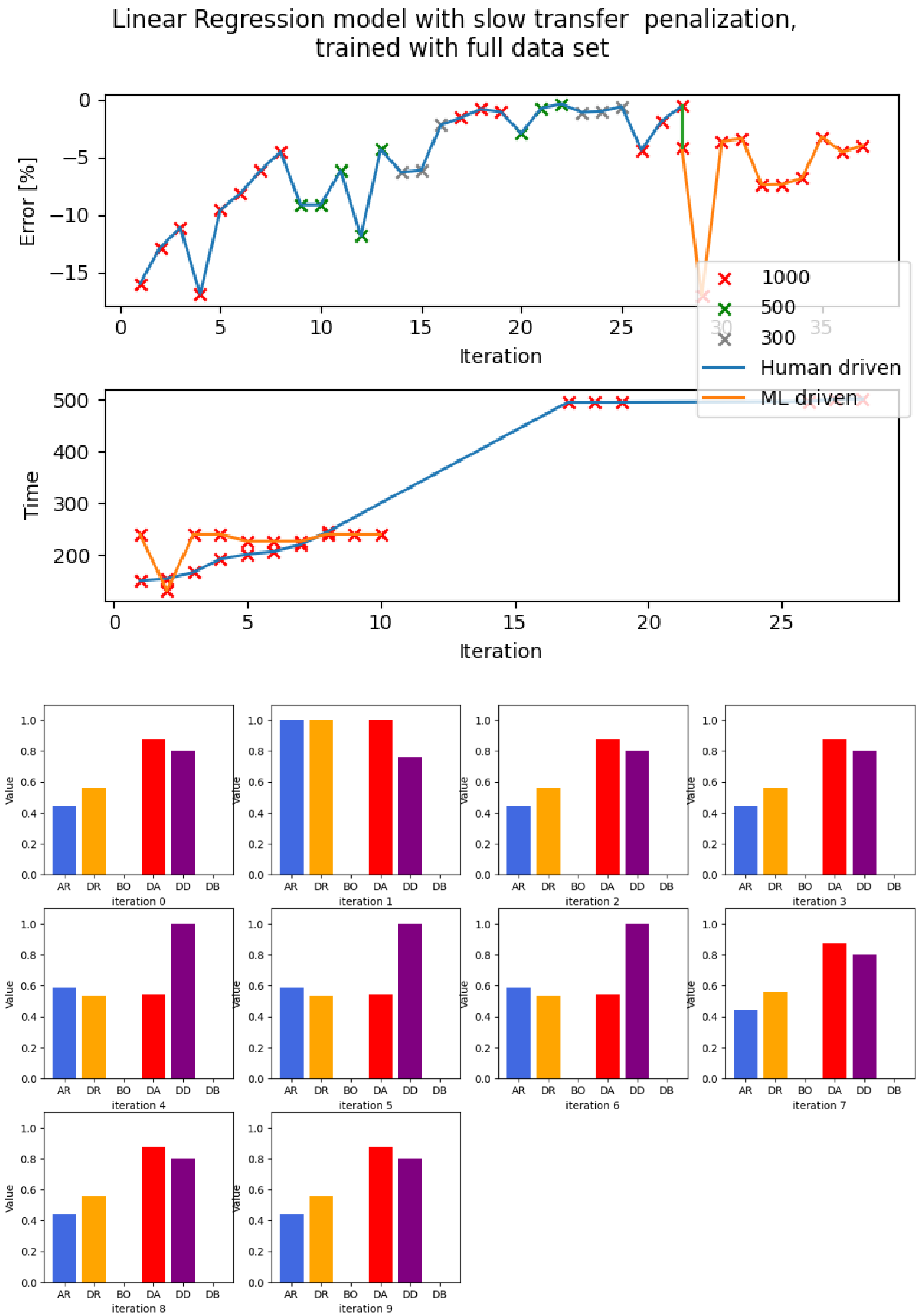




Gaussian Process Regression model with no penalization,
trained with full data set

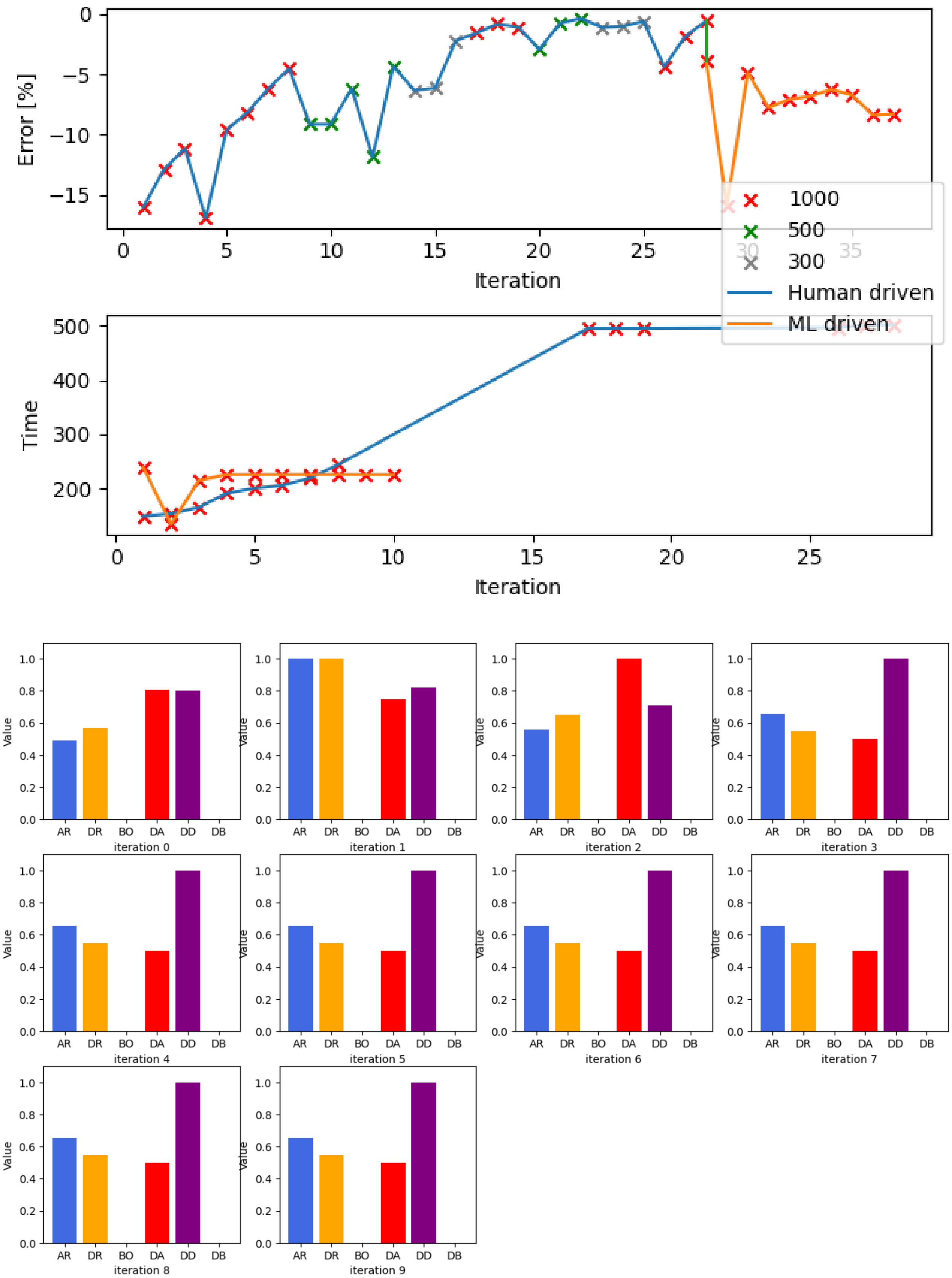






Linear Regression model with slow transfer penalization,
trained with full data set
Error [%]
Iteration
1000
500
300
Human driven
ML driven
Time
Iteration
AR
DR
BO
DA
DD
DB
Value
iteration 0
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
iteration 9

Linear Regression model with no penalization,
trained with full data set



Set: Half (unorderedT)

<u>Observation of trends</u>

Observation for LIN - scaling: none

Some variations in the parameters generated but most trials share the same set of parameters: 15.3, 7.0, 2.6, 4.5, with a percentage error of -6 to -7%.

Observation for LIN - scaling: multiply

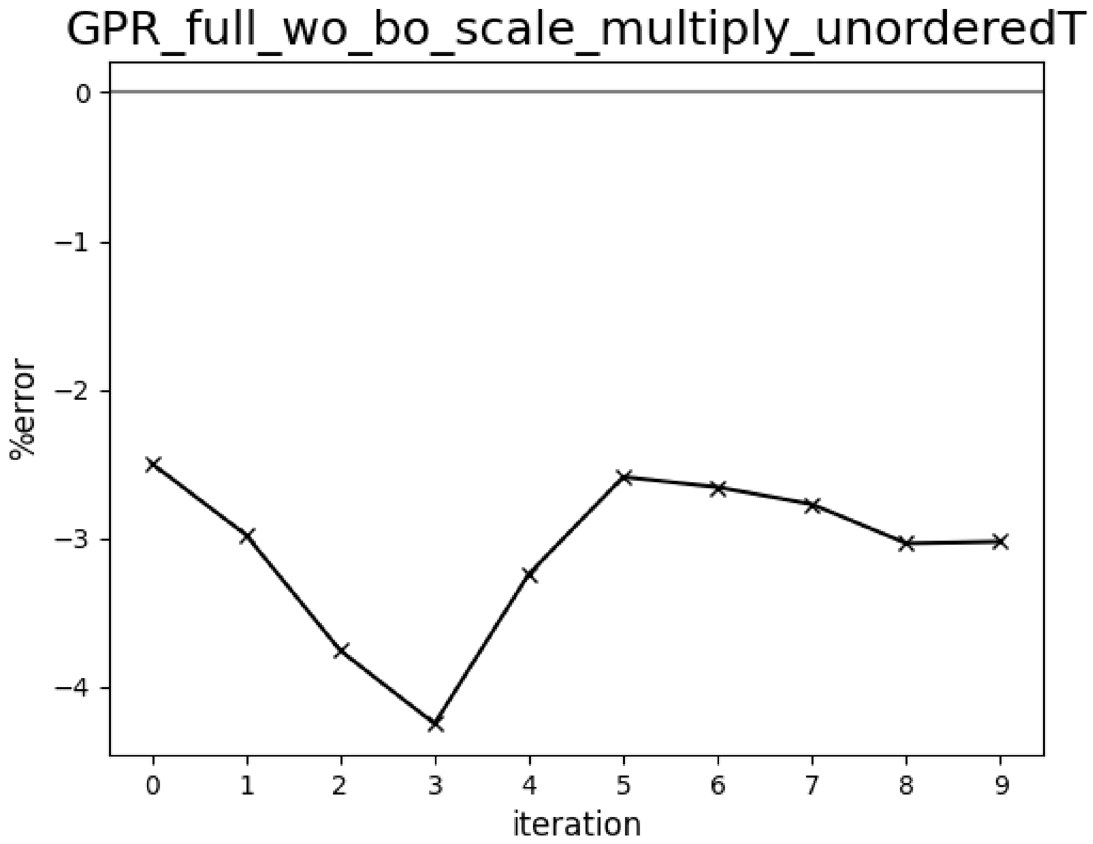
Some variations between the 2 sets of parameters: 11.5, 7.3, 4.2, 3.6 (percentage error around -3 to -4%) and 15.3, 7.0, 2.6, 4.5.

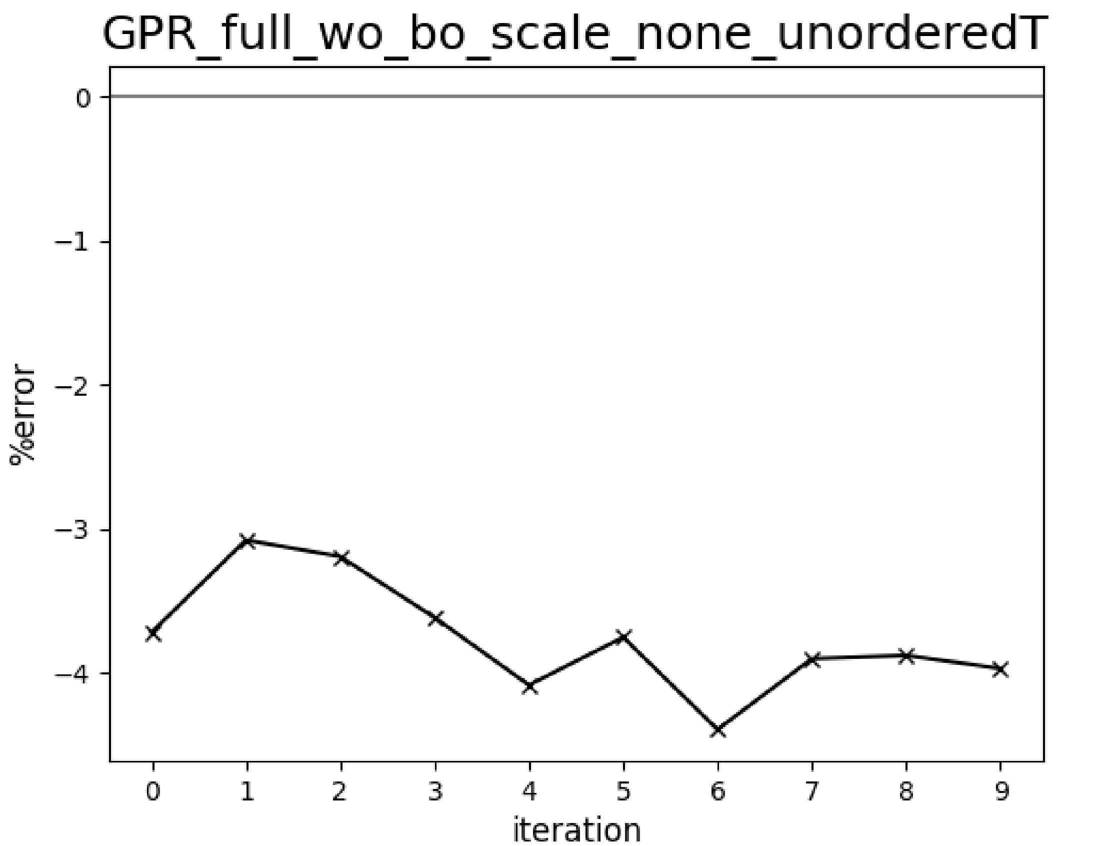
Observations for GPR - scaling: none

Very slight variations in parameters, most test trials share the same parameters: 11.1, 13.7, 4.1, 5.5.

Observations for GPR - scaling: multiply

Some variations between the 2 sets of parameters: 11.5, 7.3, 4.2, 3.6 (percentage error around -3 to -4%) and 15.3, 7.0, 2.6, 4.5.


GPR_full_wo_bo_scale_multiply_unorderedT
0
−1
−2
−3
−4
%error
0
1
2
3
4
5
6
7
8
9
iteration


GPR_full_wo_bo_scale_none_unorderedT
0
−1
−2
−3
−4
%error
0
1
2
3
4
5
6
7
8
9
iteration


LIN_full_wo_bo_scale_multiply_unorderedT
0
−1
−2
−3
−4
%error
0 1 2 3 4 5 6 7 8 9
iteration

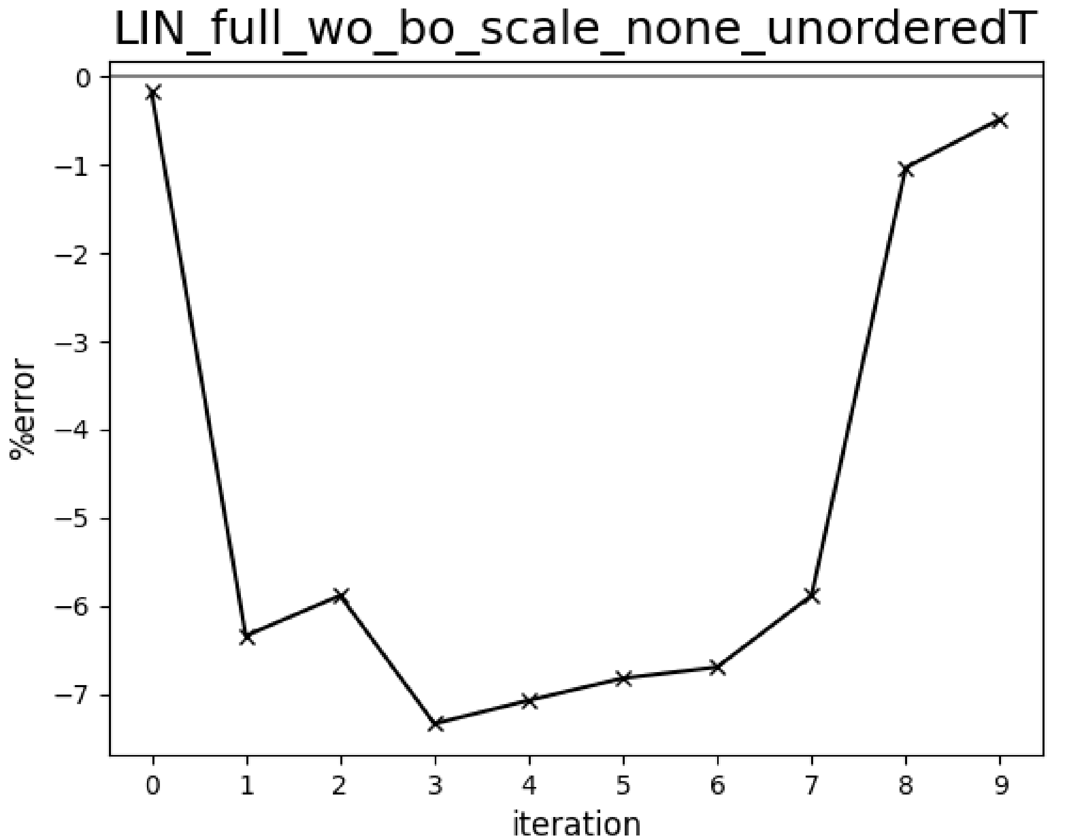
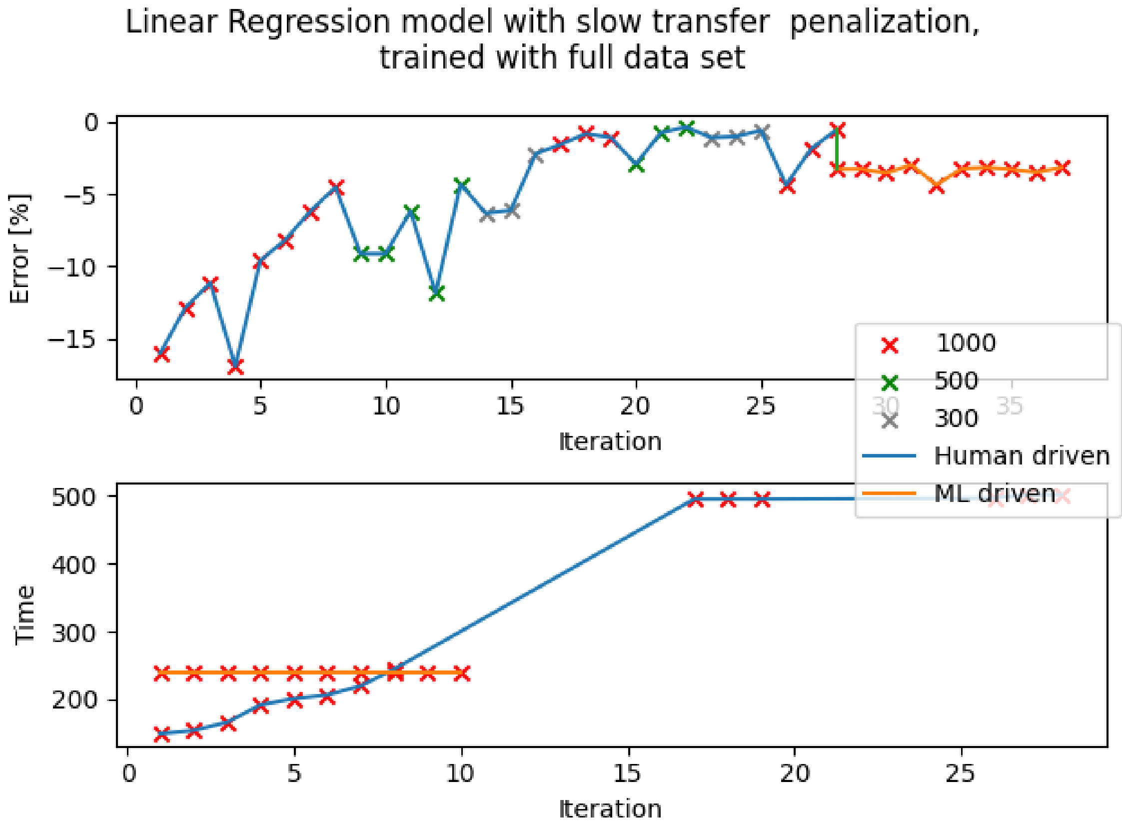
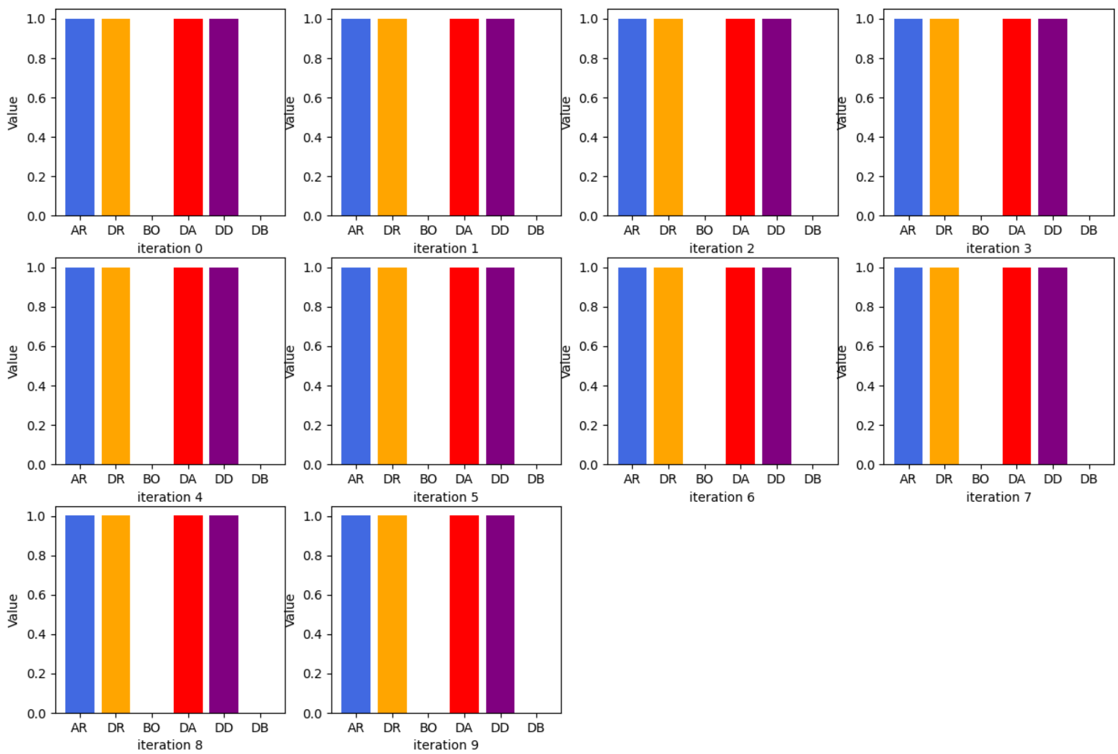
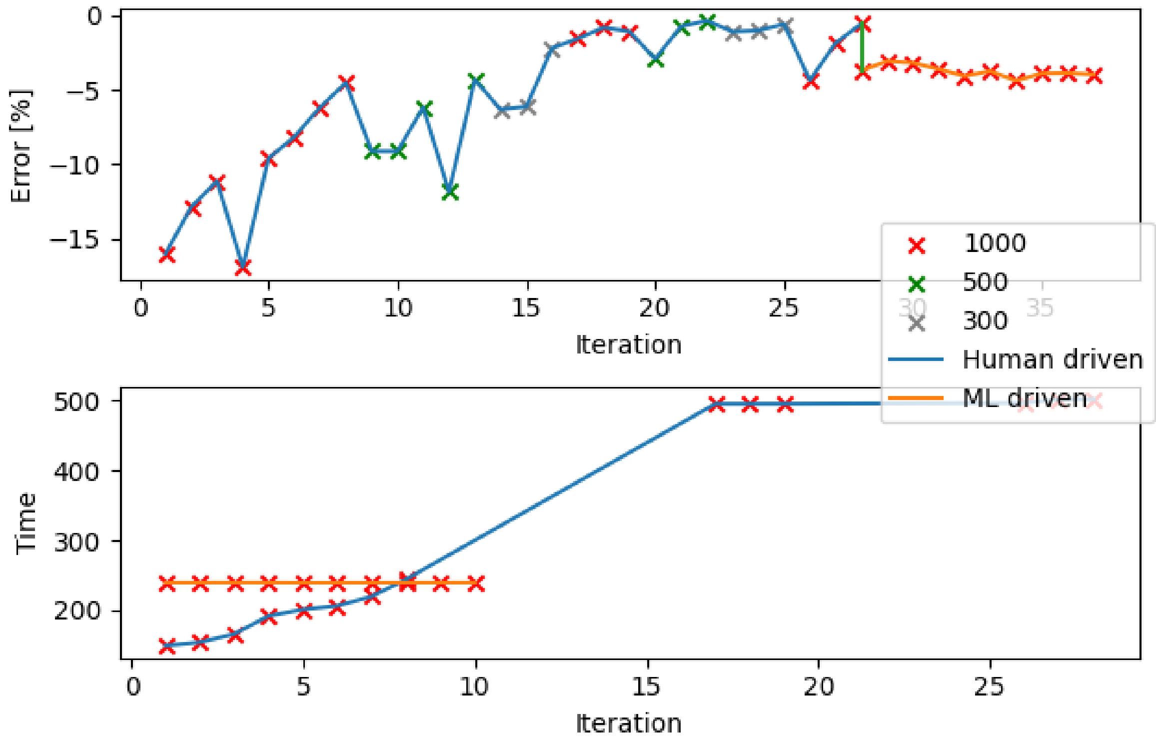

LIN_full_wo_bo_scale_none_unorderedT
0
−1
−2
−3
−4
−5
−6
−7
%error
0 1 2 3 4 5 6 7 8 9
iteration

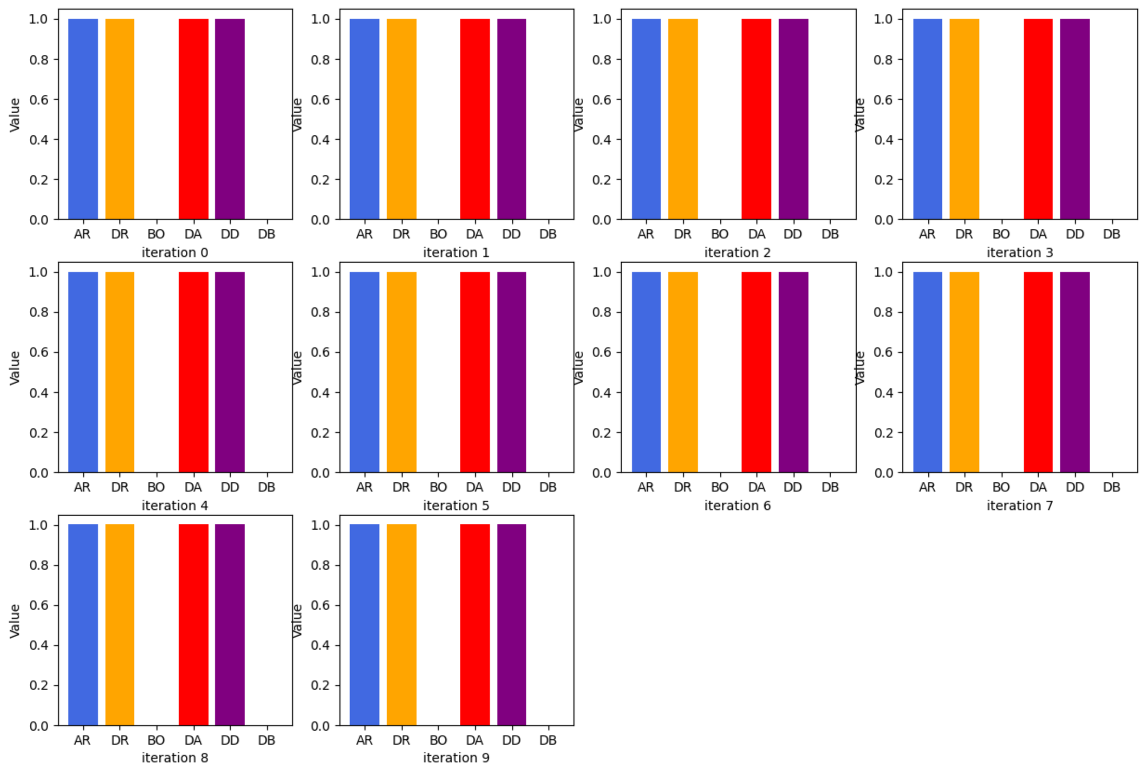
Diagram comparing human-driven and ML test trials (1275 cP, set:half, unorderedT)

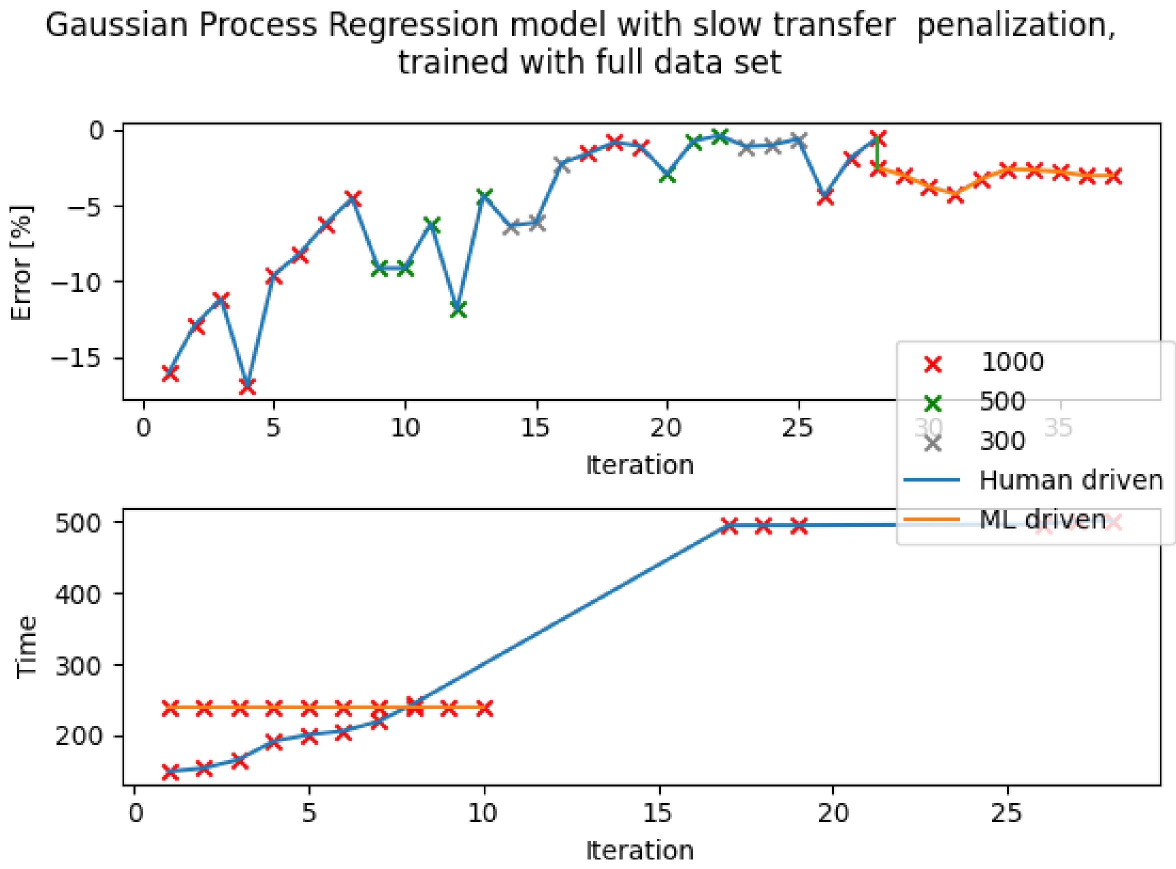


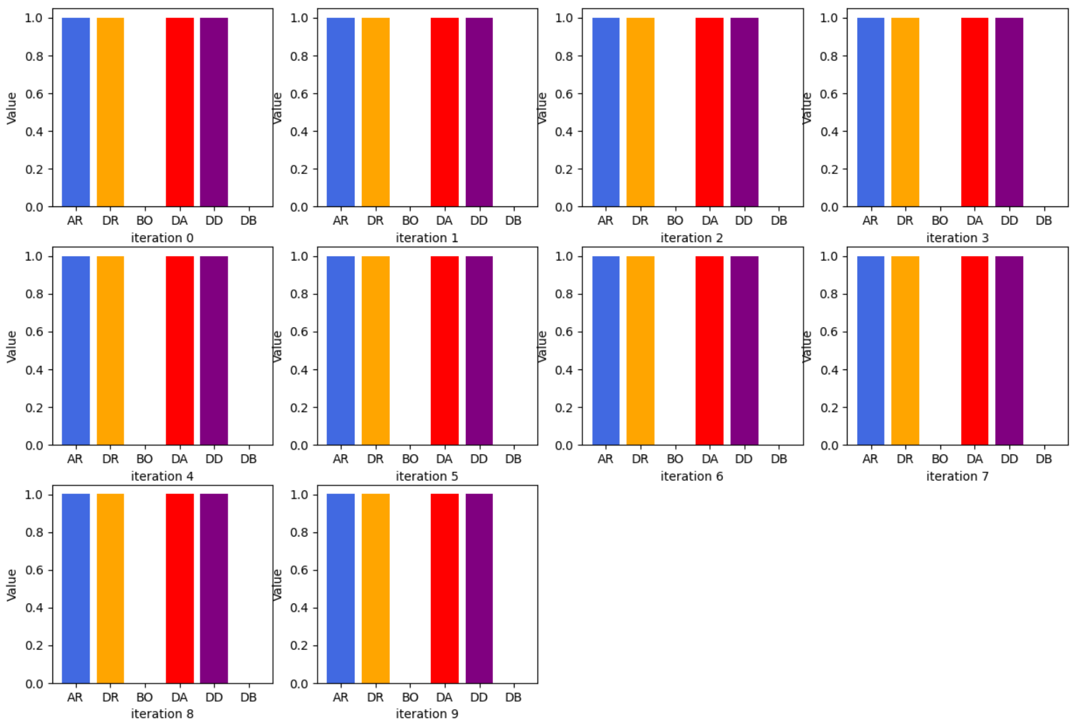


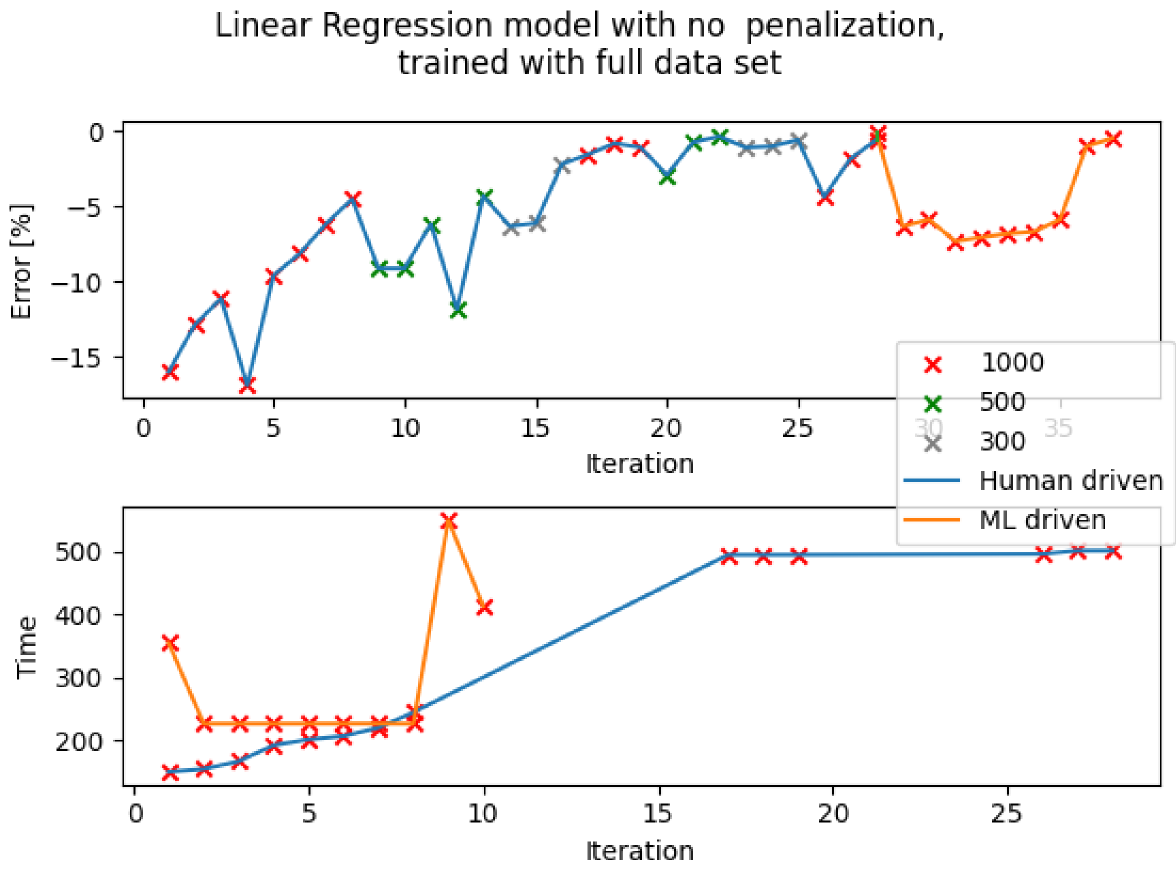
Gaussian Process Regression model with no penalization, trained with full data set

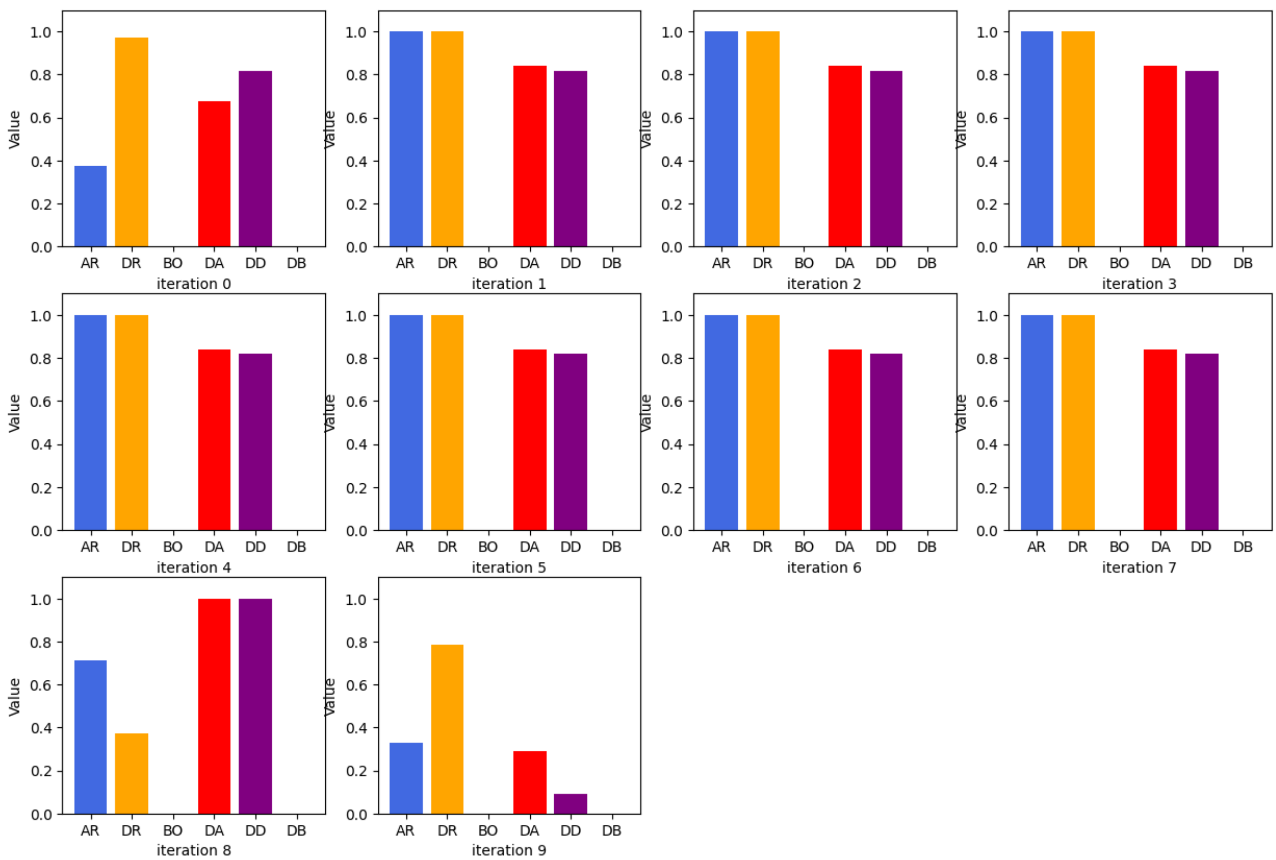





Gaussian Process Regression model with slow transfer penalization,
trained with full data set
Error [%]
Iteration
1000
500
300
Human driven
ML driven
Time
Iteration


Value
AR DR BO DA DD DB
iteration 0
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
iteration 9


Linear Regression model with no penalization,
trained with full data set
Error [%]
Iteration
1000
500
300
Human driven
ML driven
Time
Iteration


Value
AR
DR
BO
DA
DD
DB
iteration 0
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
iteration 9

Set: 1

Observation of trends

Observation for all iterations:

The parameters generated are more focused on quick transfer times than percentage error. Even though the transfer process will take around 160 seconds as derived from the standard calibration documents, most transfers only took between 90 to 140 seconds, with an average of 100 seconds.

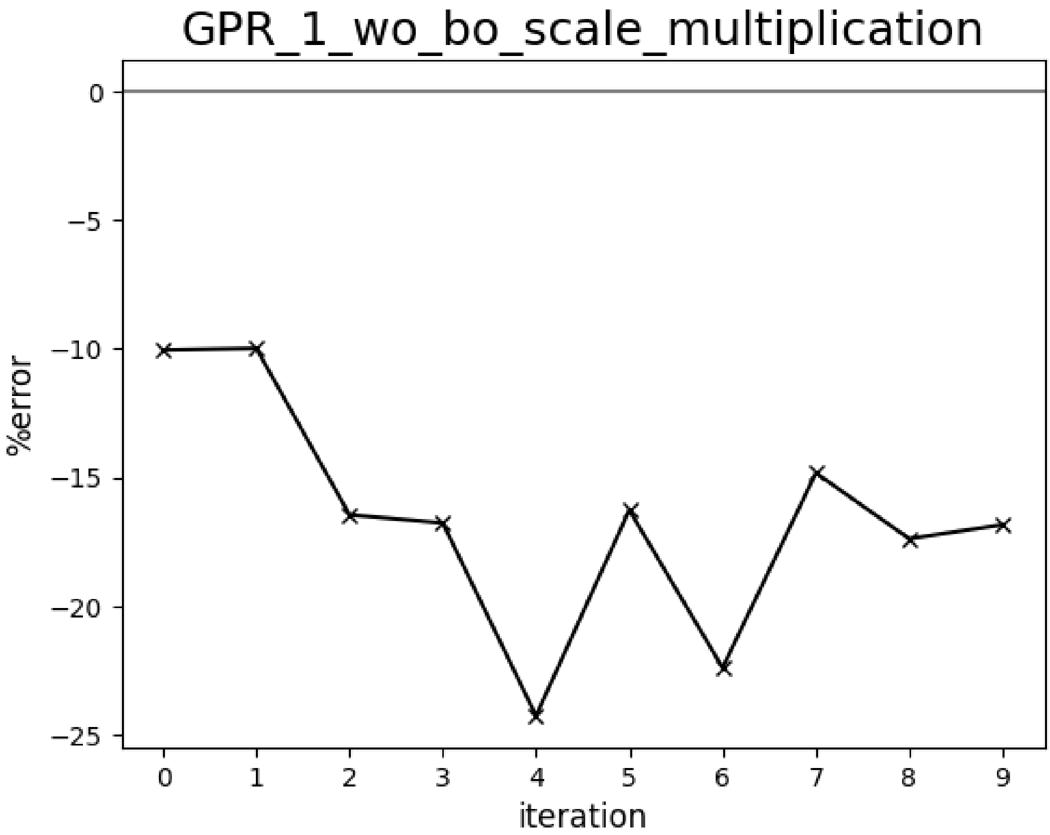
Suggestion for improvement:

When I was doing the test trials, it was discovered that the first reading acquired will not be accurate. This is because the pipette tip has not been used and it is harder to aspirate and dispense the liquid as compared to one that is already “rinsed” with the standard. Since this is only discovered in the later part of this project, I did not manage to make amendments to the csv when I am still finding the standard calibrations. Thus, by using the first reading as the only point of reference, it might feed the system with the wrong information.

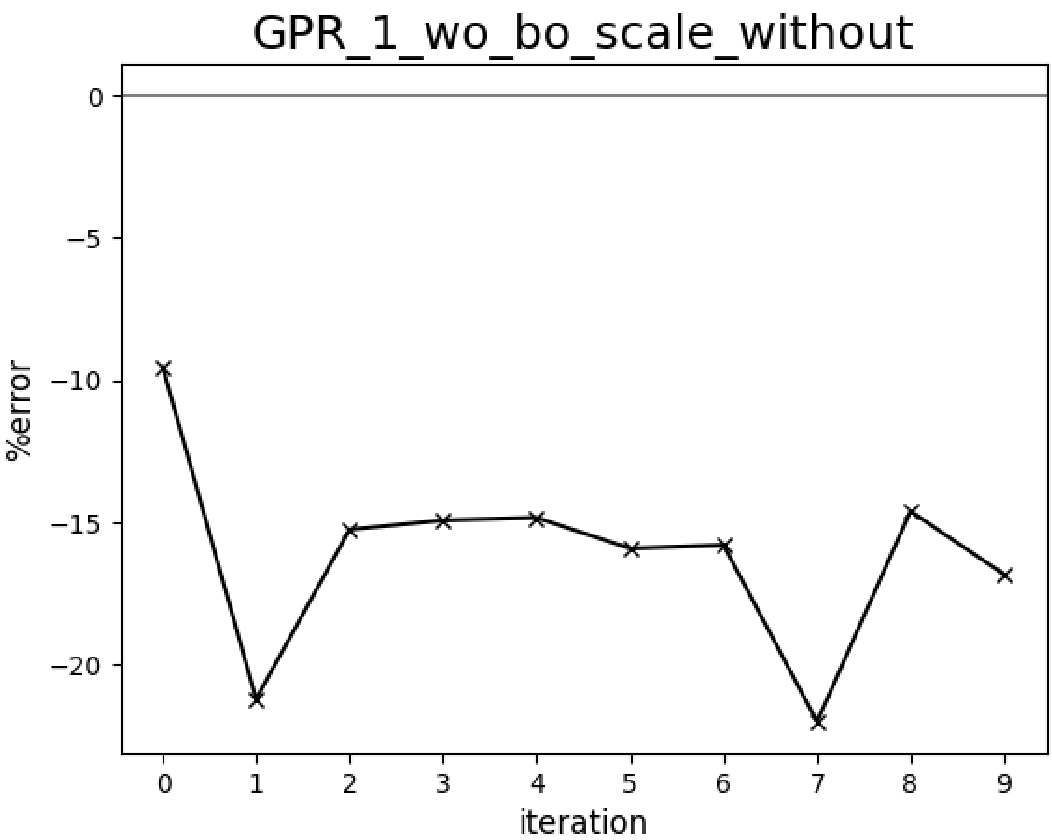
It is also worth noting that most trials have a percentage error of around -8% which is around the same error percentage that is caused by the “clean pipette error”.

One can try to use the second set of data collected instead. Although I’m not sure if it will work because even the trials with the half set of data as the reference points did not give very good parameters, but maybe the first data is prioritised (?) by the algorithm.

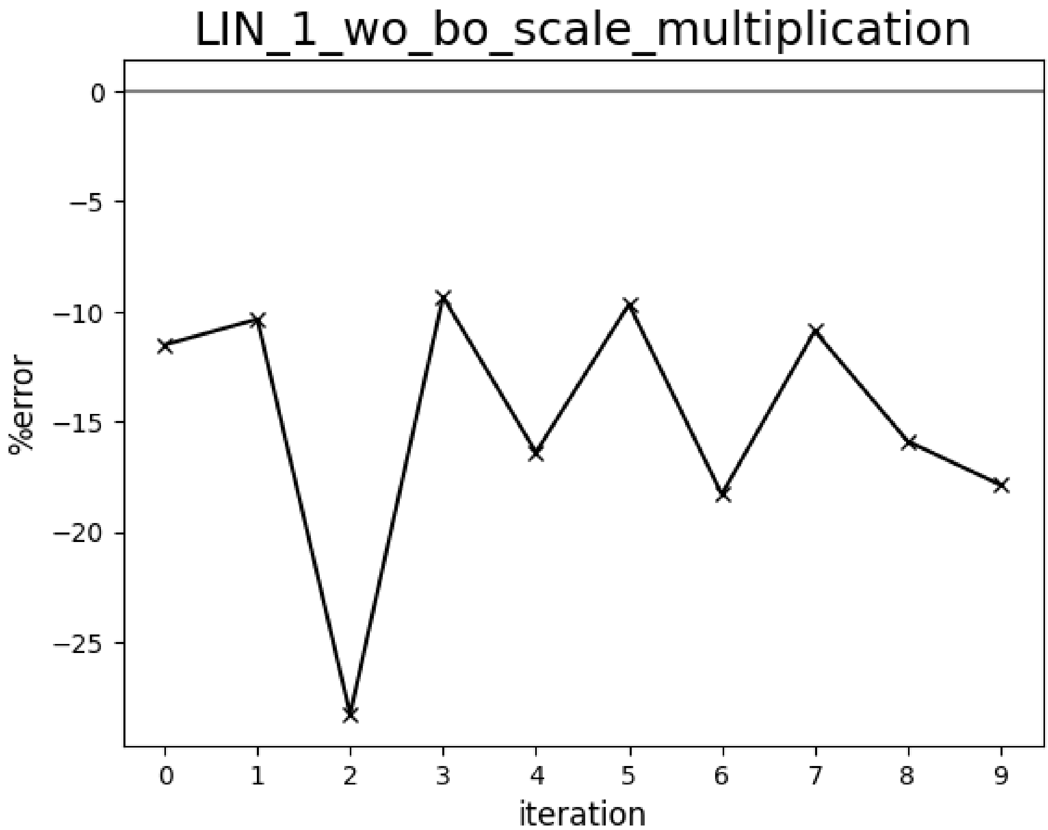
GPR scaling: multiplication



GPR scaling: without



LIN scaling: multiplication



LIN scaling: without

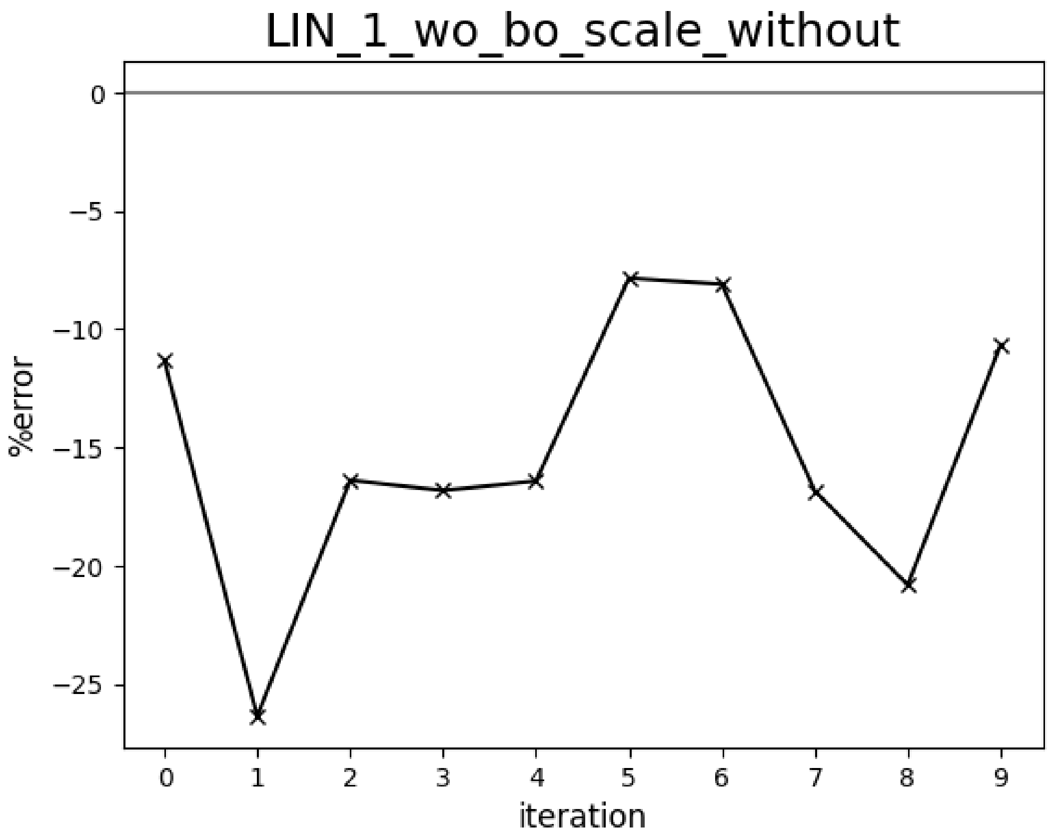
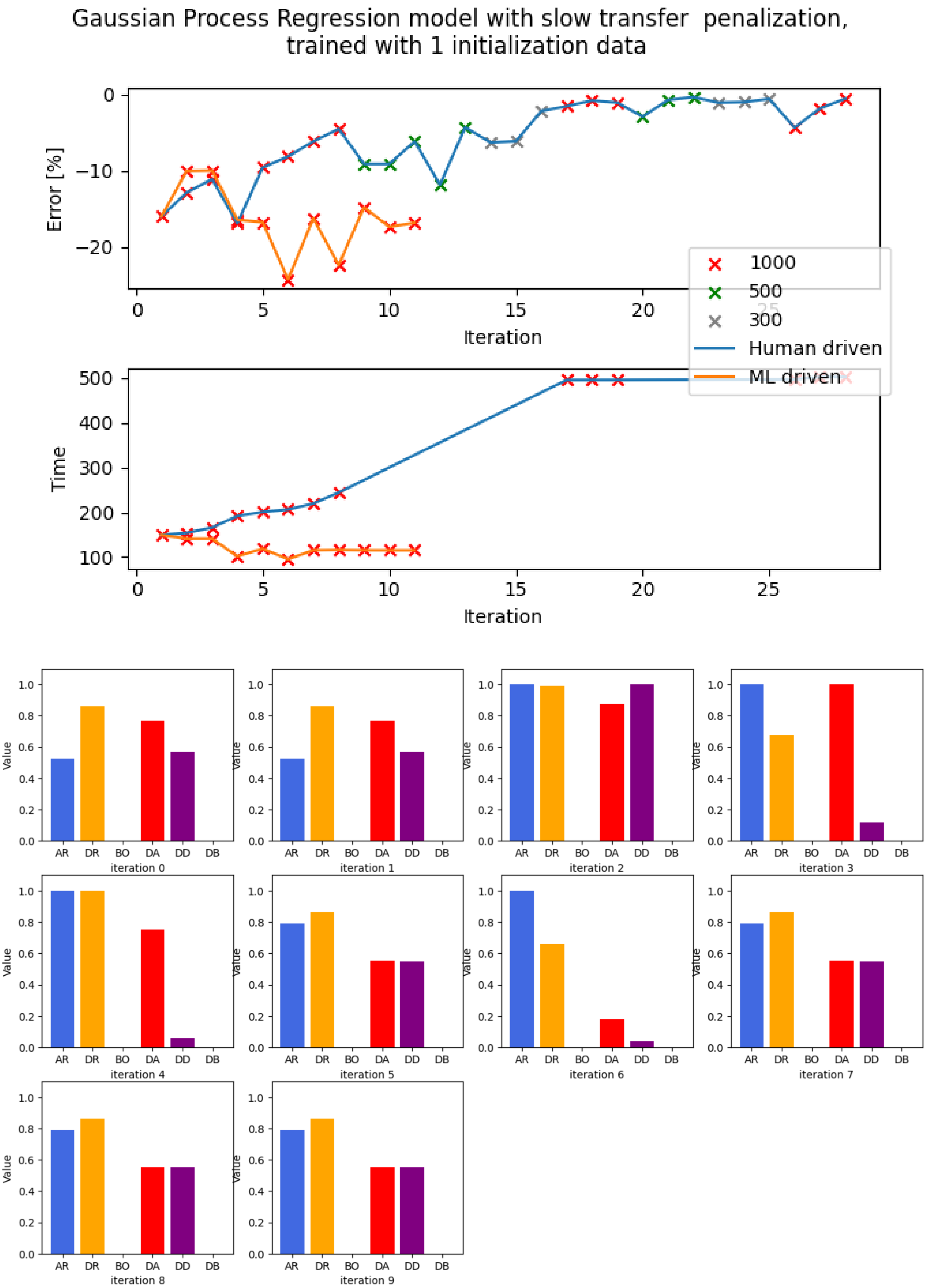
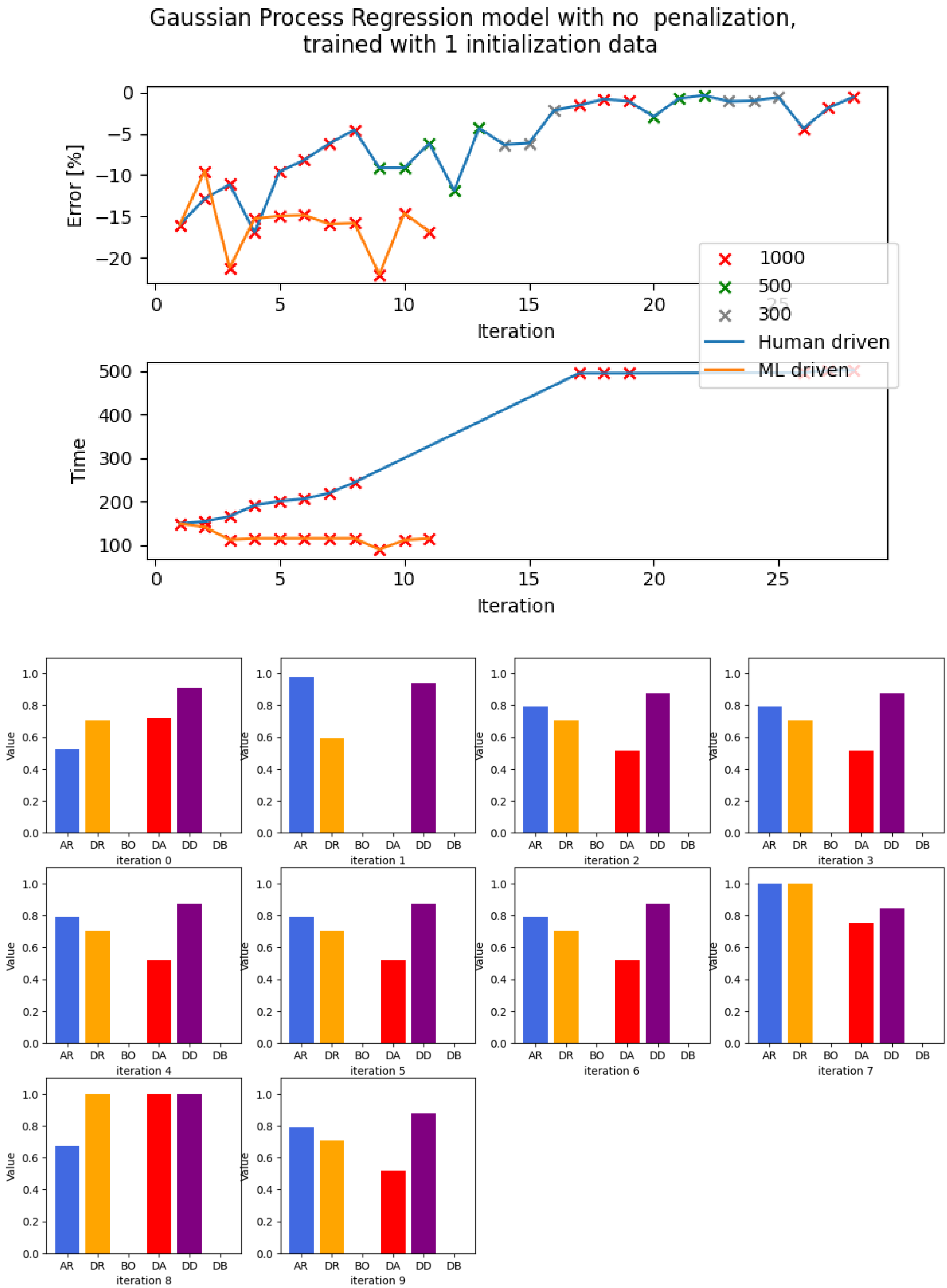


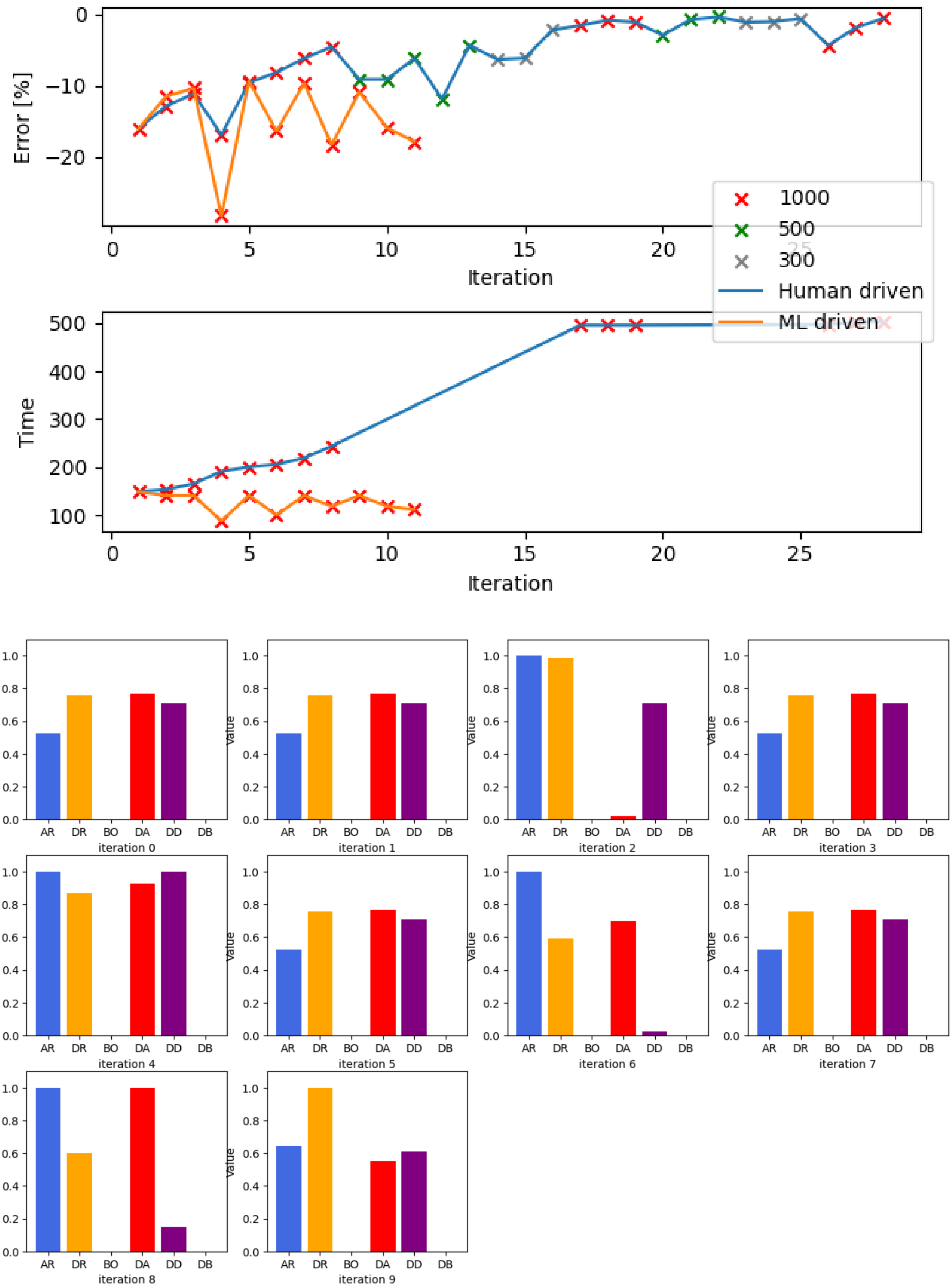
Diagram comparing human-driven and ML test trials (1275 cP, set:1)

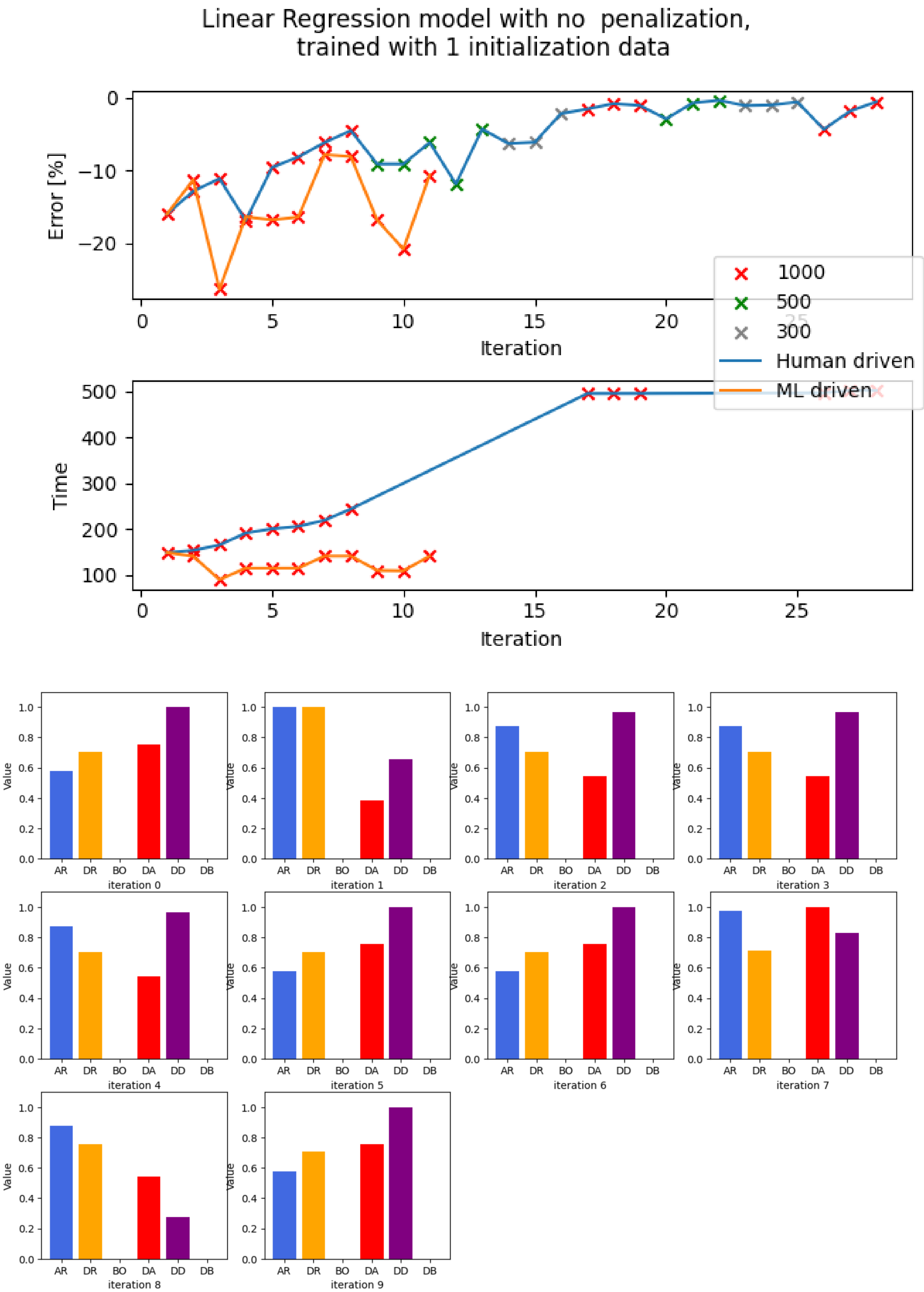




Gaussian Process Regression model with no penalization,
trained with 1 initialization data
Error [%]
Iteration
1000
500
300
Human driven
ML driven
Time
Iteration
Value
AR
DR
BO
DA
DD
DB
iteration 0
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
iteration 9

Linear Regression model with slow transfer penalization,
trained with 1 initialization data





Linear Regression model with no penalization,
trained with 1 initialization data
Error [%]
Iteration
1000
500
300
Human driven
ML driven
Time
Iteration
Value
AR DR BO DA DD DB
iteration 0
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
iteration 9

Set: 1 (unorderedT)

<u>Observation of trends</u>

Observation for LIN - scaling: none

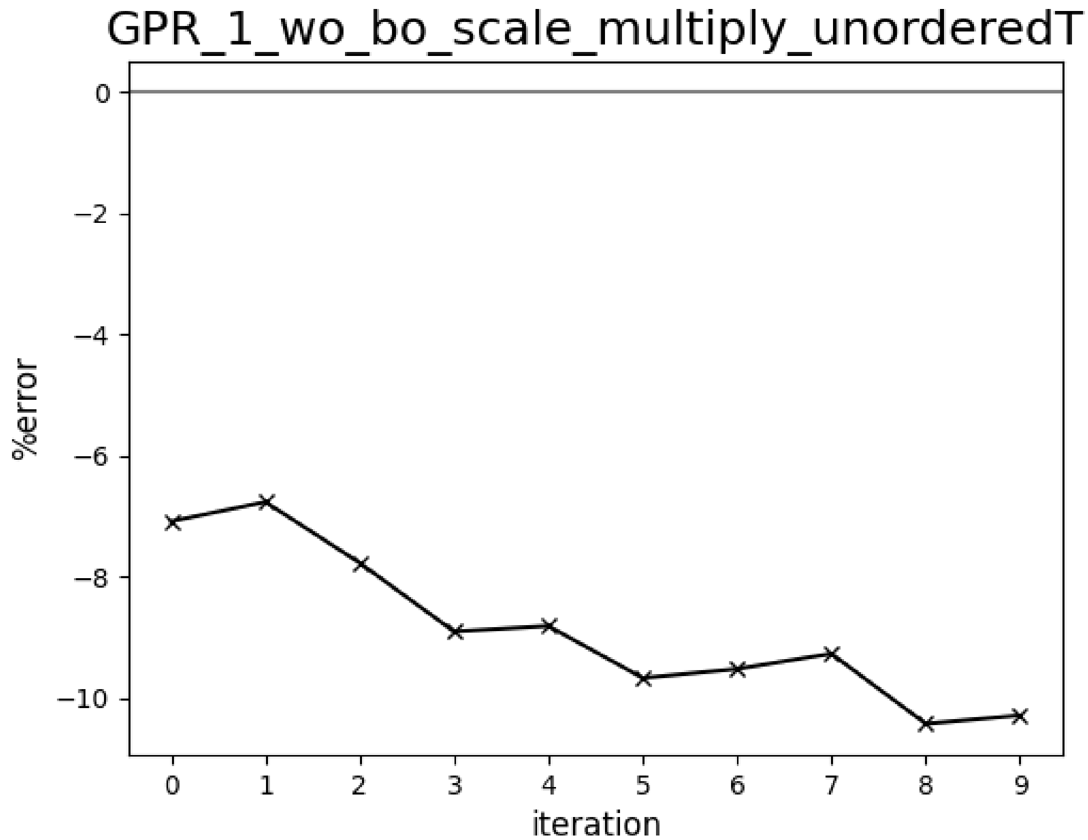
Mostly between these 2 sets of parameters: 13.8, 18.6, 4.3, 2.9 (with a percentage error of around -8 to -11%) and 20.9, 18.6, 3.1, 2.8 (with a percentage error of around -18 to -20%)

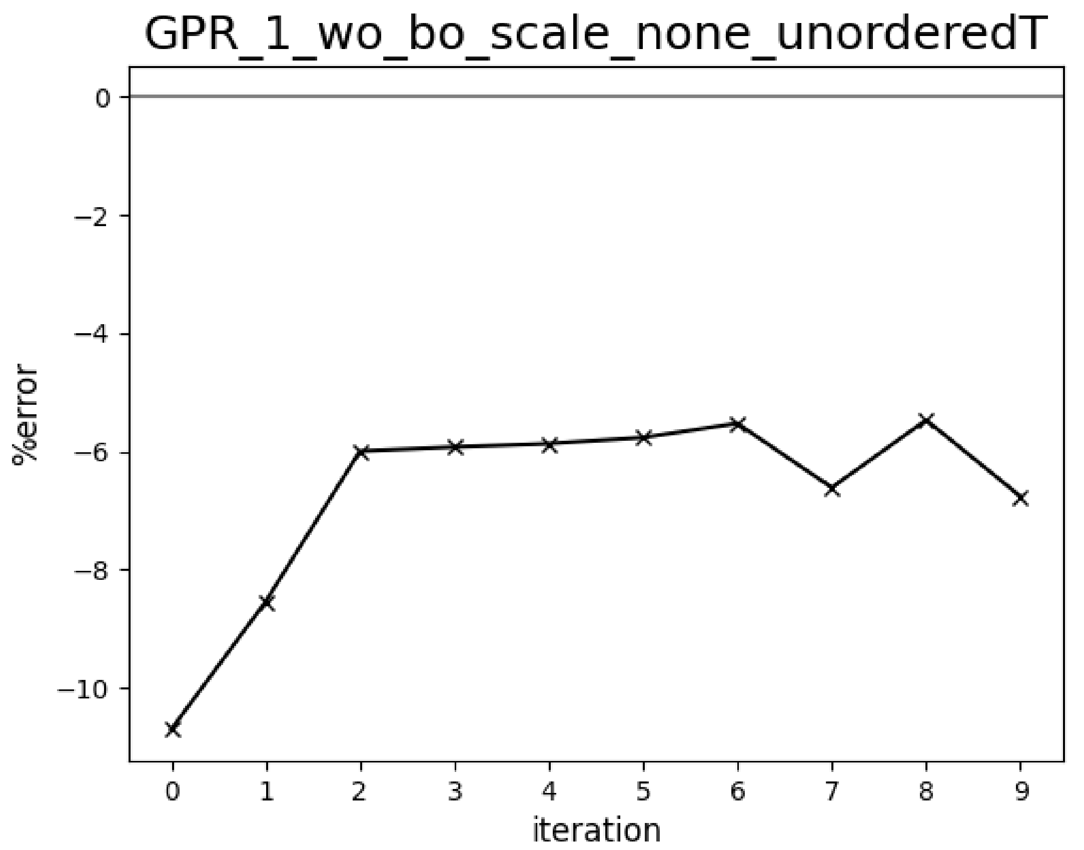
Observation for LIN - scaling: multiply and GPR - scaling: none

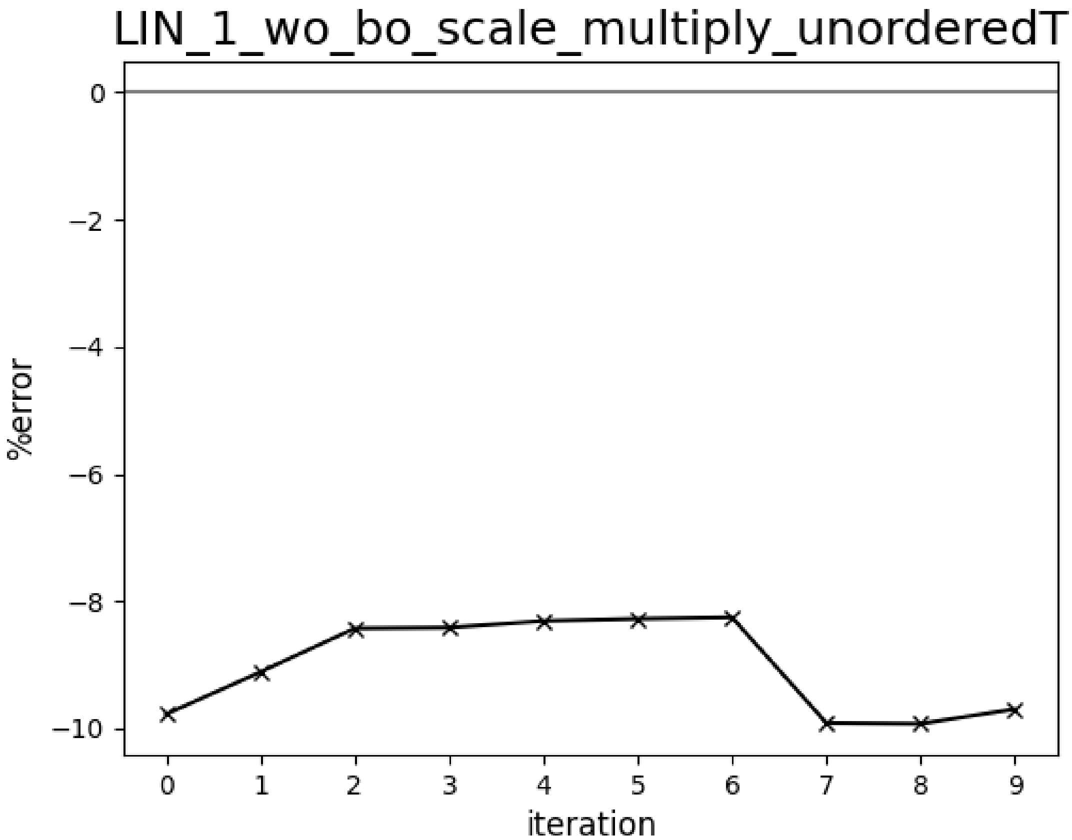
Most test trials use this set of parameters: 20.9, 18.6, 3.1, 2.8 and 26.4, 24.3, 0.1, 2.9, very inaccurate sets of parameters.

Observation for GPR - scaling: multiply

It has a blend of all 3 sets of parameters that are commonly used in set: 1 test trials.
13.8, 18.6, 4.3, 2.9 & 20.9, 18.6, 3.1, 2.8 & 26.4, 24.3, 0.1, 2.9


GPR_1_wo_bo_scale_multiply_unorderedT
0
−2
−4
−6
−8
−10
%error
0
1
2
3
4
5
6
7
8
9
iteration


GPR_1_wo_bo_scale_none_unorderedT
0
−2
−4
−6
−8
−10
%error
0
1
2
3
4
5
6
7
8
9
iteration


LIN_1_wo_bo_scale_multiply_unorderedT
%error
0
−2
−4
−6
−8
−10
0
1
2
3
4
5
6
7
8
9
iteration

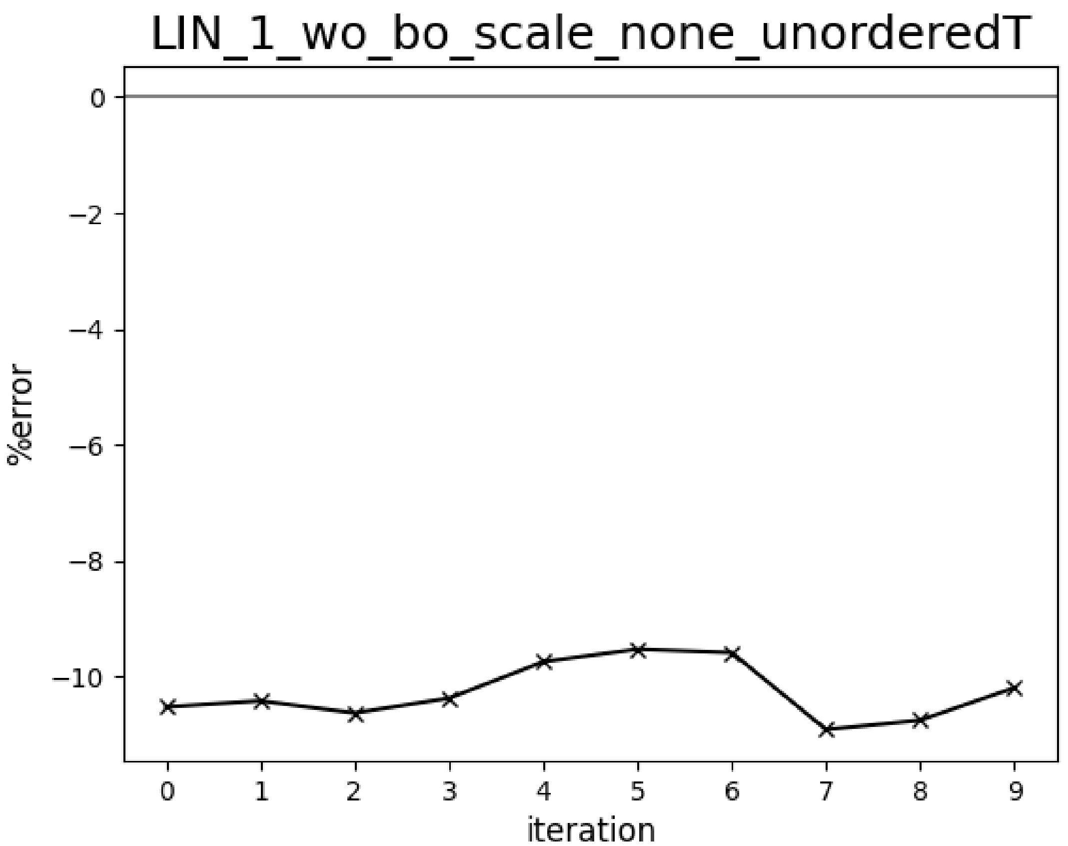
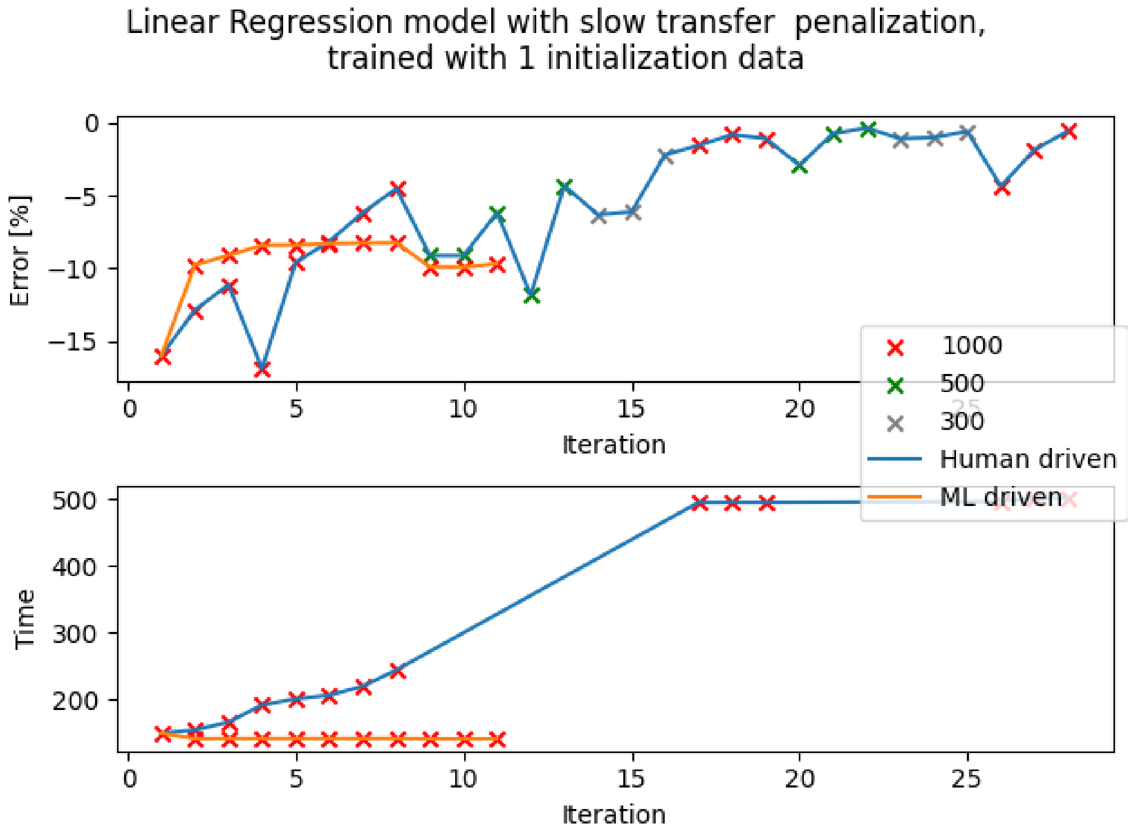
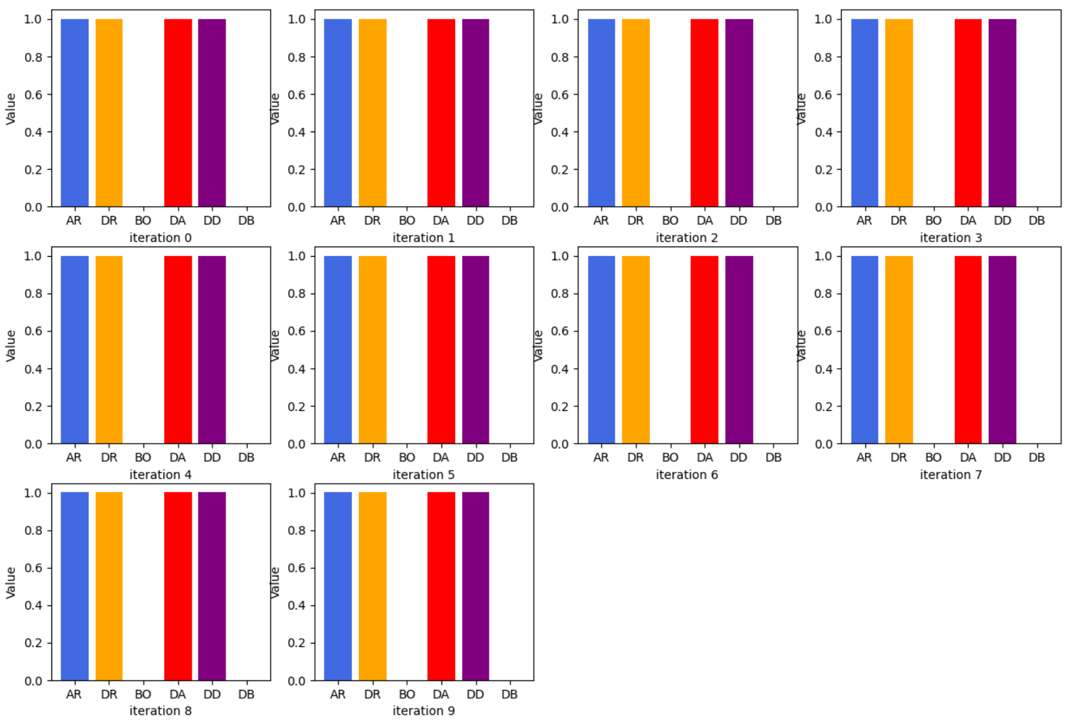
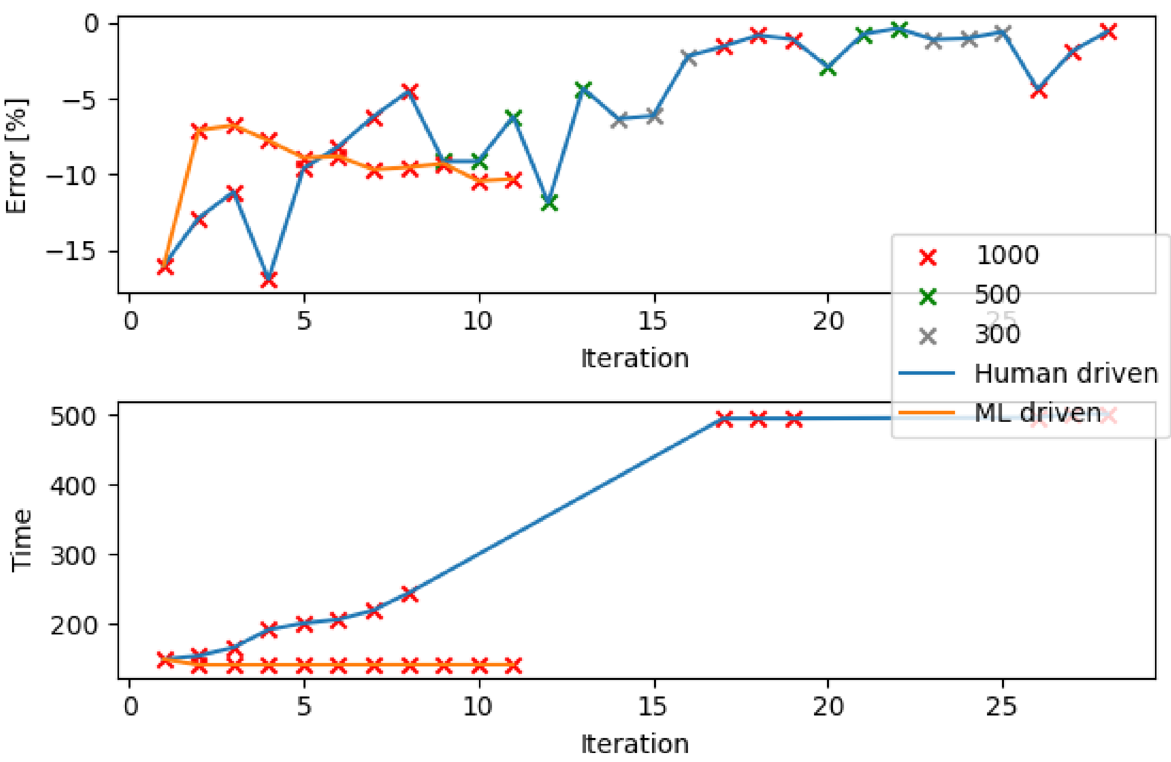

LIN_1_wo_bo_scale_none_unorderedT
%error
0
−2
−4
−6
−8
−10
0
1
2
3
4
5
6
7
8
9
iteration

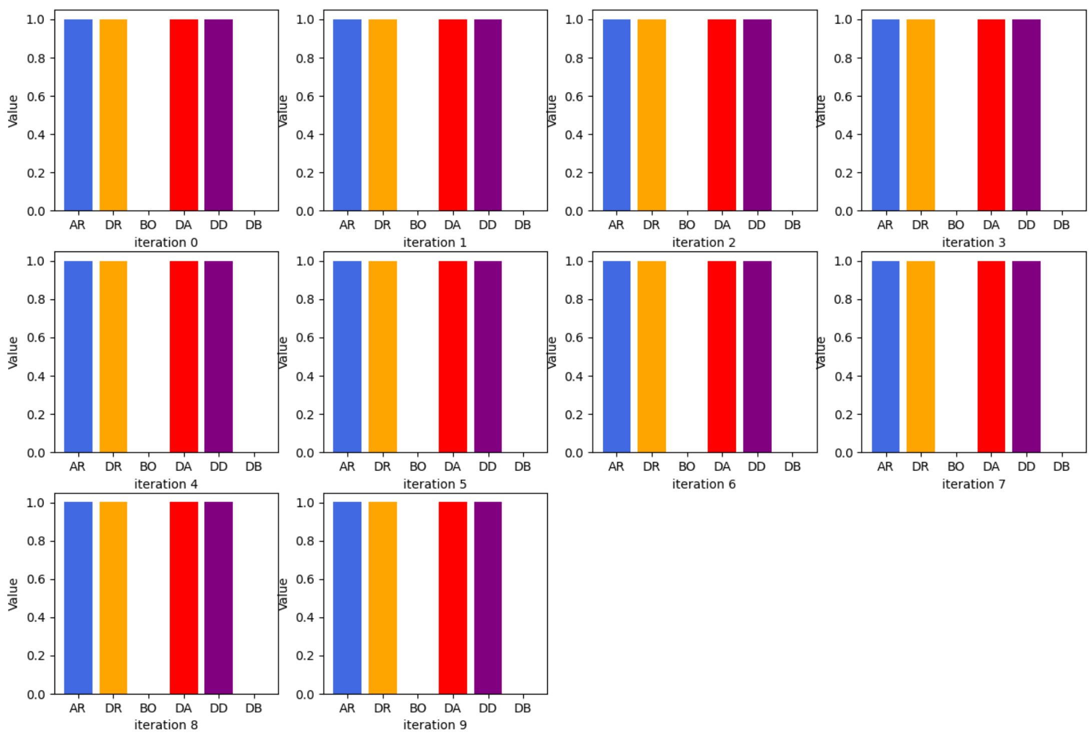
Diagram comparing human-driven and ML test trials (1275 cP, set:1, unorderedT)

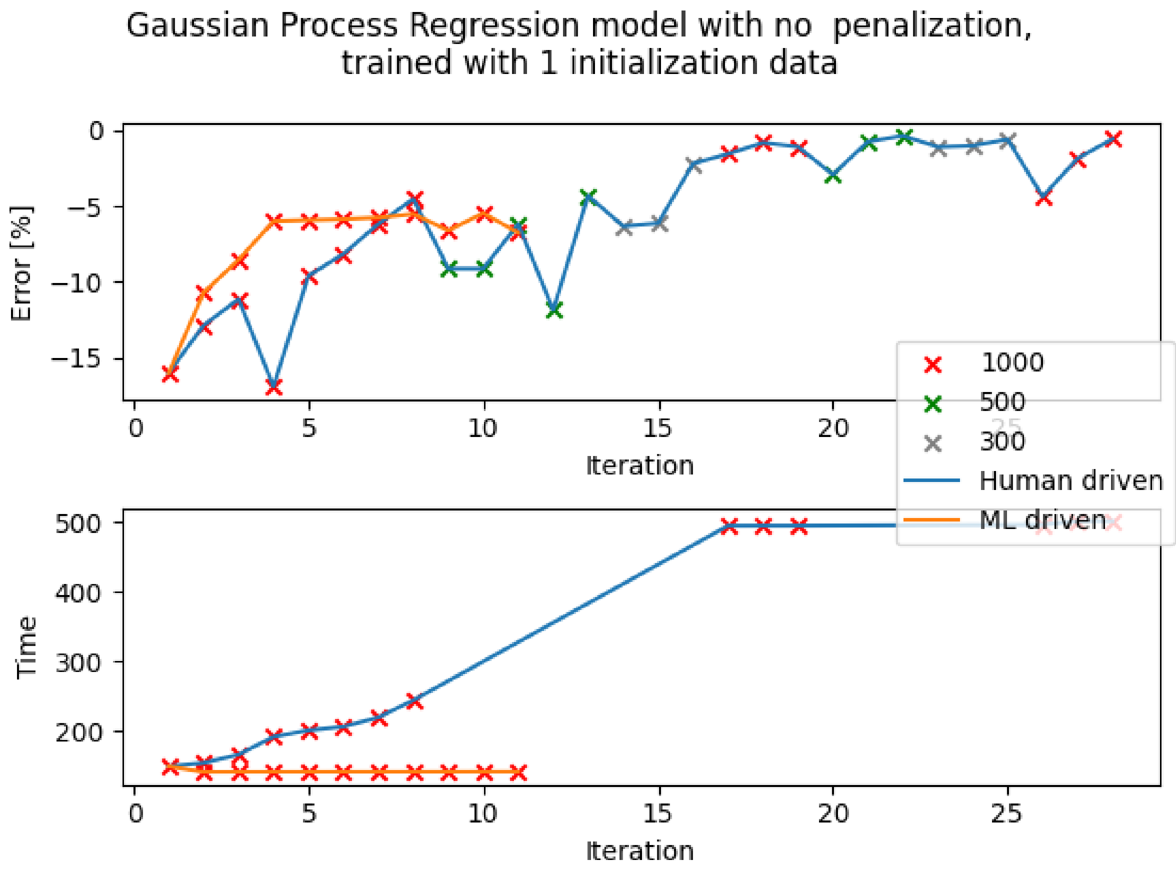




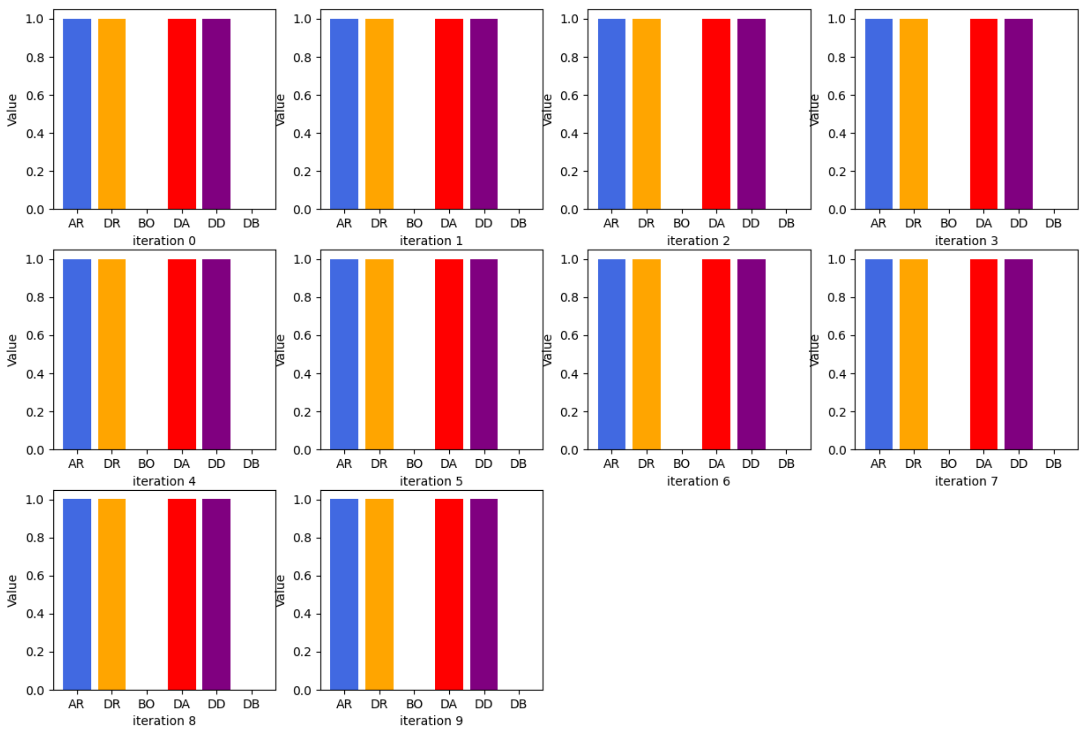
Gaussian Process Regression model with slow transfer penalization, trained with 1 initialization data



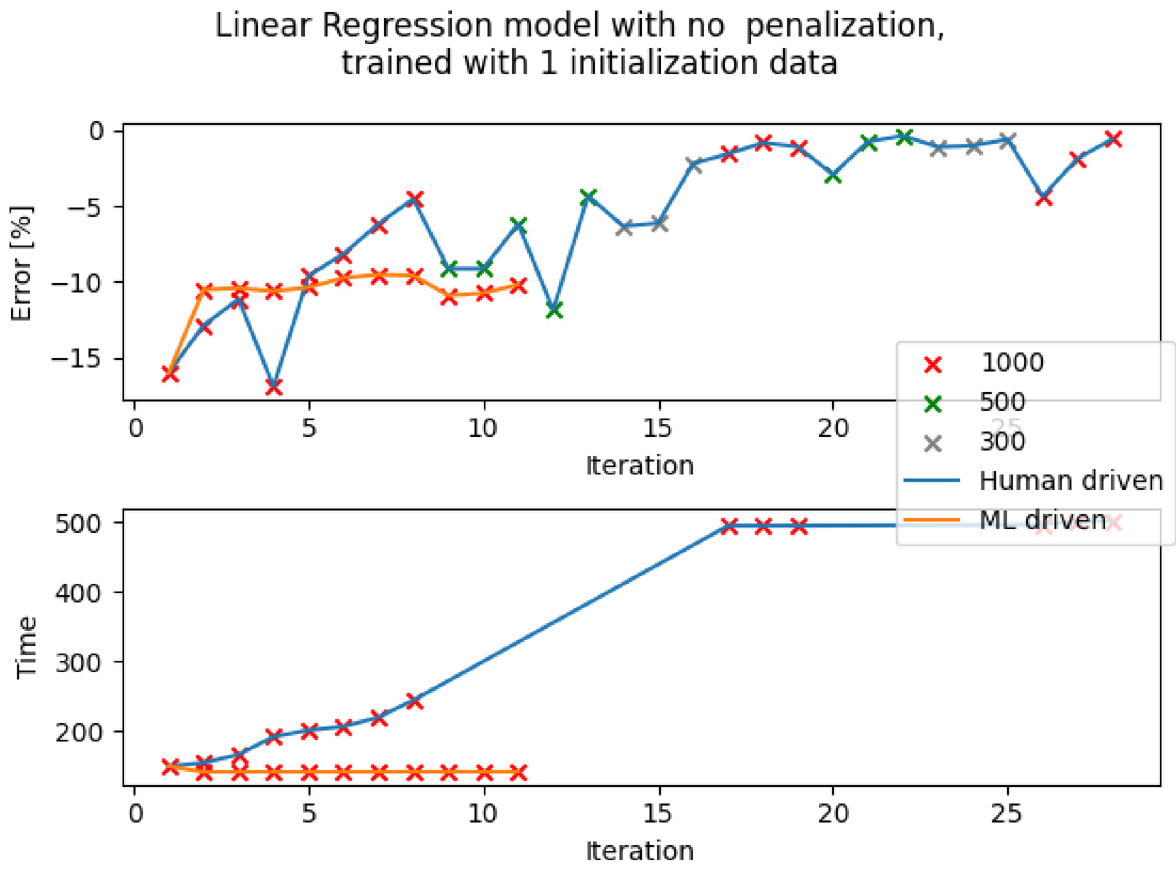


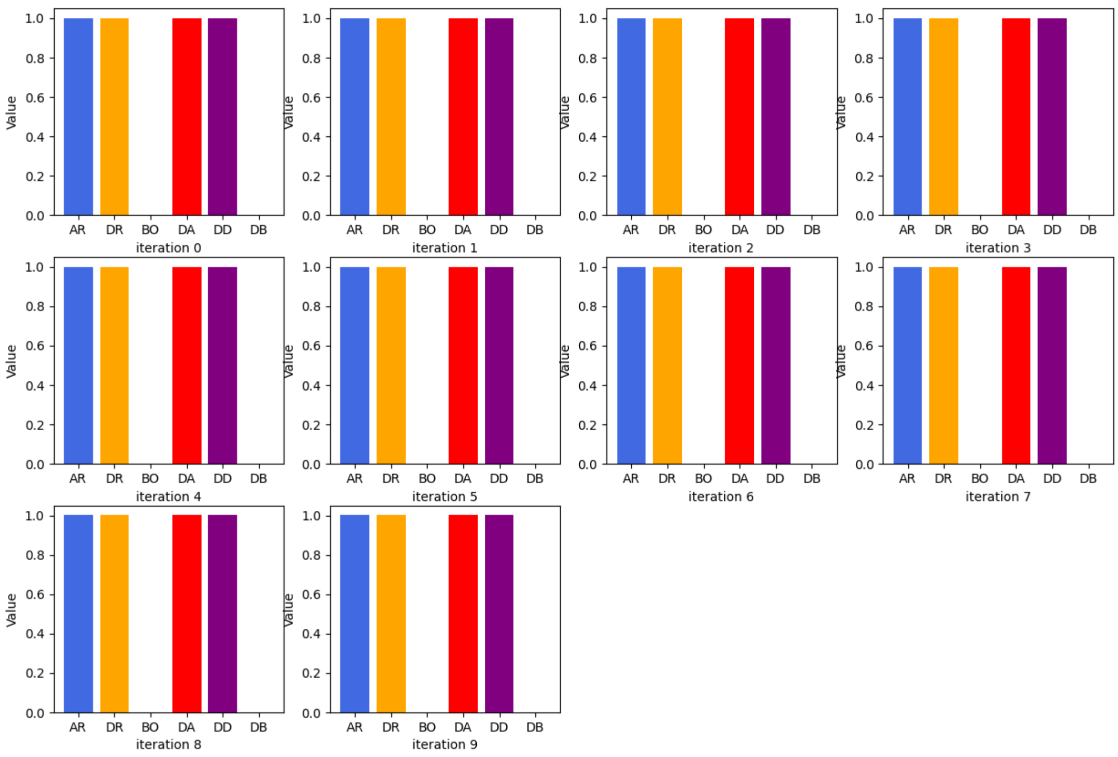


Gaussian Process Regression model with no penalization,
trained with 1 initialization data
Error [%]
0
−5
−10
−15
Iteration
1000
500
300
Human driven
ML driven
Time
500
400
300
200
0
5
10
15
20
25
Iteration



Value
AR DR BO DA DD DB
iteration 0
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
iteration 9


Linear Regression model with no penalization,
trained with 1 initialization data
Error [%]
Iteration
Time
Iteration
1000
500
300
Human driven
ML driven


Value
AR
DR
BO
DA
DD
DB
iteration 0
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
iteration 9

Set: 1 (ver4)

Observation of trends

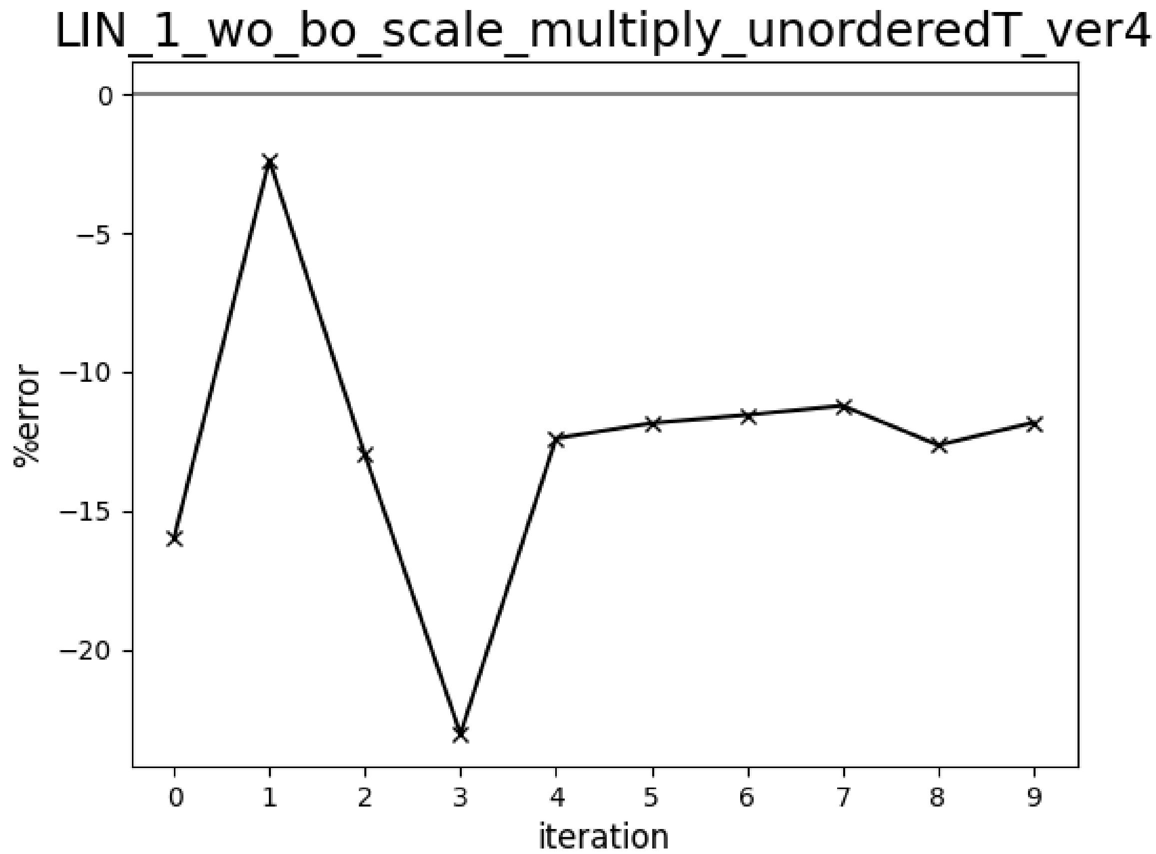
Observation 1: LIN - scaling: multiply

The parameters suggested with slow time penalization clearly prioritise fast transfer duration a lot more than the previous experiment with no time penalization (141s < 245s -> none). However, the percentage error is much higher than the preferred -2 to 2%. It falls between -11 to -12%.

Observation 2: LIN - scaling: none

The parameters suggested are much slower than the test trials with slow transfer penalization, however, the percentage error is slightly lower than LIN - scaling: multiply, around -7 to -8% on average.

Percentage error against iteration diagram:



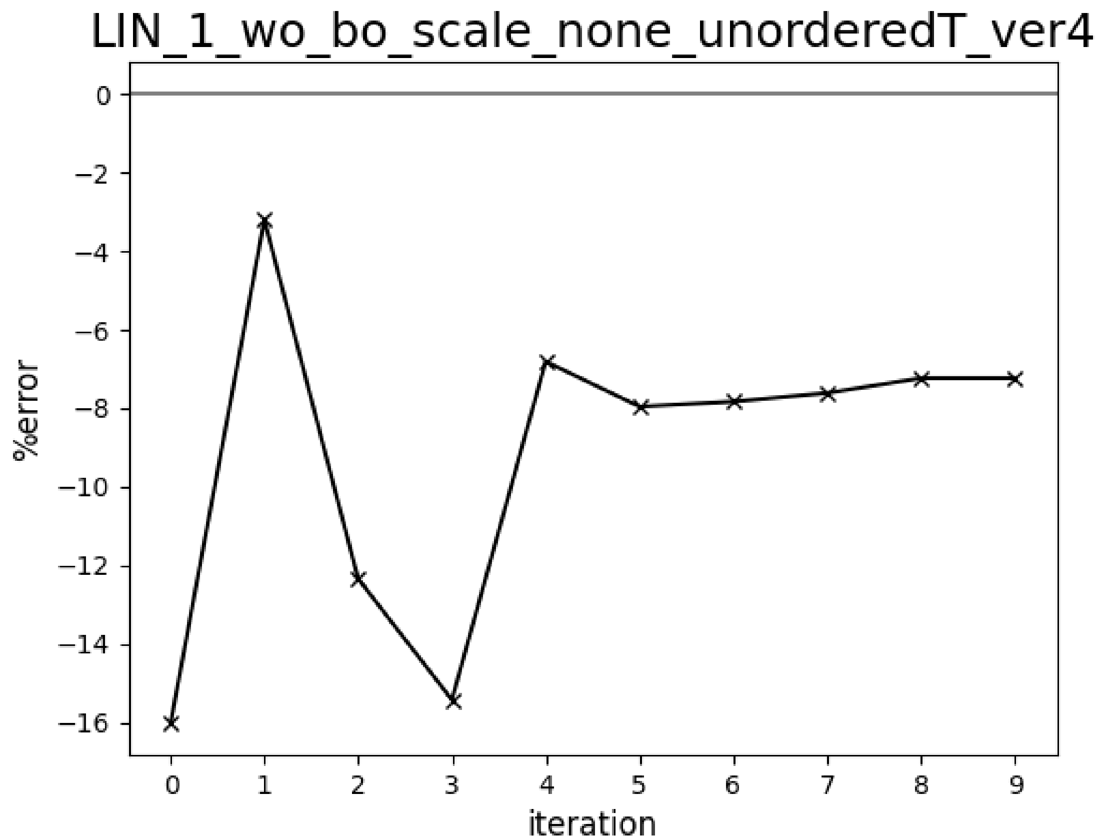
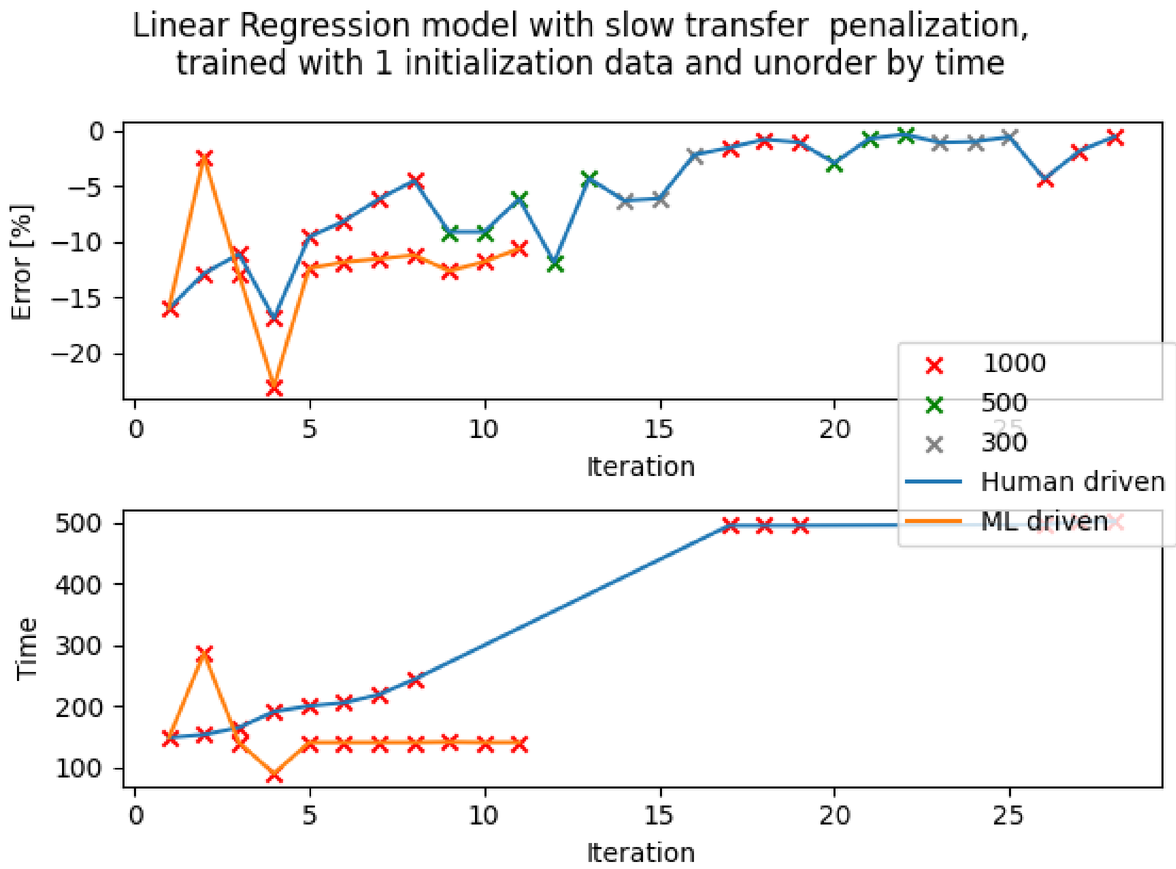
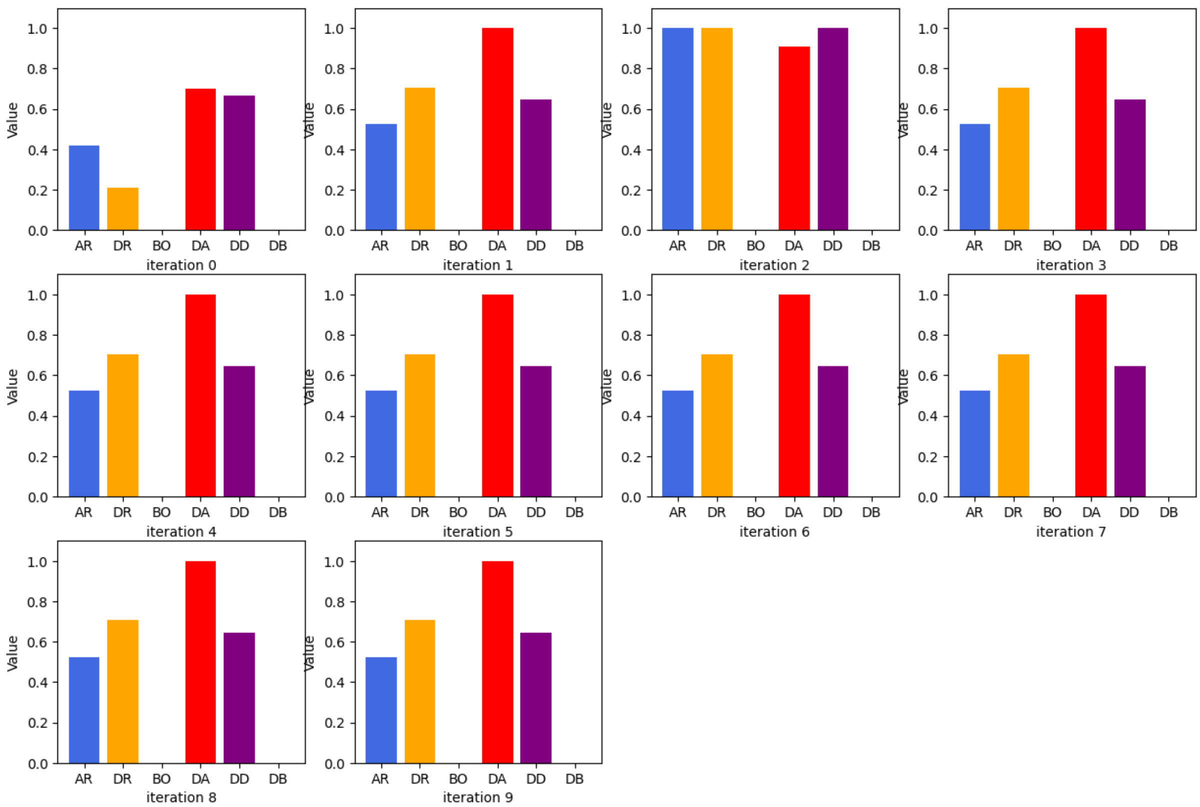
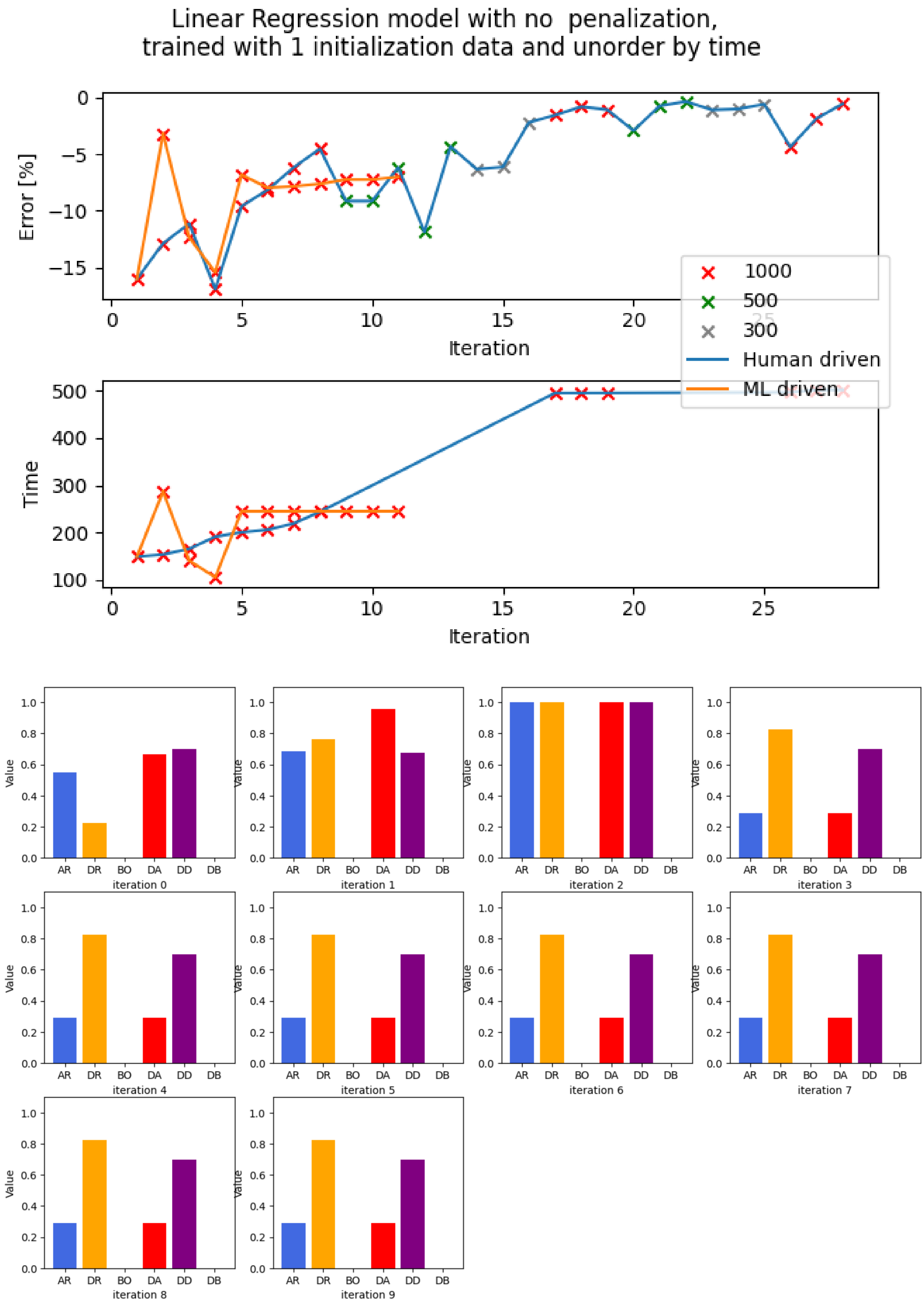


Diagram comparing human-driven and ML test trials (1275 cP, set:1, ver4)






Linear Regression model with no penalization,
trained with 1 initialization data and unorder by time
Error [%]
Iteration
1000
500
300
Human driven
ML driven
Time
Iteration
Value
AR
DR
BO
DA
DD
DB
iteration 0
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
iteration 9

Set: 1 (sampling: distributed)

<u>Observation of trends</u>

3 April
Scaling: none, Order: unorderedT

The test trials done using the updated algorithm provide a good variation of parameters that are much more distributed than when the sampling method is set to “random". Although the percentage error generated is still quite a distance from the preferred range of -2 to 2 %, it is clear that the algorithm is trying different types of parameters, such as parameters with both fast aspiration and dispense rates, with slow aspiration and fast dispense rates, and slow aspiration and dispense rate. (the dispense rate is rarely faster aspiration rate)

For this updated algorithm, one can observe a more accurate judgement of percentage error from the algorithm, and it can be seen improving when the same set of parameters are generated.

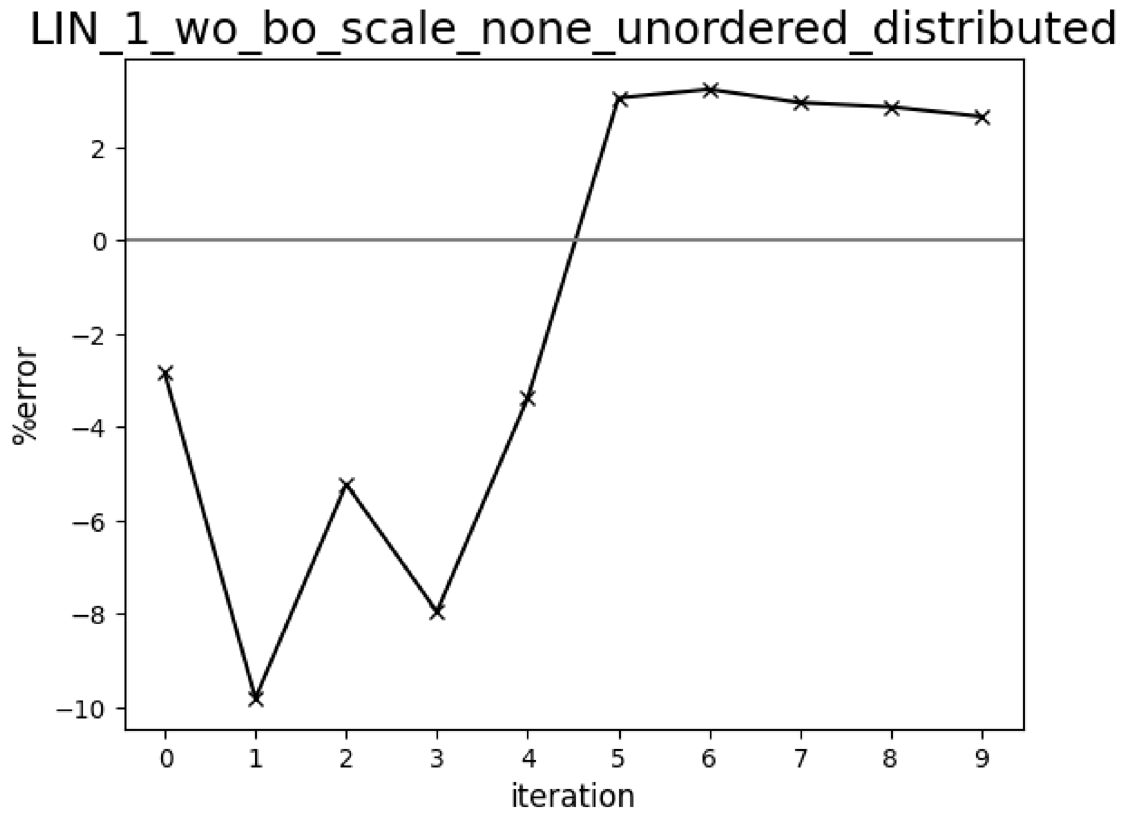
For example, when a set of parameters is first generated, it thinks that the %error is -7%. But when the mass measured does not align with it, it generates the parameters again thinking that it gives 0.3%. When the mass is still a small margin away from the expected mass, it generates the parameters again and it gets closer and closer to the actual percentage error.

4 April
Scaling: multiply, Order: unorderedT

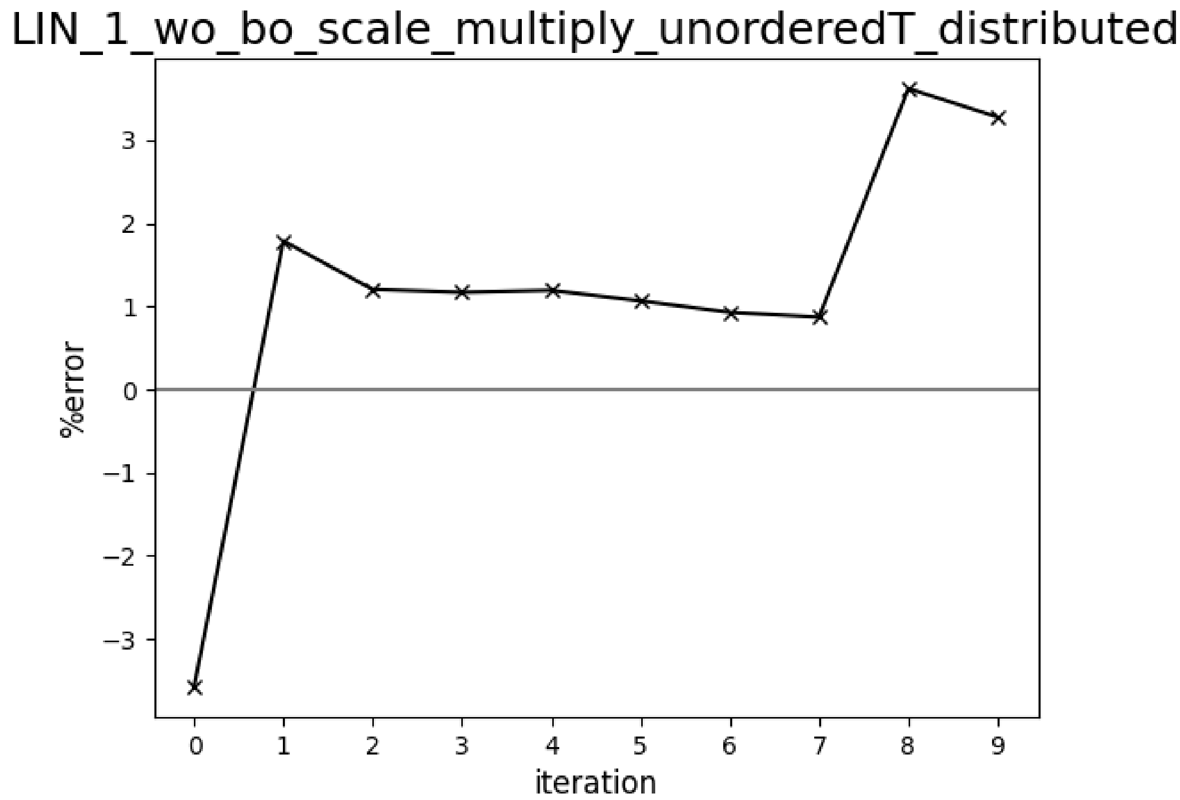
Although there are slow transfer time penalties, it decided to choose a set of parameters that are slow and accurate (within tolerance of 1000uL). On the other hand, the penalization worked for ver4 multiplication trials. It worked in the opposite direction, even suggesting parameters that are capable of transferring liquid of much higher viscosity.

6 April
Scaling: divide, Order: unorderedT

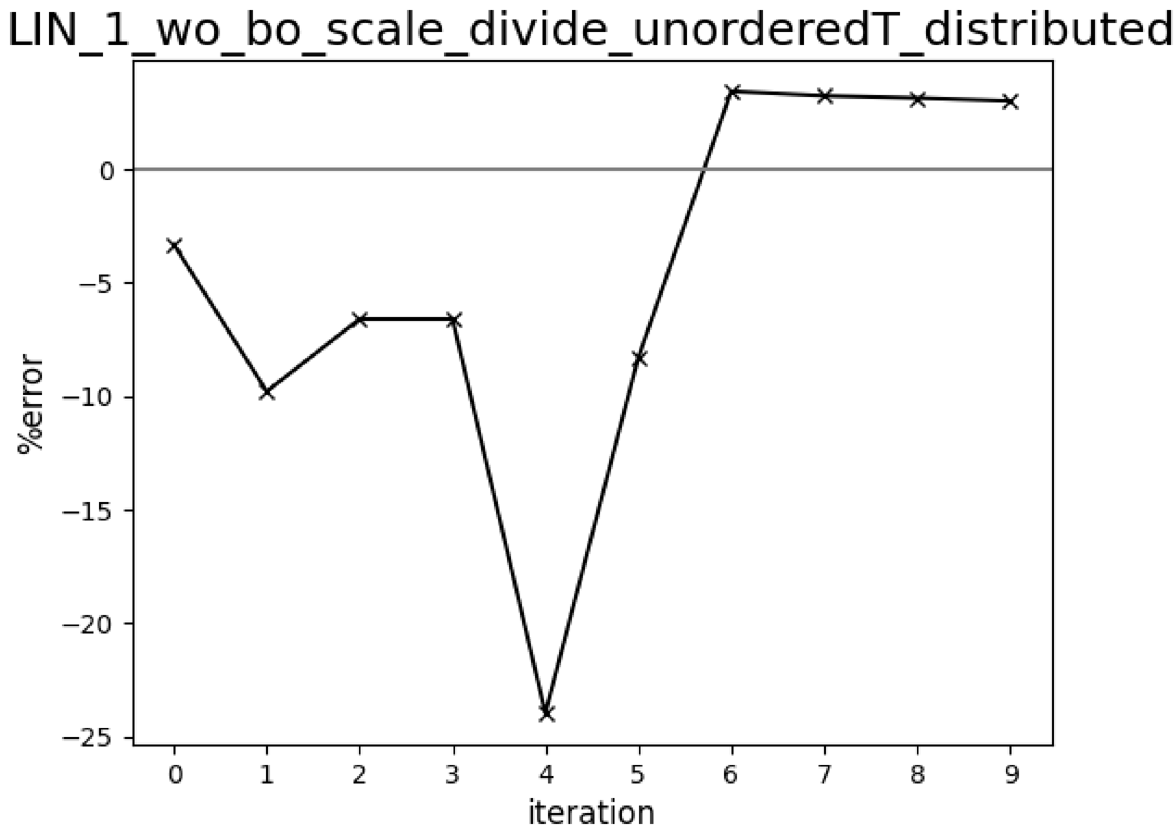
There are more variations in the sets of parameters generated but the last 4 sets are repeated. Transfer time measured is around the same as that of standard calibration and the percentage error is around 3%.



LIN_1_wo_bo_scale_none_unordered_distributed
%error
2
0
−2
−4
−6
−8
−10
0
1
2
3
4
5
6
7
8
9
iteration

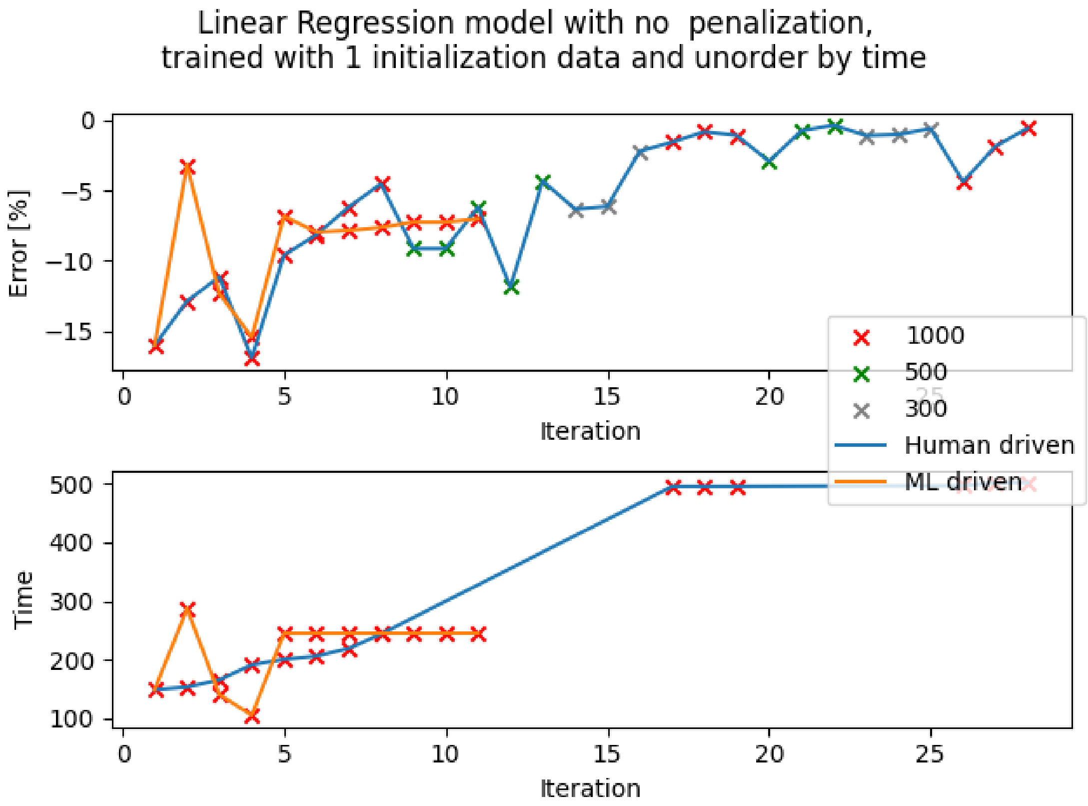


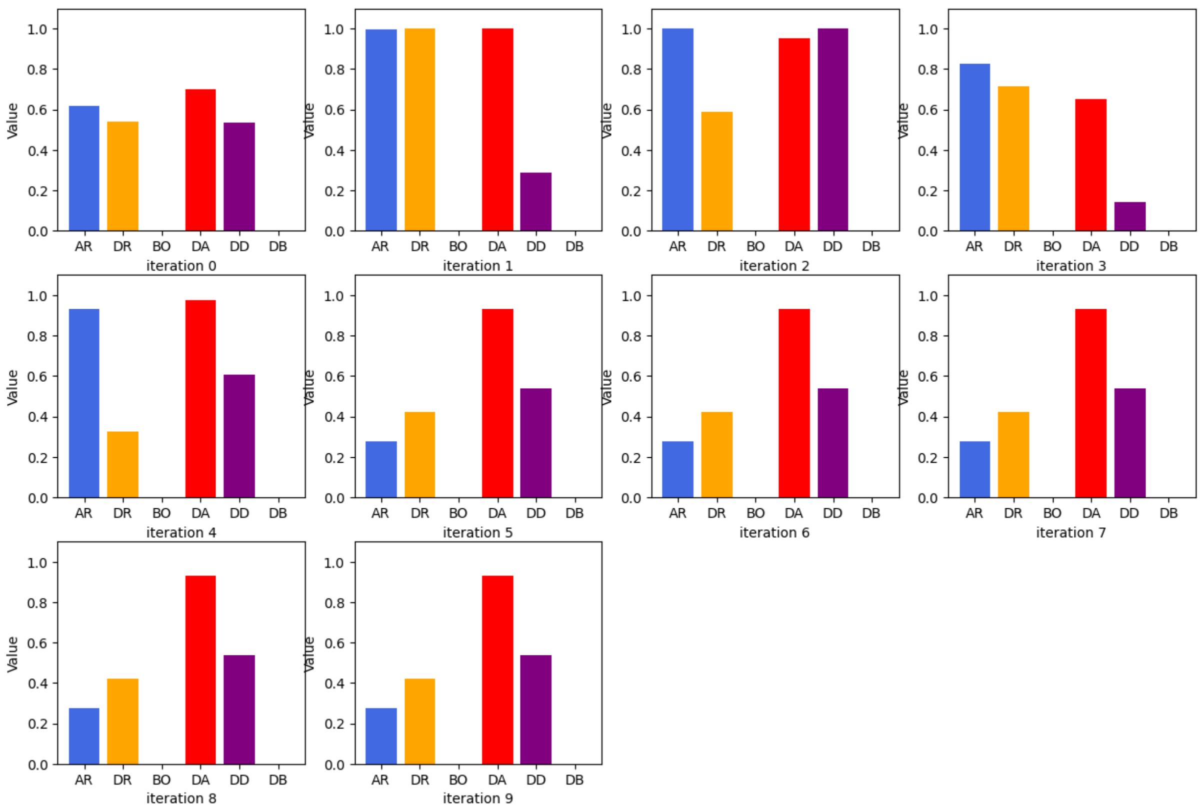
LIN_1_wo_bo_scale_multiply_unorderedT_distributed
%error
3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9
iteration



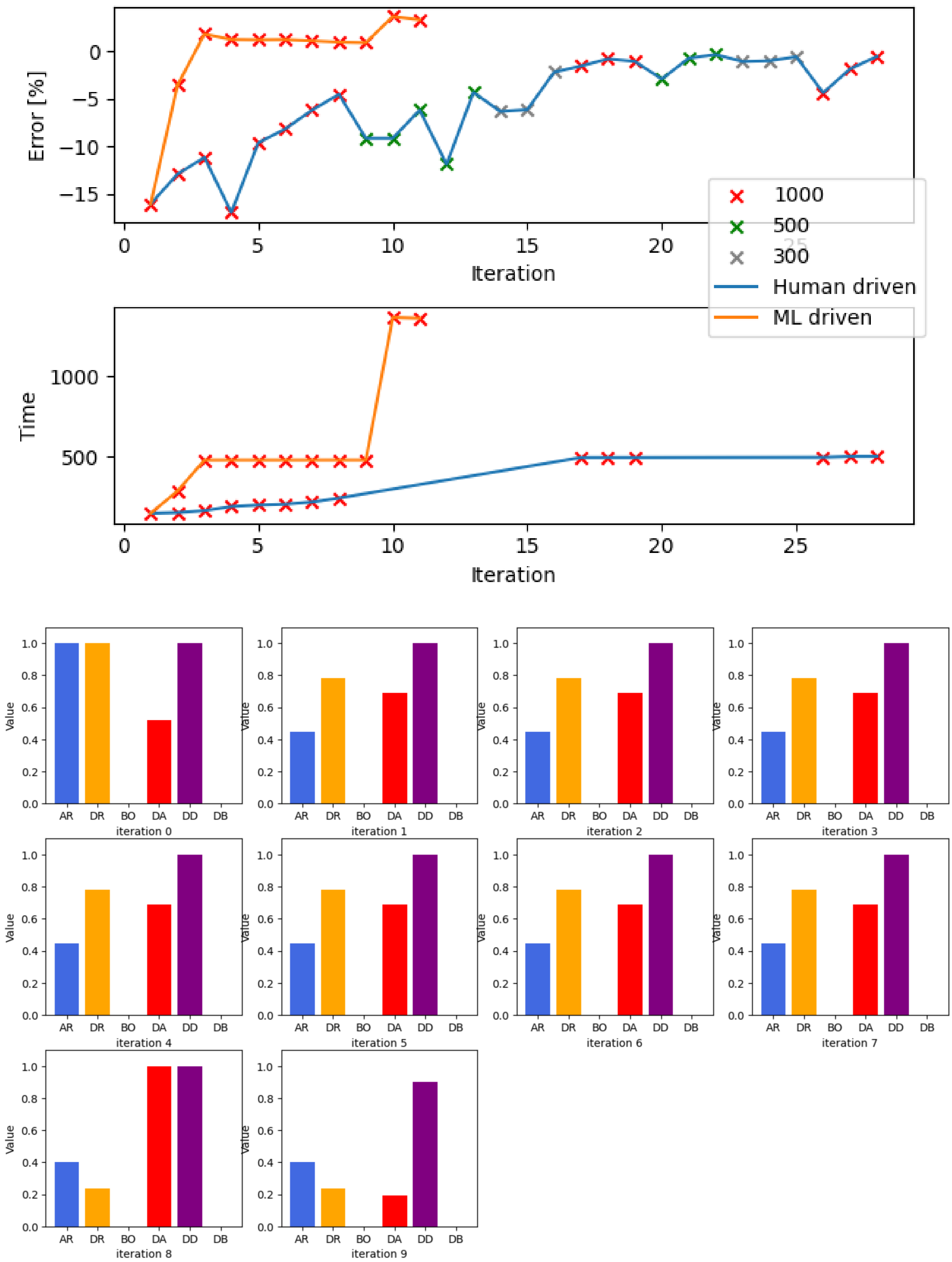
LIN_1_wo_bo_scale_divide_unorderedT_distributed
0
−5
−10
−15
−20
−25
%error
0
1
2
3
4
5
6
7
8
9
iteration

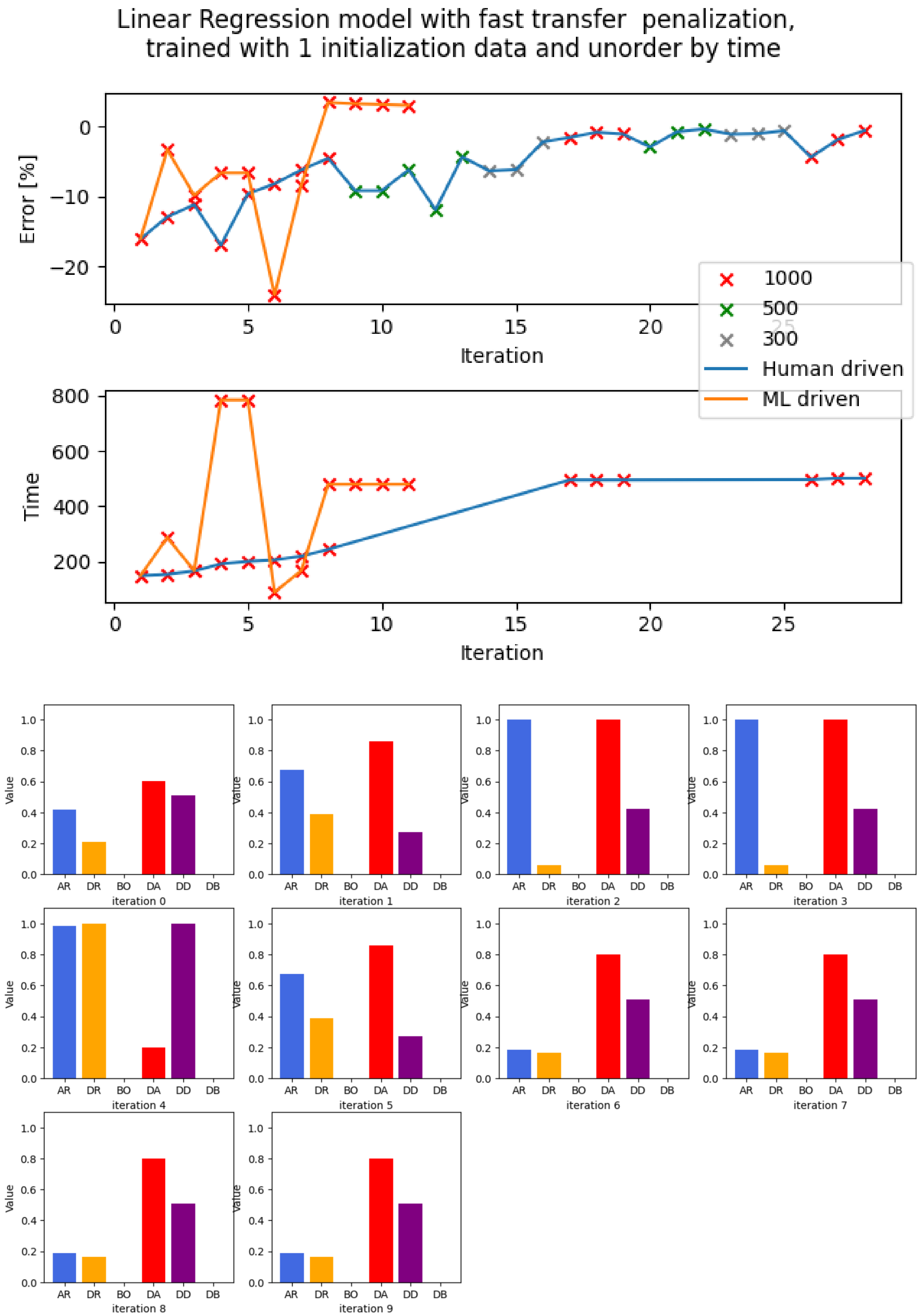
Diagram comparing human-driven and ML test trials (1275 cP, set:1, distributed)





Linear Regression model with slow transfer penalization,
trained with 1 initialization data and unorder by time




Linear Regression model with fast transfer penalization, trained with 1 initialization data and unorder by time
Error [%]
Iteration
Time
Iteration
1000
500
300
Human driven
ML driven
Value
AR
DR
BO
DA
DD
DB
iteration 0
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
iteration 9

Set: 1 (amended - real LIN)

Observation of trends

Observation 1: LIN - scaling: divide, Order: unorderedT

The first 5 trials are quite random (100 to 600s and -2 to -13% for percentage error). After that, it starts repeating (around 20s faster than the standard calibration, percentage error around 2.5%). The penalisation did not work too for this series of test trials.

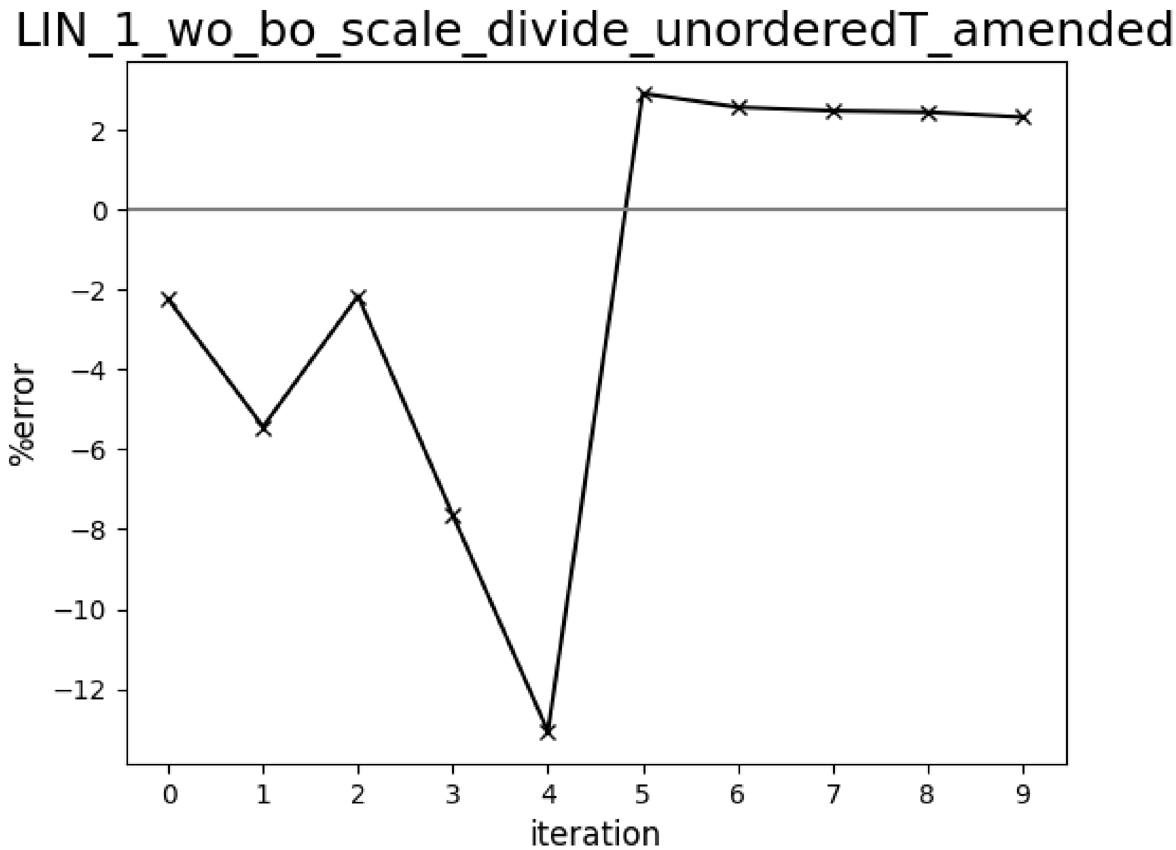
Observation 2: LIN - scaling: multiply, Order: unorderedT

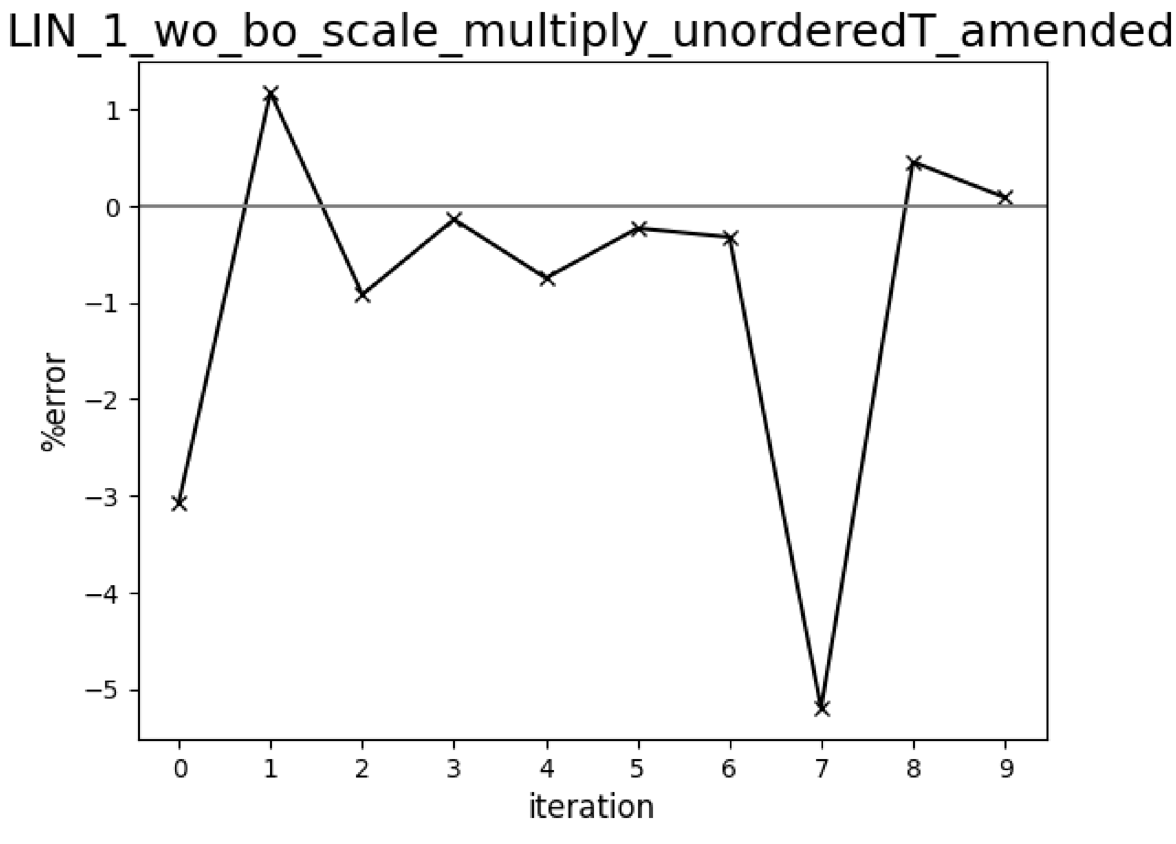
The trials are all really accurate, the percentage error fluctuates very close to 0% for most trials with a few outliers of -5% and 1%. There's one trial that took really long (around 1200s) but the rest are around 400 to 600s on average.

Observation 3: LIN - scaling: none, Order: unorderedT

The percentage error fluctuates beyond the preferred boundary of -2 to 2% and lacks consistency. The transfer times are scattered between 300s to 600s since there is no penalization applied to it. But overall, a converging trend can be seen in the percentage error across trials.

Percentage error against iteration graphs:





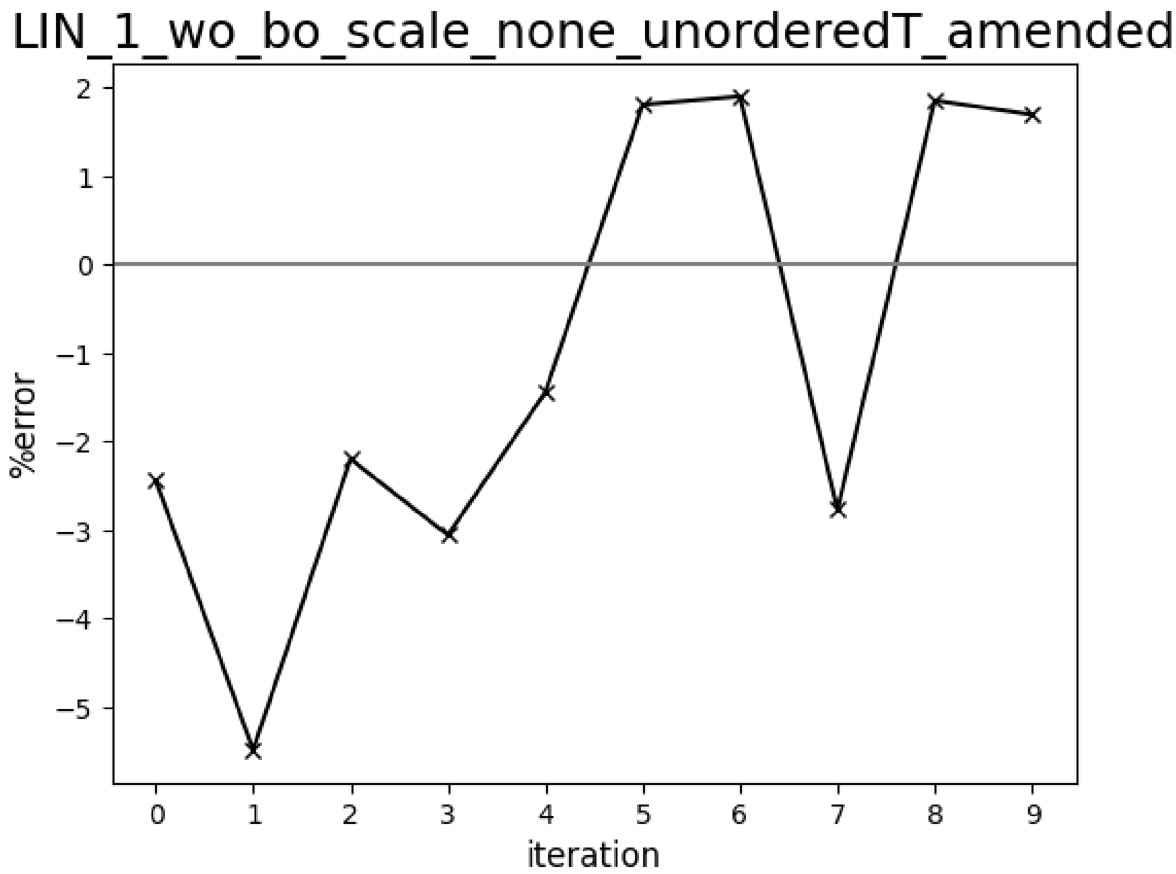
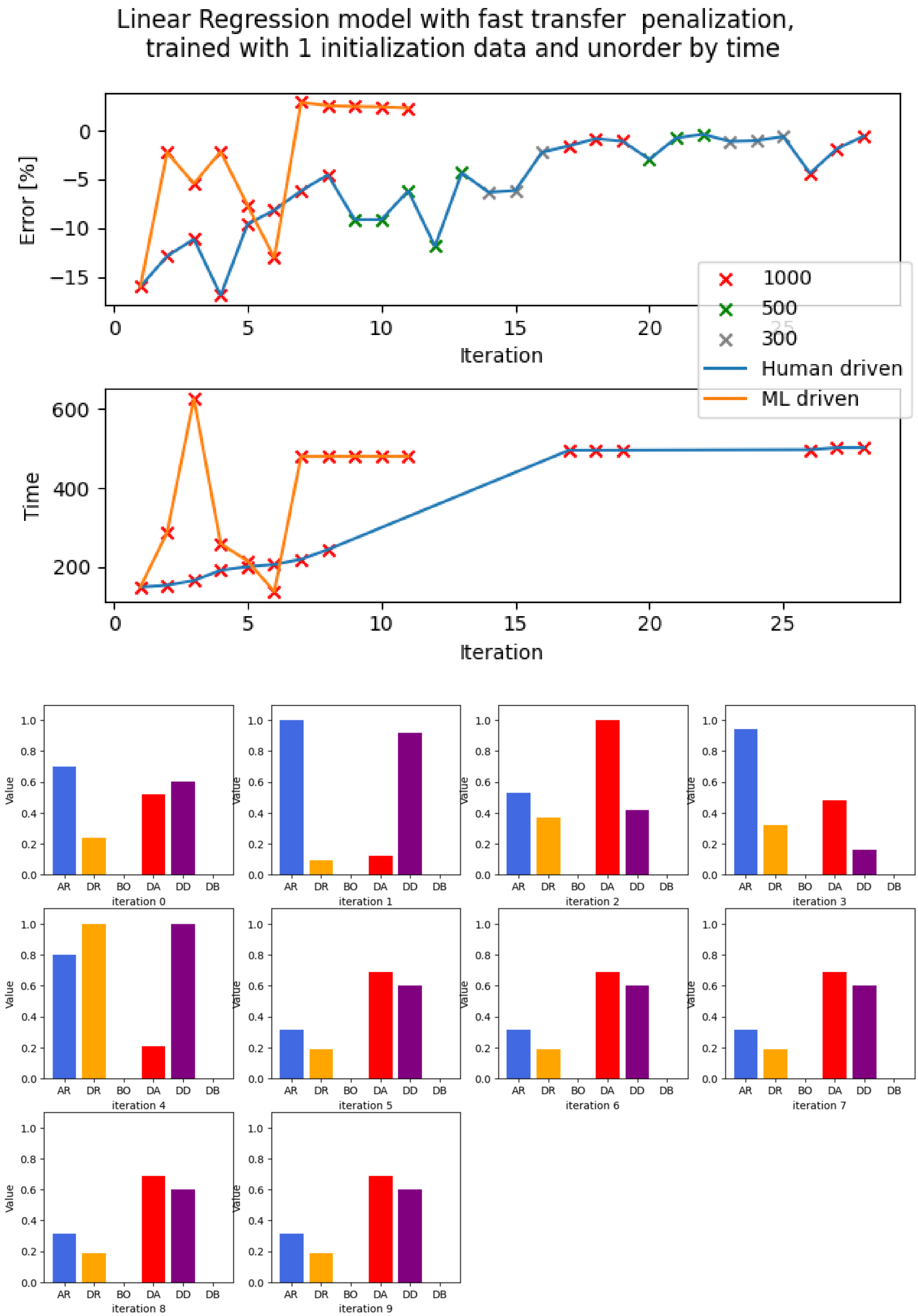
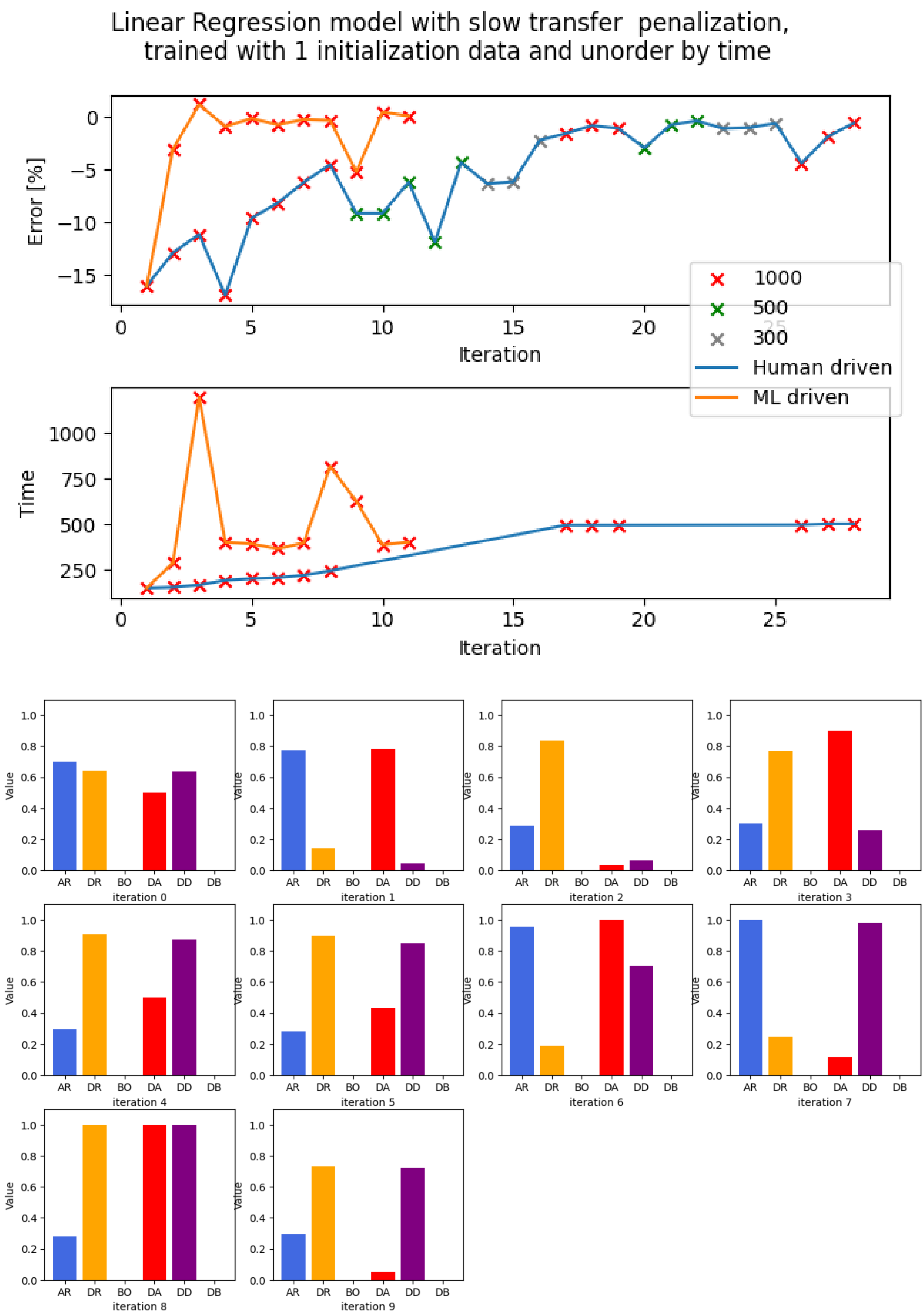
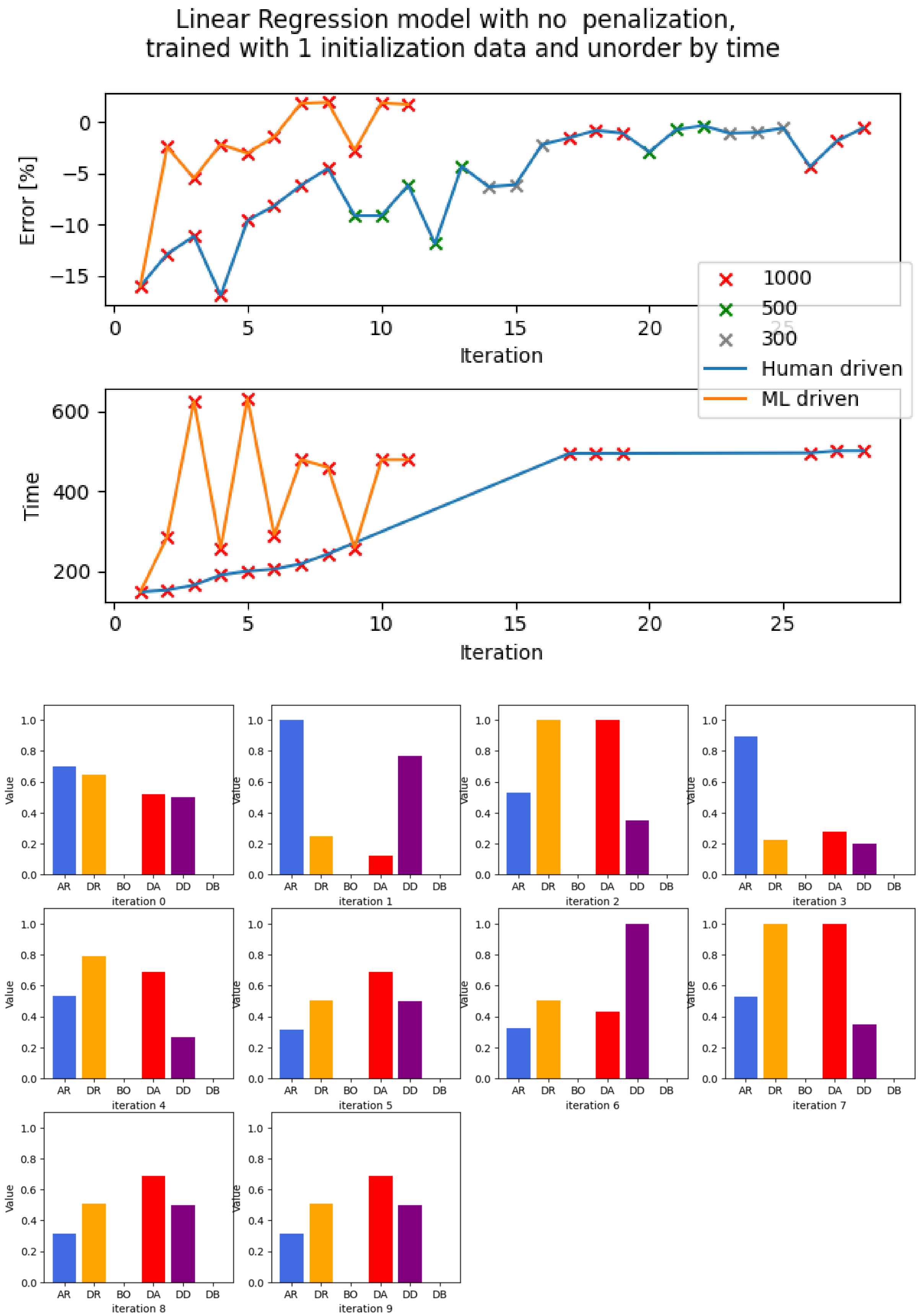

LIN_1_wo_bo_scale_none_unorderedT_amended
%error
iteration
2
1
0
−1
−2
−3
−4
−5
0
1
2
3
4
5
6
7
8
9

Diagram comparing human-driven and ML test trials (1275 cP, set:1, amended: LIN)




Linear Regression model with slow transfer penalization,
trained with 1 initialization data and unorder by time
Error [%]
Iteration
1000
500
300
Human driven
ML driven
Time
Iteration
Value
AR DR BO DA DD DB
iteration 0
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
iteration 9


Linear Regression model with no penalization,
trained with 1 initialization data and unorder by time
Error [%]
Iteration
Time
Iteration
1000
500
300
Human driven
ML driven
Value
AR DR BO DA DD DB
iteration 0
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
iteration 9

Set: 1 (absolute)

<u>Observation of trends</u>

Observation 1: GPR - scaling: multiply

For most of the test trials, the percentage error falls within the preferred boundary of -2 to 2% and the average transfer time is slightly faster than the one derived in standard calibration. Thus, in this case, the slow time penalization seemed to work. Most of the trials are repetitions but it's alright since both the time and percentage error is favourable.

Observation 2: GPR - scaling: none

Error against iteration graphs:

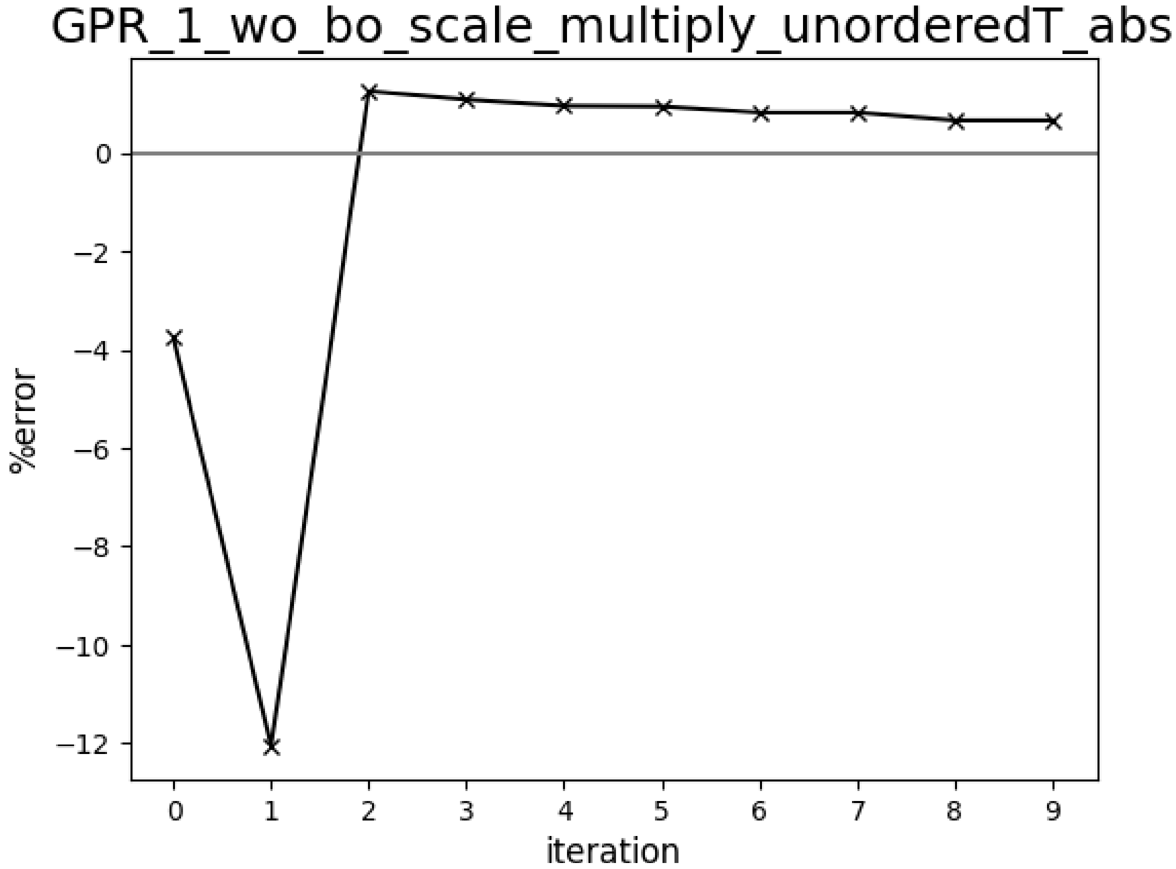


Diagram comparing human-driven and ML test trials (1275 cP, set:1, absolute)