Viscosity project summary_Section 2_ML 398 cP

Table of content

Standard 398.5 cP

Machine learning segment

Set: Half

<u>Observation of trends</u>

Observation 1: GPR - scaling: Division and LIN - scaling: Multiplication and Division

Once an 'optimal' set of parameters are attained, it will double check one time to confirm, then continue to explore parameters. The 'confirmation' stage usually occurs in trial 5 or 6 of the 10 trials and it can continue to suggest parameters that deviate quite far from the 'optimal' set of parameters.
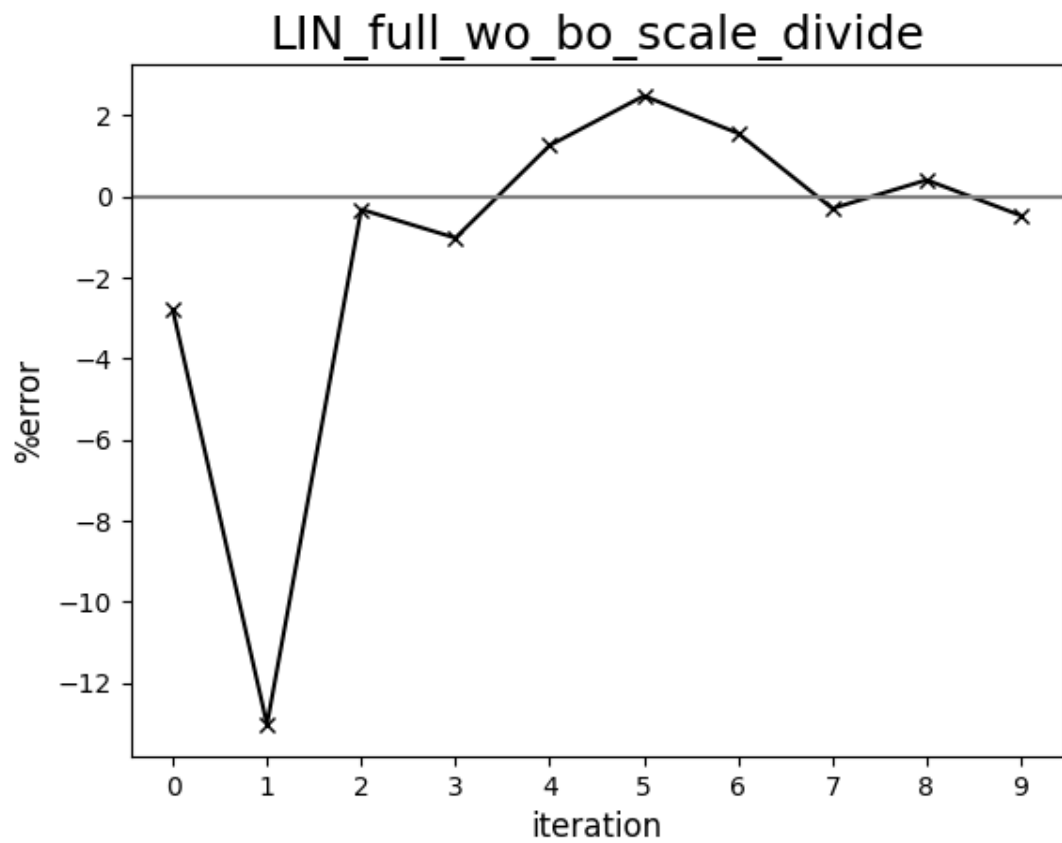
Observation 2: GPR - scaling: Multiplication

Once an 'optimal' set of parameters are attained, it will remain the same for the rest of the 10 trials. The 'optimal' set of data is usually attained in trial 5 or 6 too, similar to when division is used for scaling.
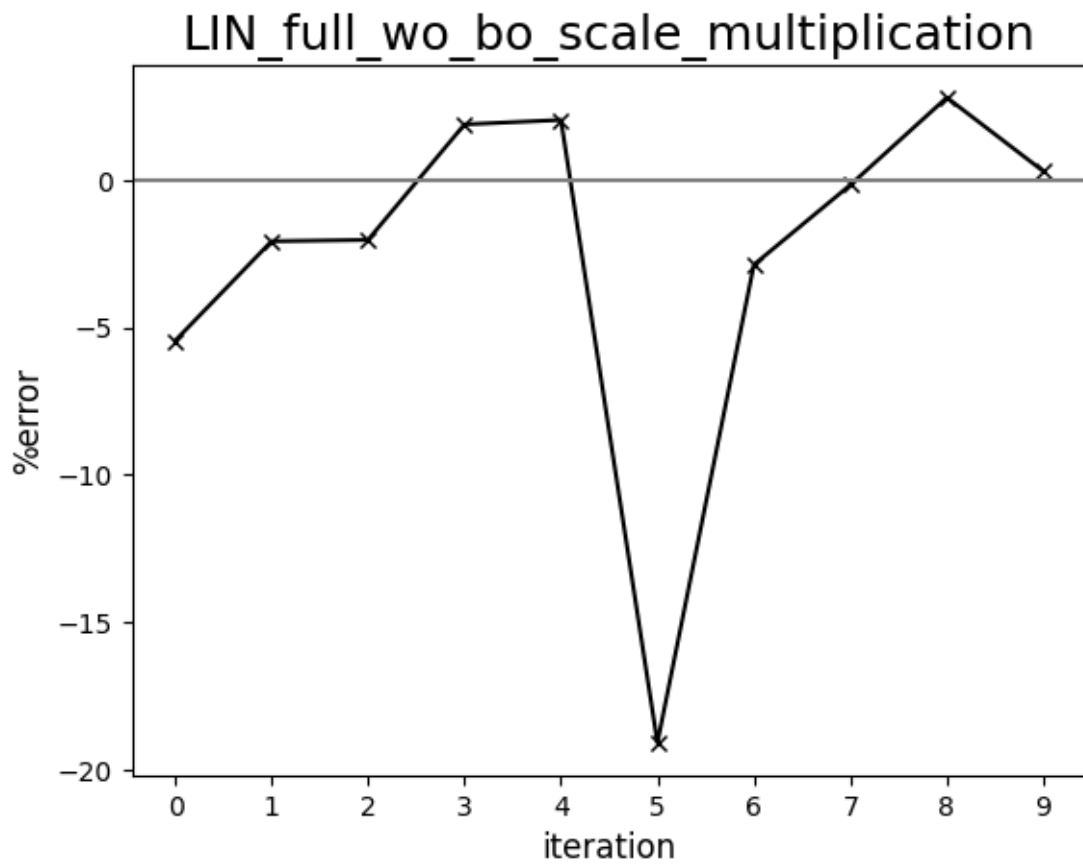
Observation 3: GPR - scaling: Without and LIN - scaling: Without

A converging trend is observed for percentage error. The first 2 trials' %error usually falls between 5-6% but it converges towards 0 from trial 2-9. However, trial 10 deviated from the consistency observed between trial 2-9. Unlike the trials with GPR-multiplication used in the scaling code (stays the same after 'optimal' parameters are found), the parameters continue to change slightly after the 'optimal' set of parameters are attained. For that, GPR without scaling can be said to be somewhere between when division and multiplication are used for the scaling parameters (continues to 'explore' moderately after 'optimal set' is found).
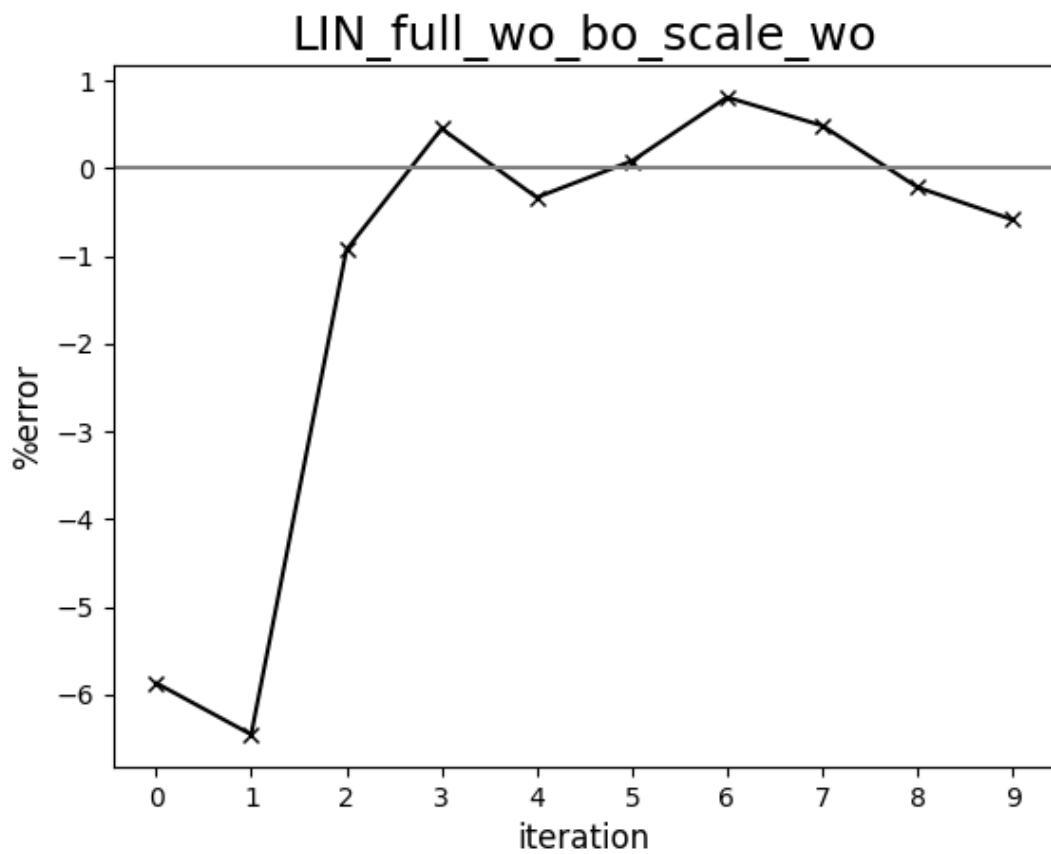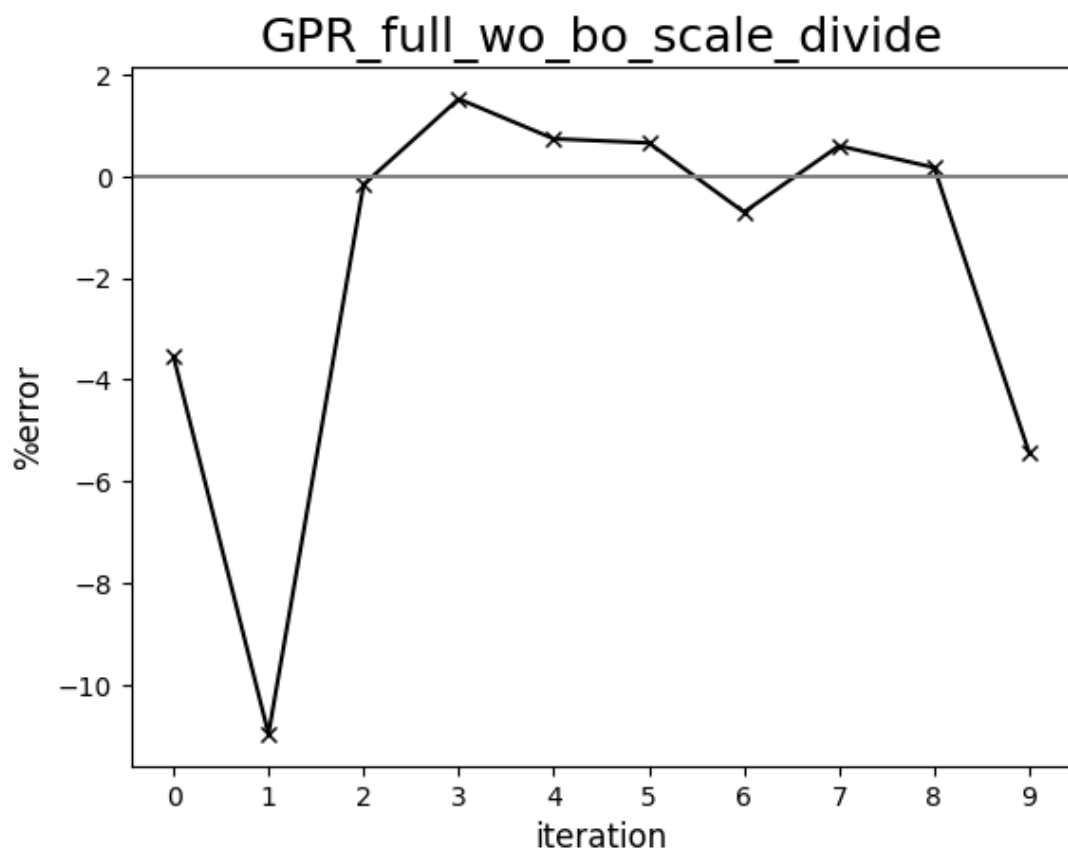
LIN scaling: division

LIN_full_wo_bo_scale_divide
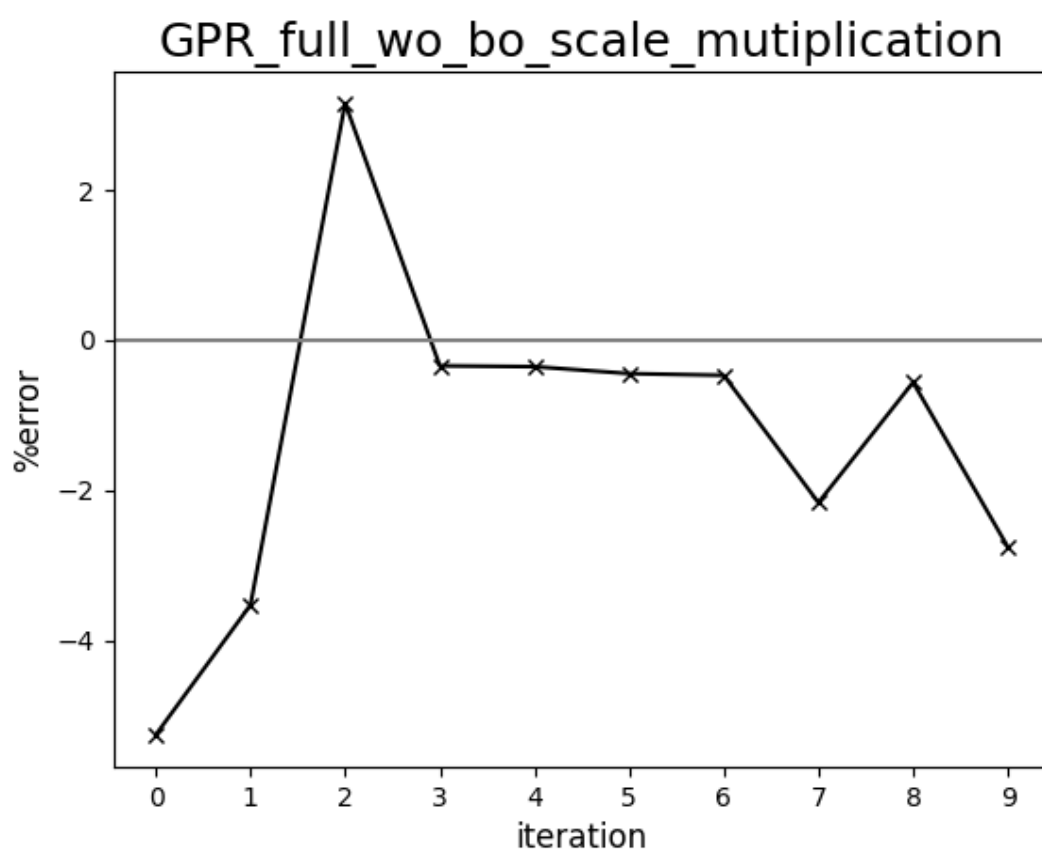

LIN scaling: multiplication

LIN_full_wo_bo_scale_multiplication

LIN scaling: without

LIN_full_wo_bo_scale_wo

GPR scaling: division

GPR_full_wo_bo_scale_divide

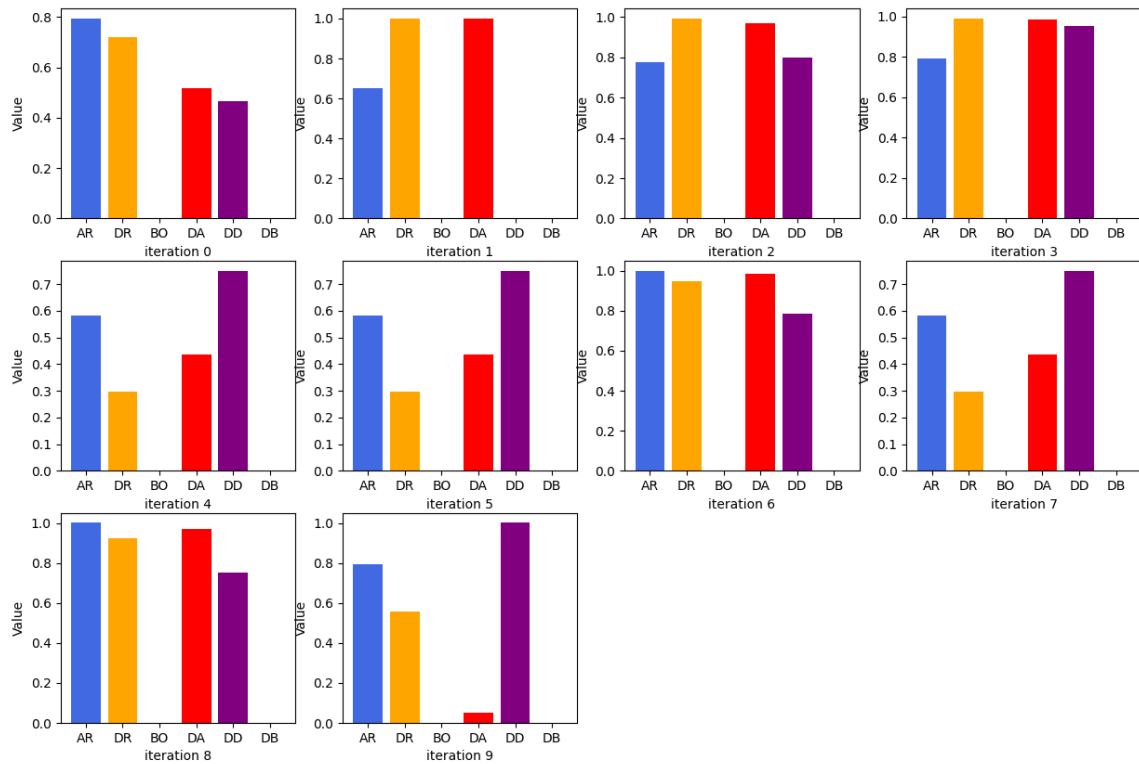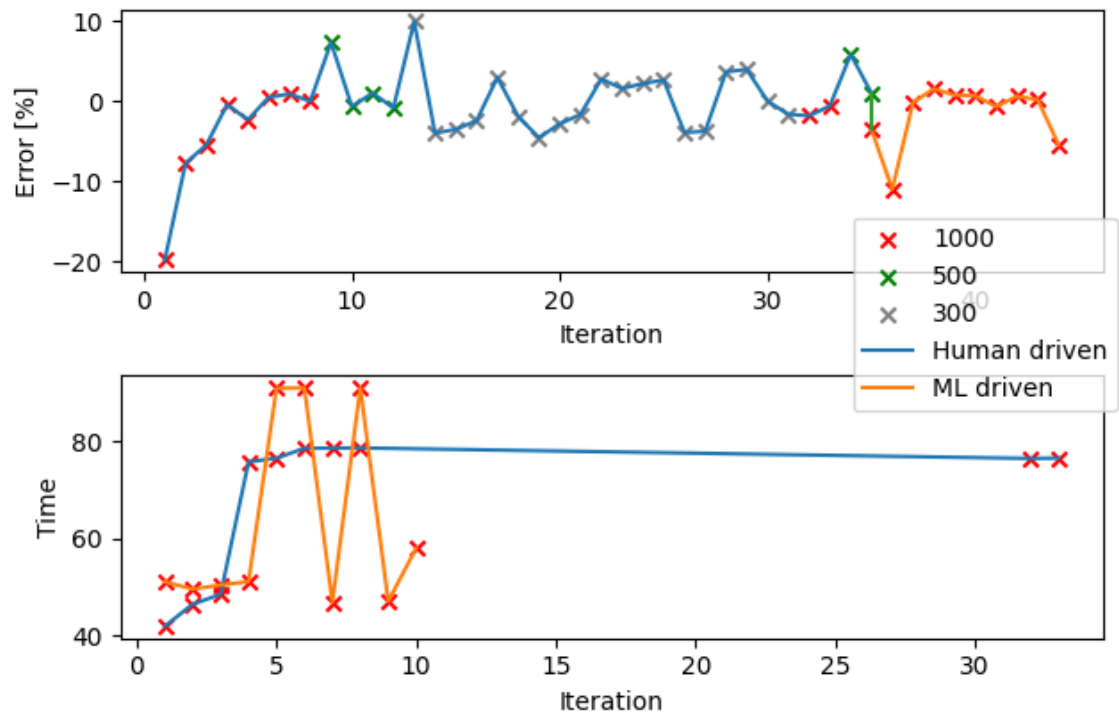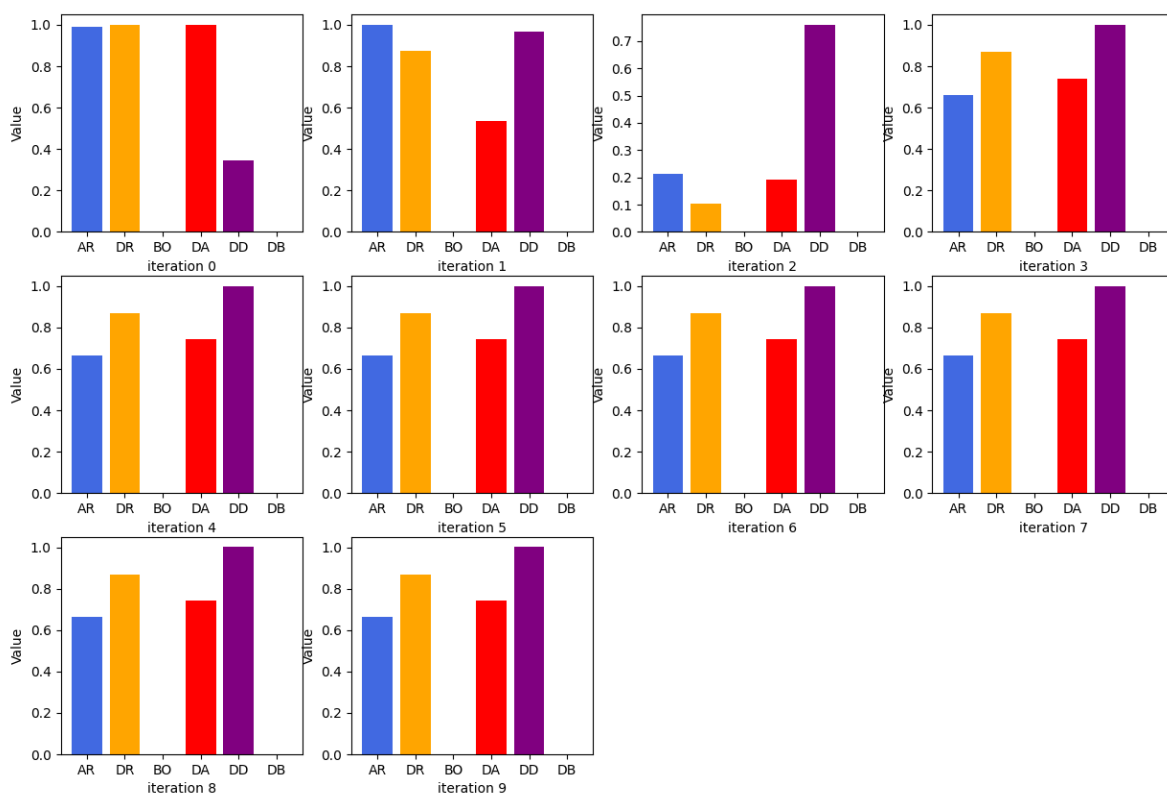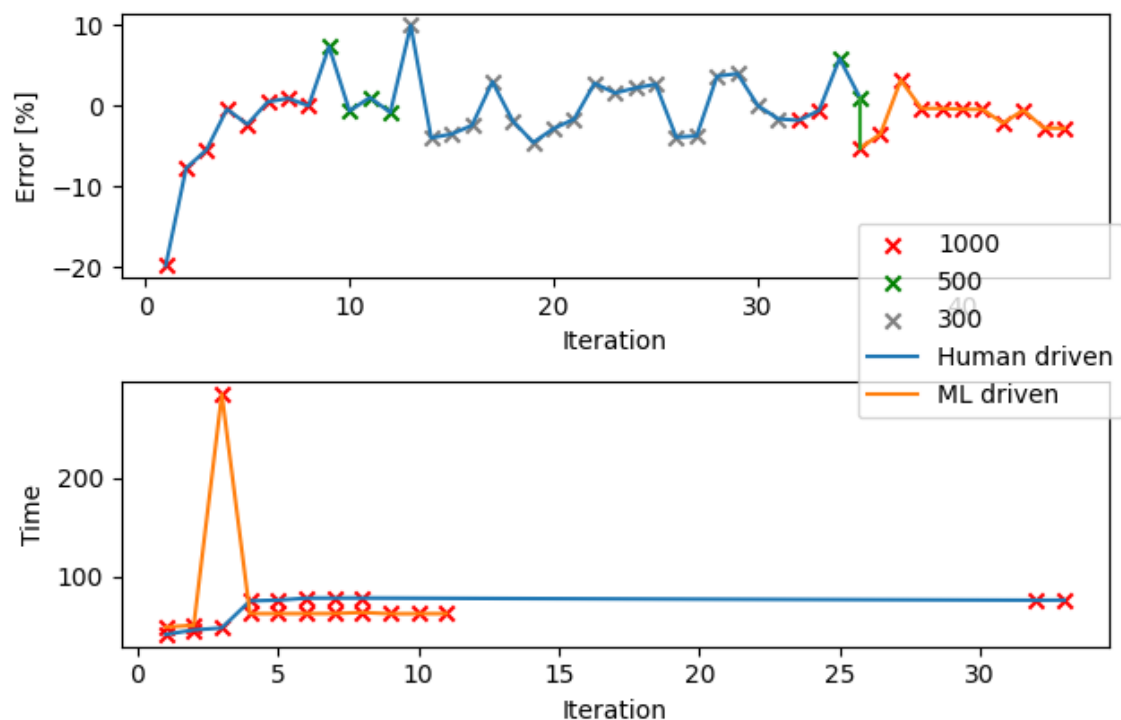GPR scaling: multiplication

GPR scaling: without
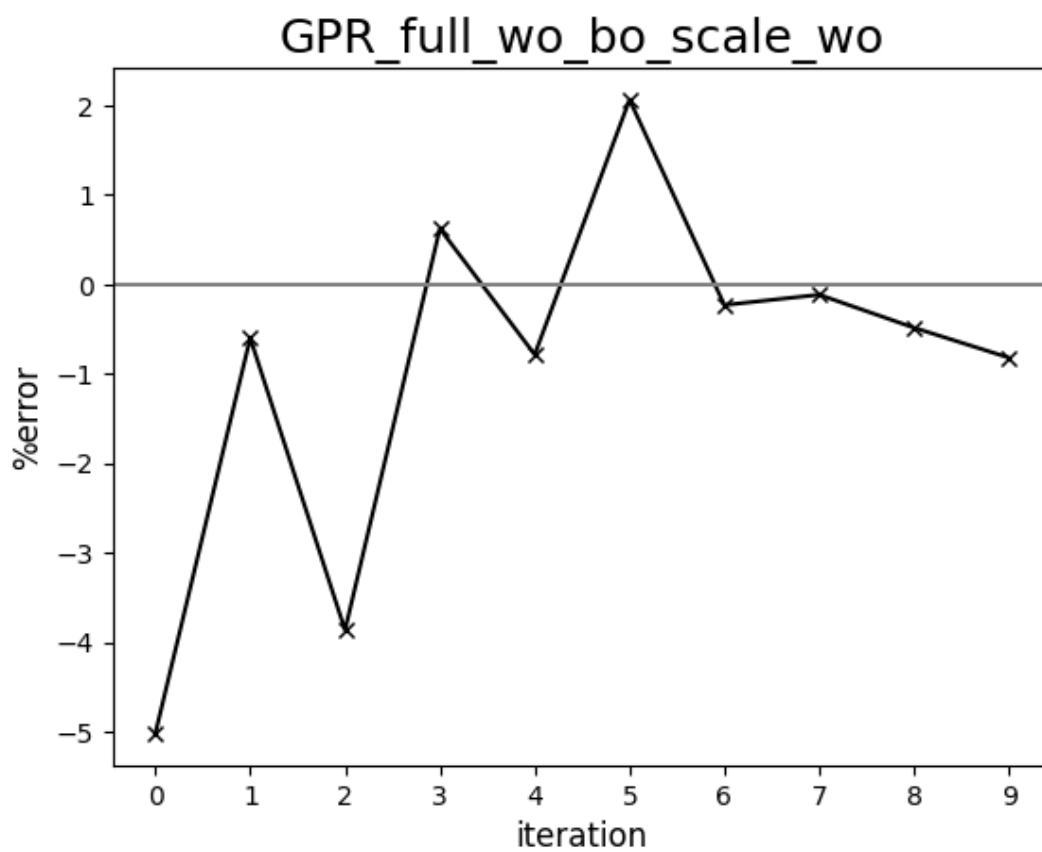
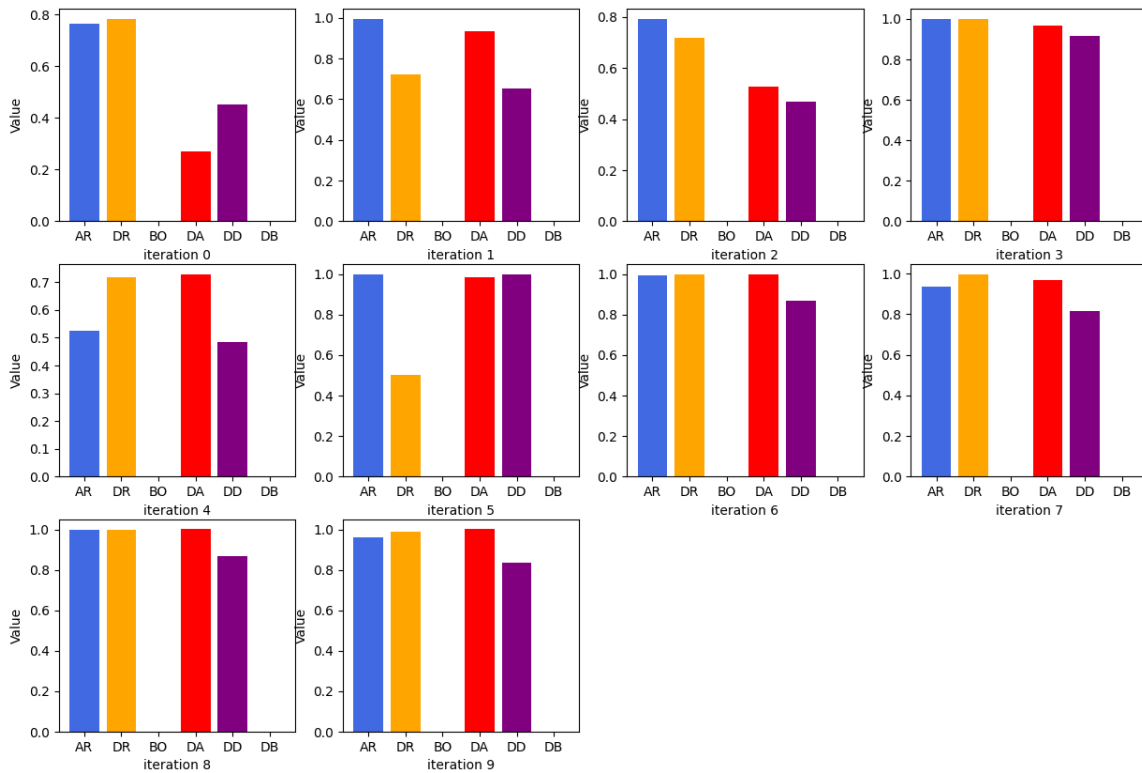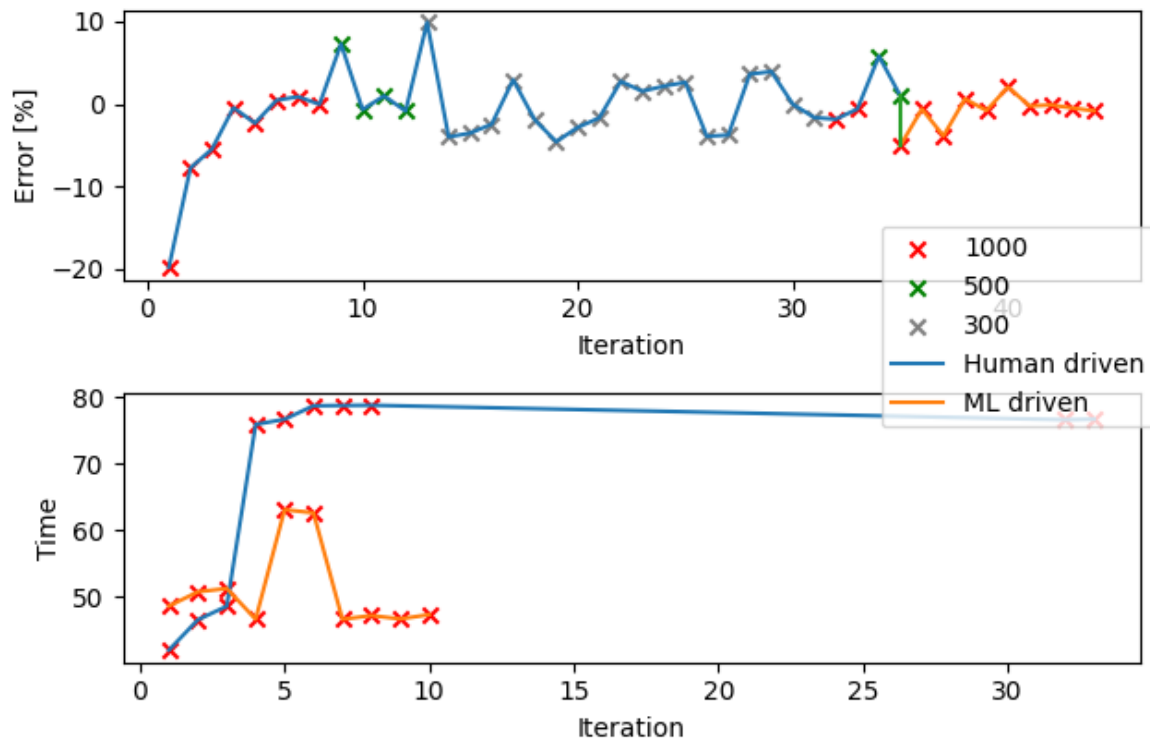# Diagram comparing human-driven and ML test trials (398 cP, set:half)

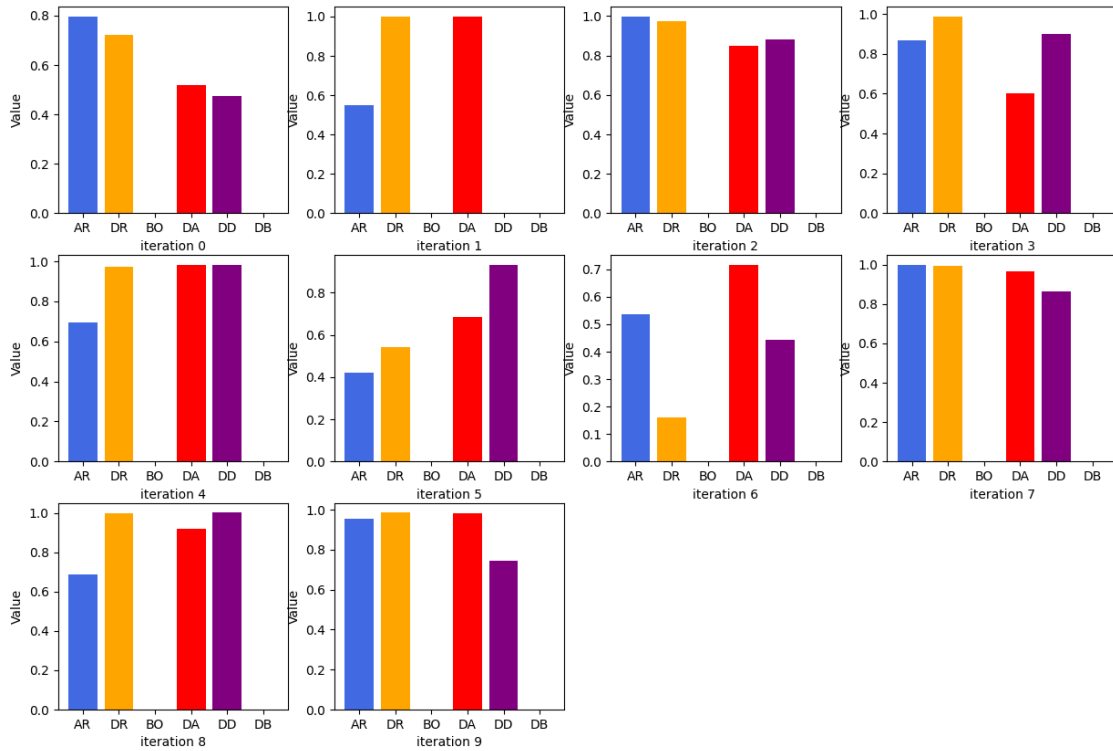## Gaussian Process Regression model with fast transfer penalization, trained with half data set
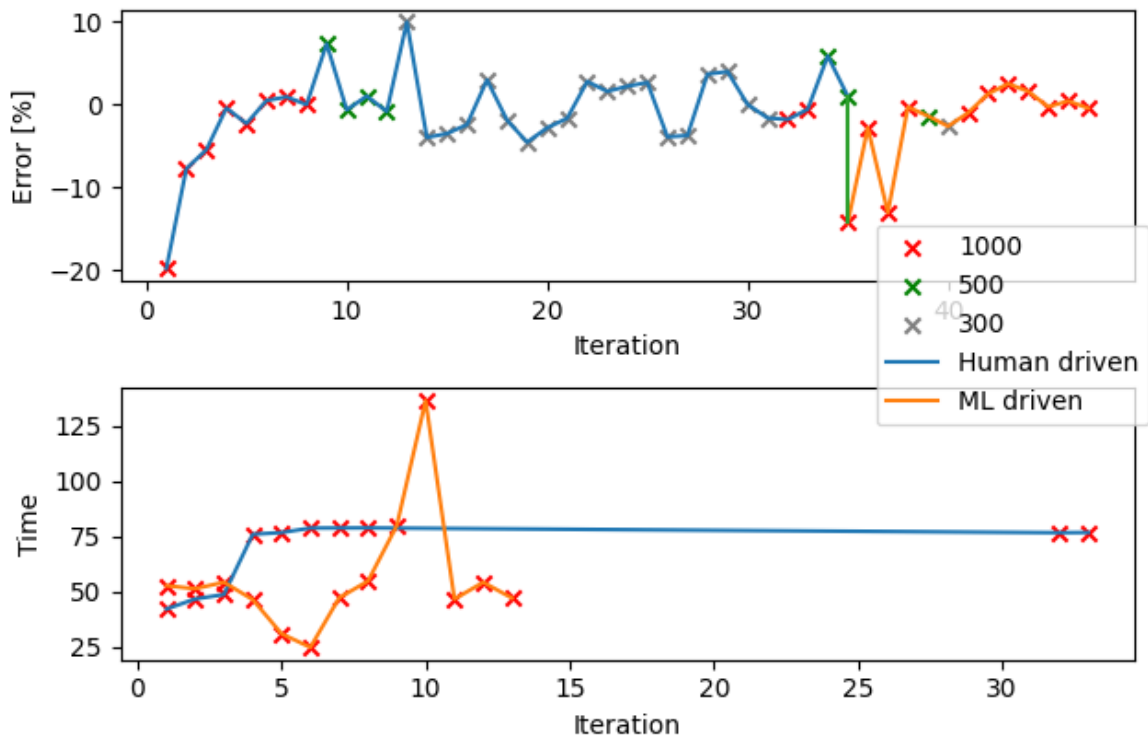
Gaussian Process Regression model with slow transfer penalization, trained with full data set
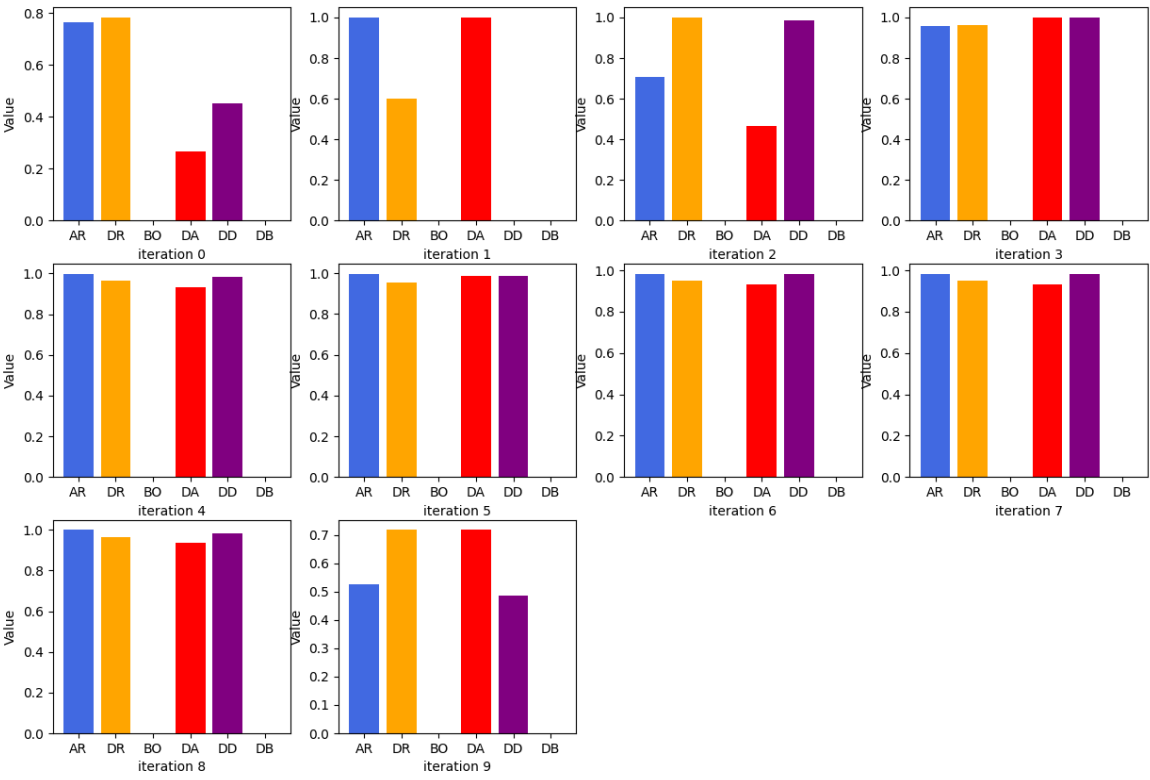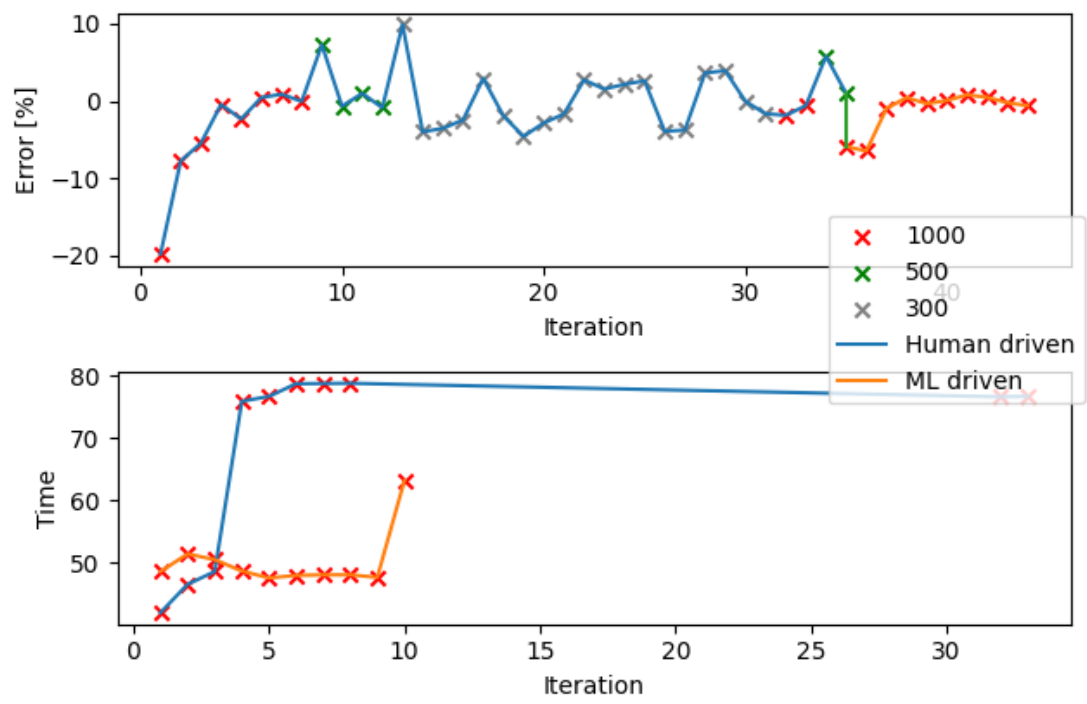
Gaussian Process Regression model with no penalization,
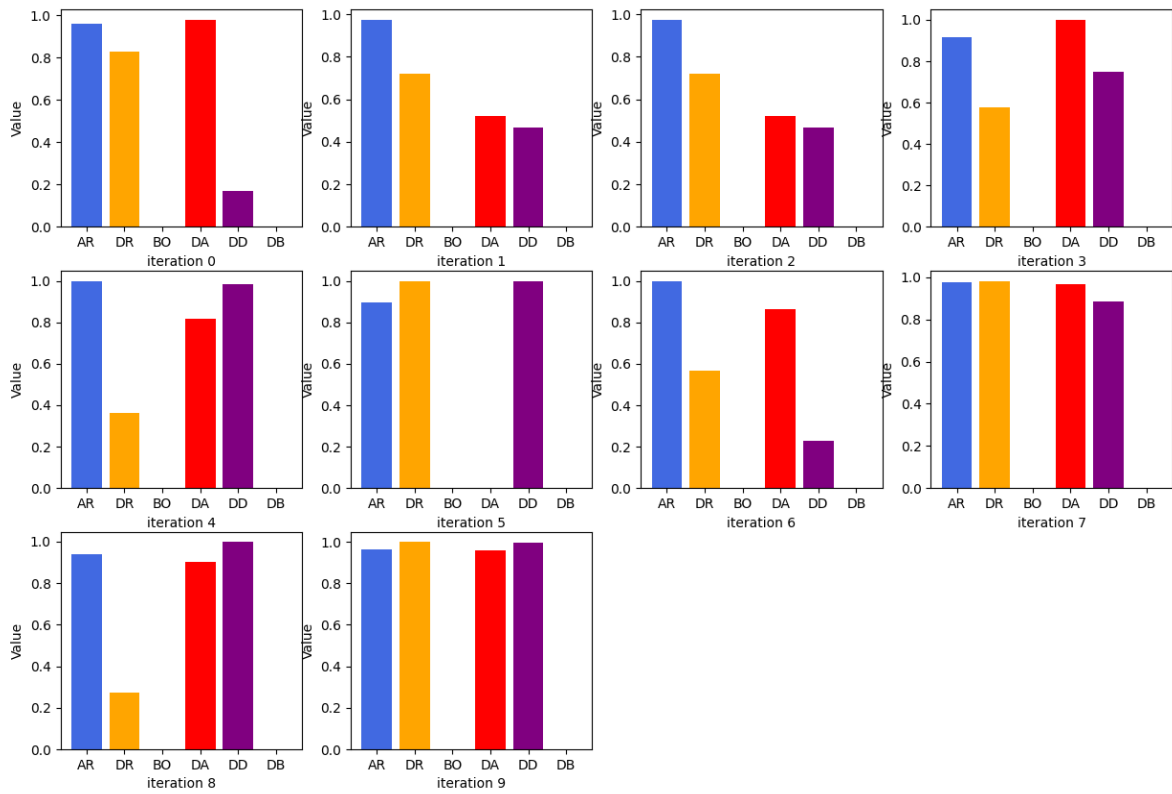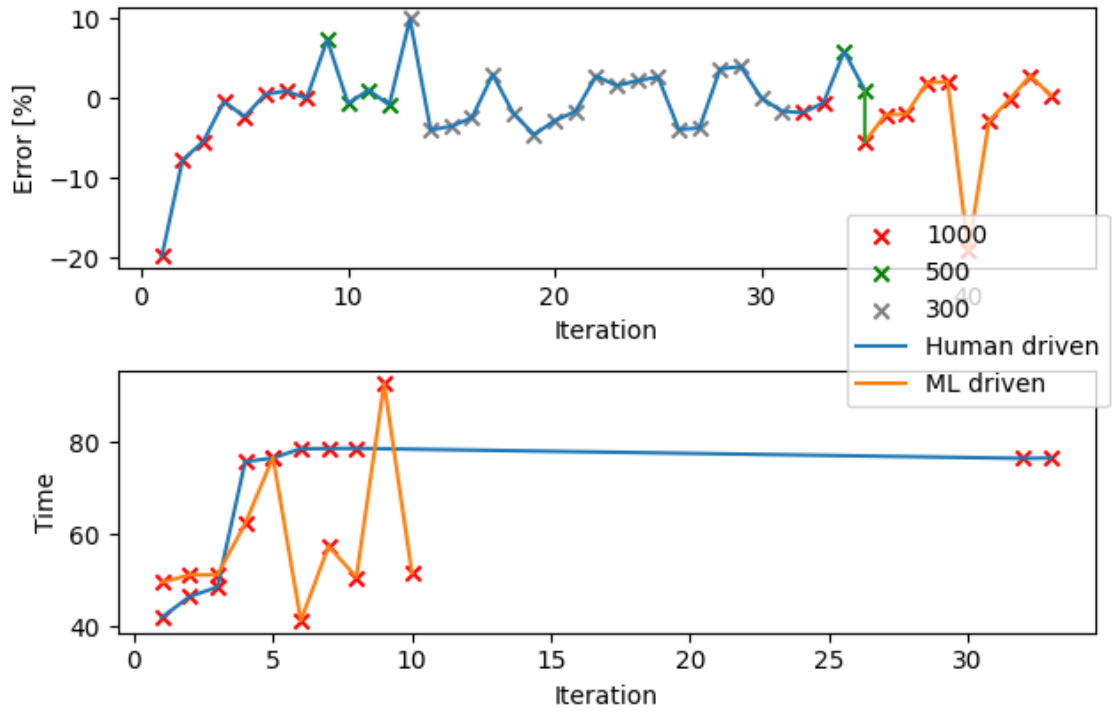trained with full data set

Linear Regression model with fast transfer penalization, trained with full data set

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

Linear Regression model with slow transfer penalization, trained with full data set

Set: Half (unorderedT)

<u>Observation of trends</u>

Observation 1: GPR - scaling: multiply

Some repetitions in the sets of parameters generated but it does show a decent amount of variation (set: 1 trials often have more repetition, GPR especially - sometimes all 10 sets of parameters are the same). However, the sets of parameters generated do not seem to align with the penalization used. When it is changed to multiply, it should penalise slower transfer times but this iteration has more long time trials as compared to GPR - none, which does not have the penalization. The percentage error is not so bad,

Observation 2: GPR - scaling: none

In general, fast and relatively accurate trials, around 7 out of 10 trials are within tolerance of 1000uL. The time taken for transfer is similar to the average time taken for human driven section trials (standard calibration).

Observation 3: LIN - scaling: multiply

Slow transfer penalization seems to work on LIN. Overall short transfer time, not very accurate as seen in the data, only 5 to 6 out of 10 trials are within tolerance of 1000uL.

Observation 4: LIN - scaling: none

Accuracy is similar to LIN - multiply but much longer trials on average, a few trials even reached 200 seconds even though the standard calibration indicates around 80 seconds. However, at the last 5 trials, the model is able to discover sets of parameters that give very accurate transfer even though the 5 sets of parameters are very different, time ranges from 56s to 200s. (This shows that this iteration is prioritising low percentage error instead of shortening transfer times) Thus, one can suspect, in the long run this model can provide interesting parameters that are accurate.
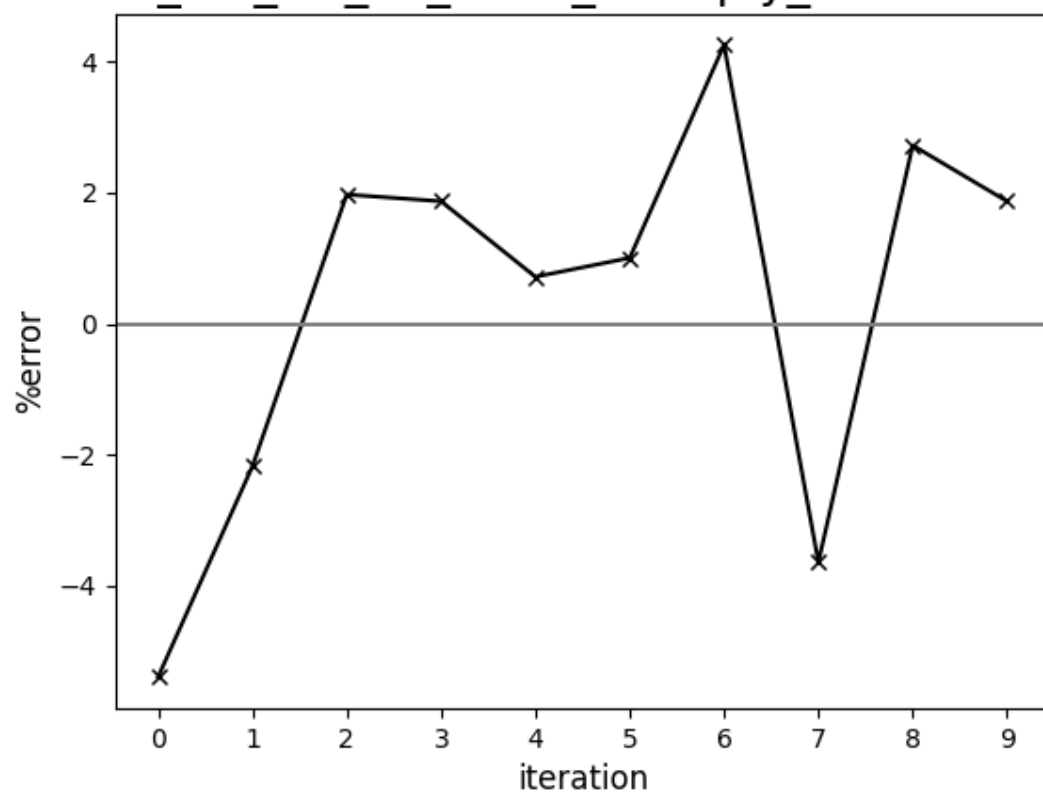
GPR_full_wo_bo_scale_multiply_unorderedT
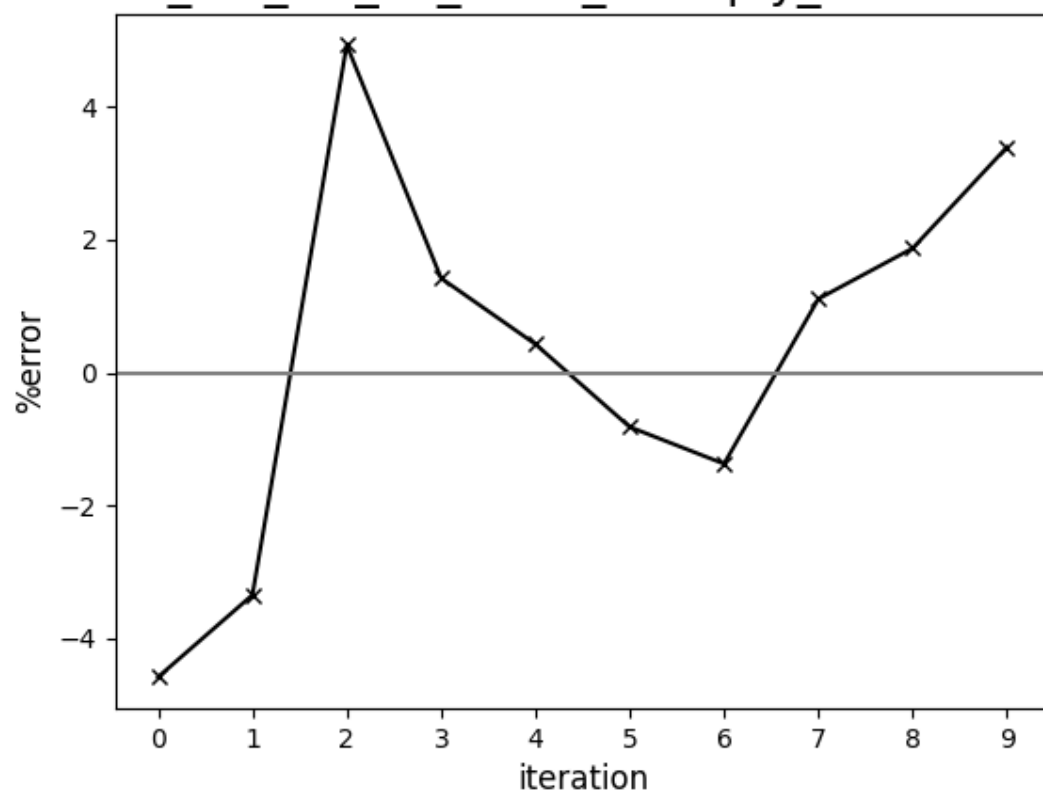
GPR_full_wo_bo_scale_none_unorderedT
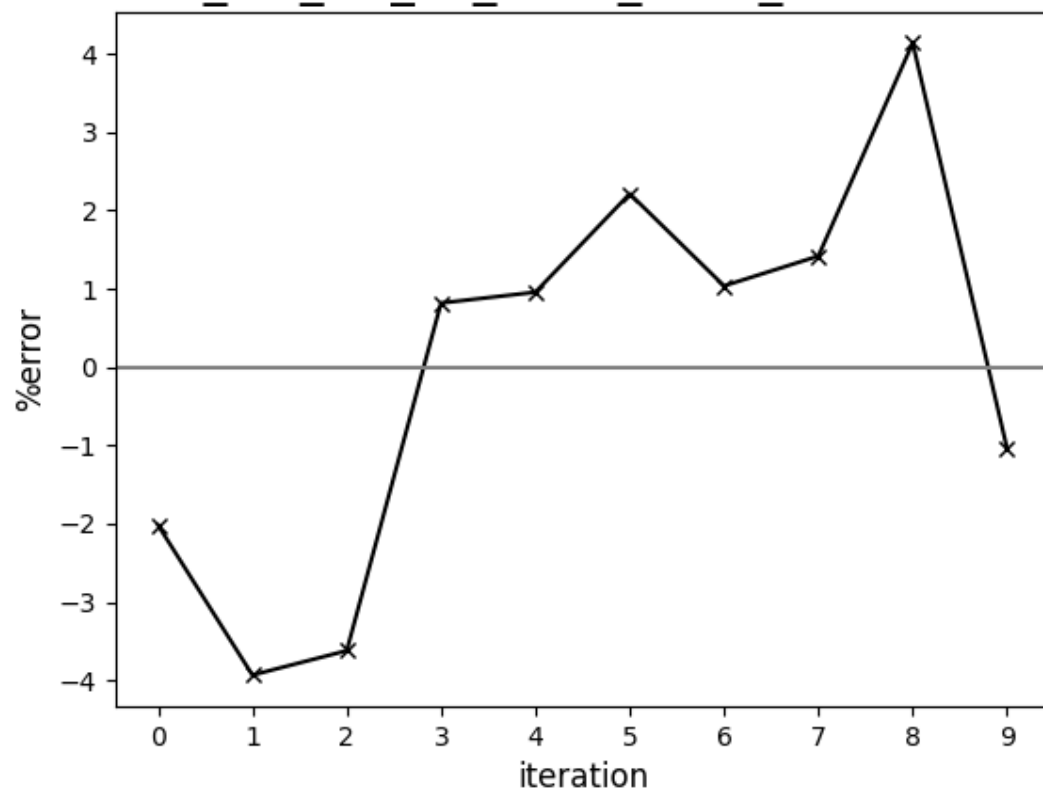
LIN_full_wo_bo_scale_multiply_unorderedT
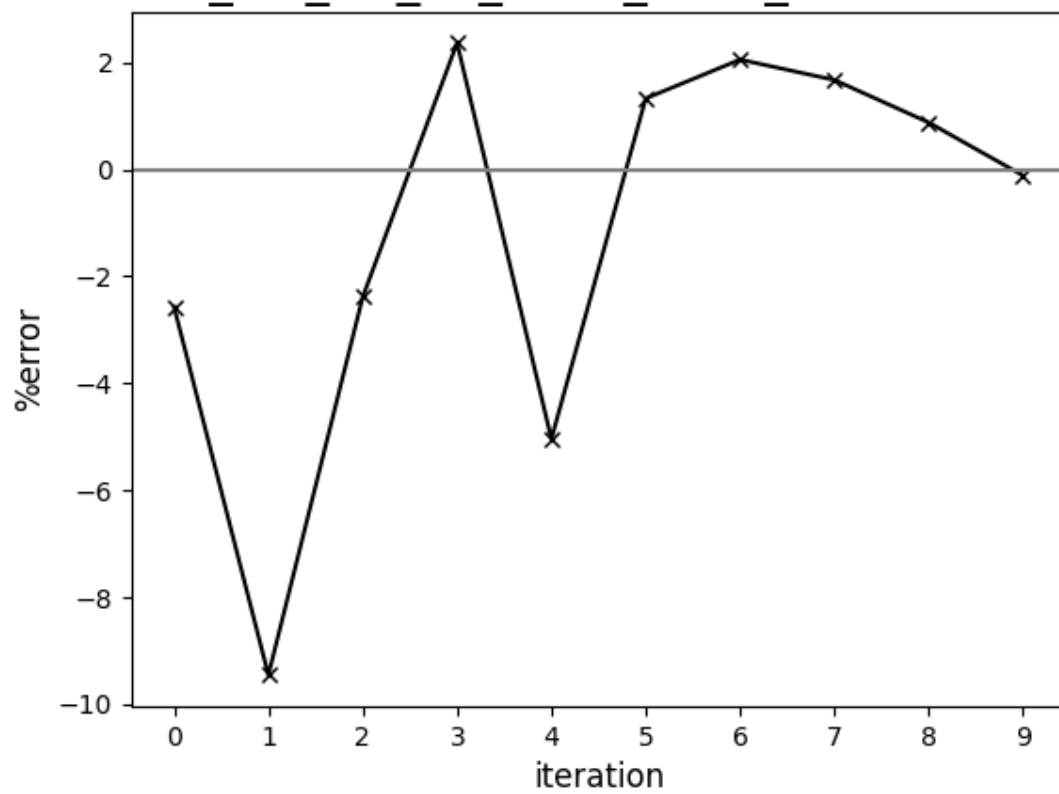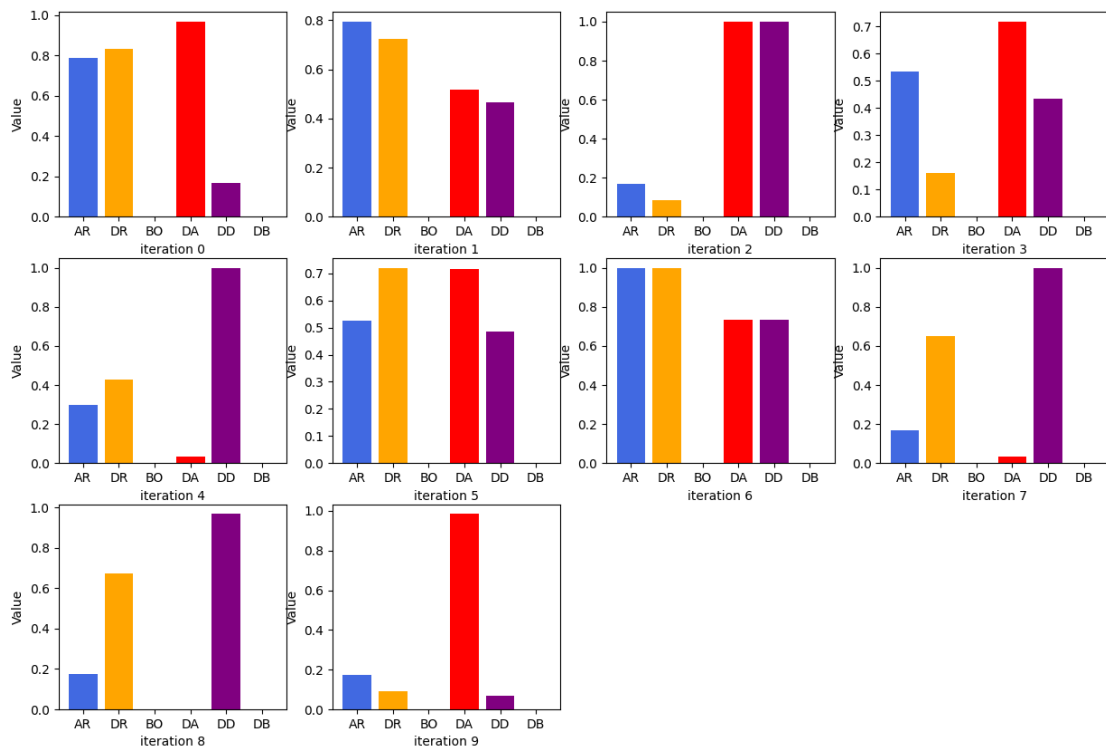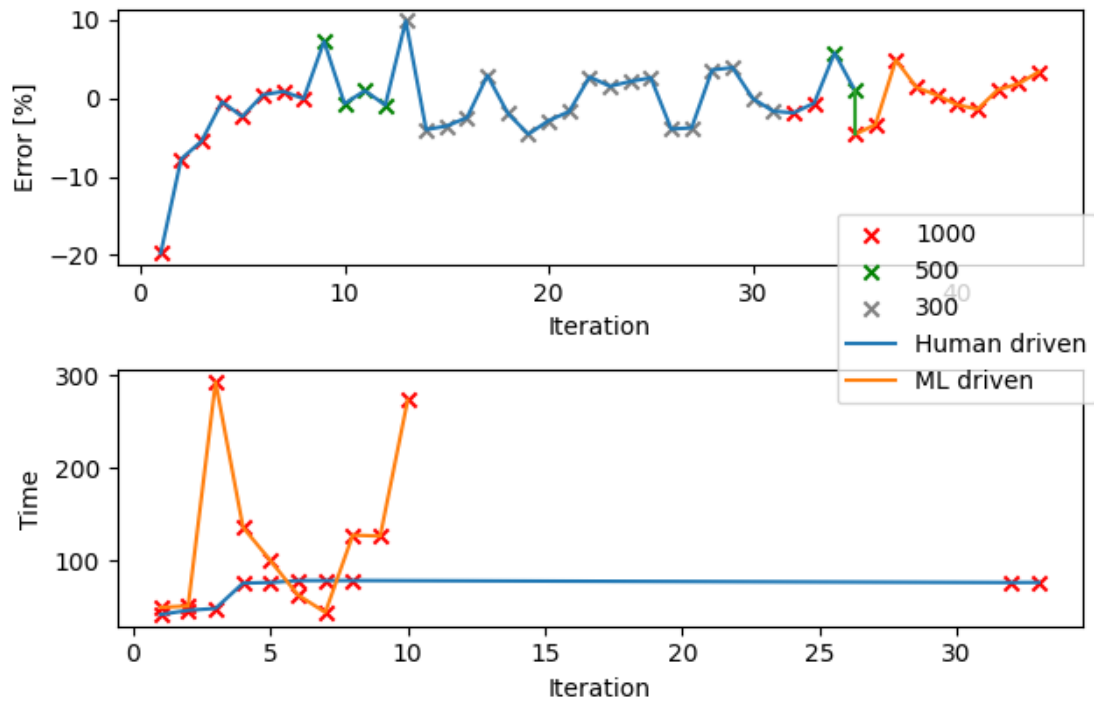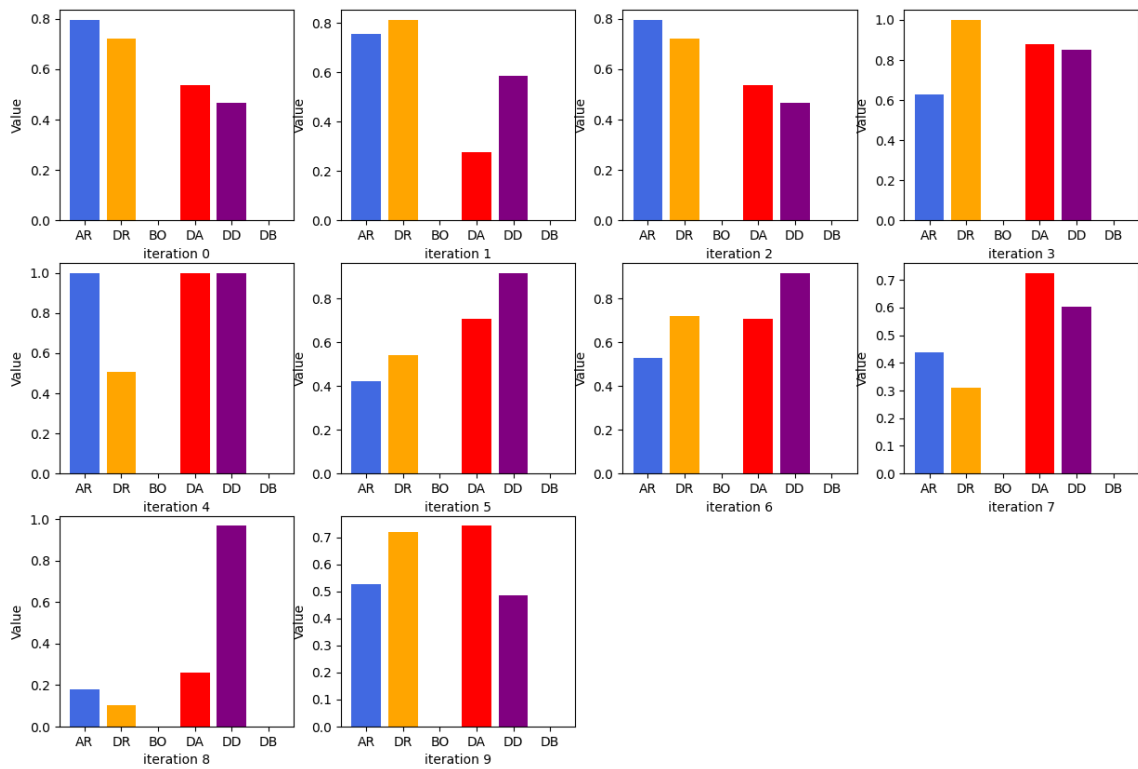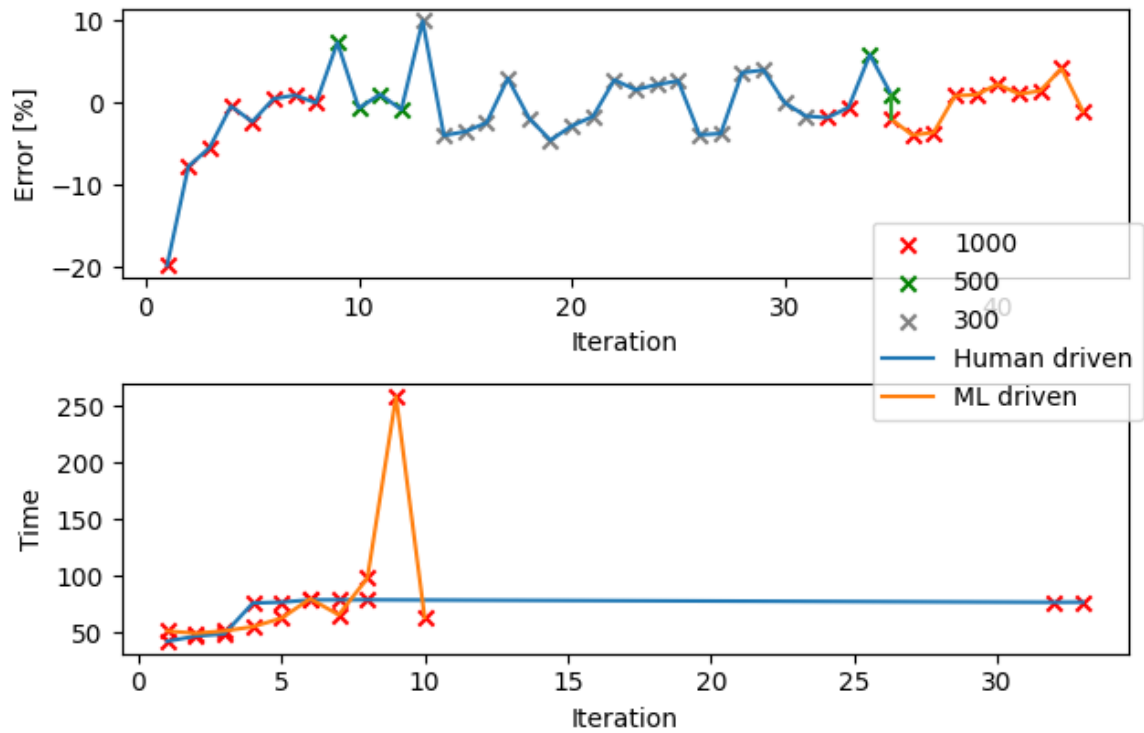
LIN_full_wo_bo_scale_none_unorderedT

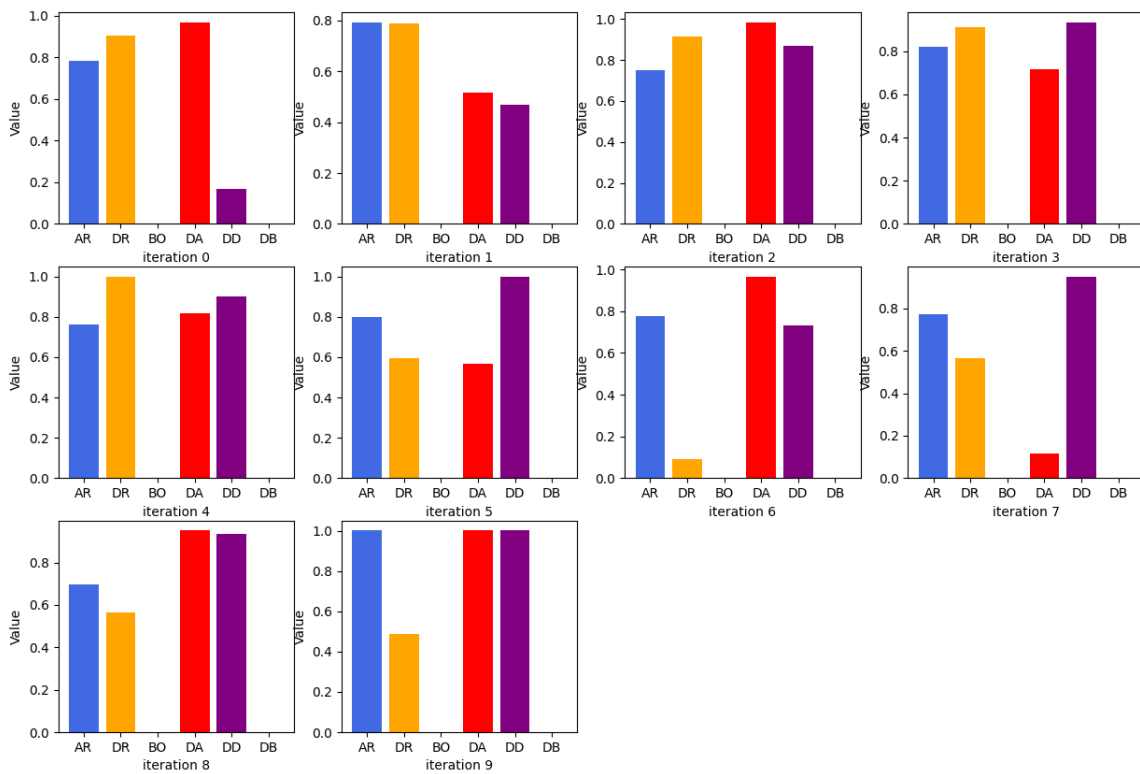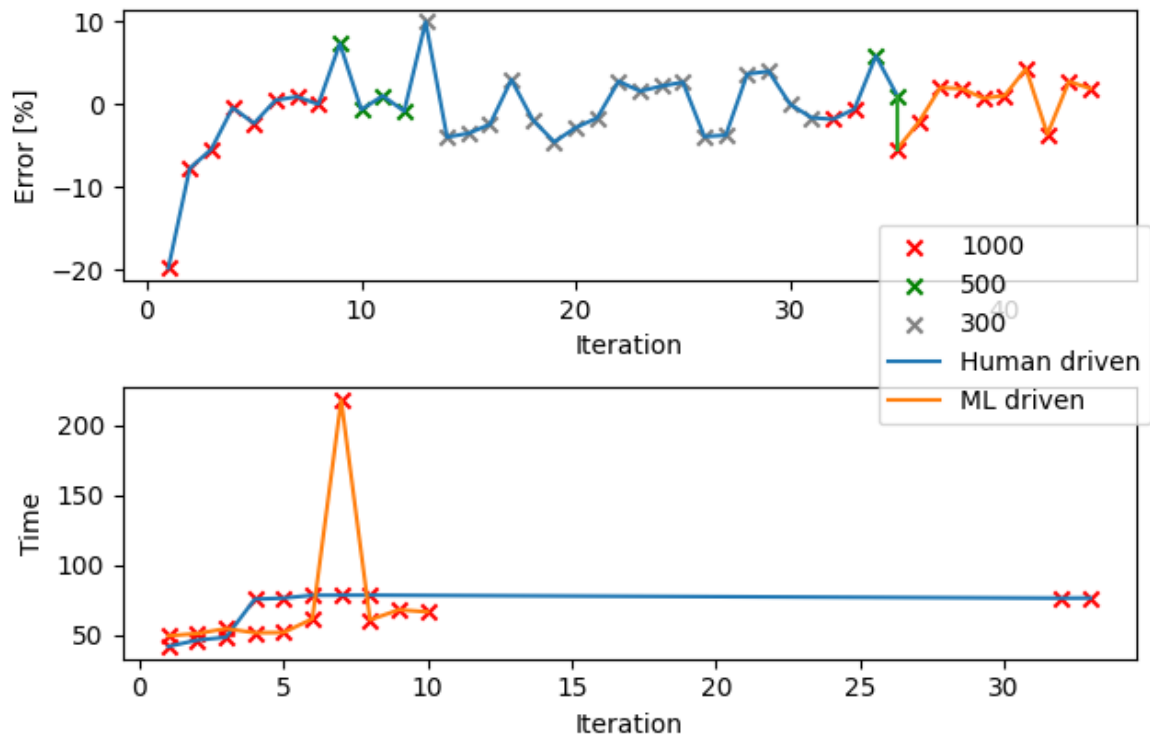# Diagram comparing human-driven and ML test trials (398 cP, set:half, unorderedT)

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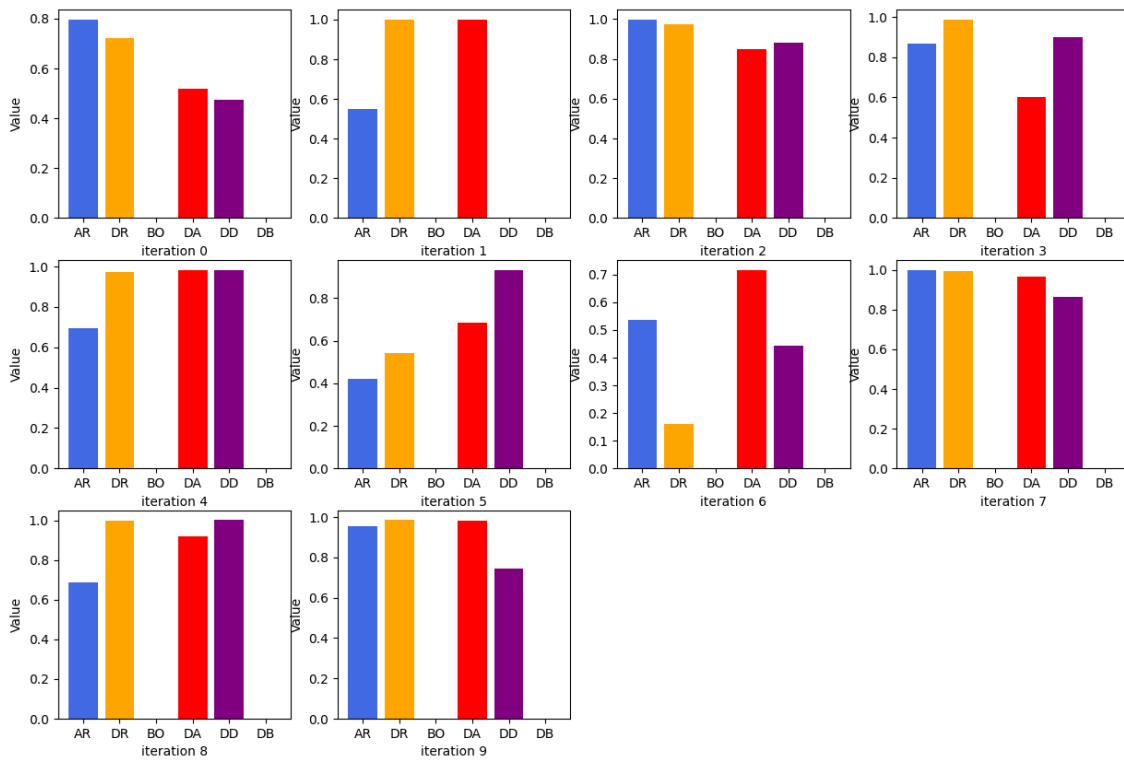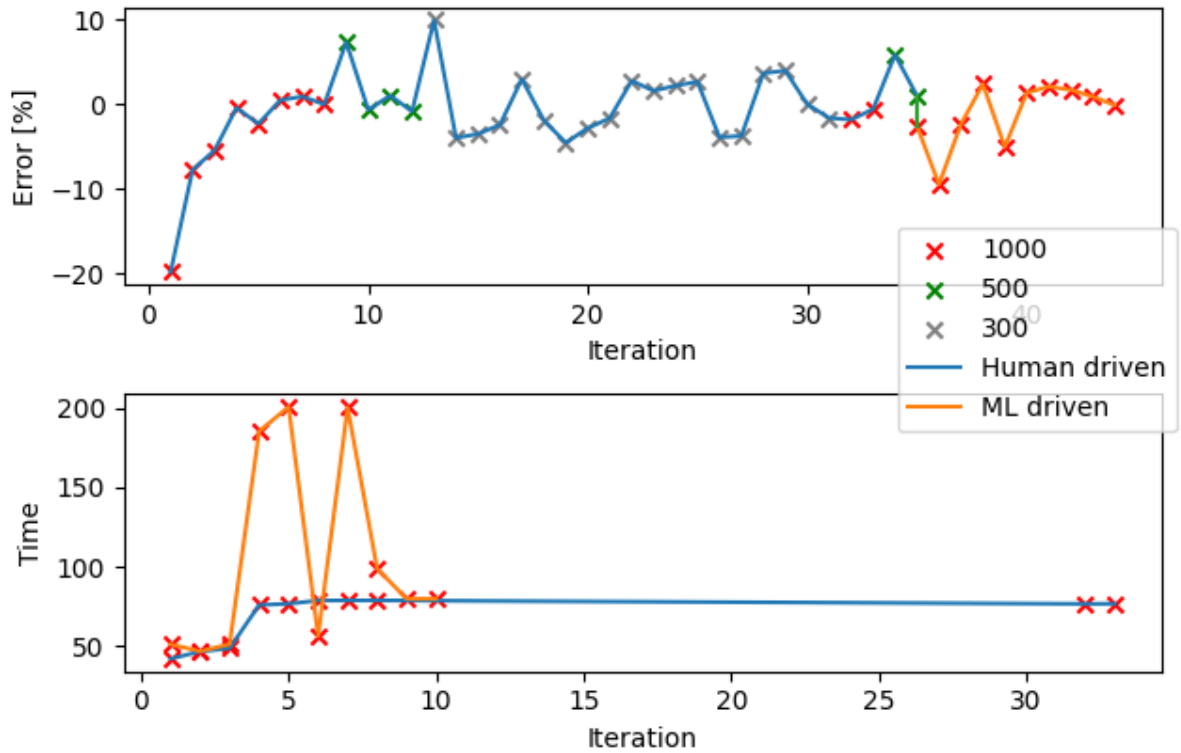
Gaussian Process Regression model with no penalization, trained with full data set

Linear Regression model with slow transfer penalization, trained with full data set

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

Set: 1

<u>Observation of trends</u>

Observation 1: LIN - scaling: Without and LIN - scaling: Multiplication

Once an 'optimal' set of parameters are attained, it will double check once or twice to confirm, then continue to explore parameters. The 'confirmation' stage usually occurs in trial 5 or 6 of the 10 trials and it can continue to suggest parameters that deviate quite far from the 'optimal' set of parameters.

The error percentage fluctuates slightly between -2% to 2%, which is within tolerance of 1000uL. The sets of parameters are the most random out of the 3 observations (but does not have very big percentage error)

Observation 2: GPR - scaling: Division and LIN - scaling: Division

The sets of parameters generated are all prioritising faster transfer duration, neglecting the percentage error generated. On average, the percentage errors are all around -11 to -12% for both LIN and GPR models. The parameters do change but very slightly, all targeting at faster transfer rate.

Observation 3: GPR - scaling: Without and GPR - scaling: Multiplication
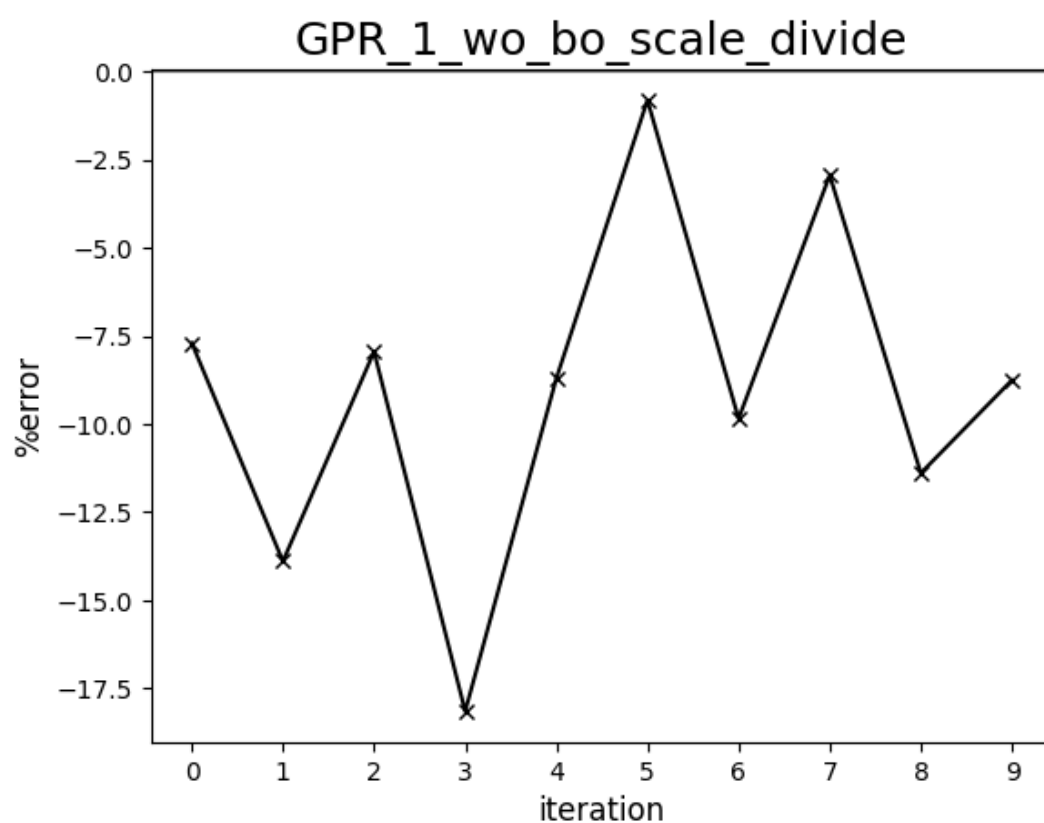
Both iterations show a lot of repetition in the parameters suggested. For scaling: without, the parameters are the same from trial 1 to trial 2 and also the same from trial 3 to trial 10. For scaling: without, the percentage error is an average of -3% while scaling: multiplication has percentage error of -2%. The sets of parameters suggested for GPR scaling: without and multiplication can be said to be almost identical.

Repetition is commonly observed for trials that use the GPR model, this pattern is consistent for both set: 1 and set: half.

It is worth noting that there are some sets of parameters that are repeated for many different iterations. It might help in determining the optimal set of data in the future. A few of such set of parameters are: (AR,DR,AD,DD)
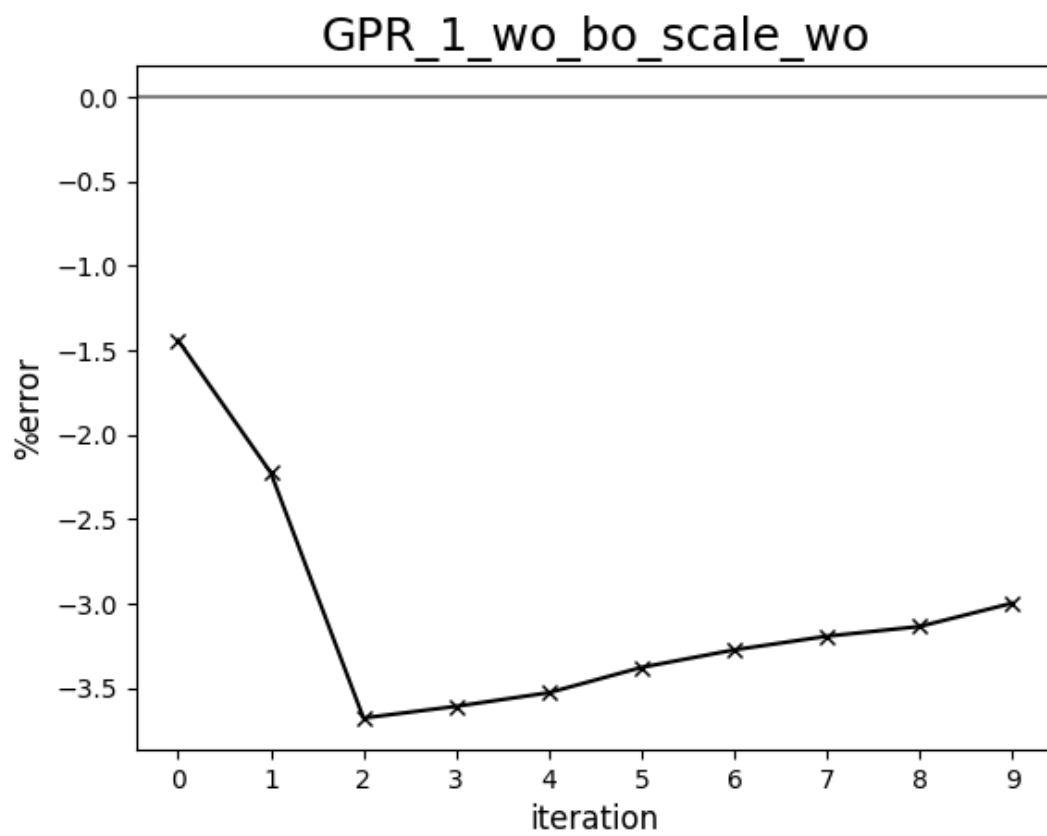- 52.3, 47.5, 3.1, 2.8
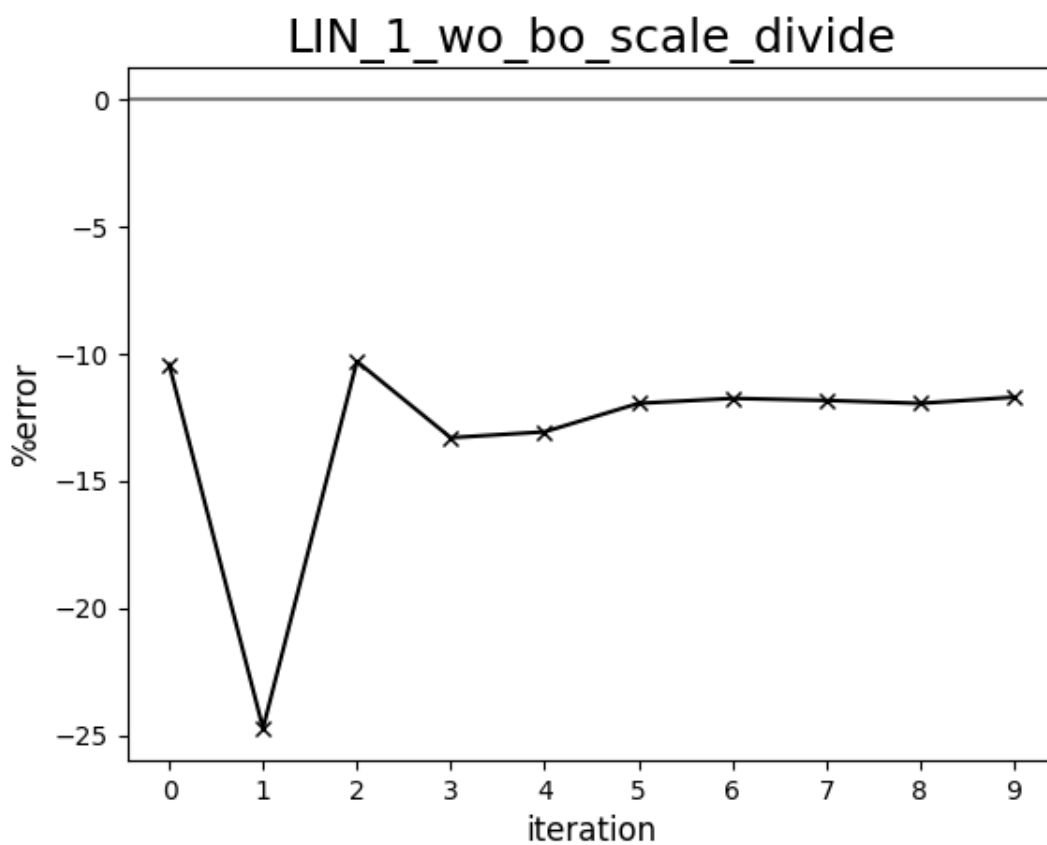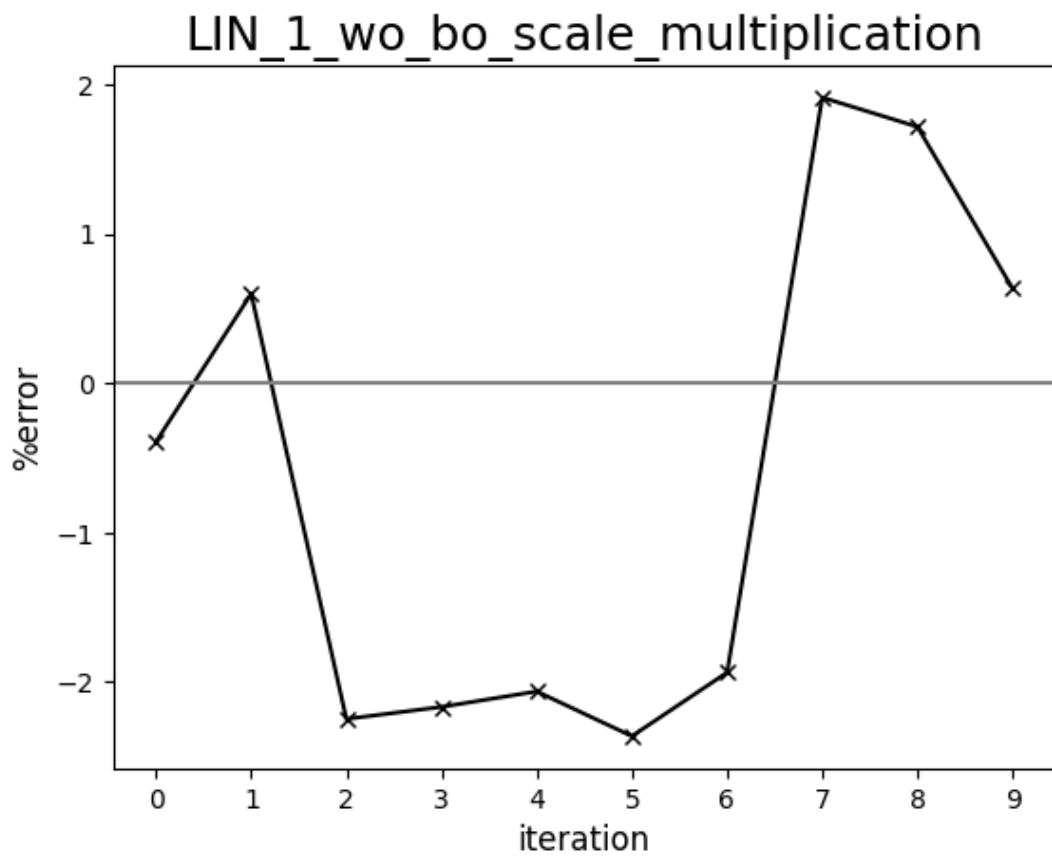- 66.0, 66.0, 0.0, 6.0

GPR scaling: divide

GPR_1_wo_bo_scale_divide

GPR scaling: multiplication

GPR_1_wo_bo_scale_multiplication

GPR scaling: without

GPR_1_wo_bo_scale_wo
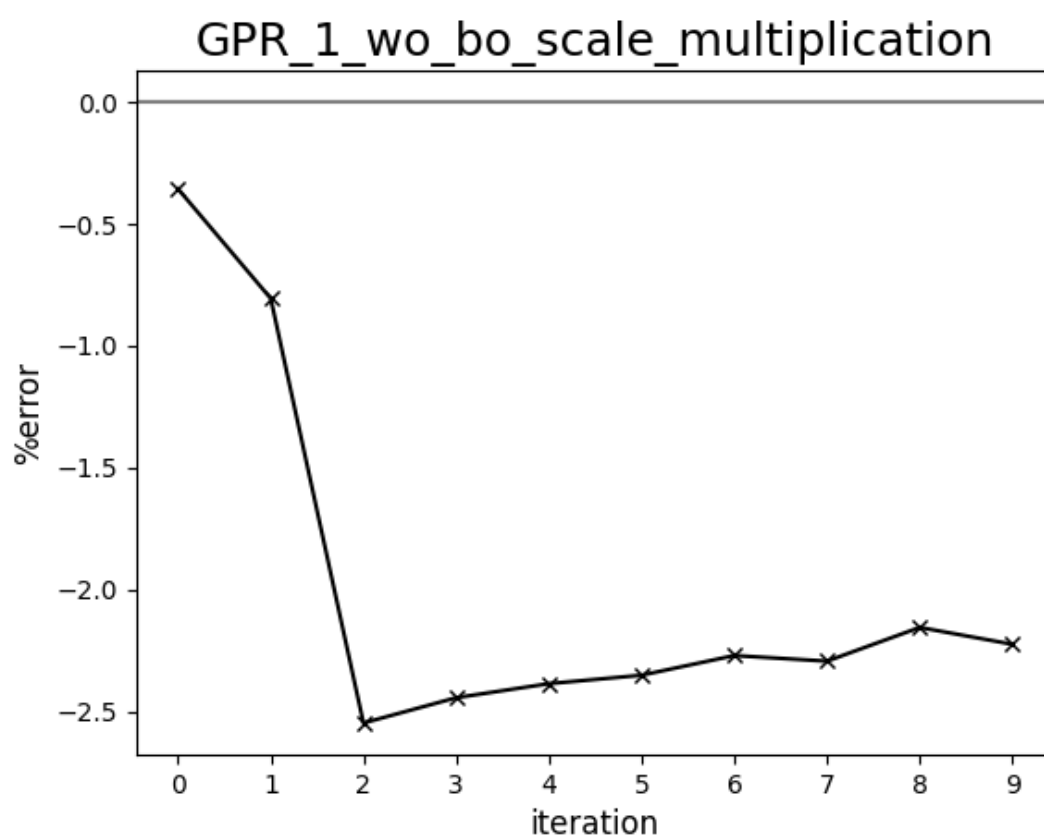
LIN scaling: divide

LIN_1_wo_bo_scale_divide
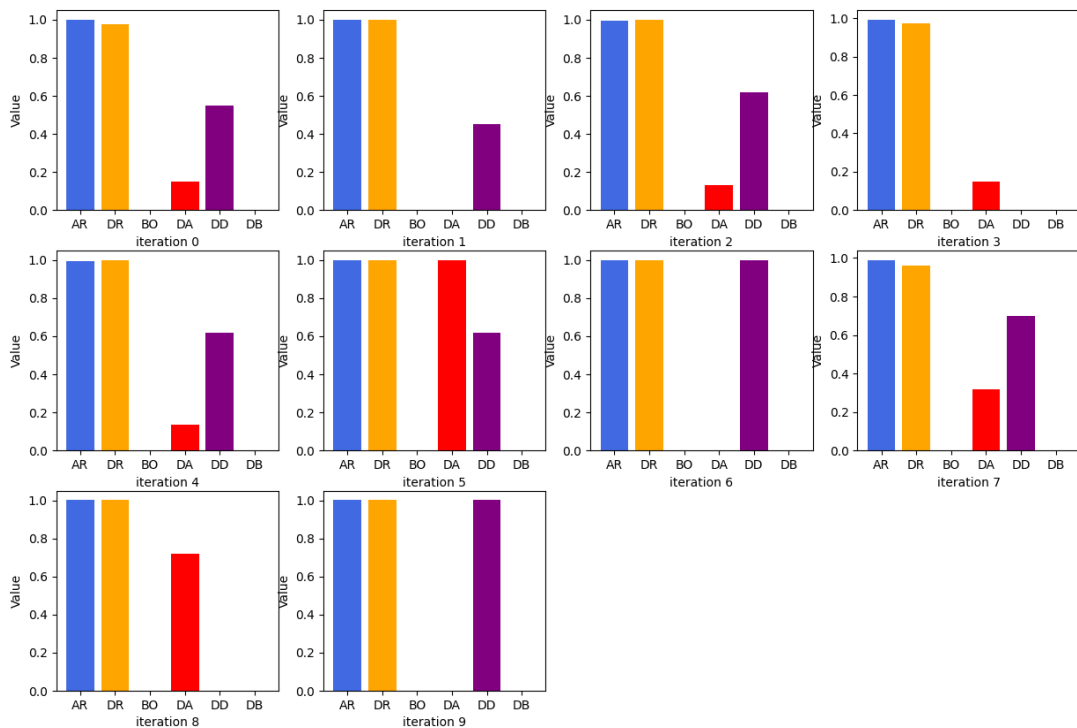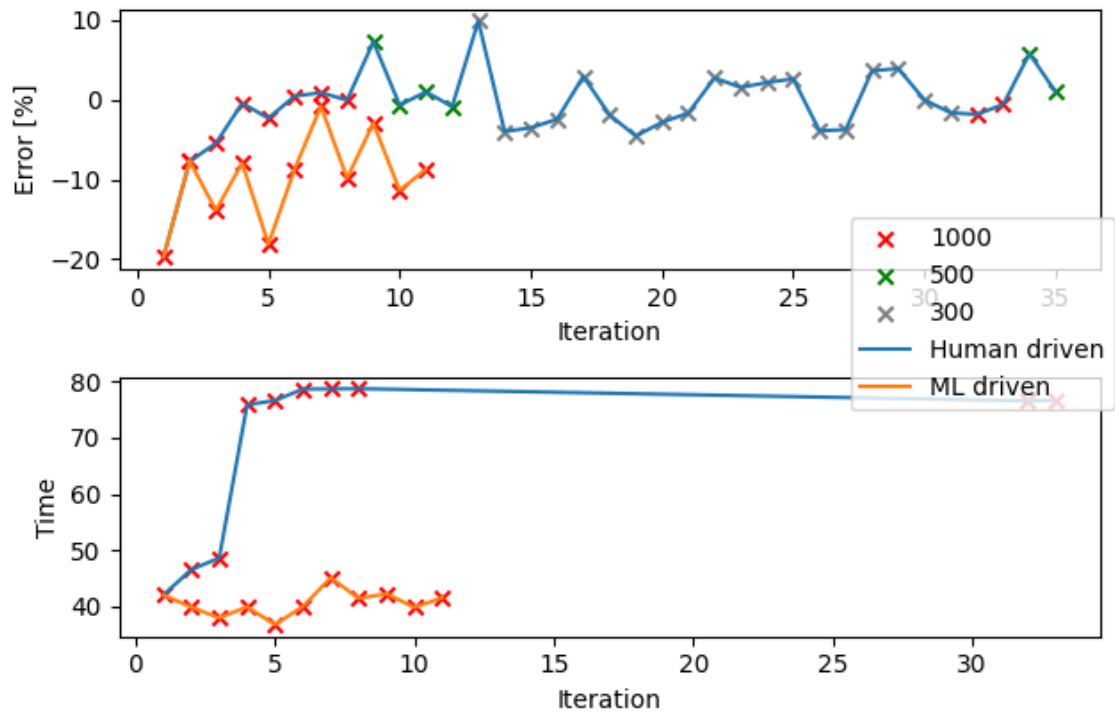
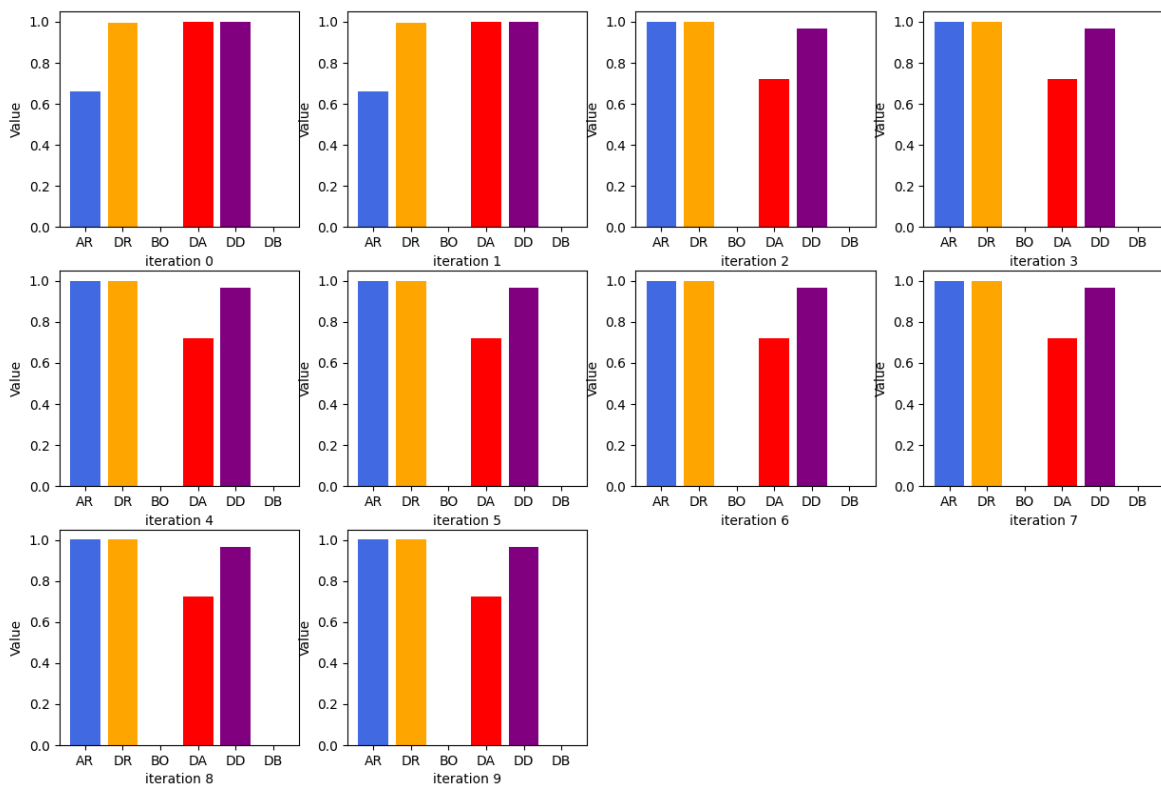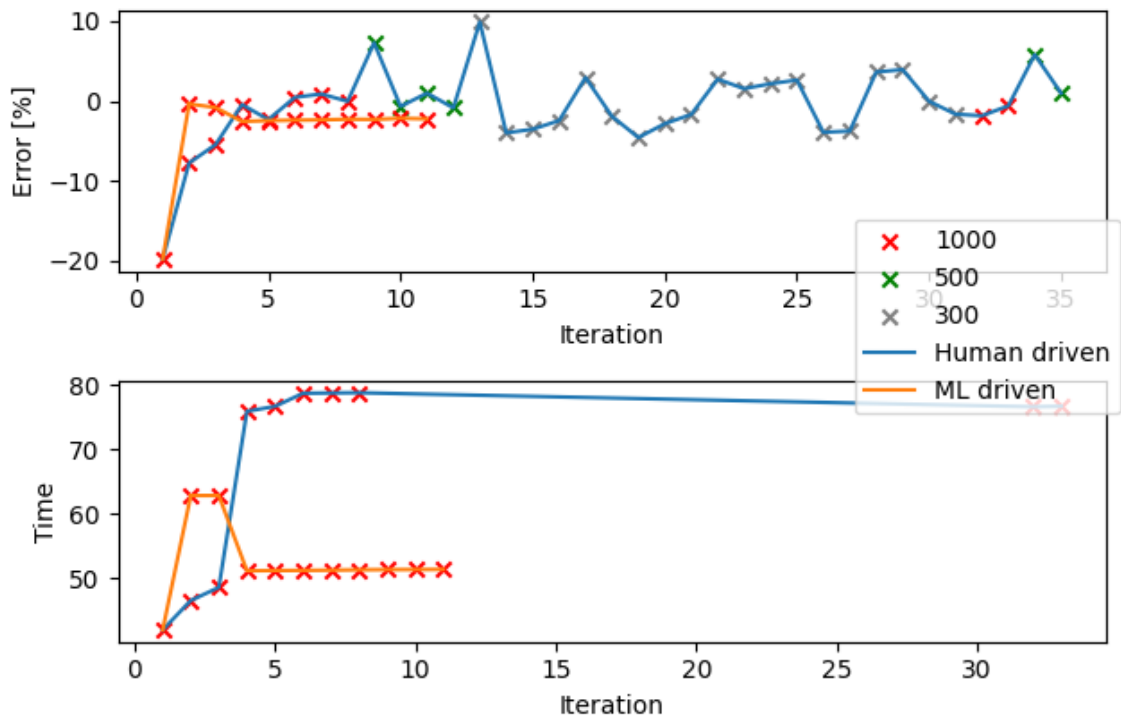LIN scaling: multiplication

LIN scaling: without
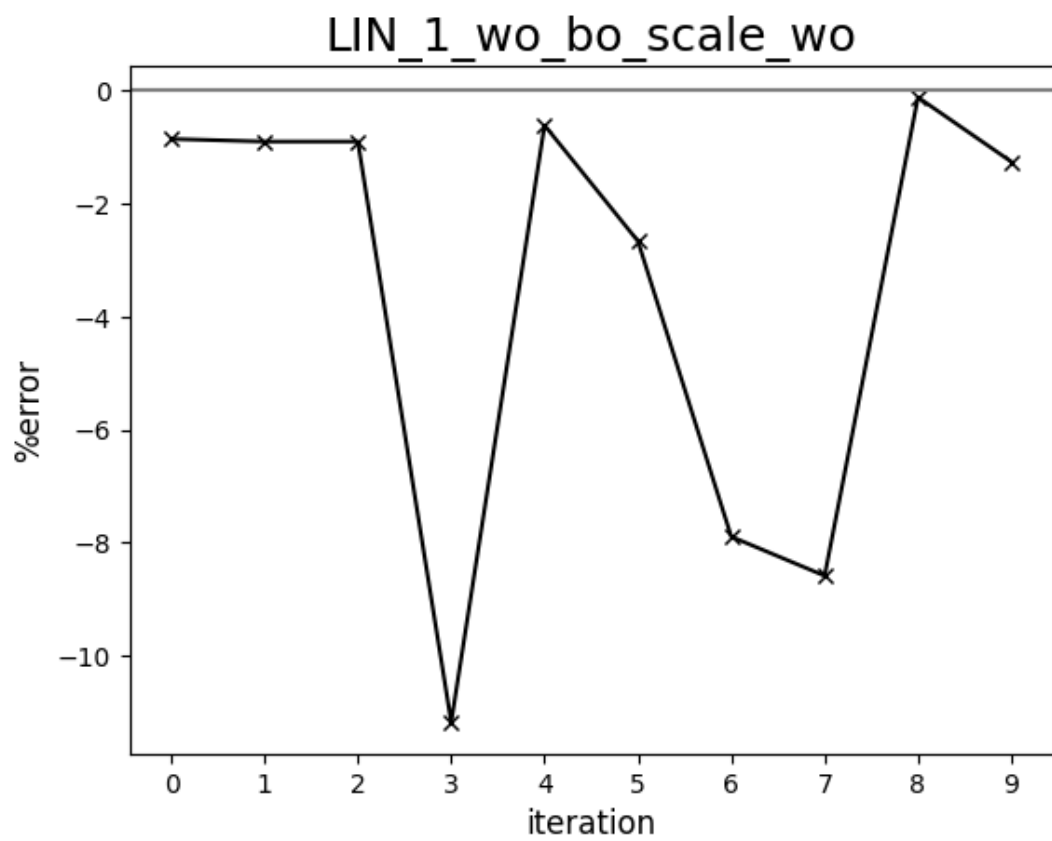
Diagram comparing human-driven and ML test trials (398 cP, set:1)

Gaussian Process Regression model with fast transfer penalization, trained with 1 initialization data
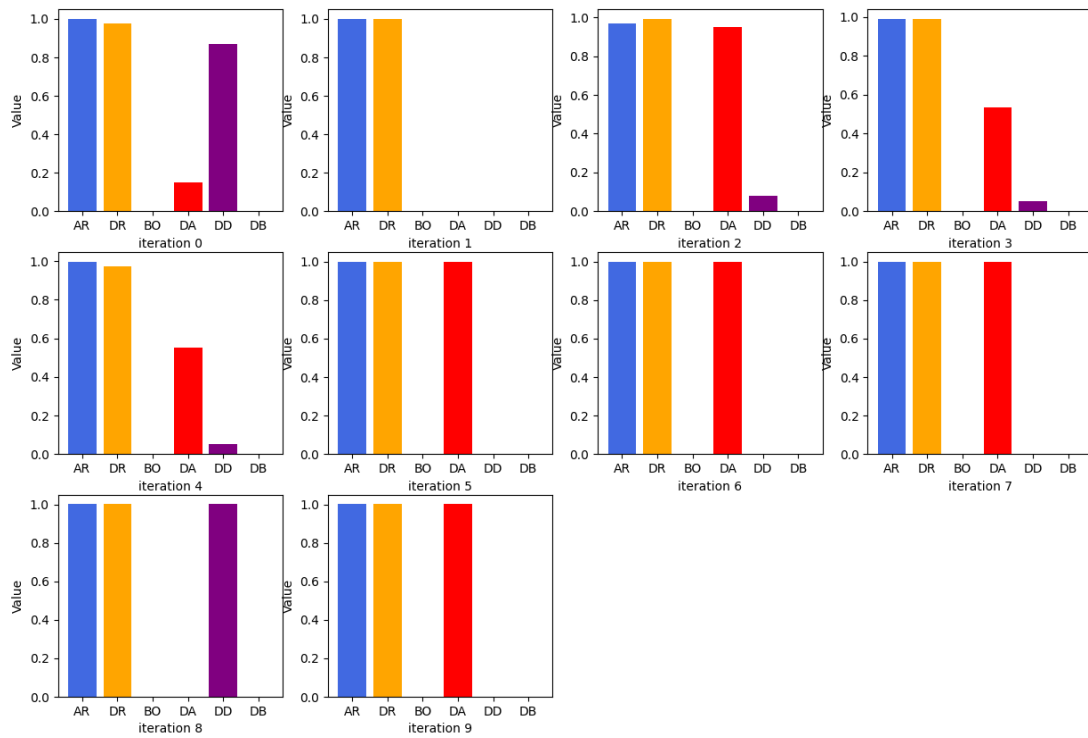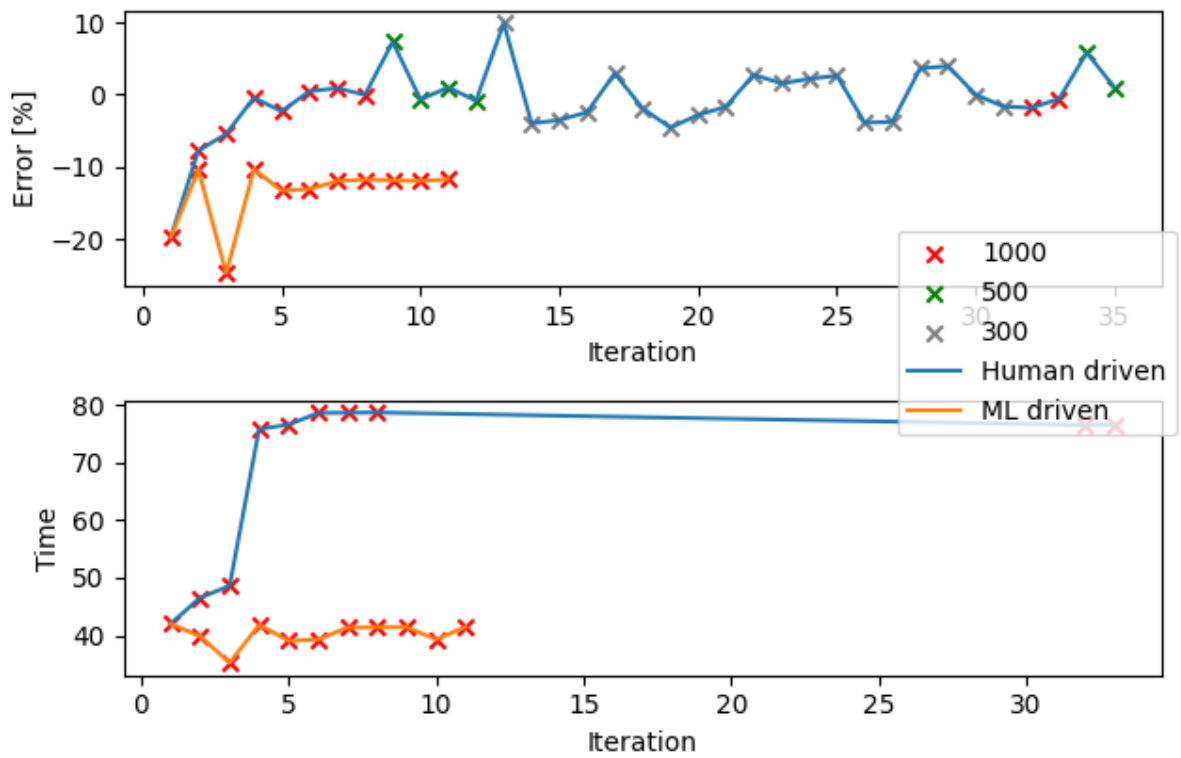
Gaussian Process Regression model with slow transfer penalization, trained with 1 initialization data
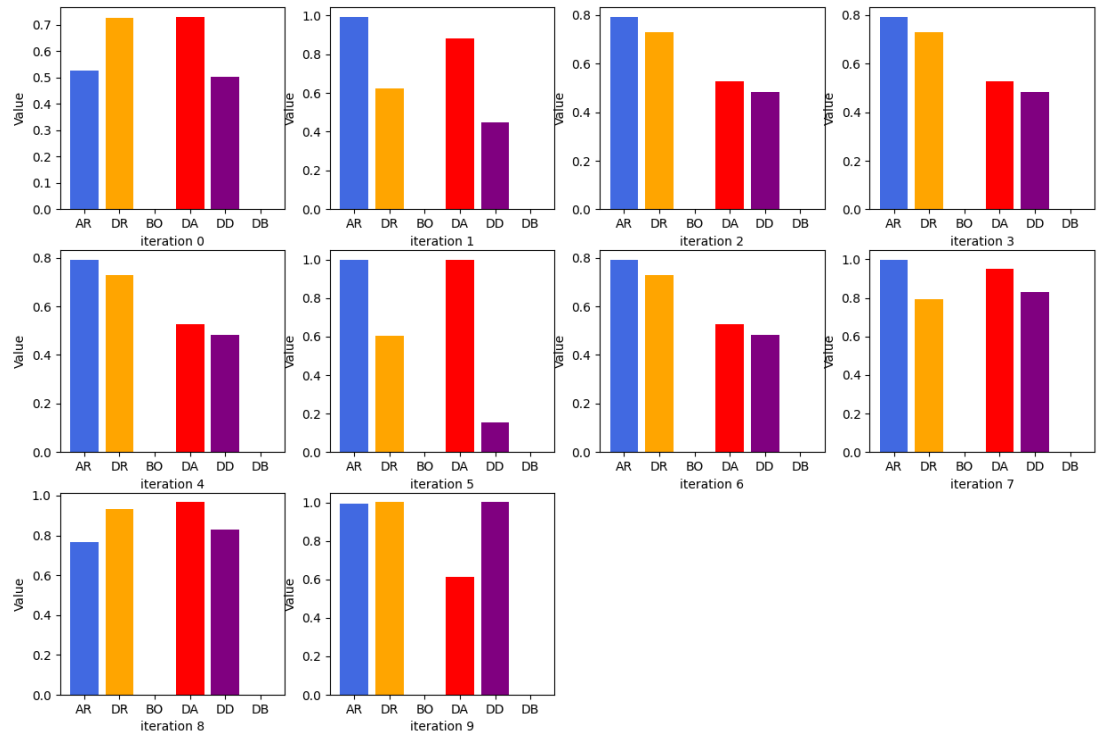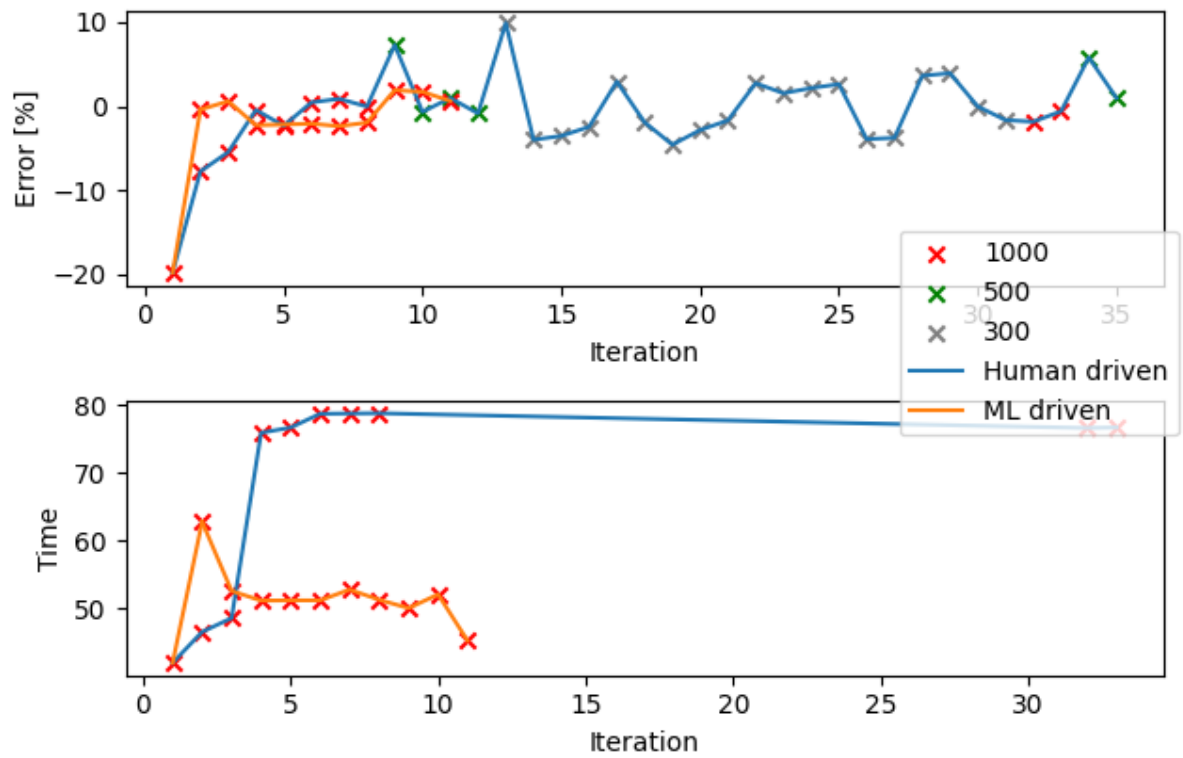
Gaussian Process Regression model with no penalization, trained with 1 initialization data
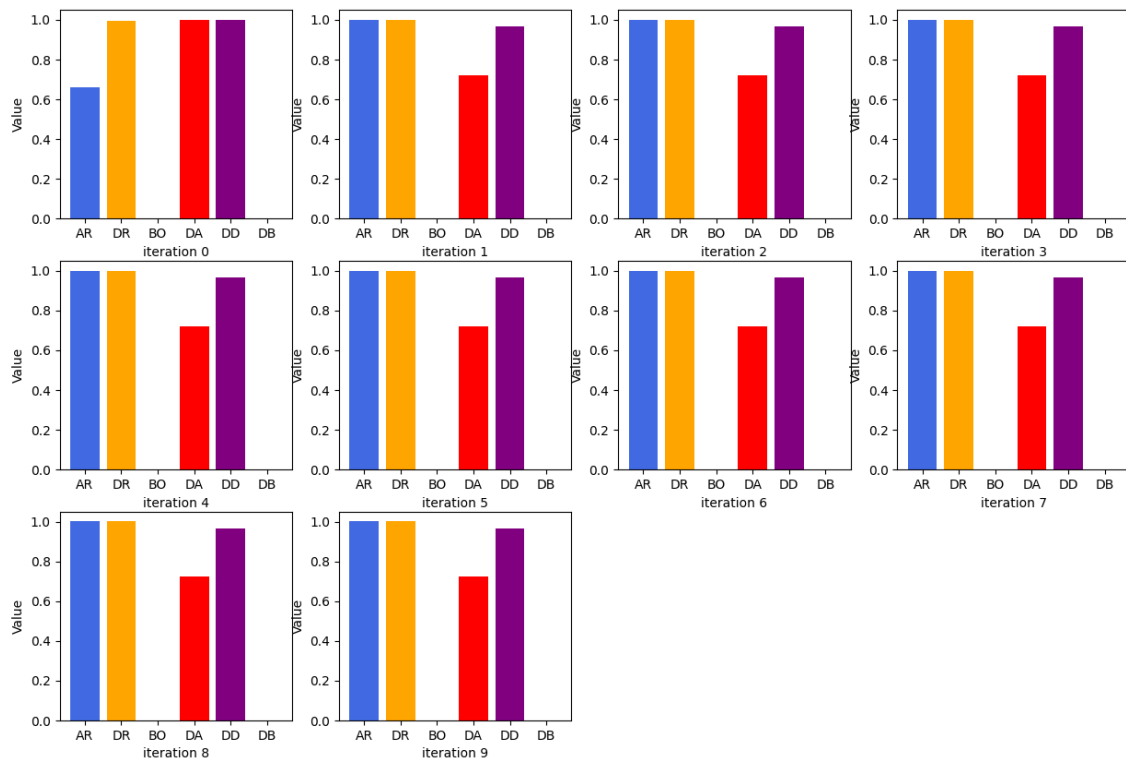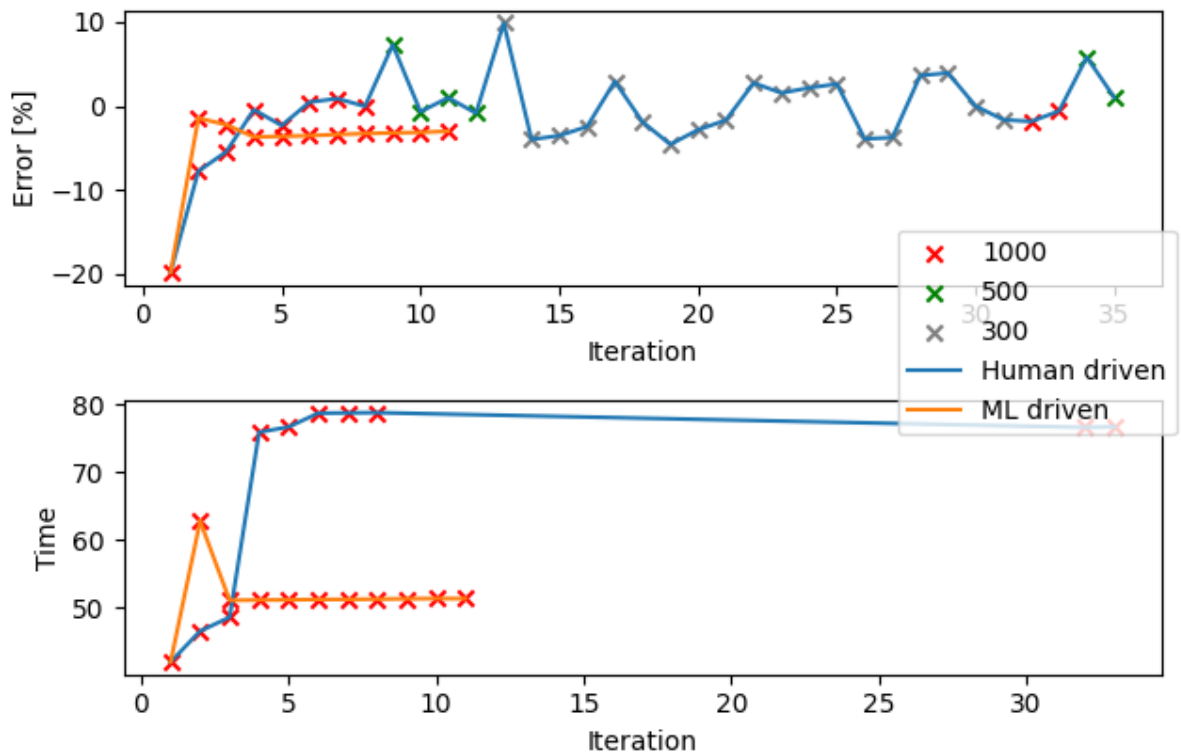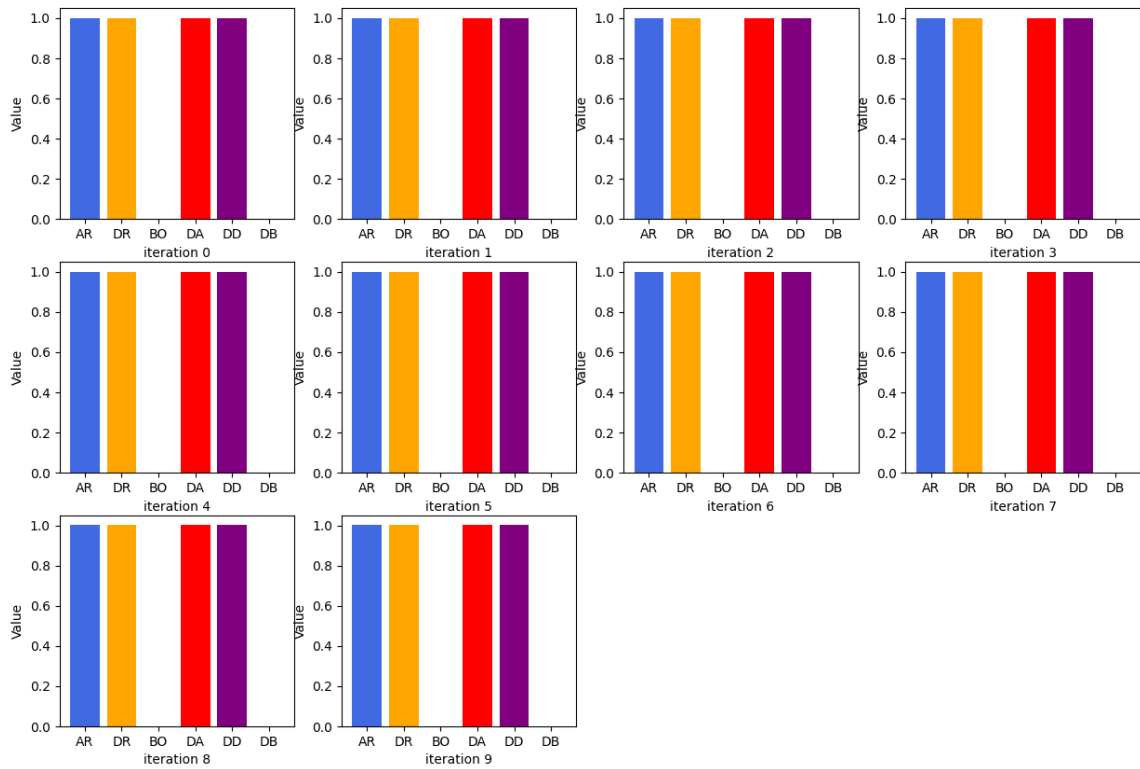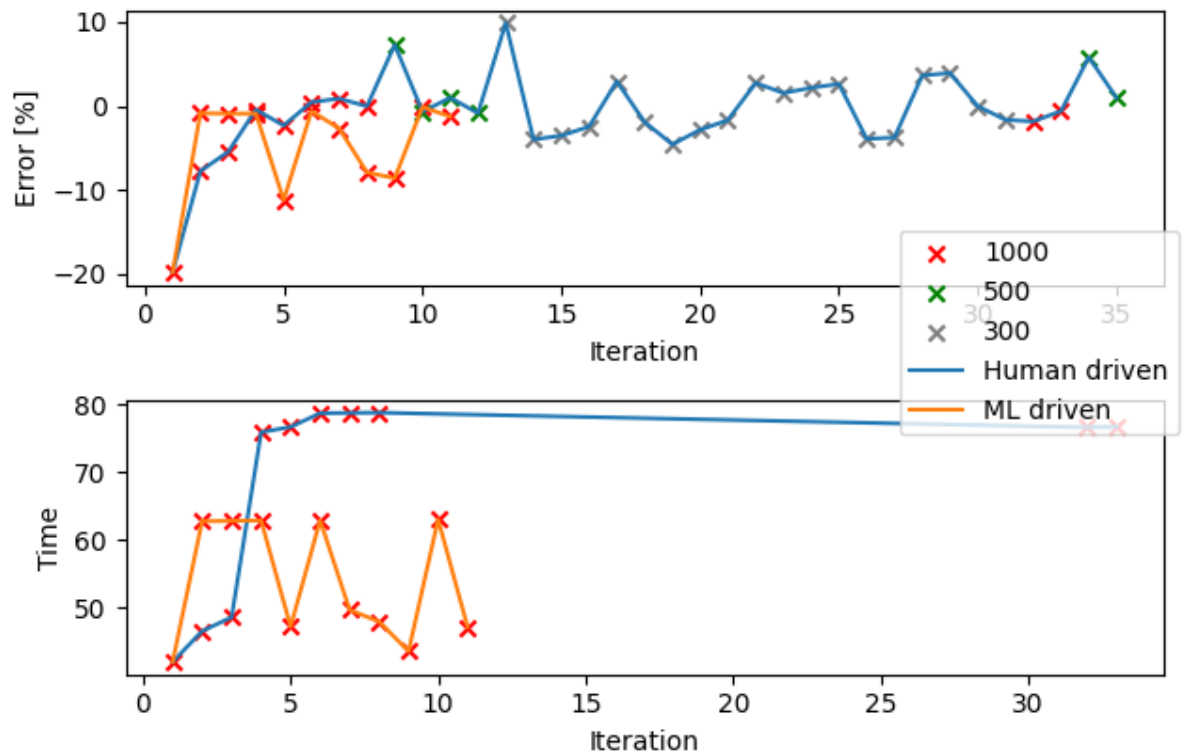
Linear Regression model with fast transfer penalization, trained with 1 initialization data

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

Linear Regression model with no penalization, trained with 1 initialization data
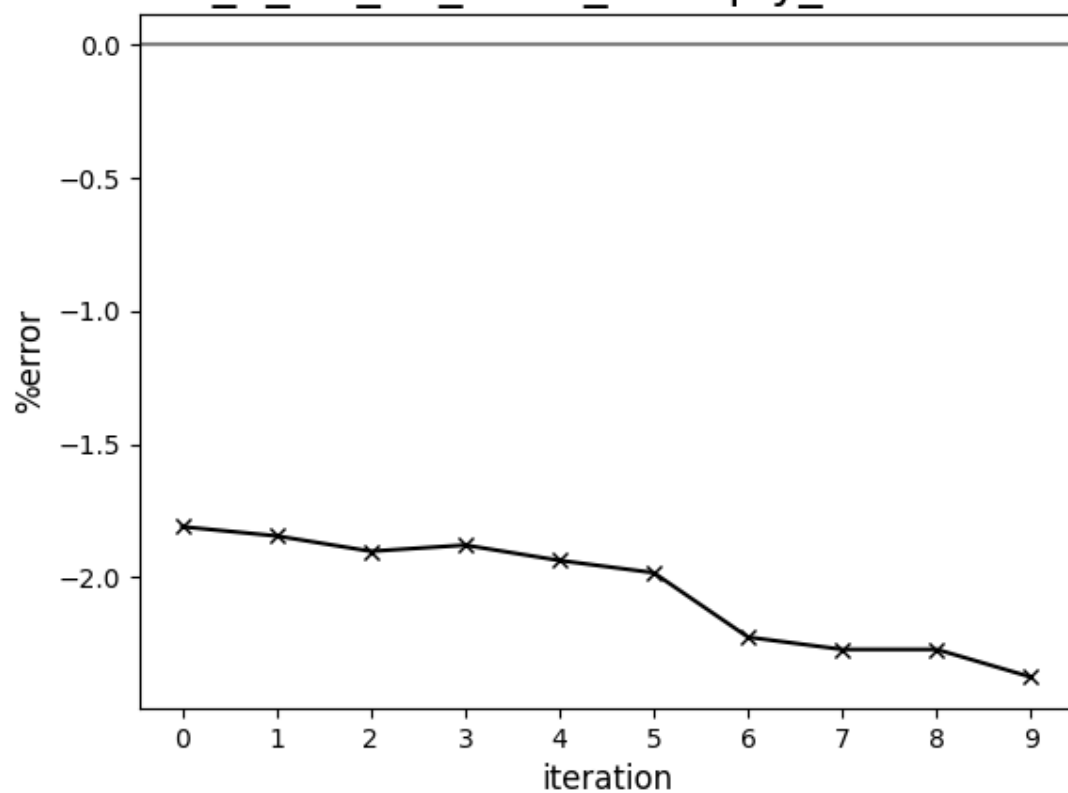
Set: 1 (unorderedT)

<u>Observation of trends</u>

All iterations

All 4 iterations have sets of parameters that are exactly the same. Basically, the only set of parameters used is: (AR, DR, DA, DD)
34.6, 47.3, 4.3, 2.9

Despite the fact that all 40 trials are identical, the transfers are quite accurate. The percentage error falls between -1 to -2%. All trials took around 62 to 63 seconds, as compared to 80s derived from the standard calibration.
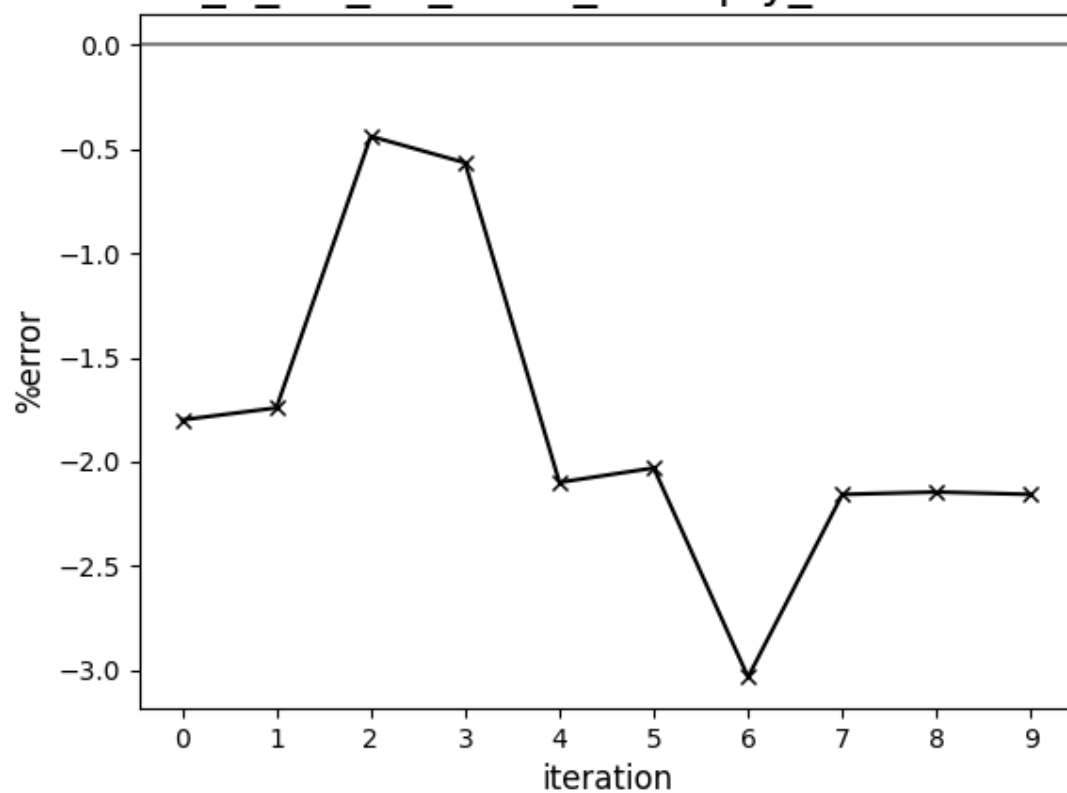
GPR_1_wo_bo_scale_multiply_unorderedT

GPR_1_wo_bo_scale_none_unorderedT
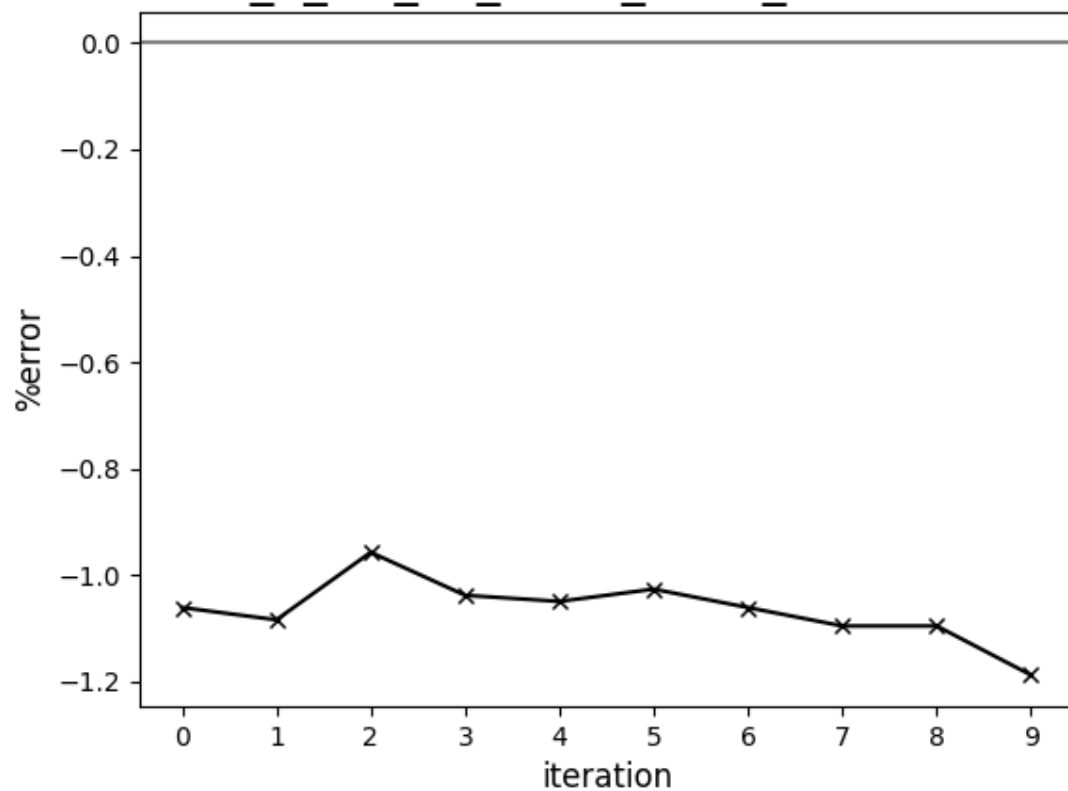
LIN_1_wo_bo_scale_multiply_unorderedT
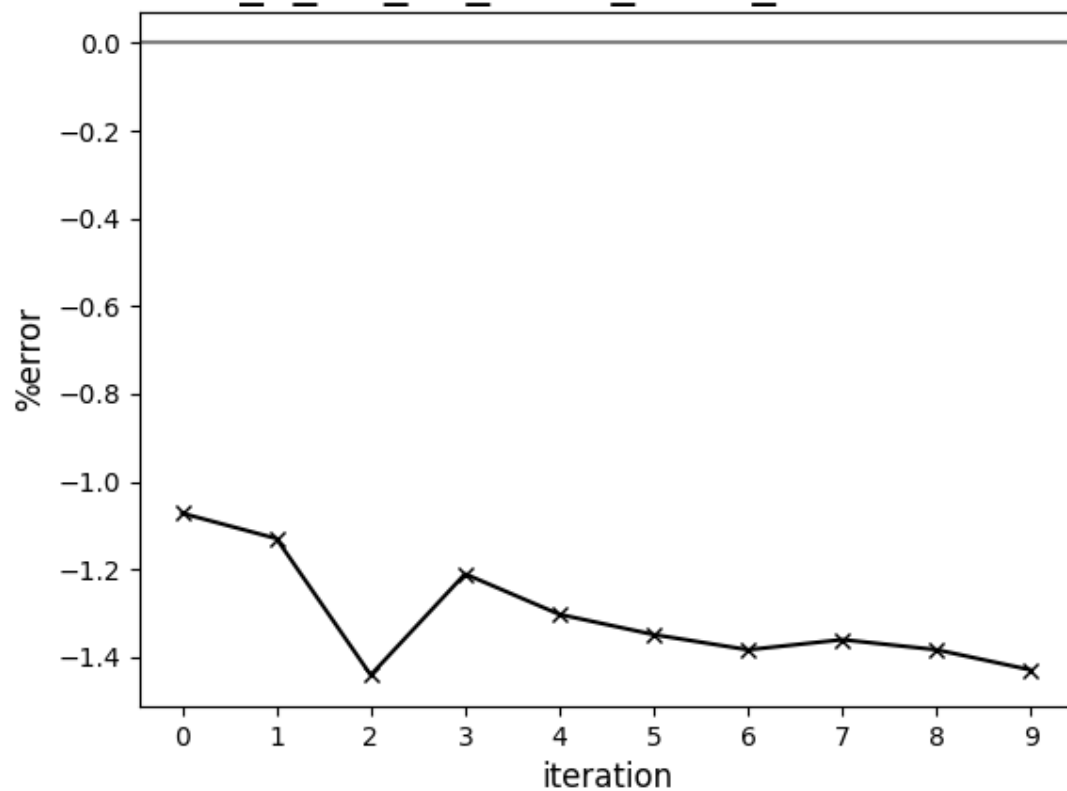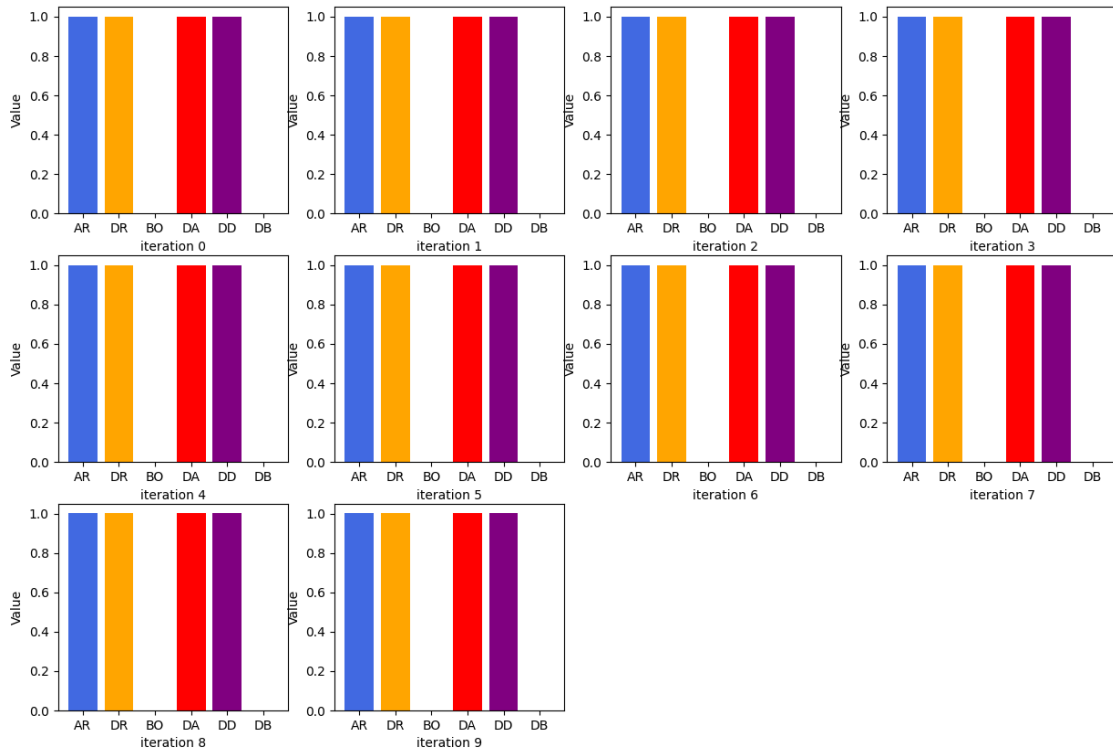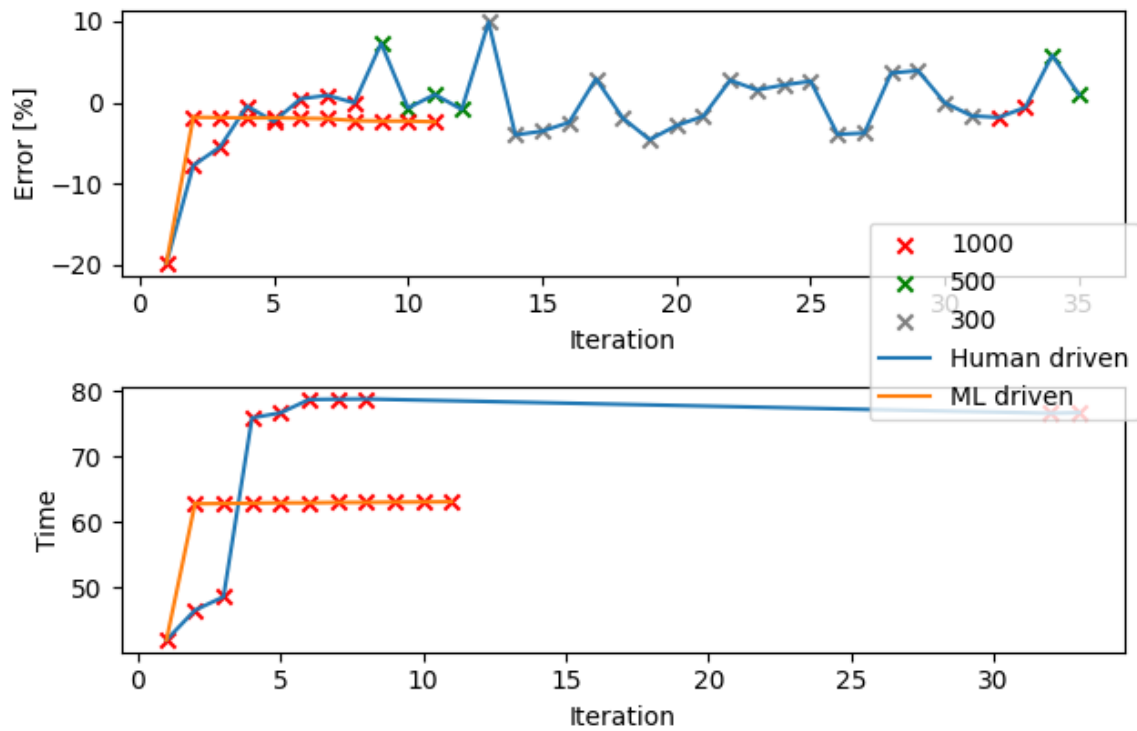
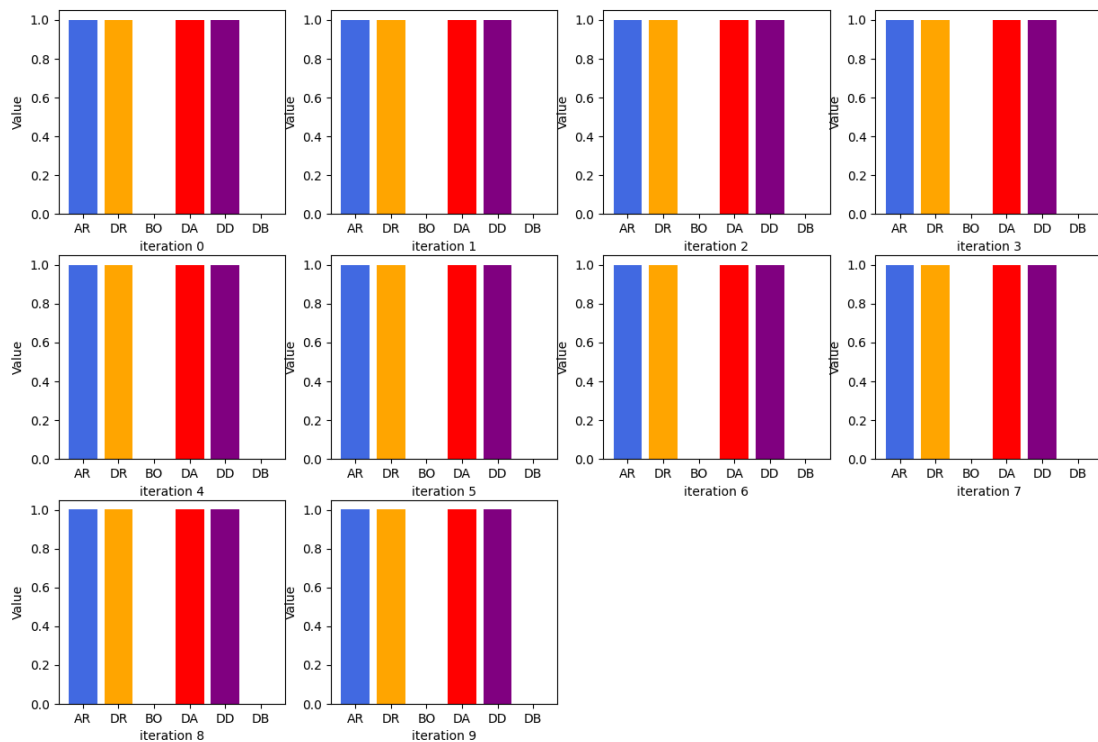

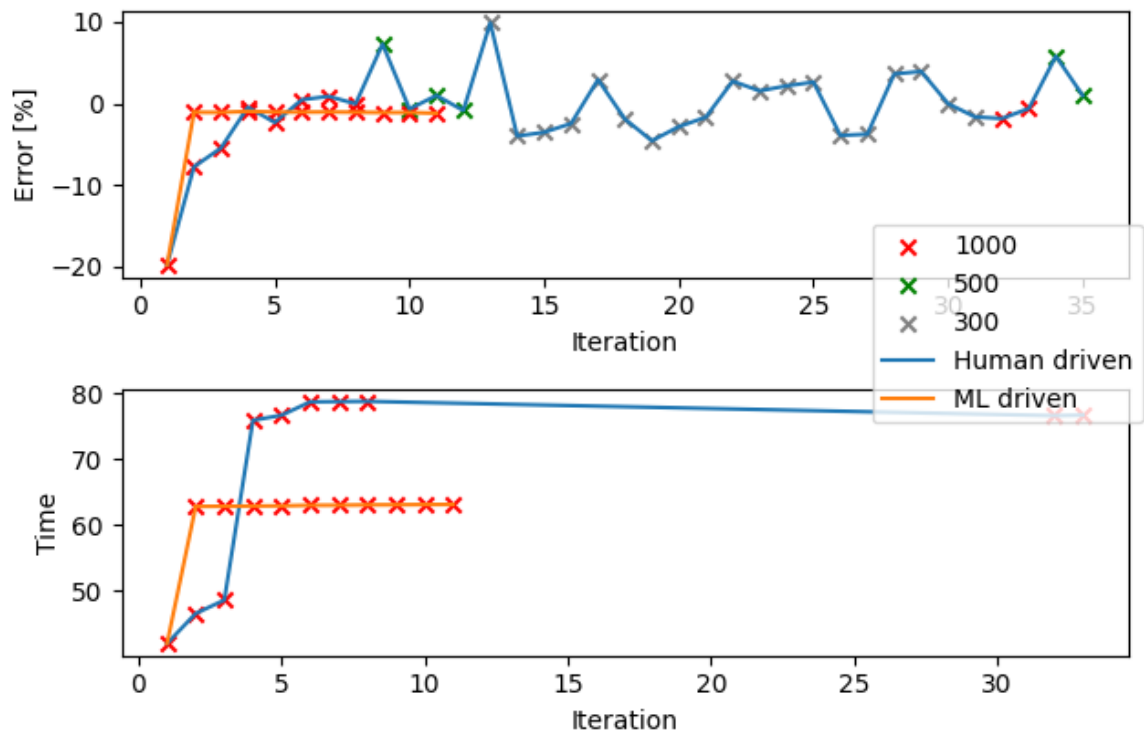# LIN_1_wo_bo_scale_none_unorderedT

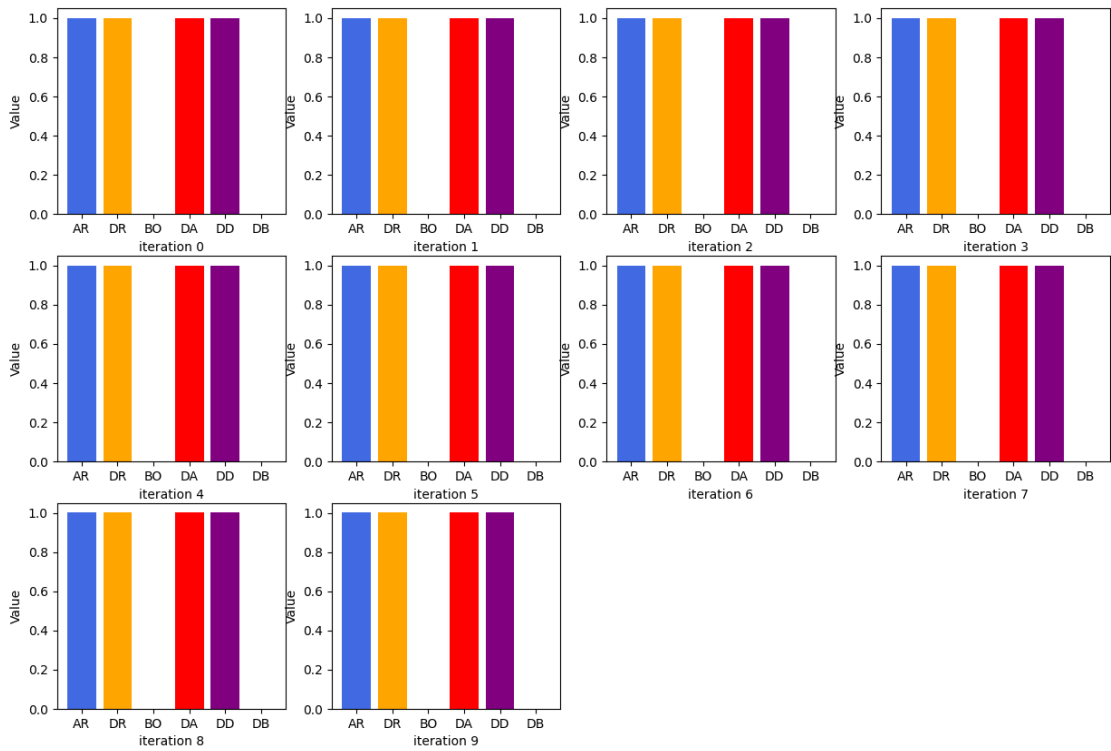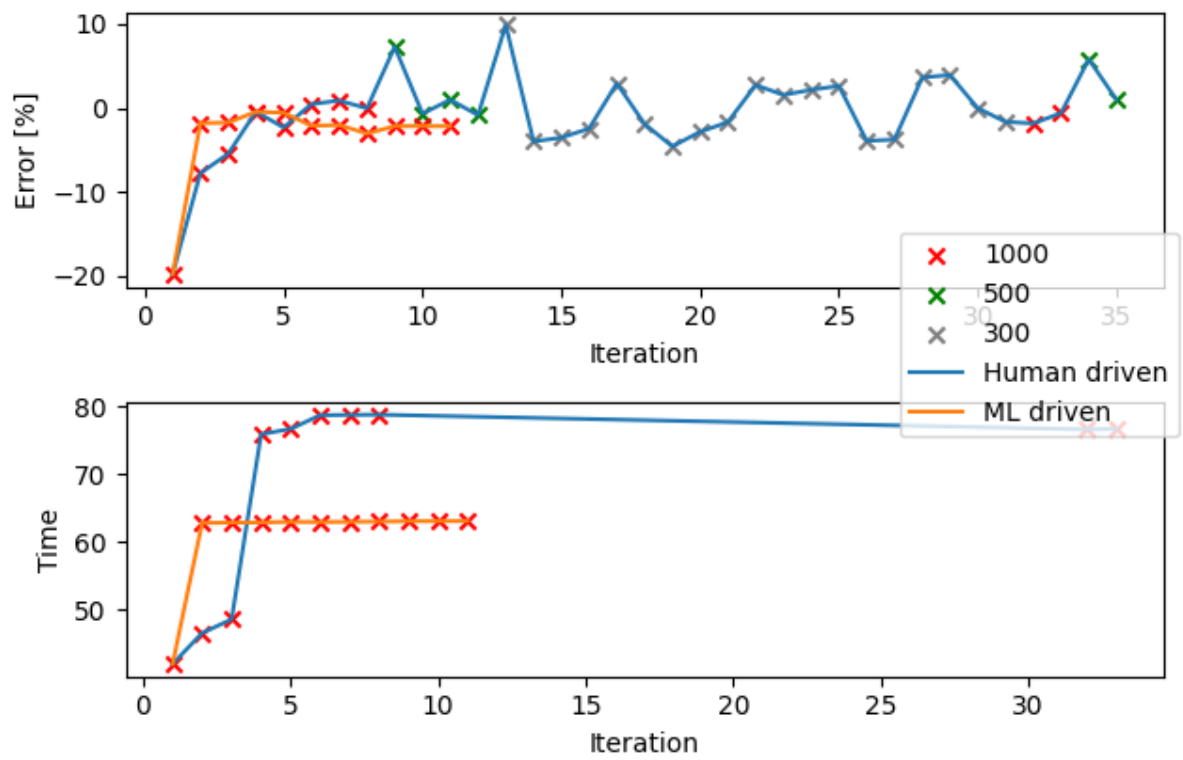# Diagram comparing human-driven and ML test trials (398 cP, set:1, unorderedT)

Gaussian Process Regression model with slow transfer penalization, trained with 1 initialization data
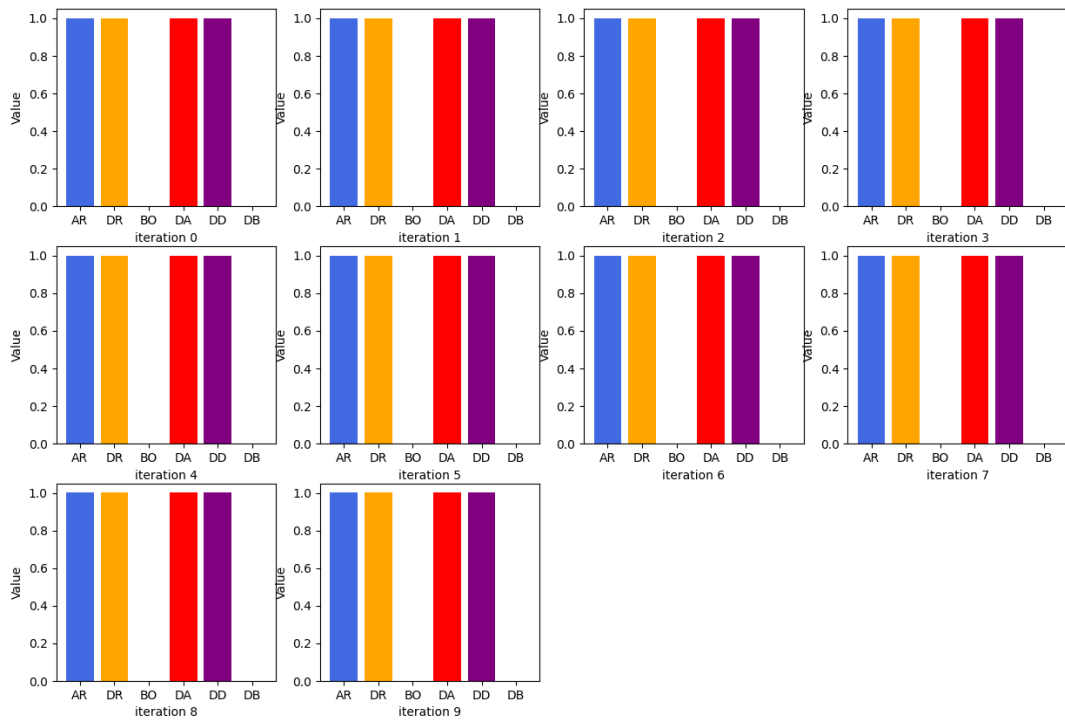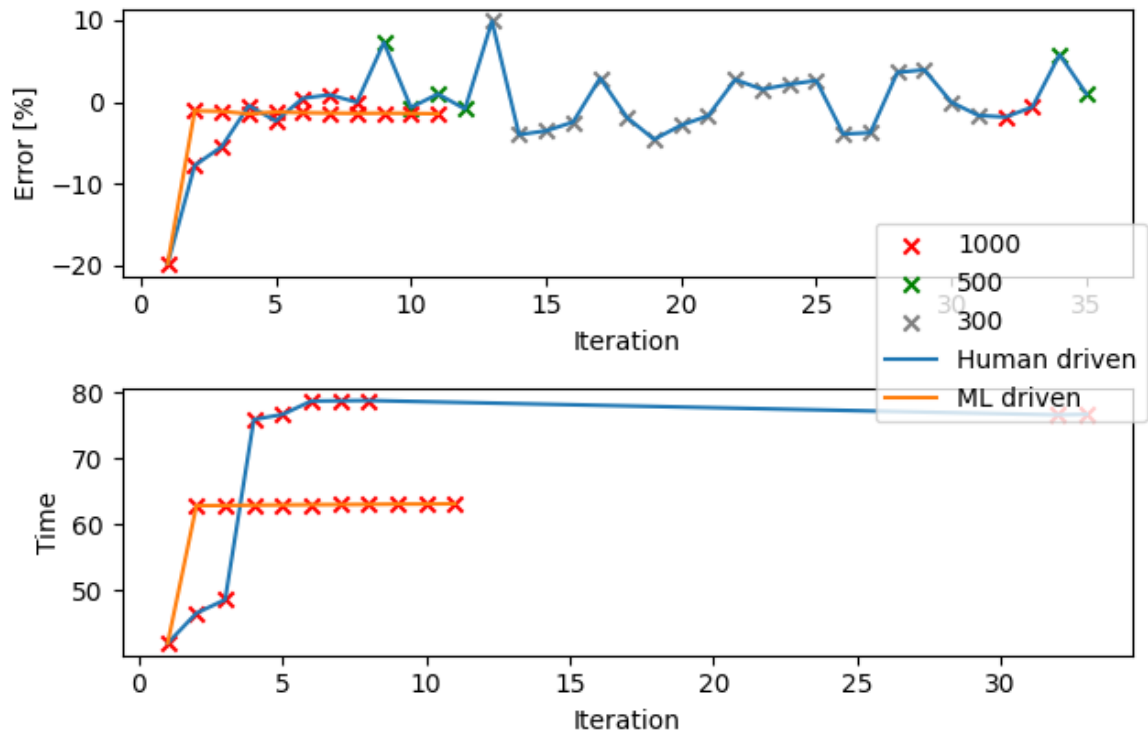
Gaussian Process Regression model with no penalization, trained with 1 initialization data

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

Linear Regression model with no penalization, trained with 1 initialization data

Set: 1 (LCB)

<u>Observation of trends</u>

Observation 1: LIN - scaling: multiply

All 10 sets of parameters are identical but the percentage error falls within the set boundary of -2 to 2%.

Observation 2: LIN - scaling: none

Similar as above, all 10 sets of parameters are identical but the percentage error falls within the set boundary of -2 to 2%.
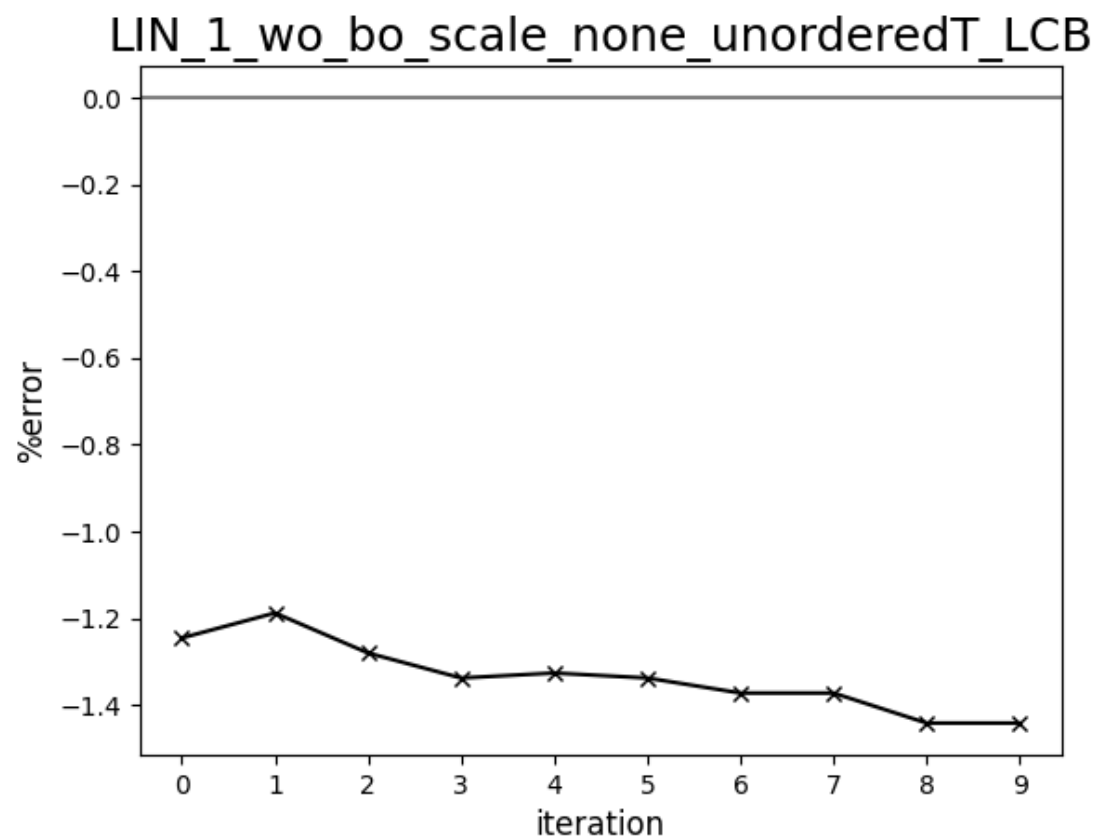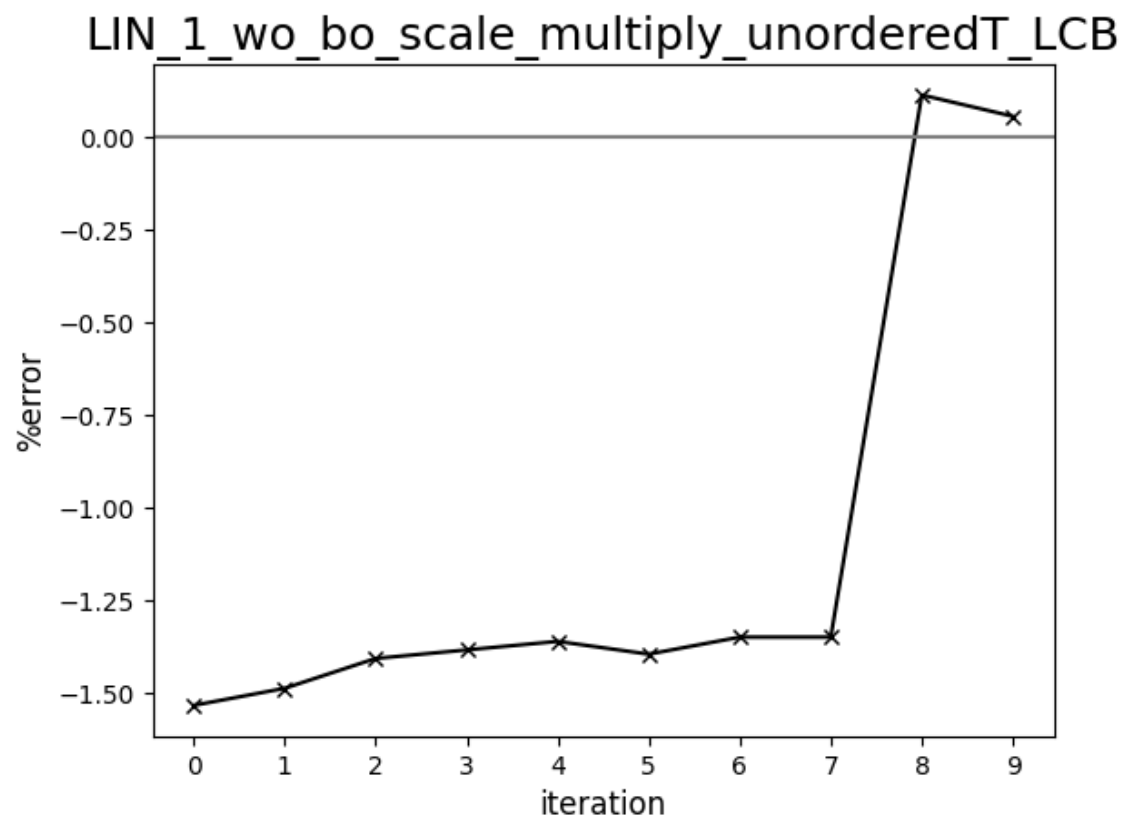
Error against iteration diagrams:

LIN_1_wo_bo_scale_multiply_unorderedT_LCB
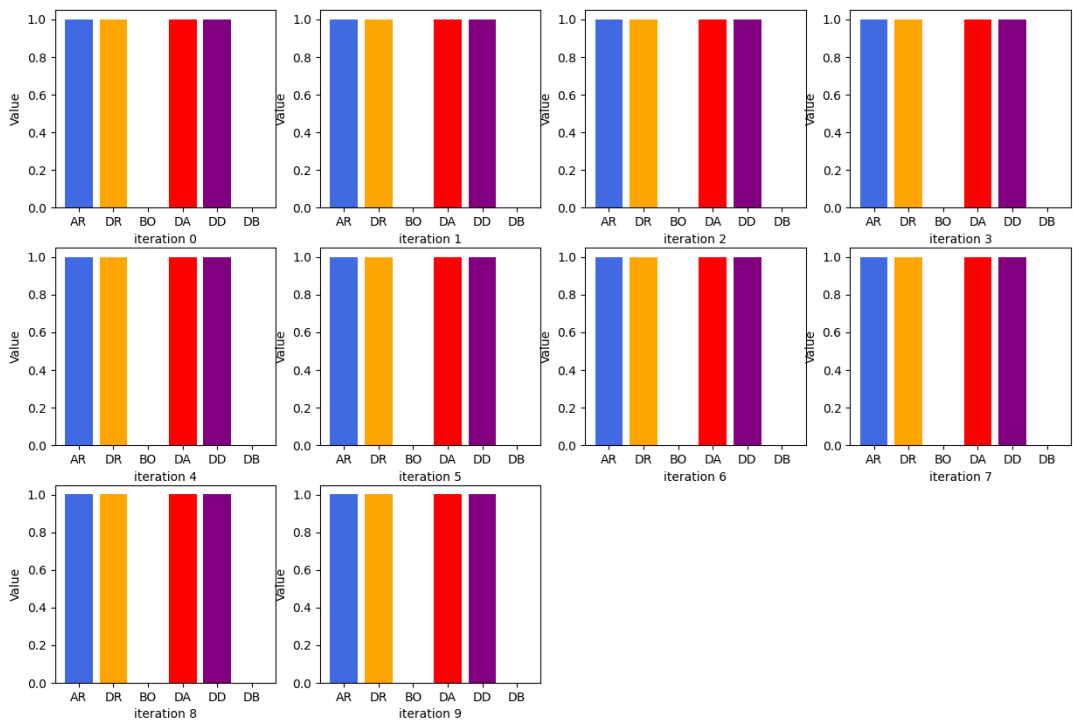

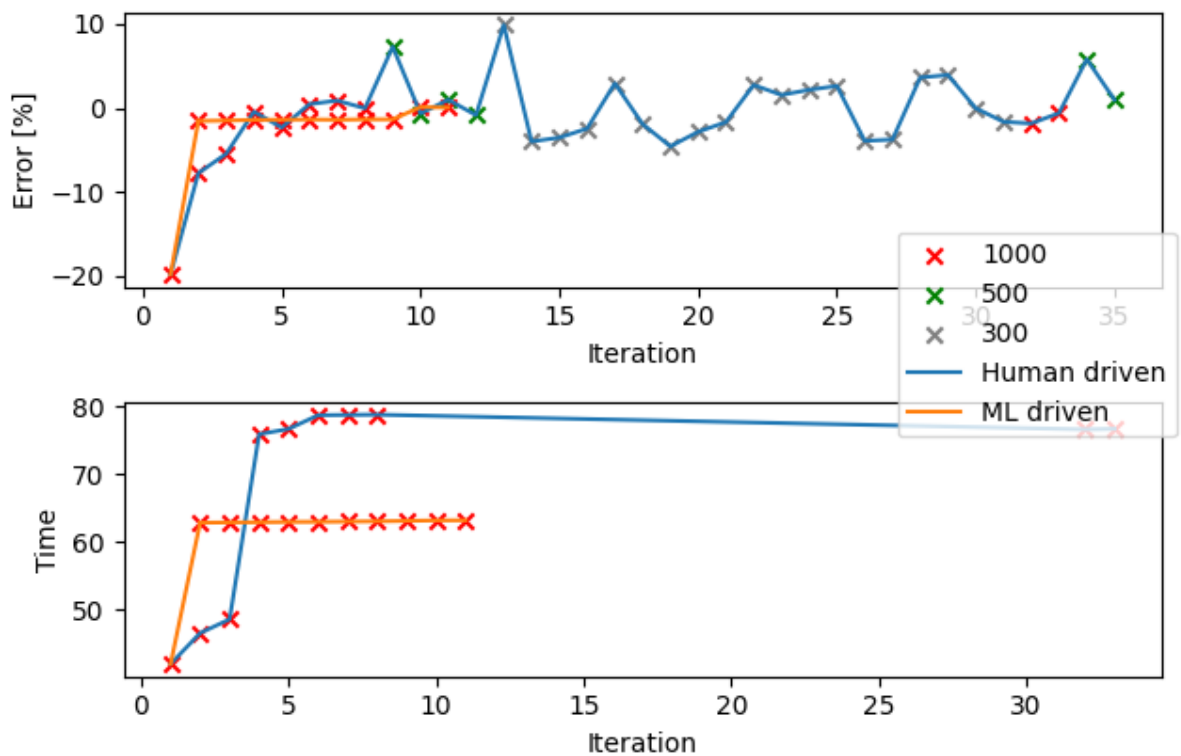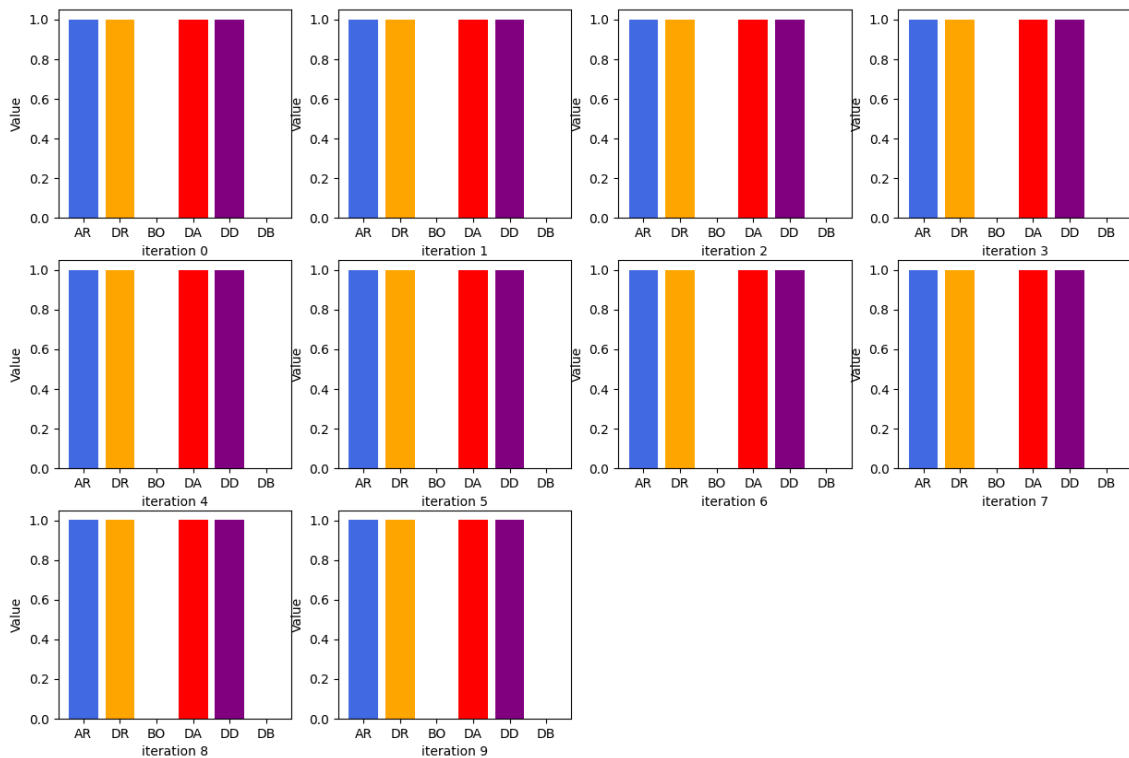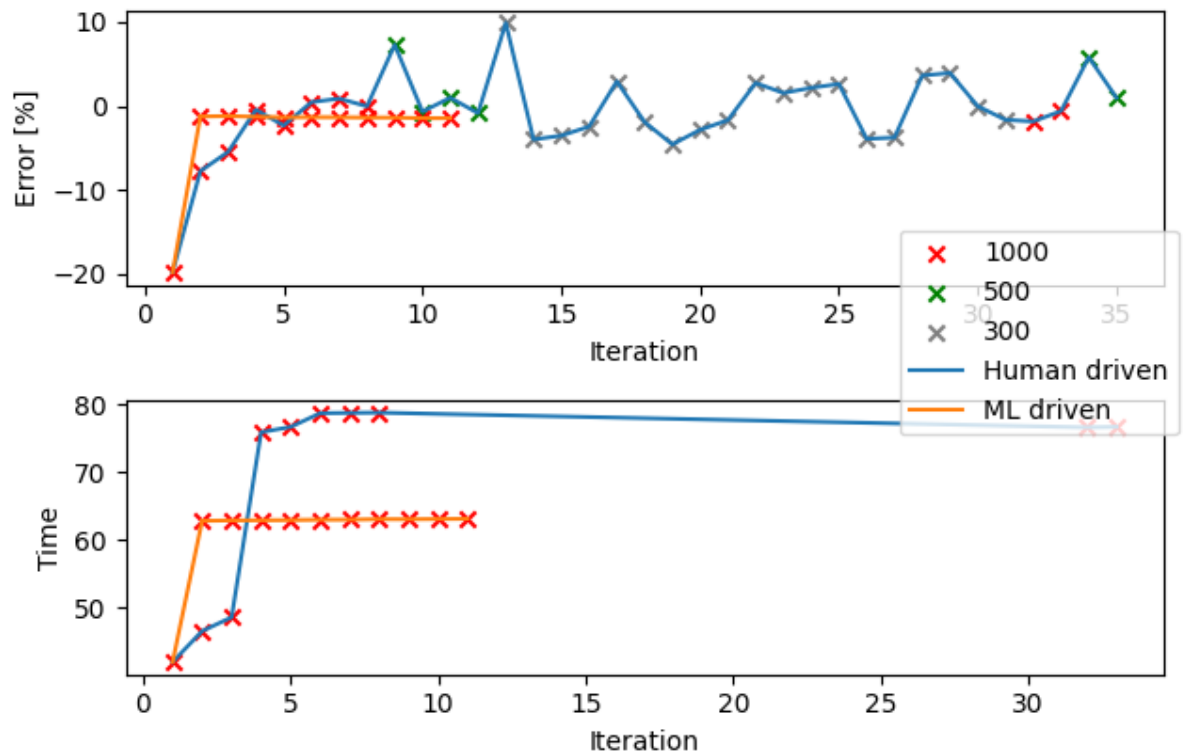## LIN_1_wo_bo_scale_none_unorderedT_LCB

Diagram comparing human-driven and ML test trials (398 cP, set:1, LCB)

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

Linear Regression model with no penalization,
trained with 1 initialization data

Set: 1 (ver4)

Observation of trends

Observation: LIN - scaling: none

The first 2 'pre-determined' sets of parameters at the start are not very accurate but the rest of the parameters generated parameters that are within tolerance of 1000uL.

The first 2 sets of parameters:
- first: first set from standard calibration
- Second: half of the first set of parameters

Percentage error against iteration diagram:

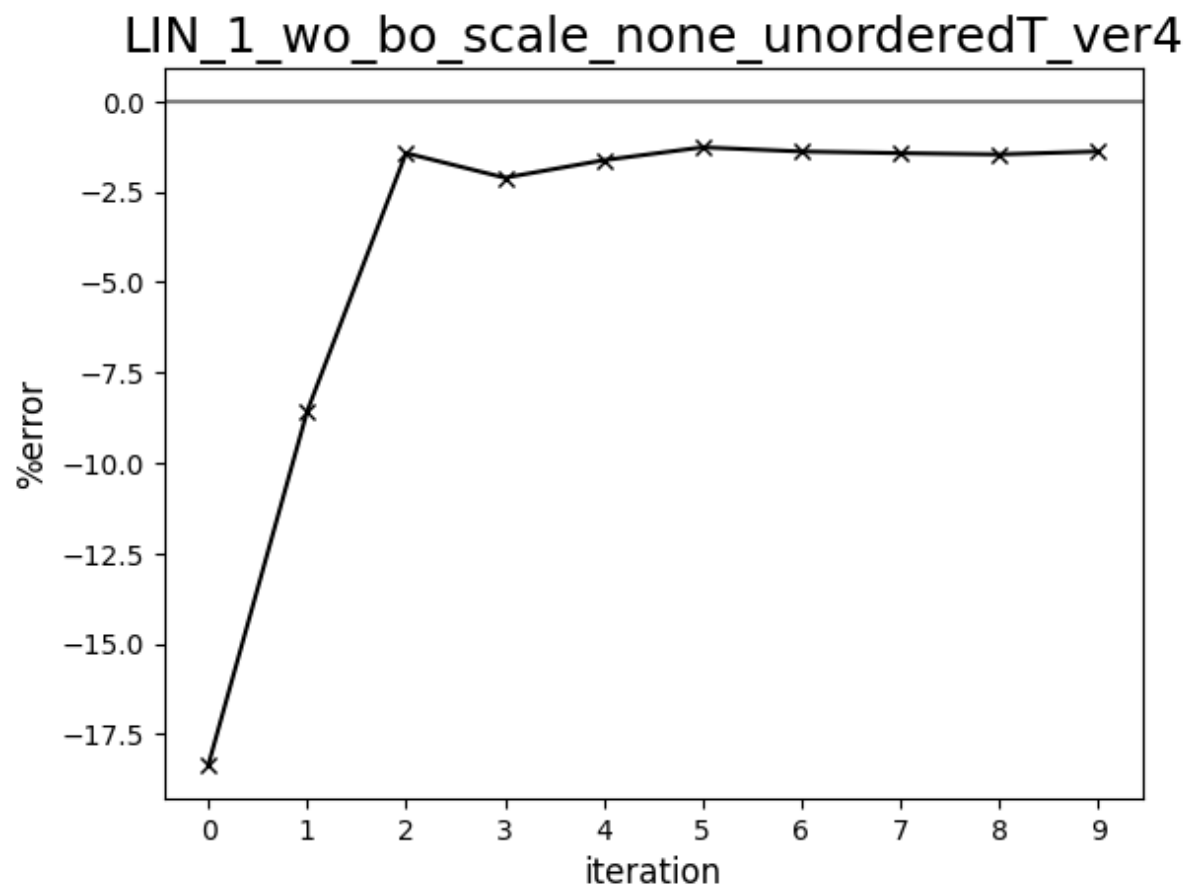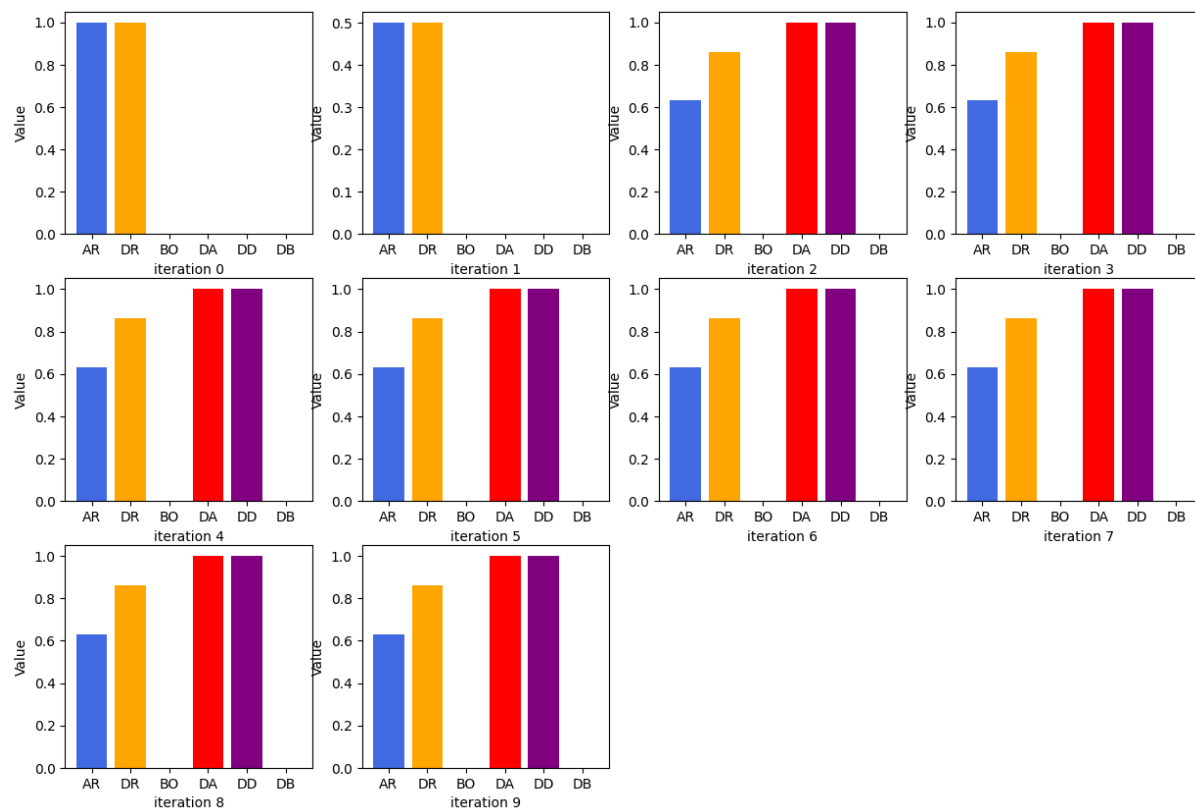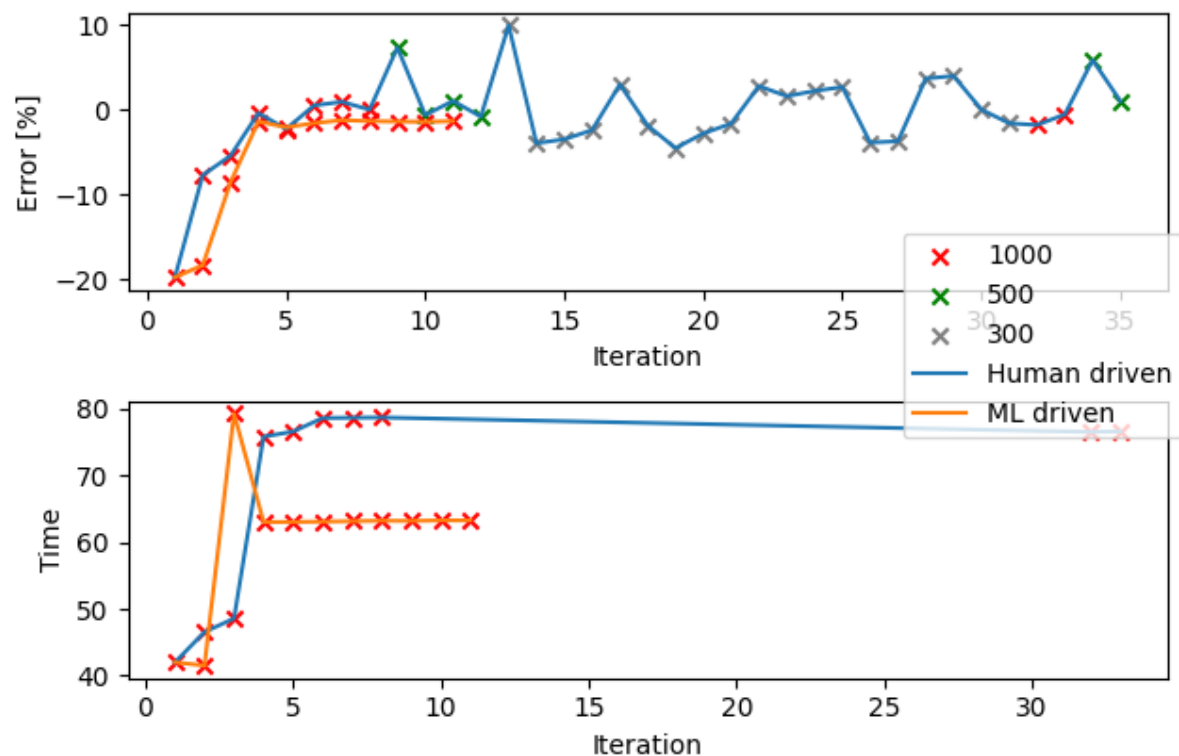LIN_1_wo_bo_scale_none_unorderedT_ver4

Diagram comparing human-driven and ML test trials (398 cP, set:1, ver4)

Linear Regression model with no penalization,
trained with 1 initialization data and unorder by time

Set: 1 (distributed)

<u>Observations</u>

Observation 1: LIN - scaling: multiply (slow transfer time penalisation)

Quite a lot of repetition in the sets of parameters generated, maybe it is because it already found the best possible set of parameters. The transfer time is very long (average of 200s) as compared to standard calibration but the percentage error falls within the preferred range of -2 to 2% (1.4 to 1.6%). The slow transfer time penalisation does not seem to work for iteration either.

Observation 2: LIN - scaling: none

Similar as above, a lot of repetition is observed in the sets of parameters generated but this iteration generates a set of parameters that is faster and more accurate than the iteration above. This iteration has an average of -1% and an average transfer time of around 70s.

Percentage error against iteration diagrams:

LIN_1_wo_bo_scale_multiply_unorderedT_distributed
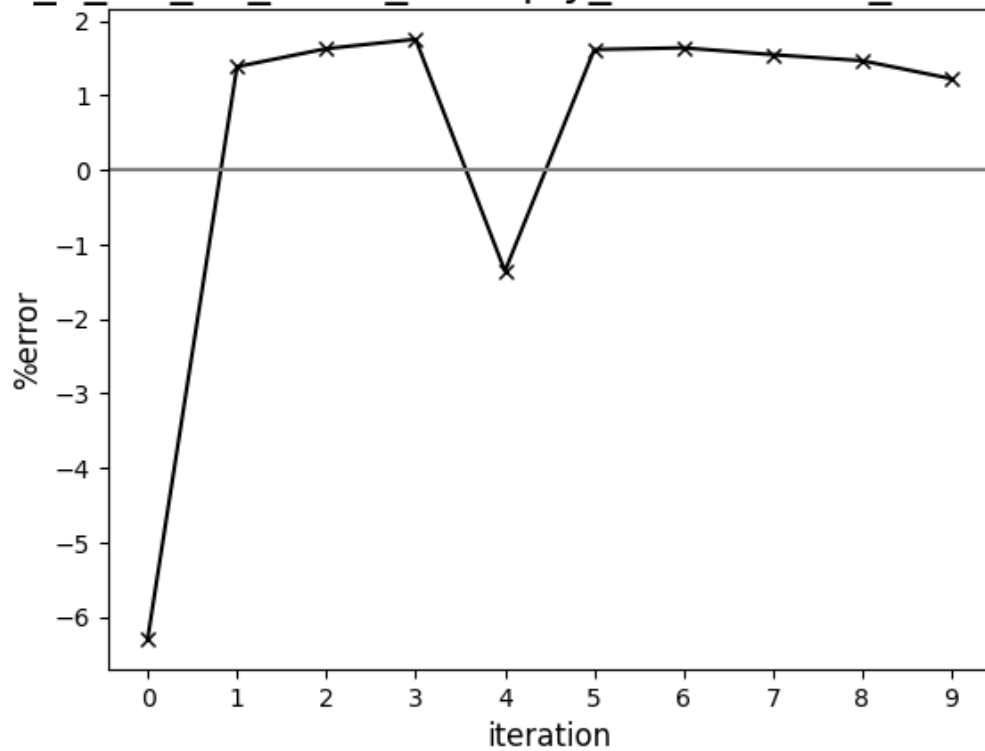
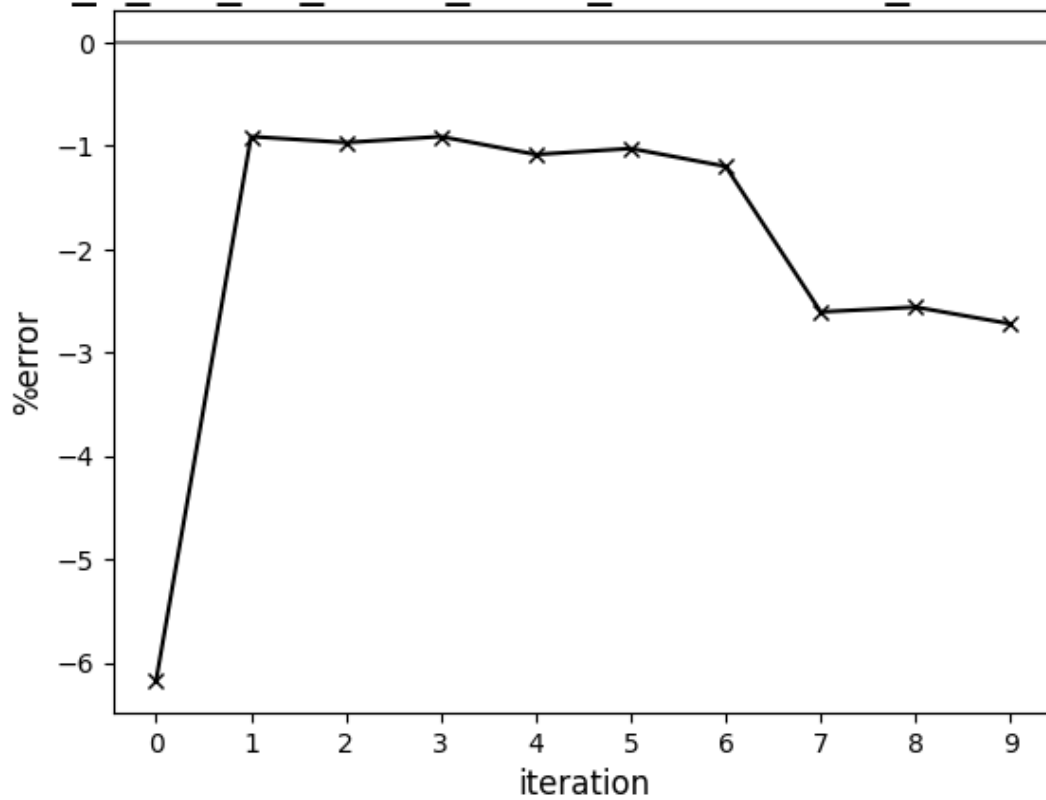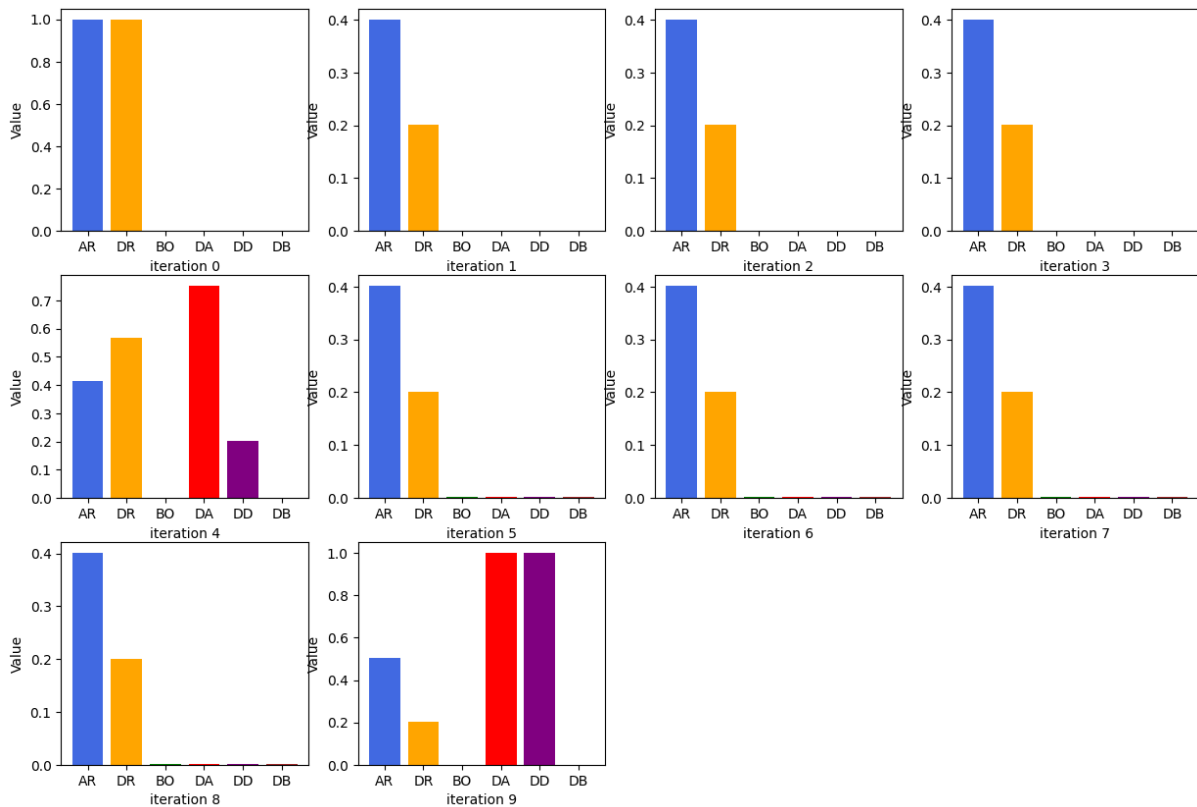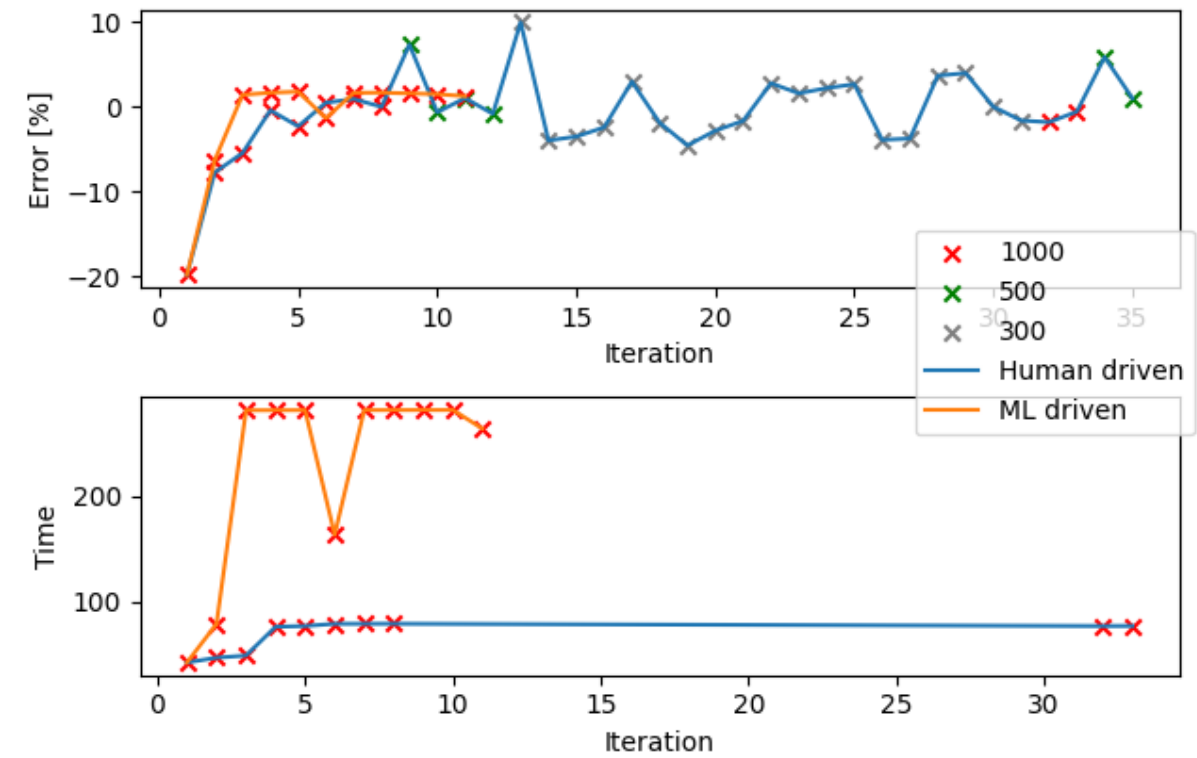LIN_1_wo_bo_scale_none_unorderedT_distributed

Diagram comparing human-driven and ML test trials (398 cP, set:1, distributed)

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

Linear Regression model with no penalization,
trained with 1 initialization data and unorder by time