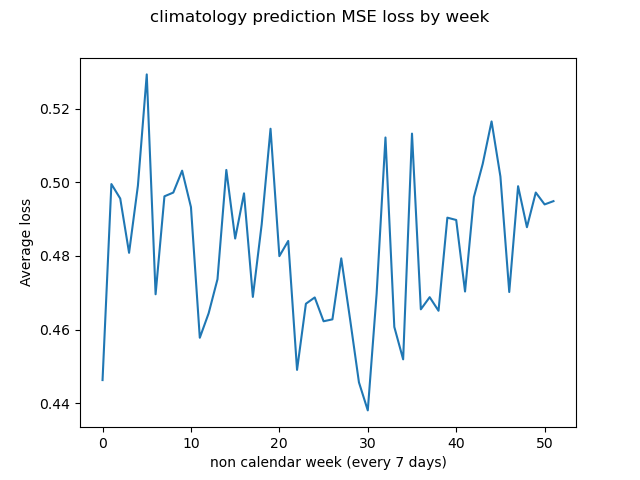
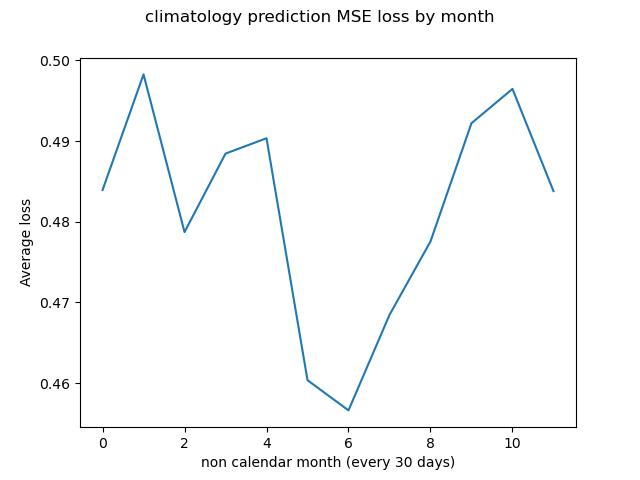
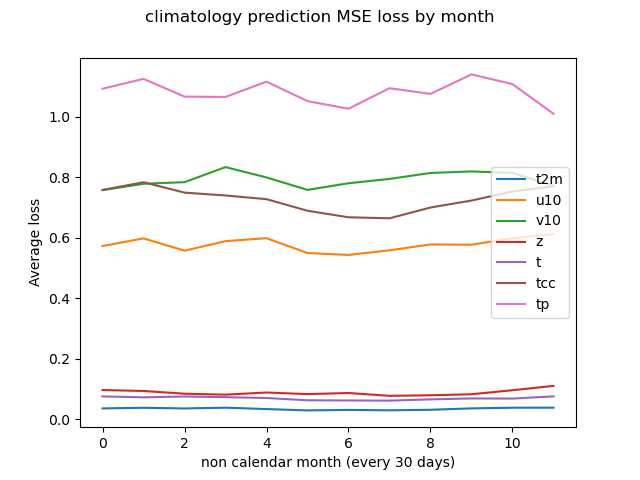
**REPORT 3**

1. CLIMATOLOGY MODELS

We have developed a climatology model as a baseline for prediction strength. If our models can perform better than this baseline, we can claim them to be effective. In climatology, we first take the mean over specified time intervals for the whole training data. Then we use these means as the prediction for the corresponding intervals. Two models have been observed. Weekly climatology is more specialized because it predicts different values for each week, but it is more likely to overfit the training data because of the smaller sample size for the means. That’s why we also look at the monthly climatology which is less specialized but also less likely to overfit. We look at different types of loss functions, namely MSE, RMSE, MAE.



In both models we observe lower losses during the summer season for the northern hemisphere. If we assume that the climate is more stable in the ocean regions compared to the land because the air above the land heats up more quickly (more explanations can be given), we can claim that the loss changes are more dependent on the land predictions. Because 70% of the land is in the Northern Hemisphere, then the lower loss might be related to the summer climate of the Northern Hemisphere. Since the standard deviation of the parameters is lower in the summer, this could mean that summer climate is more stable and predictable than the other seasons. In the end the losses for both models are very similar in average.

Chart, line chart

Description automatically generated

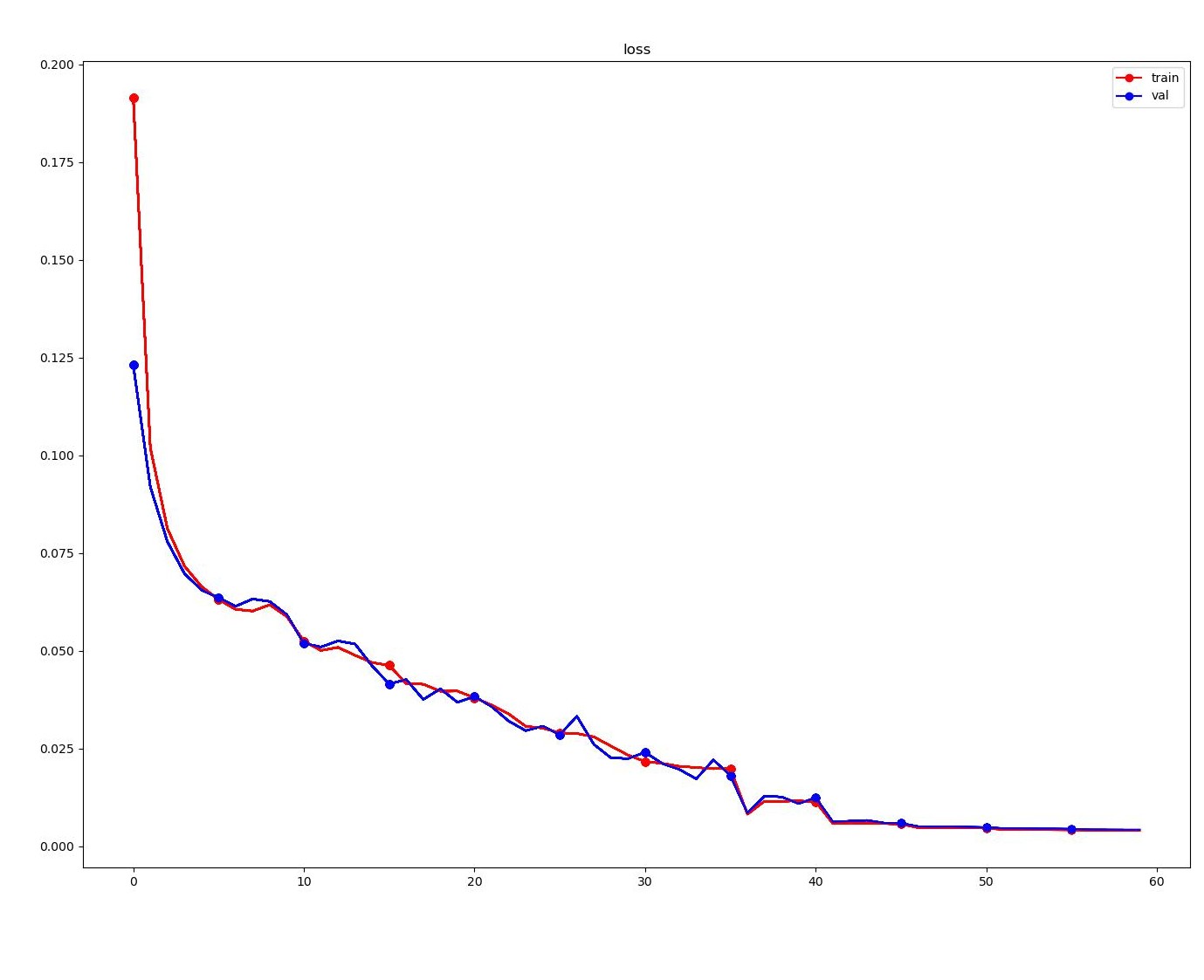
When observing MSE losses per parameter, we see that total precipitation has the most deviation from the means. This can be correlated with the frequency of showers, because if there are few extreme precipitation events in an interval and most other days have no precipitation, then the predictions will be swayed by those extreme events.

Lower losses for 850 hPa temperature (t) and 500 hPa geopotential (z) is expected. The higher in atmosphere you go, the more stable are the climate parameters such as temperature.

Another point to mention is that the L1 loss per parameter performs differently compared to MSE loss. (I do not know if this needs further explanation. The fact that L1 tp loss is lower confirms that extreme prediction differences are strengthened when computing MSE losses). In the next section, I will use L1 loss when comparing climatology and our model, because I don’t want to put too much weight on the outliers.

1. UNET MODEL AND MANY-TO-ONE MODELS LOSS CURVES

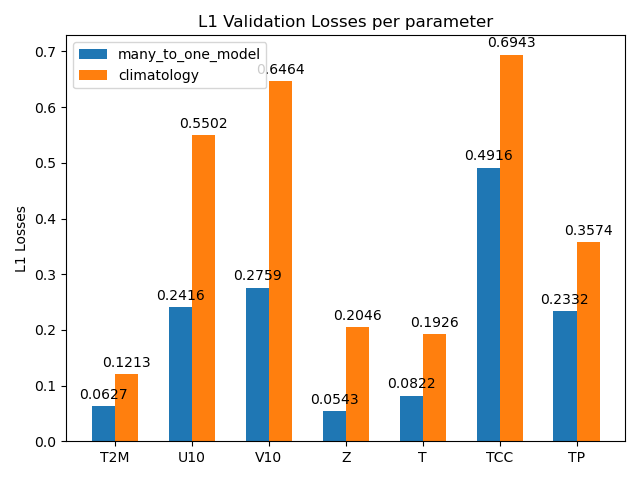
The training process for many-to-many model has been further optimized. I was observing some fluctuations in validation data and a plateau being reached after approximately 35 epochs. I applied an exponential scheduler with a gamma of 0.5 starting with 35. epoch every 5 epochs. I saw a very big decrease in both the training loss and validation loss. (From 0.015 to 0.003).



When looking at individual parameter losses during the training process, I recognized that every parameter loss has a marginal decrease at the start except the total precipitation loss. TP loss is approximately 5-10 times higher at the start. It gets stuck around 0.25 loss for many epochs and as the model tries to adapt the weights more for tp prediction we see a sudden decrease in the tp loss at one point, and a small increase in the other parameters. This seems to stabilize very quickly, and the model keeps updating weights such that every parameter slowly goes down. At the end all parameter losses are relatively close to each other.

This matches the results from the climatology model because TP standard deviation is high. Also, TP values are really small (10-e05 mean) before normalization which could also affect the predictions after normalization.

Lastly, we compare our many-to-one UNet models to the baseline climatologies. The many-to-one UNet models predicts an individual parameter as output from the remaining parameters as input. So, we have 7 different models in total, each predicting a single parameter. The training process is different than the original many-to-many model. Many-to-one models seem to converge faster. With a learning rate of 1-e3, the validation loss reaches a plateau after 6 epochs. At that point a scheduler is applied every 2 epochs with a gamma of 0.25. Validation Loss (L1) is computed for each parameter. Following Graph compares the monthly climatology model and many-to-one UNet model.



We can clearly see that our UNet model has in most cases losses that are less than half. Most noteworthy parameters are TCC, TP and Z. Z performs extremely well compared to climatology. Our many-to-one models still seem to struggle with TCC and TP values as they are closer to the baseline values. The analysis of these differences will be done later.