Norwegian University
of Life Sciences

Master's Thesis 2024 30 ECTS
Faculty of Science and Technology

A comparative study of soil temperature models, including machine learning models using few parameters

Mats Hoem Olsen
Master of Science in Data Science

Foreword

Denne oppgaven tok ikke 3 måneder å skrive, det tok 5 år med arbeid, læring, møter med fantastiske mennesker, egen læring, og dedikasjon.

Jeg vil ønsker å takke alle mine venner som har støttet meg under skrivingen, og trøstet meg i de tunge tidene. En stor takk til Mahrin Tasfe, Emma Sofie Rikheim & Celine Hagen, Vegard Eriksen Sæther, Bee, og Sindre Stokke for å passe på at jeg ikke overarbeidet meg, og kunne få en puste pause under denne intense perioden.

jeg vil også strekke ut en takk til Spillforeningen Både Kort og Bredt, og Rollespill gruppen som jeg deltar i for å gi meg noen morsomme øyeblikk og noe å glede meg til når oppgaven er levert inn.

Jeg ønsker også å gi en stor takk til Studentlivssenteret i Ås og Ås Helsestasjon for den hjelpen og de tilbudene de gir til studentene og som jeg fikk bruke. Mental helse er helse og det viste seg under skrivingen, og er evig taknemelig for gratis tilbudene de tilbyr.

Jeg vil også takke veilederne mine; Mareile Wolff, Berit Nordskog, og Brita Linnstad for å ta meg imot og være tolmodig med meg mens jeg fant ut av ting underveis dette semesteret. Dere har vært flinke veiledere.

Contents

List of Algorithms

List of Figures

List of Tables

Todo list

Abstract

This study focuses on 3 models that have been used in the literature to predict soil temperatures. The depths chosen as targets are 10cm, and 20cm in 4 regions; Innlandet, Østfold, Vestfold, and Trøndelag. In each region there are 4 stations spread along the region to get the most coverage of area to be the most representative of the local area. The models chosen had been used in the literature to predict soil temperatures. The data used in this study was collected from Kilden with the features; hourly air temperature from 2m height, soil temperature from 10cm depth, soil temperature from 20cm depth.

The models chosen are LSTM, BiLSTM, GRU, linear regression, and linear regression model modified by Plauborg. All the models performed within $2°C$ RMSE and an absolute bias of $1.5|°C|$ except linear regression that had an RMSE of 4.5°C RMSE and absolute bias of

Oppsumering

Denne studien ser på modeler som predikerer jordtemperaturer

Keywords: LSTM, GRU, RNN, Soil temperature, Machine learning, regression,hourly, weather forecasting data

1 Introduction

In agriculture soil temperature is one of the important parameters to put into consideration when thinking about pest prevention, conservation, and yield prediction. The reasoning for this is that knowing the soil temperature is gaining useful insight into gain important info for water management [2], potential yields [3], calculation of plant-growth [4], and predicting hatching insect eggs[5, 6]. Being able to predict the soil temperature a few days in advance does give insight into potential flooding and erosions[7], when seeds start to sprout [4], nitrogen processes [8] in the soil. Due to climate change it is more important to know soil temperatures at given depths.

If it's important, why don't institutions measure it everywhere? There are several reasons for this, but a common reason is that it's expensive to install new equipment on old weather stations. Furthermore, it is unfeasible to install sensors absolutely everywhere at any depth, however it is not necessary with full coverage of an area as it is sufficient to have a few samples here and there to get an overview of the current state of the soil. Another thing is that it might be impractical to install sensors in some areas due to climate, soil quality (or lack there of), or the misrepresentation of the area if it's a geographical or meteorological special case.

Sometimes the weather station do have the sensors in the fields reading soil temperature at given levels, but due to technical misadventures and unforeseen phenomenons there might be gaps or misreadings that need to be replaced with approximations or NULL values[1].

Previous research has investigated soil heat conductivity, leading to the formulation of differential equations [9]. However, these mathematical statements, which involve heat transfer, are computationally demanding and challenging to simulate or calculate [9, 10]. Numerical solutions are not the only obstacle; the dynamic nature of heat within the soil also plays a crucial role. For instance, frost in Scandinavian countries significantly alters soil heat conductivity [7], further complicating accurate calculations. As part of this study, data will be collected from Norway, situated within the Scandinavian region.

Deeplearning models

[1]These values are different from 0 as they represent "no data" and can't be used to do calculations.

A beneficial model would be one using the fewest number of parameters as possible while returning results within acceptable tolerances. This study will consider models that can use only time and air temperature as those two features are the most common measurements measured at weather stations, since soil temperature is not necessarily calculated as stated earlier. A good metric in this study will be considered to be a combination of Root Mean Square Error and Explained Variance (see section 3.5).

This study aims to address the following key questions:

- Achieving Good Results with Minimal Parameters: Can satisfactory predictions be obtained using a limited set of meteorological and chronological parameters?
- Deep Learning Models for Soil Temperature Prediction: Is it feasible to employ deep learning models for predicting soil temperatures?
- Complexity of Deep Learning Models: Is it necessary to utilize complex deep learning architectures when predicting soil temperatures?
- Suitable model for Nordic climate: Is there a model that fits for the Scandinavian climate?

Regarding deep learning models, this study primarily focuses on Recurent Neural Network networks and explores various compositions of this technology. The definition of a ”good result” will be relative to the performance of other models in the field and to similar studies that employ comparable architectures. Additionally, the Gated Recurent Unit (GRU) has been considered as an alternative to LSTM in this context due to its simplicity, and yet mechanically similar to the LSTM.

2 Theory

This section discusses the theory behind the models used in the study, with the first section being general information about the soil and use cases of soil temperature in agriculture.

2.1 Soil temperature

The difficult in predicting the soil temperature comes from that the environment is highly variant and radically different from each other. A farmer in Sunndal in Middle-Norway would have to do different considerations than a farmer in Karasjok in Nord-Norge or a farmer in Bergen in Sør-Norge simple due to differn climat and soil profile. There exist methods to help farmers get an local estiamte of current soil state; Methods such as soil texture, soil smell, soil feel, and colour. These methods works to make on-the-spot desition to when plant crops, water the crops, or when to harvest. A better approch is to have a model to predict the upcoming temperatures so farmers have a window of time to prepere for crop harvest or planting.

remove, and replace with just the practical application

An application of the soil temperature is used as a mean of agricultural advise is the study conducted by LMT where they look at the mean five-day air temperature compaired to the mean five-day soil temperature to asses when it is useful to start sowing the seeds[11].

rewrite so it is more accurate but still general, and change chose of place and method

A way to messure soil temperature is to insert a rod into the ground and messure, however to messure multiple depths there are usally three ways to do it; one rod to messure all the depths, one rod for each depth, and a hybrid solution of the first and the second. There are two sensors that are being used

1. Model 107 Temperature Probe[12]
   - Tolerance: $\pm 0.2°C$ (over the $0°C$ to $50°C$ range)
   - Measuring Range: $-50°C$ to $+100°C$
   - Probe Diameter: 0.76 cm (0.3 in.)
2. PT500 temperature sensor[13]
   - Measuring Range: They have a broad measuring range, usually from $-50°C$ to $+400°C$
   - Long-term Stability: PT500 sensors are known for their high long-term stability, making them reliable for continuous use over long periods.
   - Construction: These sensors are constructed with platinum, which contributes to their precise measurements and robustness.

There are a few depths to choose to monitor, on of those ranges are 5 cm to 15 cm range that is the root zone[14]. The root zone is an range where the roots of plants where the highest density of roots are, and at the end of the root zone researchers can gather information about droughts and snow-melts that are filling up og depleting the root system.

Add something here between generic and formula. How messure, Why these depths

A simple naive way to predict soil temperature would be to use the equation found by [15]

$$\text{daily soil temperature} \approx \widetilde{T}_{\text{soil,year}} + e^{-z/D} \sin(\omega t - z/D + \phi) \tag{1}$$

This analytical formula has its limitations as it does not take into account rain fall, snow melt, freezing, and re-freezing. Further more, the formula does not incorporate the importance of the surface temperature and its inpact on the soil layers over time. An ideal formula would incorporate all of these elements and possibly more, but that would require more computation power than currently available.

An expantion of this soultion was expanded by [16] by including the solar movement to predict daily soil temperatures. The sun does heat up the soil differently depending on it angle over horizon and the cloud blocking or not hiding the sun. It is commented by the author of [16] that this simple model does not describe the soil temperature the effect of snow cover or precipitations effects.

From current understanding of soil physics a modern model researches in cooperate the saturated hydraulic conductivity and the unfrozen water content to their equations[7]. Some of these formulas contains nested exponentials[7]. This formulation introduces numerical limitations as the estimation at the center of the formula would be amplified as the computation continues. A commonly used approch is Finite Difference Method (FDM) where a differential equation gets decomposed to several equation that gets mapped to a grid with boundry conditions[8, 17, 18].

Discuss the problem with predicting soil temperature

2.2 Linear regression

gi grunnlag til hvorfor denne modelen

Air temperature has a direct connection to soil temperature as the main source of thermal energy next to solar radiation. A primitiv relation between air temperature and soil temperature at a given depth would be

$$T_{\text{Soil, n cm}} \approx \beta_{\text{n cm}} T_{\text{Air}} + \vec{varepsilon} \tag{2}$$

The $\beta_{\text{n cm}}$ represent the scaling factor for the air-soil relation. The regression model will be for the sake of convenience be expressed as the following expression

$$\vec{F}(\mathbf{A})\vec{\beta} = \vec{y} + \vec{\varepsilon} \tag{3}$$

Where $\vec{F}$ is a vector function with following domain $\vec{F} : \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times p}$ where $m, n, p \in \mathbb{N}$, $\mathbf{A}$ is the data in matrix form with dimensions $\mathbb{R}^{n\times m}$, $\vec{\beta}$ is the regression terms with shape $\mathbb{R}^{p\times 1}$, $\vec{y}$ is the target (TJM10 or TJM20) with shape $\mathbb{R}^{n\times 1}$, and $\vec{\varepsilon}$ is the residual error with the same shape as$\vec{y}$.

This basic model to express the linearity of the components to soil temperature. This will function as the base model for regression models.

2.3 Plauborg linear regression model with Fourier terms

The model developed by [1] was trained in Denmark hand has shown promising results for a Scandinavian model.

An improvement over an time independent linear regression model would be a time dependent linear regression model that takes not only current time into account of the calculations but also previous measurements. It is current knowledge that soil temperatures depends on previous temperatures and meteorological phenomenons. In the paper Plauborg [1] extend the features from only air temperature at current time to include also previus days of year and the air temperature from those days as an extention of [16]. This means the following F function that Plauborg used would be

$$\vec{F} := [air_t, air_{t-1}, air_{t-2}, air_{t-3}, \sin(\omega t), \cos(\omega t), \sin(2*\omega t), \cos(2*\omega t)]^T$$

Where air_t is the air temperature at time t expressed in day of the year (0-365), ω is the angular frequency in radians per hour or radians per day, depending on the time unit. The sine/cosine elements in the F function represent the variations through the day by fitting $\vec{\beta}$ to the yearly variation. To adapt the authors model to an hourly time unit would be to either

1. Extend the F function to include a larger ω coefficient to reflect hourly oscillations in conjunction with daily fluctuation
2. Refit the Fourier terms with a larger ω coefficient to make the oscillations more representative of daily temperature changes.

The larger coefficient could be expressed as $\pi/12$ while the smaller ω for daily values would be rescaled to $2\pi/365$.

The problem with this approach would be Fourier Sine-Cosine series approximation which would suggest that Plauborg's method could be subject to overfitting with addition of more terms on a small dataset. On the other hand it gives us a way to compute the coefficients α_i and γ_i for sine and cosine terms respectively, though it would be more numerically stable with a pseudo-inverse computation or a max log likelihood approach. However the python module used in this study utilizes a different algorithm described in this paper [19] that performs an iterative method of solving $\underset{\beta}{\operatorname{argmin}} \left| X\vec{\beta} - \vec{y} \right|$.

2.4 Long Short Term-Memory (LSTM) model

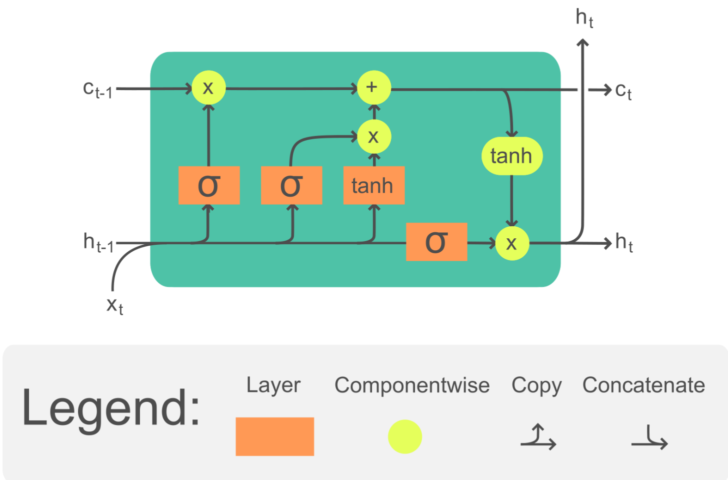


Figure 1: LSTM cell, From: Chevalier [20]

When modeling soil temperature it is important to know the air temperatures at the previous hours or days to predict the next soil temperature time step, for this a natural selection for a data driven model is a recurrent network. This type of deep learning models makes prediction based on previous time steps in the data, however the longer timespan the model takes into account the less important are the earlier time-steps become according to the model. The LSTM model has been tried in Türkiye [21], Belgium [22], United States of America [4] and their findings suggest it is a good model for predicting soil temperature, however it has also been used on a broader dateset spanning 3 continents [23].

LSTM[24] was developed in the field of Economy to predict the rise and fall of stocks, but has shown to be applicable to other problems that relies on time-series. It has been used to predict soil temperatures[4, 22, 23, 25–30] and utelises a method of storing information across the timesteps to preserve information better when predicting, and sending information backwards during training so the earlier weights get adjusted more efficiently. This method is an attempt to solve the problem of the vanishing gradient problem.

The most common problem in training Neural networks is the vanishing gradient problem where updating the first few layers of a large network becomes exponentially more difficult since the adjustments gets smaller and smaller for each layer towards the start rather than the reverse. Long Short Term-Memory changes this by carrying information from the previous cells forward thereby allowing updating earlier cells with bigger impact than the standard approach[24]. LSTM is part of a family of Recurent Neural Network's that passes information to other cells in the same layer.

LSTMs have proven effective in various tasks outside of economics such as natural language processing, speech recognition, and time series prediction. They provide a powerful mechanism for modeling sequential data while mitigating the vanishing gradient problem commonly encountered in basic RNNs.

2.5 Gated Recurent Unit (GRU)

The complexity of soil temperatures and its dependency of previous time-steps make Recurent Neural Network's a natural chose of a deep learning model, but the intricacy of an LSTM makes it difficult to fine tune. An alternative to LSTM is the GRU model[31] that has fewer parameters to adjust however it has a memory mechanism that allows it to forget and remember information that is passes to other cells in the model. It has been used in Turkey[21], and China[32]

The Gated Recurent Unit (GRU)[31] was developed in the field of natural language modeling to make translation predictions, however it has been shown that this model is also applicable in other applications than language translation.

GRU is a simplification of the LSTM cell with fewer total gates, and no output gate. This makes it quicker to train and better for memory deficient computers/servers.

GRU shares similarities with LSTM networks but simplifies the architecture by using two gates: the update gate and the reset gate. These gates allow GRU to selectively retain relevant information from previous time steps while avoiding keeping unnecessary information. The update gate determines how much past information should be passed to the future, while the reset gate controls how much past information to forget. GRU has been effective in various sequence modeling tasks.

EXPAND, and add formulas

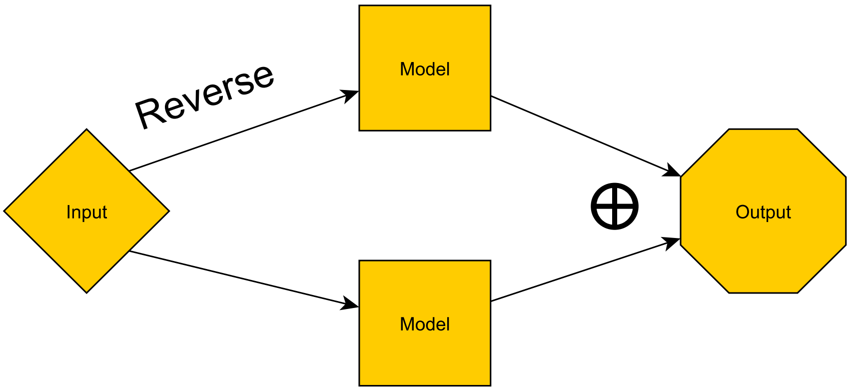


Figure 3: A diagram of the bidirectional method. The reverse indicates that the input data get reversed before it gets inserted into the model, while the bottom model gets the data. At the end the operator $\bigoplus$ get invoked that combines the output data from both model to a single output. The specifics of this operator can be any operation that combines two vectors to a single vector. In this study the operation is averaging the values.

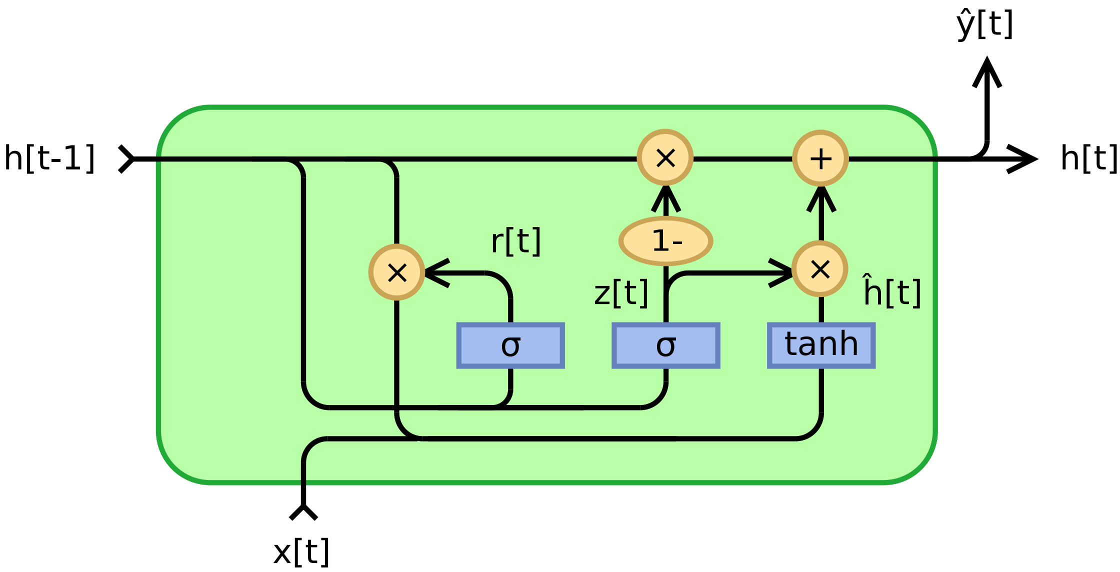


Figure 2: The GRU model where the memory cell gets more efficiently adjusted by the update gate (z) and appended information via the reset gate (r). From: [33]

The GRU model offers several advantages over LSTM. It is smaller, converges more efficiently due to the integrated memory and prediction layer, and has fewer gates. These fewer tunable parameters make it faster to train and store the model weights.

2.6 Bidirectional models

When making predictions with air temperatures as input, it would be useful to not only make predictions based only on air temperature from t_0 to t_{max} as there is information in the reverse time direction, t_{max} to t_0. This technique has been applied in other studies with noteworthy performance increase, therefore will be used in this study for both LSTM model and GRU model using the TensorFlow Keras module.

For the rest of the study a model with the prefix ”Bi” indicates that it is a bidirectional model.

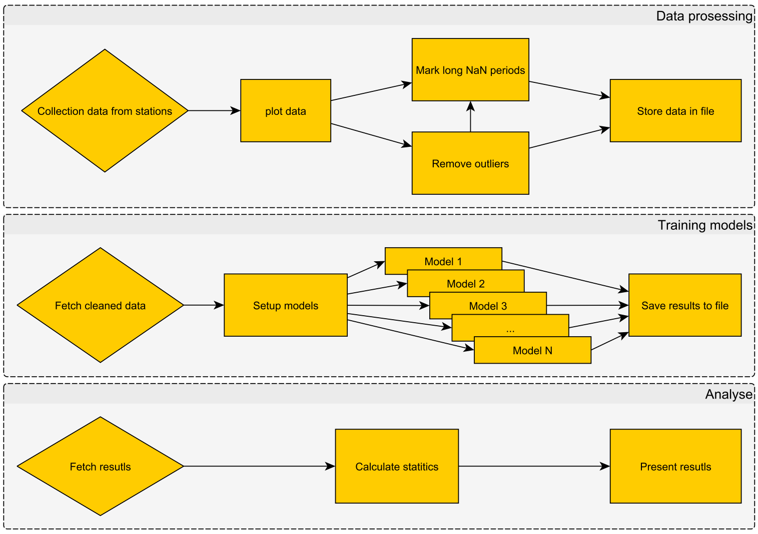


Figure 4: A surface level diagram of the methodology.

3 Method

Complete diagram

3.1 Source of data

For this comparative study the following data sources will be used

1. Norwegian Institute of Bioeconomy Research LandbruksMeteorologisk service (LMT) 2. Norwegian Institute of Bioeconomy Research Kilden (Kilden) 3. The Norwegian Meteorological Institute (MET)

3.2 Dataset

The dataset is chosen from four regions in Norway; Innlandet, Vestfold, Trøndelag, and Østfold. From each region are four stations picked:

All stations are sampled from the date[2] 03-01 to 10-31 from 2016 to 2020. The features mean hourly soil temperature at 10cm (TJM10), mean hourly soil temperature at 20cm (TJM20), and mean hourly air temperature at 2m (TM) are sampled from the LMT database.

3.2.1 Selection process

An array of stations was provided by LMT based on their possession of the necessary data. All stations were reviewed, checked for missing data, and those with excessive gaps were removed from the list or replace with another station. After compiling a list of stations, each one was re-examined to identify outliers present in the data and eliminate them accordingly. If certain stations had an excessive number of missing values after the outlier check, nearby stations were sought, and the affected station was replaced and the outlier check was re-done. Table 5 is showing Fåvang before treatment, and table 6 shows the same station cleaned and ready for being used as training/testing data.

[2]Format month-day

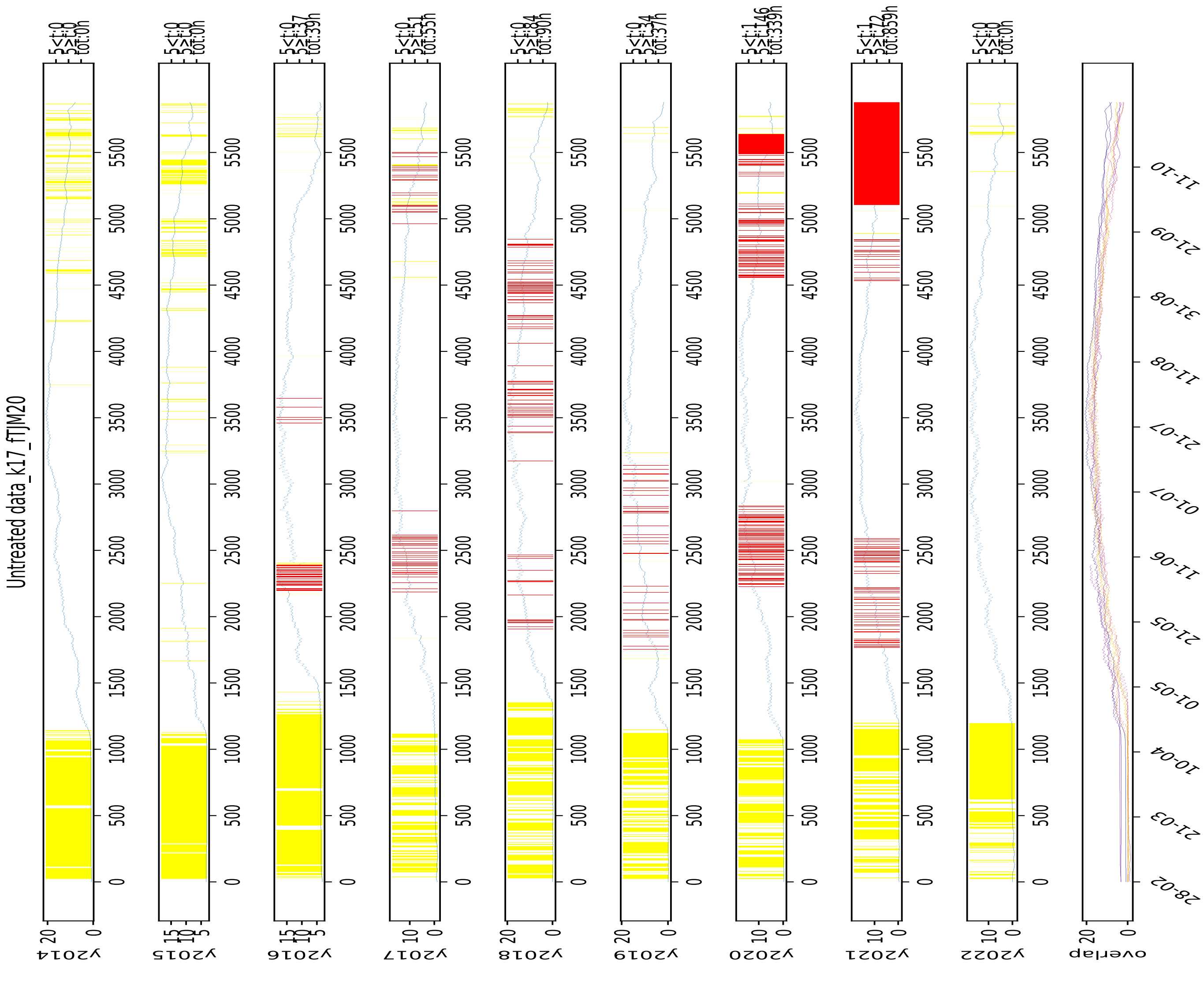


Figure 5: Visual representation of missing values at Fåvang from 2014 to 2022 at the parameter "TJM20". The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

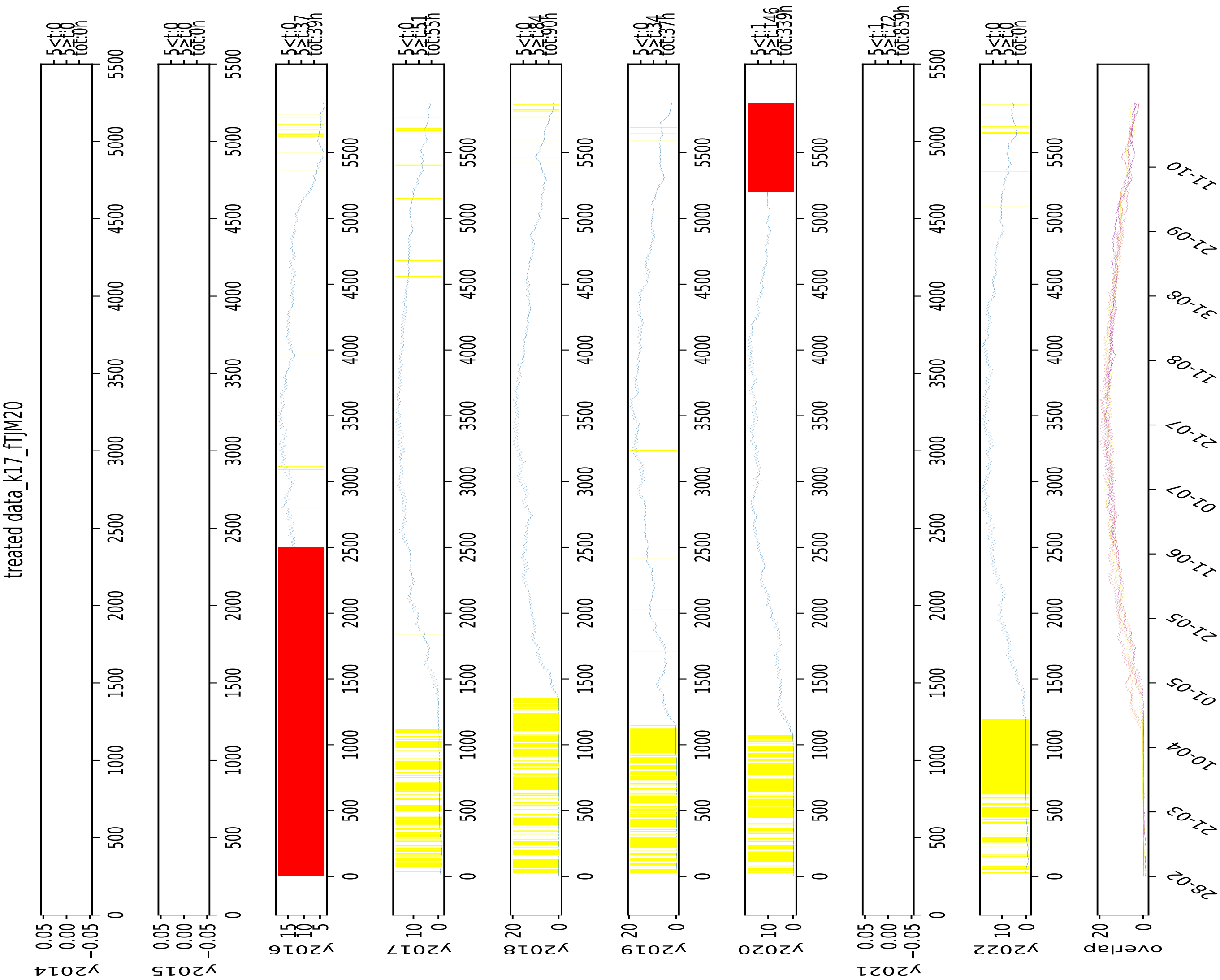


Figure 6: Visual representation of missing values at Fåvang from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL"). The years 2014 and 2015 was removed from the dataset but not coloured in due to technical limitations of reusable code, and furthermore half of 2016 was removed due to suspicion of misreading from sensor at this station.

Region	Name	ID	Drain type	MET name	Latitude	Longdetude
Innlandet	Apelsvoll	11	Selvdrenert	SN11500	60,70024	10,86952
Innlandet	Fåvang	17	Selvdrenert	SN13150	61,45822	10,18720
Innlandet	Ilseng	26	Selvdrenert	SN12180	60,80264	11,20298
Innlandet	Kise	27	Vannmettet	SN12550	60,77324	10,80569
Trøndelag	Kvithamar	57	Vannmettet	SN69150	63,48795	10,87994
Trøndelag	Frosta	15	Selvdrenert	SN69655	63,56502	10,69298
Trøndelag	Mære	34	Selvdrenert	SN71320	63,94244	11,42527
Trøndelag	Rissa	39	Vannmettet	SN71320	63,58569	9,97007
Vestfold	Lier	30	Vannmettet	SN19940	59,79084	10,25962
Vestfold	Sande	42	Vannmettet	SN26990	59,61620	10,22339
Vestfold	Tjølling	50	Selvdrenert	SN27780	59,04641	10,12513
Vestfold	Ramnes	38	Vannmettet	SN27315	59,38081	10,23970
Østfold	Rakkestad	37	Vannmettet	SN3290	59,38824	11,39042
Østfold	Rygge	41	Selvdrenert	SN17380	59,39805	10,75427
Østfold	Tomb	52	Vannmettet	SN17050	59,31893	10,81449
Østfold	Øsaker	118	Vannmettet	SN3370	59,31936	11,04221

Table 1: Station information from stations used in this study. The MET names was found by looking at the coordinates of the station and finding the closest MET station coordinates.

Name	version	description
Powershell	7.3.11	Windows native scripting language
Curl	8.4.0 (windows) libcurl/8.4.0 Schannel WinIDN	Commandline tool to communicate with servers using http.
python	3.9.11	popular Scripting language

Table 2: The description the software used in this study

The data plots (see figure 5 and 6 as example) shows all the raw data plotted as a blue line from 03-01 to 10-31. The yellow markings is placed there by computer algorithmns (see section 3.3.1 for in-depth explanation of the outlier detection methods) as potential outliers in the data. These markings have been looked over and verified weather or not they are genuine outliers or not. Further more the red lines are indicators of missing values, the number of missing values longer or shorter than 5 hours[3] are noted on the side bare with the total number of missing values regardless of length. The bottom bar are all the years laid on top of each other to highlight any years or periods that deviates for a particulate year compared to all the other years. There will be two versions of these plots, one for the untreated data and one for the treated data to show the difference the interpolation does to the data (see seciotn 3.3.2 for further details.).

FROST	Description
Station ID	Sendt a request to LMT for station information using their remote API.
LMT	Description
Meteorological data	Requested soil temperature from 10cm depth, and 20cm depth and air temperature (2m), from 2014-03-01 to 2022-10-31.

Table 3: Description of what was requested from each server (MET, LMT).

3.2.2 Collection of data

In the tabel 2 are the softwares used in this study and in the collection and treatment of the data. The program used in the collection of meteorological messurements is Powershell in combination with Curl. Using hyperlinks gathered from inspecting Kilden's web page using the browser (Microsoft edge) built-in inspector tool to get the relevant links to send data requests. The presise URL's can be reviewed on GitHub in the studies GitHub repository[4]. For a more surfac level description on what was requested of the servers see table 3.

3.2.3 Storage of data

The storage of the data is done through two data structures; Hashmap and DataFrame from the package pandas. The transformation of data is done with a customized data-type called "DataFileHandler" which is converted to a module for convenience. The keys for the hashmap is chosen by the naming of the data files and the pattern given to the class. To escalate the loading of the data it will also be exported to a binary file for faster retrieval.

Technical overview of custom data structure The data structure used to store the data from the different stations is called "DataFileHandler" and stores the data in a nested dictionary which can be inpreted as the data structure "tree". The main features of "DataFileHandler" is

1. Simple syntax for partitioning the data
2. Grouping the data after loading
3. Transforming all the data with the same function
4. loading and unloading of a large collection of data

This tree-structure uses recursion to search the dictonary to find the appropriate dataset to output.

3.3 Data cleaning and treatment

To use the data in this study it must be cleaned and treated for training. Though the data has been examined by the supplier, however it still had outliers that needed to be treated before modelling. For this reason several steps and methods is utilized in the prepossessing steps. The selection process for finding these station can be compiled into these steps

[3]The threshold for rain is 3 hours due to the high variance.

[4]Link: https://github.com/ConAltDelete/MT24CS

1. Recommendation from Norwegian Institute of Bio-economy Research
2. Compute the missing values in the data
3. Missing values analyse
4. Searching LMT database for alternative station candidates if current data is insufficient
5. If some station was replaced the repeat step 2

3.3.1 Outlier detection and removal

Revise so it reflects what you actually do

Though the data fetched from LMT is treated and controlled the external data from MET might not be, and this research project incorporated raw, untreated data from LMT to fill inn missing values.

The method to quickly find obvious outliers was to look at the following condition

$$|z(|\Delta T|)| = \left| \frac{|\Delta T| - E(|\Delta T|)}{Var(|\Delta T|)} \right| > 4°C \tag{4}$$

Make more clear

Where the $E()$ is the expected value, and $Var()$ is the variance. This condition looks at the absolute difference between consecutive measurements and calculates the z-score for each observation. It is expected that the change in temperature can't be too rapid. Further methods used to highlight potential outliers is

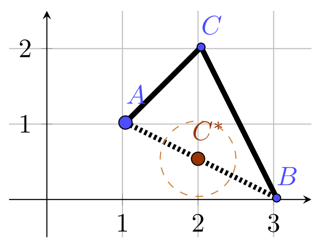


Figure 7: An simple outlier detection method utilizing a simple line to estimate where the expected point (C^*,red dotted circle) is supposed to be. If observed point C falls outside the tolerance level (green circle) then it is marked as an outlier.

3.3.2 Missing value imputation

The data has missing values, in particular during early Autumn when there were sub-zero temperatures meaning any rain measurements done during this period would have unpredictable fluctuations since at negative temperatures water can freeze, get clogged up with residual biomaterial from the surrounding area . When interpolating values the method chosen is a linear interpolation with a maximum period of 5 consecutive hours for soil temperatures, and 3 consecutive hours for air temperatures. The reasoning for this is that the soil temperatures are more reliable making it safer to interpolate without loosing too much information, while air temperatures has a higher variance making it more difficult to interpolate without cutting values.

Rewrite this part to reflect what is going on

3.3.3 Summary of data

After going to the procedure of removing outliers and interpolating missing values, the current form of the station are

stations	TM		TJM10		TJM20	
50	μ:11.43	max:31.3	μ:11.108	max:21.9	μ:10.709	max:19.3
	σ:6.164	min:-14.0	σ:5.346	min:-0.3	σ:5.157	min:0.0
42	μ:11.041	max:33.0	μ:11.268	max:23.2	μ:11.156	max:20.6
	σ:6.831	min:-14.5	σ:5.931	min:-0.4	σ:5.622	min:-0.1
38	μ:11.114	max:32.3	μ:11.13	max:22.8	μ:10.959	max:21.7
	σ:6.805	min:-19.7	σ:5.892	min:0.1	σ:5.713	min:0.3
30	μ:11.114	max:32.5	μ:11.243	max:27.6	μ:11.142	max:23.5
	σ:6.922	min:-16.8	σ:5.98	min:-3.3	σ:5.662	min:-3.3
118	μ:11.146	max:33.1	μ:10.613	max:21.8	μ:10.417	max:20.3
	σ:6.608	min:-16.6	σ:5.468	min:-0.9	σ:5.218	min:-0.7
52	μ:11.006	max:32.6	μ:11.706	max:23.2	μ:11.672	max:24.2
	σ:6.777	min:-18.0	σ:5.505	min:-1.0	σ:5.411	min:-0.6
41	μ:11.241	max:33.7	μ:11.648	max:25.9	μ:11.319	max:22.2
	σ:6.683	min:-19.2	σ:6.019	min:-0.8	σ:5.703	min:-0.3
37	μ:10.151	max:31.6	μ:10.589	max:22.0	μ:10.493	max:20.5
	σ:6.839	min:-20.1	σ:5.583	min:-1.5	σ:5.465	min:-0.7
39	μ:9.62	max:31.4	μ:9.286	max:19.6	μ:9.178	max:18.5
	σ:5.936	min:-15.1	σ:4.9	min:-0.9	σ:4.793	min:-0.3
34	μ:9.336	max:32.7	μ:8.792	max:19.3	μ:8.687	max:17.9
	σ:6.554	min:-19.5	σ:4.9	min:-1.3	σ:4.727	min:-0.2
57	μ:9.73	max:32.8	μ:9.387	max:22.2	μ:9.131	max:19.1
	σ:6.408	min:-21.4	σ:5.225	min:-1.3	σ:5.068	min:-0.9
15	μ:10.069	max:33.5	μ:9.188	max:20.2	μ:9.066	max:18.7
	σ:6.098	min:-15.4	σ:4.763	min:-0.2	σ:4.601	min:0.0
27	μ:9.971	max:33.0	μ:10.554	max:23.9	μ:10.205	max:21.4
	σ:6.992	min:-23.2	σ:6.25	min:-2.7	σ:5.967	min:-2.2
26	μ:9.64	max:32.4	μ:9.533	max:22.6	μ:9.515	max:20.6
	σ:7.308	min:-26.3	σ:6.171	min:-3.0	σ:6.104	min:-2.2
17	μ:9.192	max:29.9	μ:9.64	max:22.8	μ:9.35	max:19.9
	σ:7.633	min:-25.9	σ:6.441	min:-2.4	σ:6.094	min:-1.4
11	μ:9.761	max:30.2	μ:10.153	max:23.3	μ:10.004	max:21.6
	σ:6.903	min:-21.4	σ:6.172	min:-1.6	σ:5.913	min:-1.1

Table 4: The table shows the statistics of each station for each feature except for Time as it is a strictly monotonic sequence.

Parameter	value	Parameter	value
activation	”tanh”	kernel_constraint	None
recurrent_activation	”sigmoid”	recurrent_constraint	None
use_bias	True	bias_constraint	None
kernel_initializer	”glorot_uniform”	dropout	0.0
recurrent_initializer	”orthogonal”	recurrent_dropout	0.0
bias_initializer	”zeros”	seed	None
unit_forget_bias	True	return_sequences	False
kernel_regularizer	None	return_state	False
recurrent_regularizer	None	go_backwards	False
bias_regularizer	None	stateful	False
activity_regularizer	None	unroll	False
use_cudnn	”auto”		

Table 5: All the hyperparameters that are available to the Keras LSTM Layer and the standard options that this study chose to remain unchanged

3.4 Setup of models

In general, write more on all subsections

The models are set up in according to the relevant paper the model is fetched from, alternatively reuse the code made by the author. When importing the data to the model there will be modifying to the original code to facilitate for the model as far as it goes. Any modifications will be in the appendix under section **??**. For the convenience of the reader all code is using the sklearn estimator class to make all the models discuses in this study more user friendly and compatible with sklearns other functions. The details of the models will be discussed in section 2, this section discusses the setup and implementation of the models.[5]

3.4.1 Basic Linear model

The linear model (section 2.2) utilises in the study is created from the python model sklearn (or scikit-learn according to pythons package manager). The model is setup with standard parameters, and the data is fed into the model without scaling with fitted intercept coefficient.

3.4.2 Plauborg

The Plauborg regression will be formulated as a linear regression problem so that the 'LinearRegression' function in the Sci-kit module can be used. For the parameters used in the paper[1] the F function defined in section 2.3 will be formulated with loops to give rise 3 more parameters for fine-tuning the model. NULL-values generated from the procedure get replaced with 0, since the data fed to the model is significantly larger than 10h (the minimum for the training is 24h).

3.4.3 LSTM

rewrite or possibaly move

The LSTM used in this study came from the keras module using standard settings.

[5]Caution to the reader; The code used was run on the Linux subsystem (Debian) on windows due to the fact that the current version of tensorflow can't run on Windows.

Parameter	value	Parameter	value
activation	"tanh"	recurrent_constraint	None
recurrent_activation	"sigmoid"	bias_constraint	None
use_bias	True	dropout	0.0
kernel_initializer	"glorot_uniform"	recurrent_dropout	0.0
recurrent_initializer	"orthogonal"	seed	None
bias_initializer	"zeros"	return_sequences	False
kernel_regularizer	None	return_state	False
recurrent_regularizer	None	go_backwards	False
bias_regularizer	None	stateful	False
activity_regularizer	None	unroll	False
kernel_constraint	None	reset_after	True
use_cudnn	"auto"		

Table 6: All the hyperparameters that are available to the Keras GRU Layer and the standard options that this study chose to remain unchanged

Parameter	value
merge_mode	"ave"
weights	None
backward_layer	None

3.4.4 GRU

The GRU model used in this study is fetched from TensorFlow-Keras python module with standard settings. The model used in this study is the Keras default GRU layer.

The standard model settings are

3.4.5 Bi-Directional layer

rewrite

The Bi-directional layer used in this study came from the keras module using standard settings except for "merge_mode" which is set to "ave" for averaging the output values.

3.5 Metrics

The metrics used in this study are

• Root Mean Square Error (RMSE) • Mean Absolut Error (MAE) • Explained Variance (R^2) • Bias • Log Condition number ($\log(\kappa)$) • digit sensitivity

Soil temperature as a different behavior than air temperature since energy (temperature) though the soil gets dampen and delayed. Since the data used in this study has outliers that was not caught during data treatment, which has been addressed, the author of this study decided to include two more metrics that are not usually included in the evaluation; The log condition number, and digit sensitivity. Both metrics are based on the calculation of the condition number defined as

$$\kappa = \lim_{\varepsilon \to 0^+} \sup_{|\partial x| \leq \varepsilon} \frac{|f(x + \partial x) - f(x)|}{|f(x)|} * \frac{|x|}{|\partial x|} \tag{5}$$

Calculating this directly is impossible due to the limitations of handling infinitesimally small numbers in simulations. However, this paper uses a specific algorithm (referred to as κ) to estimate this value for all the models.

Algorithm 1: Method for calculating κ. $\mathcal{U}$ is a uniform random distribution in a range.

Data: Data
Result: $\log(\kappa)$
Let κ_f be the function 5;
$\kappa \leftarrow 0$;
for $i \in 1 \ldots |Data|$ **do**
  $\partial x \leftarrow \mathcal{U}_{[-\sqrt{\varepsilon/|Data|}, \sqrt{\varepsilon/|Data|}]}$;
  $k \leftarrow$ calculate with κ_f from x and $x + \partial x$;
  if $k > \kappa$ **then**
    $\kappa \leftarrow k$;
  end if
end for
return κ

The digit sensitivity is included to give an intuitive understanding of κ and is computed simply as $\log_e(\kappa)+1$. This number tells us the significant digit generated from the model. If the number is less than 0 then it's the nth digit after the decimal point.

For the rest of the metrics, they are defined as follows

- RMSE $= \sqrt{\frac{\sum(y_{\text{pred}} - y_{\text{truth}})^2}{n}}$
- MAE $= \frac{\sum|y_{\text{pred}} - y_{\text{truth}}|}{n}$
- bias $= \frac{\sum(y_{\text{pred}} - y_{\text{truth}})}{n}$
- $R^2 = 1 - \frac{\sum(y_{\text{pred}} - y_{\text{truth}})^2}{\sum(y_{\text{pred}} - \vec{y})^2}$

Where $\vec{y}$ is the mean of the target, y_{pred} is the predicted data, and y_{truth} is the observed soil temperature.

3.5.1 Model training

The models get trained on air temperature, however the precise input for each model is not the same for all. The features used for each model are described in table 7 and their transformation in table 7.

The models get a sample of the training data at the time due to the size and the amount for missing data (for example figure 5) The algorithm used to fetch reliable indexes are demonstrated at algorithm 2.

After the missing values has been identified rows get removed in such a way that if any of the features or target values is missing then that row get removed. At the end all the rows get concatenated , so there is just one complete dateset for the training.

Algorithm 2: Find Non-NULL Ranges (Abstract)

Input : Input data data
Output: List of tuples: ranges

```
FindNonNULLRanges(data) ranges ← empty list;
start ← None;
for item in data do
    if item is not NULL then
        if start is None then
            start ← item;
        end if
    end if
    else
        if start is not None then
            Add (start, item index - 1) to ranges;
            start ← None;
        end if
    end if
end for
if start is not None then
    Add (start, Last index) to ranges;
end if
return ranges;
```

Model name	features	transformations
Linear regression	TM	—
Plauborg	Time, TM	Time get translated in two way; the current day since new year if looking at daily values, and hours since new year if looking at hourly predictions. When converting TM to daily values the hourly data get averaged in 24 hour periods from midnight to 23:00
Deep learning models	Time, TM	Time get translated to hours since new year by taking the day of the year and multiplying it by 24 and adding the hour.

Table 7: Parameters used for predicting soil temperatures at depth 10cm and 20cm.

3.6 Use of Artificial Intelligence in this paper

In this paper there has been used Artificial Intelligence (AI), specifically Bing Chat / Copilot hosted by Microsoft Cooperation according to the current guidelines for use of artificial intelligence at the faculty of The Norwegian University of Life Science (NMBU), for the following purposes:

1. Formalizing sentences and rephrasing sentences.

2. Spellchecking
3. Code generation of basic consepts and structures (tree traversal, template for generic classes)

It is important to emphasize that my engagement with AI have been actively curated and verified with known sources. All code underwent rigorous manual inspection within a dedicated testing environment. Furthermore, no confidential or sensitive information was shared with the AI; my interactions focused solely on broad topics and general inquiries. To validate the accuracy of AI-generated responses, I cross-checked them with established research papers and textbooks.

model		52	37	50	38	57	34	27	17	average
Linear model 10cm	R^2	0.221	0.473	0.584	0.388	0.295	-1.386	0.551	0.501	0.423
	MAE	2.889	3.503	2.702	3.741	3.356	3.415	3.236	3.432	3.267
	RMSE	3.679	4.501	3.611	4.815	4.293	4.262	4.272	4.518	4.231
	bias	1.226	2.07	1.766	2.502	2.775	3.176	2.078	2.506	2.303
Plauborg model (daily values) 10cm	R^2	0.664	0.873	0.938	0.845	0.85	0.397	0.895	0.875	0.861
	MAE	1.837	1.755	1.105	1.908	1.547	1.687	1.627	1.78	1.621
	RMSE	2.418	2.206	1.395	2.42	1.978	2.143	2.063	2.26	2.074
	bias	-0.636	0.373	-0.046	0.667	1.108	1.535	0.396	0.921	0.608
Plauborg model (hourly values) 10cm	R^2	0.636	0.857	0.884	0.754	0.677	-0.193	0.872	0.828	0.794
	MAE	1.964	1.804	1.472	2.368	2.263	2.306	1.724	1.979	1.926
	RMSE	2.514	2.344	1.908	3.055	2.906	3.013	2.277	2.649	2.529
	bias	-0.349	0.237	0.558	1.363	0.677	0.845	0.163	0.065	0.597
BiLSTM 10cm	R^2	0.749	0.961	0.964	0.946	0.924	0.685	0.954	0.952	0.934
	MAE	1.479	1.006	0.858	1.026	1.106	1.171	1.126	1.112	1.111
	RMSE	2.08	1.221	1.061	1.429	1.405	1.557	1.356	1.401	1.423
	bias	-1.078	-0.167	-0.475	0.178	0.419	0.749	-0.115	0.459	0.06
LSTM 10cm	R^2	0.705	0.917	0.938	0.892	0.86	0.52	0.911	0.92	0.886
	MAE	1.725	1.495	1.106	1.477	1.463	1.466	1.561	1.433	1.472
	RMSE	2.254	1.773	1.388	2.014	1.906	1.923	1.894	1.809	1.871
	bias	-0.807	0.023	-0.047	0.537	0.549	0.877	0.043	0.464	0.302
GRU 10cm	R^2	0.696	0.906	0.933	0.893	0.914	0.639	0.915	0.904	0.894
	MAE	1.686	1.528	1.183	1.569	1.19	1.315	1.486	1.54	1.418
	RMSE	2.289	1.89	1.442	2.005	1.5	1.69	1.853	1.976	1.807
	bias	-1.195	-0.155	-0.719	0.018	0.589	1.031	-0.139	0.499	0.027
BiGRU 10cm	R^2	0.69	0.928	0.947	0.916	0.884	0.621	0.922	0.93	0.904
	MAE	1.687	1.369	1.049	1.34	1.377	1.336	1.438	1.311	1.36
	RMSE	2.31	1.659	1.288	1.777	1.736	1.732	1.774	1.693	1.722
	bias	-1.149	-0.313	-0.397	0.229	0.226	0.554	-0.304	0.077	-0.037

Table 8: Statitics of the models for depth 10

4 Results

Summerize results

Make sections

Include tables

En model av gangen

4.1 Linear model

The linear model scales the air temperature by a single factor making the model follow the form of the air temperature with the exception of Autumn where it diverges until April. After April the model fits better however it still has a high variance through out the years.

4.2 Plauborg daily

The Plauborg model follows strongly the soil temperature ant thereby has good prediction, with the exception of a few staitions. There are two stations that does not follow the same trend as the other stations, those are station Fåvang in year 2021 predicting soil temperatures at 10cm (se figure **??** and **??**) and Apelsvoll in year 2022 and year 2021 (se figure **??** and **??**).

In the region Østfold year 2021 there is a divergence at predicting the 10 cm soil temperature, this arises as an effect from an old code used to plot the graph of this model. In the old code the

model		41	118	42	30	39	15	26	11	average
Linear model 10cm	R^2	0.593	0.306	0.506	0.588	0.213	-0.003	0.456	0.51	0.423
	MAE	3.07	3.486	3.525	3.106	3.193	3.141	3.651	2.713	3.267
	RMSE	3.976	4.5	4.571	4.053	4.057	3.918	4.714	3.567	4.231
	bias	1.494	2.93	2.109	1.733	2.835	2.799	2.902	1.529	2.303
Plauborg model (daily values) 10cm	R^2	0.9	0.839	0.881	0.908	0.828	0.787	0.855	0.864	0.861
	MAE	1.587	1.697	1.75	1.519	1.455	1.428	1.937	1.504	1.621
	RMSE	1.975	2.165	2.239	1.914	1.896	1.806	2.43	1.879	2.074
	bias	-0.293	1.137	0.333	-0.046	1.193	1.114	1.251	0.339	0.608
Plauborg model (hourly values) 10cm	R^2	0.903	0.706	0.852	0.892	0.633	0.562	0.843	0.823	0.794
	MAE	1.519	2.322	1.885	1.555	2.151	1.991	1.918	1.666	1.926
	RMSE	1.938	2.928	2.501	2.072	2.77	2.589	2.532	2.146	2.529
	bias	0.151	1.639	0.703	0.354	0.701	0.903	0.821	0.038	0.597
BiLSTM 10cm	R^2	0.944	0.94	0.953	0.947	0.878	0.888	0.946	0.93	0.934
	MAE	1.203	1.071	1.103	1.132	1.239	1.018	1.2	1.084	1.111
	RMSE	1.473	1.324	1.403	1.443	1.592	1.3	1.487	1.344	1.423
	bias	-0.757	0.64	-0.169	-0.559	0.496	0.445	0.744	-0.118	0.06
LSTM 10cm	R^2	0.911	0.879	0.905	0.911	0.834	0.797	0.905	0.867	0.886
	MAE	1.544	1.412	1.591	1.518	1.397	1.371	1.585	1.503	1.472
	RMSE	1.85	1.872	2.002	1.876	1.855	1.754	1.964	1.854	1.871
	bias	-0.372	1.047	0.199	-0.182	0.616	0.669	0.869	-0.048	0.302
GRU 10cm	R^2	0.898	0.911	0.899	0.91	0.884	0.87	0.893	0.885	0.894
	MAE	1.6	1.288	1.645	1.507	1.214	1.086	1.654	1.367	1.418
	RMSE	1.983	1.619	2.062	1.881	1.553	1.404	2.081	1.722	1.807
	bias	-0.937	0.439	-0.309	-0.695	0.671	0.533	0.765	-0.189	0.027
BiGRU 10cm	R^2	0.915	0.918	0.921	0.922	0.867	0.839	0.925	0.875	0.904
	MAE	1.498	1.196	1.455	1.436	1.286	1.256	1.386	1.467	1.36
	RMSE	1.81	1.552	1.818	1.76	1.659	1.562	1.739	1.796	1.722
	bias	-0.716	0.709	-0.139	-0.507	0.283	0.334	0.51	-0.375	-0.037

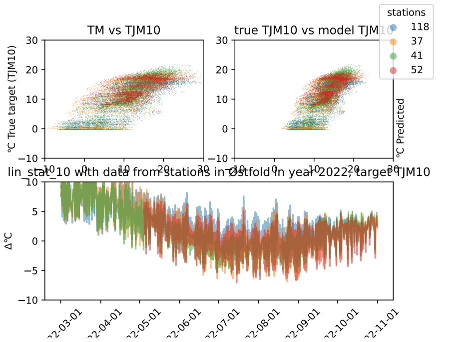
Table 9: Statitics of the models for depth 10

model		52	37	50	38	57	34	27	17	average
Linear model 20cm	R^2	0.604	0.391	0.434	0.308	0.125	-2.248	0.415	0.336	0.308
	MAE	2.841	3.675	3.025	3.682	3.636	3.675	3.547	3.84	3.474
	RMSE	3.556	4.754	4.048	4.832	4.655	4.583	4.672	5.049	4.504
	bias	0.559	2.174	2.207	2.601	3.153	3.471	2.535	2.939	2.487
Plauborg model (daily values) 20cm	R^2	0.83	0.905	0.946	0.912	0.863	0.316	0.901	0.844	0.876
	MAE	1.873	1.496	0.985	1.367	1.538	1.836	1.534	1.911	1.536
	RMSE	2.33	1.877	1.251	1.721	1.841	2.104	1.924	2.448	1.91
	bias	-1.402	0.353	0.096	0.515	1.427	1.816	0.753	1.401	0.644
Plauborg model (hourly values) 20cm	R^2	0.803	0.83	0.836	0.736	0.617	-0.547	0.839	0.762	0.756
	MAE	1.976	1.938	1.7	2.323	2.419	2.427	1.885	2.265	2.06
	RMSE	2.504	2.513	2.176	2.983	3.079	3.163	2.455	3.023	2.676
	bias	-1.2	0.067	0.815	1.149	0.744	0.797	0.335	0.137	0.528
BiLSTM 20cm	R^2	0.803	0.92	0.95	0.929	0.896	0.589	0.925	0.907	0.901
	MAE	2.067	1.439	0.986	1.244	1.233	1.233	1.385	1.508	1.349
	RMSE	2.495	1.716	1.201	1.54	1.6	1.64	1.664	1.884	1.695
	bias	-1.788	-0.255	-0.131	0.189	0.534	0.808	0.13	0.589	0.068
LSTM 20cm	R^2	0.826	0.935	0.958	0.947	0.873	0.413	0.915	0.853	0.893
	MAE	1.985	1.281	0.88	1.082	1.334	1.538	1.363	1.797	1.363
	RMSE	2.346	1.547	1.092	1.326	1.766	1.959	1.772	2.374	1.762
	bias	-1.611	0.131	-0.165	0.31	1.113	1.48	0.597	1.304	0.423
GRU 20cm	R^2	0.8	0.946	0.961	0.955	0.899	0.502	0.929	0.857	0.901
	MAE	2.148	1.146	0.854	0.985	1.194	1.489	1.2	1.634	1.294
	RMSE	2.517	1.414	1.06	1.224	1.575	1.831	1.626	2.335	1.7
	bias	-1.806	-0.036	-0.355	0.145	0.982	1.339	0.362	1.106	0.244
BiGRU 20cm	R^2	0.792	0.934	0.961	0.943	0.905	0.63	0.938	0.931	0.915
	MAE	2.131	1.283	0.825	1.088	1.198	1.228	1.205	1.208	1.236
	RMSE	2.565	1.562	1.051	1.381	1.527	1.578	1.52	1.629	1.575
	bias	-1.905	-0.397	-0.151	0.113	0.549	0.807	0.037	0.434	0.012

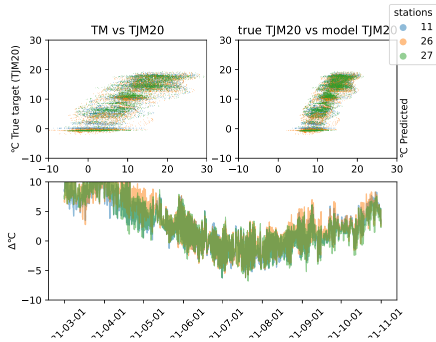
Table 10: Statitics of the models for depth 20

model		41	118	42	30	39	15	26	11	average
Linear model 20cm	R^2	0.491	0.162	0.393	0.456	0.081	-0.241	0.302	0.35	0.308
	MAE	3.286	3.654	3.741	3.433	3.39	3.342	4.009	2.924	3.474
	RMSE	4.248	4.726	4.863	4.465	4.31	4.198	5.17	3.821	4.504
	bias	1.677	3.208	2.364	2.015	3.083	3.006	3.282	1.692	2.487
Plauborg model (daily values) 20cm	R^2	0.914	0.886	0.901	0.91	0.852	0.801	0.833	0.866	0.876
	MAE	1.409	1.384	1.54	1.471	1.468	1.411	2.101	1.443	1.536
	RMSE	1.748	1.742	1.966	1.817	1.729	1.681	2.528	1.735	1.91
	bias	-0.371	1.139	0.346	-0.014	1.402	1.215	1.578	0.463	0.644
Plauborg model (hourly values) 20cm	R^2	0.872	0.656	0.807	0.859	0.615	0.484	0.801	0.755	0.756
	MAE	1.665	2.422	2.099	1.708	2.186	2.094	2.105	1.847	2.06
	RMSE	2.135	3.029	2.739	2.276	2.79	2.706	2.757	2.346	2.676
	bias	0.13	1.722	0.746	0.428	0.633	0.827	0.892	-0.023	0.528
BiLSTM 20cm	R^2	0.913	0.915	0.918	0.923	0.887	0.86	0.904	0.885	0.901
	MAE	1.455	1.162	1.463	1.379	1.123	1.08	1.567	1.333	1.349
	RMSE	1.745	1.497	1.781	1.67	1.504	1.405	1.91	1.607	1.695
	bias	-0.675	0.82	-0.034	-0.396	0.464	0.435	0.894	-0.207	0.068
LSTM 20cm	R^2	0.926	0.917	0.931	0.929	0.859	0.812	0.859	0.838	0.893
	MAE	1.339	1.119	1.343	1.308	1.253	1.203	1.798	1.419	1.363
	RMSE	1.609	1.48	1.637	1.606	1.678	1.624	2.318	1.903	1.762
	bias	-0.577	0.921	0.163	-0.176	1.006	0.887	1.491	0.428	0.423
GRU 20cm	R^2	0.924	0.941	0.931	0.924	0.878	0.836	0.874	0.844	0.901
	MAE	1.358	0.955	1.307	1.325	1.217	1.138	1.595	1.344	1.294
	RMSE	1.638	1.258	1.628	1.66	1.563	1.518	2.187	1.869	1.7
	bias	-0.779	0.727	-0.032	-0.368	0.917	0.763	1.246	0.244	0.244
BiGRU 20cm	R^2	0.93	0.932	0.935	0.934	0.902	0.878	0.926	0.904	0.915
	MAE	1.319	1.063	1.259	1.275	1.095	1.052	1.314	1.199	1.236
	RMSE	1.574	1.351	1.585	1.553	1.404	1.311	1.677	1.467	1.575
	bias	-0.749	0.756	-0.096	-0.464	0.514	0.486	0.763	-0.219	0.012

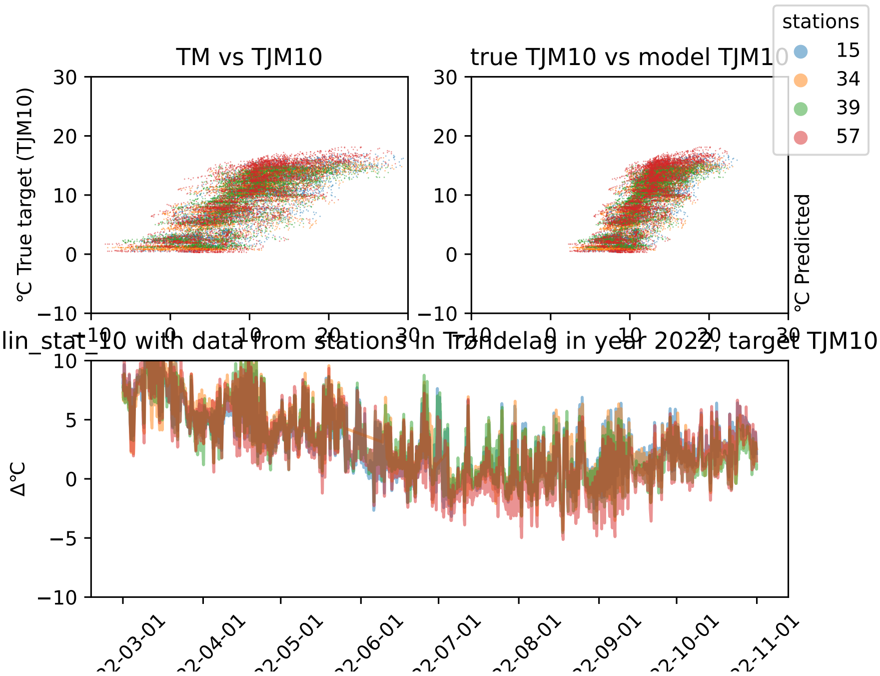
Table 11: Statitics of the models for depth 20



(a) Differnce plot for daily Plauborg model in year 2022 and region Østfold



(b) Worst plot for Linear Regression model in year 2021 and region Innlandet



(c) Average plot for Linear Regression model in year 2022 and region Trøndelag

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	4.504	3.474	2.487	-0.796	-1	0.308
region	Østfold	4.348	3.363	1.901	-0.796	-1	0.424
region	Vestfold	4.564	3.47	2.297	-0.796	-1	0.397
region	Trøndelag	4.438	3.508	3.175	-0.796	-1	-0.194
region	Innlandet	4.688	3.568	2.601	-0.796	-1	0.353
local	52	3.556	2.841	0.559	-0.796	-1	0.604
local	41	4.248	3.286	1.677	-0.796	-1	0.491
local	37	4.754	3.675	2.174	-0.796	-1	0.391
local	118	4.726	3.654	3.208	-0.796	-1	0.162
local	50	4.048	3.025	2.207	-0.796	-1	0.434
local	42	4.863	3.741	2.364	-0.796	-1	0.393
local	38	4.832	3.682	2.601	-0.796	-1	0.308
local	30	4.465	3.433	2.015	-0.796	-1	0.456
local	57	4.655	3.636	3.153	-0.796	-1	0.125
local	39	4.31	3.39	3.083	-0.796	-1	0.081
local	34	4.583	3.675	3.471	-0.796	-1	-2.248
local	15	4.198	3.342	3.006	-0.796	-1	-0.241
local	27	4.672	3.547	2.535	-0.796	-1	0.415
local	26	5.17	4.009	3.282	-0.796	-1	0.302
local	17	5.049	3.84	2.939	-0.796	-1	0.336
local	11	3.821	2.924	1.692	-0.796	-1	0.35

(a) Performance table for Linear regression 20cm

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	4.231	3.267	2.303	-0.638	-1	0.423
region	Østfold	4.236	3.28	2.015	-0.638	-1	0.45
region	Vestfold	4.277	3.26	2.019	-0.638	-1	0.517
region	Trøndelag	4.133	3.274	2.893	-0.638	-1	0.028
region	Innlandet	4.282	3.254	2.246	-0.638	-1	0.504
local	52	3.679	2.889	1.226	-0.638	-1	0.221
local	41	3.976	3.07	1.494	-0.638	-1	0.593
local	37	4.501	3.503	2.07	-0.638	-1	0.473
local	118	4.5	3.486	2.93	-0.638	-1	0.306
local	50	3.611	2.702	1.766	-0.638	-1	0.584
local	42	4.571	3.525	2.109	-0.638	-1	0.506
local	38	4.815	3.741	2.502	-0.638	-1	0.388
local	30	4.053	3.106	1.733	-0.638	-1	0.588
local	57	4.293	3.356	2.775	-0.638	-1	0.295
local	39	4.057	3.193	2.835	-0.638	-1	0.213
local	34	4.262	3.415	3.176	-0.638	-1	-1.386
local	15	3.918	3.141	2.799	-0.638	-1	-0.003
local	27	4.272	3.236	2.078	-0.638	-1	0.551
local	26	4.714	3.651	2.902	-0.638	-1	0.456
local	17	4.518	3.432	2.506	-0.638	-1	0.501
local	11	3.567	2.713	1.529	-0.638	-1	0.51

(b) Performance table for Linear regression 10cm

Table 12: Performance table for Linear Regession at 10 cm depth and 20 cm depth.

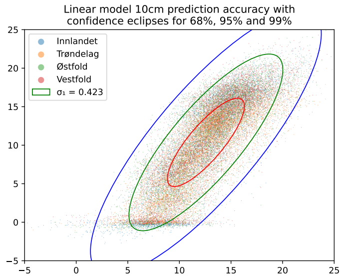
scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.91	1.536	0.644	-1.911	-2	0.876
region	Østfold	1.94	1.541	-0.073	-1.914	-2	0.885
region	Vestfold	1.71	1.341	0.236	-1.912	-2	0.915
region	Trøndelag	1.843	1.56	1.461	-1.918	-2	0.794
region	Innlandet	2.16	1.735	1.02	-1.903	-2	0.863
local	52	2.33	1.873	-1.402	-1.905	-2	0.83
local	41	1.748	1.409	-0.371	-1.919	-2	0.914
local	37	1.877	1.496	0.353	-1.912	-2	0.905
local	118	1.742	1.384	1.139	-1.914	-2	0.886
local	50	1.251	0.985	0.096	-1.916	-2	0.946
local	42	1.966	1.54	0.346	-1.911	-2	0.901
local	38	1.721	1.367	0.515	-1.912	-2	0.912
local	30	1.817	1.471	-0.014	-1.914	-2	0.91
local	57	1.841	1.538	1.427	-1.913	-2	0.863
local	39	1.729	1.468	1.402	-1.917	-2	0.852
local	34	2.104	1.836	1.816	-1.91	-2	0.316
local	15	1.681	1.411	1.215	-1.913	-2	0.801
local	27	1.924	1.534	0.753	-1.915	-2	0.901
local	26	2.528	2.101	1.578	-1.909	-2	0.833
local	17	2.448	1.911	1.401	-1.914	-2	0.844
local	11	1.735	1.443	0.463	-1.905	-2	0.866

(a) Performance table for daily values Plauborg model 20cm

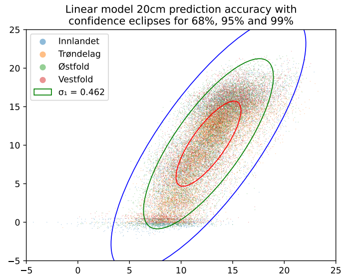
scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	2.074	1.621	0.608	-1.266	-2	0.861
region	Østfold	2.168	1.704	0.24	-1.268	-2	0.856
region	Vestfold	2.022	1.564	0.219	-1.265	-2	0.892
region	Trøndelag	1.957	1.528	1.235	-1.264	-2	0.782
region	Innlandet	2.165	1.71	0.714	-1.267	-2	0.873
local	52	2.418	1.837	-0.636	-1.265	-2	0.664
local	41	1.975	1.587	-0.293	-1.268	-2	0.9
local	37	2.206	1.755	0.373	-1.265	-2	0.873
local	118	2.165	1.697	1.137	-1.271	-2	0.839
local	50	1.395	1.105	-0.046	-1.261	-2	0.938
local	42	2.239	1.75	0.333	-1.257	-2	0.881
local	38	2.42	1.908	0.667	-1.263	-2	0.845
local	30	1.914	1.519	-0.046	-1.265	-2	0.908
local	57	1.978	1.547	1.108	-1.248	-2	0.85
local	39	1.896	1.455	1.193	-1.262	-2	0.828
local	34	2.143	1.687	1.535	-1.262	-2	0.397
local	15	1.806	1.428	1.114	-1.262	-2	0.787
local	27	2.063	1.627	0.396	-1.263	-2	0.895
local	26	2.43	1.937	1.251	-1.262	-2	0.855
local	17	2.26	1.78	0.921	-1.262	-2	0.875
local	11	1.879	1.504	0.339	-1.26	-2	0.864

(b) Performance table for Linear regression 10cm

Table 13: Performance table for daily values Plauborg model at 10 cm depth and 20 cm depth.



(a) The plot shows the ground truth against the predicted value of Linear Regression. The eclipses demonstrates the 65% (inner eclipse), 95% (middle eclipse), and 99% (outer eclipse) confidence interval for the model.



(b) The plot shows the ground truth against the predicted value of Linear Regression. The eclipses demonstrates the 65% (inner eclipse), 95% (middle eclipse), and 99% (outer eclipse) confidence interval for the model.

Figure 9: Plots showing ground truth vs predicted values from the linear Regression models with their 65%, 95%, and 99% conficence eclipses. The conficense eclipses function the same as confidence intervals but in two dimentions, meaning the area covered by the eclipse denotes where one can expect the points to be with a given confidence procence.

NULL values was set by default to 0 meaning when calculation the difference between predicted values and ground truth it in essence created a copy of the air temperature.

4.3 Plauborg hourly

The hourly Plauborg shows high variance but promising explained variance. During May and April the model shows a rise in prediction difference, likely due to the effect of snow keeping the temperature relative constant and therefore getting a copy-effect of the air temperature.

4.4 Long Short Term-Memory (LSTM)

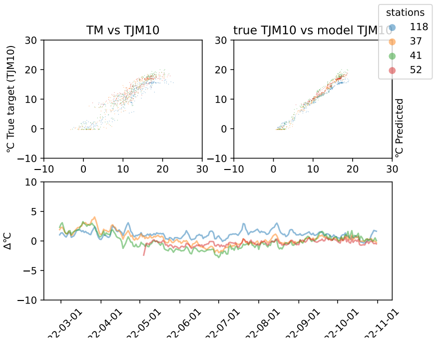
The LSTM shows a great fit to the data with the exception of the May/April month where there is a trend of constant over estimation with a constant value with few stations showing a spike in difference error in April.

4.5 Bi-Directional Long Short Term-Memory (LSTM)

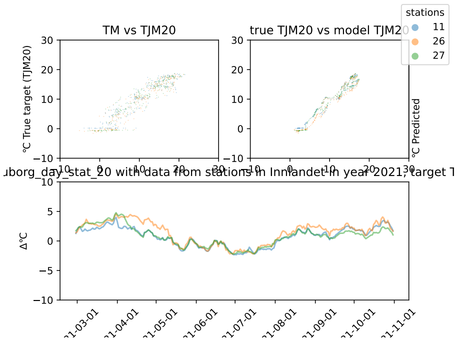
The LSTM shows a great fit to the data with the exception of the May/April month where there is a high variance in the transition from one month to the other. Østfold does have subpar performance, in particular station Øsaker and Tomb. Those two consistently either over estimate or under estimate the soil temperature compared to the other stations in the same region (Rakkestad, and Rygge).

4.6 Gated Recurent Unit (GRU)

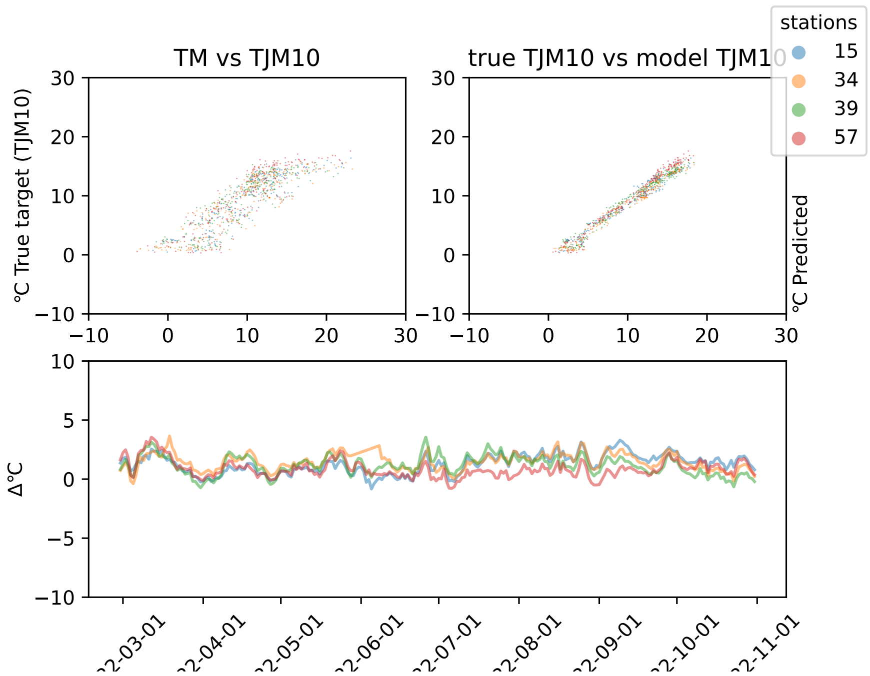
The GRU model demonstrates encouraging outcome with few notable stations and year. One of which is station Øsaker who over estimates the soil temperature at both depth (10 cm and 20 cm).



(a) Differnce plot for daily Plauborg model in year 2022 and region Østfold



(b) Worst plot for daily values Plauborg model in year 2021 and region Innlandet



(c) Average plot for daily values Plauborg model in year 2022 and region Trøndelag

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	2.676	2.06	0.528	-0.323	-1	0.756
region	Østfold	2.564	2	0.176	-0.321	-1	0.8
region	Vestfold	2.565	1.958	0.785	-0.328	-1	0.81
region	Trøndelag	2.938	2.279	0.75	-0.319	-1	0.477
region	Innlandet	2.612	1.997	0.379	-0.324	-1	0.799
local	52	2.504	1.976	-1.2	-0.322	-1	0.803
local	41	2.135	1.665	0.13	-0.327	-1	0.872
local	37	2.513	1.938	0.067	-0.33	-1	0.83
local	118	3.029	2.422	1.722	-0.327	-1	0.656
local	50	2.176	1.7	0.815	-0.328	-1	0.836
local	42	2.739	2.099	0.746	-0.329	-1	0.807
local	38	2.983	2.323	1.149	-0.328	-1	0.736
local	30	2.276	1.708	0.428	-0.327	-1	0.859
local	57	3.079	2.419	0.744	-0.33	-1	0.617
local	39	2.79	2.186	0.633	-0.328	-1	0.615
local	34	3.163	2.427	0.797	-0.319	-1	-0.547
local	15	2.706	2.094	0.827	-0.326	-1	0.484
local	27	2.455	1.885	0.335	-0.325	-1	0.839
local	26	2.757	2.105	0.892	-0.33	-1	0.801
local	17	3.023	2.265	0.137	-0.317	-1	0.762
local	11	2.346	1.847	-0.023	-0.326	-1	0.755

(a) Performance table for daily values Plauborg model 20cm

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	2.529	1.926	0.597	-0.443	-1	0.794
region	Østfold	2.448	1.894	0.512	-0.444	-1	0.816
region	Vestfold	2.412	1.81	0.733	-0.436	-1	0.846
region	Trøndelag	2.822	2.176	0.781	-0.446	-1	0.547
region	Innlandet	2.382	1.805	0.312	-0.447	-1	0.847
local	52	2.514	1.964	-0.349	-0.439	-1	0.636
local	41	1.938	1.519	0.151	-0.445	-1	0.903
local	37	2.344	1.804	0.237	-0.443	-1	0.857
local	118	2.928	2.322	1.639	-0.436	-1	0.706
local	50	1.908	1.472	0.558	-0.435	-1	0.884
local	42	2.501	1.885	0.703	-0.442	-1	0.852
local	38	3.055	2.368	1.363	-0.443	-1	0.754
local	30	2.072	1.555	0.354	-0.445	-1	0.892
local	57	2.906	2.263	0.677	-0.444	-1	0.677
local	39	2.77	2.151	0.701	-0.45	-1	0.633
local	34	3.013	2.306	0.845	-0.45	-1	-0.193
local	15	2.589	1.991	0.903	-0.441	-1	0.562
local	27	2.277	1.724	0.163	-0.443	-1	0.872
local	26	2.532	1.918	0.821	-0.444	-1	0.843
local	17	2.649	1.979	0.065	-0.447	-1	0.828
local	11	2.146	1.666	0.038	-0.435	-1	0.823

(b) Performance table for hourly values Plauborg model 10cm

Table 14: Performance table for hourly values Plauborg model at 10 cm depth and 20 cm depth.

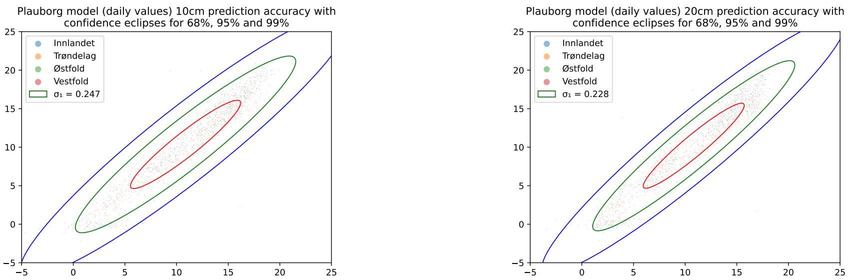
scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.695	1.349	0.068	-1.797	-2	0.901
region	Østfold	1.902	1.532	-0.479	-1.831	-2	0.889
region	Vestfold	1.563	1.268	-0.093	-1.796	-2	0.929
region	Trøndelag	1.538	1.166	0.557	-1.787	-2	0.856
region	Innlandet	1.76	1.444	0.339	-1.786	-2	0.908
local	52	2.495	2.067	-1.788	-1.819	-2	0.803
local	41	1.745	1.455	-0.675	-1.755	-2	0.913
local	37	1.716	1.439	-0.255	-1.813	-2	0.92
local	118	1.497	1.162	0.82	-1.814	-2	0.915
local	50	1.201	0.986	-0.131	-1.832	-2	0.95
local	42	1.781	1.463	-0.034	-1.814	-2	0.918
local	38	1.54	1.244	0.189	-1.808	-2	0.929
local	30	1.67	1.379	-0.396	-1.818	-2	0.923
local	57	1.6	1.233	0.534	-1.823	-2	0.896
local	39	1.504	1.123	0.464	-1.83	-2	0.887
local	34	1.64	1.233	0.808	-1.819	-2	0.589
local	15	1.405	1.08	0.435	-1.807	-2	0.86
local	27	1.664	1.385	0.13	-1.822	-2	0.925
local	26	1.91	1.567	0.894	-1.801	-2	0.904
local	17	1.884	1.508	0.589	-1.835	-2	0.907
local	11	1.607	1.333	-0.207	-1.805	-2	0.885

(a) Performance table for BiLSTM model 20cm

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.423	1.111	0.06	-1.858	-2	0.934
region	Østfold	1.483	1.154	-0.252	-1.869	-2	0.932
region	Vestfold	1.341	1.03	-0.264	-1.832	-2	0.952
region	Trøndelag	1.467	1.133	0.524	-1.864	-2	0.877
region	Innlandet	1.4	1.135	0.225	-1.896	-2	0.947
local	52	2.08	1.479	-1.078	-1.838	-2	0.749
local	41	1.473	1.203	-0.757	-1.891	-2	0.944
local	37	1.221	1.006	-0.167	-1.88	-2	0.961
local	118	1.324	1.071	0.64	-1.893	-2	0.94
local	50	1.061	0.858	-0.475	-1.859	-2	0.964
local	42	1.403	1.103	-0.169	-1.861	-2	0.953
local	38	1.429	1.026	0.178	-1.878	-2	0.946
local	30	1.443	1.132	-0.559	-1.87	-2	0.947
local	57	1.405	1.106	0.419	-1.813	-2	0.924
local	39	1.592	1.239	0.496	-1.84	-2	0.878
local	34	1.557	1.171	0.749	-1.829	-2	0.685
local	15	1.3	1.018	0.445	-1.836	-2	0.888
local	27	1.356	1.126	-0.115	-1.836	-2	0.954
local	26	1.487	1.2	0.744	-1.848	-2	0.946
local	17	1.401	1.112	0.459	-1.875	-2	0.952
local	11	1.344	1.084	-0.118	-1.869	-2	0.93

(b) Performance table for BiLSTM model 10cm

Table 15: Performance table for BiLSTM model at 10 cm depth and 20 cm depth.



(a) The plot shows the ground truth against the predicted value of daily values Plauborg model. The eclipses demonstrates the 65% (inner eclipse), 95% (middle eclipse), and 99% (outer eclipse) confidence interval for the model.

(b) The plot shows the ground truth against the predicted value of daily values Plauborg model. The eclipses demonstrates the 65% (inner eclipse), 95% (middle eclipse), and 99% (outer eclipse) confidence interval for the model.

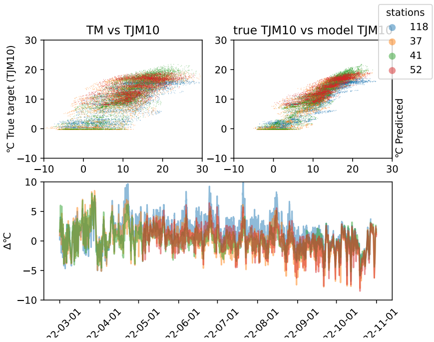
Figure 11: Plots showing ground truth vs predicted values from the daily values Plauborg models with their 65%, 95%, and 99% conficence eclipses. The conficense eclipses function the same as confidence intervals but in two dimentions, meaning the area covered by the eclipse denotes where one can expect the points to be with a given confidence procence.

All station shows an seasonal sensitivity, as all stations tend to incrementally over estimate in May before falling down towards 0 $\Delta°C$ around April month.

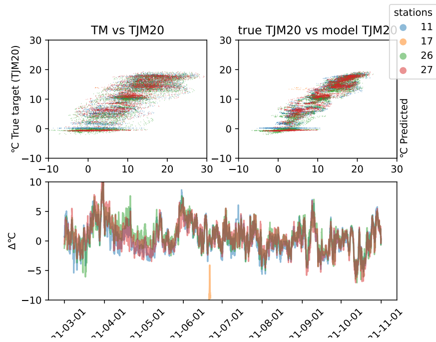
4.7 Bi-directional Gated Recurent Unit (GRU)

4.8 Deep learning models

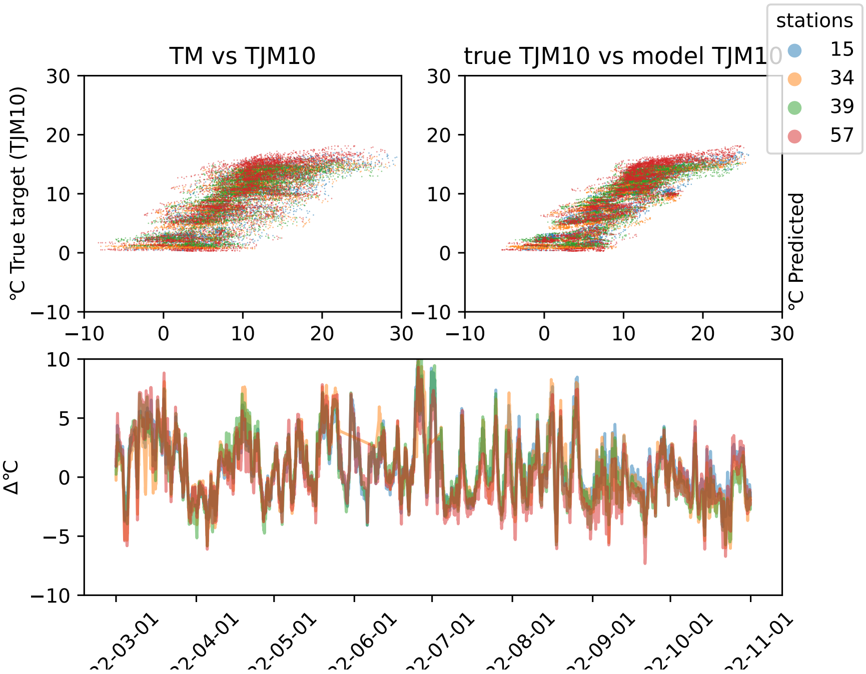
LSTM based models has a higher RMSE than the GRU model that performs similary to the Plauborg models in its performance while the rest (BiLSTM, and LSTM) inhibits an Autumn discrepancy, see section 5.1. However this discrepancy is less prevalent in the more advance models (BiLSTM, and GRU) but still visible compared to the rest of the year.



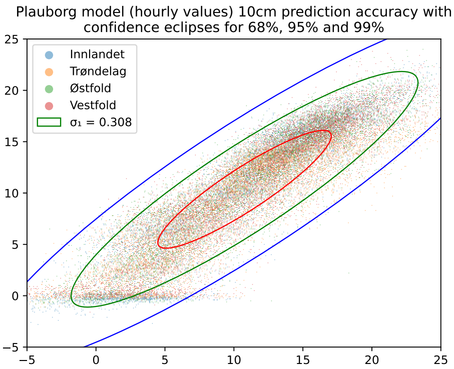
(a) Differnce plot for hourly Plauborg model in year 2022 and region Østfold



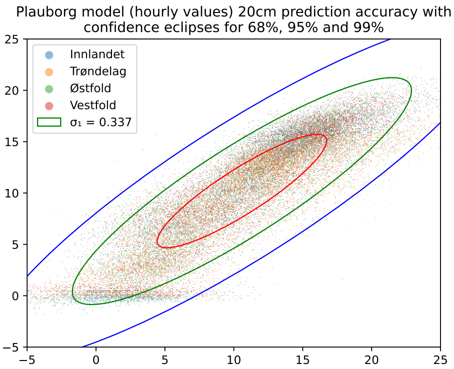
(b) Worst plot for hourly values Plauborg model in year 2021 and region Innlandet



(c) Average plot for hourly values Plauborg model in year 2022 and region Trøndelag



(a) The plot shows the ground truth against the predicted value of hourly values Plauborg model. The eclipses demonstrates the 65% (inner eclipse), 95% (middle eclipse), and 99% (outer eclipse) confidence interval for the model.



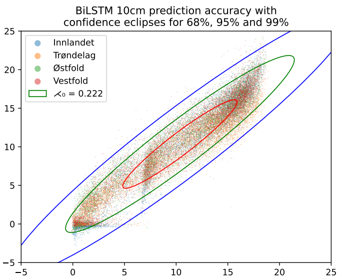
(b) The plot shows the ground truth against the predicted value of hourly values Plauborg model. The eclipses demonstrates the 65% (inner eclipse), 95% (middle eclipse), and 99% (outer eclipse) confidence interval for the model.

Figure 13: Plots showing ground truth vs predicted values from the hourly values Plauborg models with their 65%, 95%, and 99% conficence eclipses. The conficense eclipses function the same as confidence intervals but in two dimentions, meaning the area covered by the eclipse denotes where one can expect the points to be with a given confidence procence.

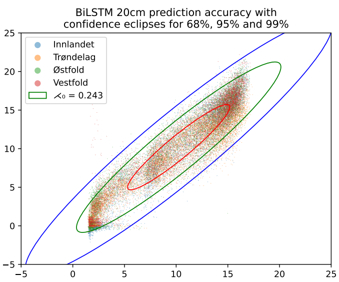
(a) Differnce plot for BiLSTM model in year 2022 and region Østfold

(b) Worst plot for BiLSTM model in year 2021 and region Innlandet

(c) Average plot for BiLSTM model in year 2022 and region Trøndelag

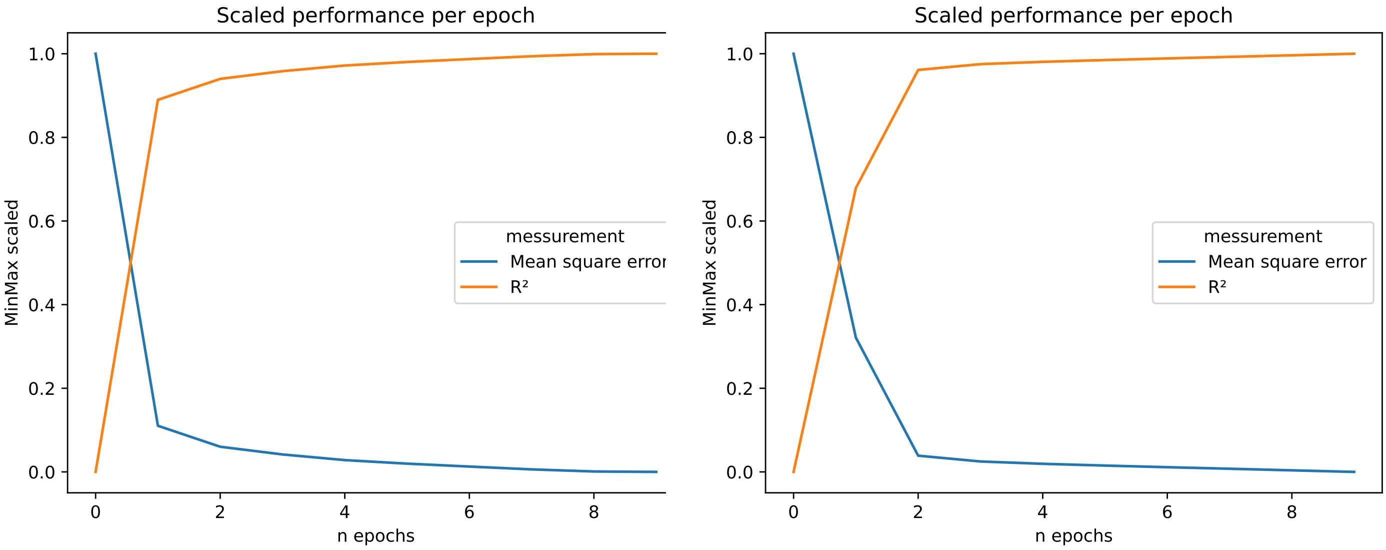


(a) The plot shows the ground truth against the predicted value of BiLSTM model. The eclipses demonstrates the 65% (inner eclipse), 95% (middle eclipse), and 99% (outer eclipse) confidence interval for the model.



(b) The plot shows the ground truth against the predicted value of BiLSTM model. The eclipses demonstrates the 65% (inner eclipse), 95% (middle eclipse), and 99% (outer eclipse) confidence interval for the model.

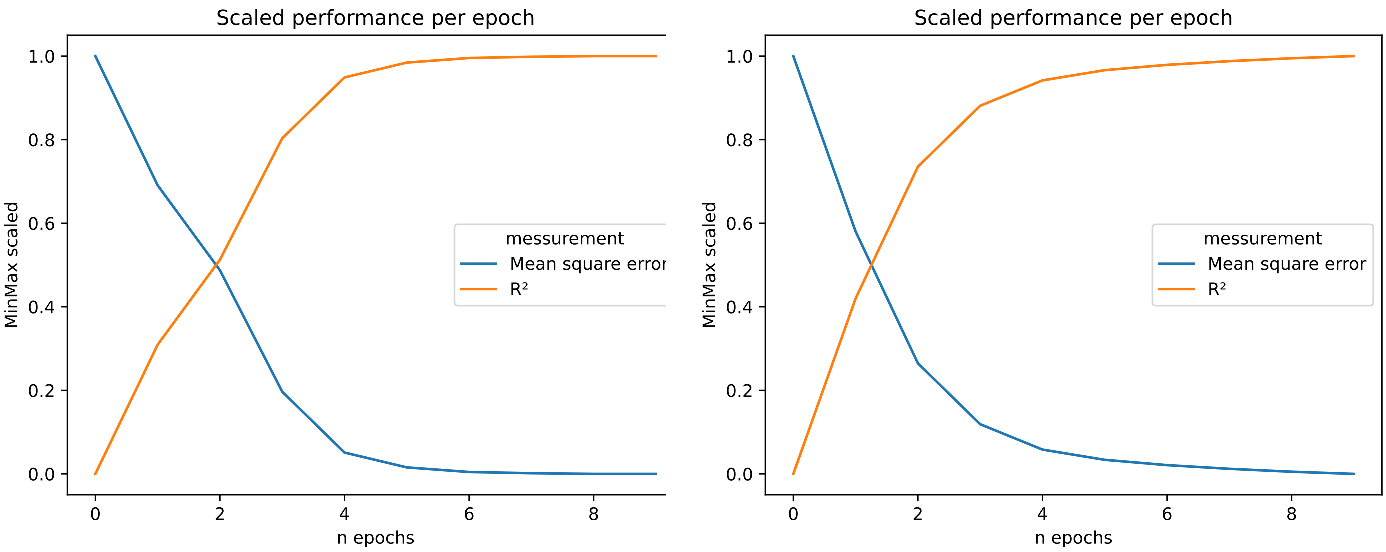
Figure 15: Plots showing ground truth vs predicted values from the BiLSTM models with their 65%, 95%, and 99% conficence eclipses. The confidence eclipses function the same as confidence intervals but in two dimensions, meaning the area covered by the eclipse denotes where one can expect the points to be with a given confidence procence.



(a) Graf of BiLSTM 10cm performance per epoch.

(b) Graf of BiLSTM 20cm performance per epoch.

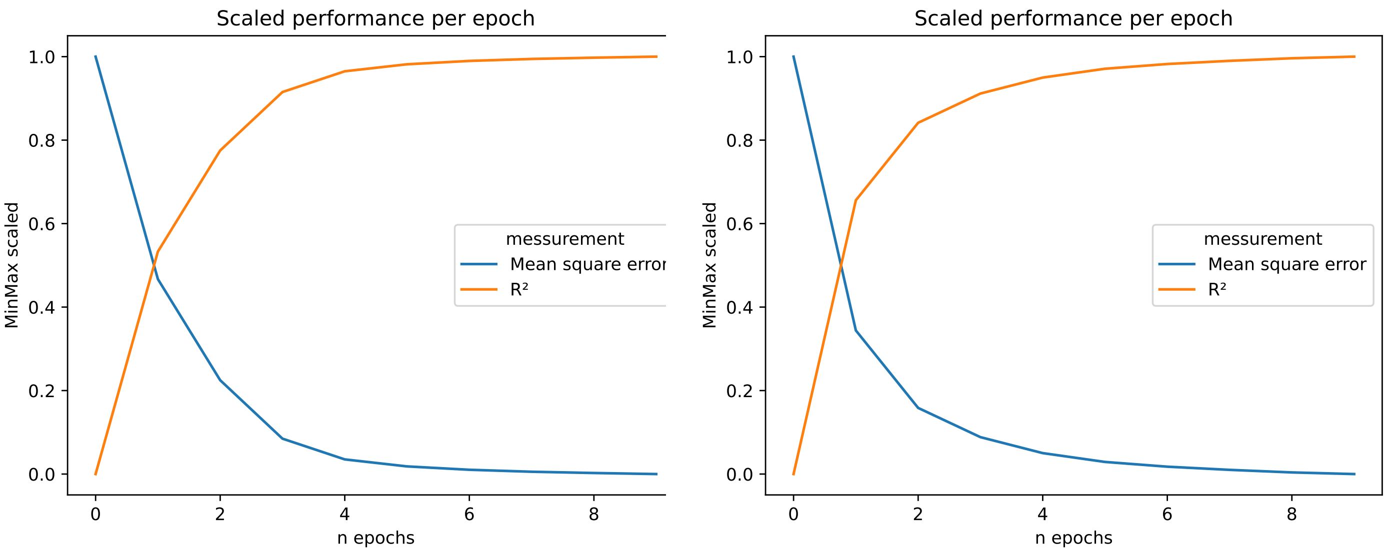
Figure 16: Performance graphs displaying the developments of Mean Square Error and Explained Variance (R^2) for each epoch.



(a) Graf of GRU 10cm performance per epoch.

(b) Graf of GRU 20cm performance per epoch.

Figure 17: Performance graphs displaying the developments of Mean Square Error and Explained Variance (R^2) for each epoch.



(a) Graf of BiGRU 10cm performance per epoch.

(b) Graf of BiGRU 20cm performance per epoch.

Figure 18: Performance graphs displaying the developments of Mean Square Error and Explained Variance (R^2) for each epoch.

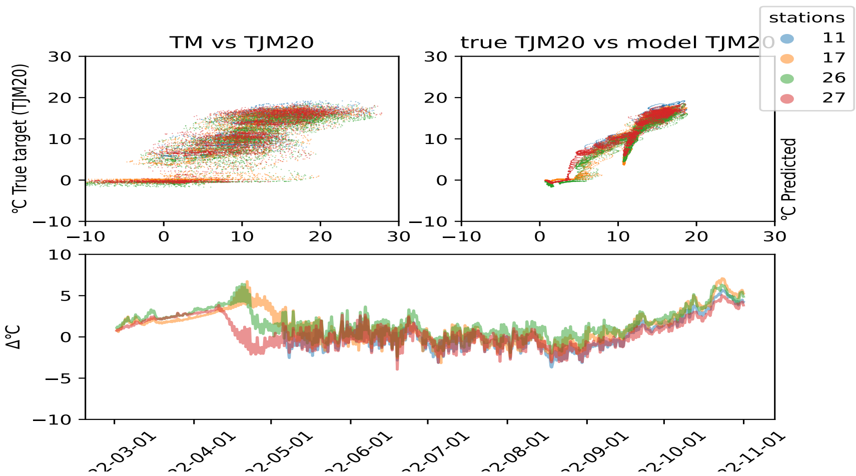


Figure 19: LSTM model applied at stations in the region Innlandet in 2022 with 10cm soil temperature as target.

The number of epochs was fixed at 10, however the performance graphs shows that after 4 epochs the performance tend to stabilise, and after 6 epochs it becomes small gains in performance.

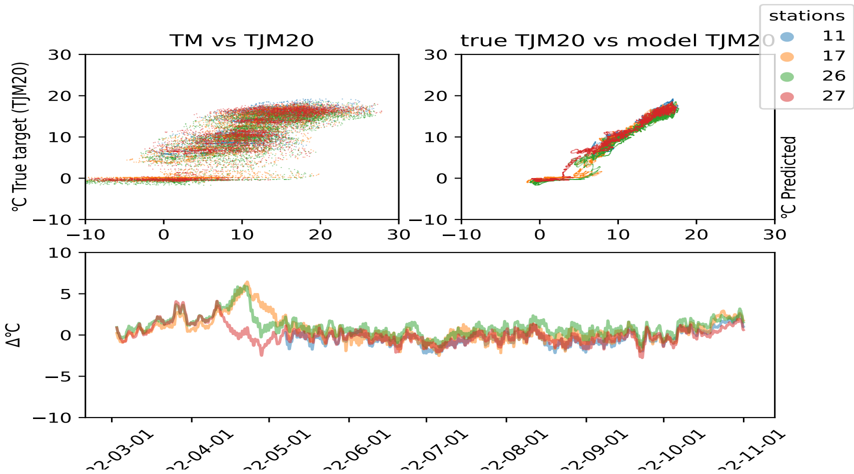


Figure 20: GRU model applied at staions in the region Innladet in 2022 with 10cm soil temperature as target.

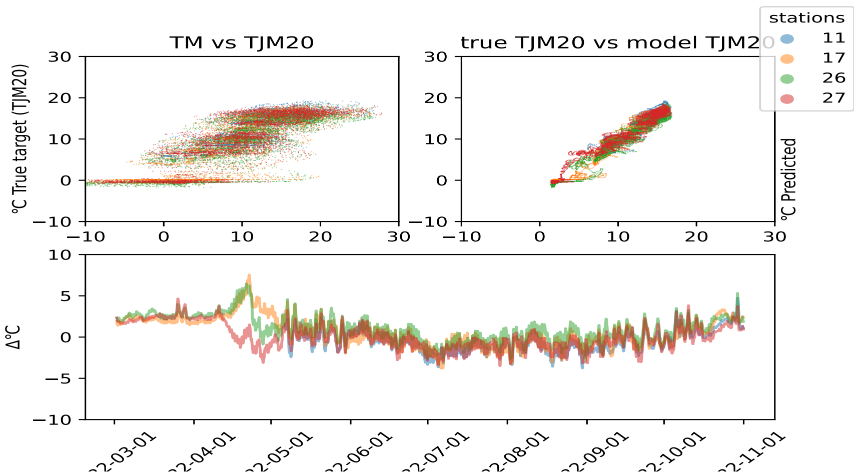


Figure 21: BiLSTM applied at stations in the region Innlandet in 2022 with 10cm soil temperature as target.

5 Discussion

Write a discussion

5.1 The Autumn discrepancy

metion annomelies and resons

Show table of parameters

A phenomenon that arose during performance evaluations was that the linear models struggles with the Autumn season. The difference graphs showed a clear over or under estimation that are larger than 10σ. When investigating the coeffisents to the model this discrepancy can be contributed to the intercept that during low temperature ($< 5°C$) giving either an over estimation or an under approximation. Further more when removing the calculation of the intersept the same phenomenon is still precent possibility due to the adaptation to Summer season.

If one looks at the deep learning models shown at figure 21 to figure 19 this effect can be observed. The GRU model is affected similarly, but to a lesser exstent where it overestimates at a few periods during the Mars month. This can be due to snow covering the soil forming an thermal isolator keeping the soil temperature constant while the air temperature is fluxiating at relative normal rates giving the models a false sense of generality when predicting this period, however the GRU model seems to interoperate a sense of season awareness allowing it to have a more constant temperature prediction in the autumn and return to "normal" operation during Summer and Spring.

5.1.1 Temperature seasons

It can be observed in the raw data and in the diff plots (see figure 21 as an example) that the seasons can be shown though the temperatures and stable periods. The flat areas is explained by snow covers that creates an thermal isolation layer that dampens the effekt of the air temperature. There excist models in the literature that takes this effect into account[8, 34] to make a more accurate prediction of soil temperature when there is snow present or frozen soil.

5.2 Plauborg

The result of the modelling (table 20 to 23) show that modelig soil temperature without the inclusion of time is an inefficient, and inaccurate method of predicting soil temperatures.

In this study the original model, that was trained for daily values was converted to predict hourly data to see if the same formulation could be used to make predictions. When comparing the results shown in table 19 and table 18 to their daily counterpart it shows similar values showing that the model proposed in Plauborg can be extended to hourliy timeseries.

5.3 Discussion of good results of Plauborg

The inclusion of previous temperatures gives an improved estimation, even on hourly basis. The coefficents for both daily and hourly are observed to be <1 making it a mean temperature and the fourier terms would estimate the soil function(6)[35].

$$E_{\text{soil,year}}(T) + e^{-z/D}\sin(\omega t - z/D + \phi) \tag{6}$$

Since the term $exp(z/D)$ is constant we would be estimating $\sin(\omega t - Q) = \sin(\omega t)\cos(Q) - \sin(Q)\cos(\omega t)$, where Q is $z/D - \phi$ and is considered constant. This will be extrapolated to a simple sum of sines and cosines as the model does. Together the Plauborg model would estimate the approximation (8). The extra terms are nessesery to include external factors that affects the temperature (rain, soil type, atmosphere, etc).

$$E_{\text{soil,year}}(T) + e^{-z/D}\sin(\omega t - z/D + \phi) \tag{7}$$
$$\approx E_{\text{air,period}}(T) + \sum \alpha_i e^{z/D}\sin(-Q_i)\cos(i\omega t_i) + \sum \beta_j e^{z/D}\cos(-Q_j)\sin(i\omega t_j) \tag{8}$$

The differences between actual and predicted values could be due to the varying soil types at different stations. To create a more accurate model, one might need to consider additional weather-related measurements (such as air pressure, humidity, soil type, and texture) or incorporate non-linear features (such as the square root of temperature or the temperature change ratio over time)

5.4 Linear regression vs Plauborg

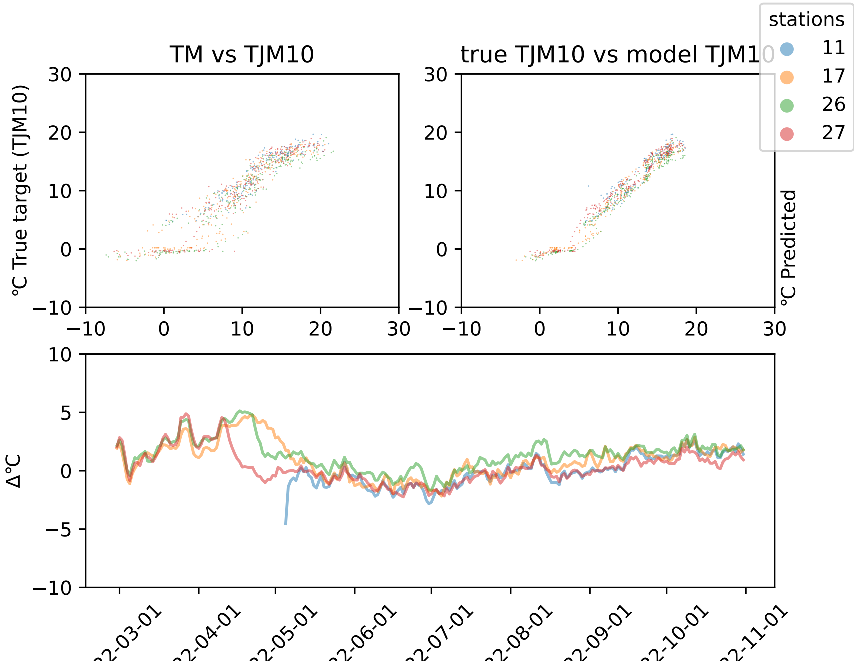
review mot tidlige i diskusjonen

The global measure for the linear regression has an average error of $2.3°C \pm 4.23°C$ while the global messure of the Plauborg daily model has an average error of $0.6°C \pm 1.96°C$. Further more Plauborg has an hight R^2 value indicating that it follows the temperature changes in the soil better than just scaling the air temperature by a scaling factor.

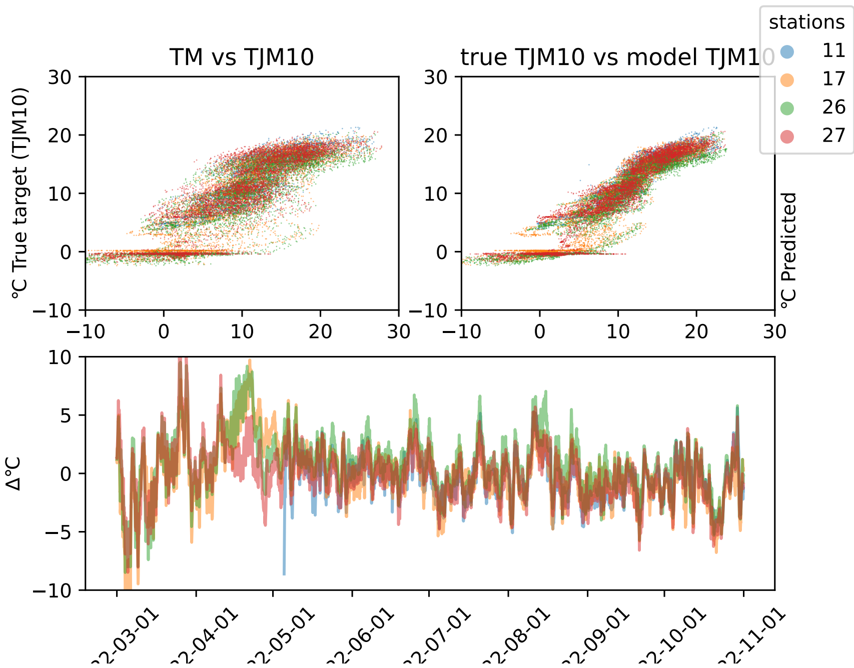
The linear model shows subpar predictive capabilities compared to Plauborg's model who uses the same technics but interoperate time dependence.

5.5 Modification of Plauborg

The Plauborg model trained in Norway was found to only need 3 days (t_0, t_{-1}, t_{-2}) compared to [1] that needed 4 days $(t_0, t_{-1}, t_{-2}, t_{-3})$. However for the Fourier terms both models (Danish model and the Norwegian model) required 2 sine and cosine terms. For the 20cm target the models diverge in the sense of quantity of terms. It was found that the 20cm model needs 14 sine terms and 2 cosine terms, however only needs 2 days.



(a) The daily model of Plauborg model. The model uses daily avergae tempratures to predict soil temperatures.



(b) The hourly model of Plauborg model. The model uses hourly temperature data.

Figure 22: Comperasion of daily versus hourly predictions

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	2.529	1.926	0.597	-0.443	-1	0.794
region	Østfold	2.448	1.894	0.512	-0.444	-1	0.816
region	Vestfold	2.412	1.81	0.733	-0.436	-1	0.846
region	Trøndelag	2.822	2.176	0.781	-0.446	-1	0.547
region	Innlandet	2.382	1.805	0.312	-0.447	-1	0.847
local	52	2.514	1.964	-0.349	-0.439	-1	0.636
local	41	1.938	1.519	0.151	-0.445	-1	0.903
local	37	2.344	1.804	0.237	-0.443	-1	0.857
local	118	2.928	2.322	1.639	-0.436	-1	0.706
local	50	1.908	1.472	0.558	-0.435	-1	0.884
local	42	2.501	1.885	0.703	-0.442	-1	0.852
local	38	3.055	2.368	1.363	-0.443	-1	0.754
local	30	2.072	1.555	0.354	-0.445	-1	0.892
local	57	2.906	2.263	0.677	-0.444	-1	0.677
local	39	2.77	2.151	0.701	-0.45	-1	0.633
local	34	3.013	2.306	0.845	-0.45	-1	-0.193
local	15	2.589	1.991	0.903	-0.441	-1	0.562
local	27	2.277	1.724	0.163	-0.443	-1	0.872
local	26	2.532	1.918	0.821	-0.444	-1	0.843
local	17	2.649	1.979	0.065	-0.447	-1	0.828
local	11	2.146	1.666	0.038	-0.435	-1	0.823

Table 16: Hourly Plauborg model results.

The modification to Plauborg's model is minor, by replacing the ω with a larger coefficient it can be used with hourly data. As seen in figure 22b the variation is stronger than 22a however the overall performance is comparable as seen in table 21 and table 19.

With modification to the model to accept hourly data it still preforms approximately as well as the daily data version. With a average error of $0.597°C \pm 2.529°C$ for TJM10 and $0.528°C \pm 2.676°C$ for TJM20. It was found that the modified Plauborg model only needs 2 sine terms to make a good prediction and 12h of air temperature which would translate to half a day instead of 3 days.

5.5.1 RNN results compared to other studies

The BiLSTM is an improvement over LSTM and the modified BiLSTM with layers is a clear indication that added complexity to a deep learning model is the way to go for . This progression of improvements has been shown in other studies[23, 25, 26, 28, 29]. None of the deep learning models has been optimised, however in according to earlier studies that focused on these types of model the authors did find that adding layers to the models does improve the model performance.

5.6 Model comparison

There is a unmistakeably distinction between the linear regression and the other models. The basic linear regression model

5.7 Future work

The models chosen in this study is not a representative sample of current knowledge of soil temperature modelling, and this study did not aim for optimizing the models beyond what

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	2.074	1.621	0.608	-1.266	-2	0.861
region	Østfold	2.168	1.704	0.24	-1.268	-2	0.856
region	Vestfold	2.022	1.564	0.219	-1.265	-2	0.892
region	Trøndelag	1.957	1.528	1.235	-1.264	-2	0.782
region	Innlandet	2.165	1.71	0.714	-1.267	-2	0.873
local	52	2.418	1.837	-0.636	-1.265	-2	0.664
local	41	1.975	1.587	-0.293	-1.268	-2	0.9
local	37	2.206	1.755	0.373	-1.265	-2	0.873
local	118	2.165	1.697	1.137	-1.271	-2	0.839
local	50	1.395	1.105	-0.046	-1.261	-2	0.938
local	42	2.239	1.75	0.333	-1.257	-2	0.881
local	38	2.42	1.908	0.667	-1.263	-2	0.845
local	30	1.914	1.519	-0.046	-1.265	-2	0.908
local	57	1.978	1.547	1.108	-1.248	-2	0.85
local	39	1.896	1.455	1.193	-1.262	-2	0.828
local	34	2.143	1.687	1.535	-1.262	-2	0.397
local	15	1.806	1.428	1.114	-1.262	-2	0.787
local	27	2.063	1.627	0.396	-1.263	-2	0.895
local	26	2.43	1.937	1.251	-1.262	-2	0.855
local	17	2.26	1.78	0.921	-1.262	-2	0.875
local	11	1.879	1.504	0.339	-1.26	-2	0.864

Table 17: Daily Plauborg model results.

the original authors have already done with the exception for base models used for comparison purposes. A more comprehensive is needed of more complex models that utilises cutting edge technologies, techniques, and theory. One of which is logic based models, for instance ASPER[36] that tries to incorporate logical descriptions of the problem and limits the model for better or equal results based on fewer samples[37]. Another approach is to use the newest deep learning method of the attention mechanism[38] combined with recurrent neural networks to elevate the accuracy and speed of the model. As the author of the paper [23] has show great promise with that approach.

Furthermore, the models presented in this study are not optimised as far as they can as there are more parameters one can include in the model, and Hyper-Parameter one can fine tune to improve the predictive capability of the models such as

- The type of loss function
- the learning rate
- the optimizer
- the activation function used
- weight regulation
- adding a dropout rate

This are a small collection of techniques that can be utilised to furter optimize the models for better predictive performance.

There have been significant developments in model types, including Answer Set Programming-enhanced Entity-Relation (ASPER). ASPER combines logical statements[6] with deep learning models to achieve results comparable to or better than "non-logical" deep learning models, but with fewer samples [36]. A study demonstrated that the ASPER model can reduce the required number of samples/observations by a factor of 1/1000 [37] and studies that uses this knowledge based approach shows to improve the predictive ability of the model to predict soil temperatures[2, 39]. In an interview with the study researcher [40], it was found that while the model requires strict rules, it is possible to incorporate Bayesian statistics to enhance its generality for various applications. By relaxing the ruleset and acknowledging that the given rules may not be 100% accurate, the model can be adapted to other applications using approximation rule-sets. Its the belife of the author of current study that this model can be adapted to soil temperature prediction when incorporating Bayesian statistics.

Additionally, attention-awareness, a method developed by the Google cooperation [38] also used in ChatGPT and other modern AI technologies, has been employed to predict soil temperatures and soil moisture[23] and has shown promising results in predicting soil temperatures by dynamically putting emphases on some of the features, particular days, and combinations of these when predicting.

[6]Statements can be thought of as formulas, natural laws, or knowledge about the solution

6 Conclusion

Soil temperature significantly impacts agriculture, influencing pest prevention, conservation, yield prediction, and more. Despite its importance, widespread measurement remains challenging due to cost limitations and technical issues. Interpolating missing data using methods like global mean approximation is common but has drawbacks including requiring a previus measurements of the soil temperature. Incorporating exogenous features can improve soil temperature estimation. Advancements in prediction and measurement are crucial for sustainable agriculture and accurate climate models.

Method used in this study are

1. Fetch avalible data to use as training and testing set
2. Compile the data and treat it to be used in the modelig
3. Train all the models on the same traning data (2014 to 2020)
4. calculate and plot relavent statitics (2021 to 2022)
5. compile the results

The results of this study show that promising results can be achieved with few parameters, however further studies need to be done to see the effekt of adding more parameters or making the models more complex by adding more structure[7]. As for regression modeling; Adding time to a regression model does improve the model predictive power over a time independent model. This make a simple model to predict soil temperatures in areas with no soil temperature measurement.

There is a clear advantage to data-driven modeling to further the investigation into deep learning models as the models shows comparable results to analytical models, as is show in other studies[4, 23, 26]. However to improve the model performance even more more features is required to make the model more general and adaptible to other local enviroments outside of the nortic climate.

6.1 Limitations

This study faced a multitude of technical difficulties including,

- Trying to get the models to run properly
- Finding proper parameters to limit memory usage and time
- Insufficient computing power for some parameters
- Getting TensorFlow to work on Windows 10/11[8]

This study would be more comprehensive if these unforseen difficulties did not occure, and it is recommended that it would be done a new study going more in depth into other type of deep learning models suitable for time-series data.

[7]Sturcture as in more layers, augmentations of input and features from feature extraction.

[8]This has later in the study that was futile as the project had moved from being Windows compatible to being exclusively runnable on Linux unless choosing an outdated version of Tensorflow.

7 Acknowledgements

The progenitor of this study presently partakes in an interval of non-occupational repose, traditionally associated with restorative diversion from vocational pursuits as of immediate effect of thesis submission.

8 Bibliography

References

[1] Finn Plauborg, "Simple model for 10 cm soil temperature in different soils with short grass," *European Journal of Agronomy*, volume 17, number 3, pages 173–179, Oct. 2002, ISSN: 11610301. DOI: 10.1016/S1161-0301(02)00006-0. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1161030102000060 (visited on 03/17/2023).

[2] Meysam Alizamir, Ozgur Kisi, Ali Najah Ahmed, Cihan Mert, Chow Ming Fai, Sungwon Kim, Nam Won Kim, and Ahmed El-Shafie, "Advanced machine learning model for better prediction accuracy of soil temperature at different depths," *PLOS ONE*, volume 15, number 4, Lei Lin, Ed., page 25, Apr. 14, 2020, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0231055. [Online]. Available: https://dx.plos.org/10.1371/journal.pone.0231055 (visited on 09/29/2023).

[3] Ha Seon Sim, Dong Sub Kim, Min Gyu Ahn, Su Ran Ahn, and Sung Kyeom Kim, "Prediction of strawberry growth and fruit yield based on environmental and growth data in a greenhouse for soil cultivation with applied autonomous facilities," *Korean Journal of Horticultural Science and Technology*, volume 38, number 6, pages 840–849, Dec. 31, 2020, ISSN: 1226-8763, 2465-8588. DOI: 10.7235/HORT.20200076. [Online]. Available: https://www.hst-j.org/articles/doi/10.7235/HORT.20200076 (visited on 10/05/2023).

[4] Cong Li, Yaonan Zhang, and Xupeng Ren, "Modeling hourly soil temperature using deep BiLSTM neural network," *Algorithms*, volume 13, number 7, page 173, Jul. 17, 2020, ISSN: 1999-4893. DOI: 10.3390/a13070173. [Online]. Available: https://www.mdpi.com/1999-4893/13/7/173 (visited on 03/17/2023).

[5] Ornela Nanushi, Vasileios Sitokonstantinou, Ilias Tsoumas, and Charalampos Kontoes, *Pest presence prediction using interpretable machine learning*, May 16, 2022. arXiv: 2205.07723[cs]. [Online]. Available: http://arxiv.org/abs/2205.07723 (visited on 05/09/2024).

[6] Scott N. Johnson, Peter J. Gregory, James W. McNicol, Yasmina Oodally, Xiaoxian Zhang, and Philip J. Murray, "Effects of soil conditions and drought on egg hatching and larval survival of the clover root weevil (*Sitona lepidus*)," *Applied Soil Ecology*, volume 44, number 1, pages 75–79, Jan. 1, 2010, ISSN: 0929-1393. DOI: 10.1016/j.apsoil.2009.10.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0929139309001887 (visited on 05/09/2024).

[7] Joris C. Stuurop, Sjoerd E.A.T.M. Van Der Zee, and Helen Kristine French, "The influence of soil texture and environmental conditions on frozen soil infiltration: A numerical investigation," *Cold Regions Science and Technology*, volume 194, page 103 456, Feb. 2022, ISSN: 0165232X. DOI: 10.1016/j.coldregions.2021.103456. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0165232X21002378 (visited on 01/31/2024).

[8] Katri Rankinen, Tuomo Karvonen, and D. Butterfield, "A simple model for predicting soil temperature in snow-covered and seasonally frozen soil: Model description and testing," *Hydrology and Earth System Sciences*, volume 8, number 4, pages 706–716, Aug. 31, 2004, ISSN: 1607-7938. DOI: 10.5194/hess-8-706-2004. [Online]. Available: https://hess.copernicus.org/articles/8/706/2004/ (visited on 03/17/2023).

[9] Tuomo Karvonen, *A model for predicting the effect of drainage on soil moisture, soil temperature and crop yield*. Otaniemi, Finland: Helsinki University of Technology, Laboratory of Hydrology and Water Resources Engineering, 1988, xvi, 215, Open Library ID: OL15197205M.

[10] Jean Baptiste Joseph Fourier and Alexander Freeman, *The analytical theory of heat*. New York: Cambridge University Press, 2009, OCLC: 880311398, ISBN: 978-1-108-00178-6.

[11] Berit Nordskog, Anne Kjersti Bakken, and Halvard Hole, *Jordtemperatur og spredning av husdyrgjødsel - Utredning av om jordtemperatur kan brukes som skranke for spredning av husdyrgjødsel på eng i Farsund kommune*. NIBIO, Jun. 13, 2018, Accepted: 2018-06-14T10:40:43Z ISSN: 2464-1162 Journal Abbreviation: Jordtemperatur og spredning av husdyrgjødsel - Utredning av om jordtemperatur kan brukes som skranke for spredning av husdyrgjødsel på eng i Farsund kommune Publication Title: 18, ISBN: 978-82-17-02114-8. [Online]. Available: `https://nibio.brage.unit.no/nibio-xmlui/handle/11250/2501555` (visited on 05/08/2024).

[12] "107 - temperature probe." (), [Online]. Available: `https://www.campbellsci.com/107` (visited on 05/12/2024).

[13] Kamstrup. "Dokumentasjon for to-tråds Pt500 temperatursensorer." (), [Online]. Available: `https://www.kamstrup.com/no-no/varmeloesninger/varmemaalere/temperature-sensors/2-wire-pt500-temperature-sensors/documents` (visited on 05/12/2024).

[14] Jones SB. "3.5 soil temperature – ClimEx handbook." (), [Online]. Available: `https://climexhandbook.w.uib.no/2019/11/05/soil-temperature-thermal-regime/` (visited on 05/12/2024).

[15] Van Wijk, WR and De Vries, DA, *Periodic temperature variations in a homogeneous soil*, in *Physics of plant environment*, 1963, pages 103–143. [Online]. Available: `https://scholar.google.com/scholar_lookup?title=Periodic%20temperature%20variations%20in%20a%20homogeneous%20soil&author=W.R.%20Van%C2%A0Wijk&publication_year=1963&pages=103-143` (visited on 05/07/2024).

[16] J. Roodenburg, "Estimating 10-cm soil temperatures under grass," *Agricultural and Forest Meteorology*, volume 34, number 1, pages 41–52, Feb. 1, 1985, ISSN: 0168-1923. DOI: `10.1016/0168-1923(85)90053-X`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/016819238590053X` (visited on 05/09/2024).

[17] Raviranjan Kumar Singh and Ram Vinoy Sharma, "Numerical analysis for ground temperature variation," *Geothermal Energy*, volume 5, number 1, page 22, Nov. 1, 2017, ISSN: 2195-9706. DOI: `10.1186/s40517-017-0082-z`. [Online]. Available: `https://doi.org/10.1186/s40517-017-0082-z` (visited on 05/07/2024).

[18] Peter John Cleall, José Javier Muñoz-Criollo, and Stephen William Rees, "Analytical solutions for ground temperature profiles and stored energy using meteorological data," *Transport in Porous Media*, volume 106, number 1, pages 181–199, Jan. 1, 2015, ISSN: 1573-1634. DOI: `10.1007/s11242-014-0395-3`. [Online]. Available: `https://doi.org/10.1007/s11242-014-0395-3` (visited on 05/07/2024).

[19] Mark H. Van Benthem and Michael R. Keenan, "Fast algorithm for the solution of large-scale non-negativity-constrained least squares problems," *Journal of Chemometrics*, volume 18, number 10, pages 441–450, 2004, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cem.889, ISSN: 1099-128X. DOI: `10.1002/cem.889`. [Online]. Available: `https://onlinelibrary.wiley.com/doi/abs/10.1002/cem.889` (visited on 05/10/2024).

[20] Guillaume Chevalier, *English: Schematic of the long-short term memory cell, a component of recurrent neural networks*, May 16, 2018. [Online]. Available: `https://commons.wikimedia.org/wiki/File:LSTM_Cell.svg` (visited on 12/10/2023).

[21] Ihsan Uluocak and Mehmet Bilgili, "Daily air temperature forecasting using LSTM-CNN and GRU-CNN models," *Acta Geophysica*, volume 72, number 3, pages 2107–2126, Jun. 1, 2024, ISSN: 1895-7455. DOI: `10.1007/s11600-023-01241-y`. [Online]. Available: `https://doi.org/10.1007/s11600-023-01241-y` (visited on 05/12/2024).

[22] Furkan Elmaz, Reinout Eyckerman, Wim Casteels, Steven Latré, and Peter Hellinckx, "CNN-LSTM architecture for predictive indoor temperature modeling," *Building and Environment*, volume 206, page 108 327, Dec. 1, 2021, ISSN: 0360-1323. DOI: `10.1016/j.buildenv.2021.108327`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0360132321007241` (visited on 04/11/2024).

[23] Qingliang Li, Yuheng Zhu, Wei Shangguan, Xuezhi Wang, Lu Li, and Fanhua Yu, "An attention-aware LSTM model for soil moisture and soil temperature prediction," *Geoderma*, volume 409, page 115 651, Mar. 2022, ISSN: 00167061. DOI: `10.1016/j.geoderma.2021.115651`. [Online]. Available: `https://linkinghub.elsevier.com/retrieve/pii/S001670612100731X` (visited on 10/05/2023).

[24] Sepp Hochreiter and Jürgen Schmidhuber, "Long short-term memory," *Neural Computation*, volume 9, number 8, pages 1735–1780, Nov. 1, 1997, ISSN: 0899-7667, 1530-888X. DOI: `10.1162/neco.1997.9.8.1735`. [Online]. Available: `https://direct.mit.edu/neco/article/9/8/1735-1780/6109` (visited on 10/18/2023).

[25] Hatice Citakoglu, "Comparison of artificial intelligence techniques for prediction of soil temperatures in turkey," *Theoretical and Applied Climatology*, volume 130, number 1, pages 545–556, Oct. 2017, ISSN: 0177-798X, 1434-4483. DOI: `10.1007/s00704-016-1914-7`. [Online]. Available: `http://link.springer.com/10.1007/s00704-016-1914-7` (visited on 09/28/2023).

[26] Yu Feng, Ningbo Cui, Weiping Hao, Lili Gao, and Daozhi Gong, "Estimation of soil temperature from meteorological data using different machine learning models," *Geoderma*, volume 338, pages 67–77, Mar. 2019, ISSN: 00167061. DOI: `10.1016/j.geoderma.2018.11.044`. [Online]. Available: `https://linkinghub.elsevier.com/retrieve/pii/S0016706118311200` (visited on 03/17/2023).

[27] Janani Kandasamy, Yuan Xue, Paul Houser, and Viviana Maggioni, "Performance of different crop models in simulating soil temperature," *Sensors*, volume 23, number 6, page 2891, Mar. 7, 2023, ISSN: 1424-8220. DOI: `10.3390/s23062891`. [Online]. Available: `https://www.mdpi.com/1424-8220/23/6/2891` (visited on 11/07/2023).

[28] Qingliang Li, Huibowen Hao, Yang Zhao, Qingtian Geng, Guangjie Liu, Yu Zhang, and Fanhua Yu, "GANs-LSTM model for soil temperature estimation from meteorological: A new approach," *IEEE Access*, volume 8, pages 59 427–59 443, 2020, ISSN: 2169-3536. DOI: `10.1109/ACCESS.2020.2982996`. [Online]. Available: `https://ieeexplore.ieee.org/document/9045947/` (visited on 12/05/2023).

[29] Xuezhi Wang, Wenhui Li, Qingliang Li, and Xiaoning Li, "Modeling soil temperature for different days using novel quadruplet loss-guided LSTM," *Computational Intelligence and Neuroscience*, volume 2022, Xin Ning, Ed., pages 1–17, Feb. 17, 2022, ISSN: 1687-5273, 1687-5265. DOI: `10.1155/2022/9016823`. [Online]. Available: `https://www.hindawi.com/journals/cin/2022/9016823/` (visited on 12/05/2023).

[30] Saeid Mehdizadeh, Farshad Ahmadi, and Ali Kozekalani Sales, "Modelling daily soil temperature at different depths via the classical and hybrid models," *Meteorological Applications*, volume 27, number 4, e1941, 2020, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/met.1941, ISSN: 1469-8080. DOI: 10.1002/met.1941. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/met.1941 (visited on 05/08/2024).

[31] Kyunghyun Cho, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio, *Learning phrase representations using RNN encoder-decoder for statistical machine translation*, Sep. 2, 2014. arXiv: 1406.1078[cs,stat]. [Online]. Available: http://arxiv.org/abs/1406.1078 (visited on 05/07/2024).

[32] Xinxing Li, Lu Zhang, Xiangyu Wang, and Buwen Liang, "Forecasting greenhouse air and soil temperatures: A multi-step time series approach employing attention-based LSTM network," *Computers and Electronics in Agriculture*, volume 217, page 108 602, Feb. 1, 2024, ISSN: 0168-1699. DOI: 10.1016/j.compag.2023.108602. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168169923009900 (visited on 05/12/2024).

[33] Jeblad, *English: Gradient recurrent unit, type 2. variation on example in recurrent neural network (RNN) – part 5: Custom cells, to encompass the type difference from gate-variants of gated recurrent unit (GRU) neural networks.* Feb. 8, 2018. [Online]. Available: https://commons.wikimedia.org/wiki/File:Gated_Recurrent_Unit,_type_2.svg (visited on 05/12/2024).

[34] Joris C. Stuurop, Sjoerd E. A. T. M Van Der Zee, Clifford I. Voss, and Helen Kristine French, "Simulating water and heat transport with freezing and cryosuction in unsaturated soil: Comparing an empirical, semi-empirical and physically-based approach," *Advances in Water Resources*, volume 149, page 103 846, Mar. 2021, ISSN: 03091708. DOI: 10.1016/j.advwatres.2021.103846. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0309170821000014 (visited on 01/31/2024).

[35] T. R. H. Holmes, M. Owe, R. A. M. De Jeu, and H. Kooi, "Estimating the soil temperature profile from a single depth observation: A simple empirical heatflow solution," *Water Resources Research*, volume 44, number 2, 2007WR005994, Feb. 2008, ISSN: 0043-1397, 1944-7973. DOI: 10.1029/2007WR005994. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2007WR005994 (visited on 04/29/2024).

[36] Trung Hoang Le, Huiping Cao, and Tran Cao Son, *ASPER: Answer set programming enhanced neural network models for joint entity-relation extraction*, version: 1, May 24, 2023. arXiv: 2305.15374[cs]. [Online]. Available: http://arxiv.org/abs/2305.15374 (visited on 03/12/2024).

[37] Fadi Al Machot, *Bridging logic and learning: A neural-symbolic approach for enhanced reasoning in neural models (ASPER)*, Dec. 18, 2023. arXiv: 2312.11651[cs]. [Online]. Available: http://arxiv.org/abs/2312.11651 (visited on 03/12/2024).

[38] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin, *Attention is all you need*, 2017. arXiv: 1706.03762[cs]. [Online]. Available: http://arxiv.org/abs/1706.03762 (visited on 04/16/2024).

[39] Olufemi P. Abimbola, George E. Meyer, Aaron R. Mittelstet, Daran R. Rudnick, and Trenton E. Franz, "Knowledge-guided machine learning for improving daily soil temperature prediction across the united states," *Vadose Zone Journal*, volume 20, number 5, e20151, 2021, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/vzj2.20151, ISSN: 1539-1663. DOI: 10.1002/vzj2.20151. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/vzj2.20151 (visited on 05/08/2024).

[40] *A discussion of ASPER and possible further developments.* In collaboration with Fadi Al Machot, Mar. 7, 2024.

Glossary

D | H | L | R

D

DataFrame
: A table of values. The name is from the python library Pandas used in this study.. 10

H

Hashmap
: A list of items where their unique placmnt in the list is detemend by their unique refrence key using a function that maps the key to a placement in the list.. 10

L

Long Short Term-Memory
: A Recurent Neural Network with a memory cell to distribute information along the other RNN cells.. 4

R

Recurent Neural Network
: A Neural network that passes information between cells in the same layers.. 2, 4, 5, E

Acronyms

Symbols | F | G | K | L | M | N | R

Symbols

R^2
: Explained Variance. 15

$\log(\kappa)$
: Log Condition number. 15

F

FDM
: Finite Difference Method. 3

G

GRU
: Gated Recurent Unit. 2, 5, 14, 23, III

K

Kilden

Norwegian Institute of Bioeconomy Research Kilden. 1, 6, 9

L

LMT

Norwegian Institute of Bioeconomy Research LandbruksMeteorologisk service. 6, 10, 11

LSTM

Long Short Term-Memory. 4, 5, 14, III

M

MAE

Mean Absolut Error. 15

MET

The Norwegian Meteorological Institute. 6, 10, 11

N

NMBU

The Norwegian University of Life Science. 17

R

RMSE

Root Mean Square Error. 15, 23

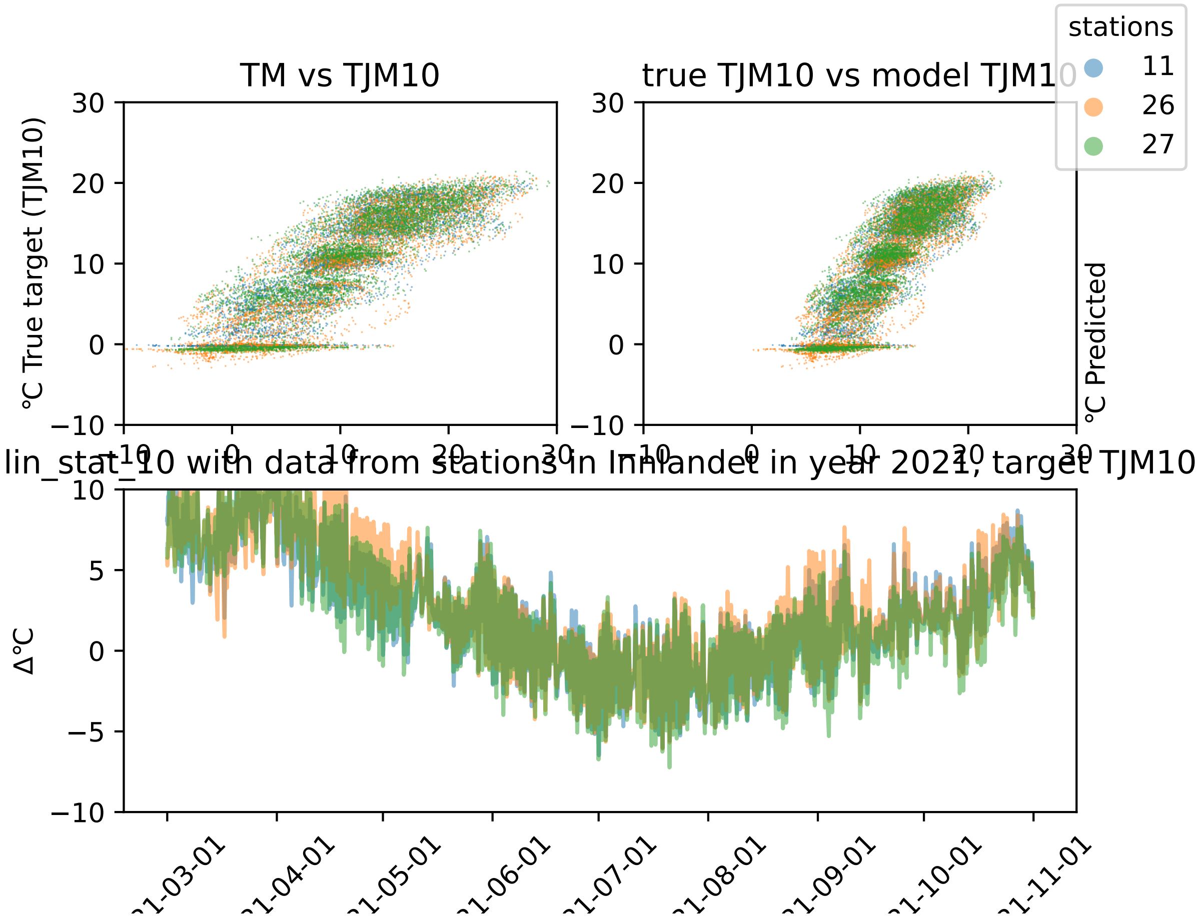


Figure 23: Difference plot for linear regression model in year 2021 and region Innlandet

A Plots

A.1 Difference plots of model performance per region

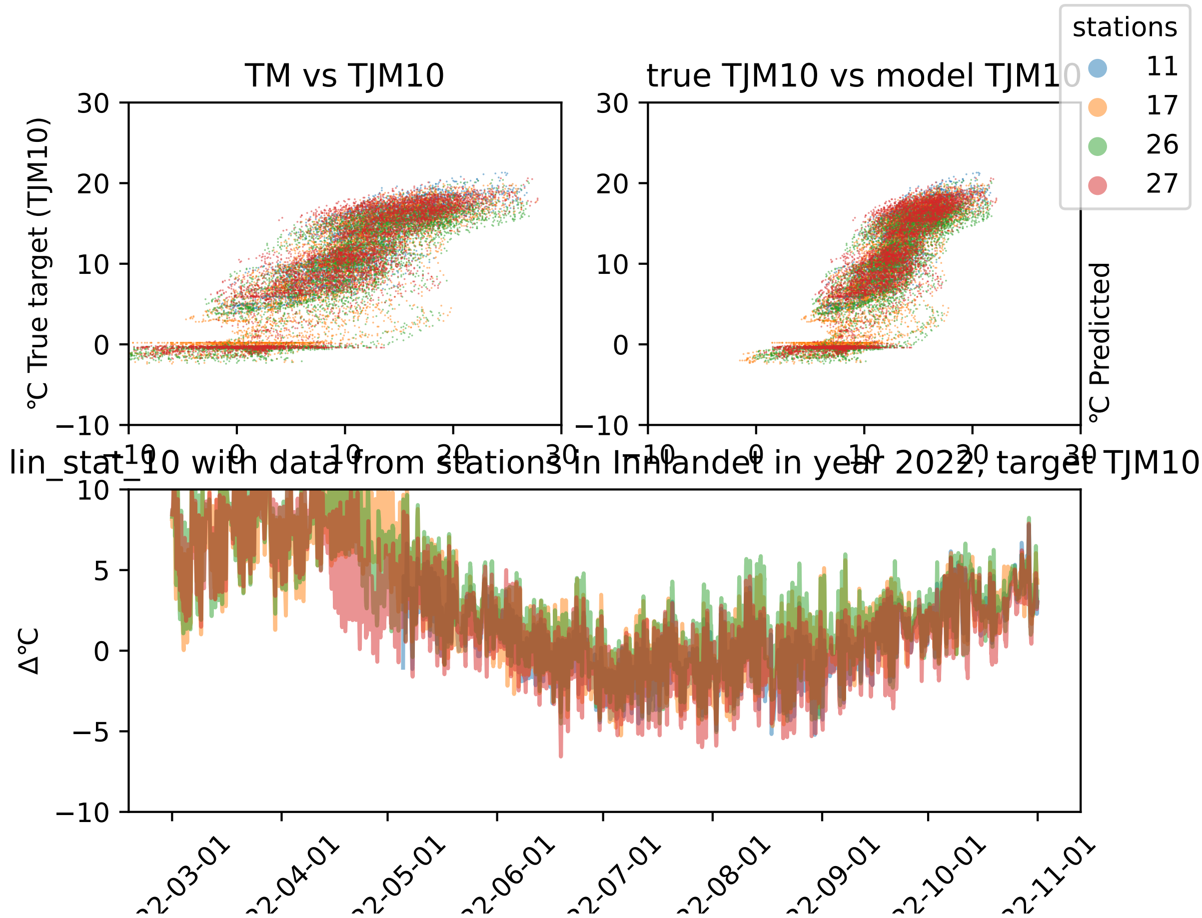


Figure 24: Difference plot for linear regression model in year 2022 and region Innlandet

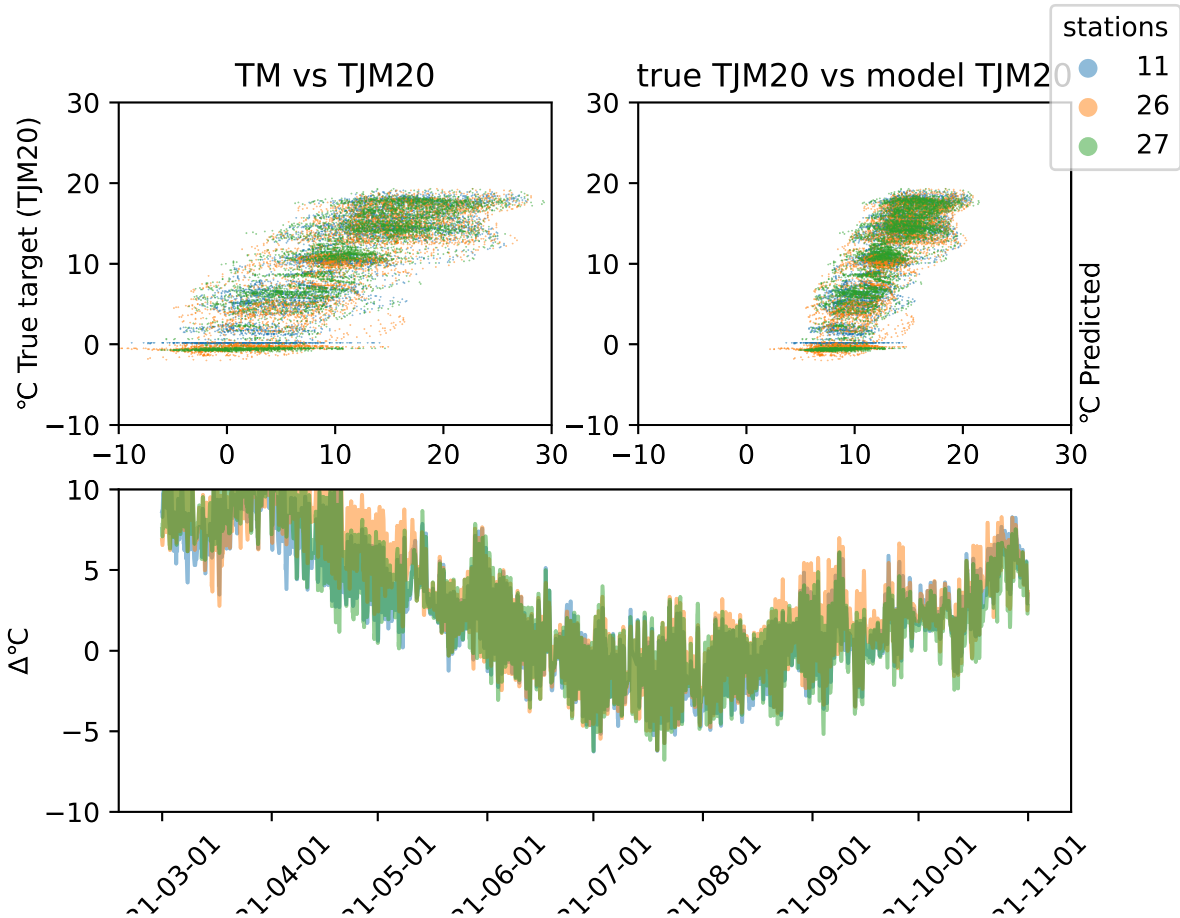


Figure 25: Difference plot for linear regression model in year 2021 and region Innlandet

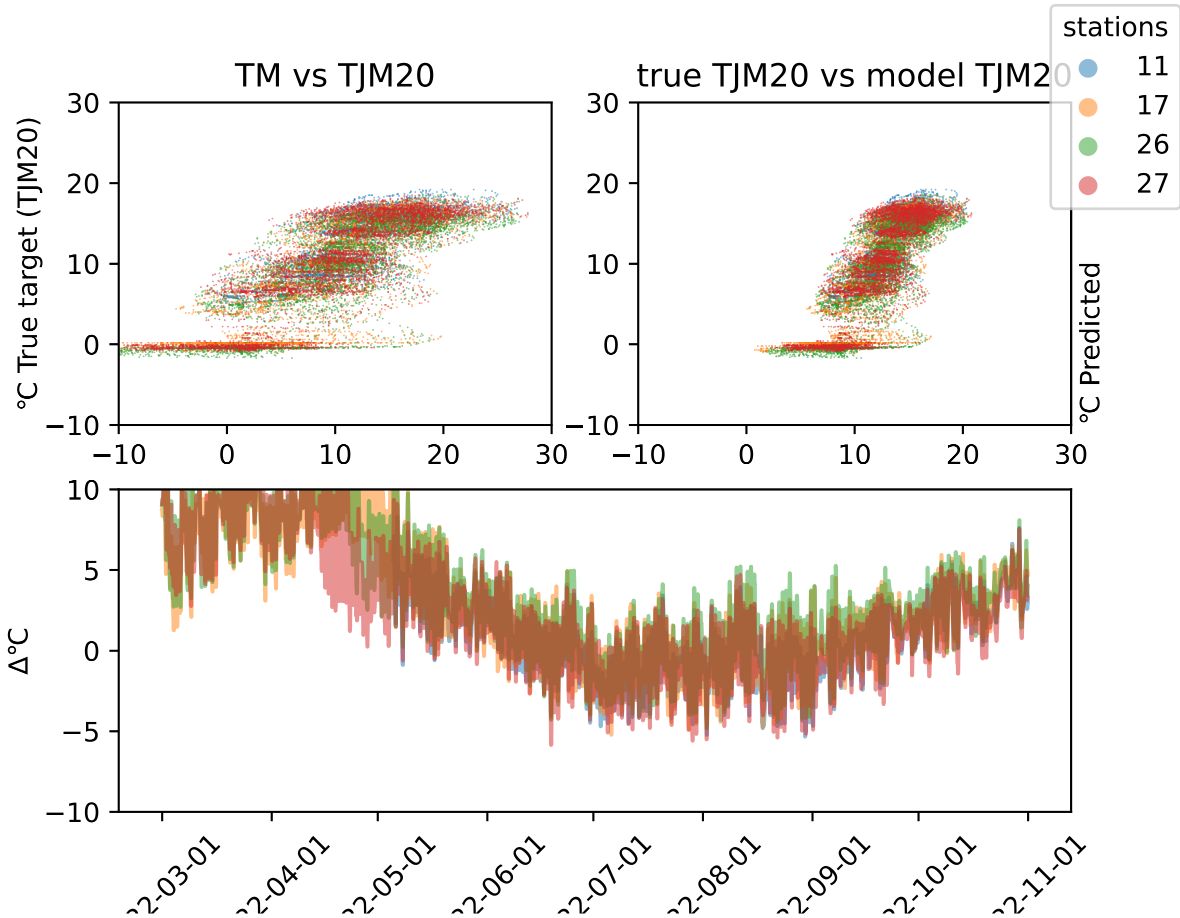


Figure 26: Difference plot for linear regression model in year 2022 and region Innlandet

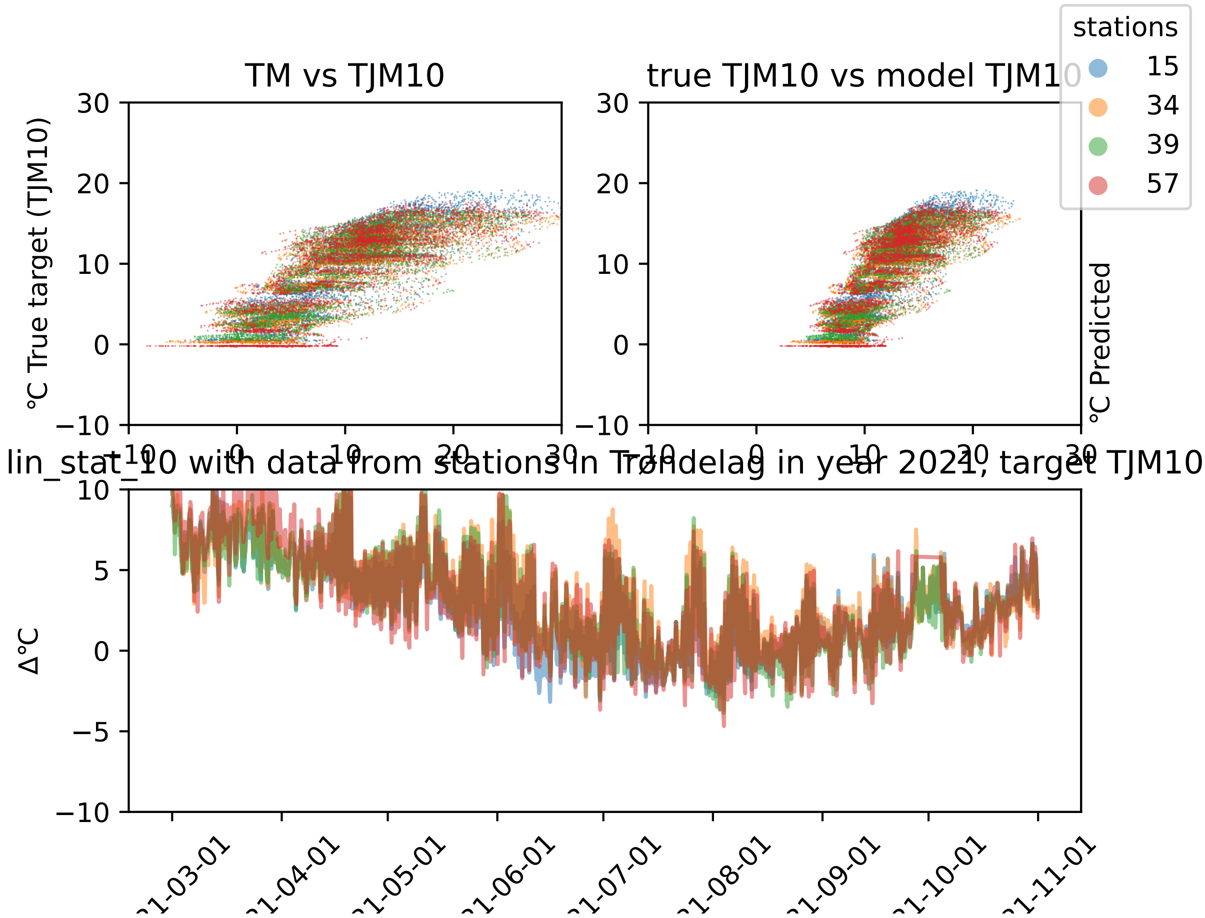


Figure 27: Difference plot for linear regression model in year 2021 and region Trøndelag

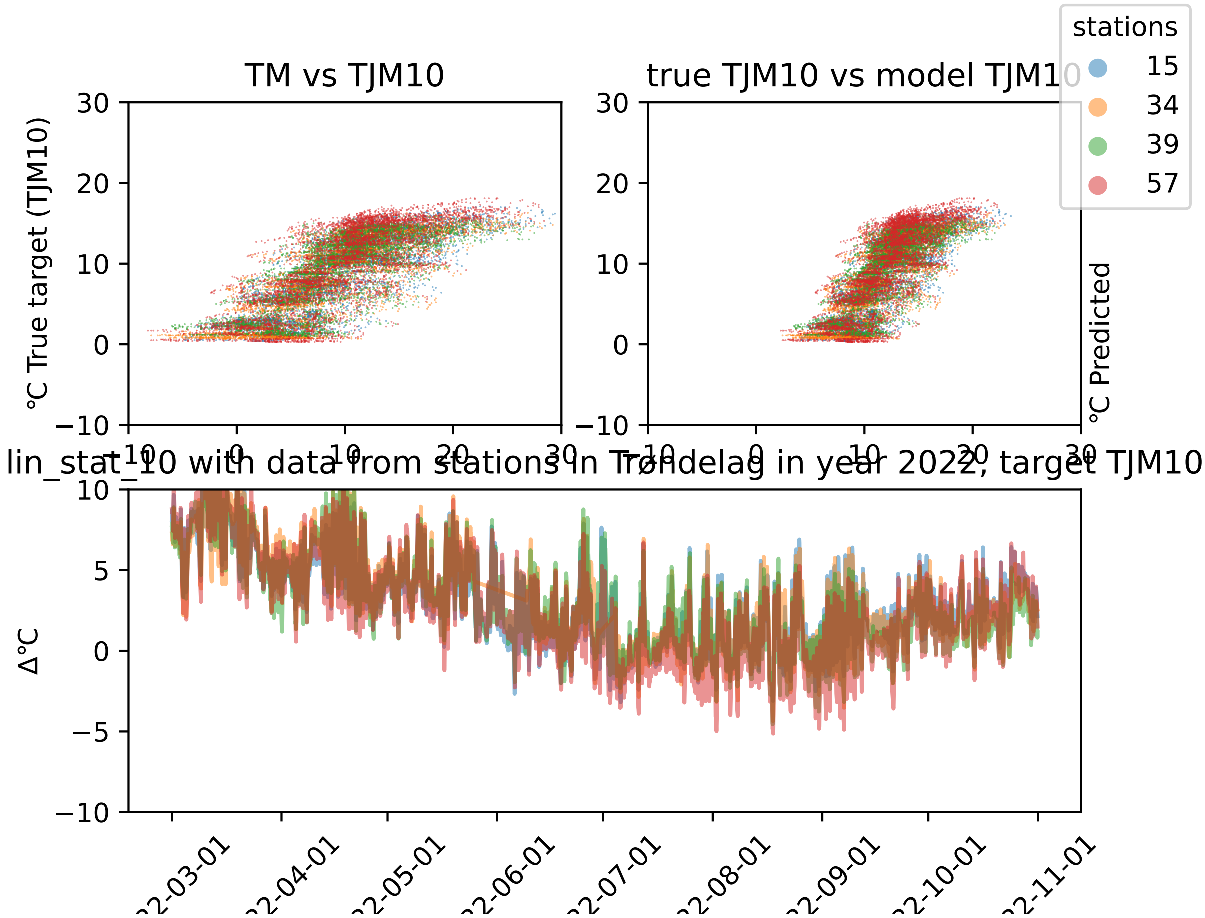


Figure 28: Difference plot for linear regression model in year 2022 and region Trøndelag

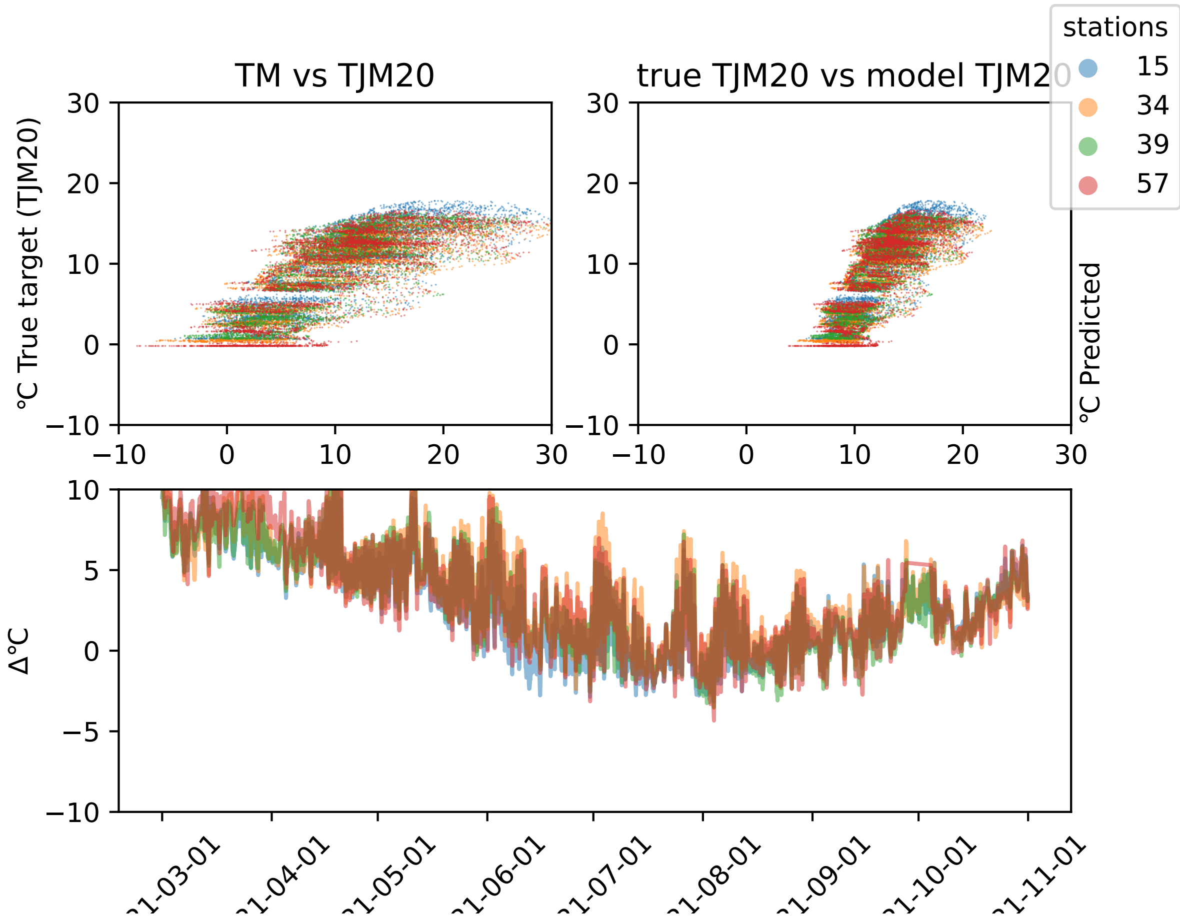


Figure 29: Difference plot for linear regression model in year 2021 and region Trøndelag

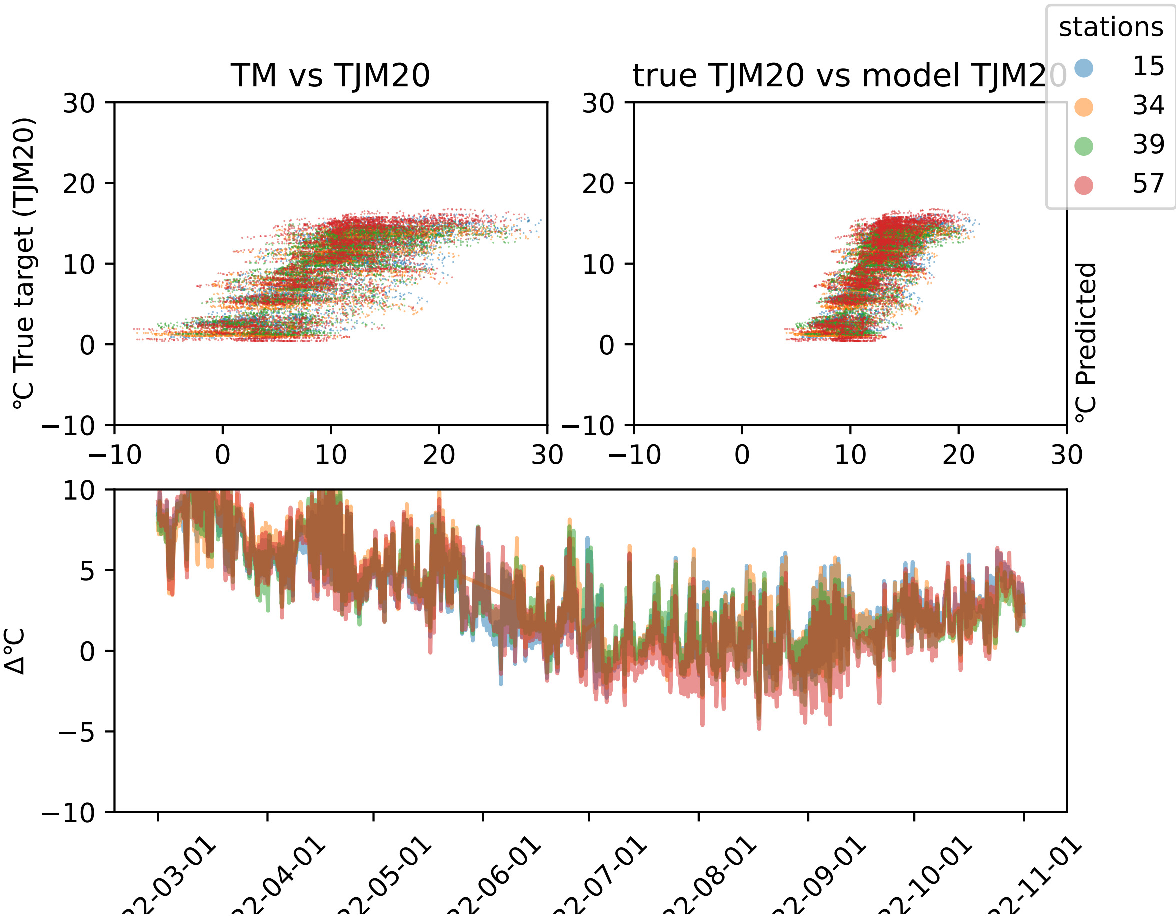


Figure 30: Difference plot for linear regression model in year 2022 and region Trøndelag

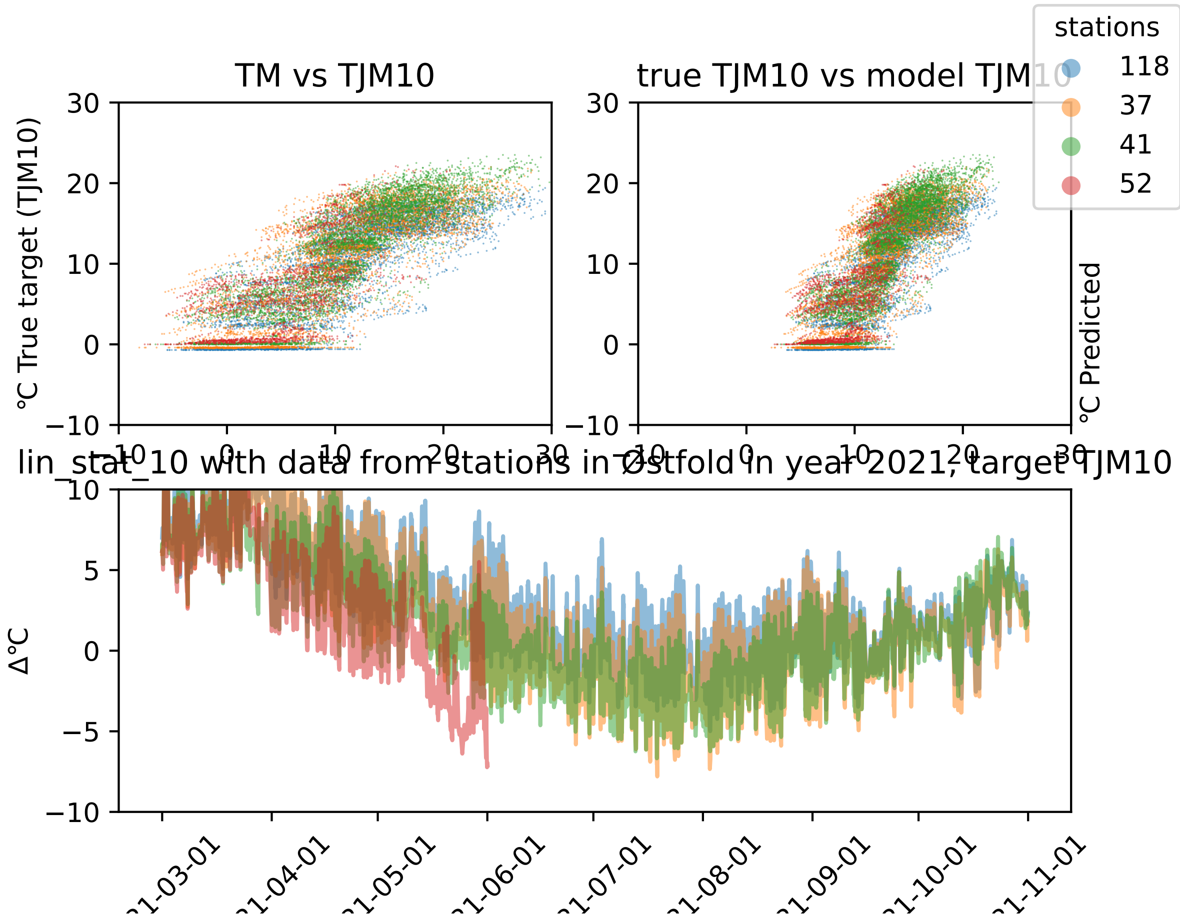


Figure 31: Difference plot for linear regression model in year 2021 and region Østfold

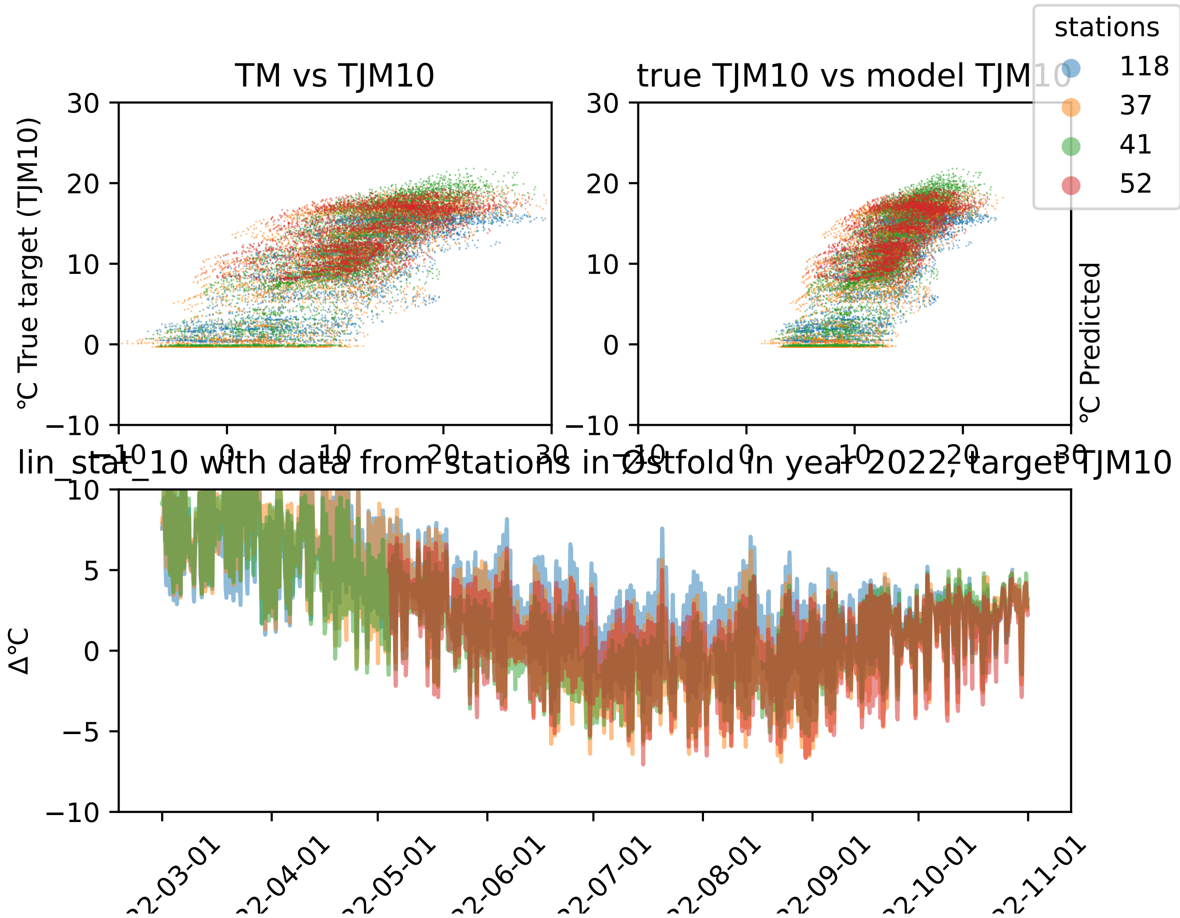


Figure 32: Difference plot for linear regression model in year 2022 and region Østfold

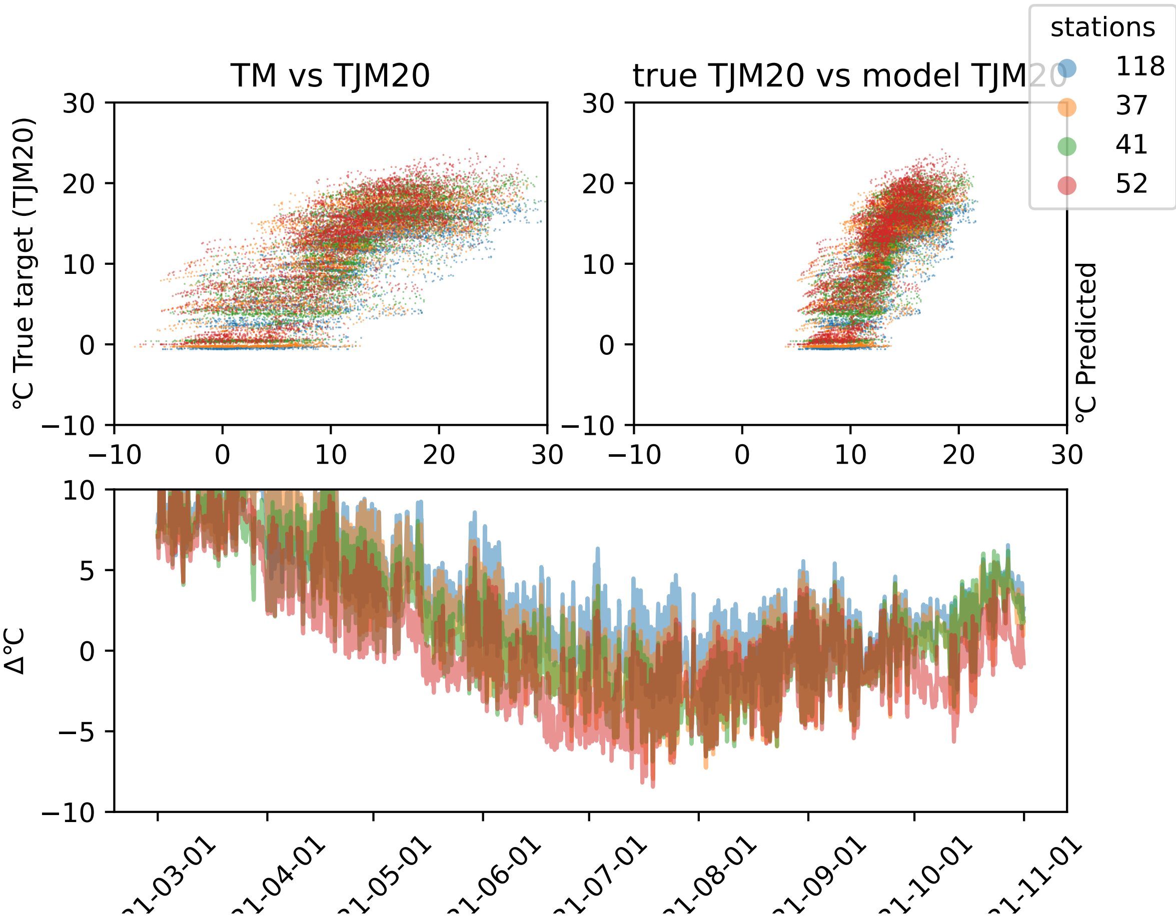


Figure 33: Difference plot for linear regression model in year 2021 and region Østfold

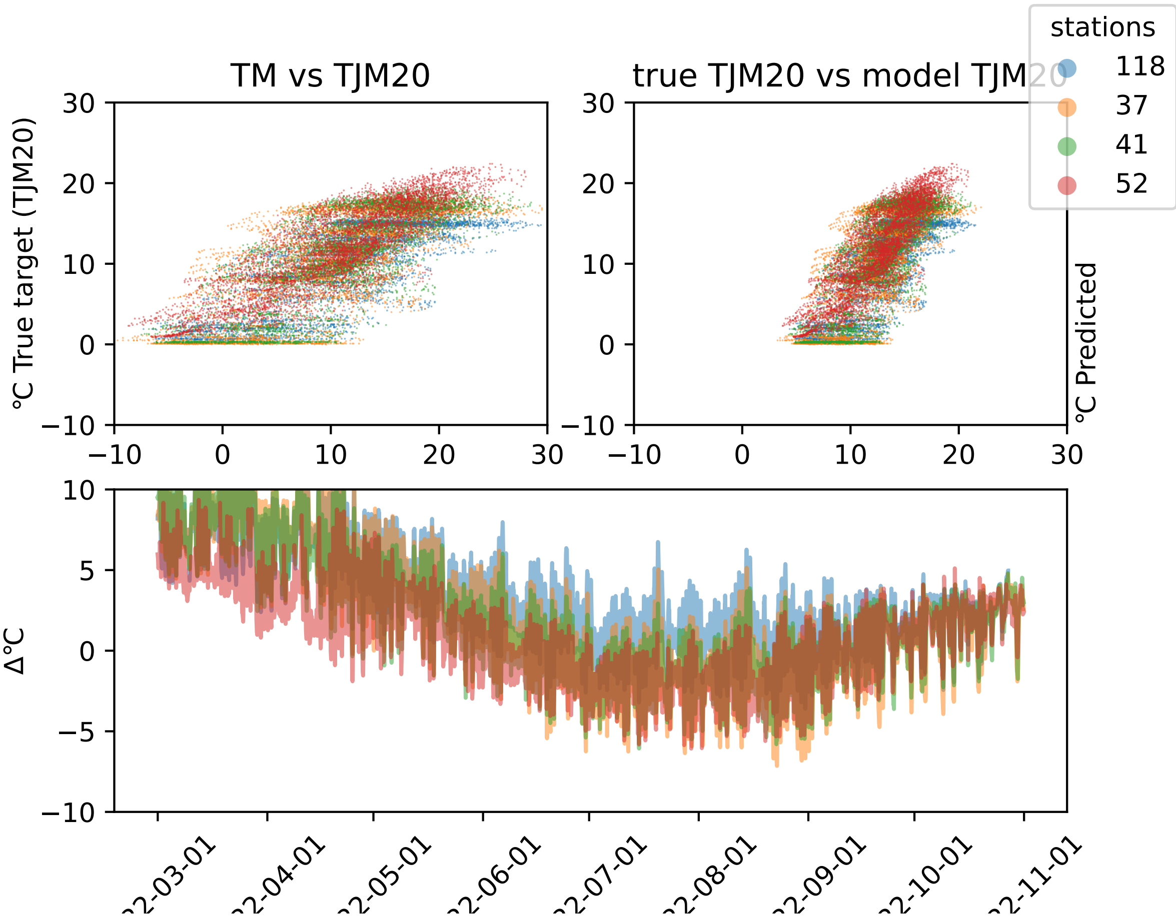


Figure 34: Difference plot for linear regression model in year 2022 and region Østfold

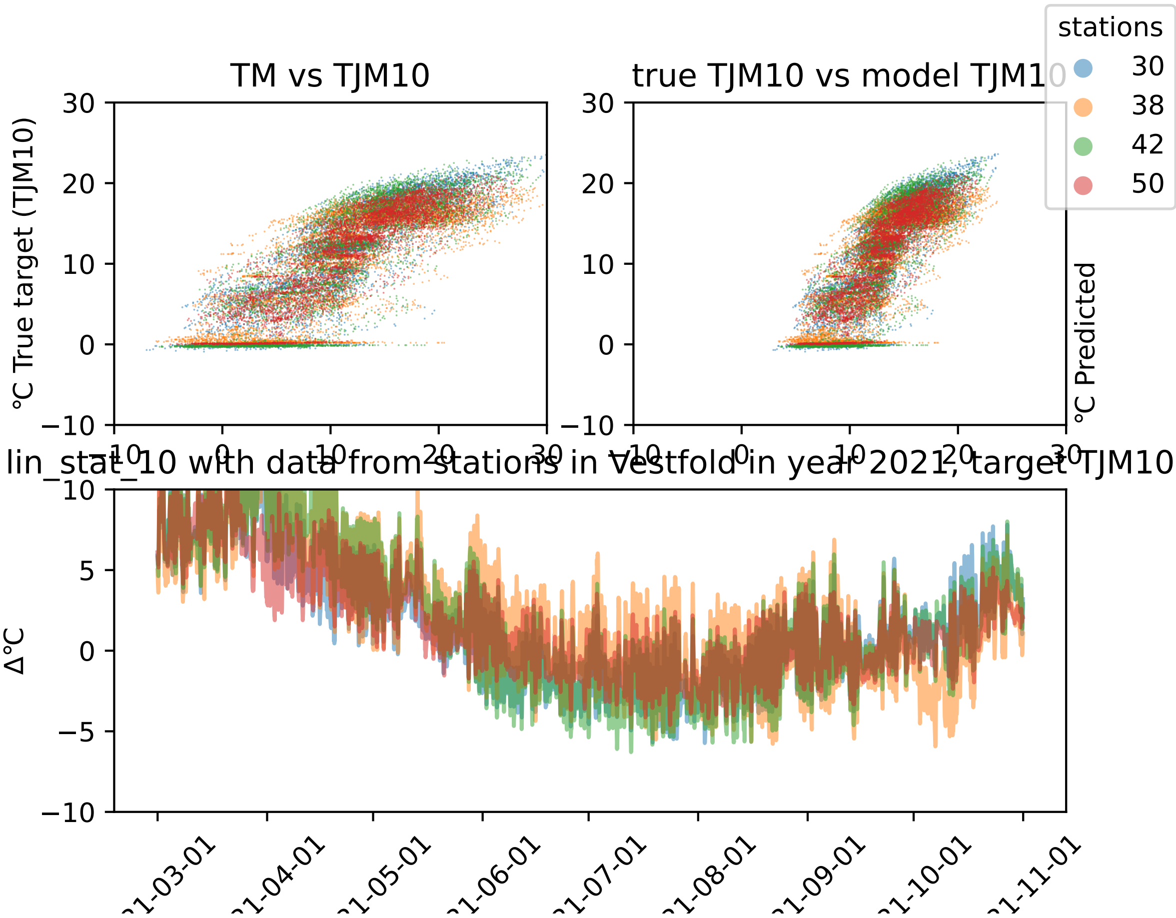


Figure 35: Difference plot for linear regression model in year 2021 and region Vestfold

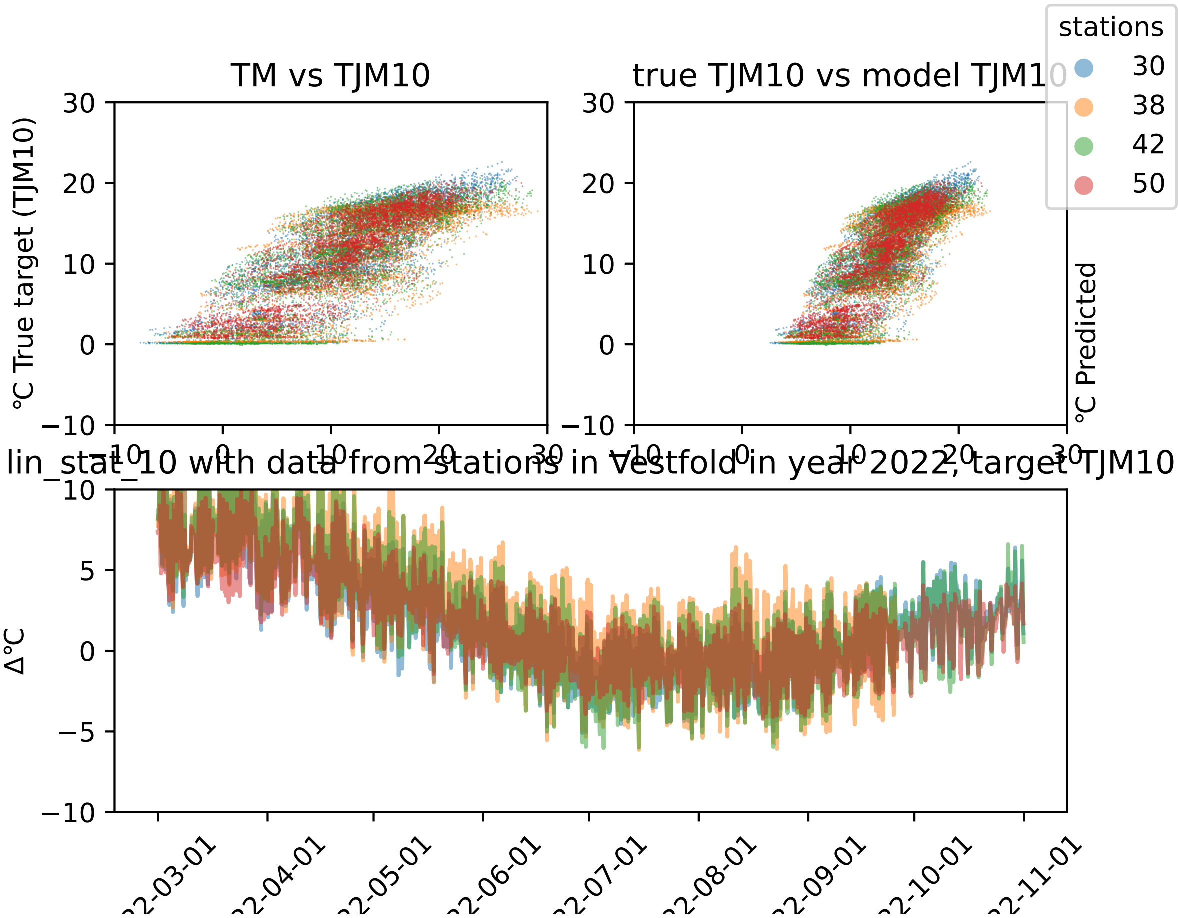


Figure 36: Difference plot for linear regression model in year 2022 and region Vestfold

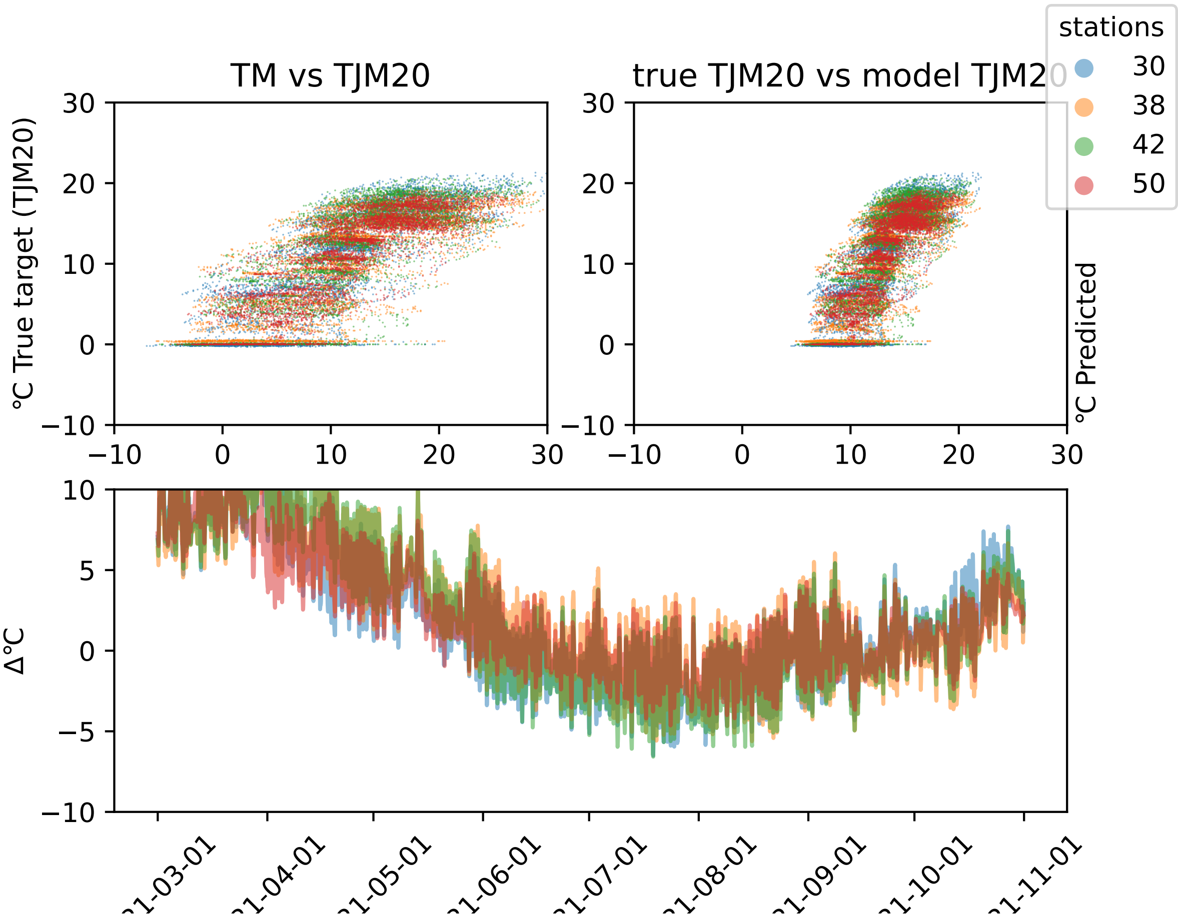


Figure 37: Difference plot for linear regression model in year 2021 and region Vestfold

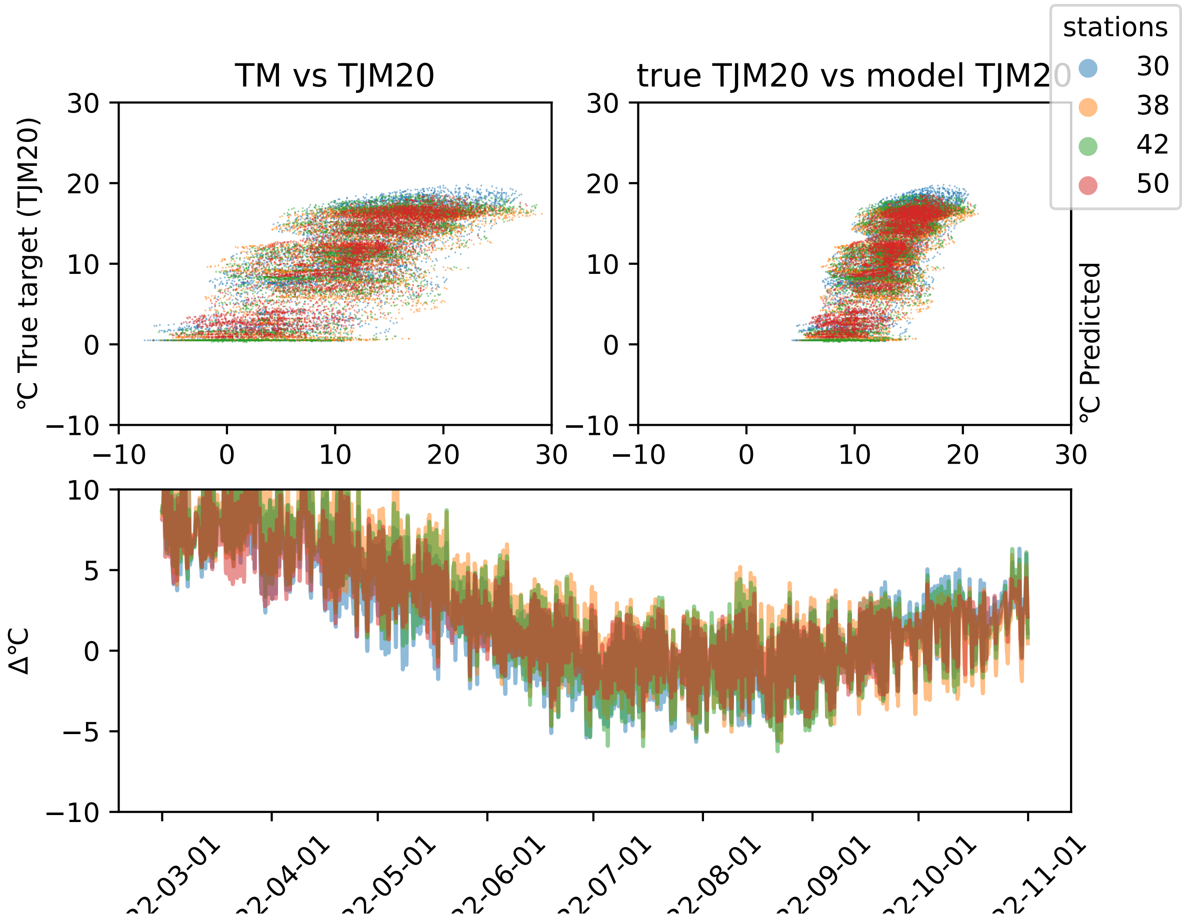


Figure 38: Difference plot for linear regression model in year 2022 and region Vestfold

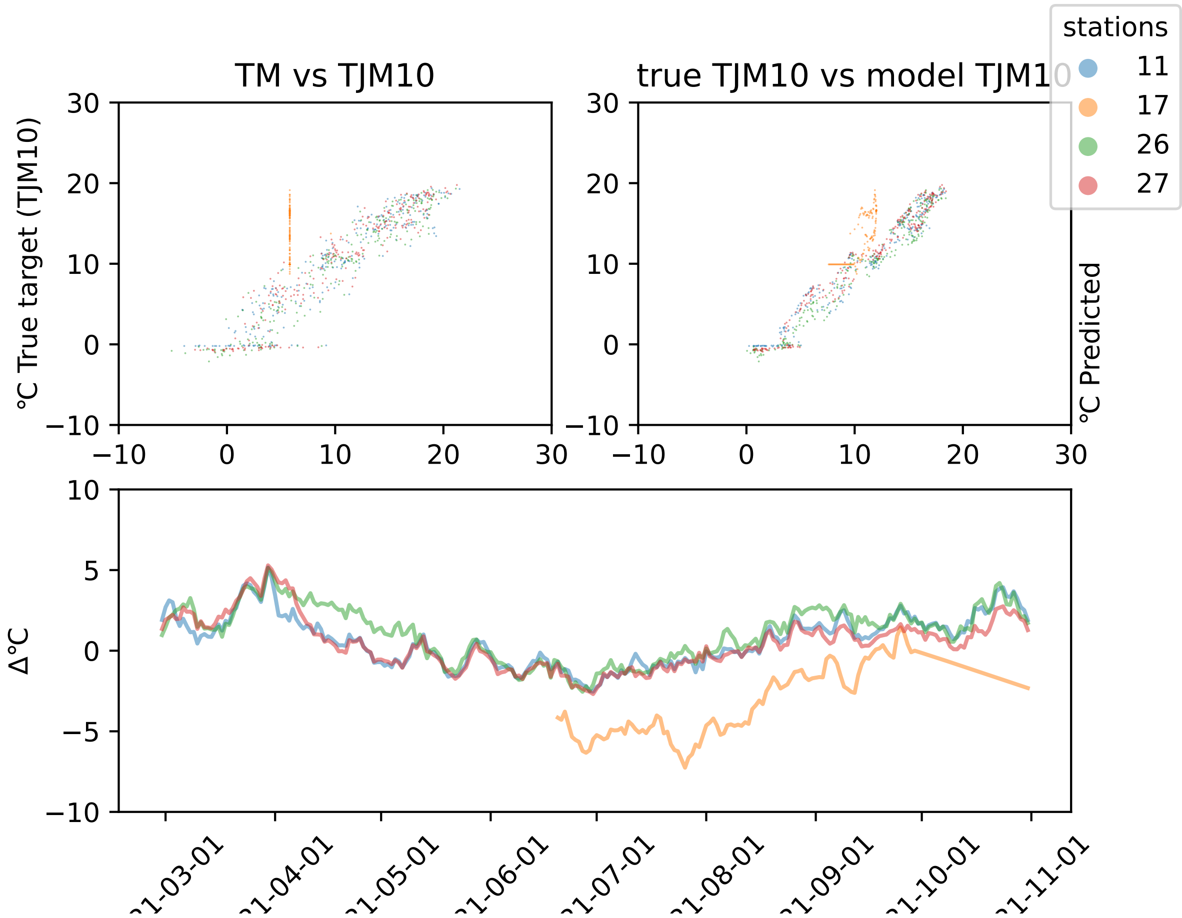


Figure 39: Difference plot for daily Plauborg model in year 2021 and region Innlandet

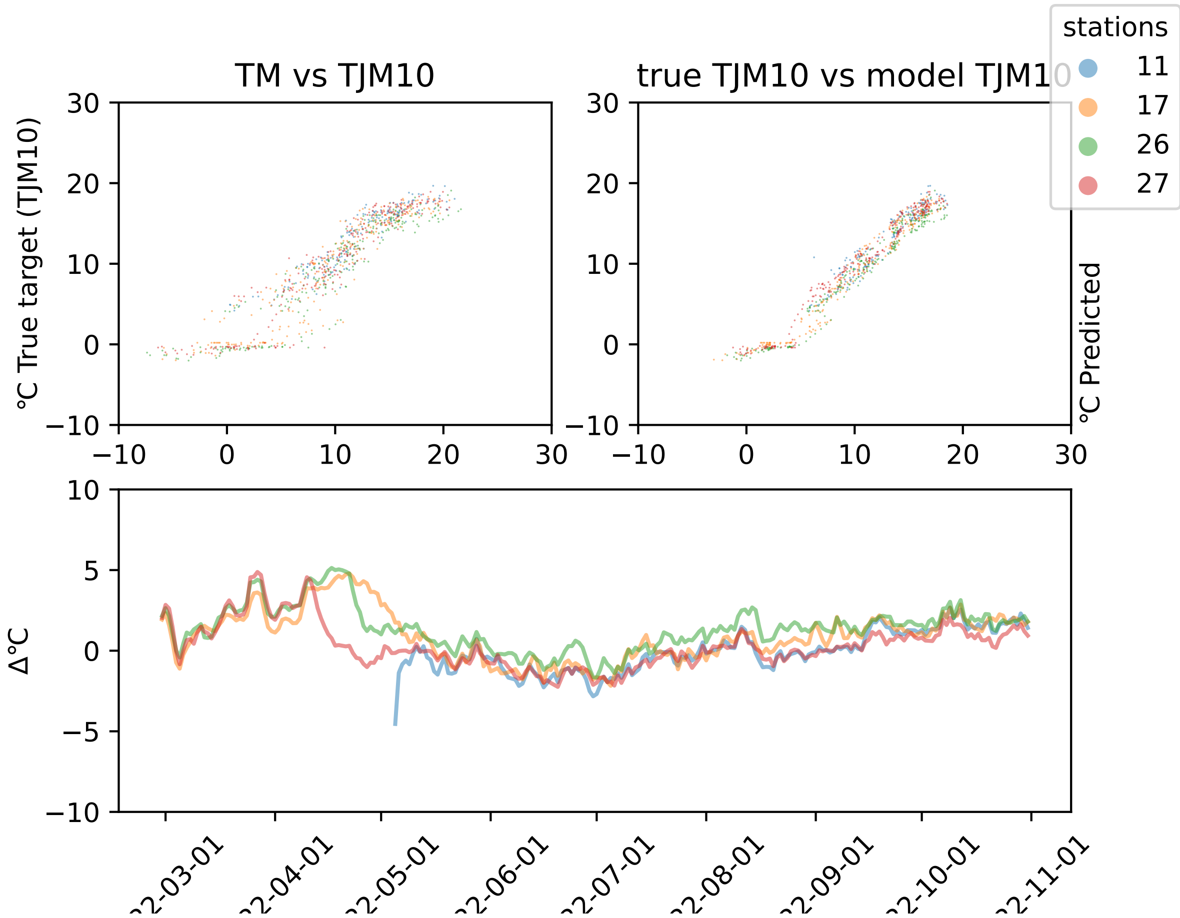


Figure 40: Difference plot for daily Plauborg model in year 2022 and region Innlandet

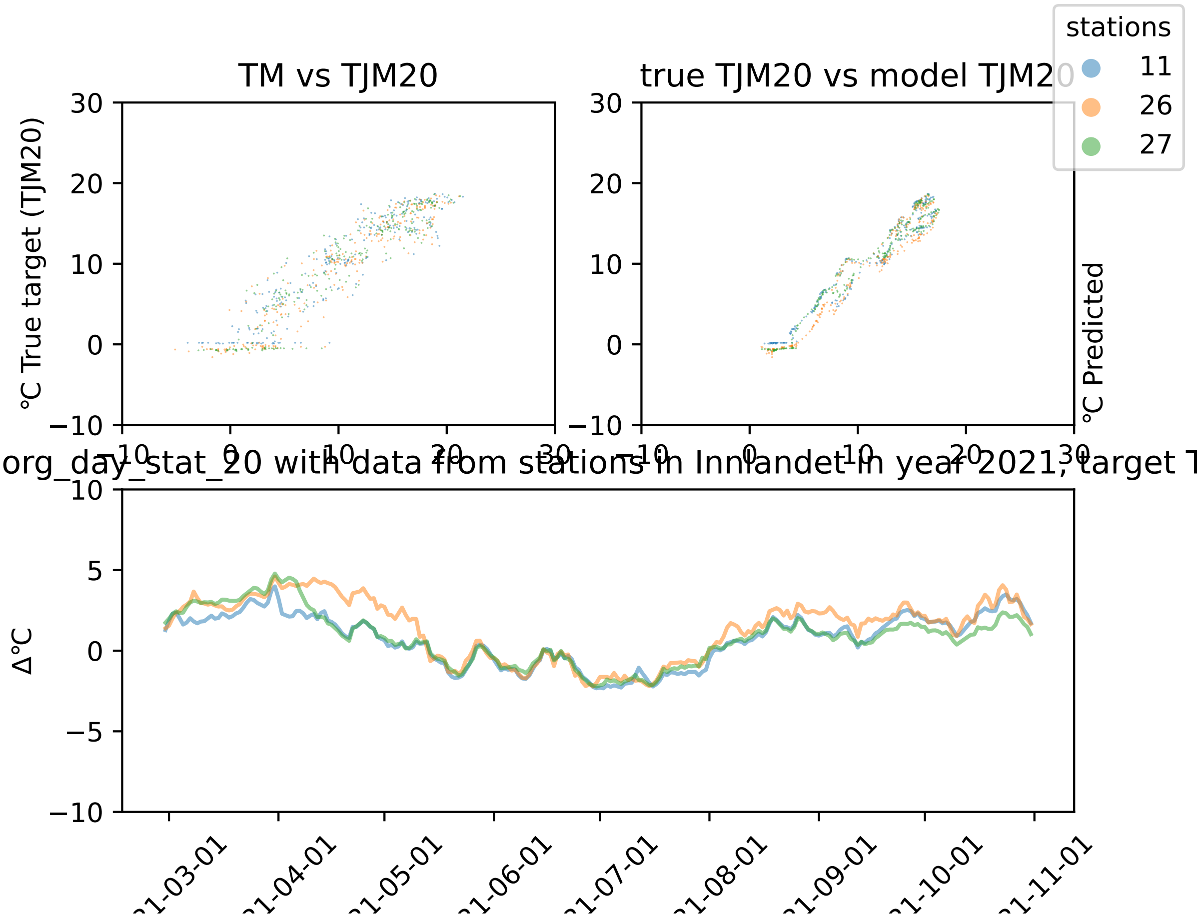


Figure 41: Difference plot for daily Plauborg model in year 2021 and region Innlandet

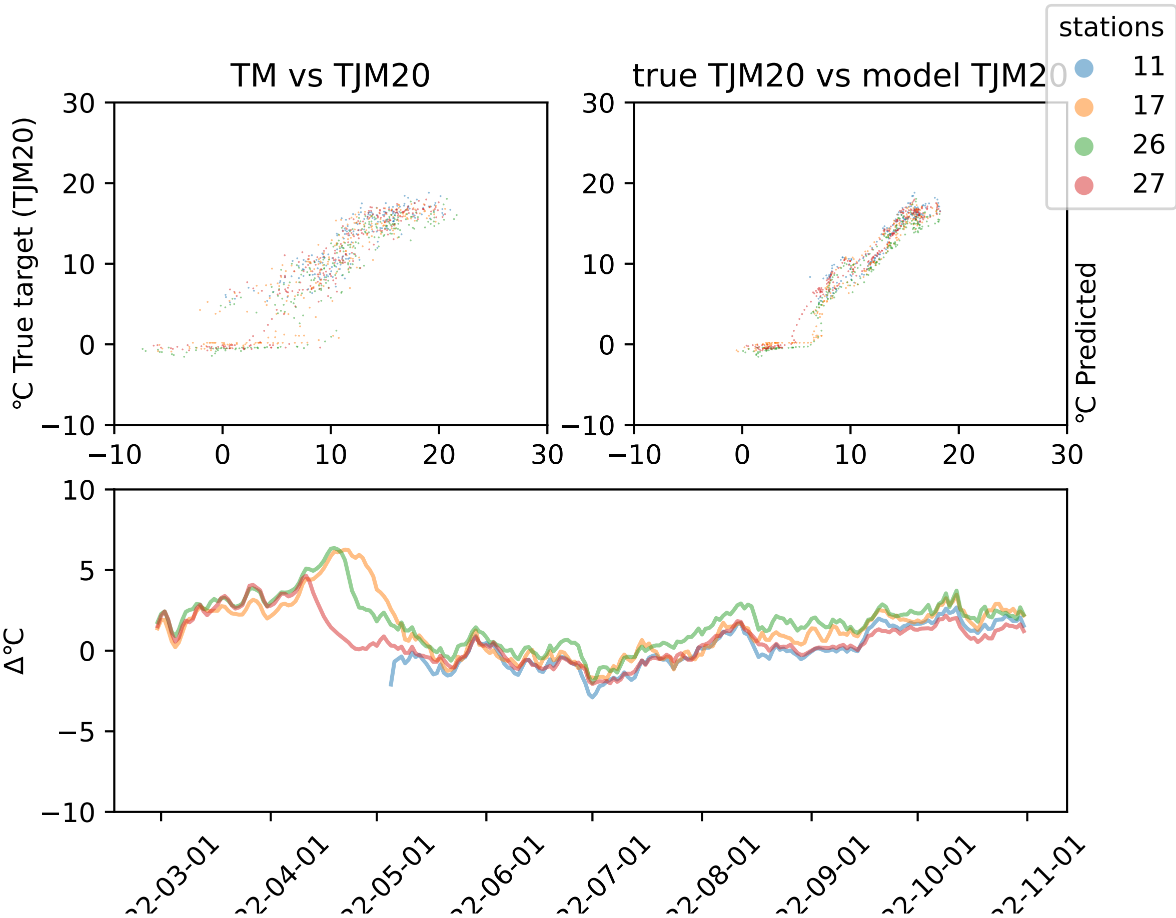


Figure 42: Difference plot for daily Plauborg model in year 2022 and region Innlandet

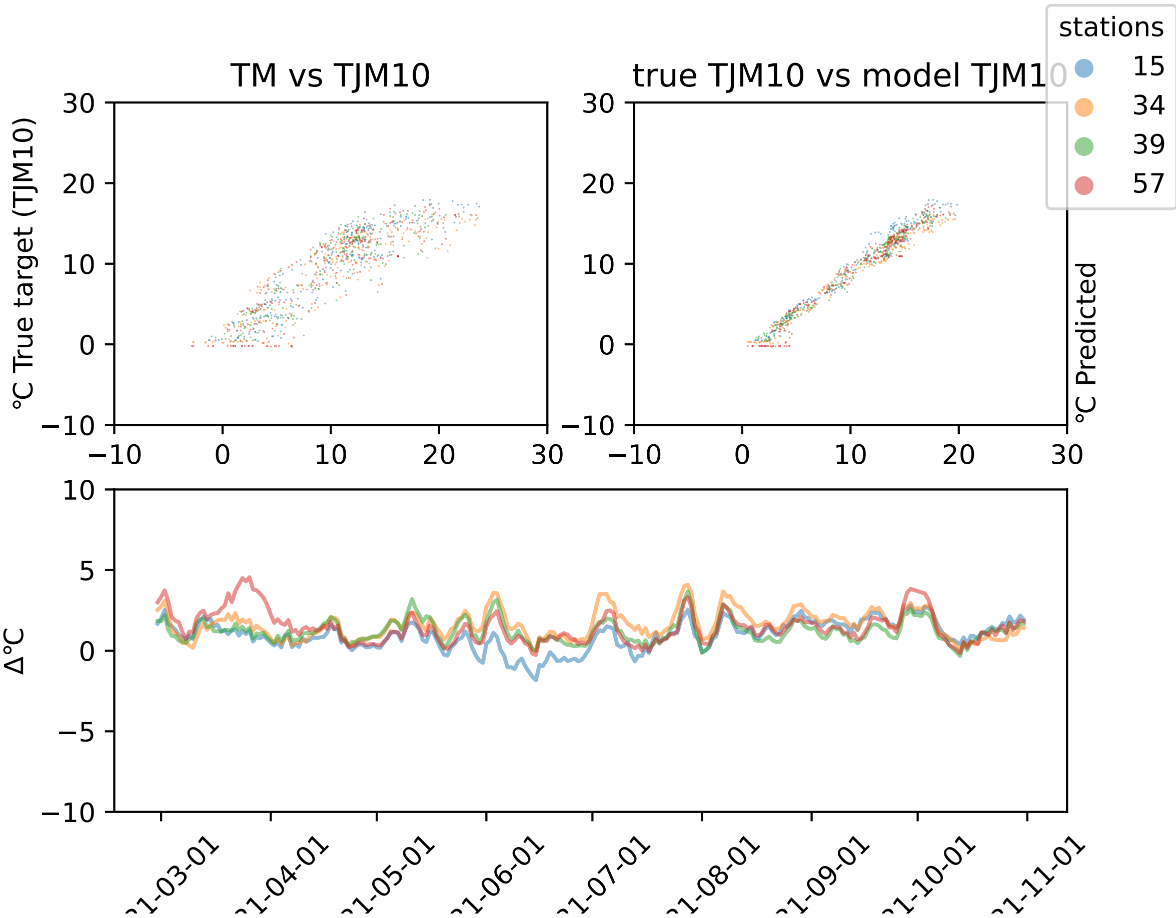


Figure 43: Difference plot for daily Plauborg model in year 2021 and region Trøndelag

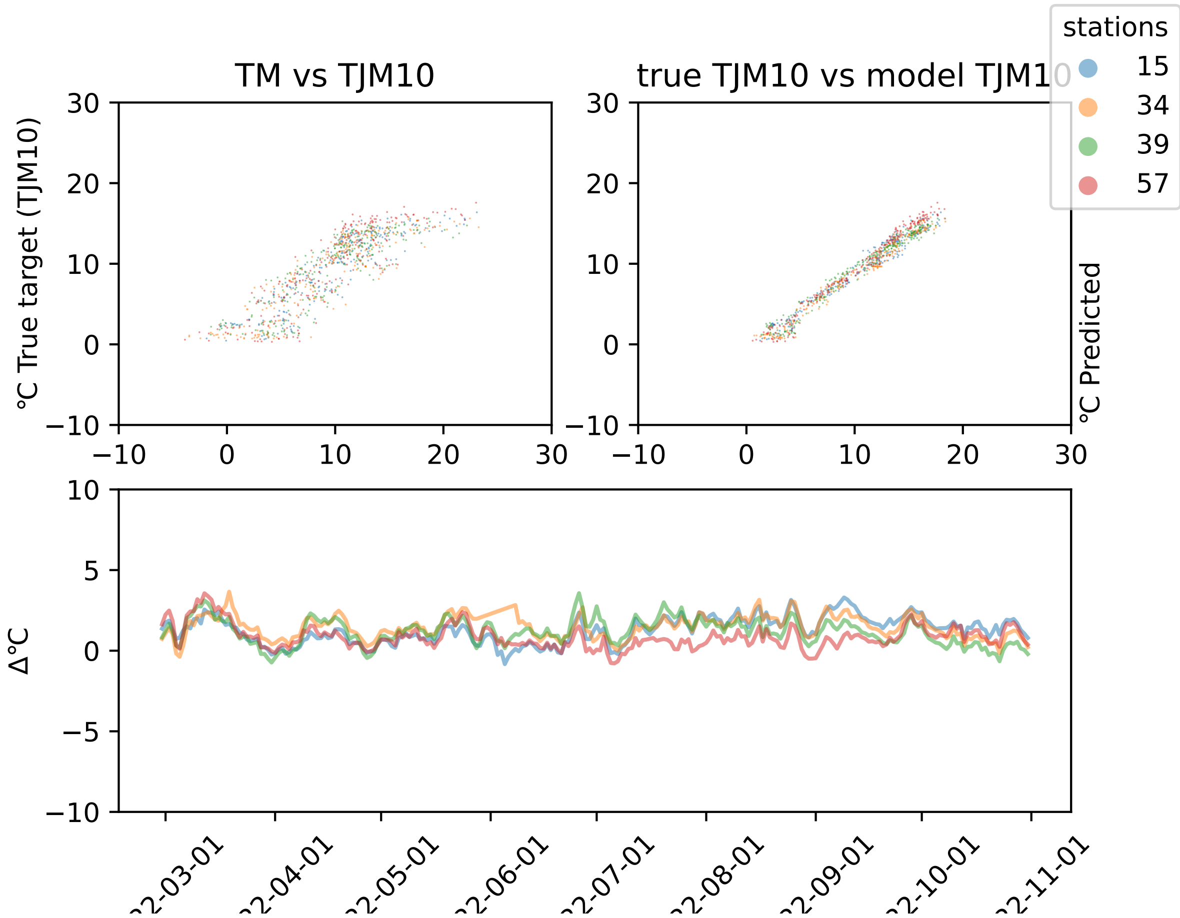


Figure 44: Difference plot for daily Plauborg model in year 2022 and region Trøndelag

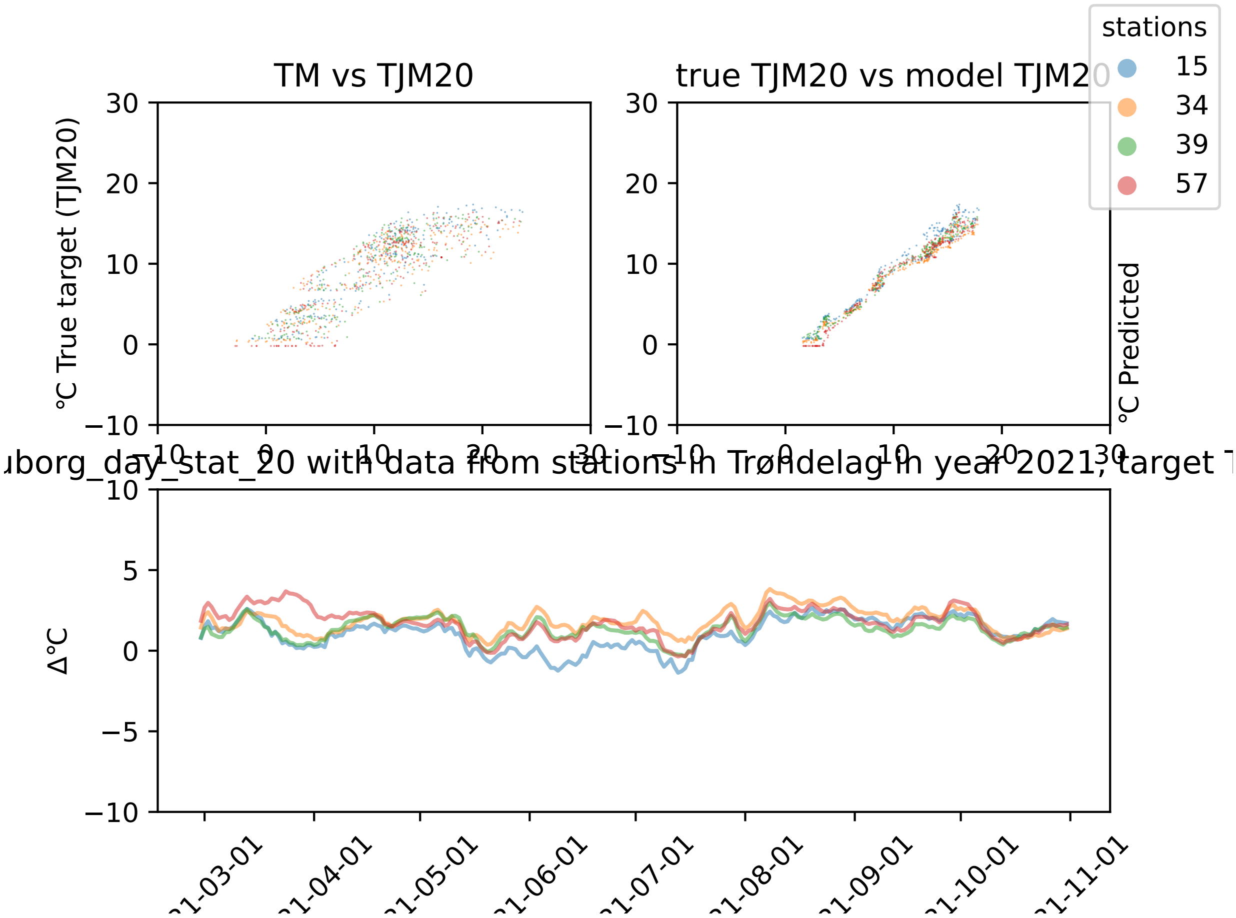


Figure 45: Difference plot for daily Plauborg model in year 2021 and region Trøndelag

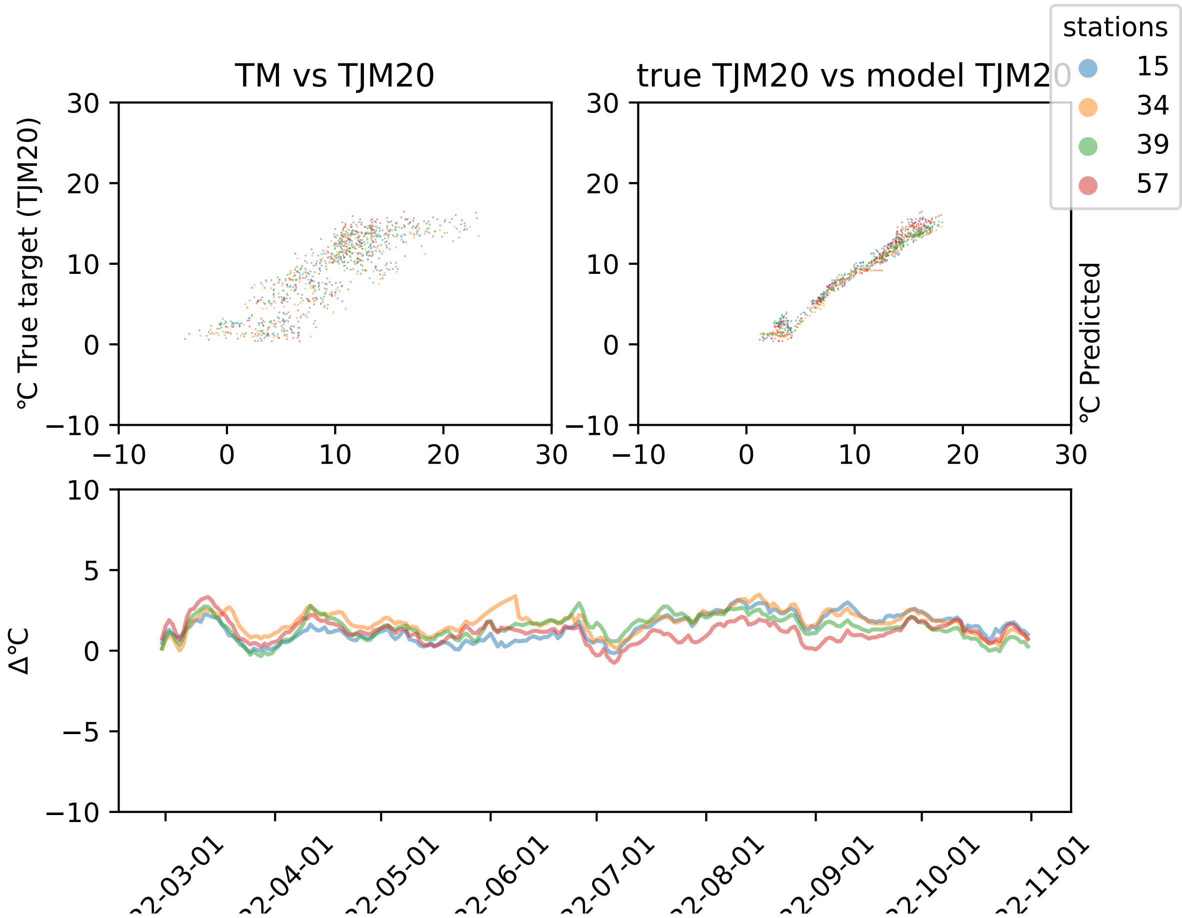


Figure 46: Difference plot for daily Plauborg model in year 2022 and region Trøndelag

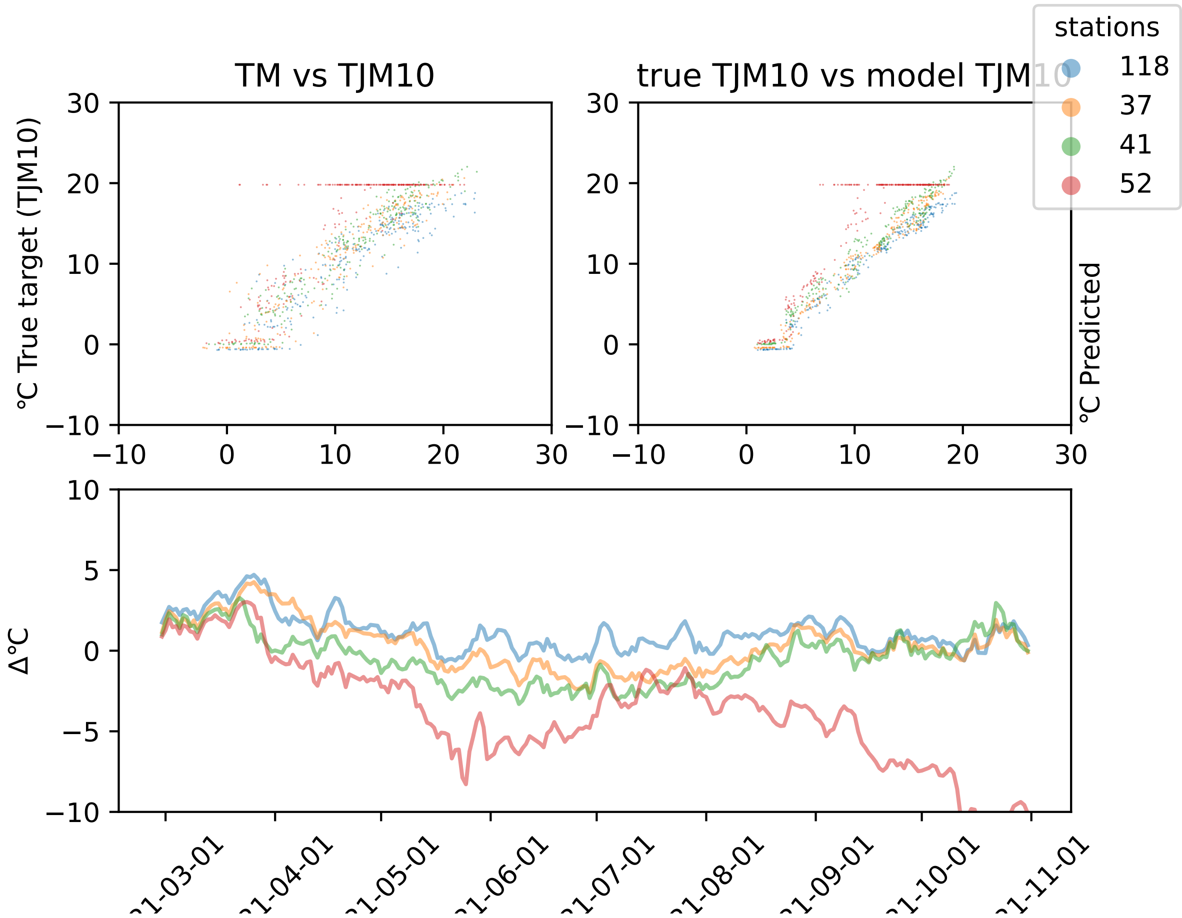


Figure 47: Difference plot for daily Plauborg model in year 2021 and region Østfold

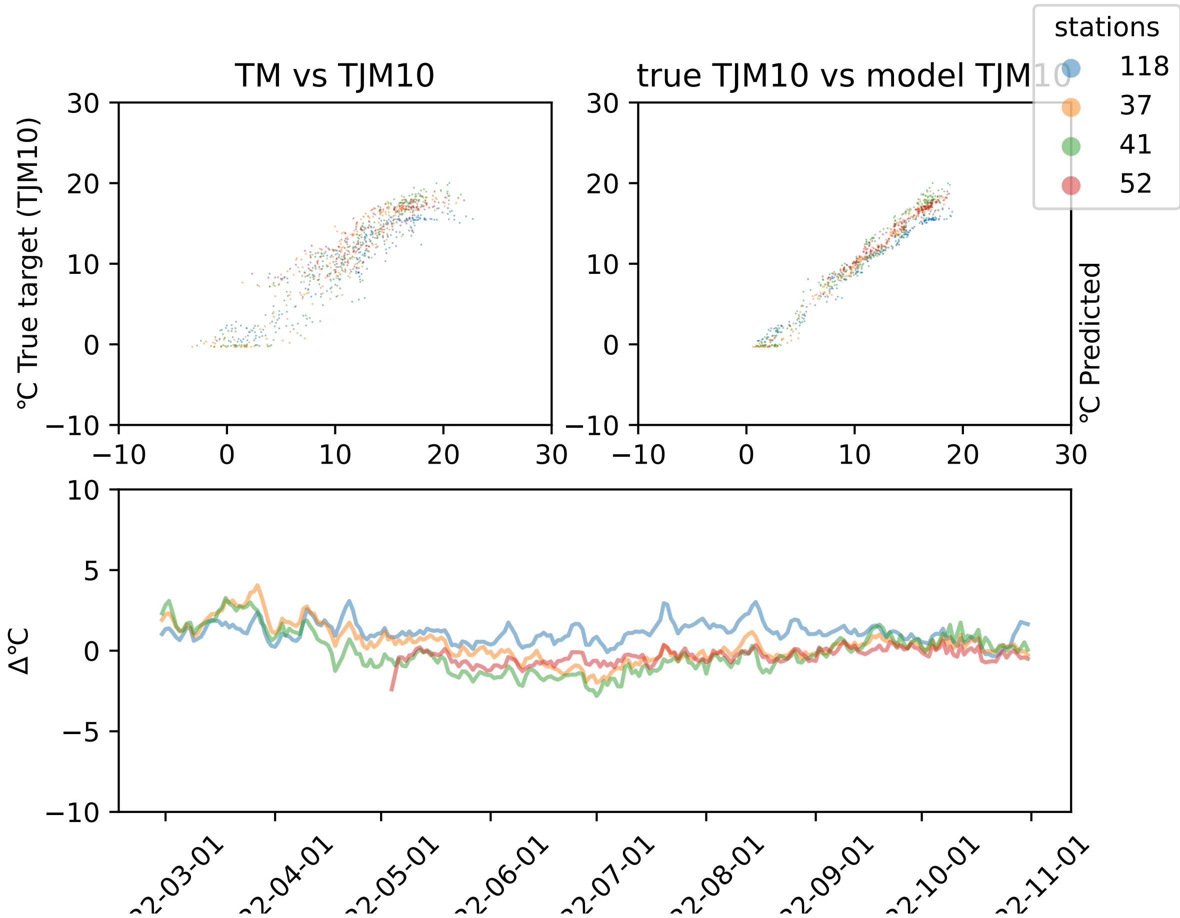


Figure 48: Difference plot for daily Plauborg model in year 2022 and region Østfold

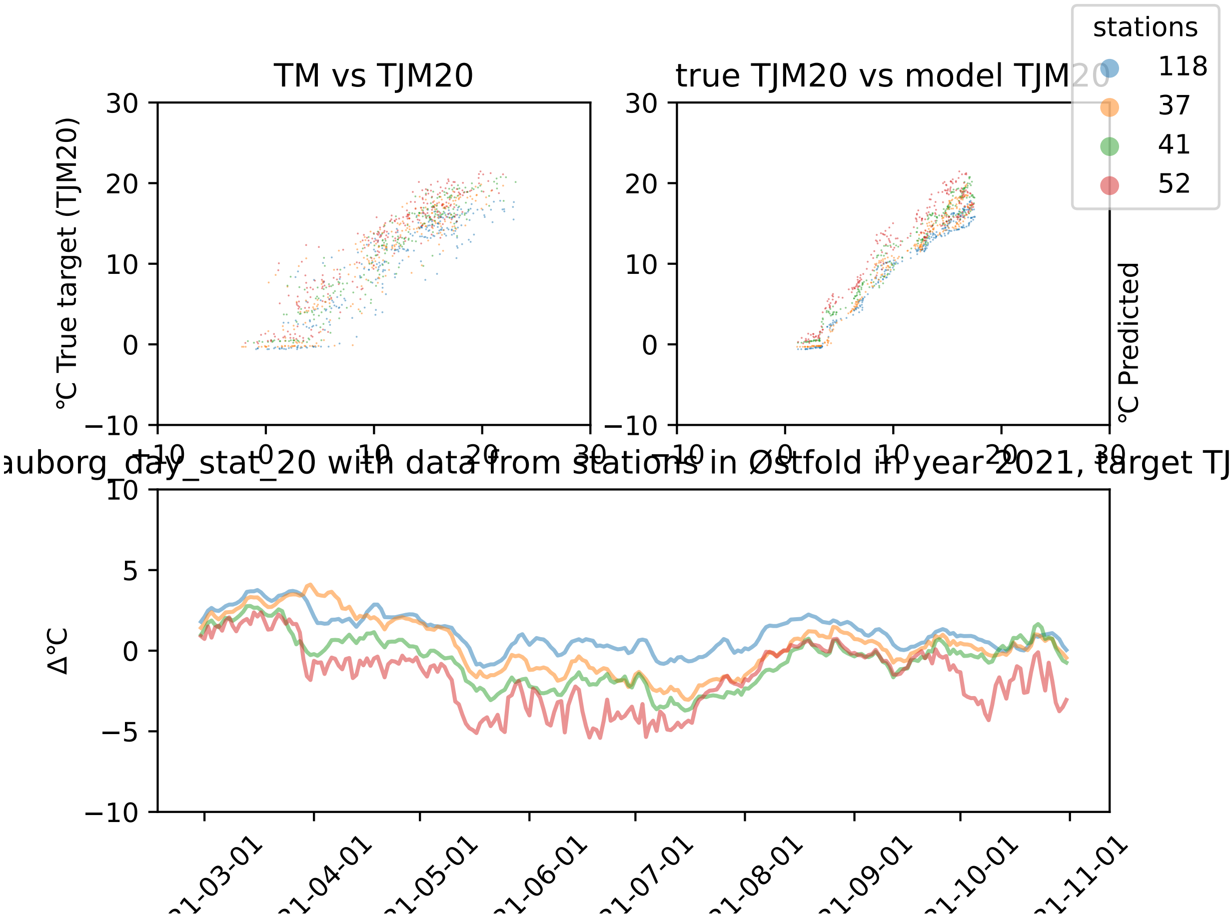


Figure 49: Difference plot for daily Plauborg model in year 2021 and region Østfold

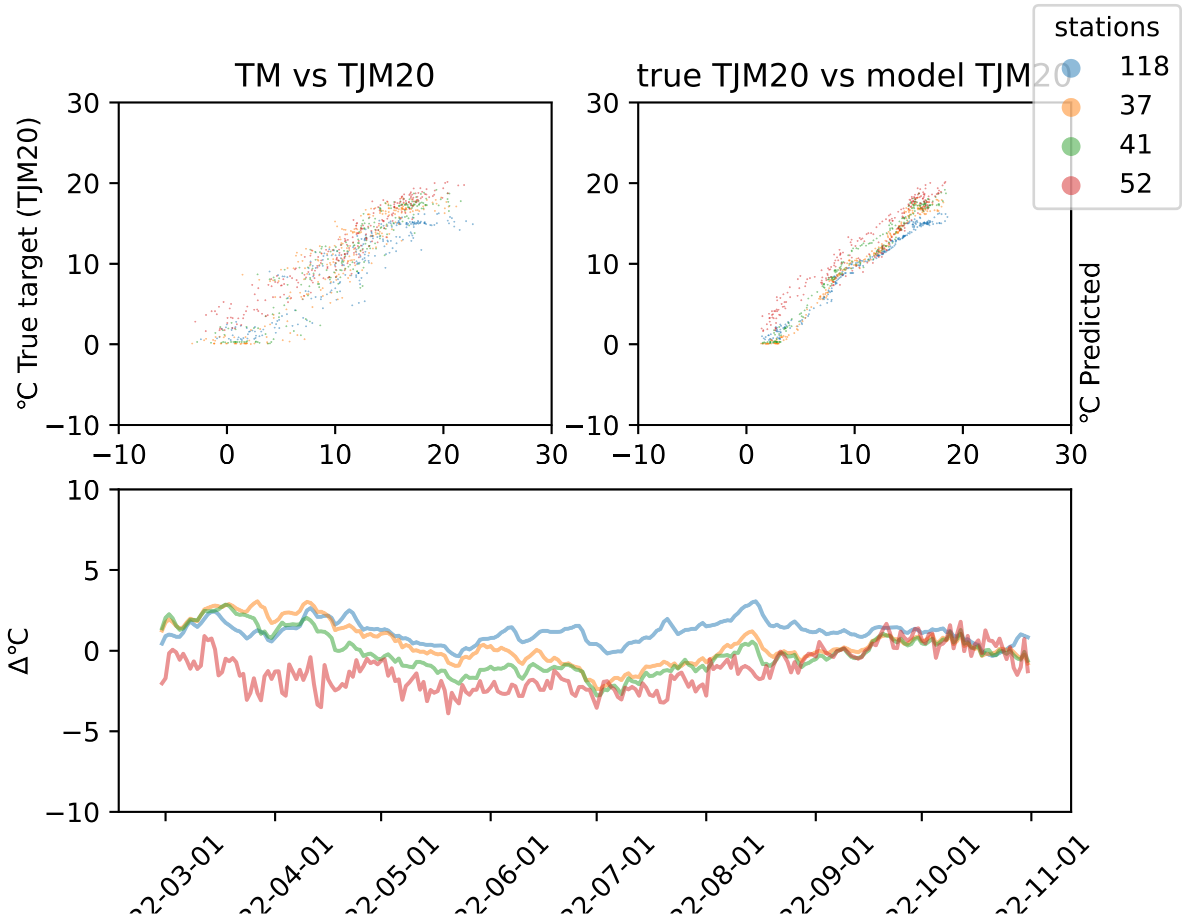


Figure 50: Difference plot for daily Plauborg model in year 2022 and region Østfold

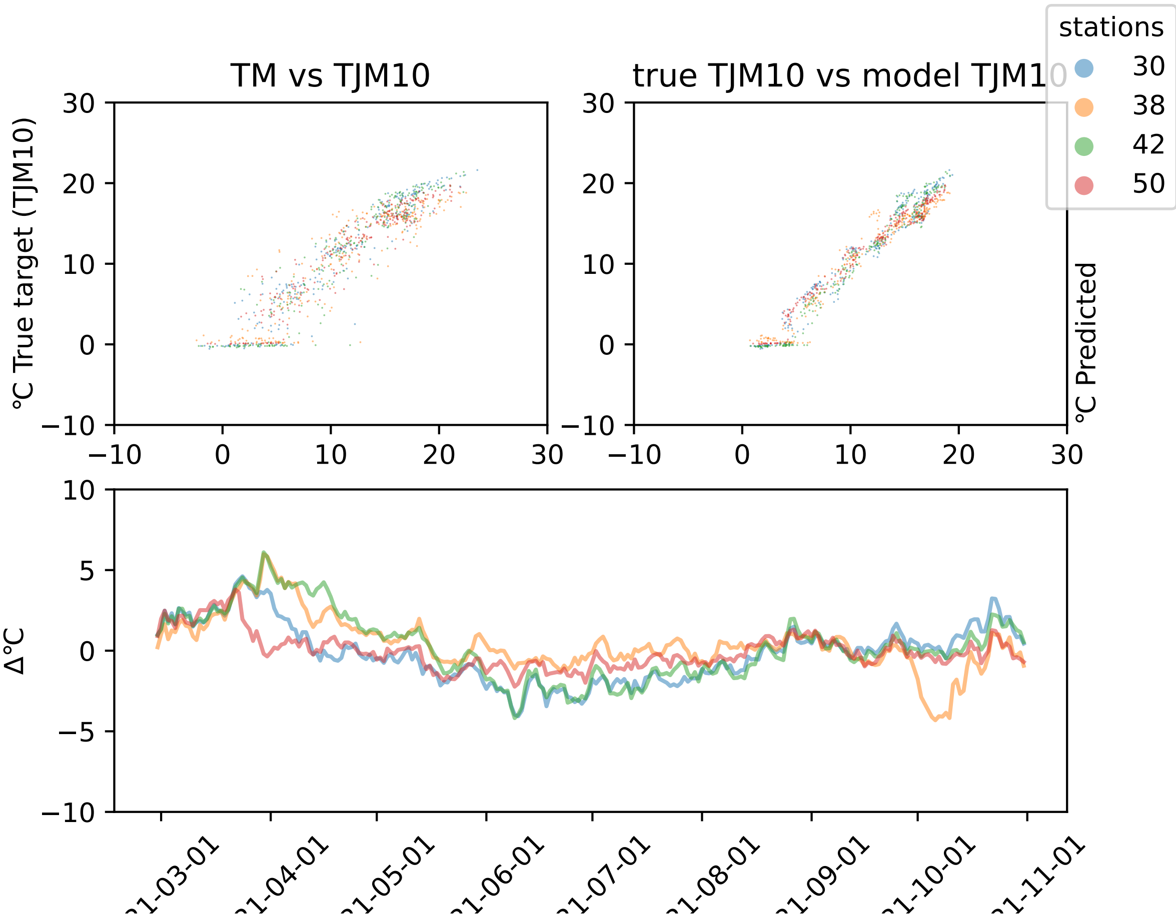


Figure 51: Difference plot for daily Plauborg model in year 2021 and region Vestfold

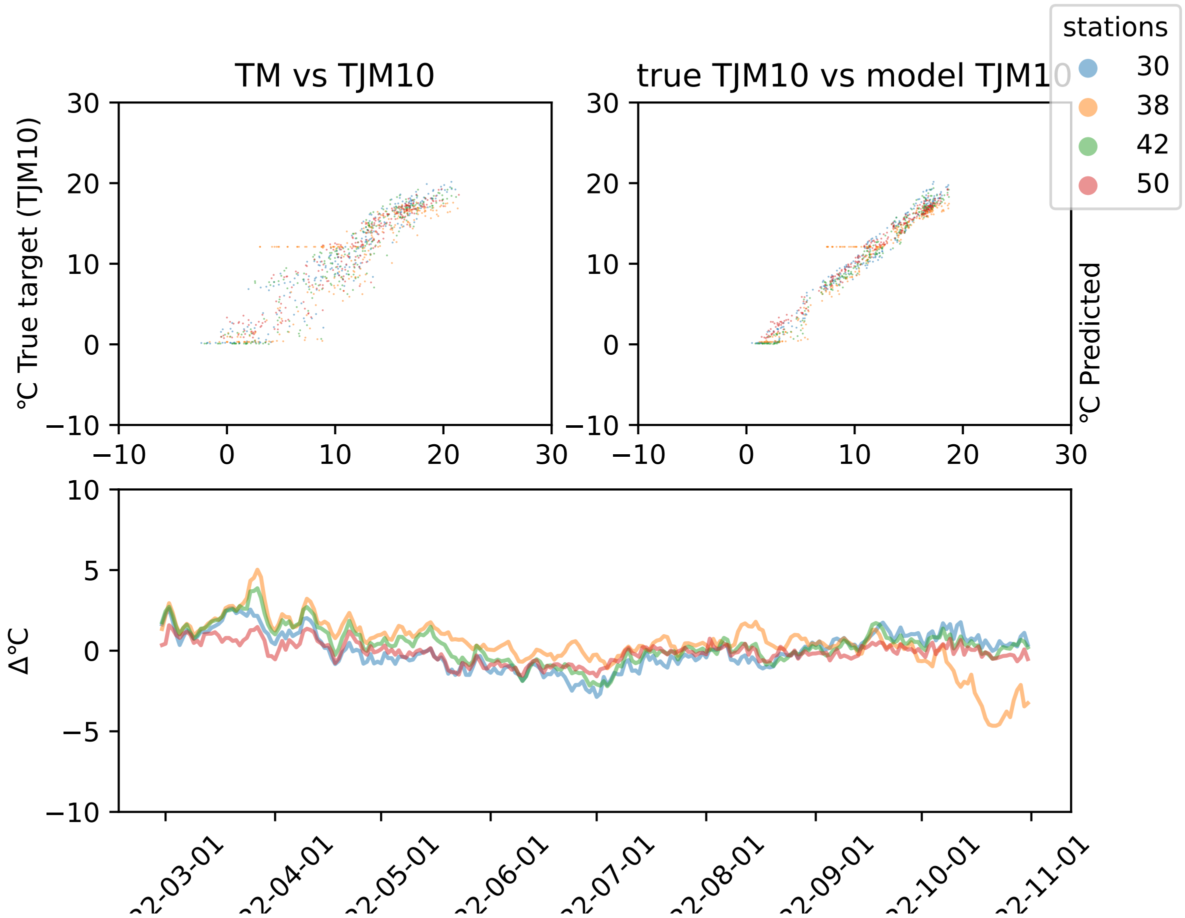


Figure 52: Difference plot for daily Plauborg model in year 2022 and region Vestfold

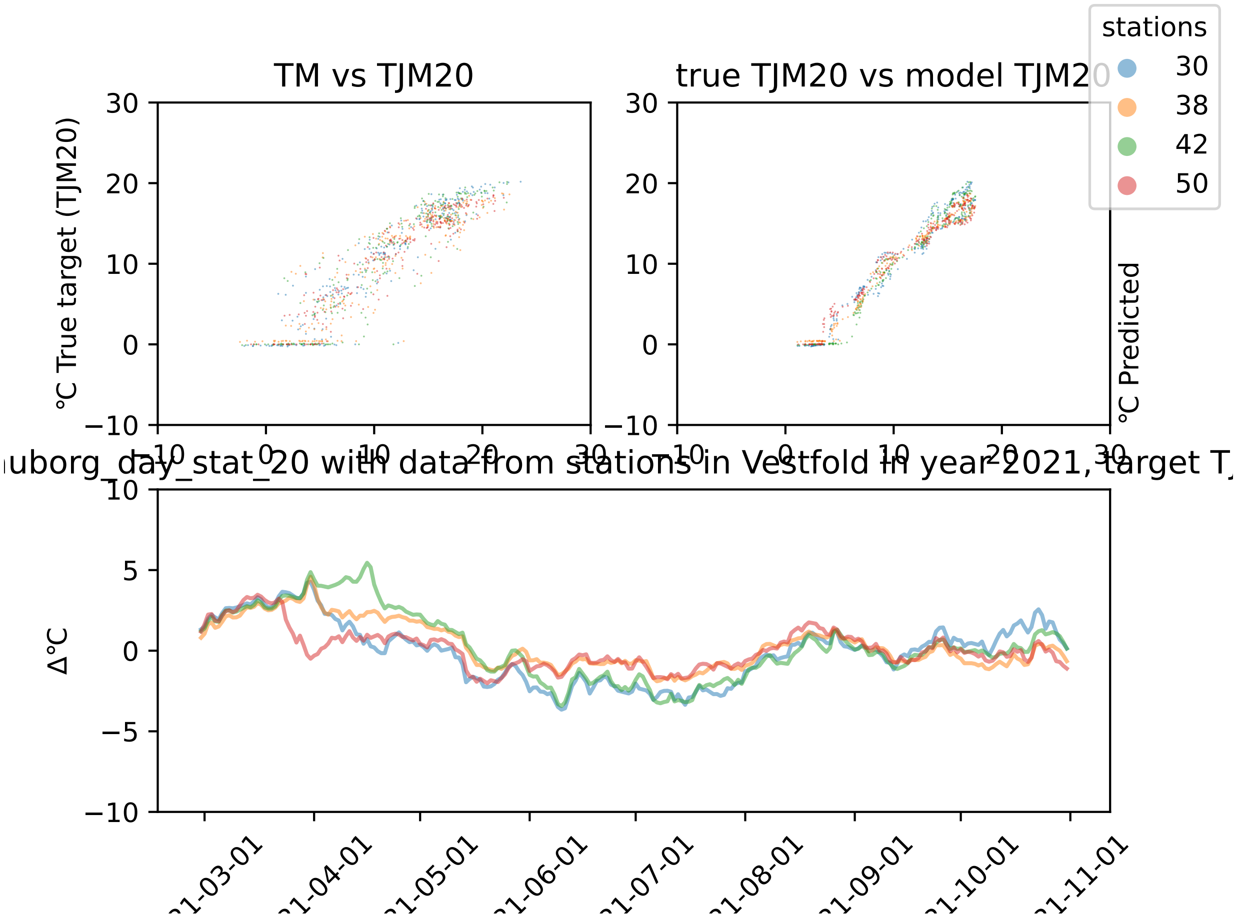


Figure 53: Difference plot for daily Plauborg model in year 2021 and region Vestfold

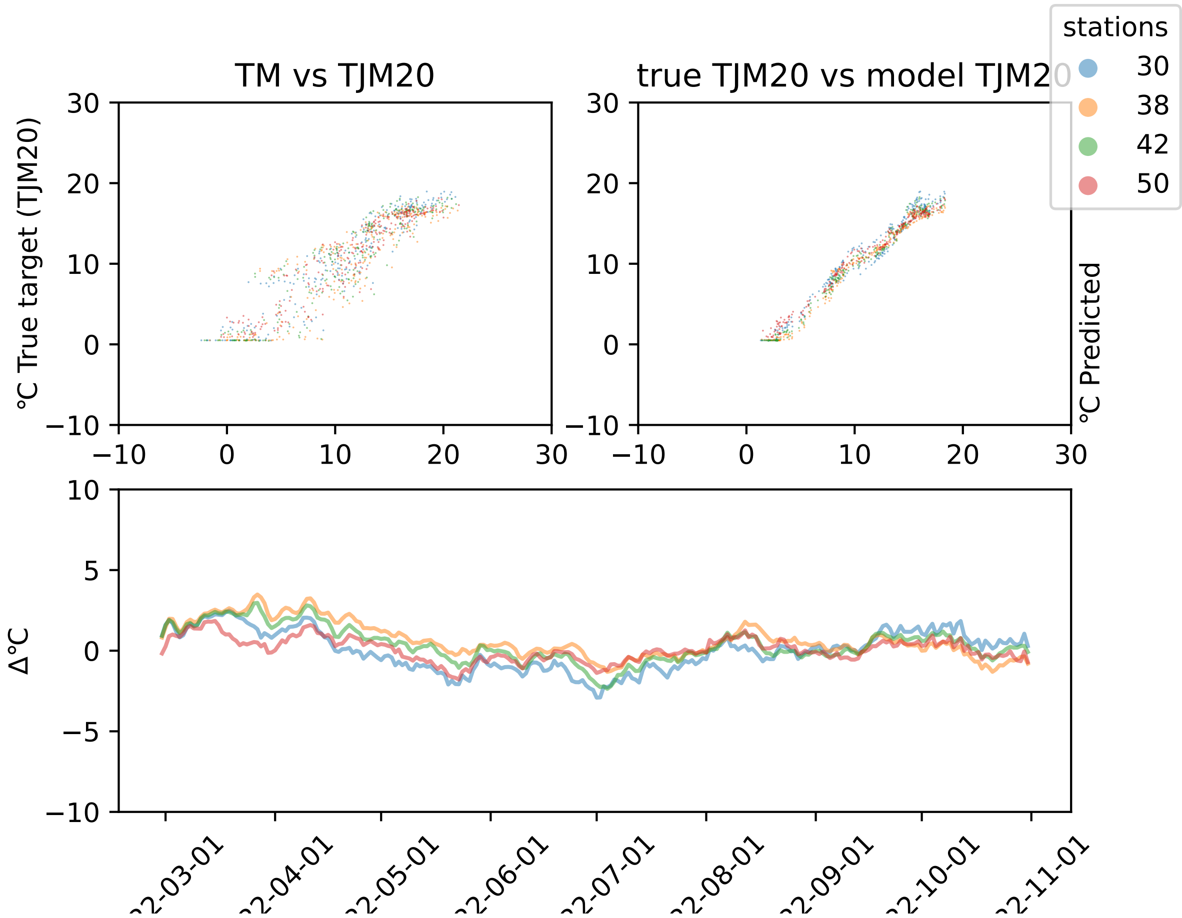


Figure 54: Difference plot for daily Plauborg model in year 2022 and region Vestfold

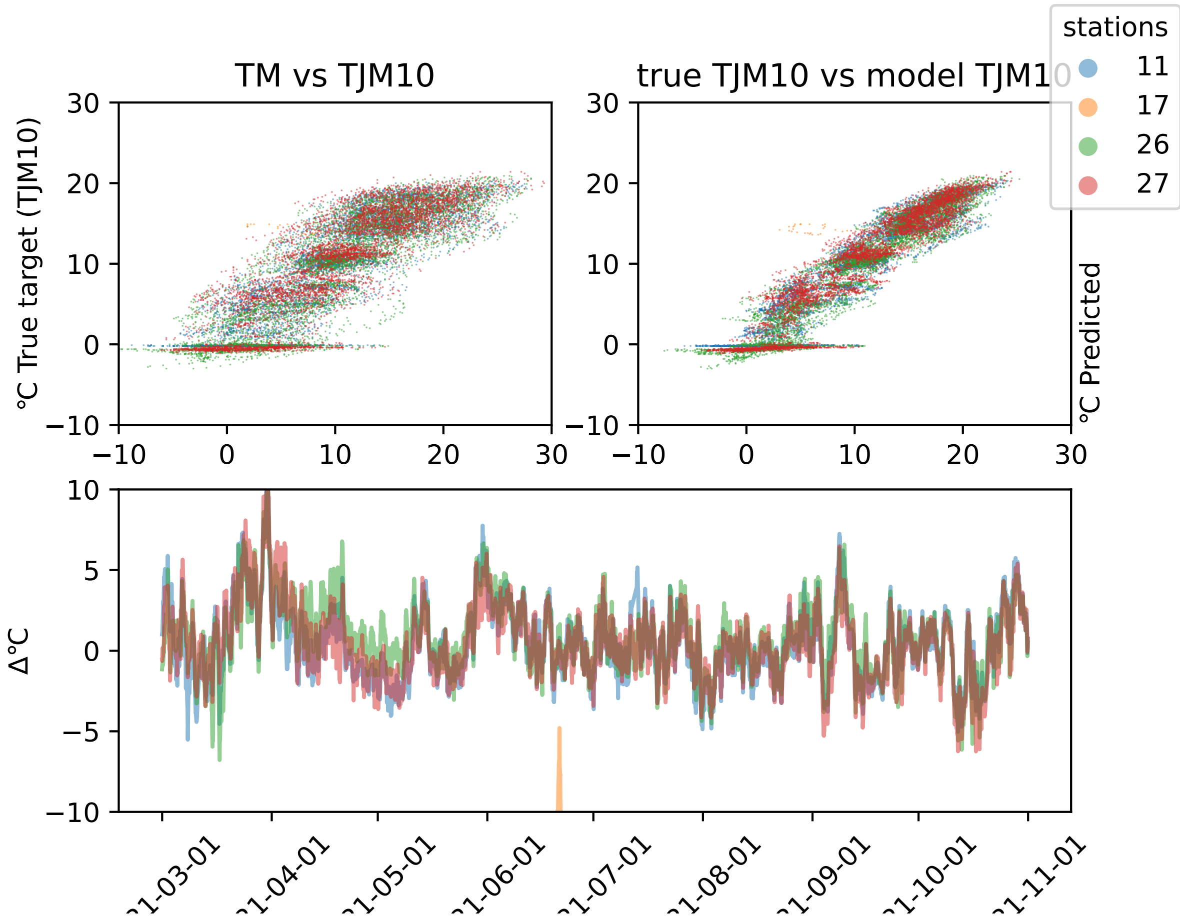


Figure 55: Difference plot for hourly Plauborg model in year 2021 and region Innlandet

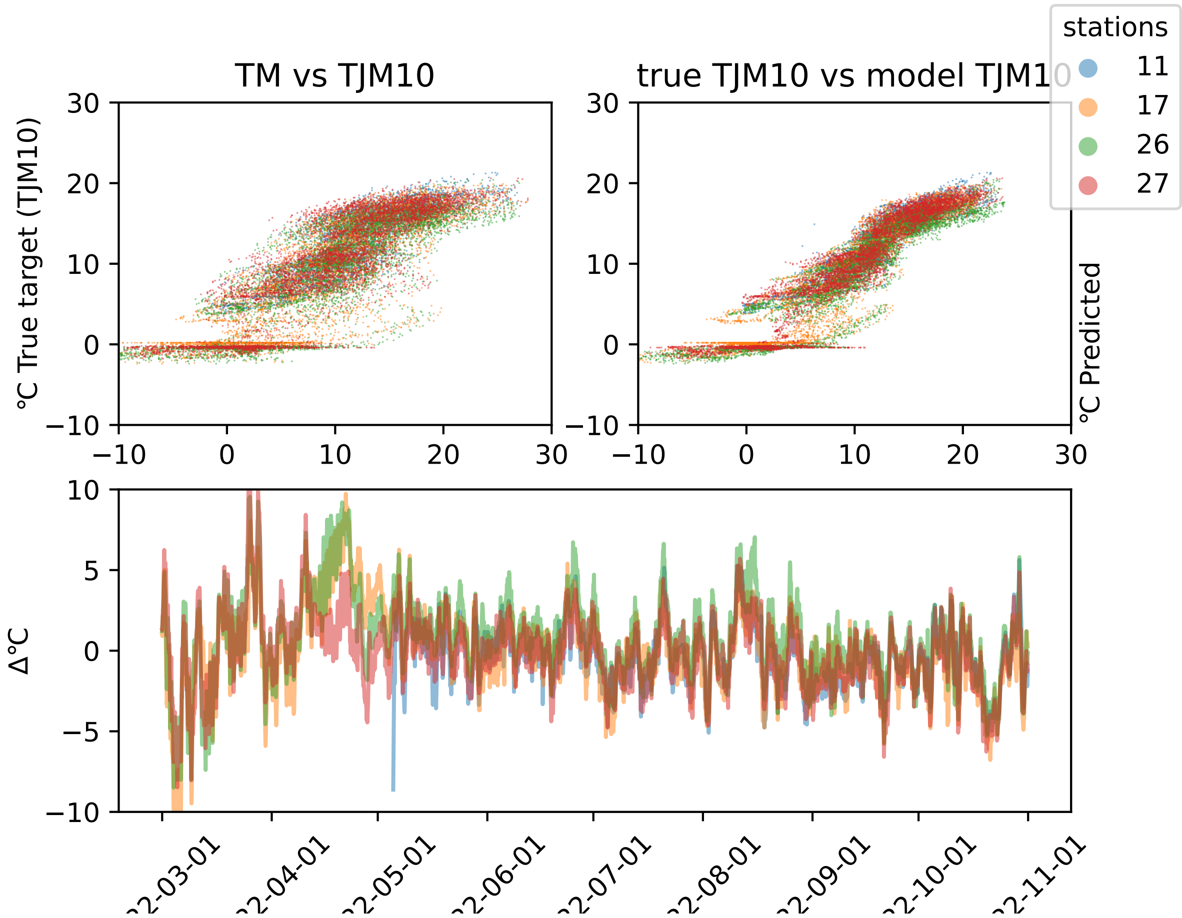


Figure 56: Difference plot for hourly Plauborg model in year 2022 and region Innlandet

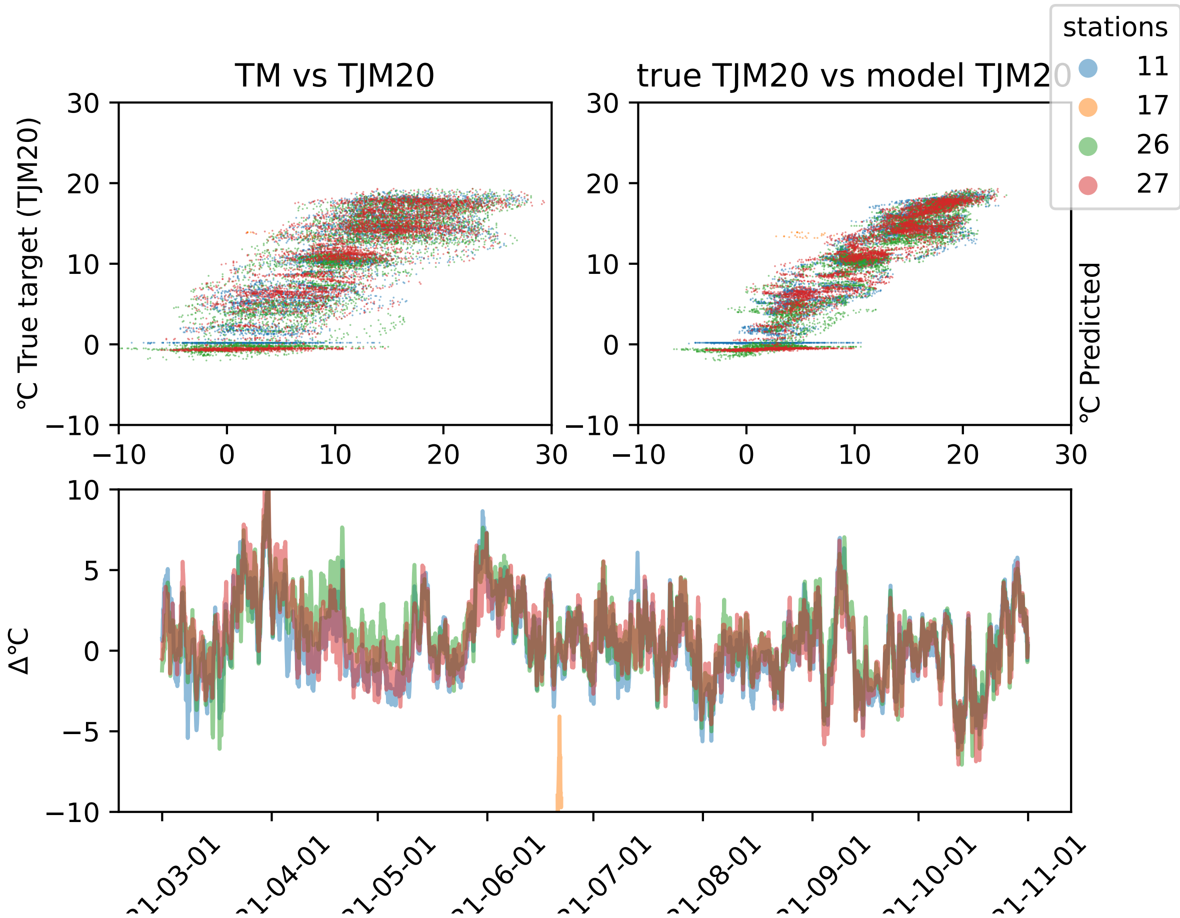


Figure 57: Difference plot for hourly Plauborg model in year 2021 and region Innlandet

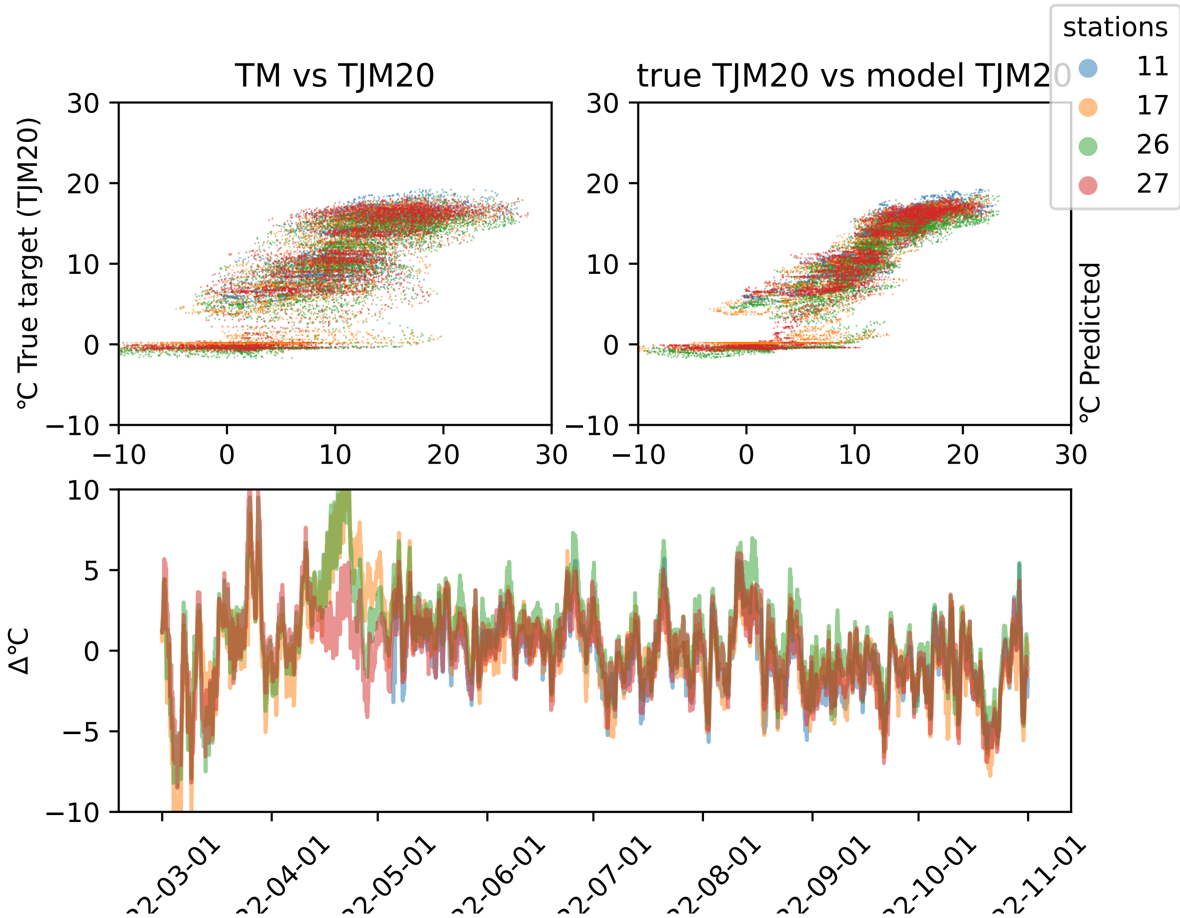


Figure 58: Difference plot for hourly Plauborg model in year 2022 and region Innlandet

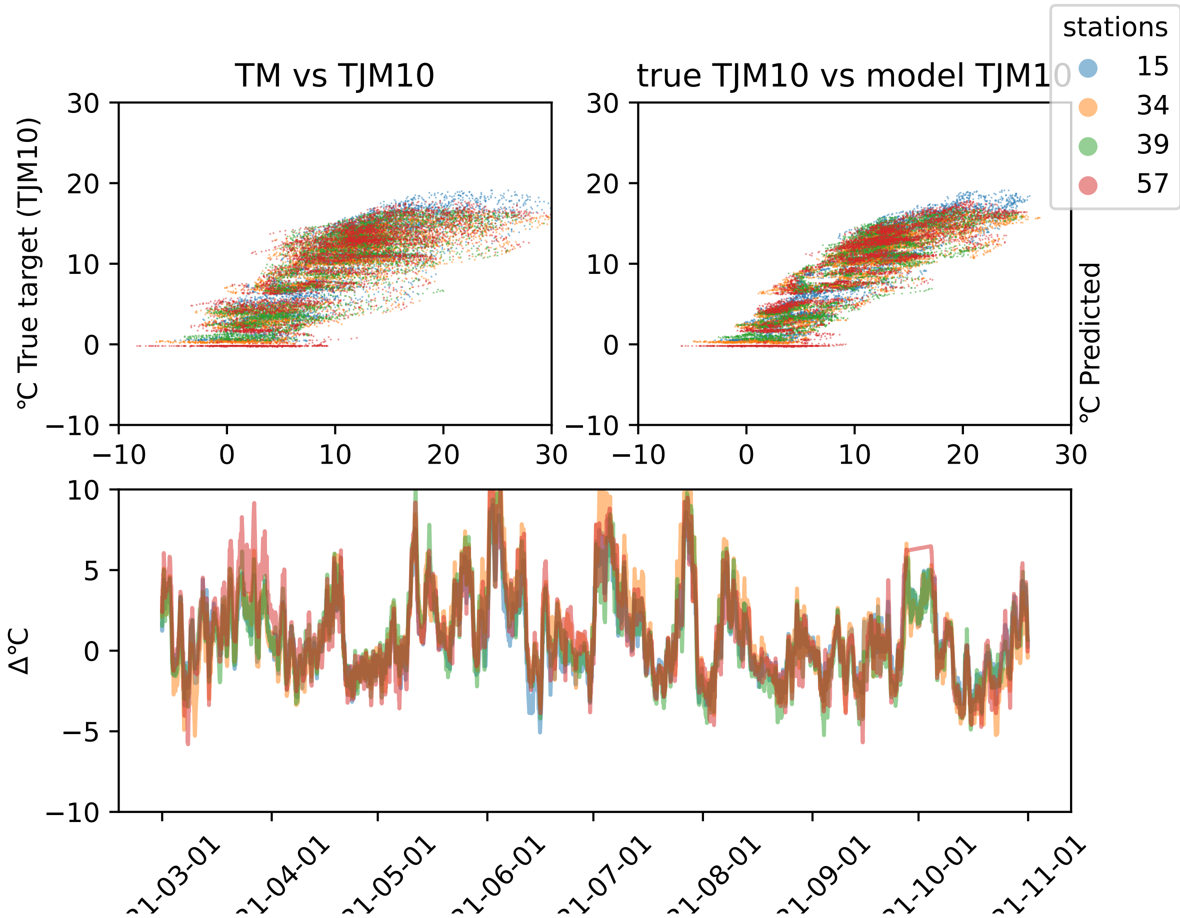


Figure 59: Difference plot for hourly Plauborg model in year 2021 and region Trøndelag

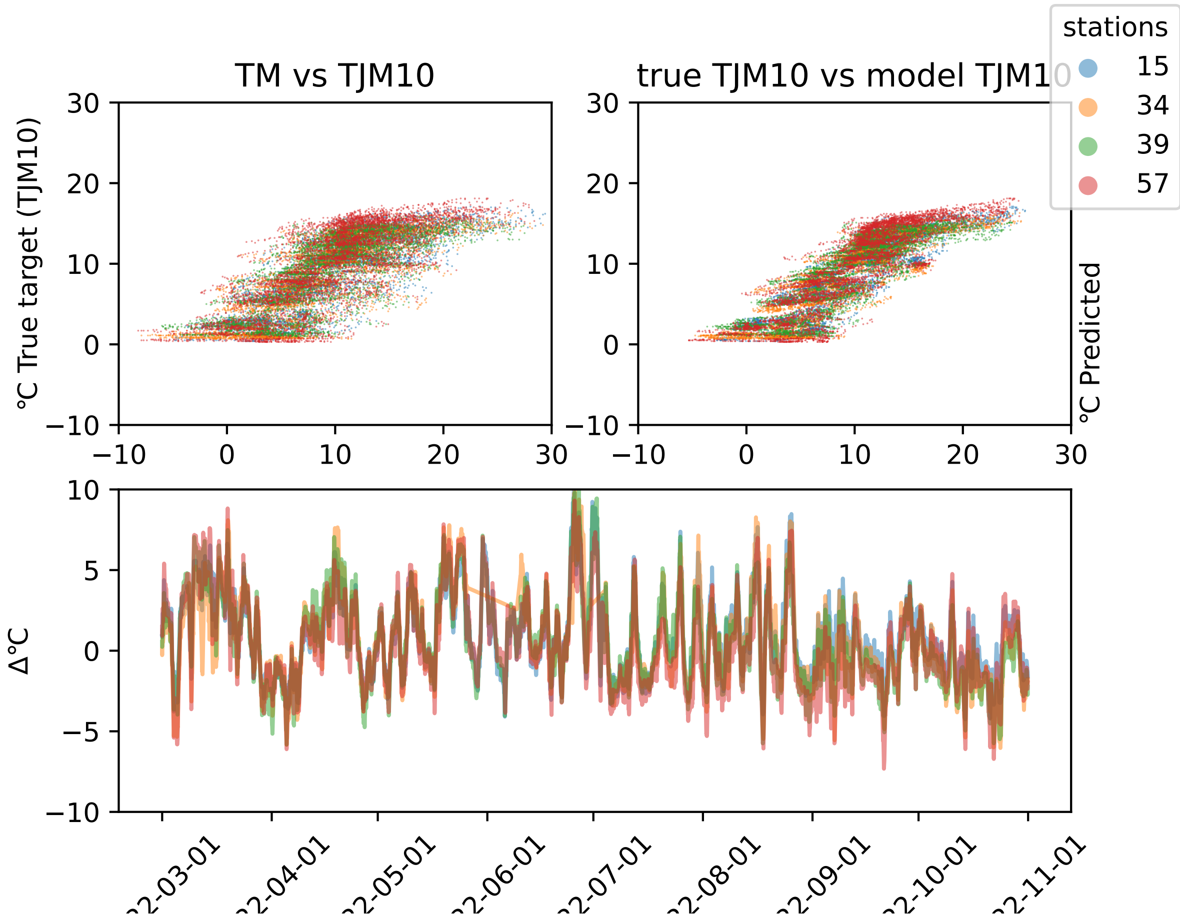


Figure 60: Difference plot for hourly Plauborg model in year 2022 and region Trøndelag

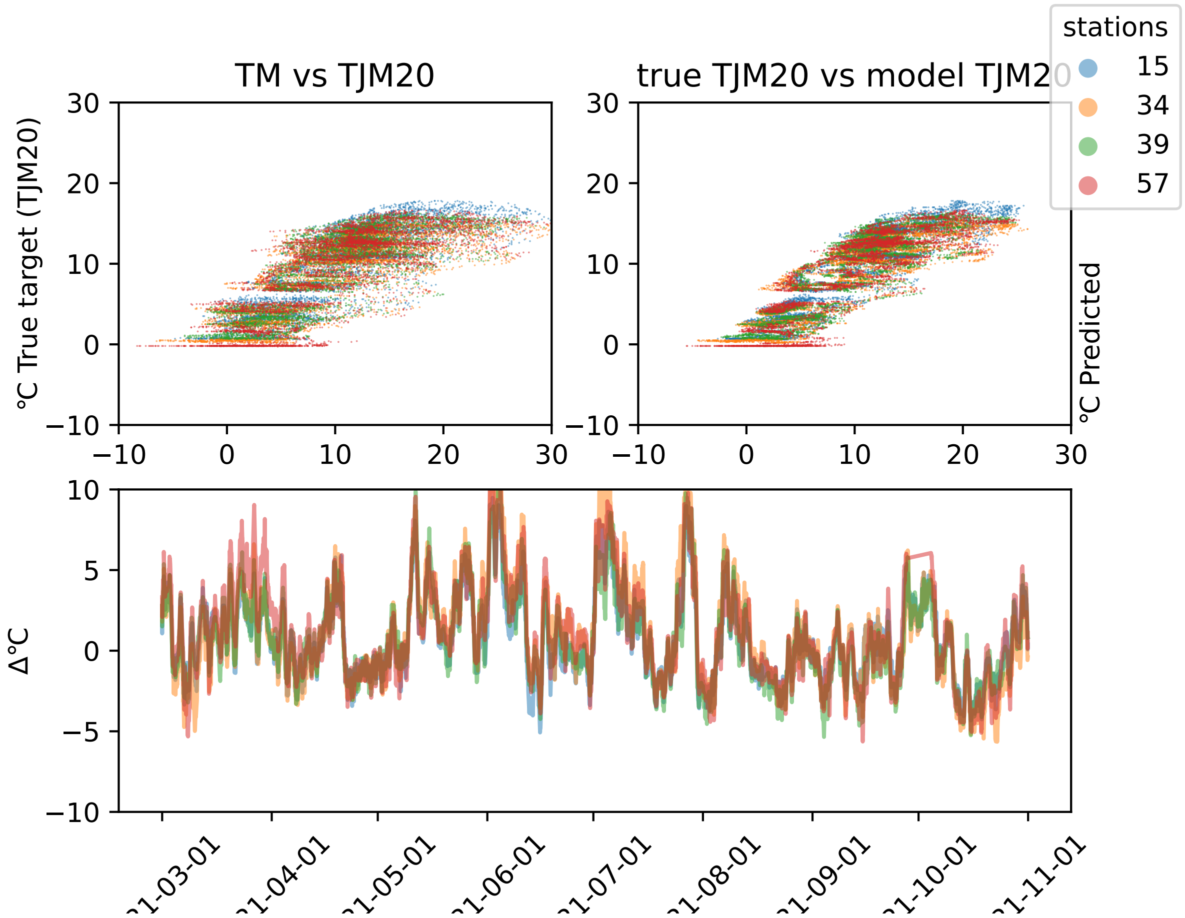


Figure 61: Difference plot for hourly Plauborg model in year 2021 and region Trøndelag

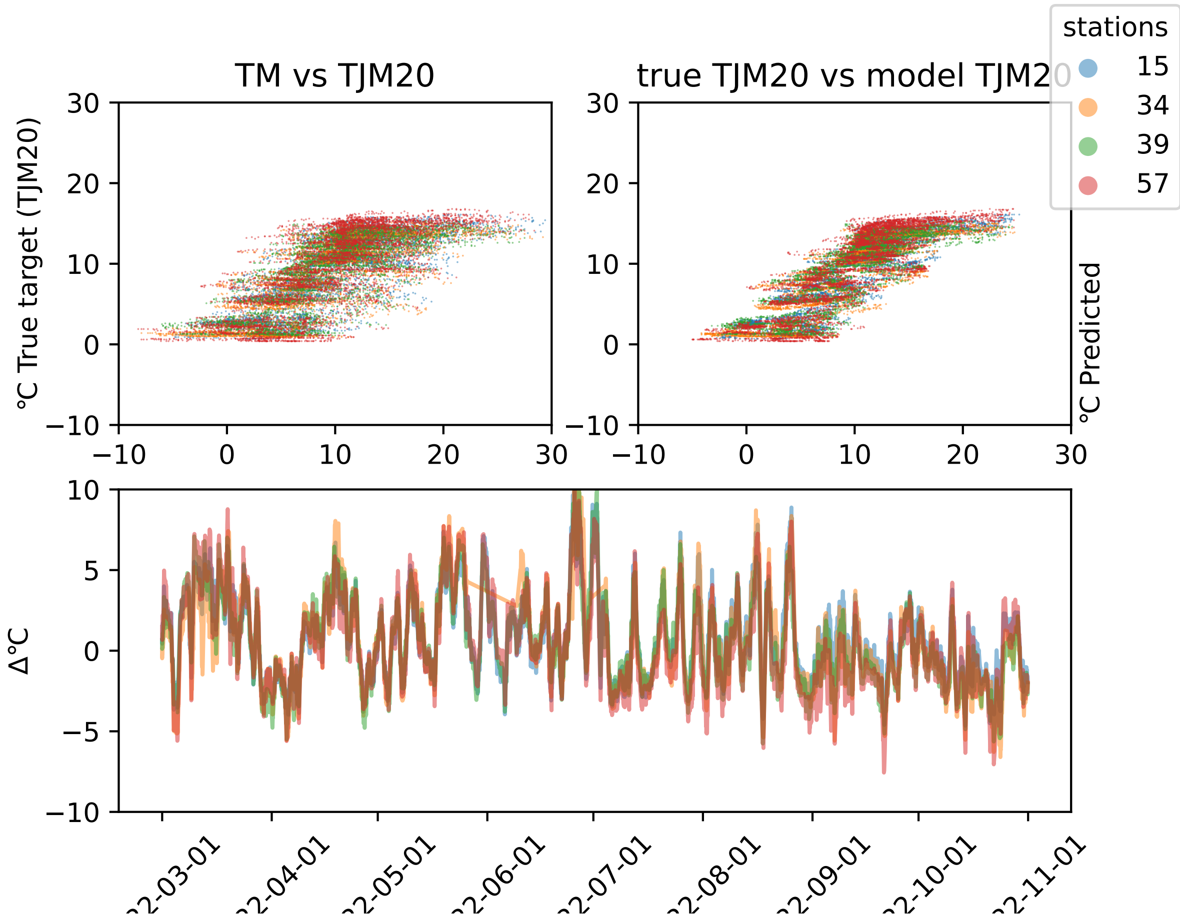


Figure 62: Difference plot for hourly Plauborg model in year 2022 and region Trøndelag

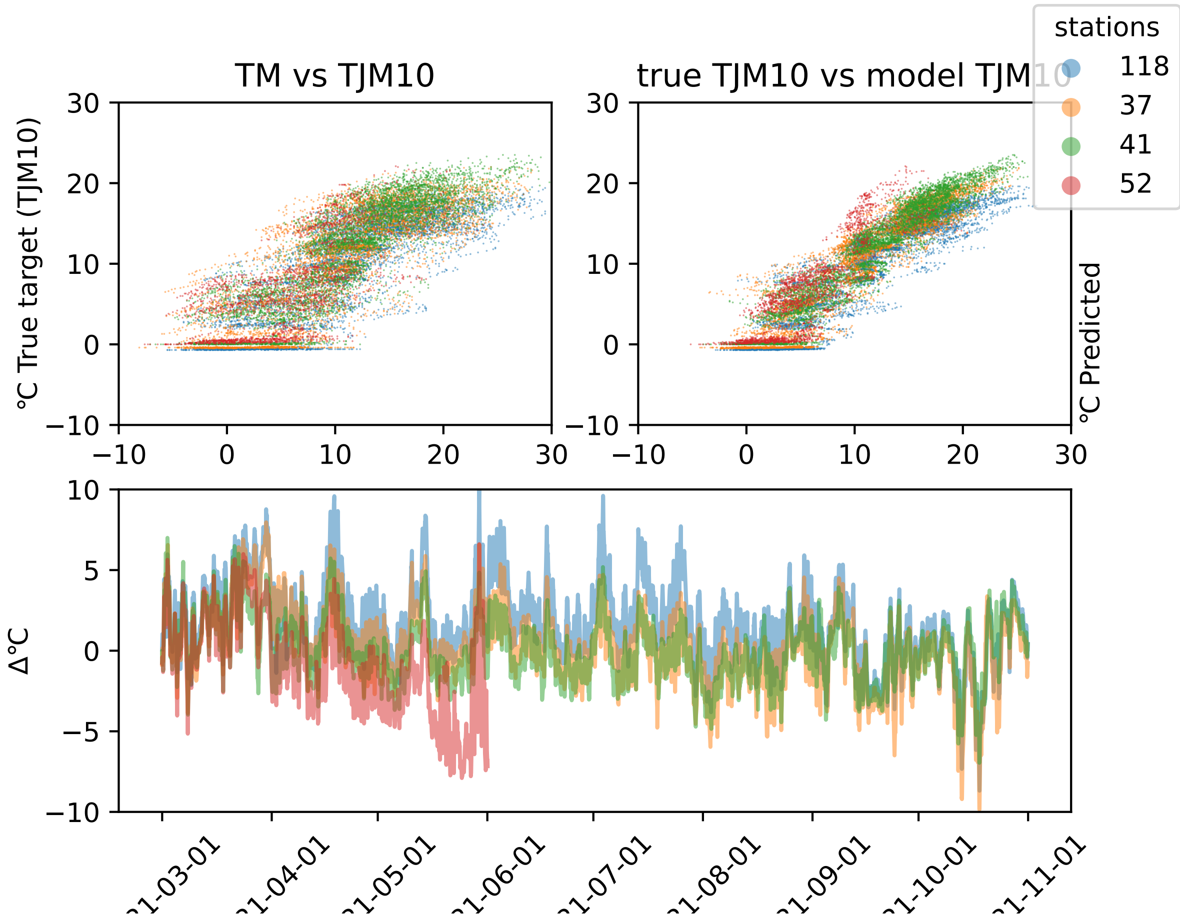


Figure 63: Difference plot for hourly Plauborg model in year 2021 and region Østfold

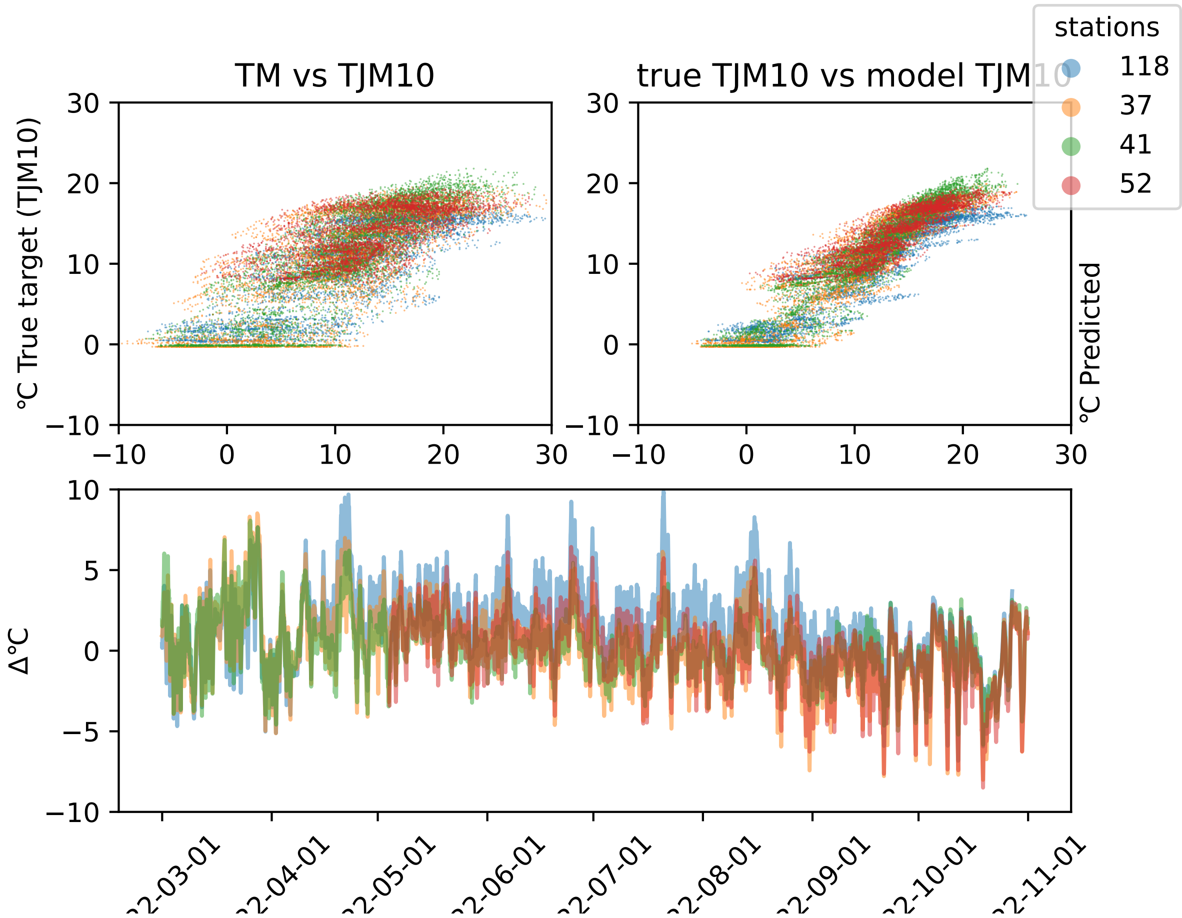


Figure 64: Difference plot for hourly Plauborg model in year 2022 and region Østfold

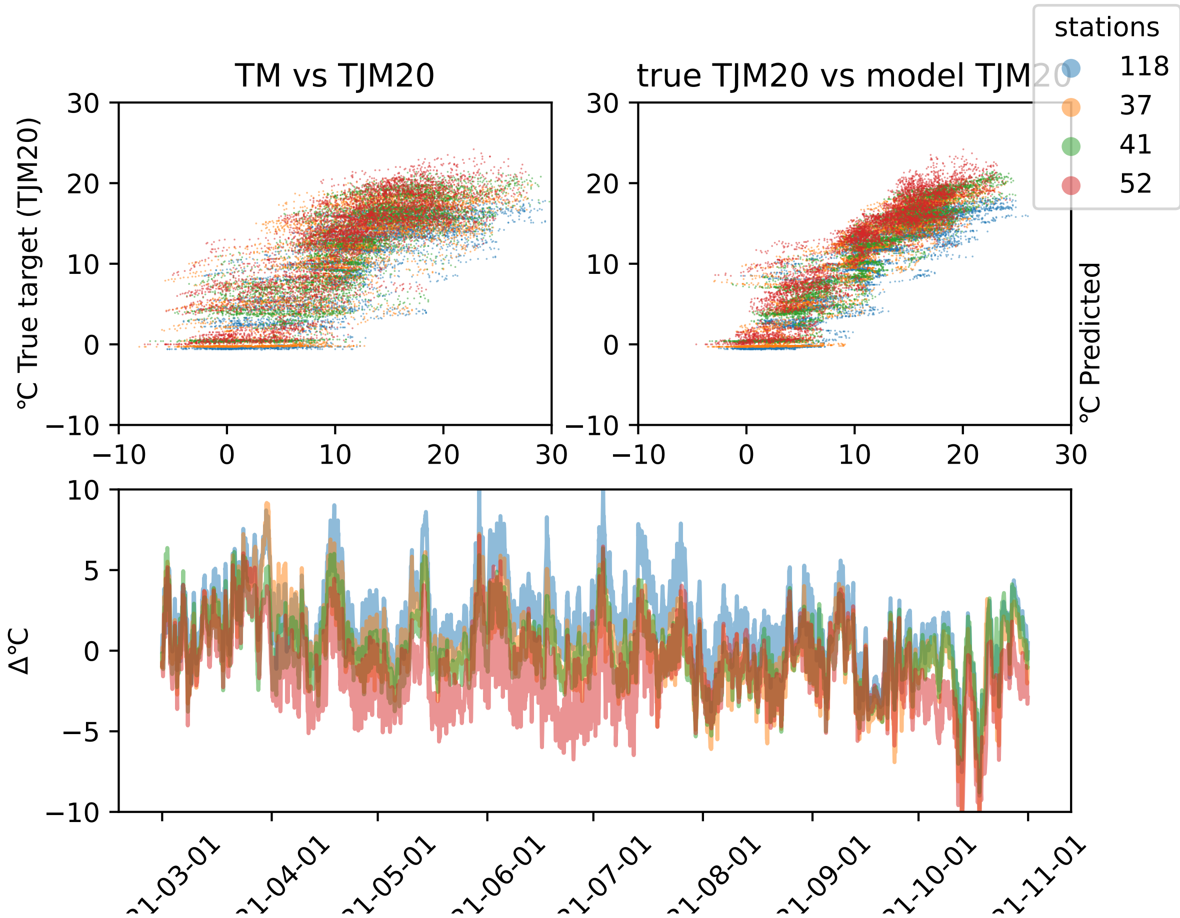


Figure 65: Difference plot for hourly Plauborg model in year 2021 and region Østfold

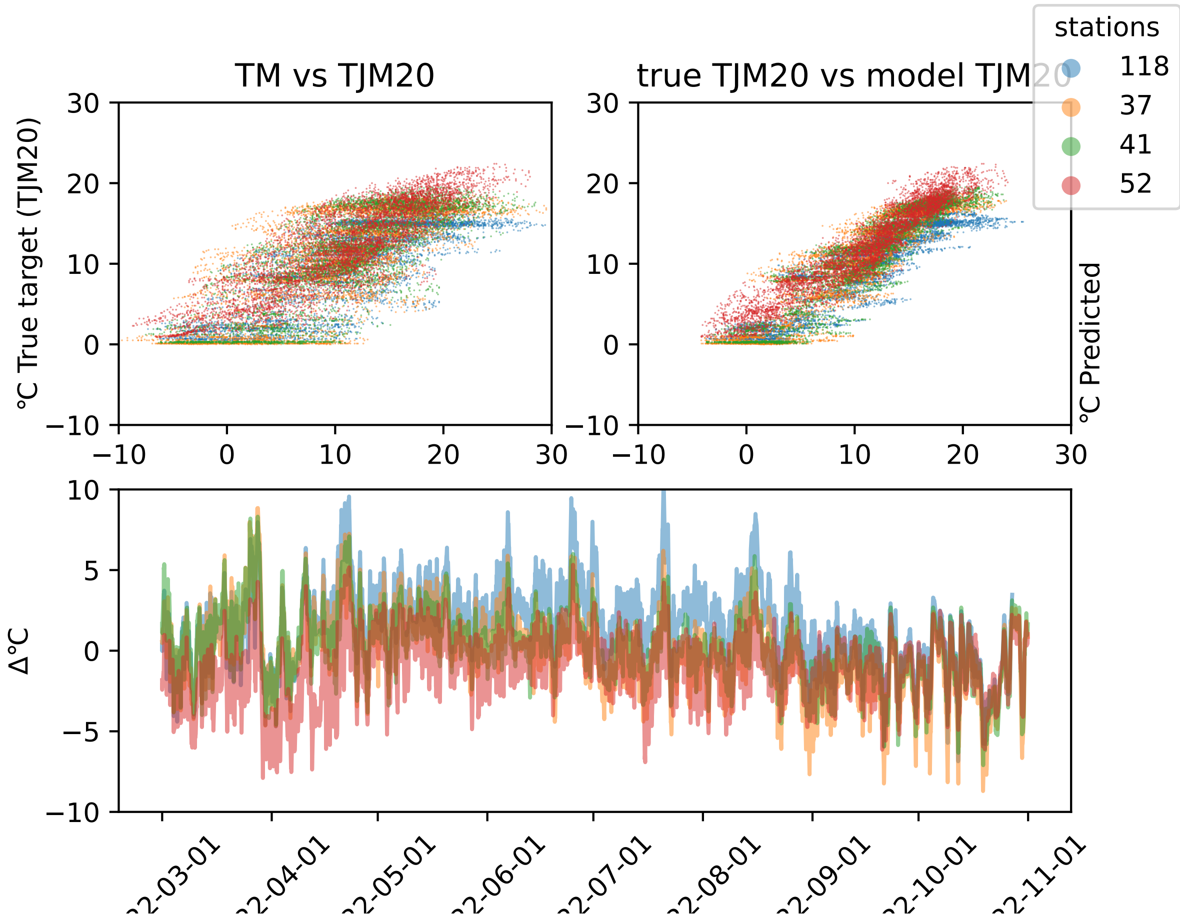


Figure 66: Difference plot for hourly Plauborg model in year 2022 and region Østfold

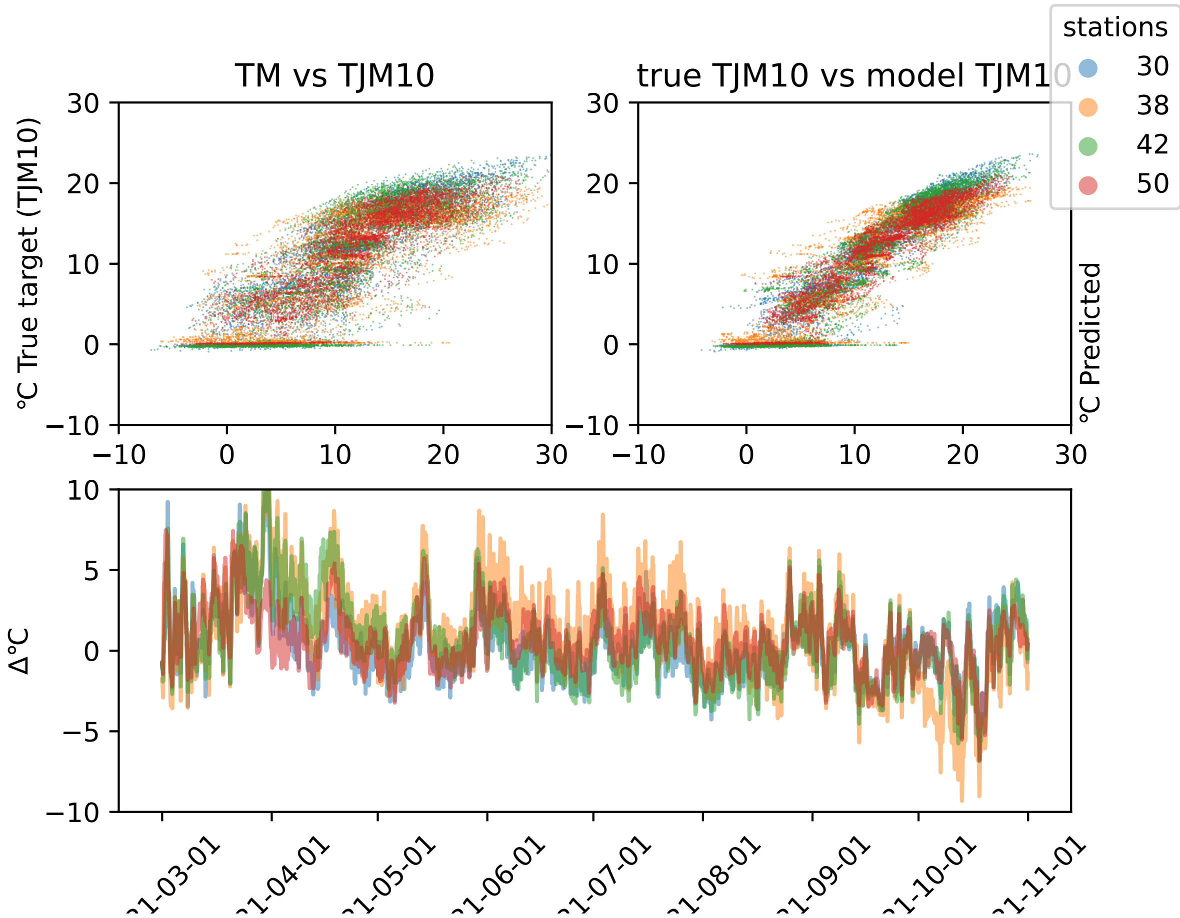


Figure 67: Difference plot for hourly Plauborg model in year 2021 and region Vestfold

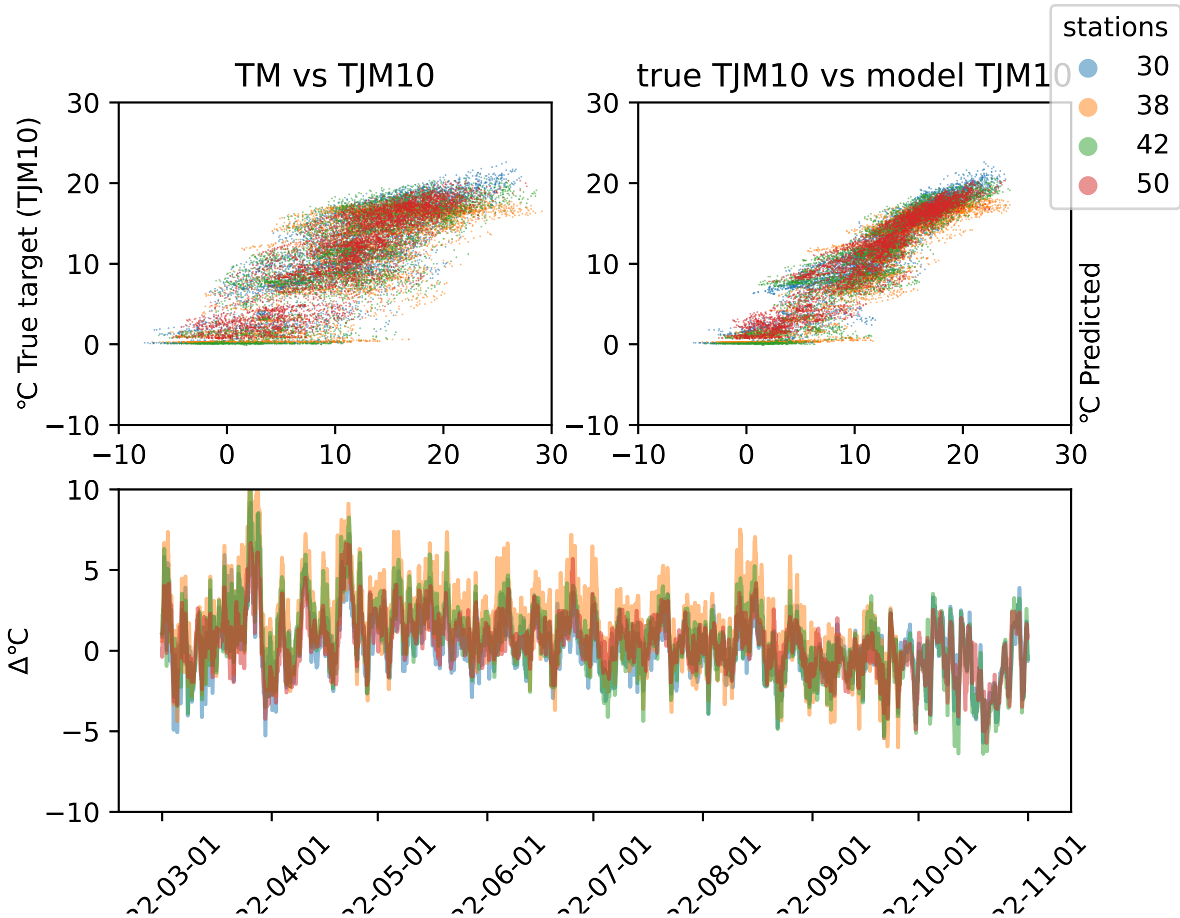


Figure 68: Difference plot for hourly Plauborg model in year 2022 and region Vestfold

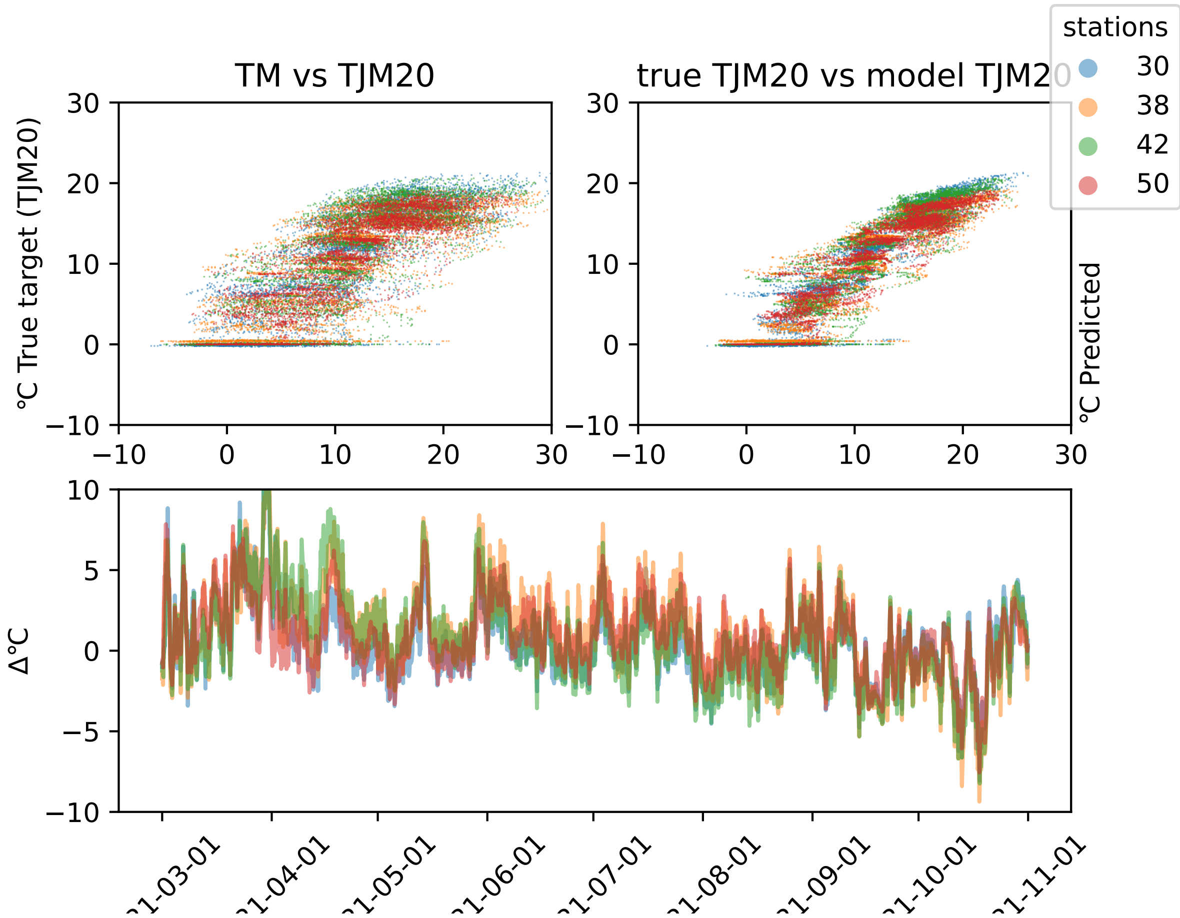


Figure 69: Difference plot for hourly Plauborg model in year 2021 and region Vestfold

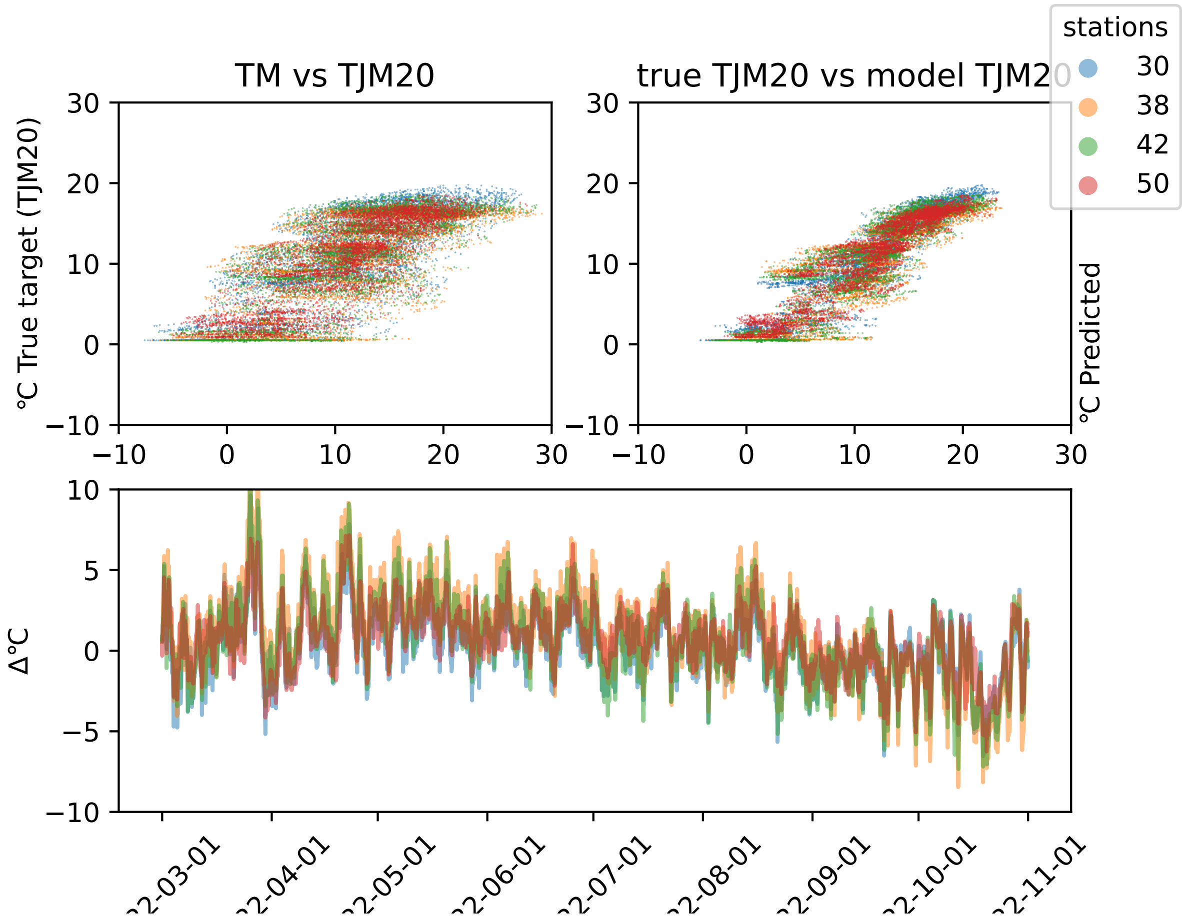


Figure 70: Difference plot for hourly Plauborg model in year 2022 and region Vestfold

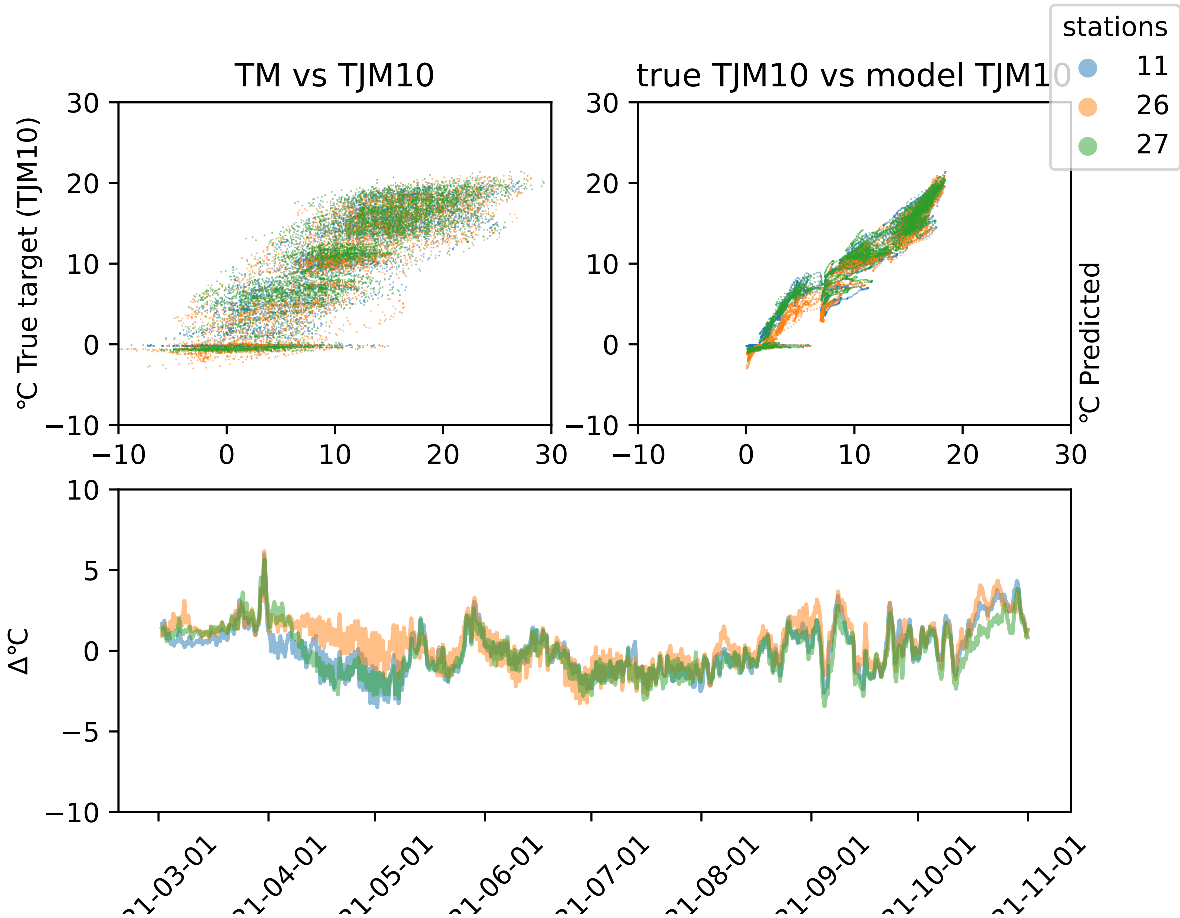


Figure 71: Difference plot for BiLSTM model in year 2021 and region Innlandet

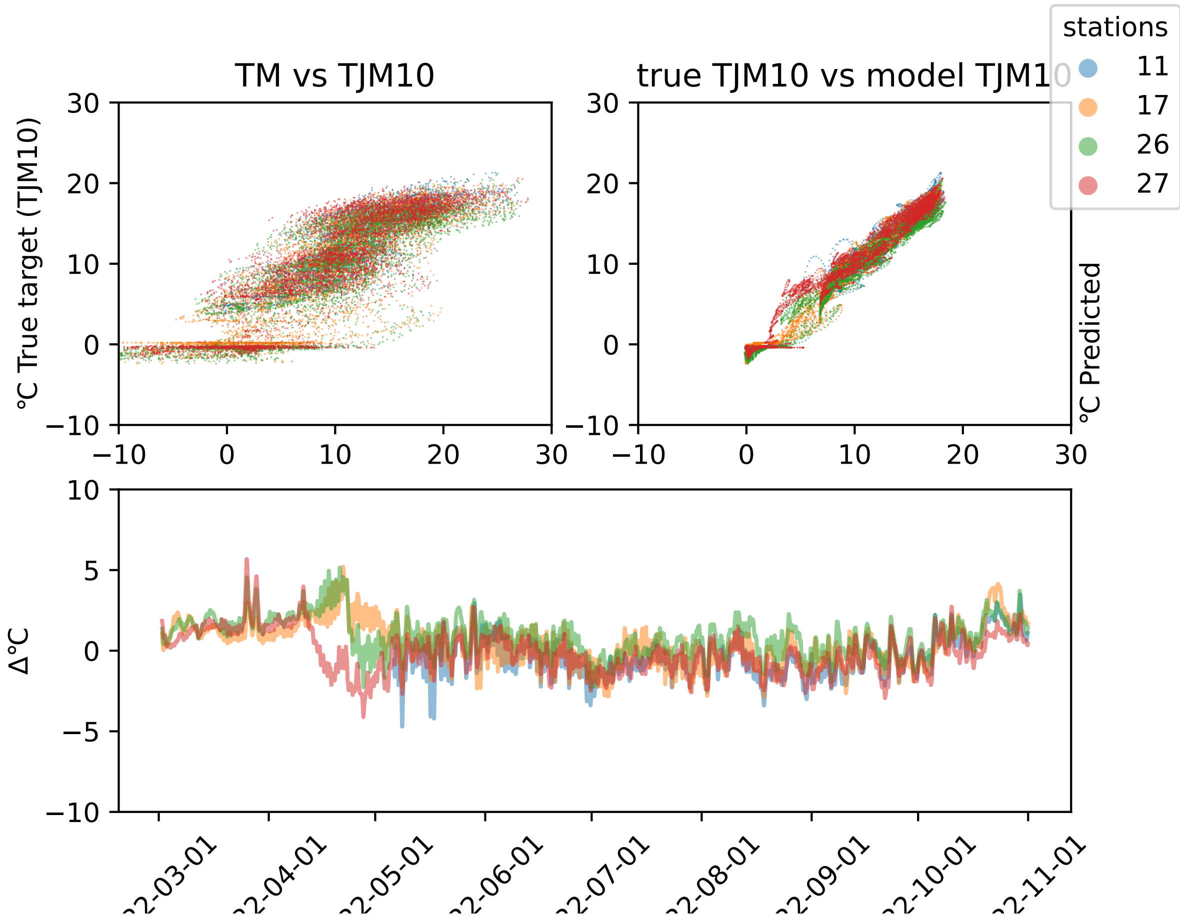


Figure 72: Difference plot for BiLSTM model in year 2022 and region Innlandet

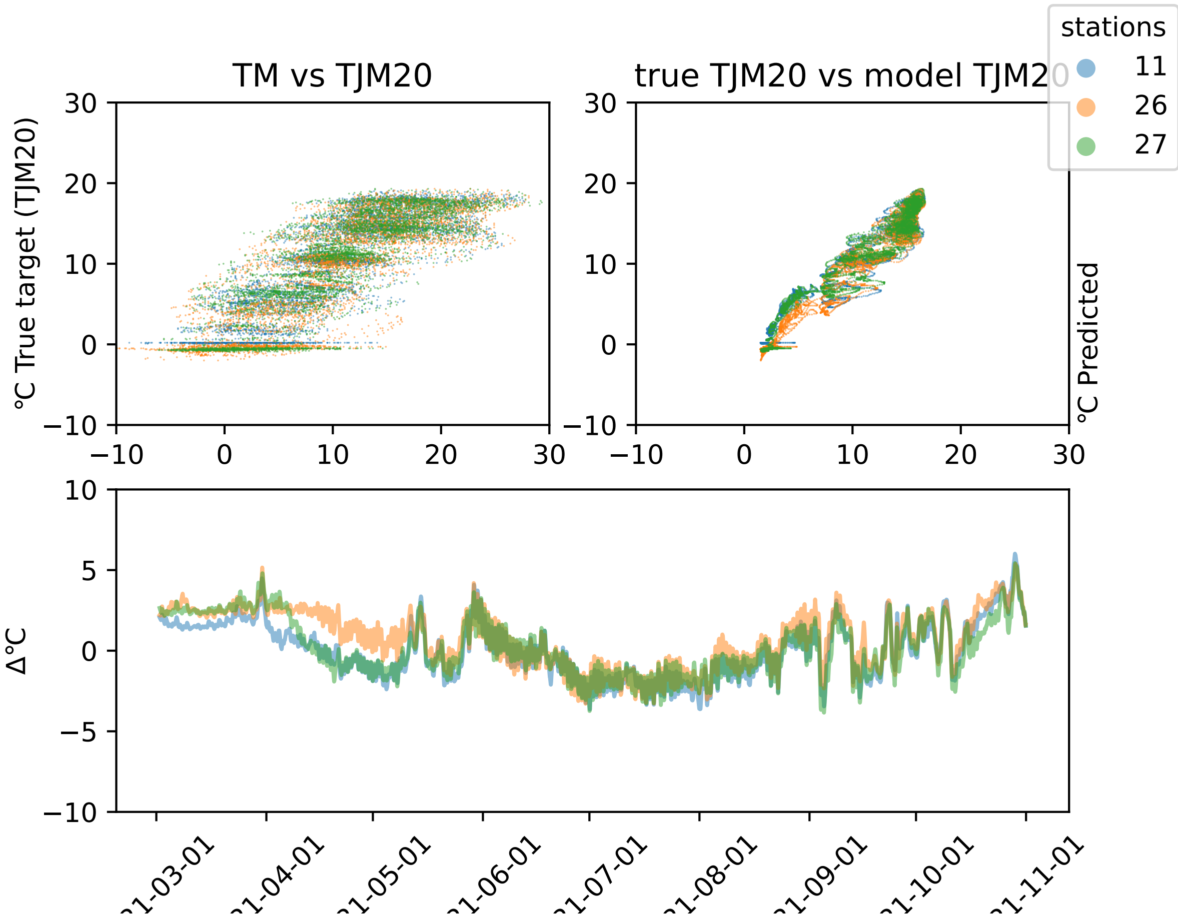


Figure 73: Difference plot for BiLSTM model in year 2021 and region Innlandet

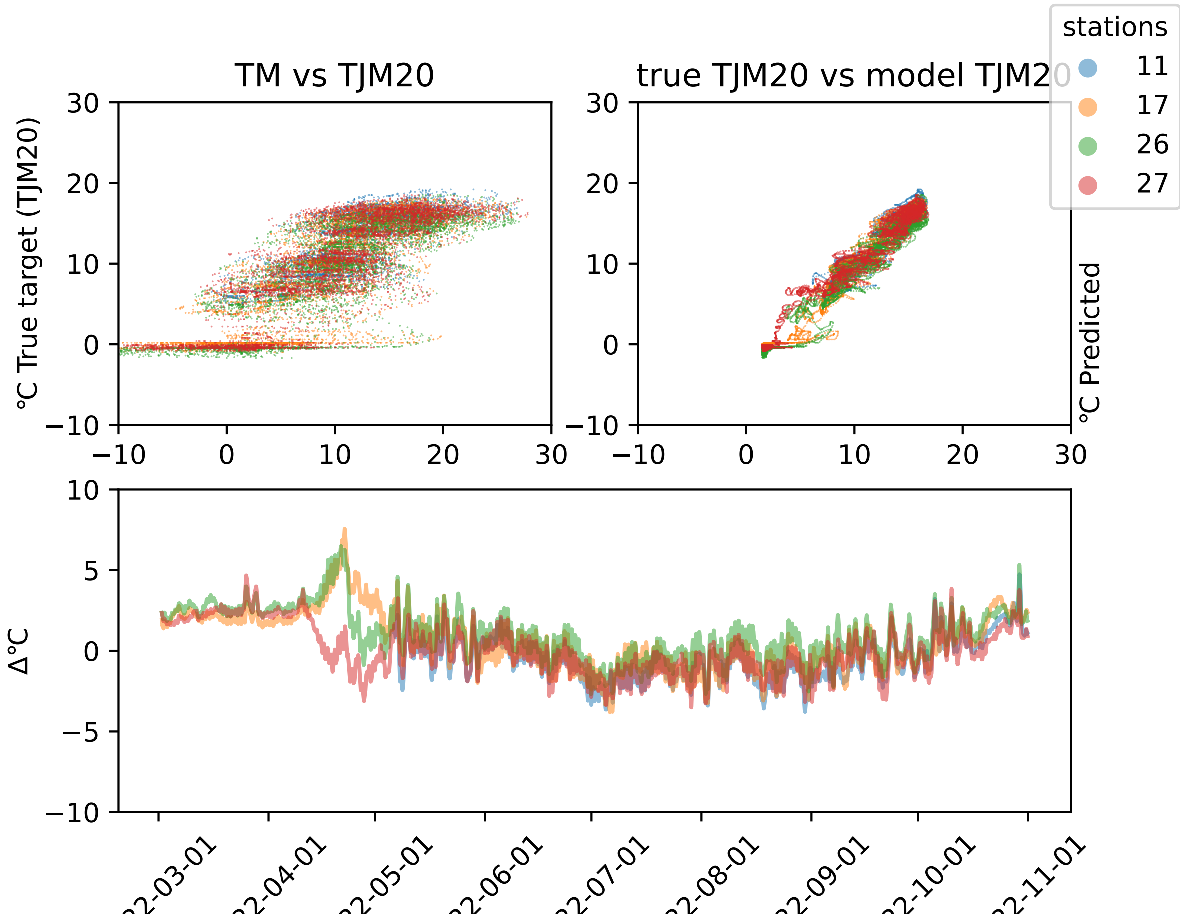


Figure 74: Difference plot for BiLSTM model in year 2022 and region Innlandet

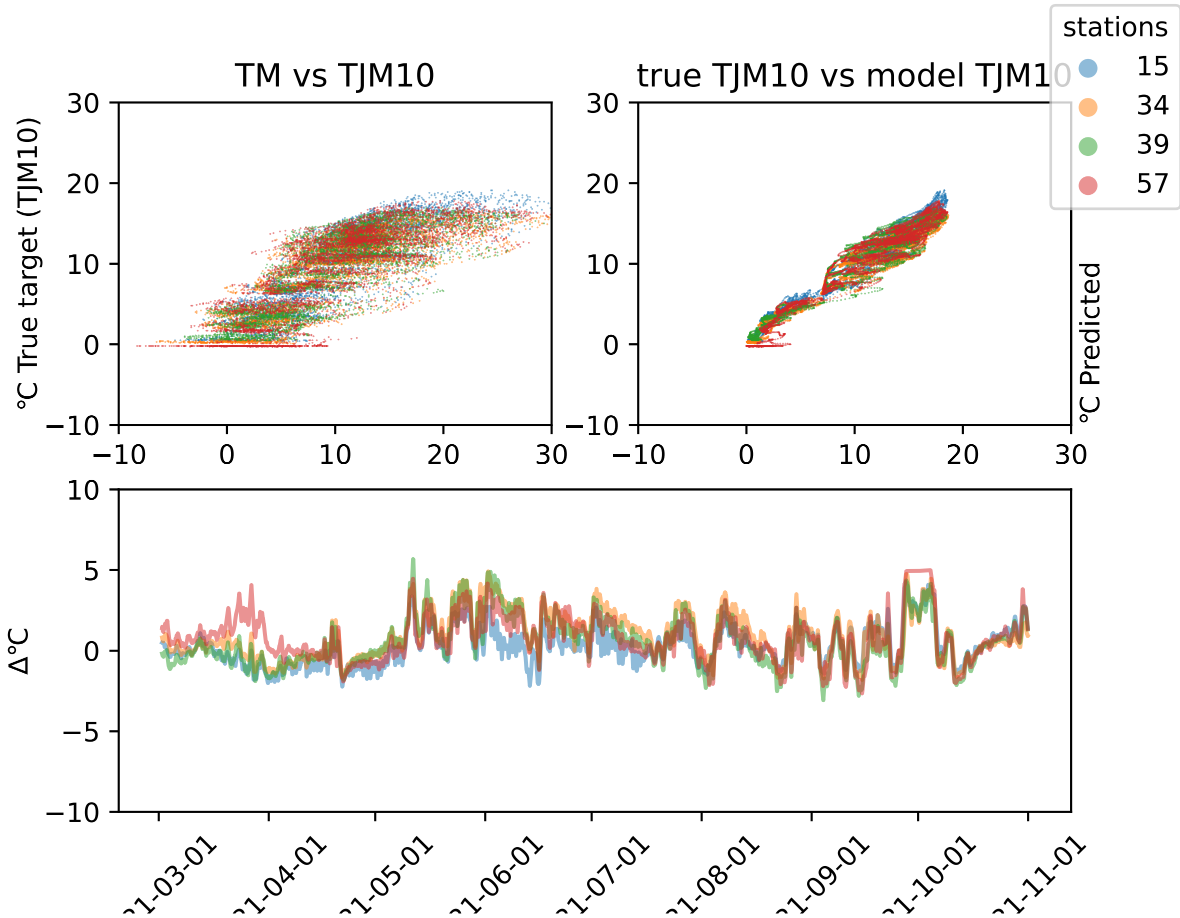


Figure 75: Difference plot for BiLSTM model in year 2021 and region Trøndelag

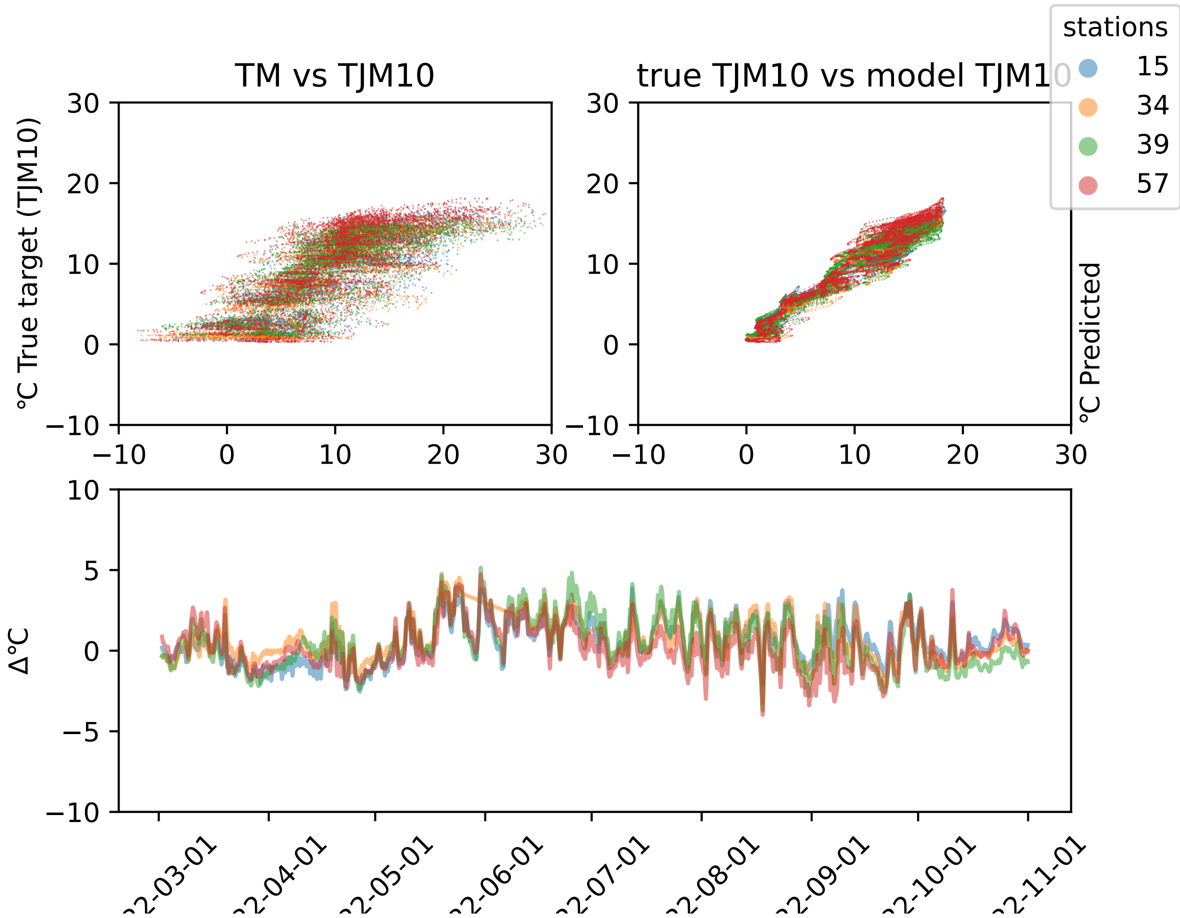


Figure 76: Difference plot for BiLSTM model in year 2022 and region Trøndelag

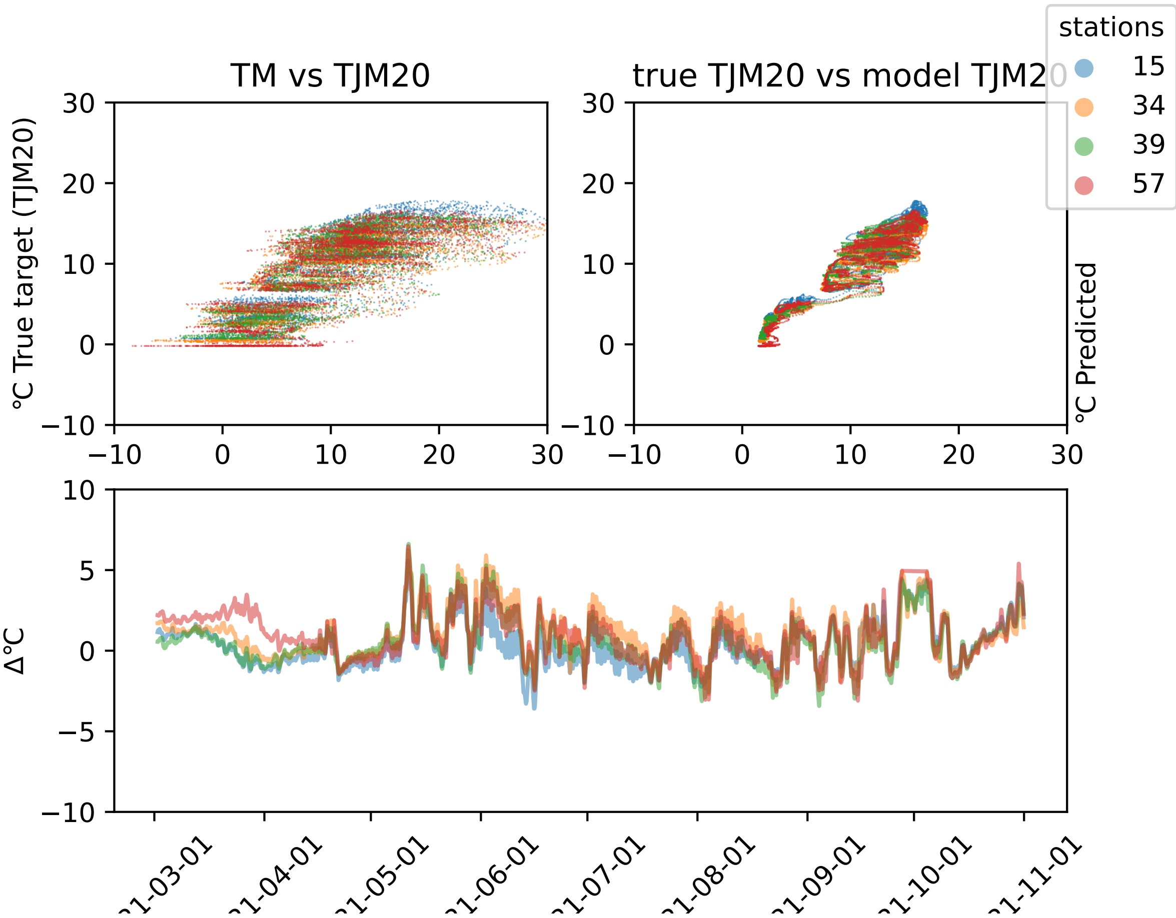


Figure 77: Difference plot for BiLSTM model in year 2021 and region Trøndelag

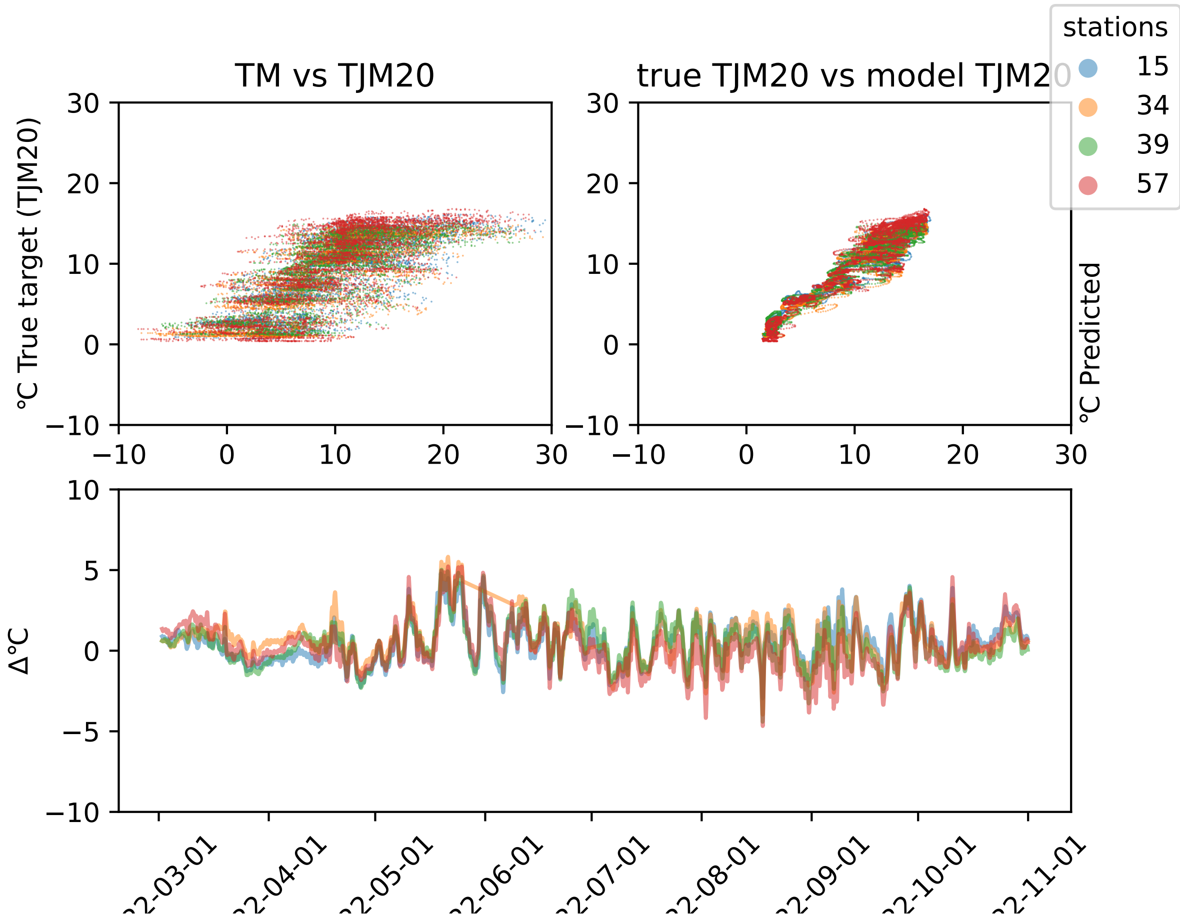


Figure 78: Difference plot for BiLSTM model in year 2022 and region Trøndelag

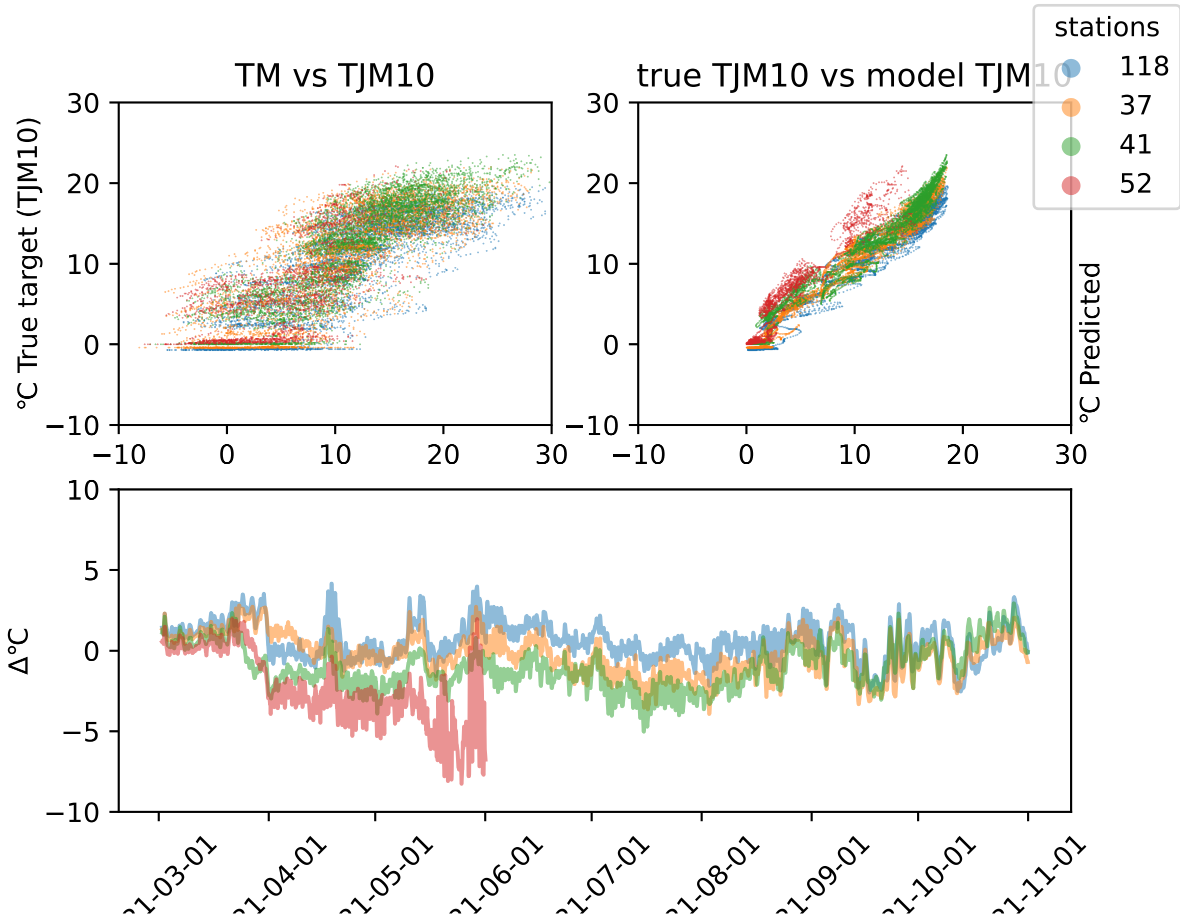


Figure 79: Difference plot for BiLSTM model in year 2021 and region Østfold

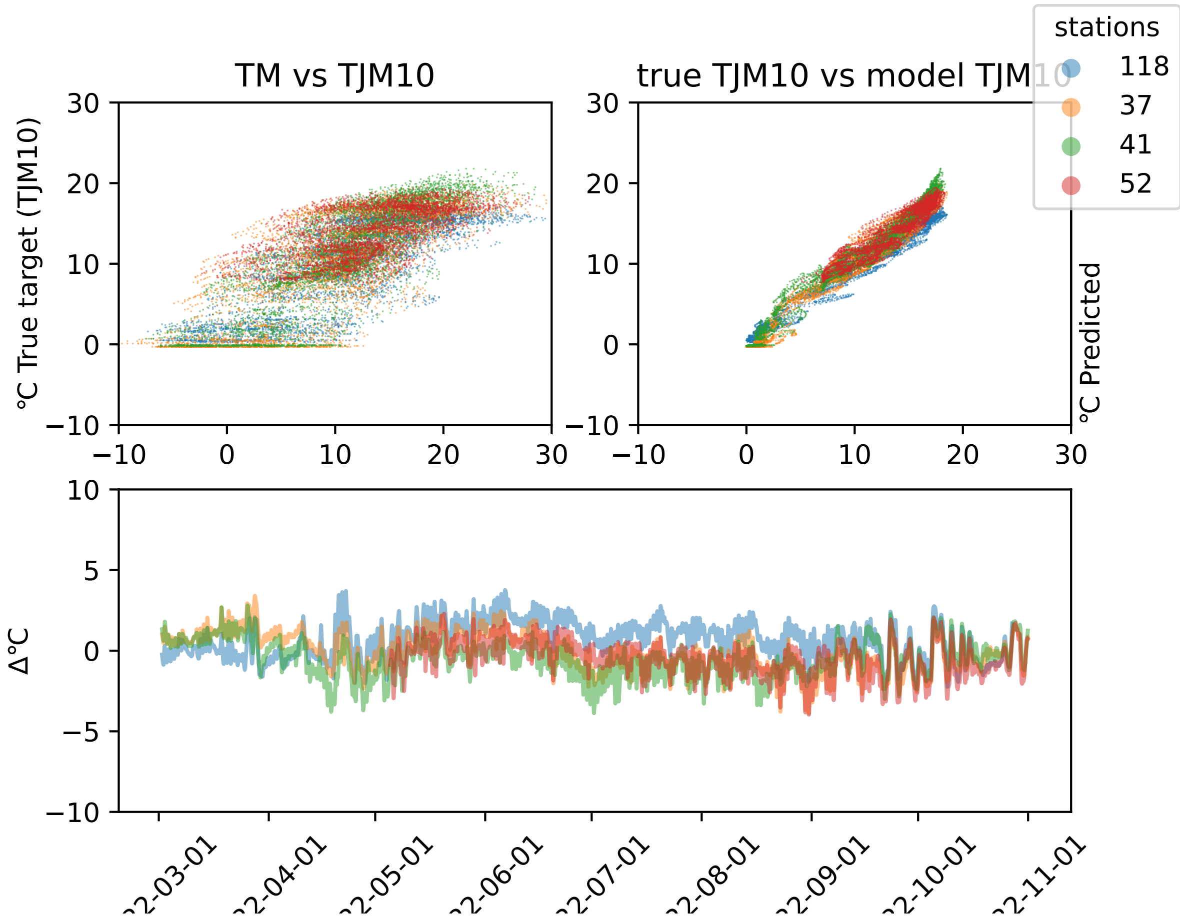


Figure 80: Difference plot for BiLSTM model in year 2022 and region Østfold

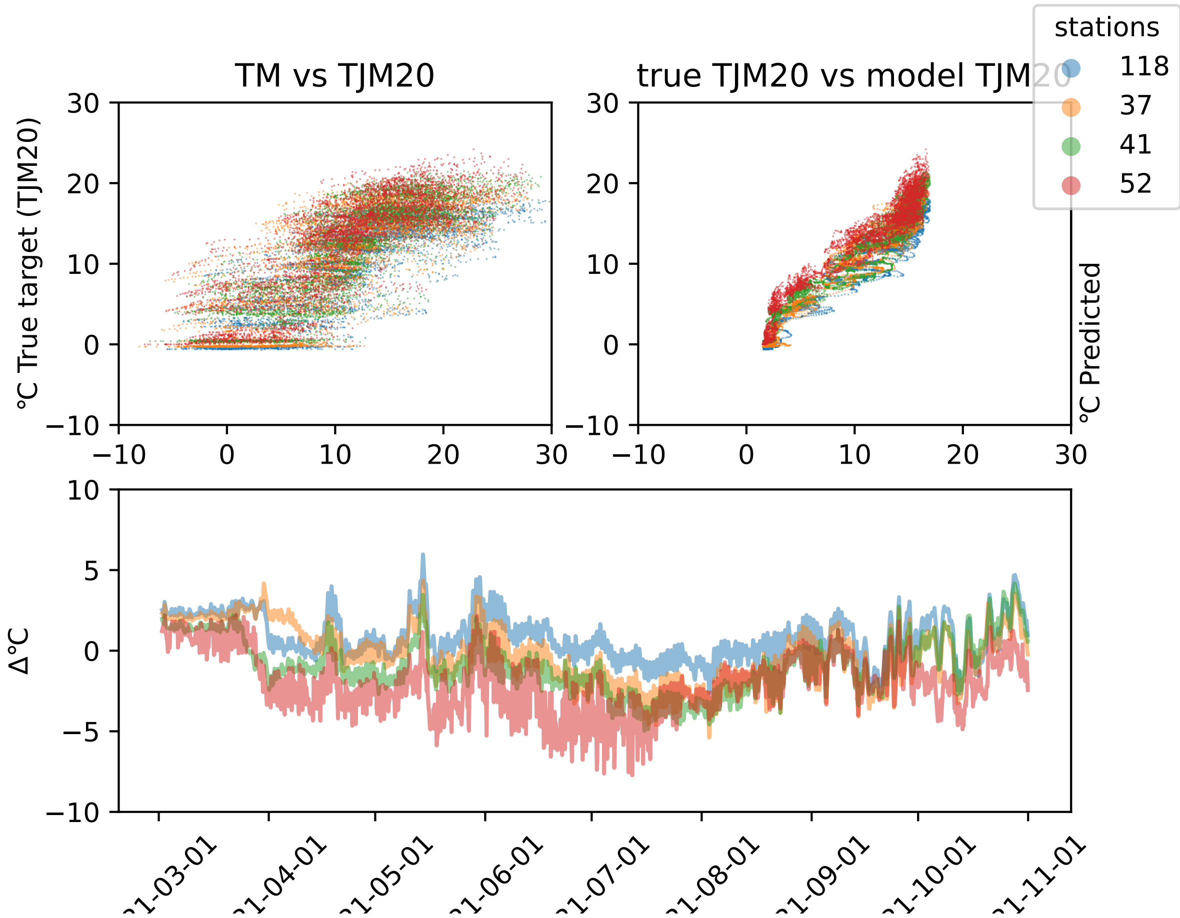


Figure 81: Difference plot for BiLSTM model in year 2021 and region Østfold

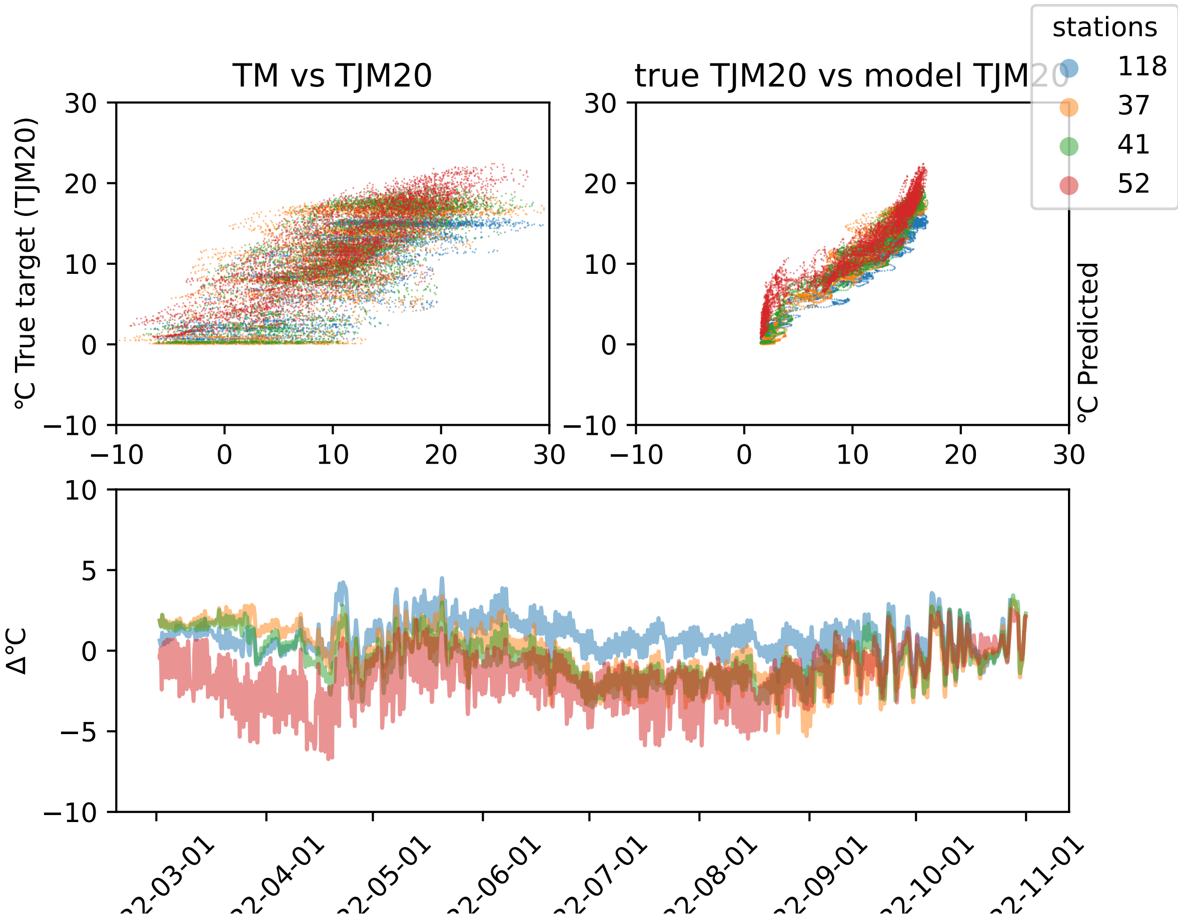


Figure 82: Difference plot for BiLSTM model in year 2022 and region Østfold

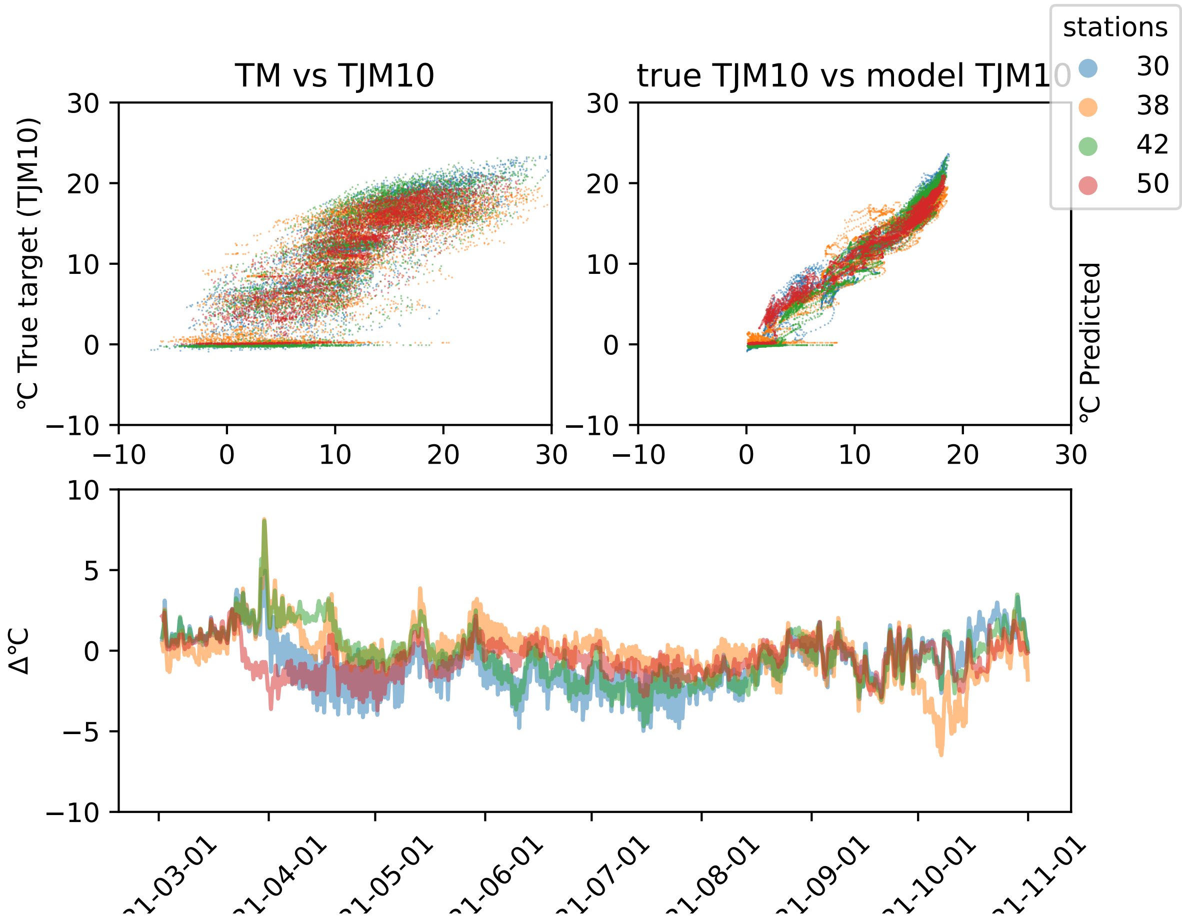


Figure 83: Difference plot for BiLSTM model in year 2021 and region Vestfold

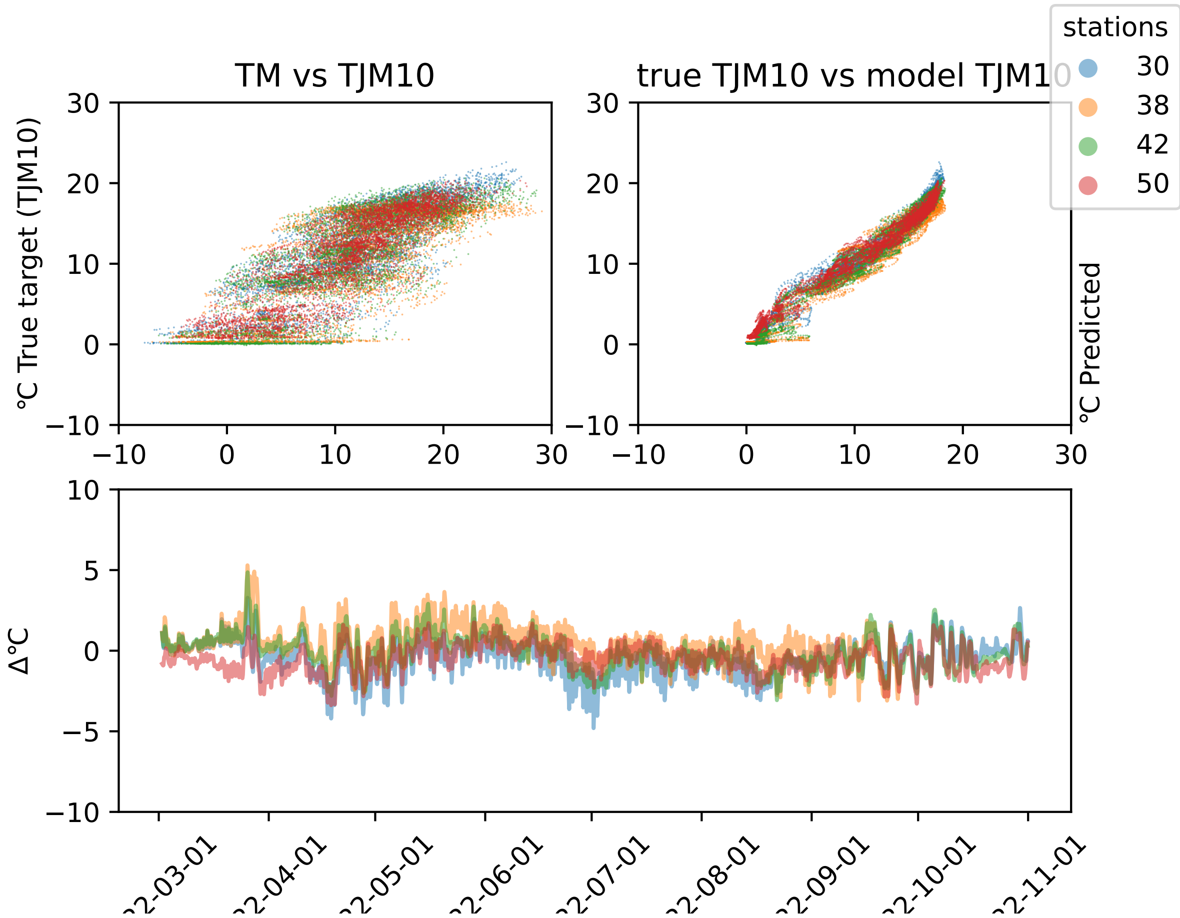


Figure 84: Difference plot for BiLSTM model in year 2022 and region Vestfold

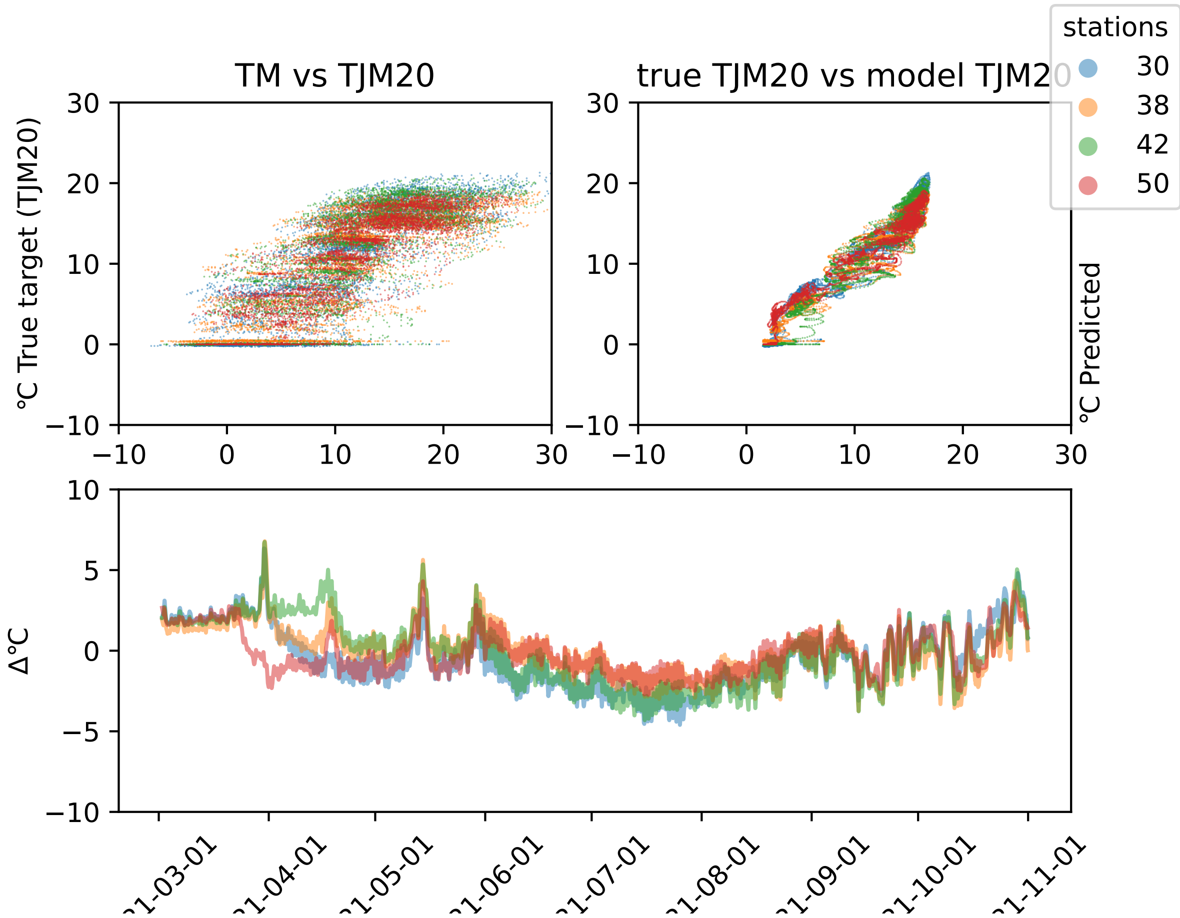


Figure 85: Difference plot for BiLSTM model in year 2021 and region Vestfold

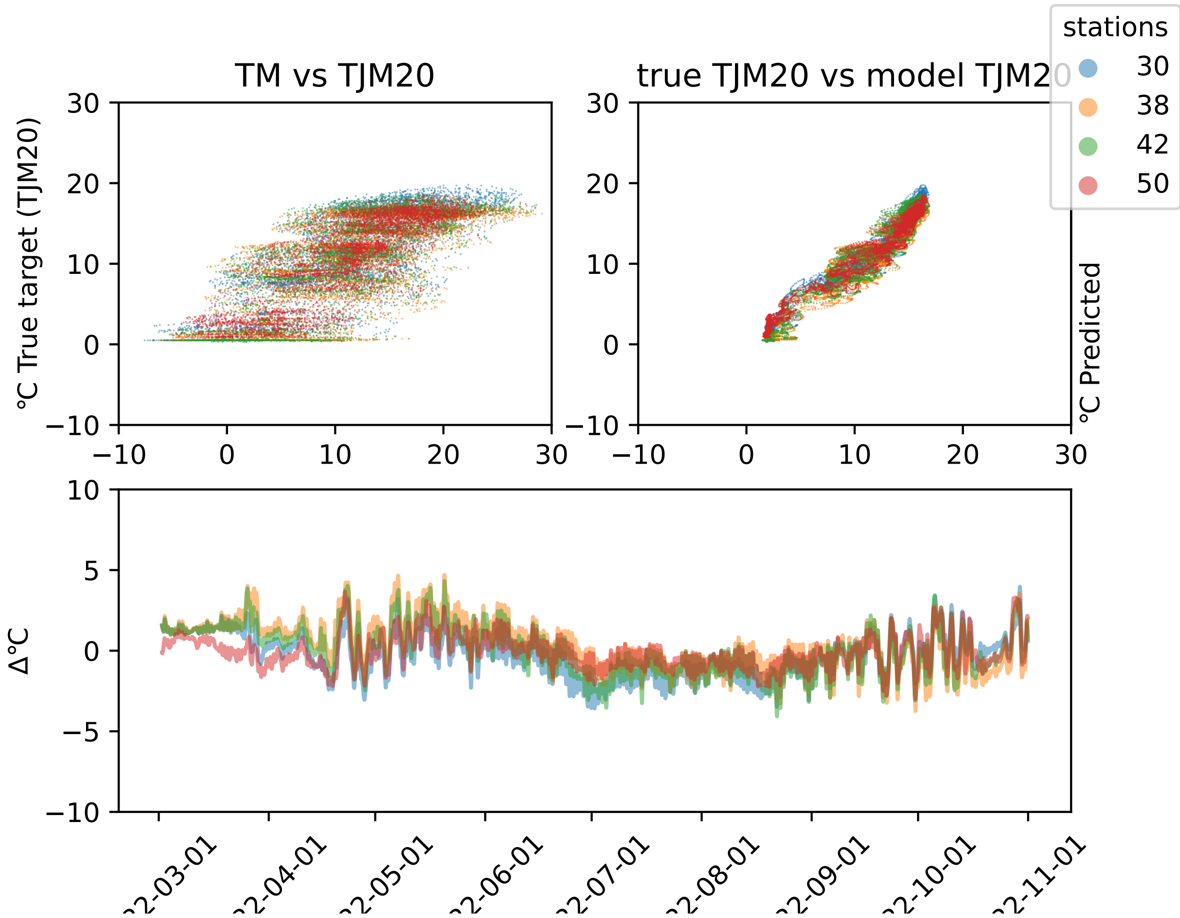


Figure 86: Difference plot for BiLSTM model in year 2022 and region Vestfold

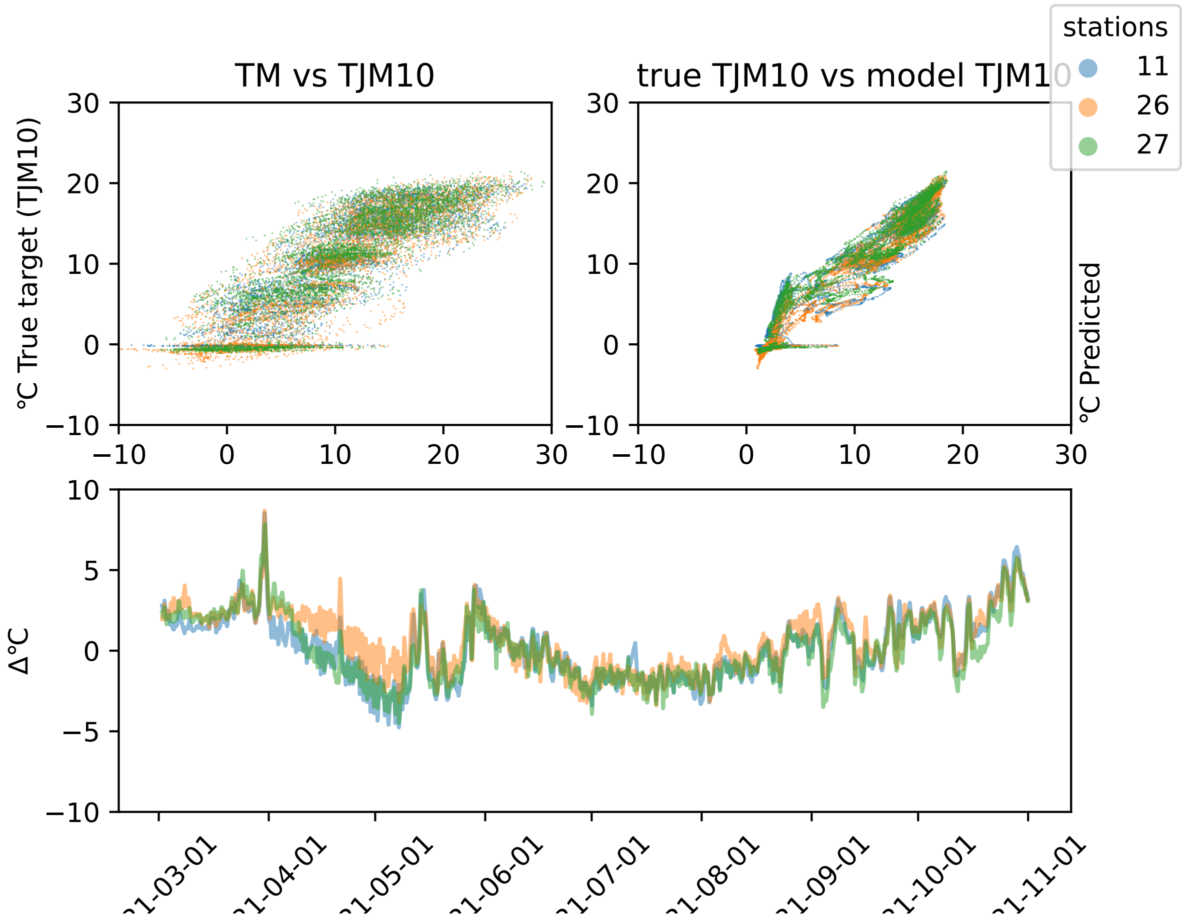


Figure 87: Difference plot for LSTM model in year 2021 and region Innlandet

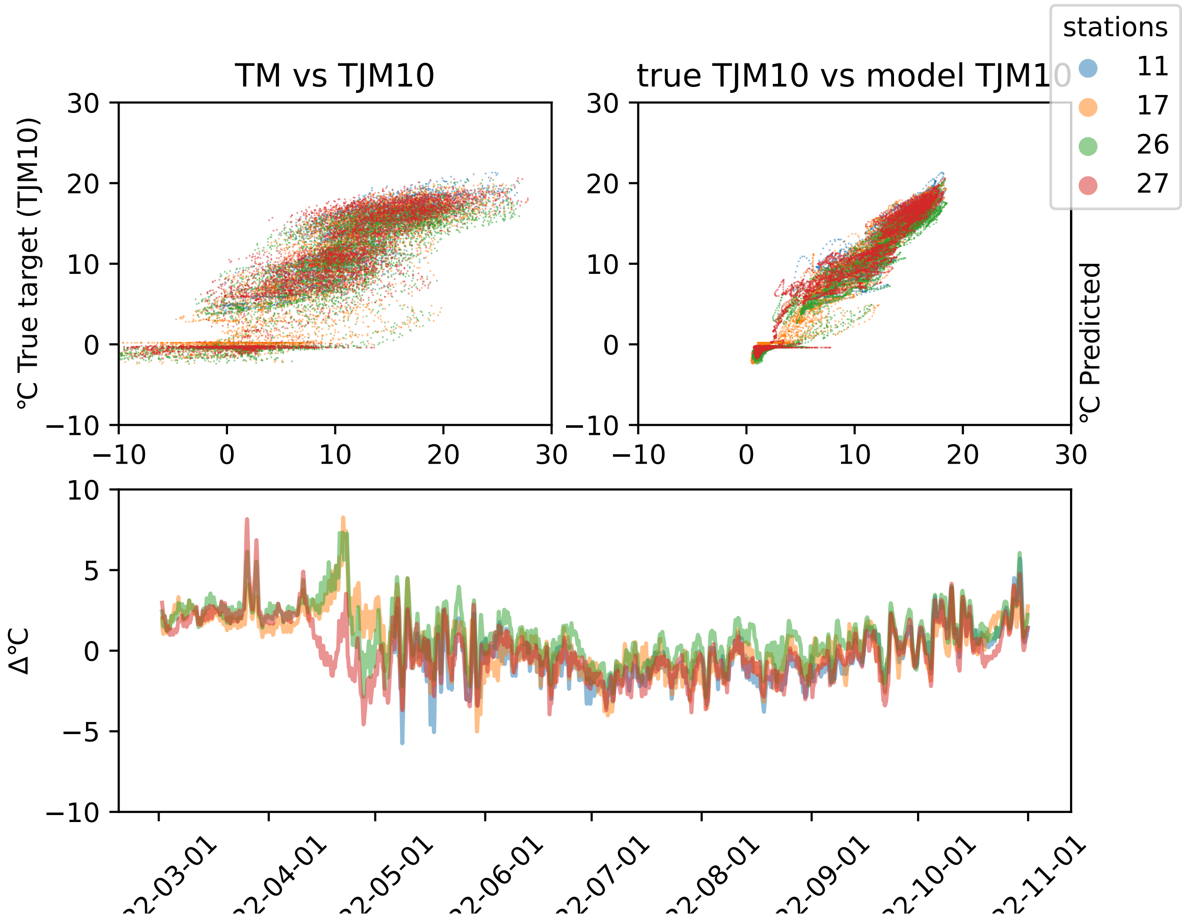


Figure 88: Difference plot for LSTM model in year 2022 and region Innlandet

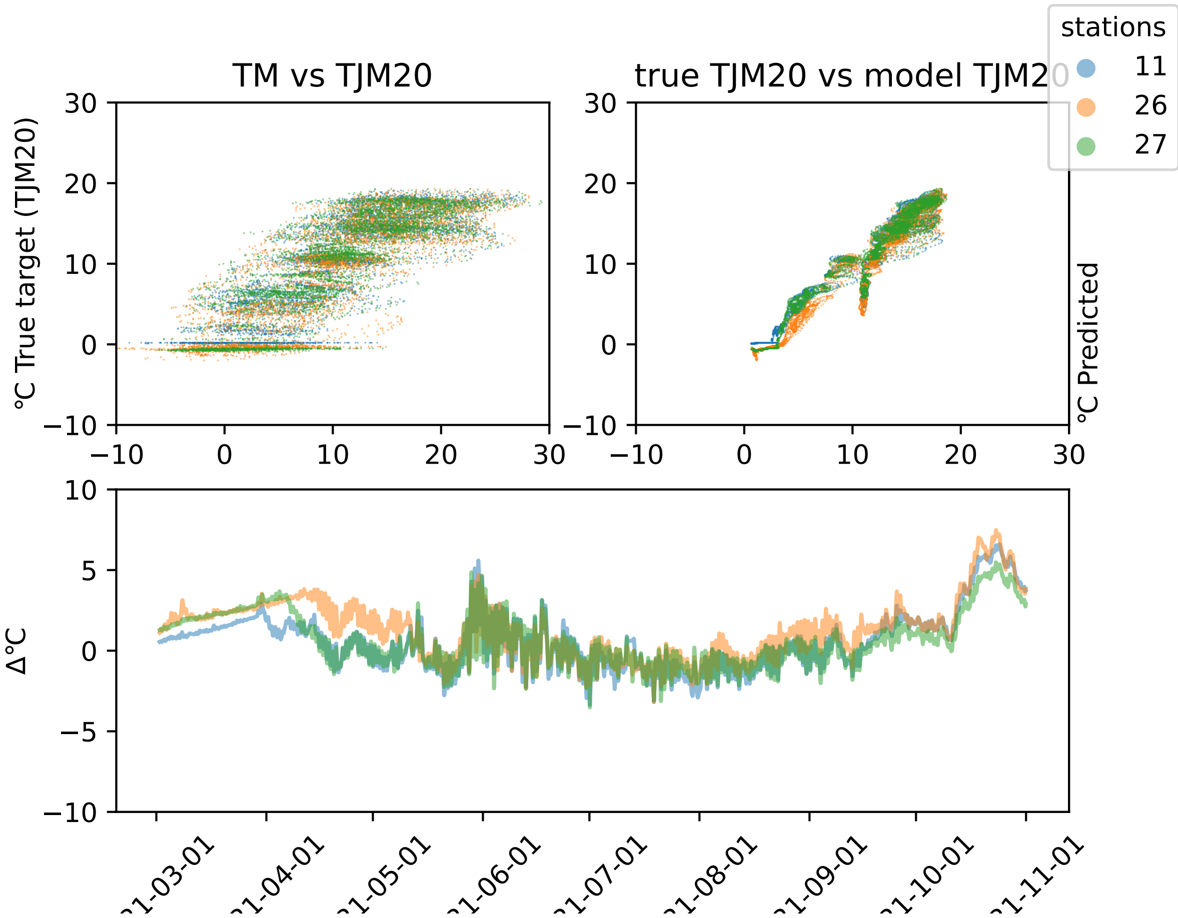


Figure 89: Difference plot for LSTM model in year 2021 and region Innlandet

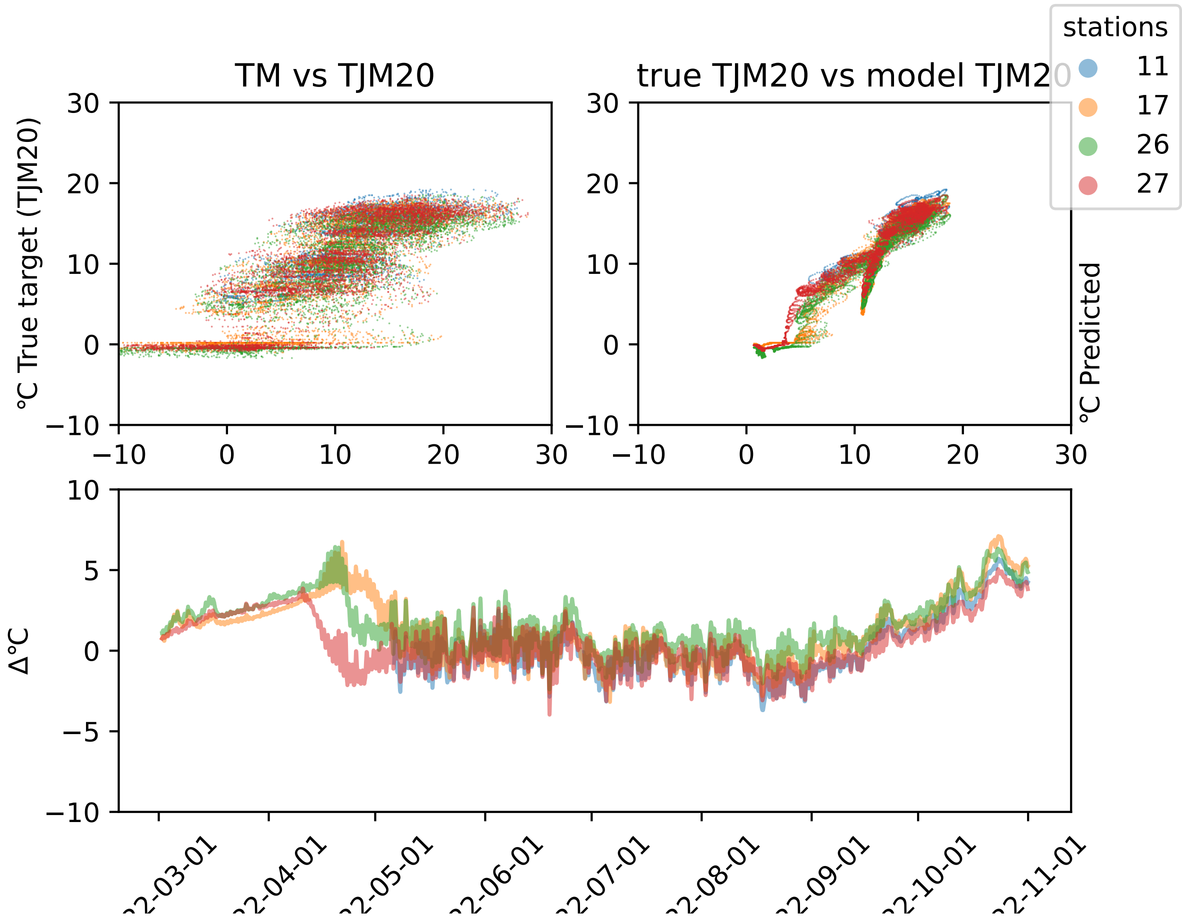


Figure 90: Difference plot for LSTM model in year 2022 and region Innlandet

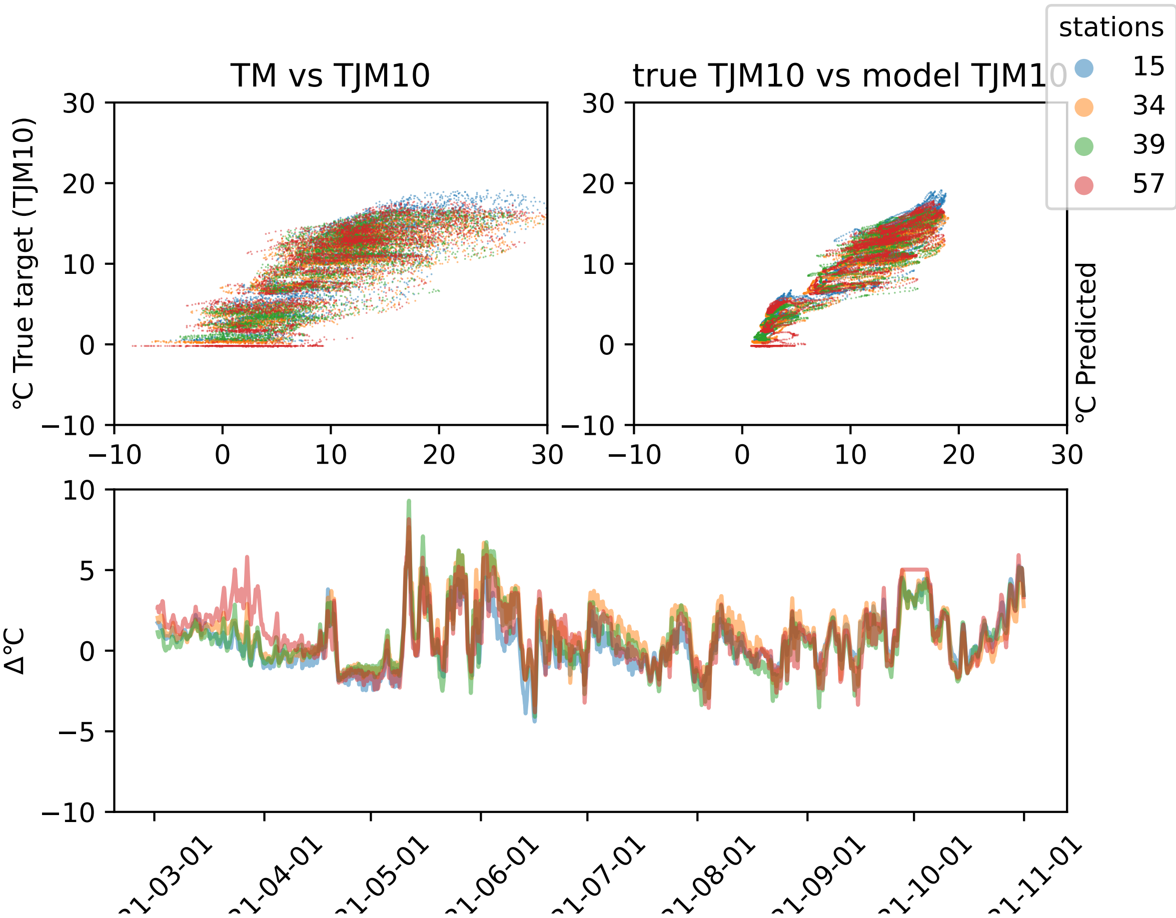


Figure 91: Difference plot for LSTM model in year 2021 and region Trøndelag

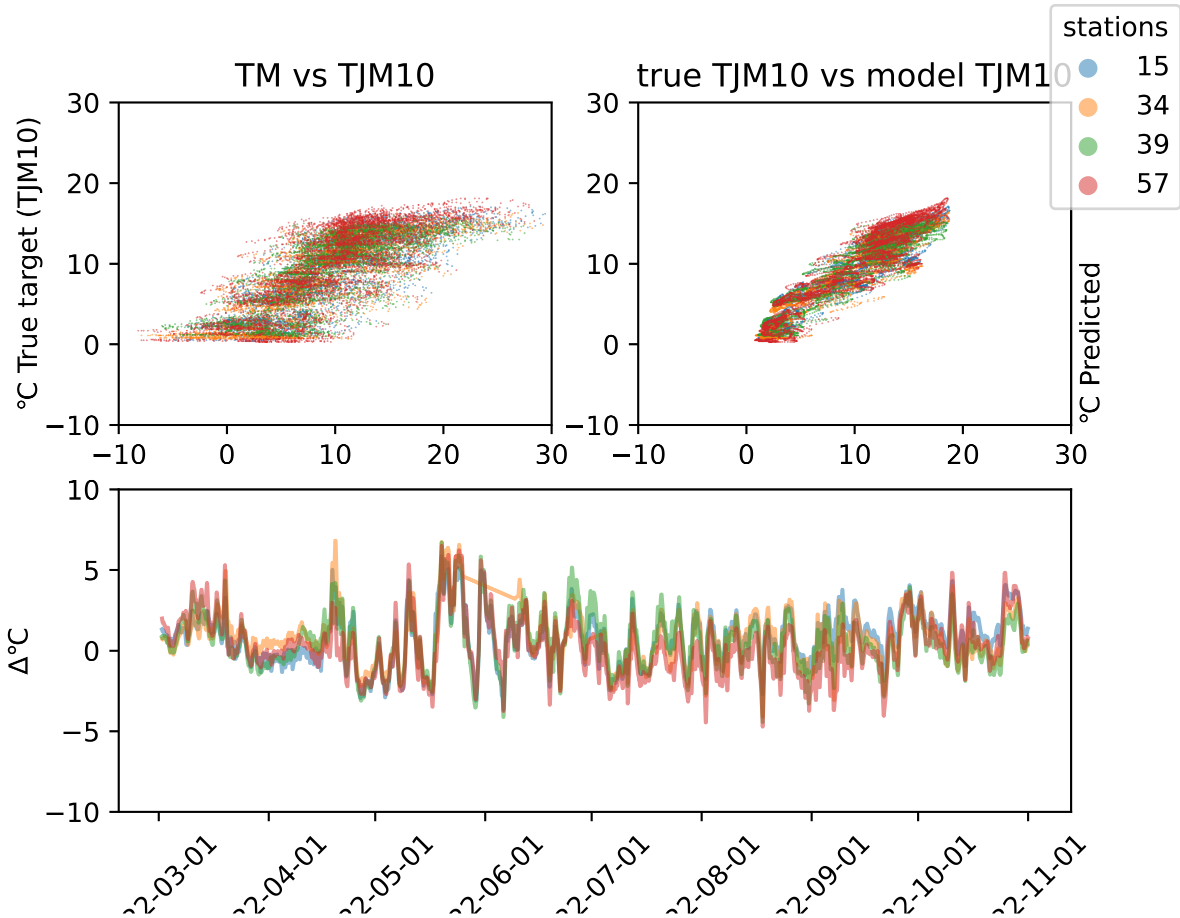


Figure 92: Difference plot for LSTM model in year 2022 and region Trøndelag

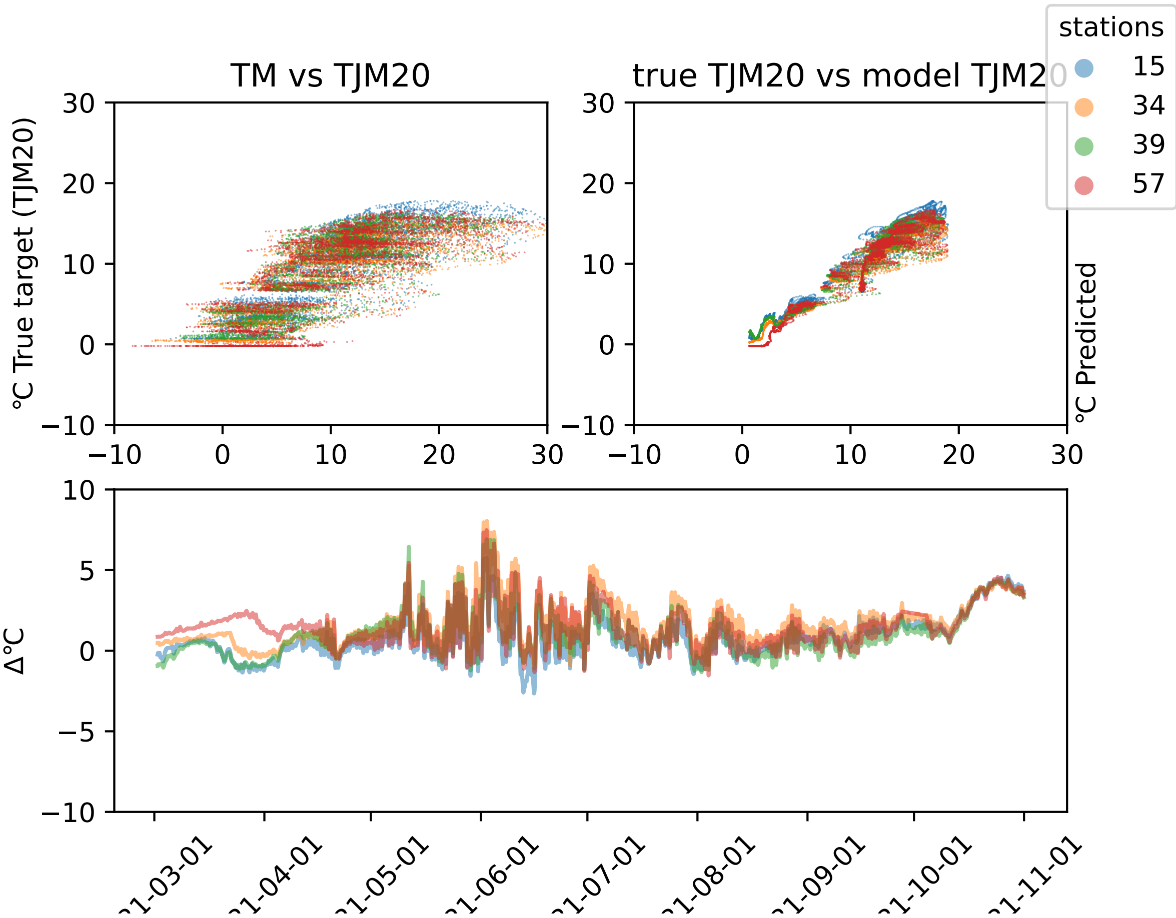


Figure 93: Difference plot for LSTM model in year 2021 and region Trøndelag

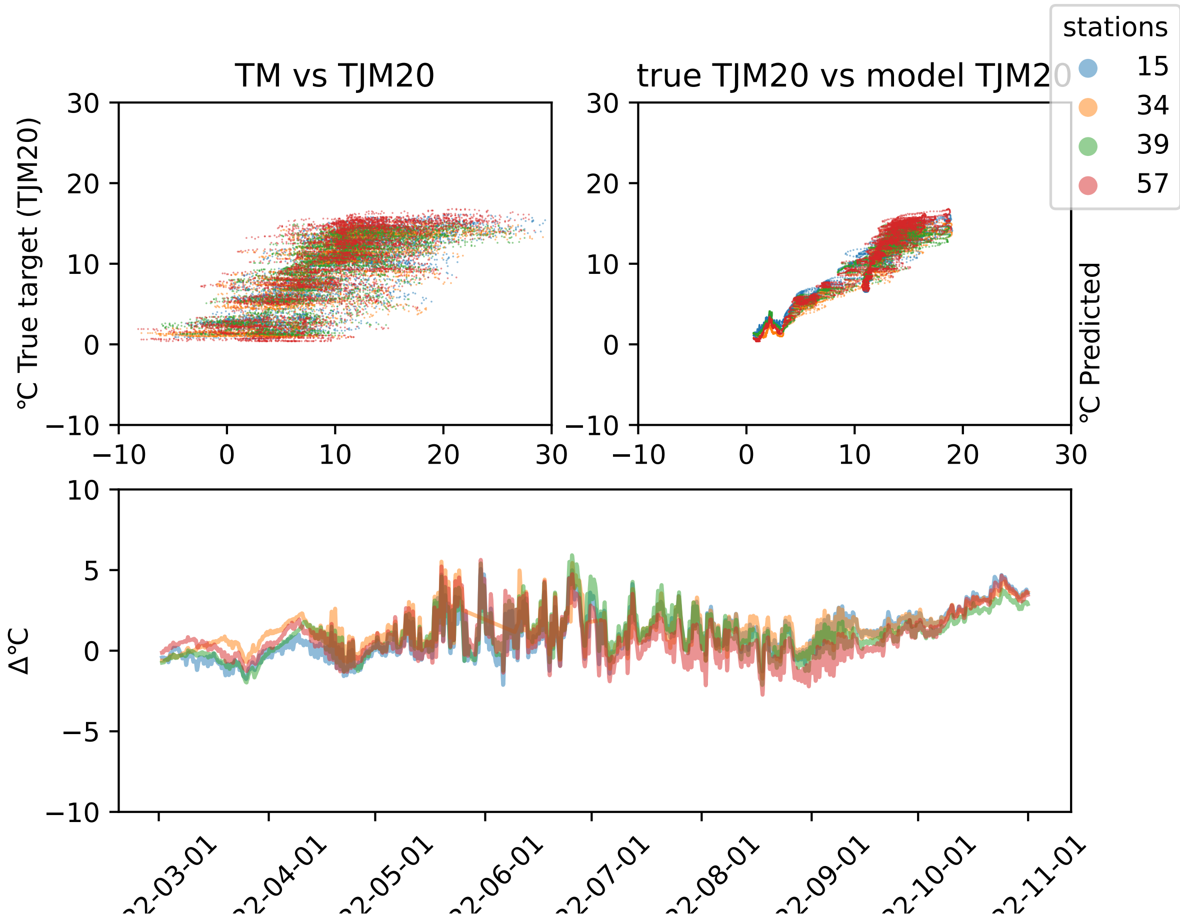


Figure 94: Difference plot for LSTM model in year 2022 and region Trøndelag

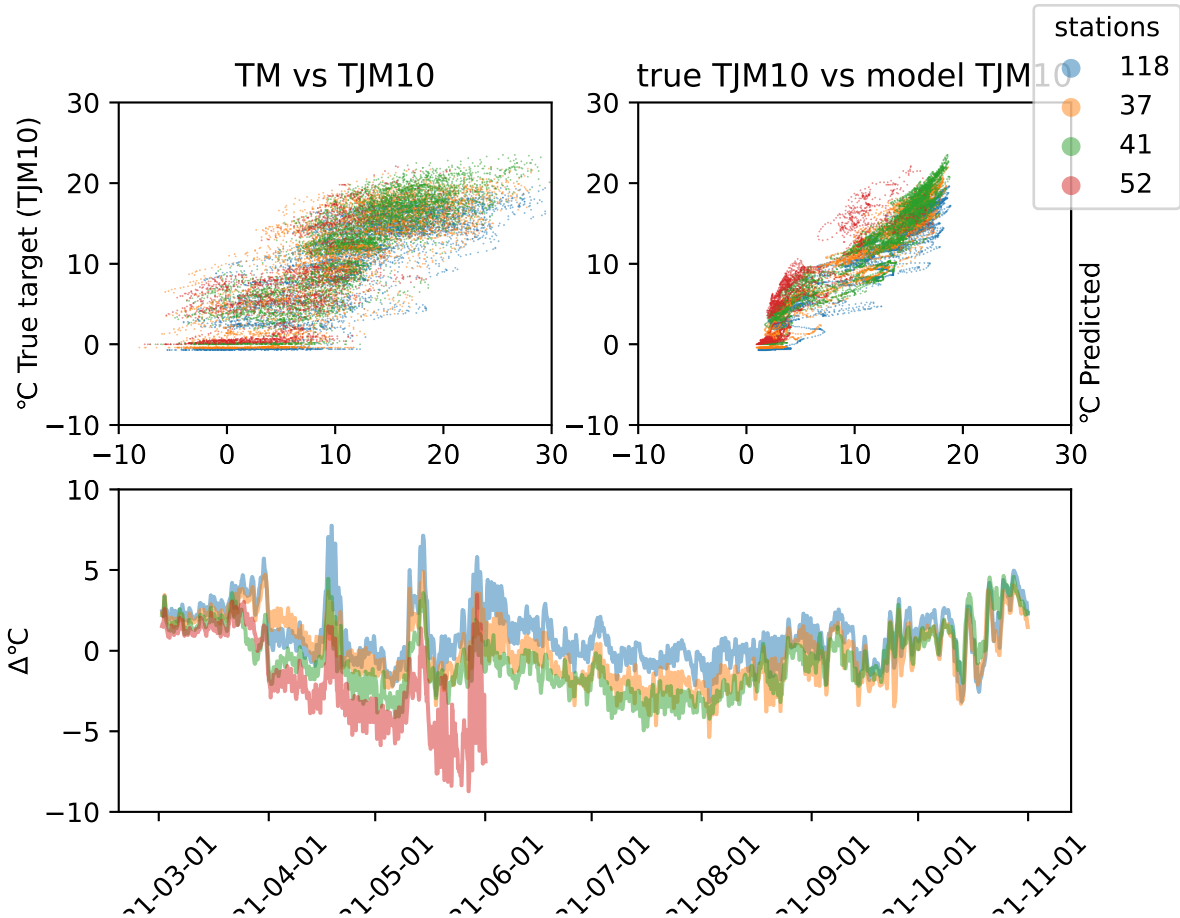


Figure 95: Difference plot for LSTM model in year 2021 and region Østfold

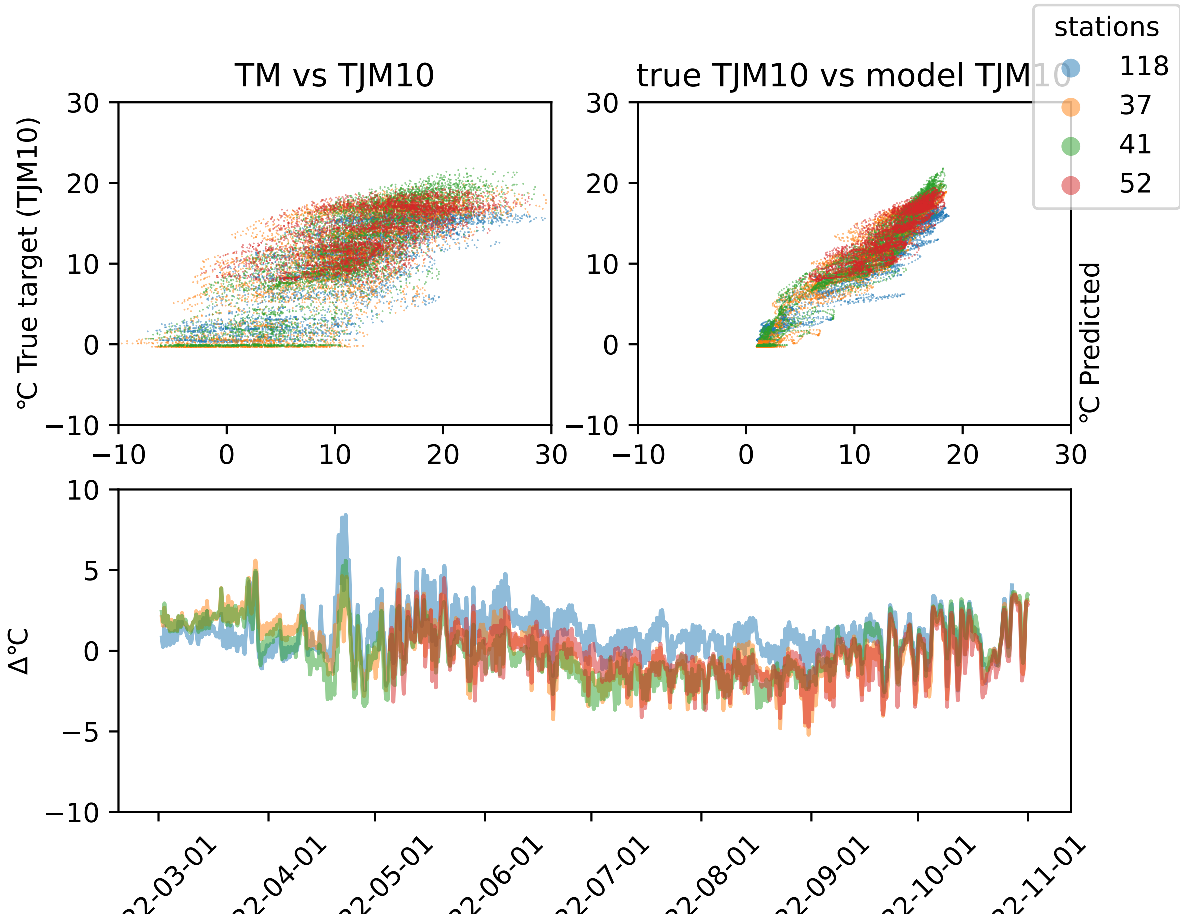


Figure 96: Difference plot for LSTM model in year 2022 and region Østfold

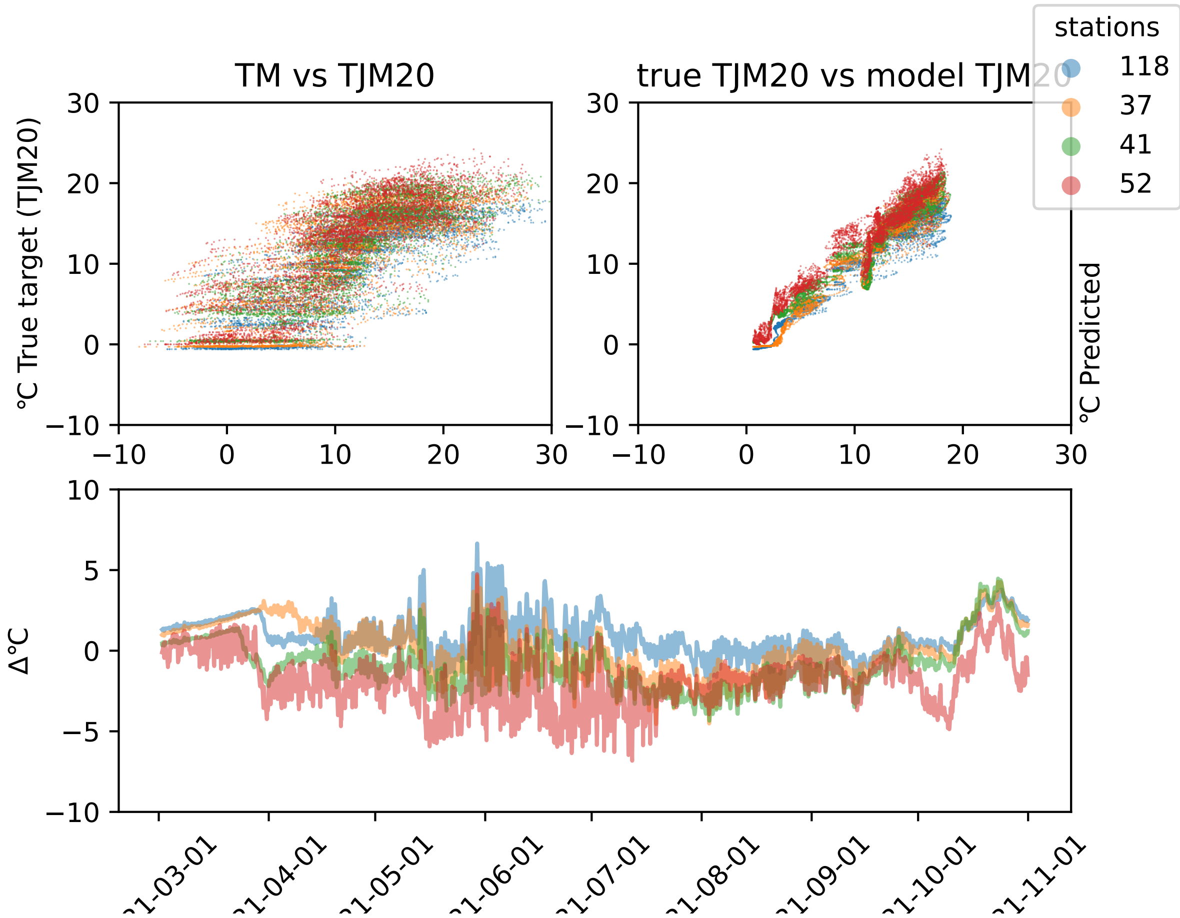


Figure 97: Difference plot for LSTM model in year 2021 and region Østfold

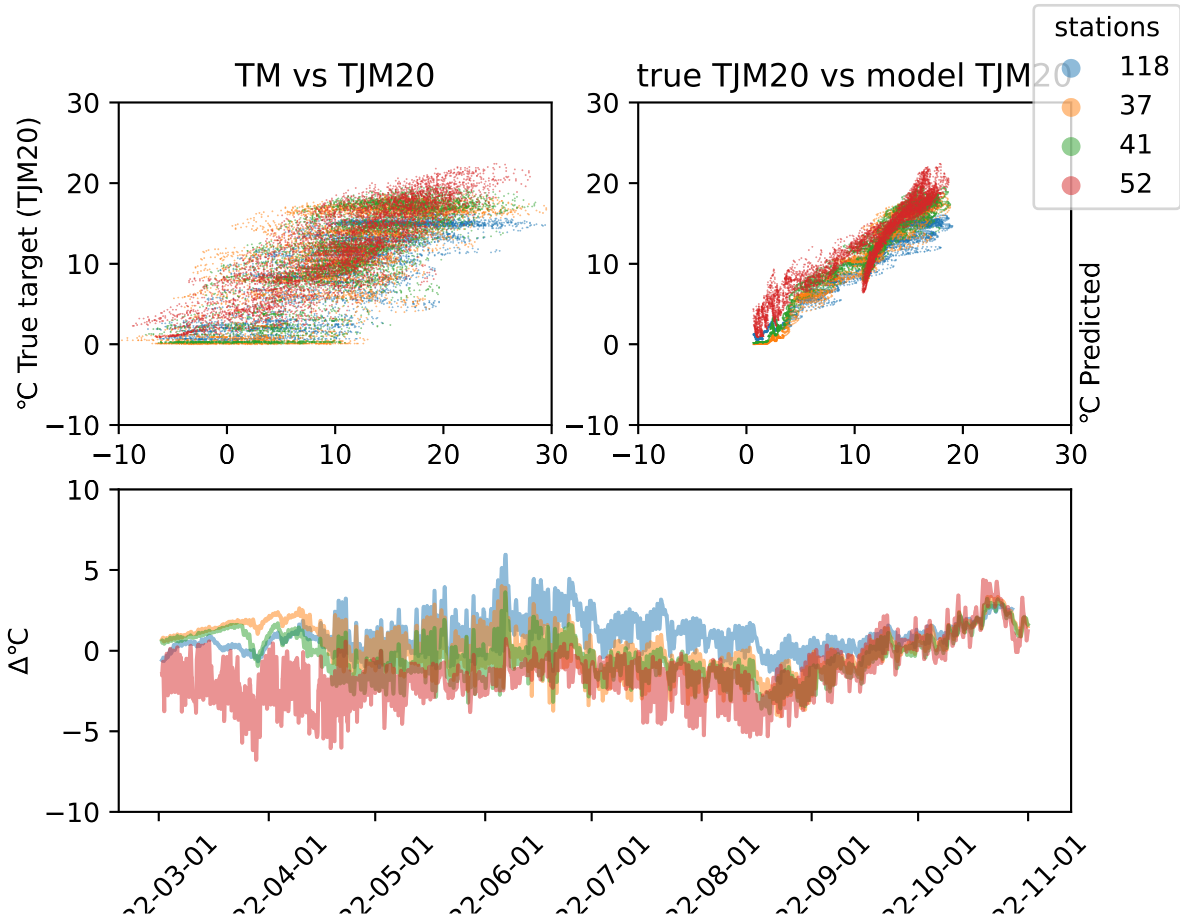


Figure 98: Difference plot for LSTM model in year 2022 and region Østfold

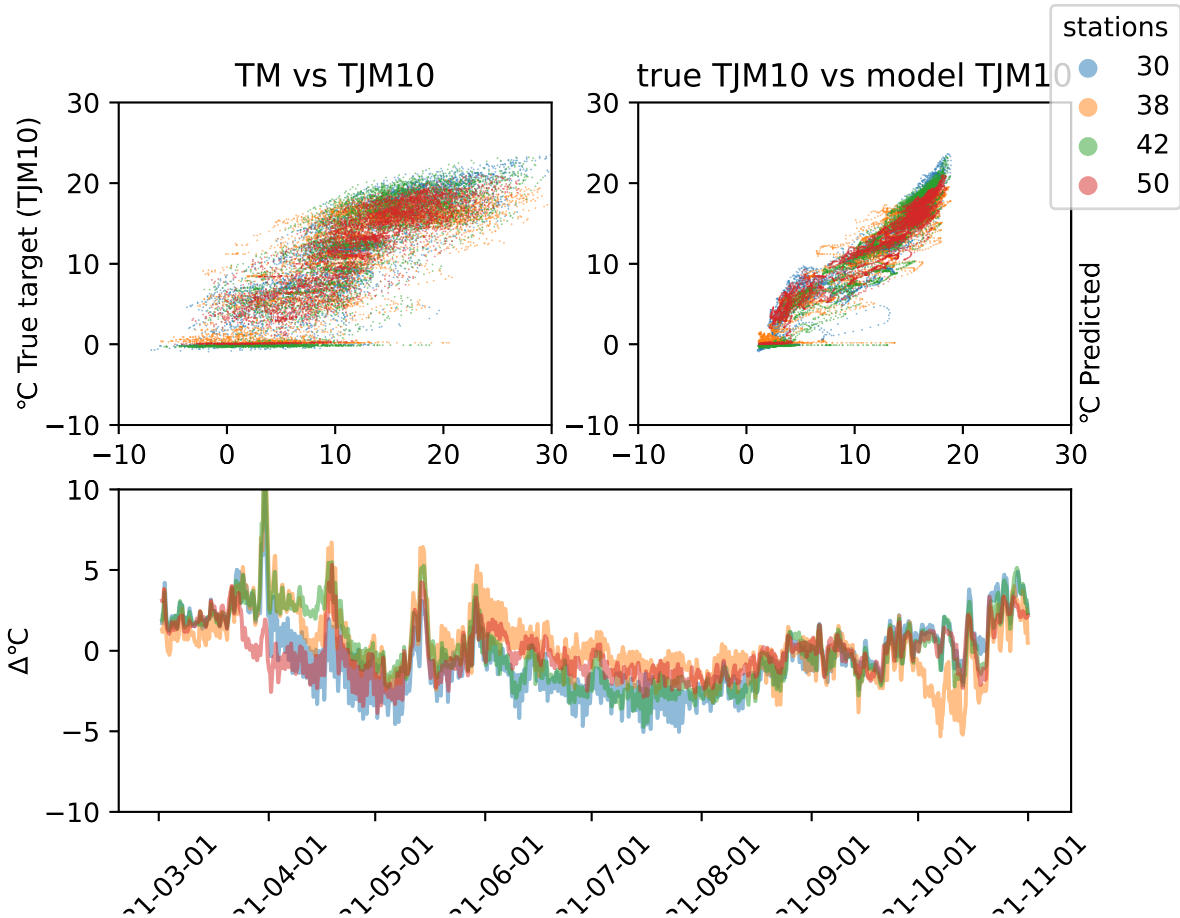


Figure 99: Difference plot for LSTM model in year 2021 and region Vestfold

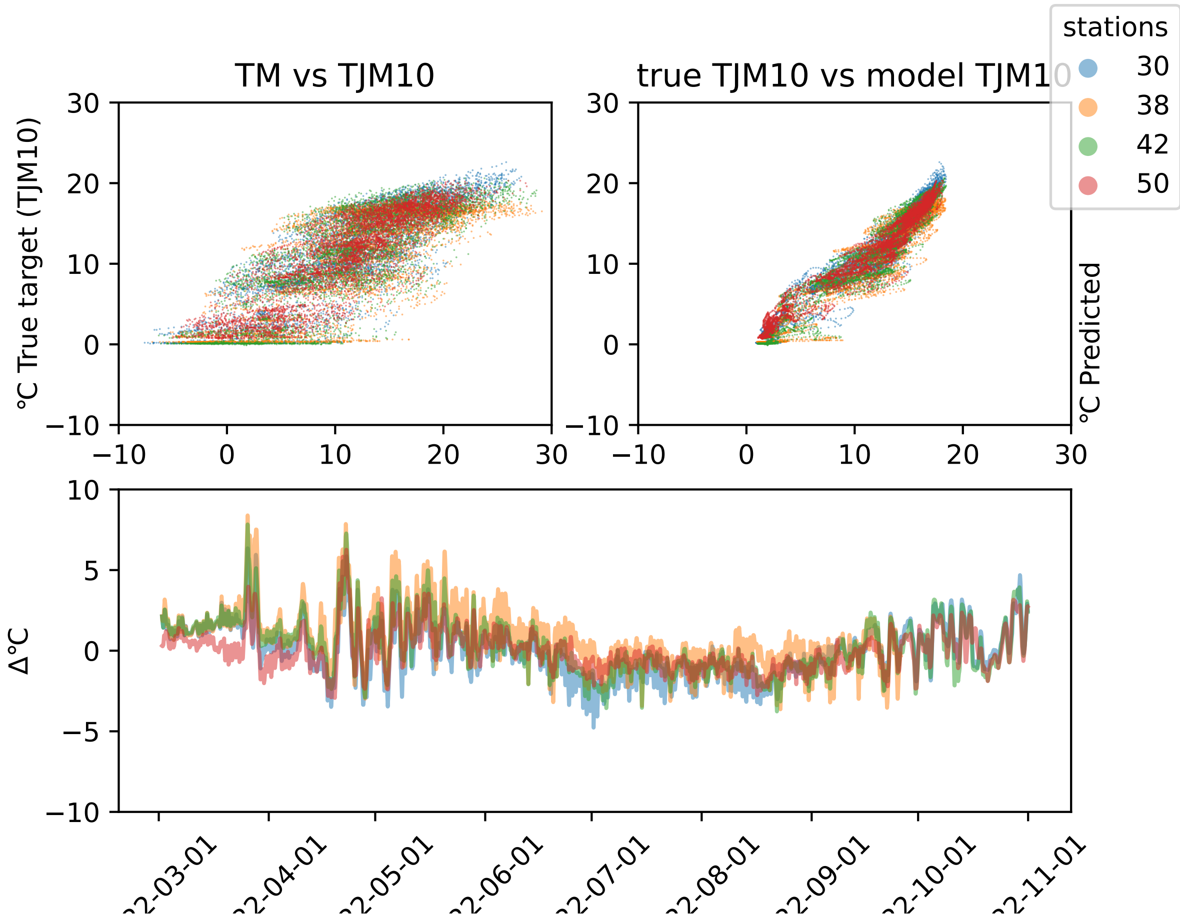


Figure 100: Difference plot for LSTM model in year 2022 and region Vestfold

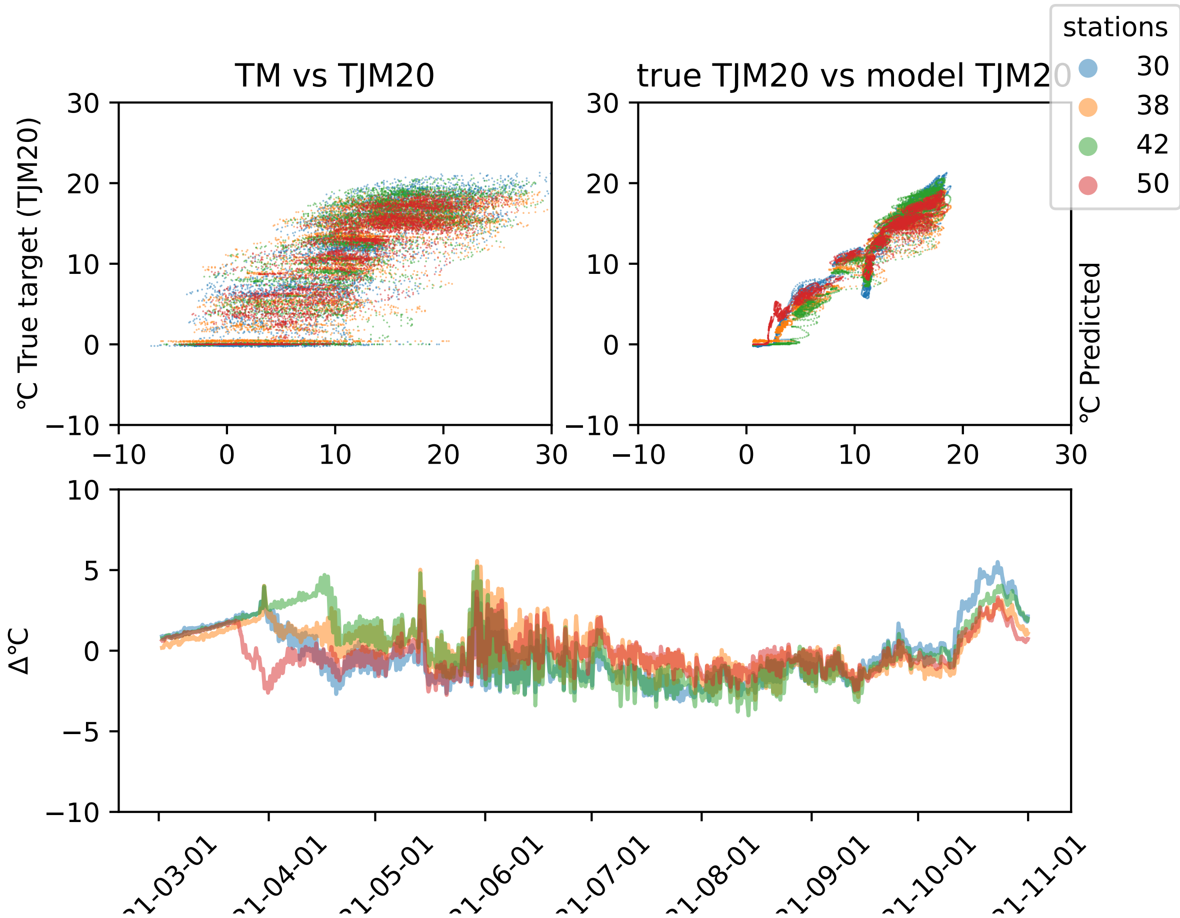


Figure 101: Difference plot for LSTM model in year 2021 and region Vestfold

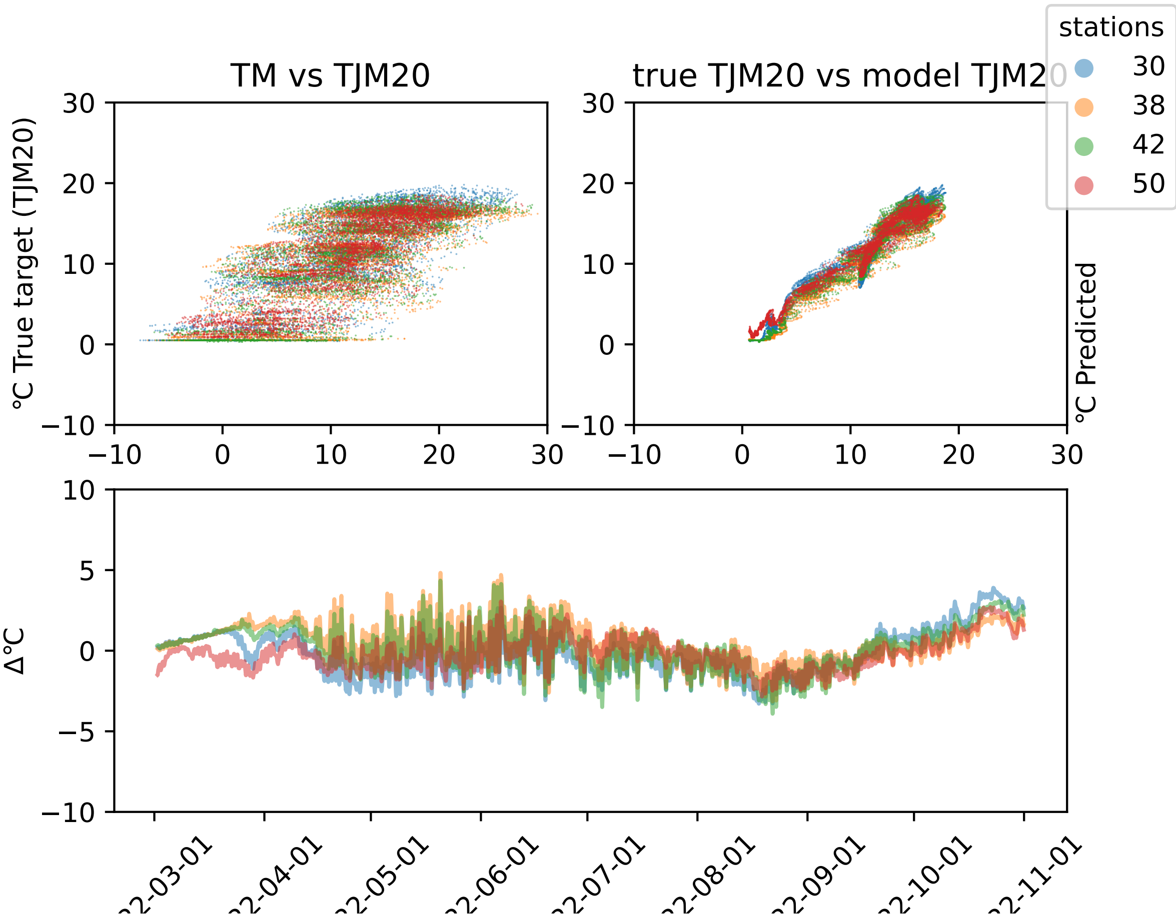


Figure 102: Difference plot for LSTM model in year 2022 and region Vestfold

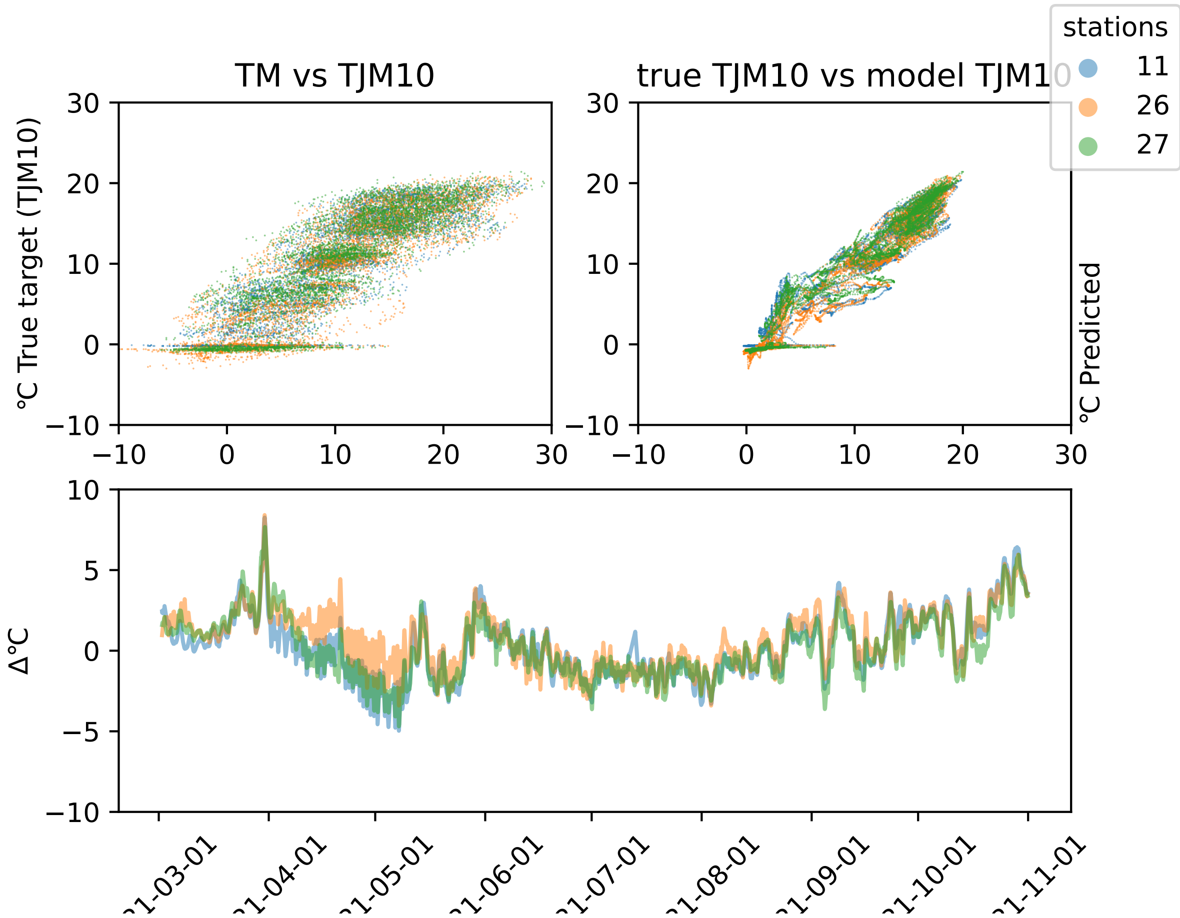


Figure 103: Difference plot for GRU model in year 2021 and region Innlandet

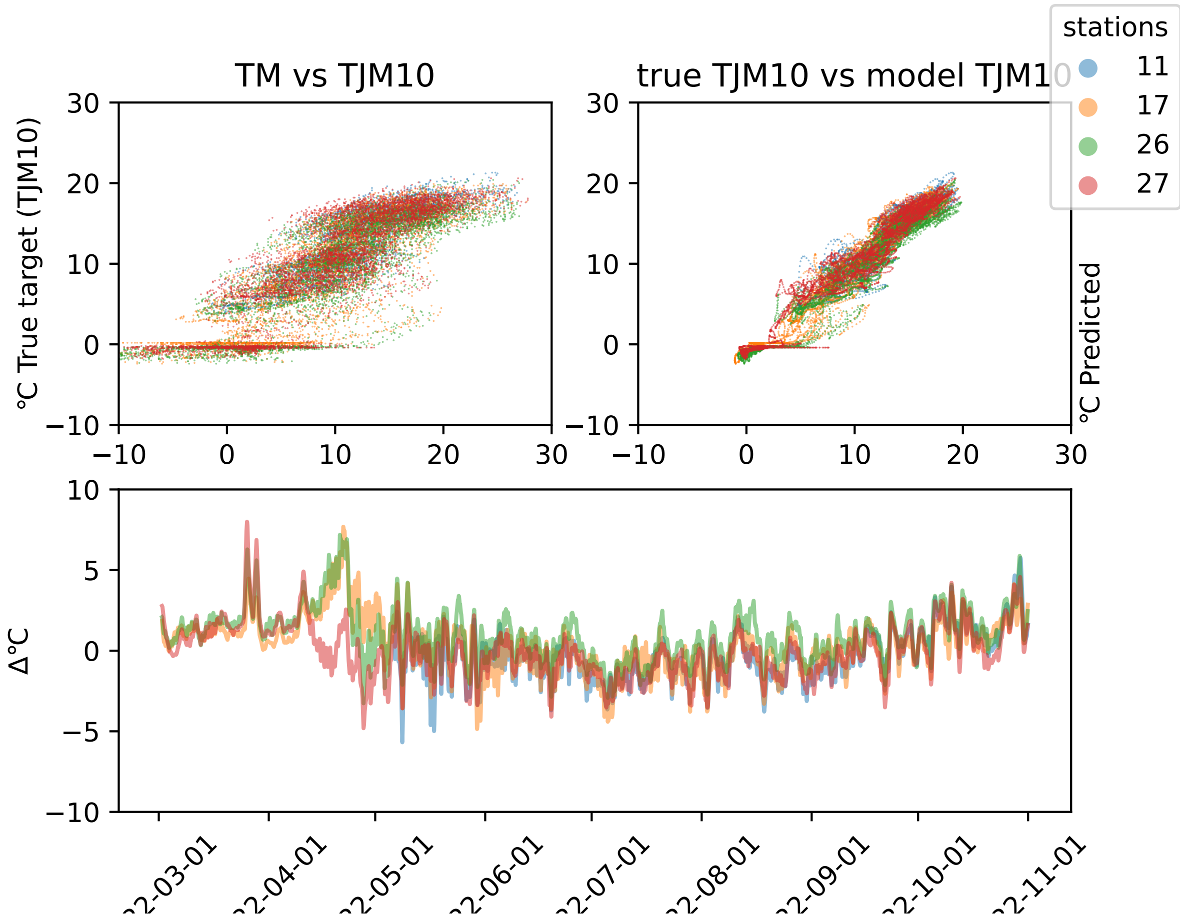


Figure 104: Difference plot for GRU model in year 2022 and region Innlandet

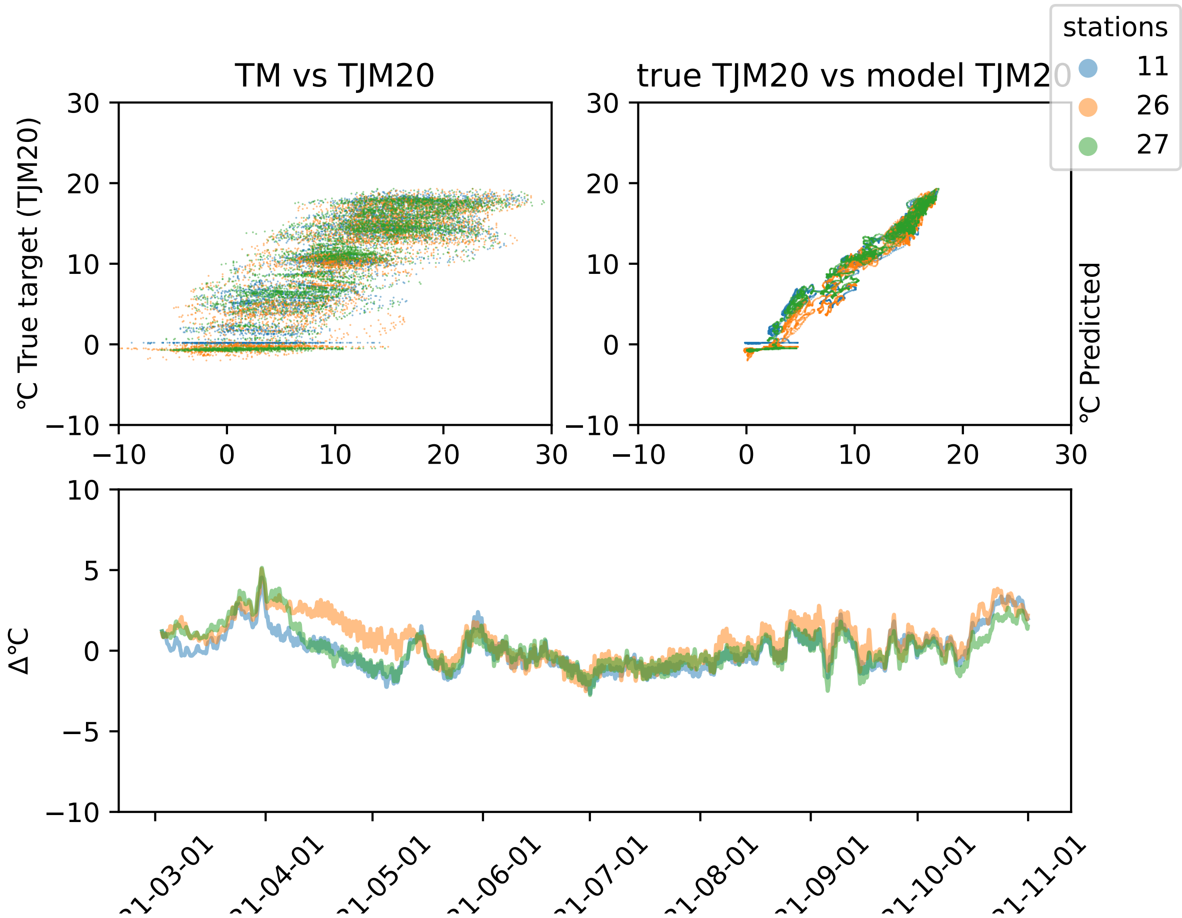


Figure 105: Difference plot for GRU model in year 2021 and region Innlandet

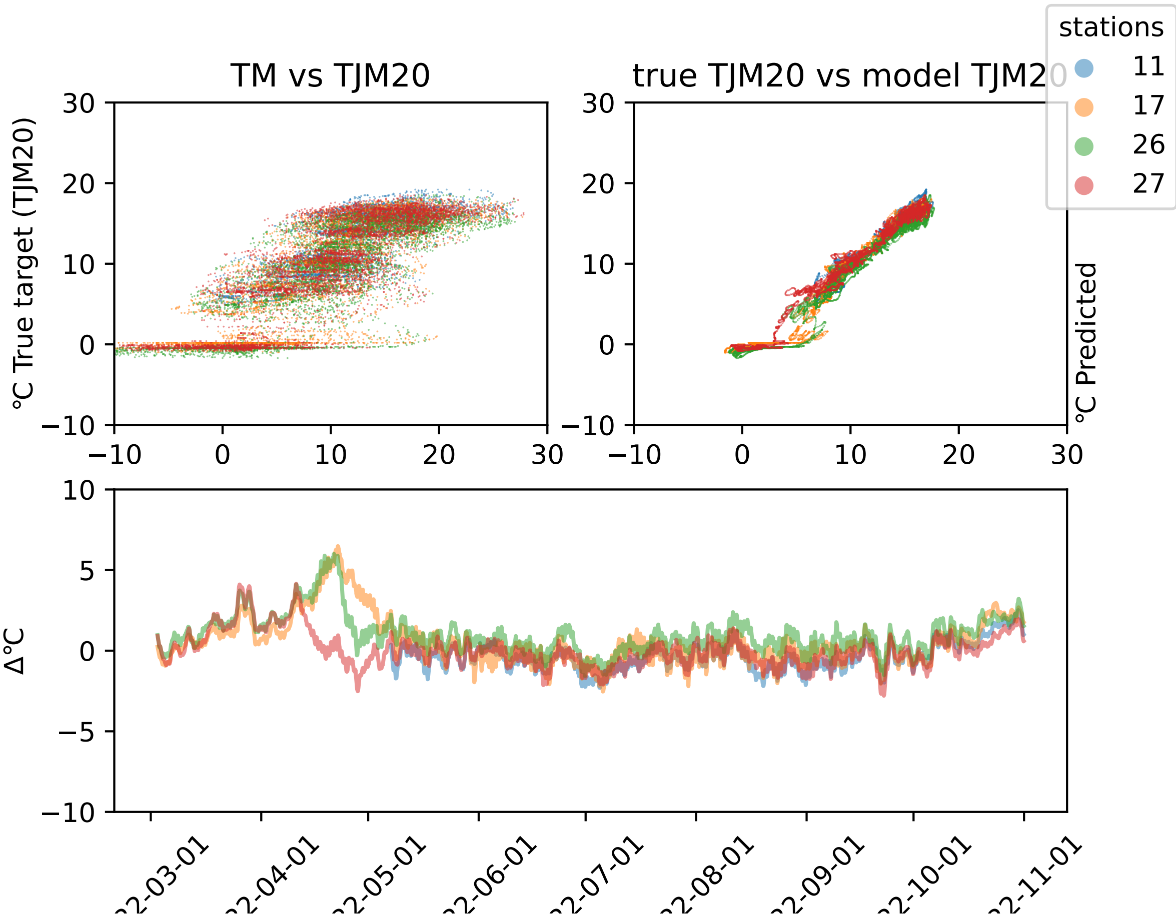


Figure 106: Difference plot for GRU model in year 2022 and region Innlandet

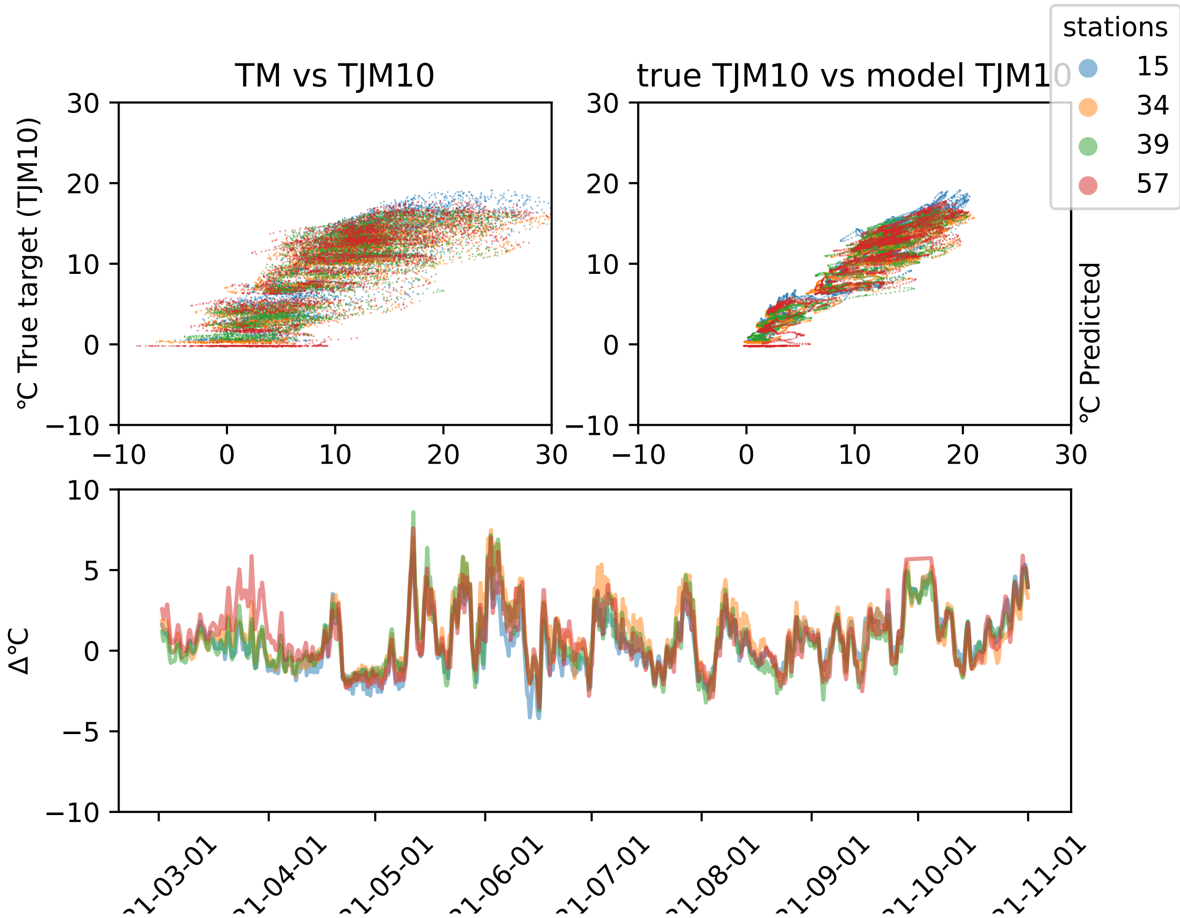


Figure 107: Difference plot for GRU model in year 2021 and region Trøndelag

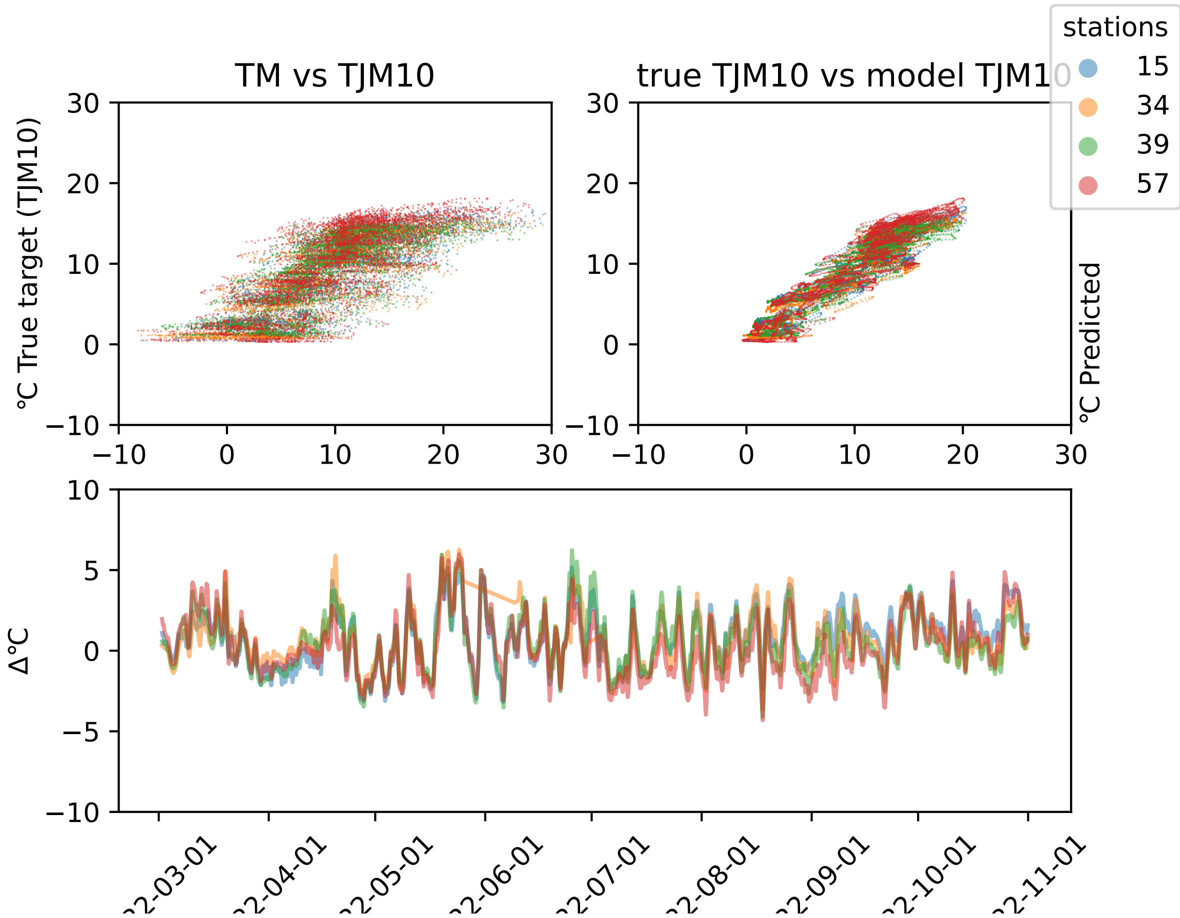


Figure 108: Difference plot for GRU model in year 2022 and region Trøndelag

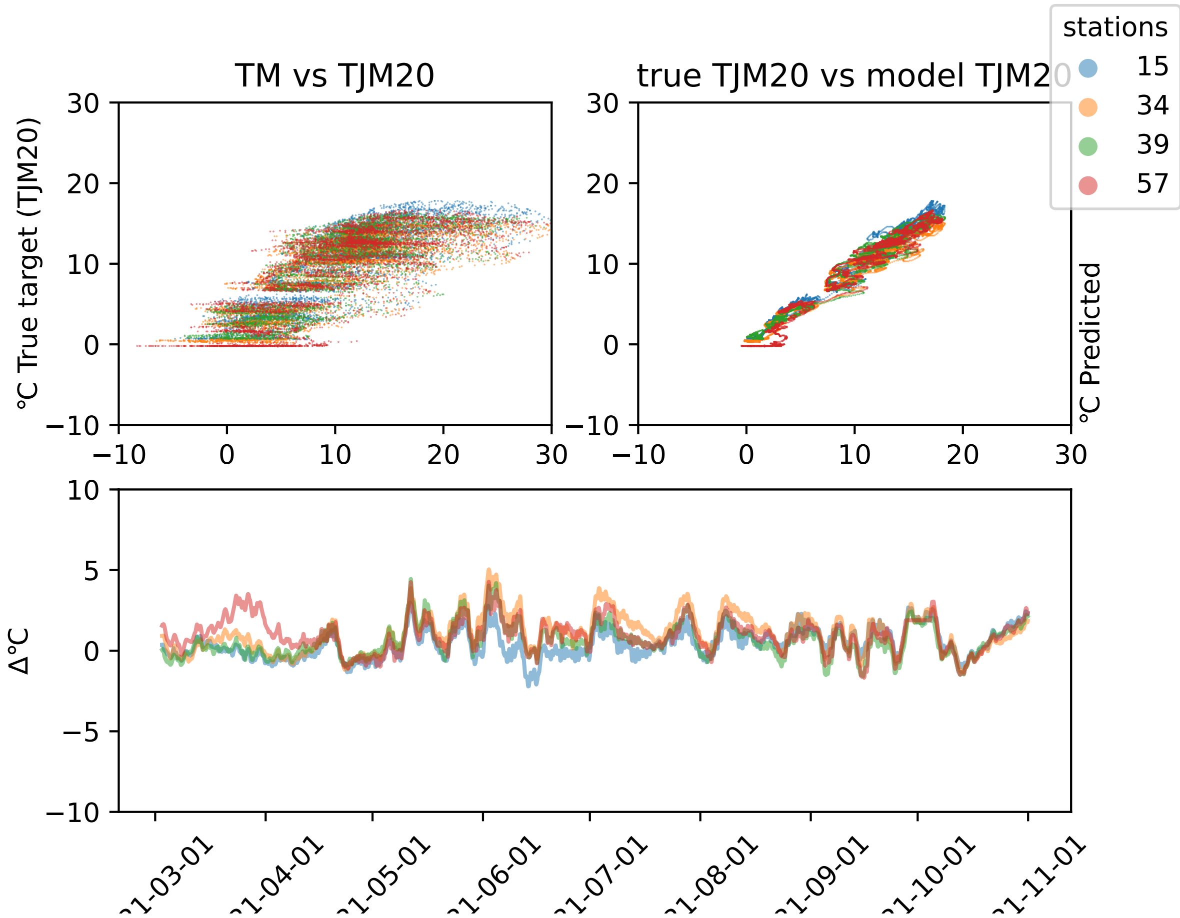


Figure 109: Difference plot for GRU model in year 2021 and region Trøndelag

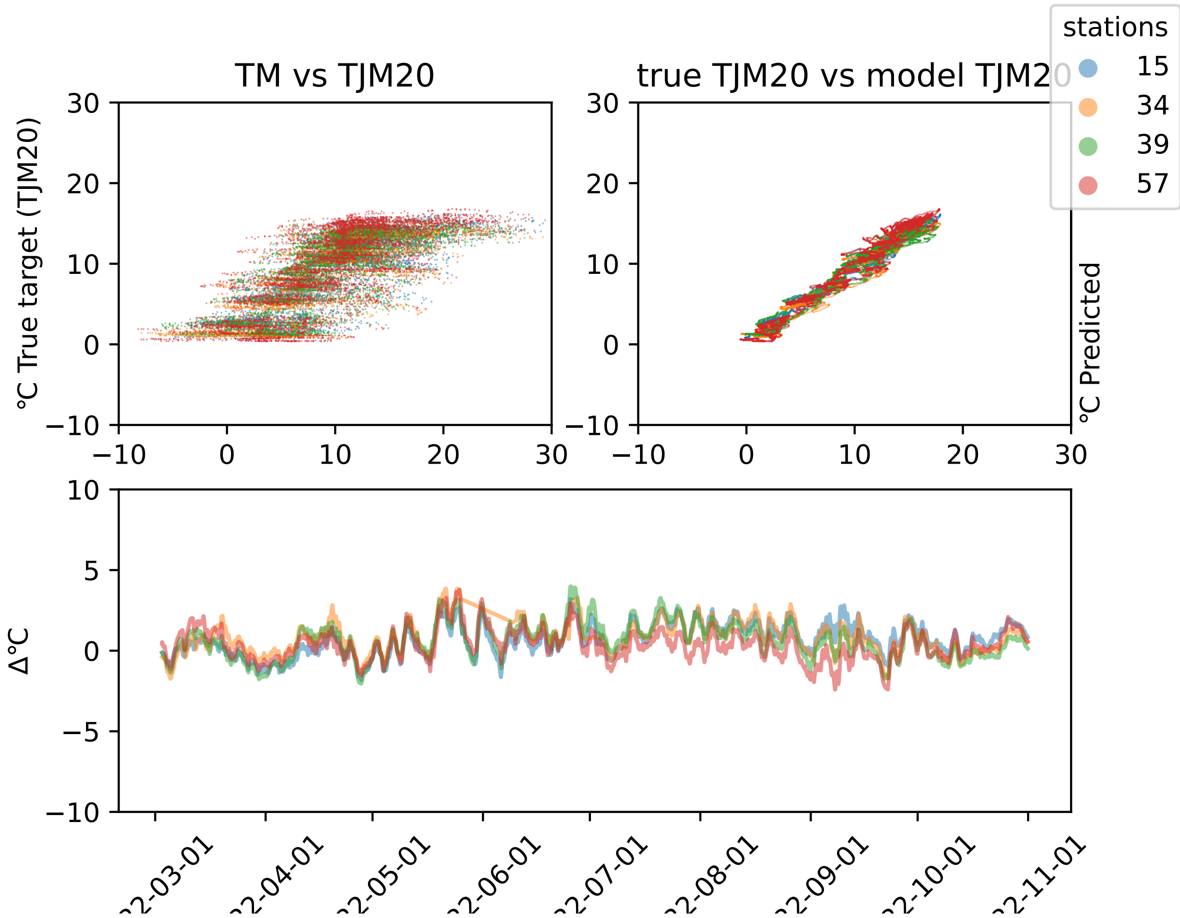


Figure 110: Difference plot for GRU model in year 2022 and region Trøndelag

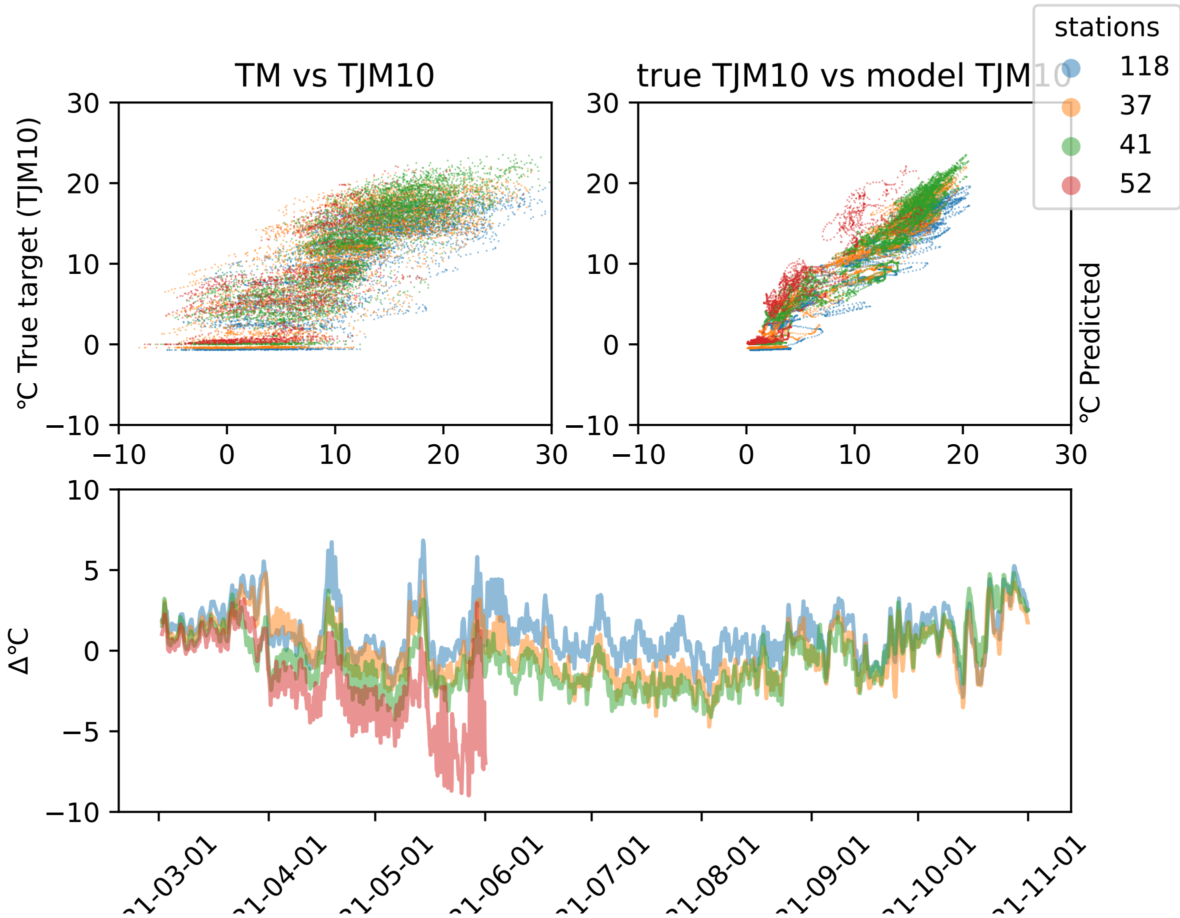


Figure 111: Difference plot for GRU model in year 2021 and region Østfold

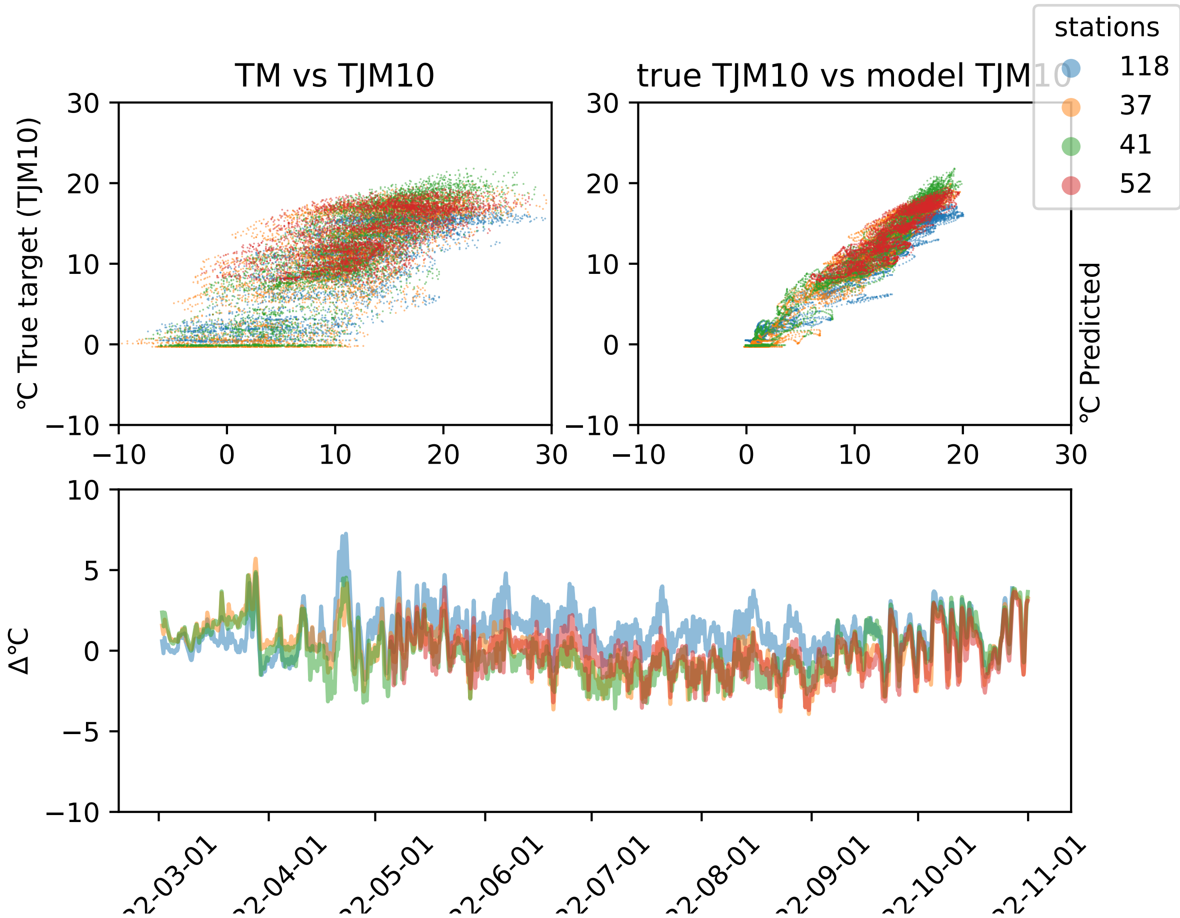


Figure 112: Difference plot for GRU model in year 2022 and region Østfold

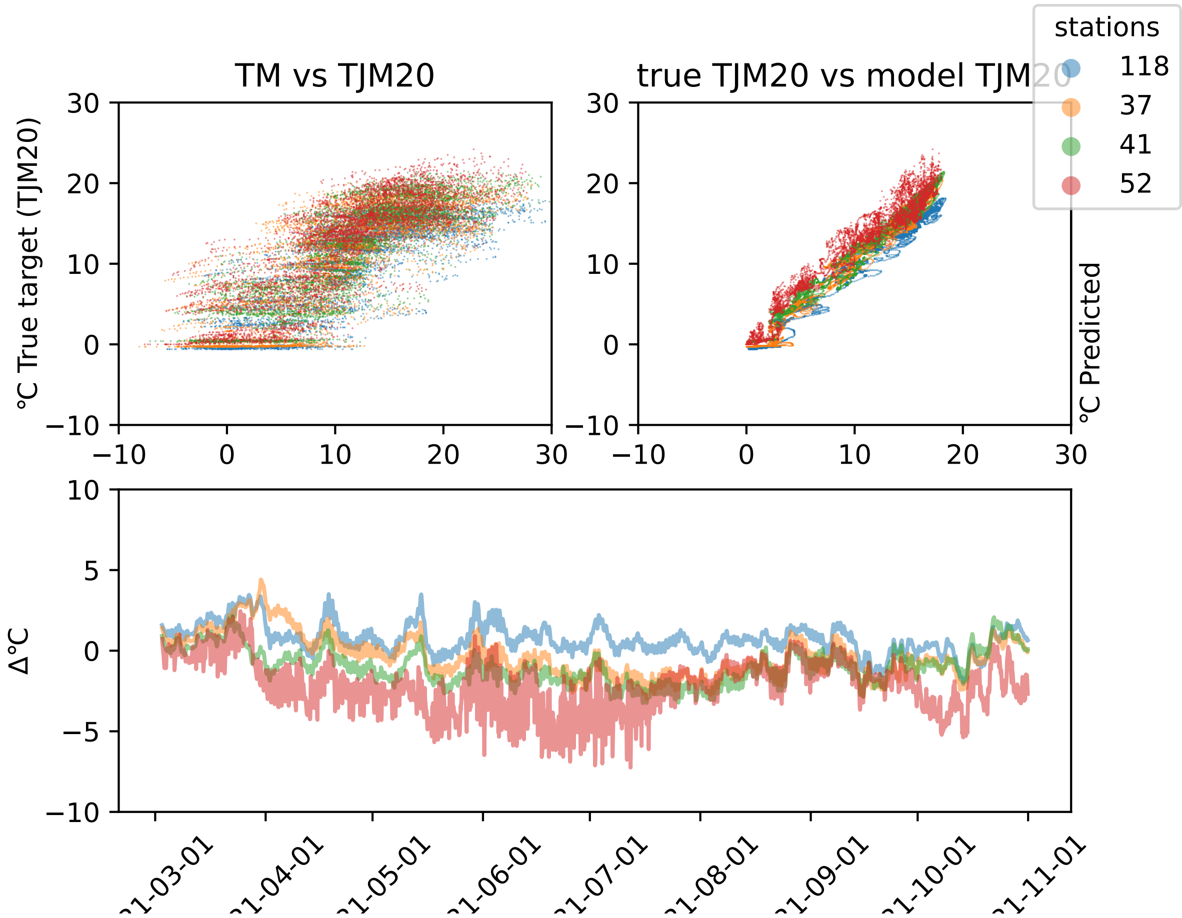


Figure 113: Difference plot for GRU model in year 2021 and region Østfold

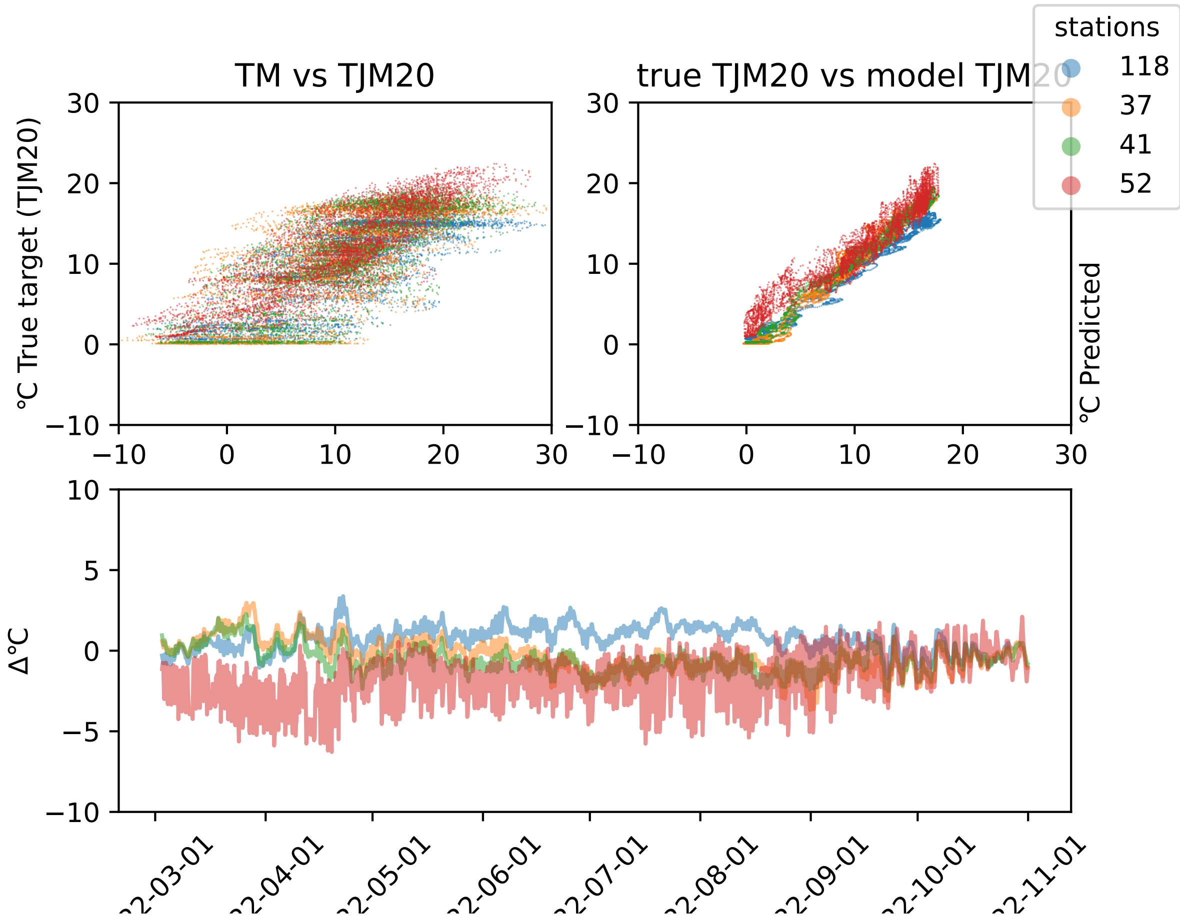


Figure 114: Difference plot for GRU model in year 2022 and region Østfold

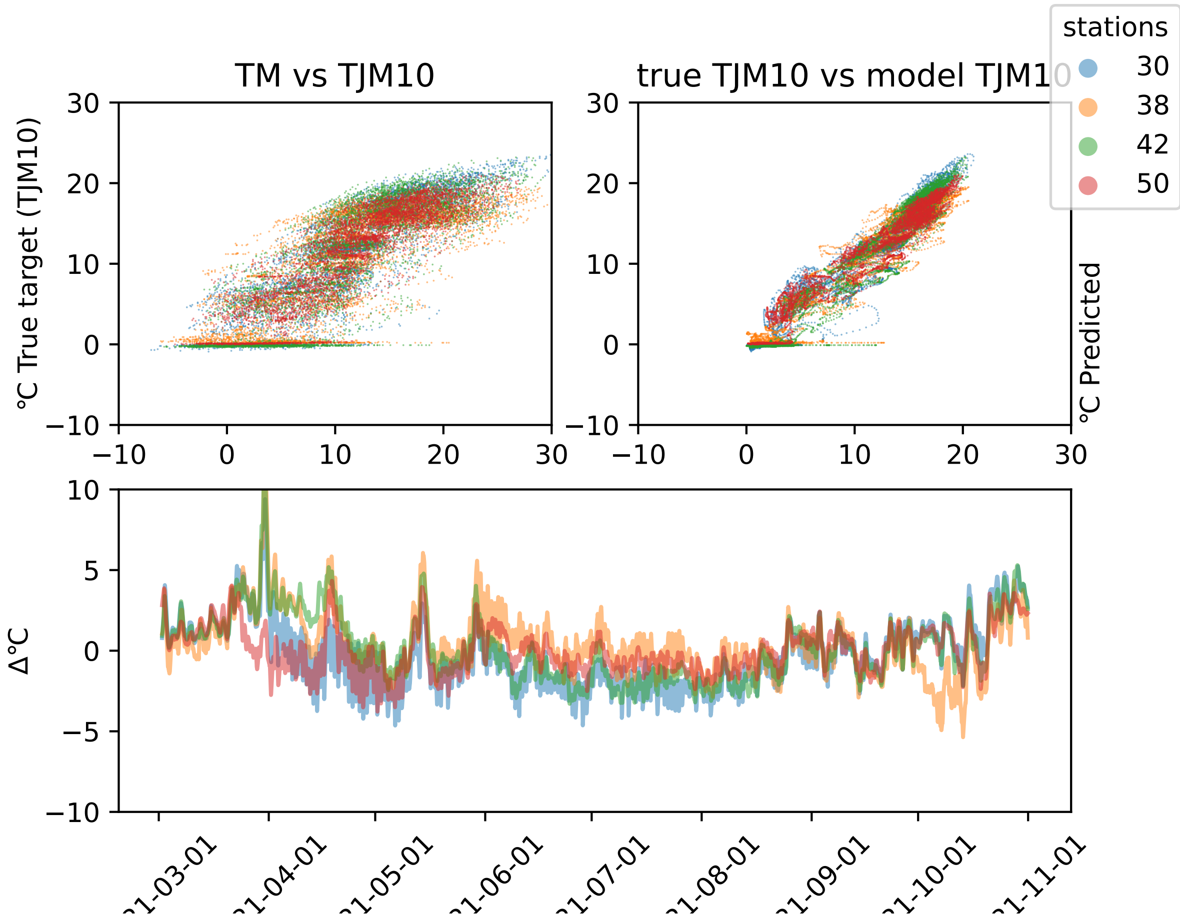


Figure 115: Difference plot for GRU model in year 2021 and region Vestfold

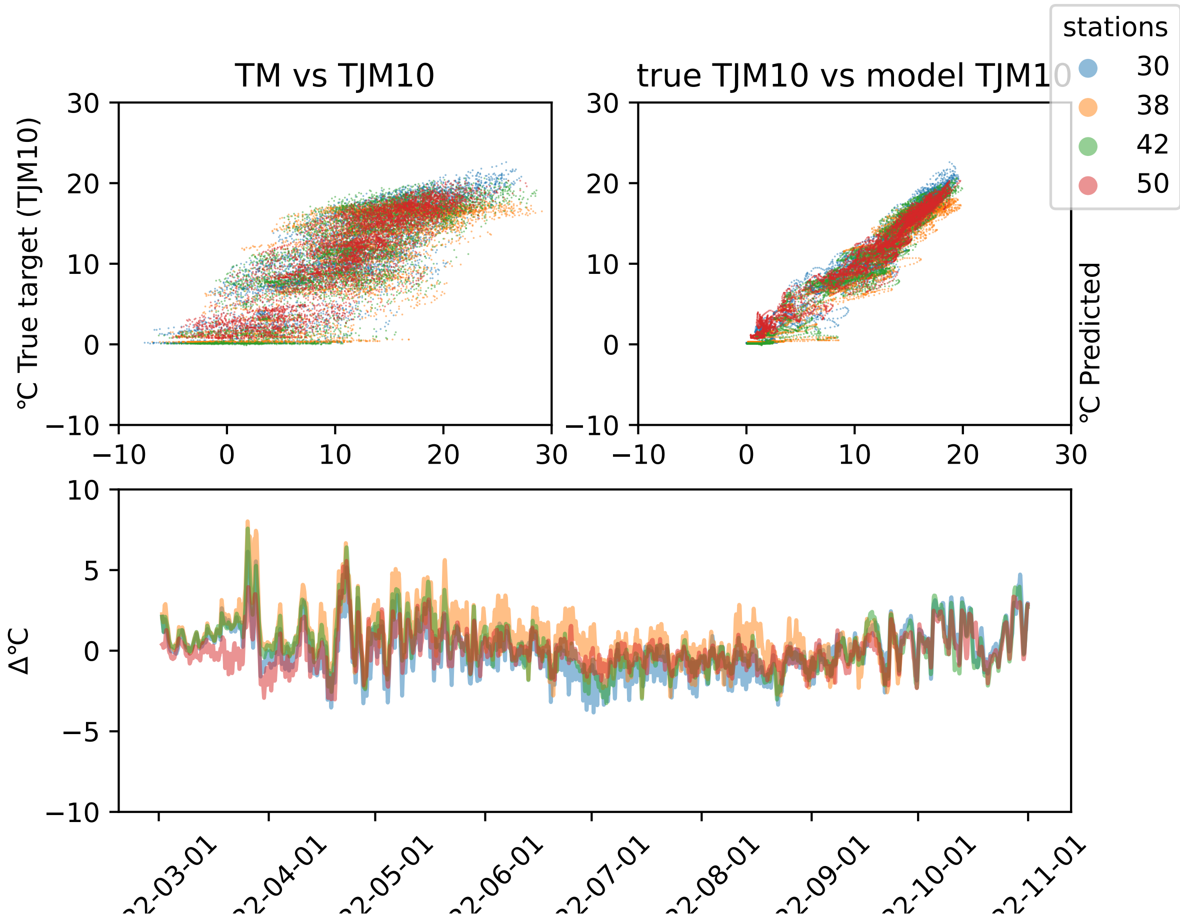


Figure 116: Difference plot for GRU model in year 2022 and region Vestfold

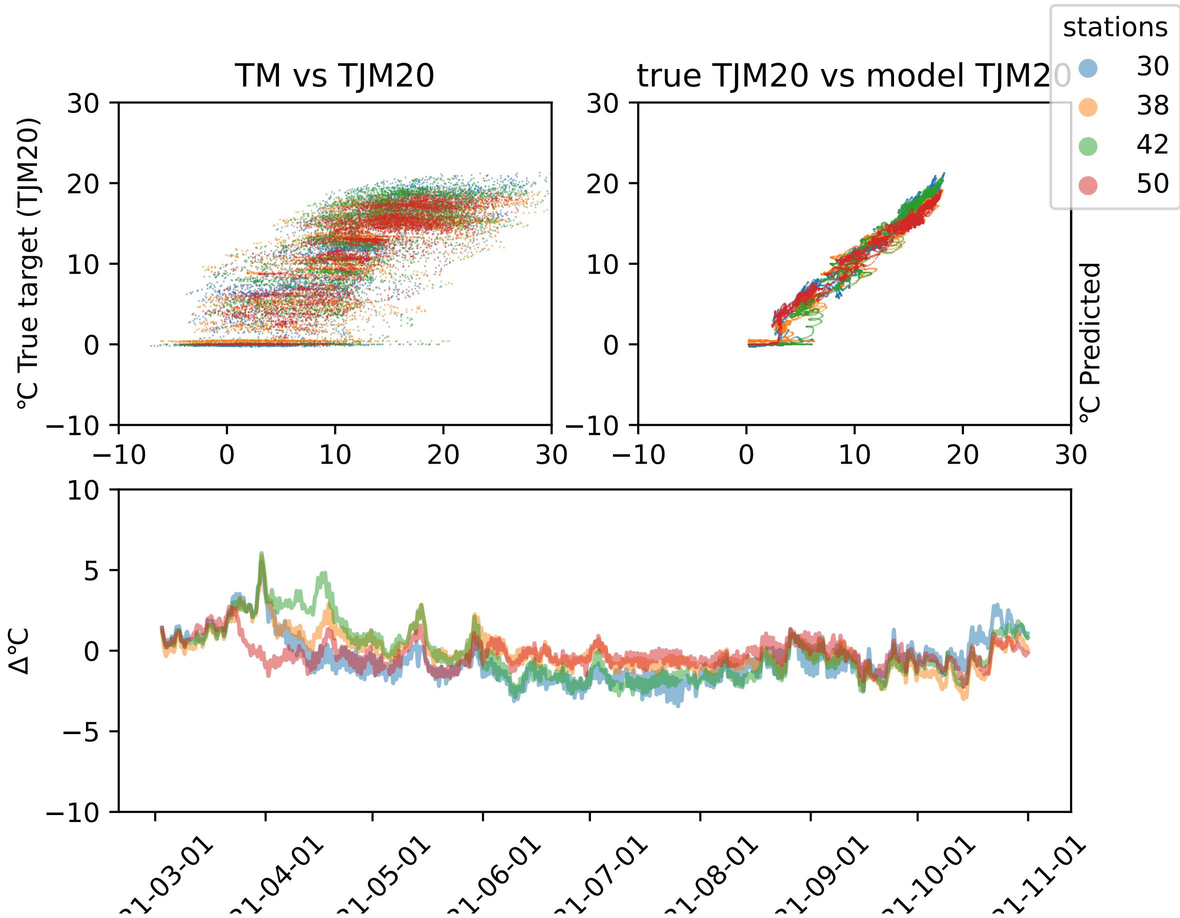


Figure 117: Difference plot for GRU model in year 2021 and region Vestfold

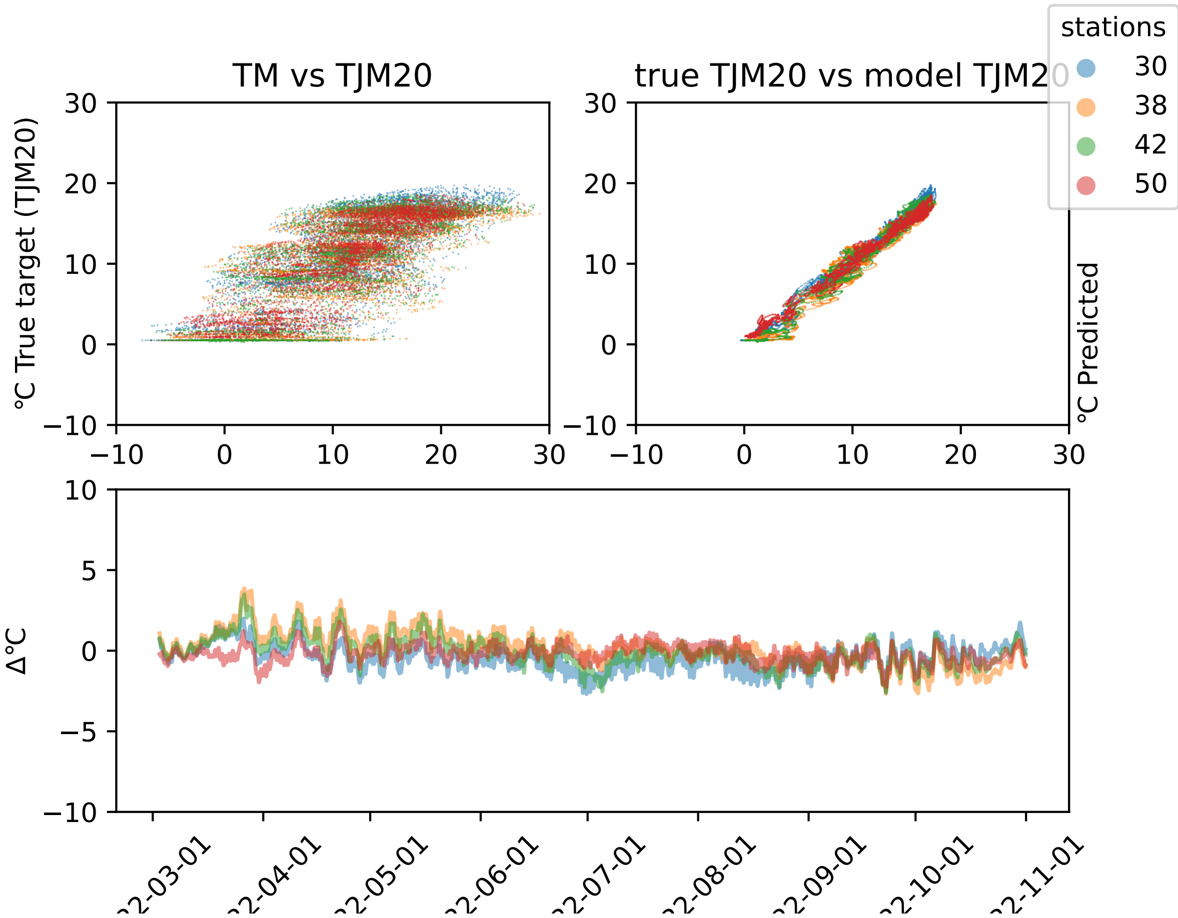


Figure 118: Difference plot for GRU model in year 2022 and region Vestfold

A.2 Data visiulation of data before treatment

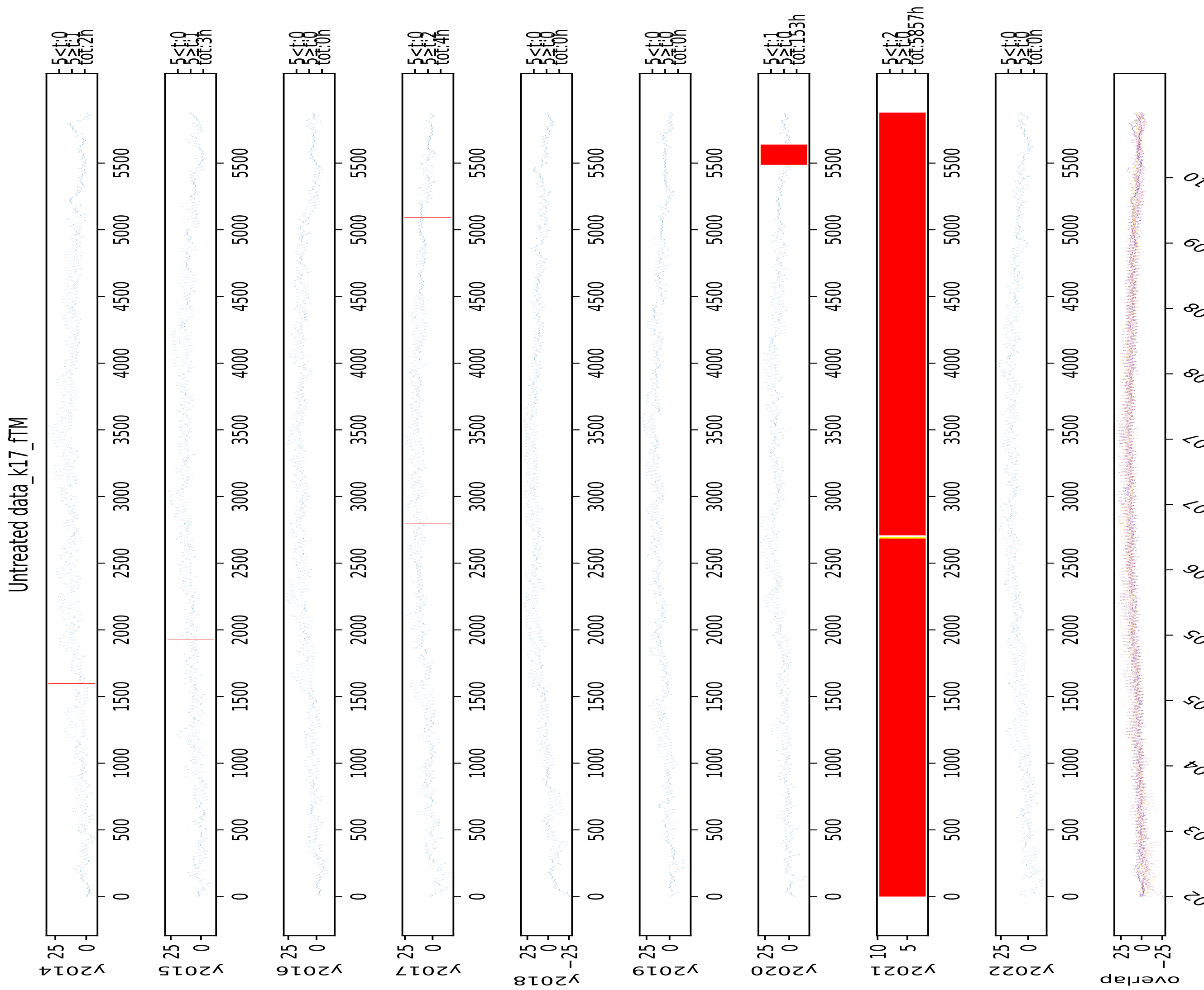


Figure 119: Visual representation of missing values at station 17 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

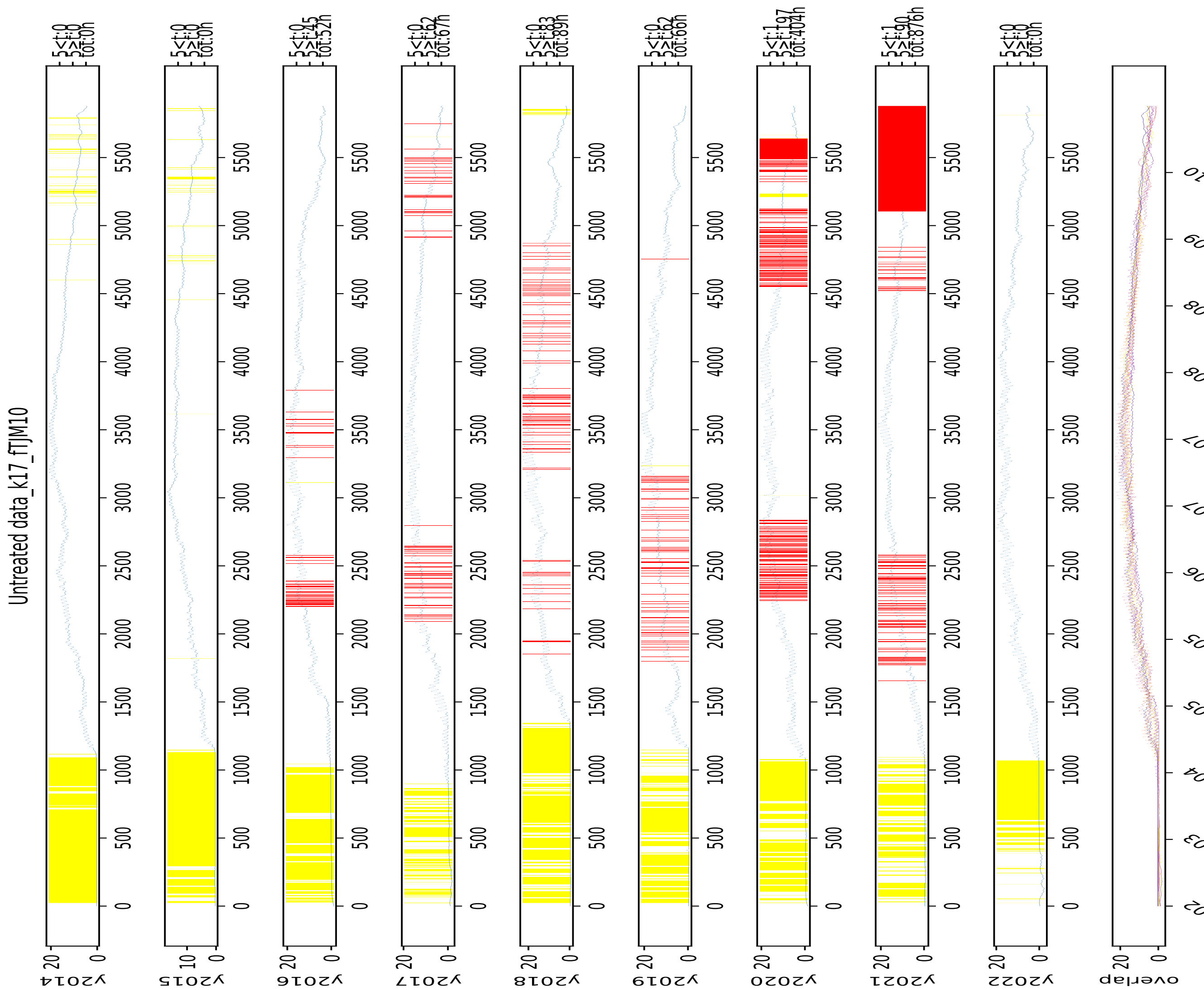


Figure 120: Visual representation of missing values at station 17 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

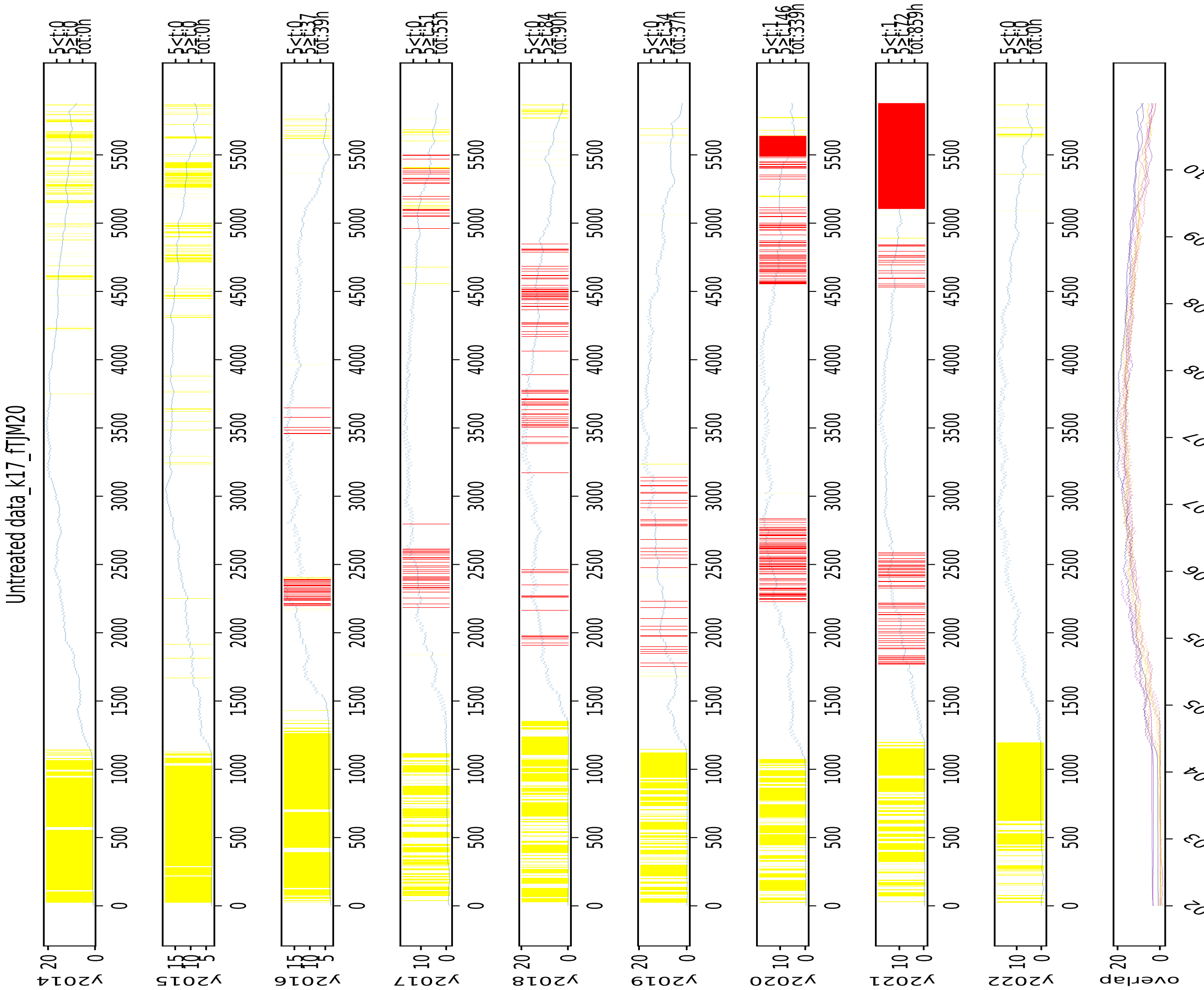


Figure 121: Visual representation of missing values at station 17 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

Figure 122: Visual representation of missing values at station 11 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

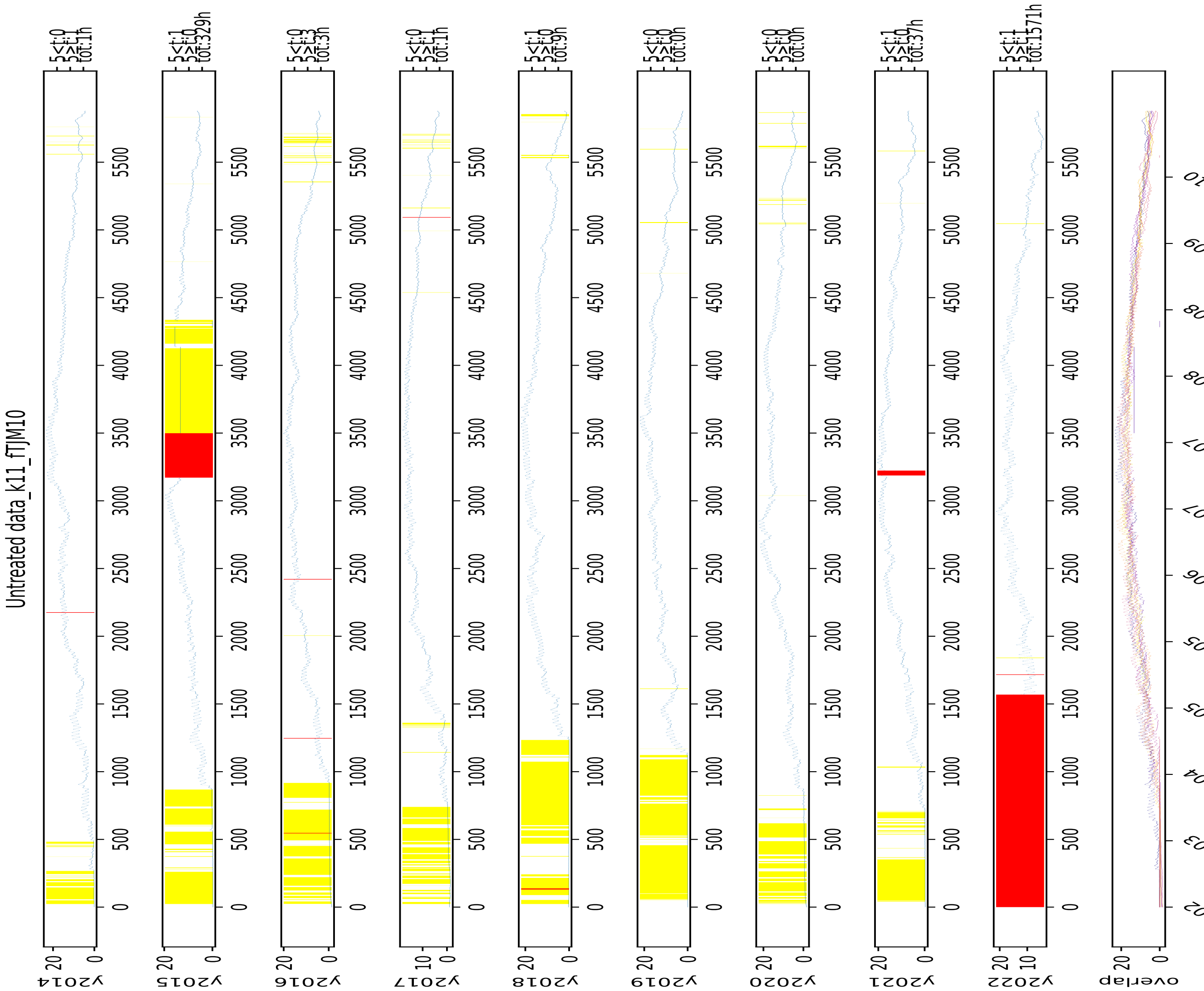


Figure 123: Visual representation of missing values at station 11 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

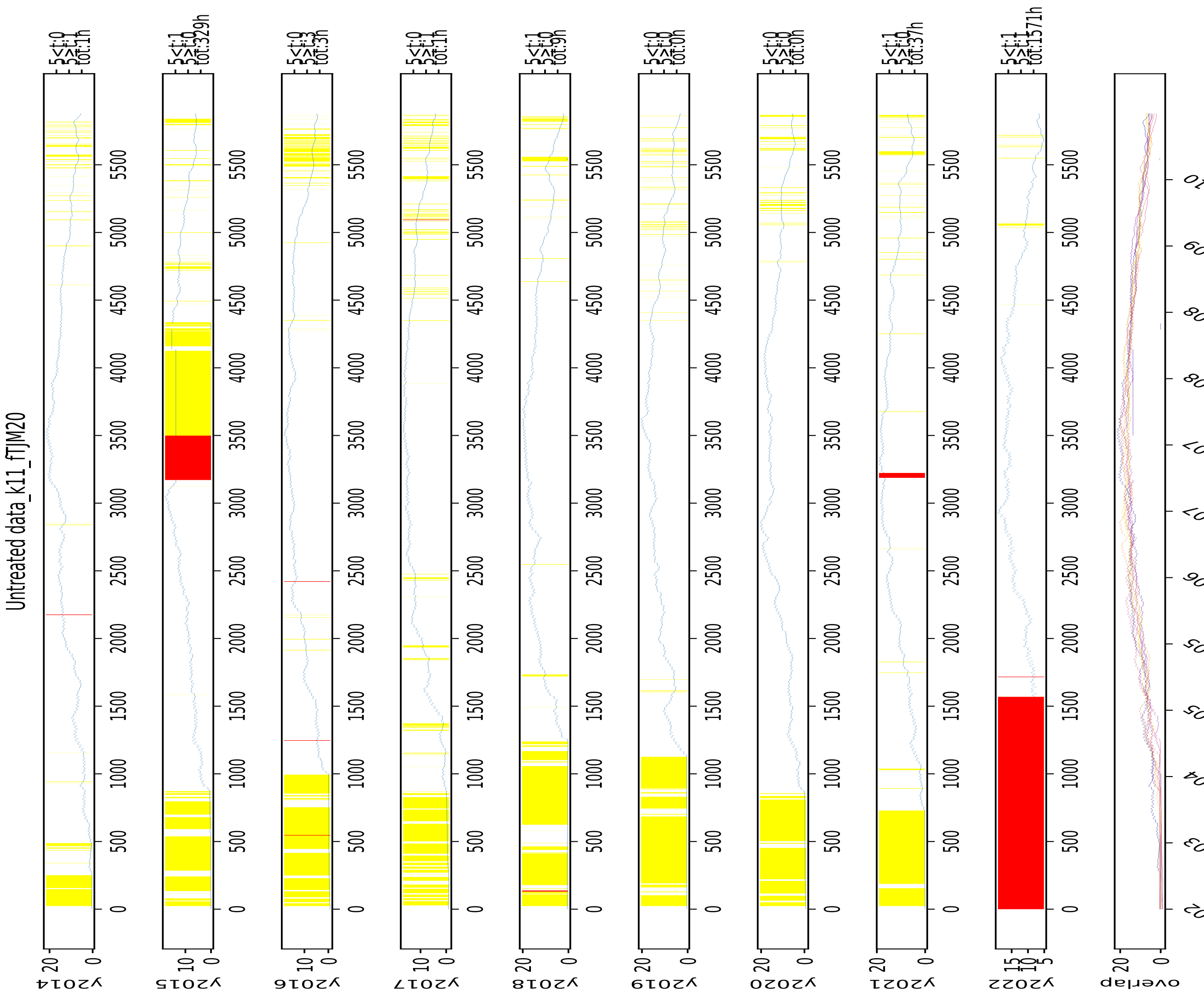


Figure 124: Visual representation of missing values at station 11 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

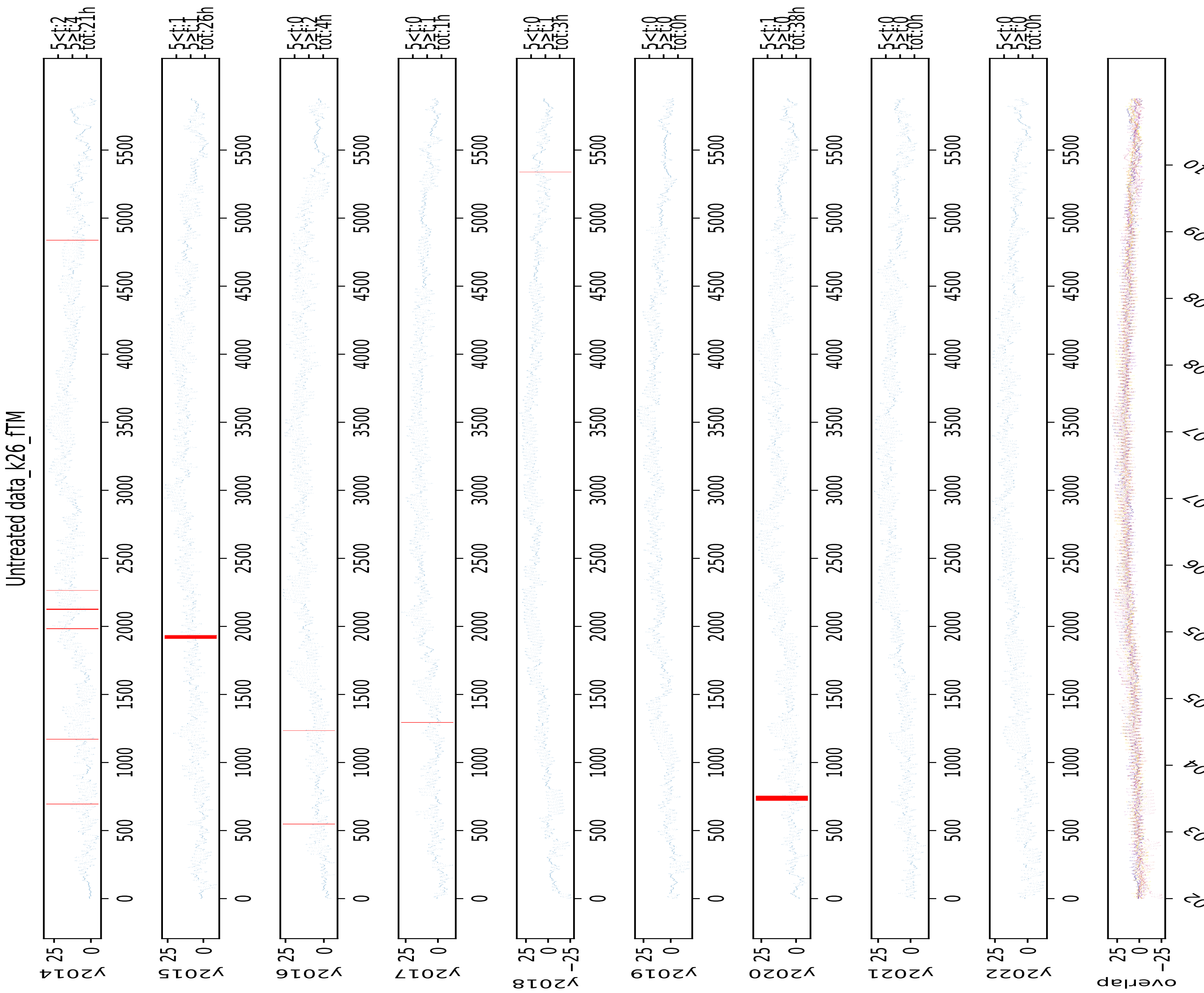


Figure 125: Visual representation of missing values at station 26 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

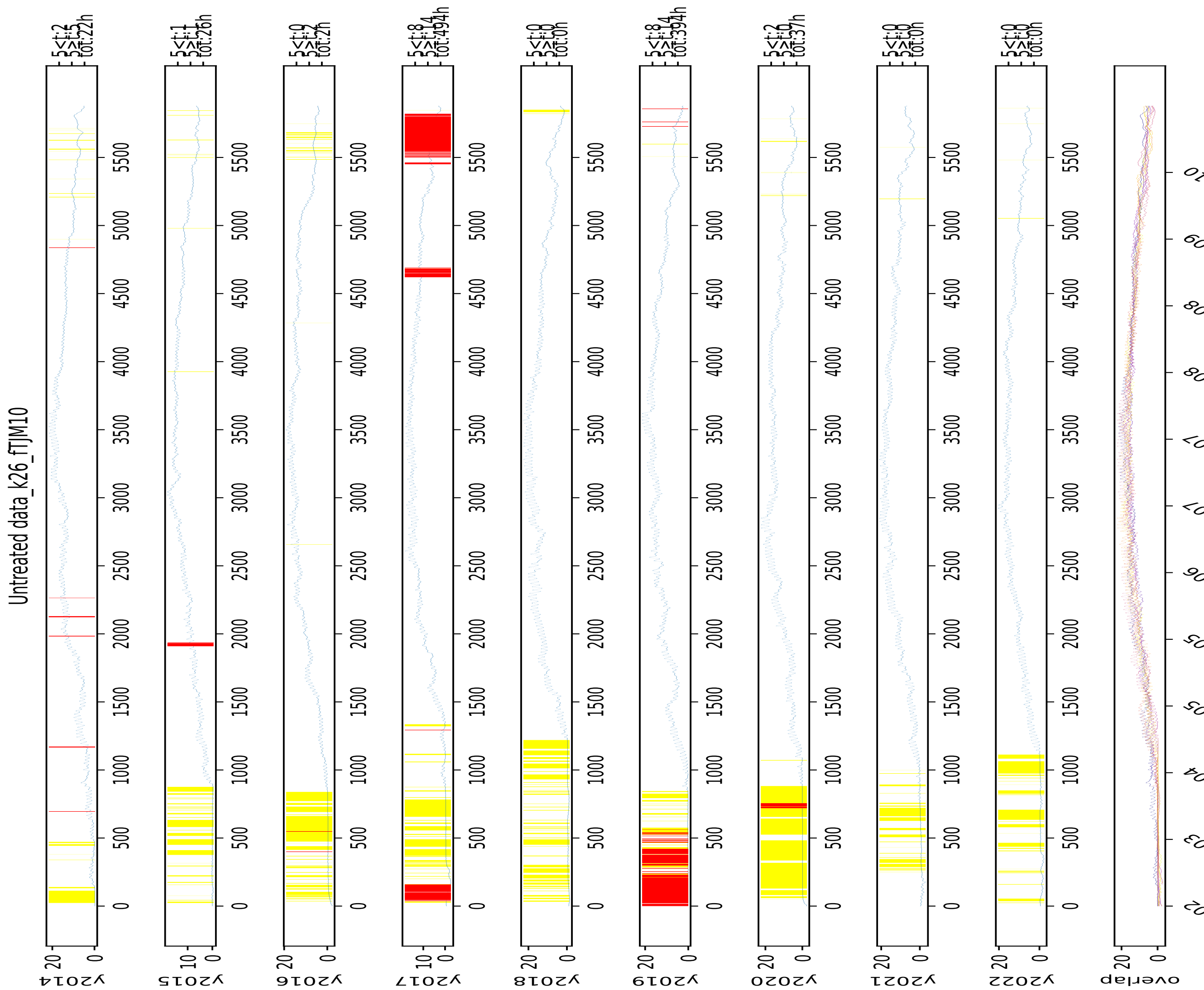


Figure 126: Visual representation of missing values at station 26 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

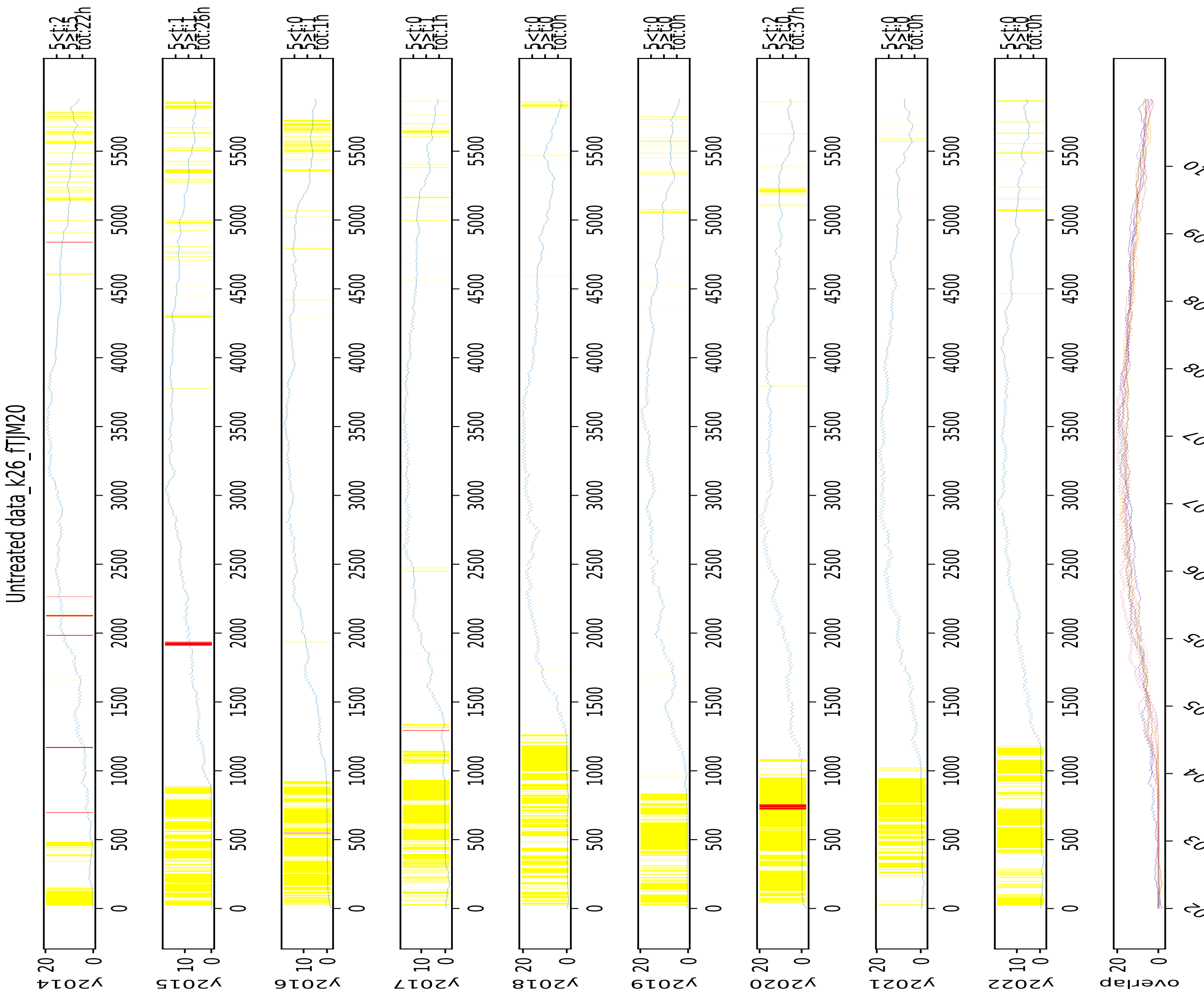


Figure 127: Visual representation of missing values at station 26 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

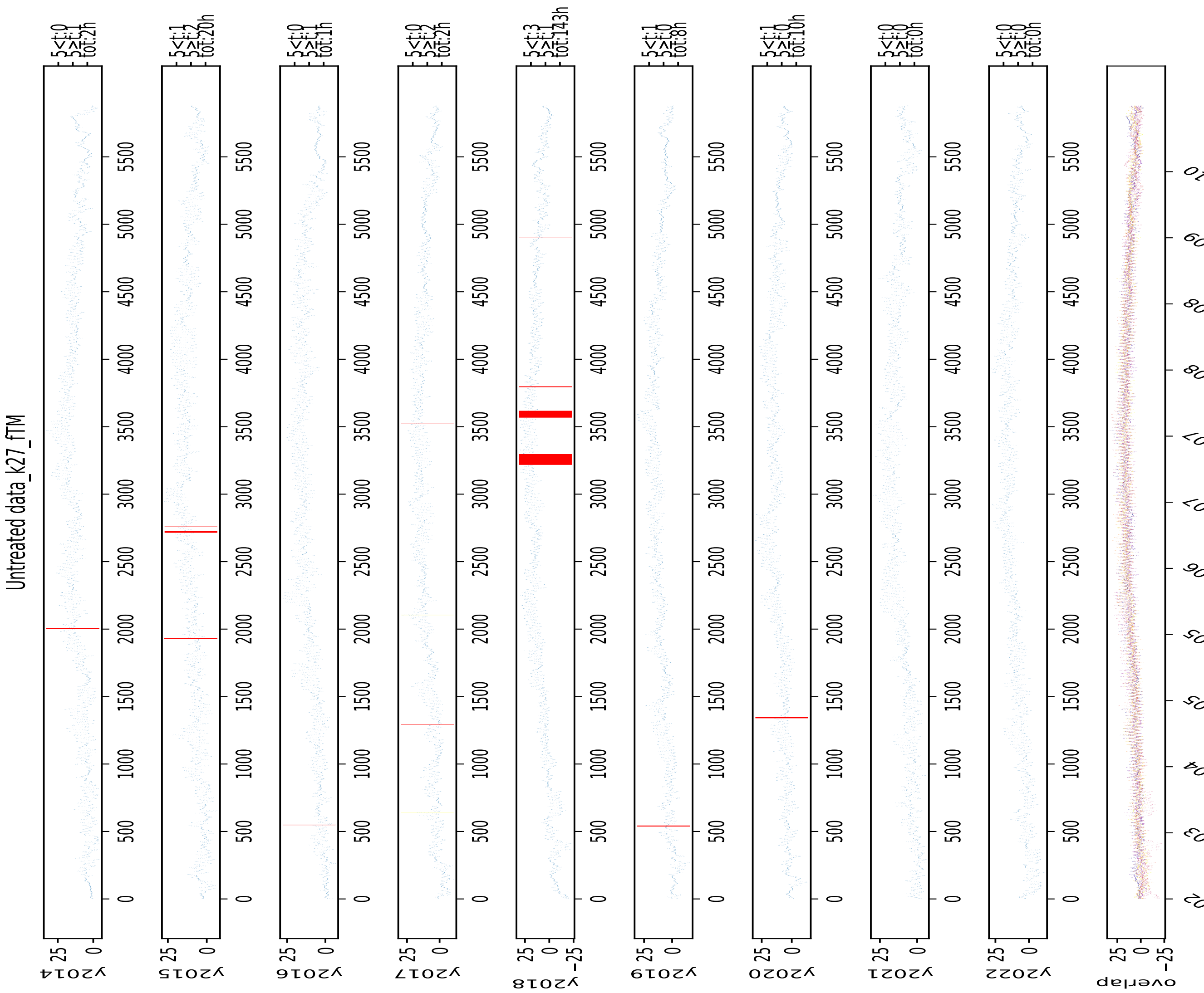


Figure 128: Visual representation of missing values at station 27 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

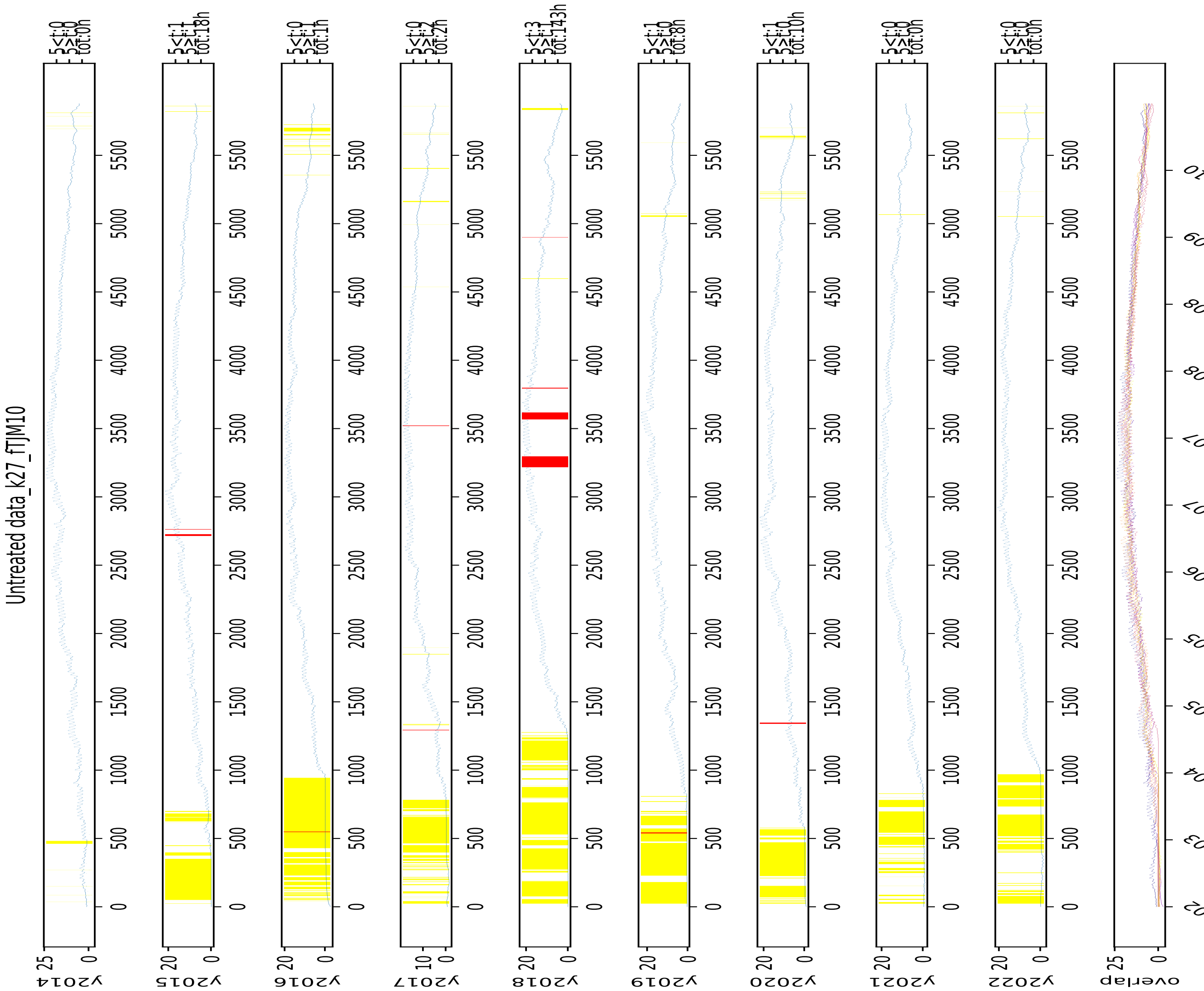


Figure 129: Visual representation of missing values at station 27 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

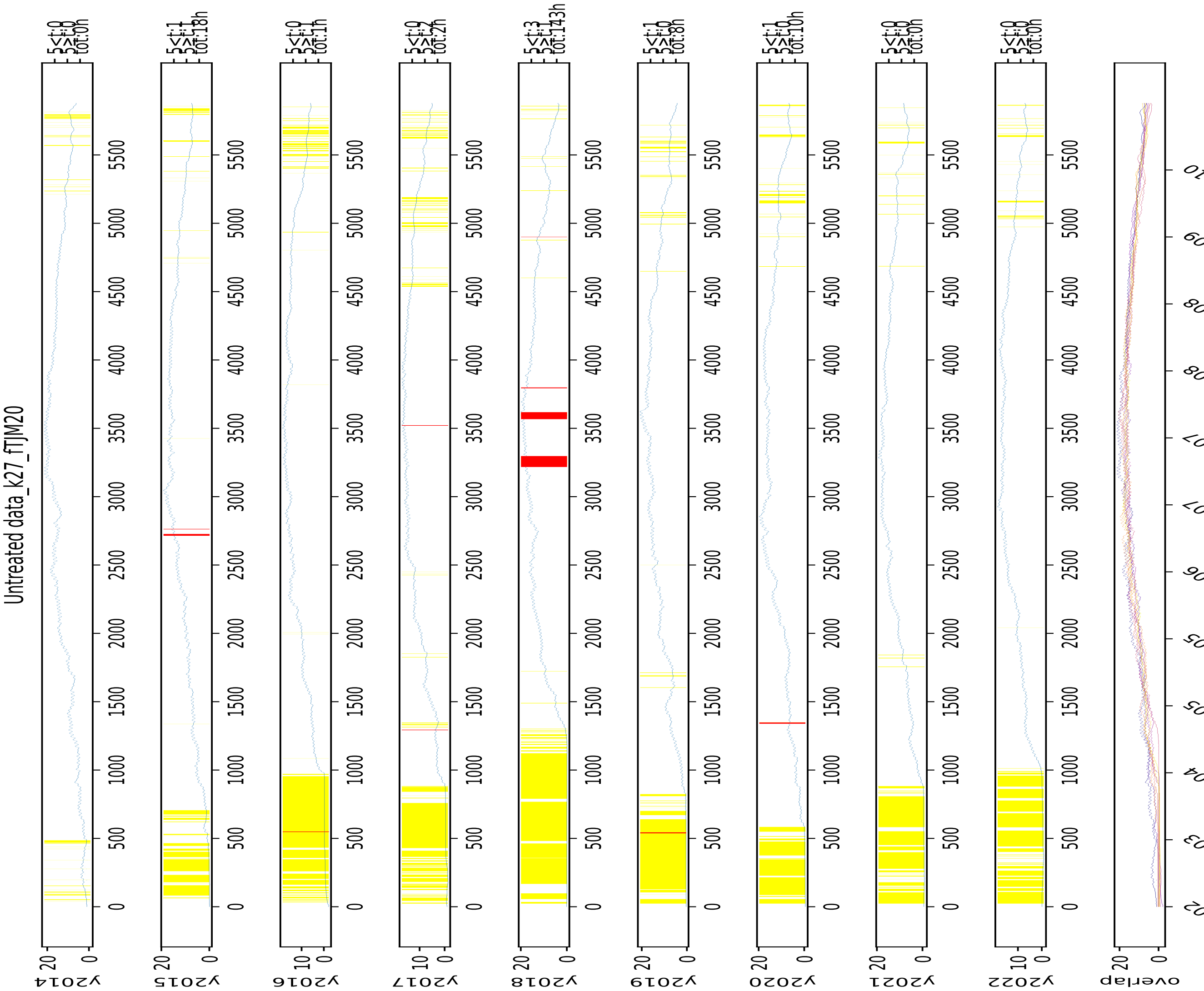


Figure 130: Visual representation of missing values at station 27 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

Figure 131: Visual representation of missing values at station 15 from 2014 to 2022 at the parameter ”TM” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

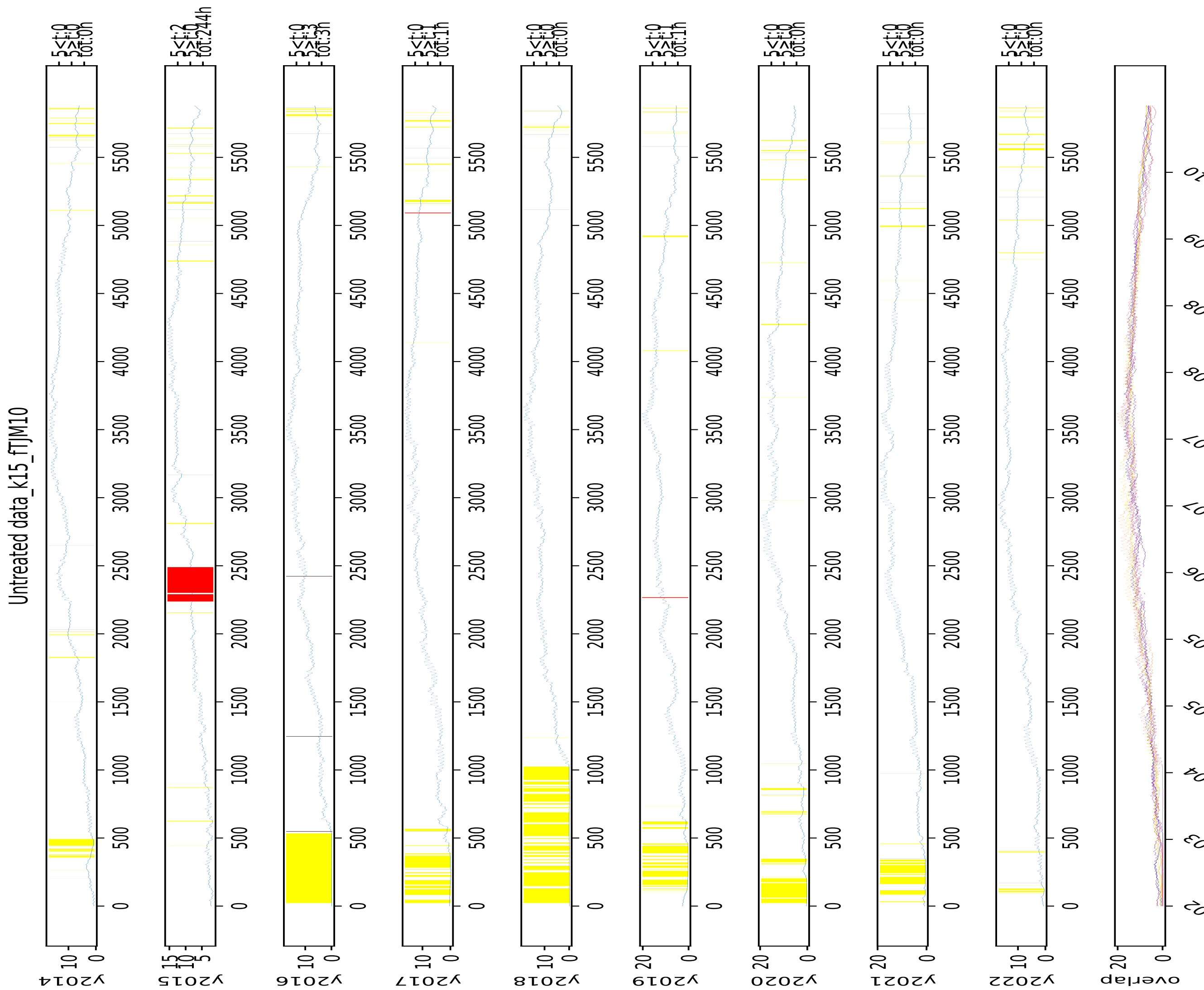


Figure 132: Visual representation of missing values at station 15 from 2014 to 2022 at the parameter ”TJM10” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

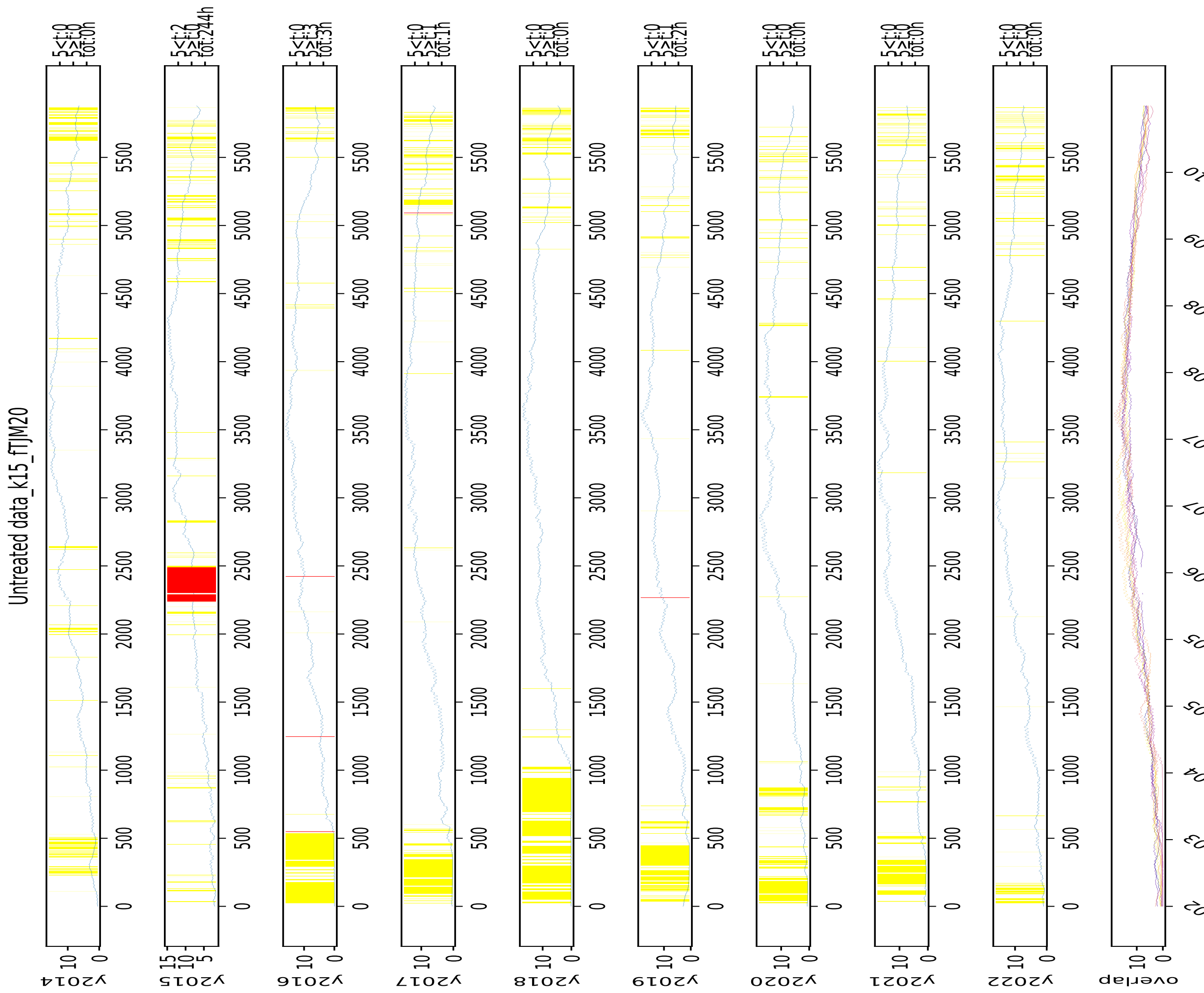


Figure 133: Visual representation of missing values at station 15 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

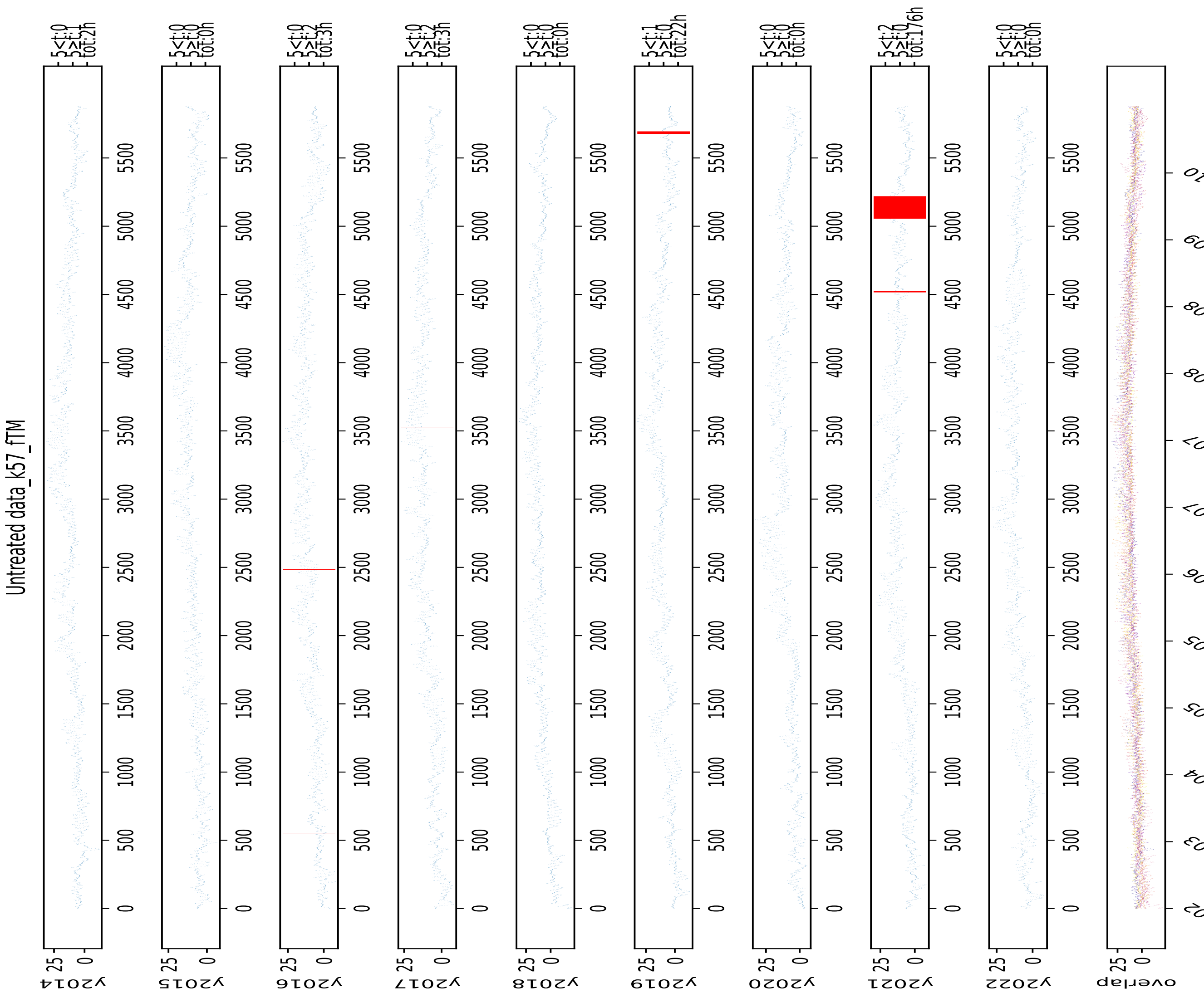


Figure 134: Visual representation of missing values at station 57 from 2014 to 2022 at the parameter ”TM” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

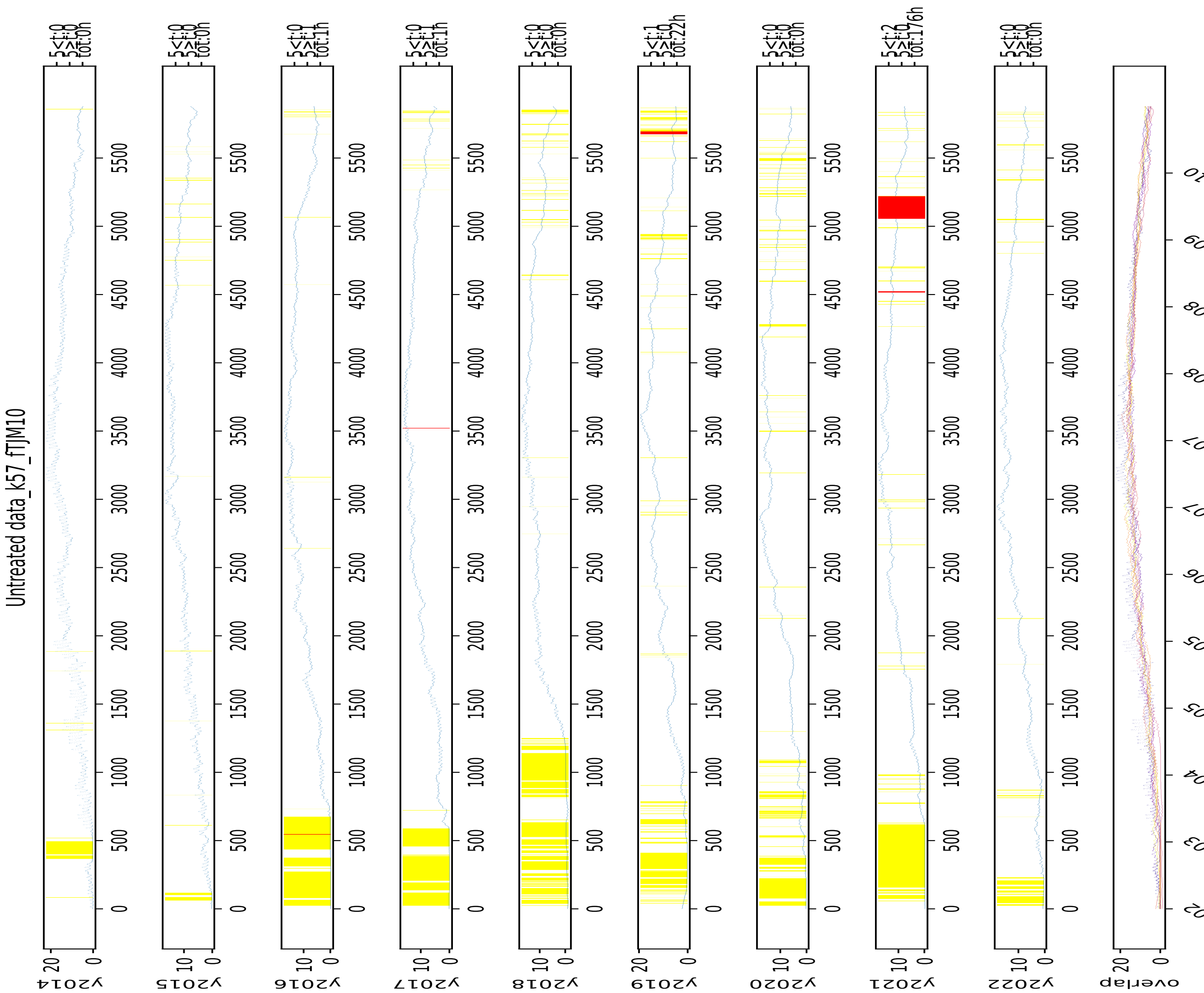


Figure 135: Visual representation of missing values at station 57 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

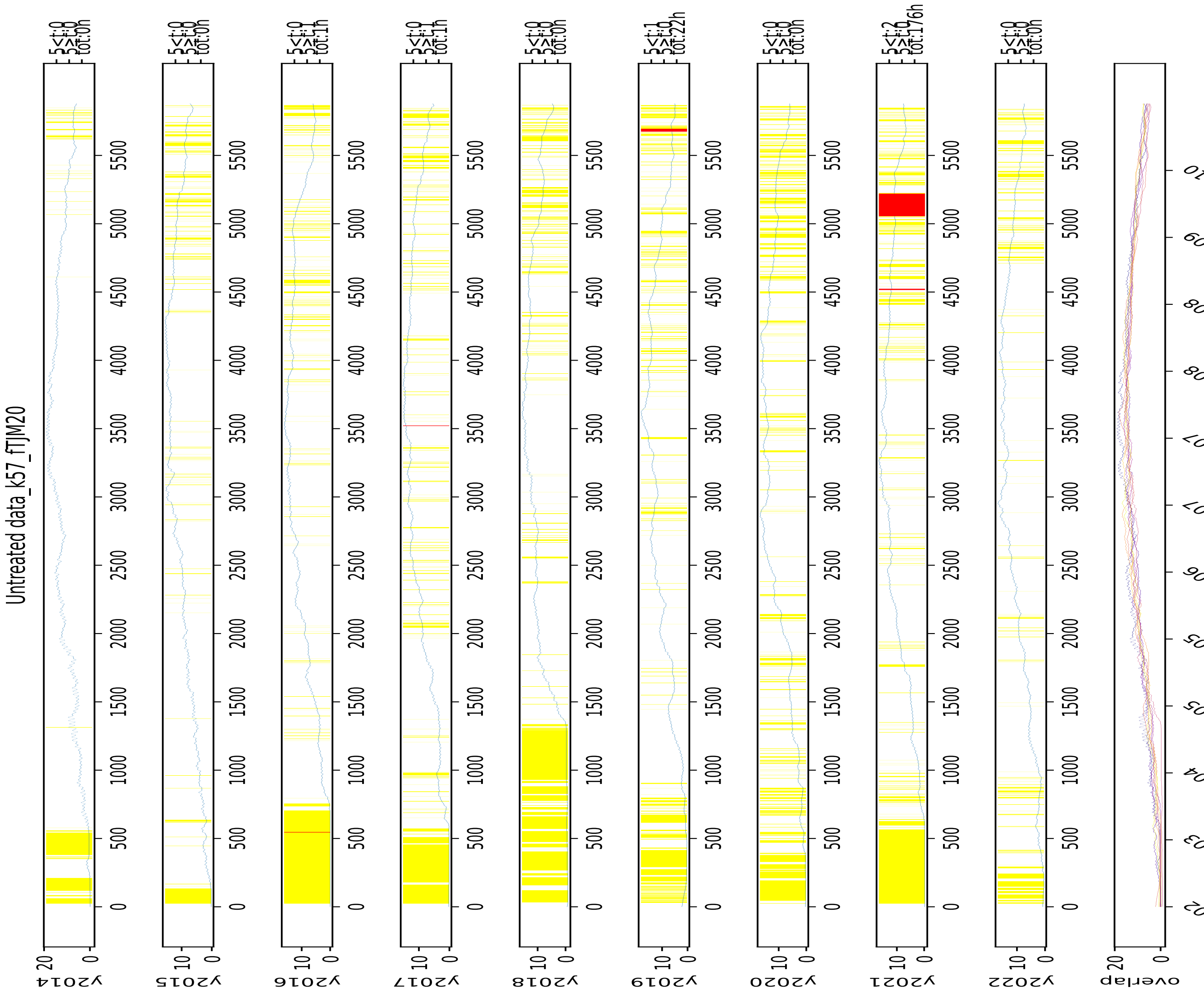


Figure 136: Visual representation of missing values at station 57 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

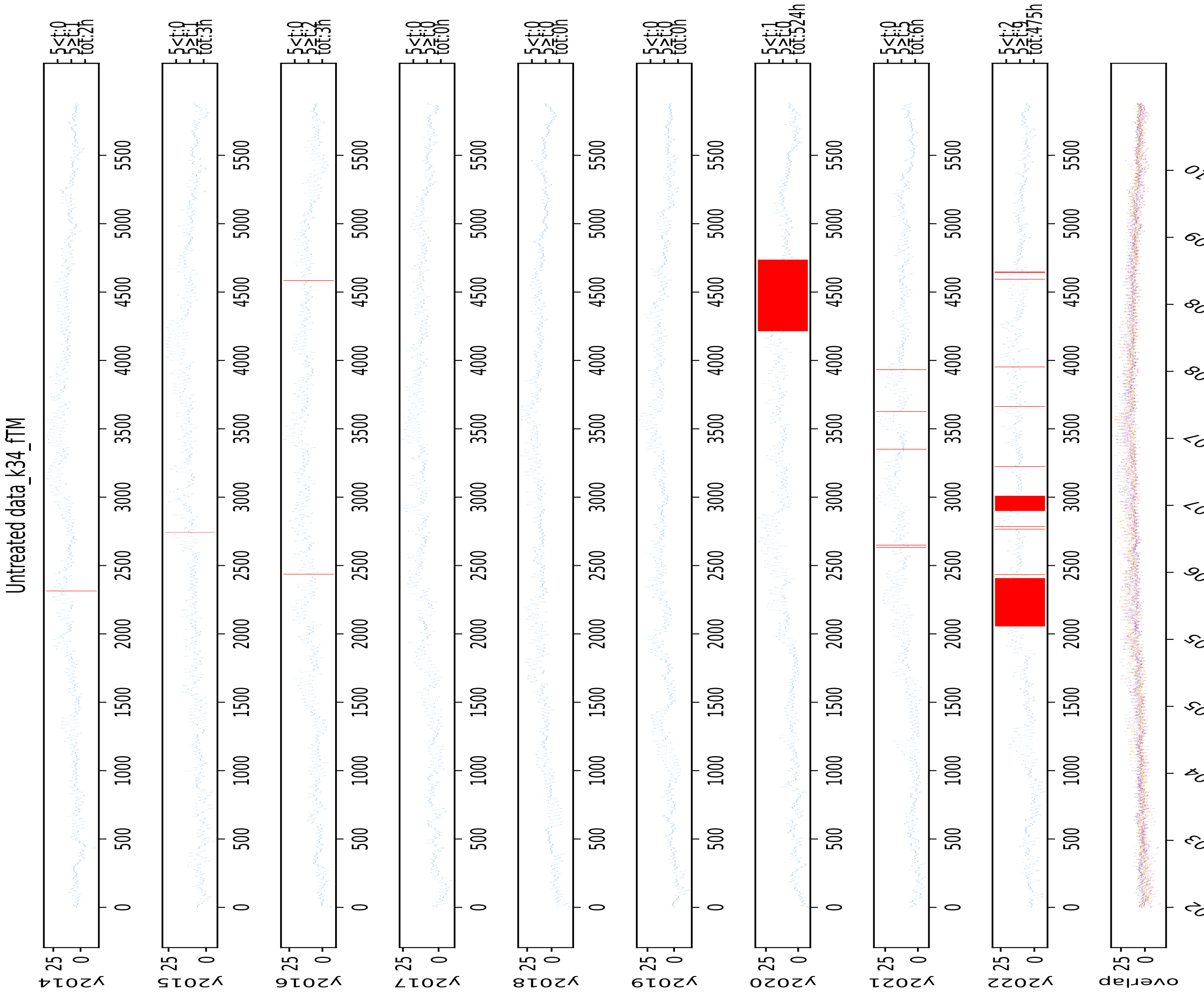


Figure 137: Visual representation of missing values at station 34 from 2014 to 2022 at the parameter ”TM” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

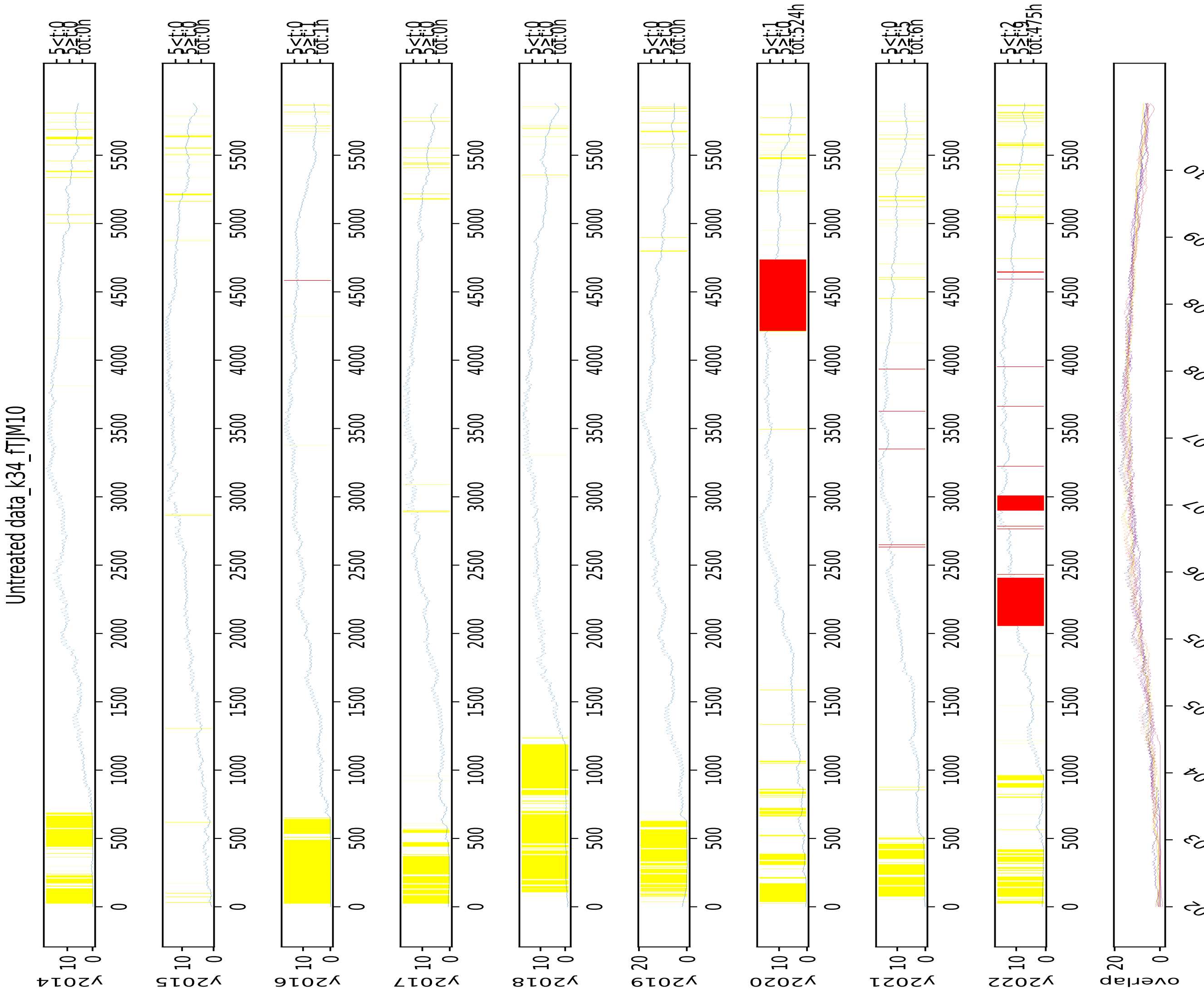


Figure 138: Visual representation of missing values at station 34 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

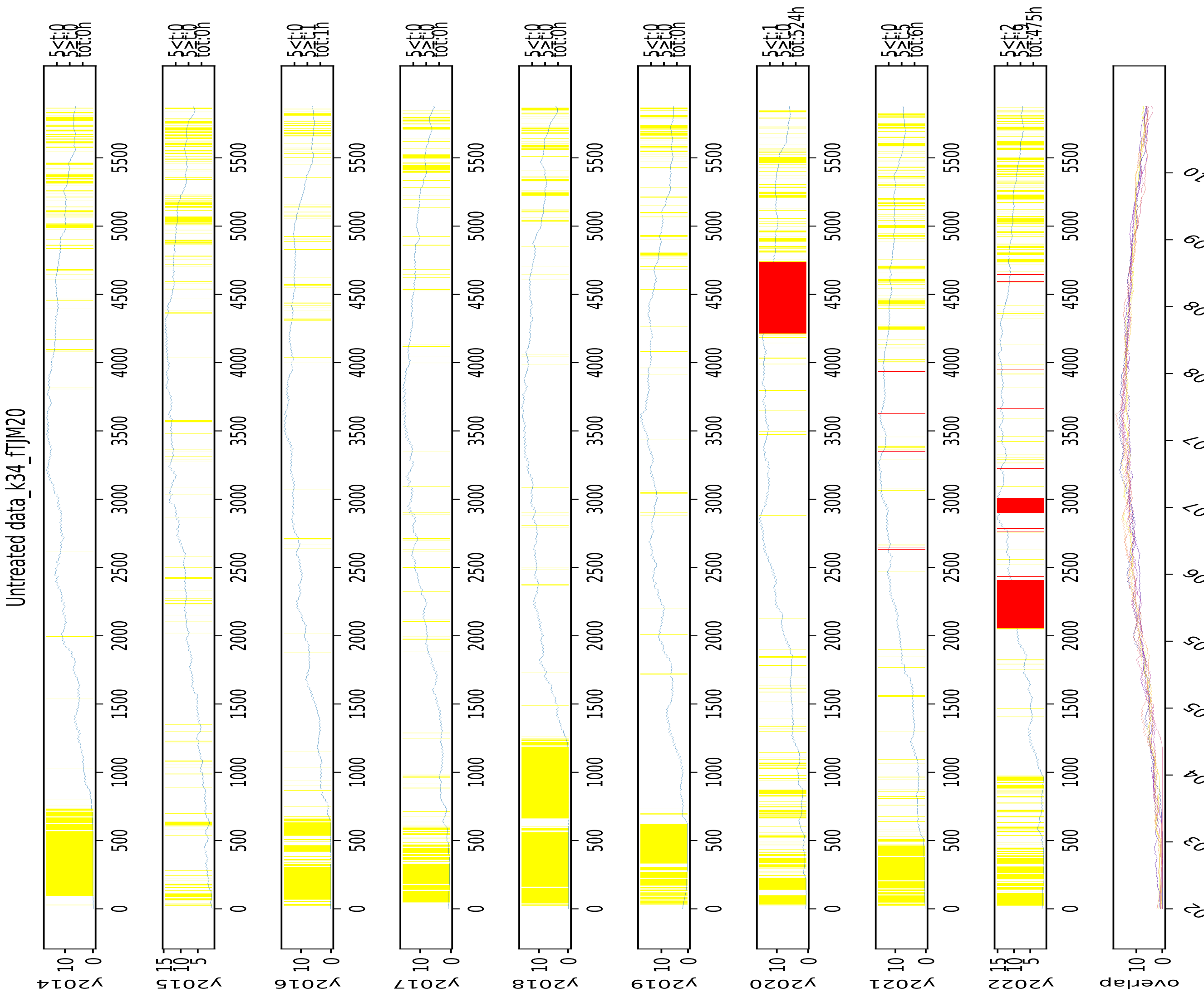


Figure 139: Visual representation of missing values at station 34 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

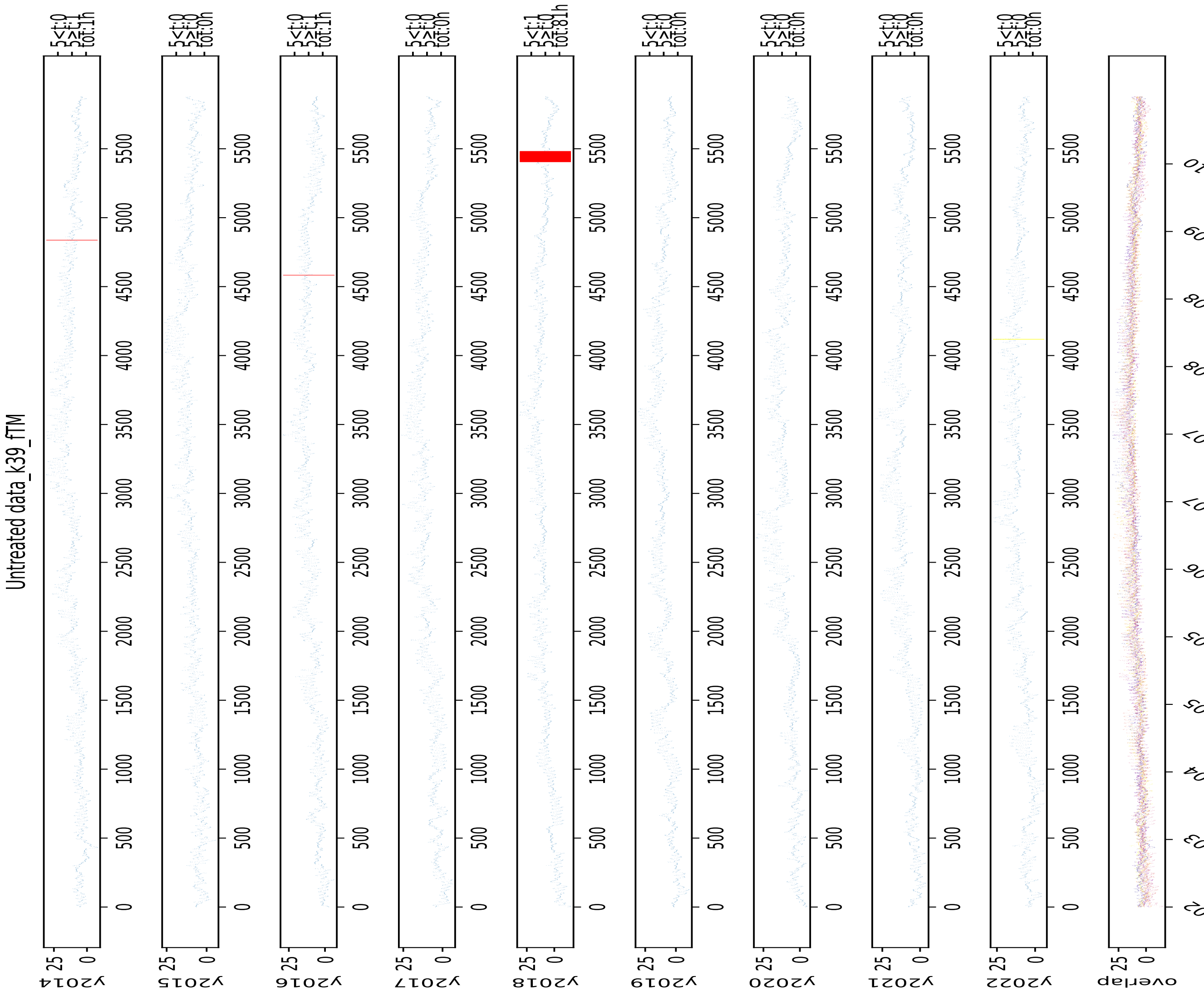


Figure 140: Visual representation of missing values at station 39 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

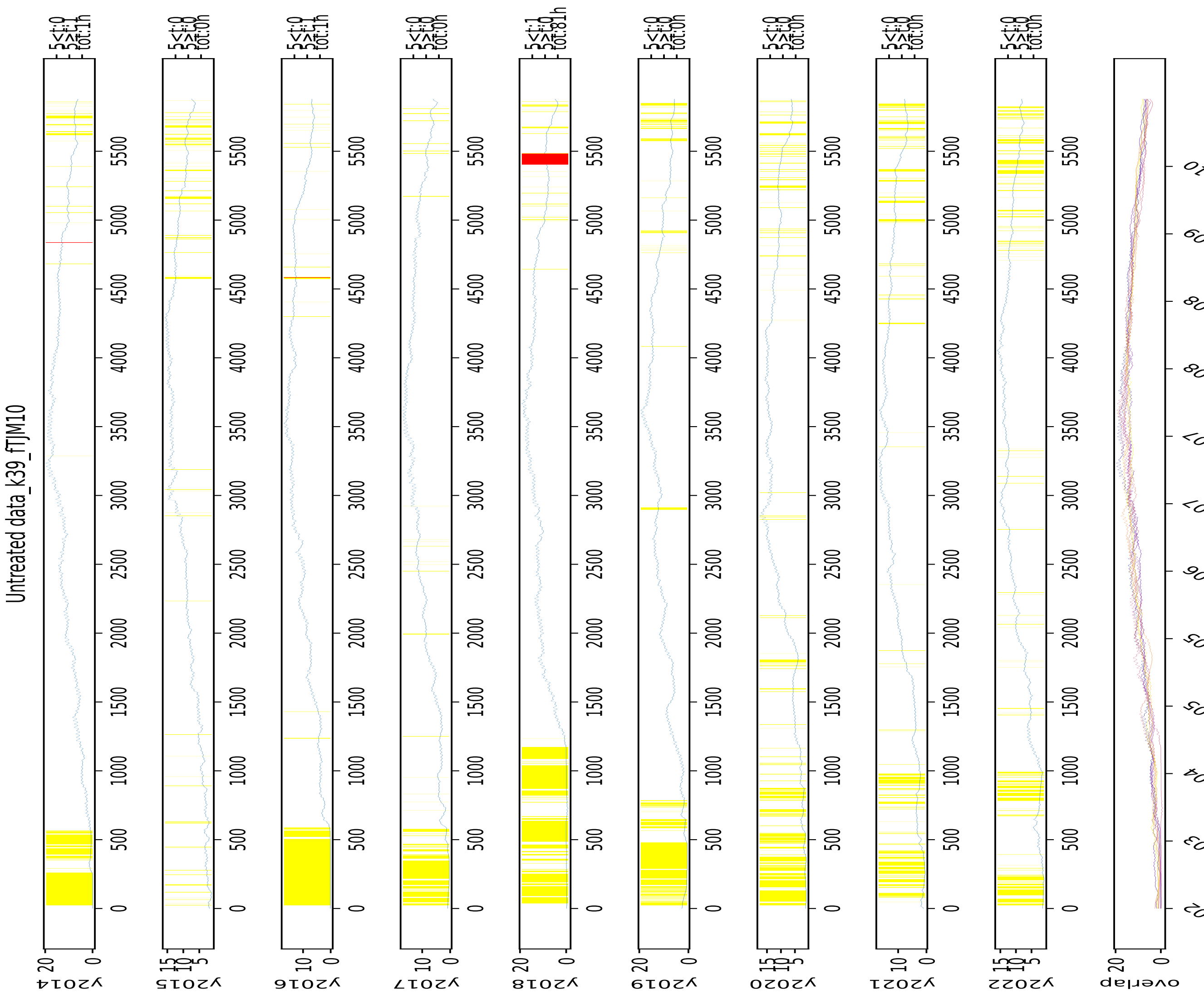


Figure 141: Visual representation of missing values at station 39 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

Figure 142: Visual representation of missing values at station 39 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

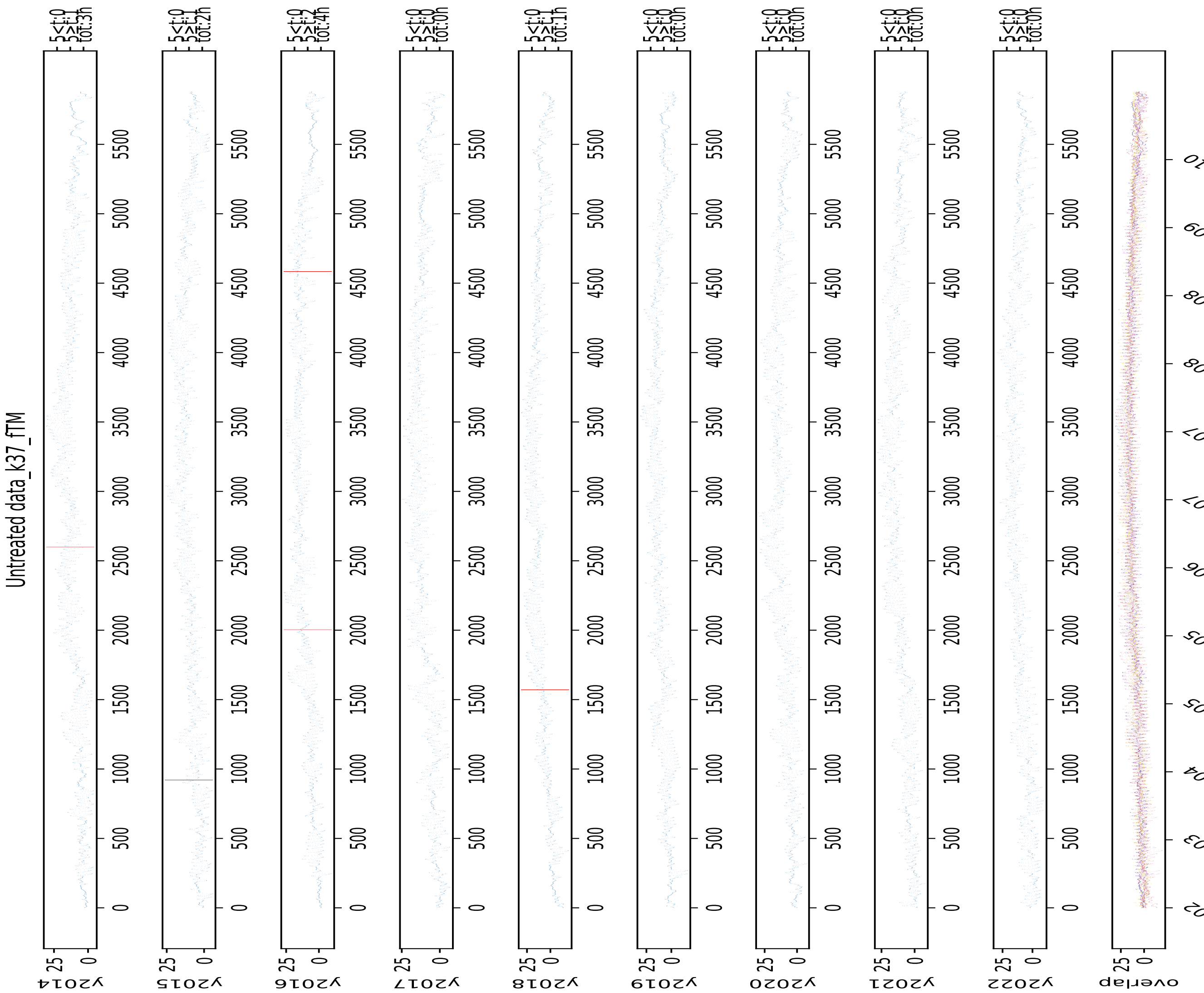


Figure 143: Visual representation of missing values at station 37 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

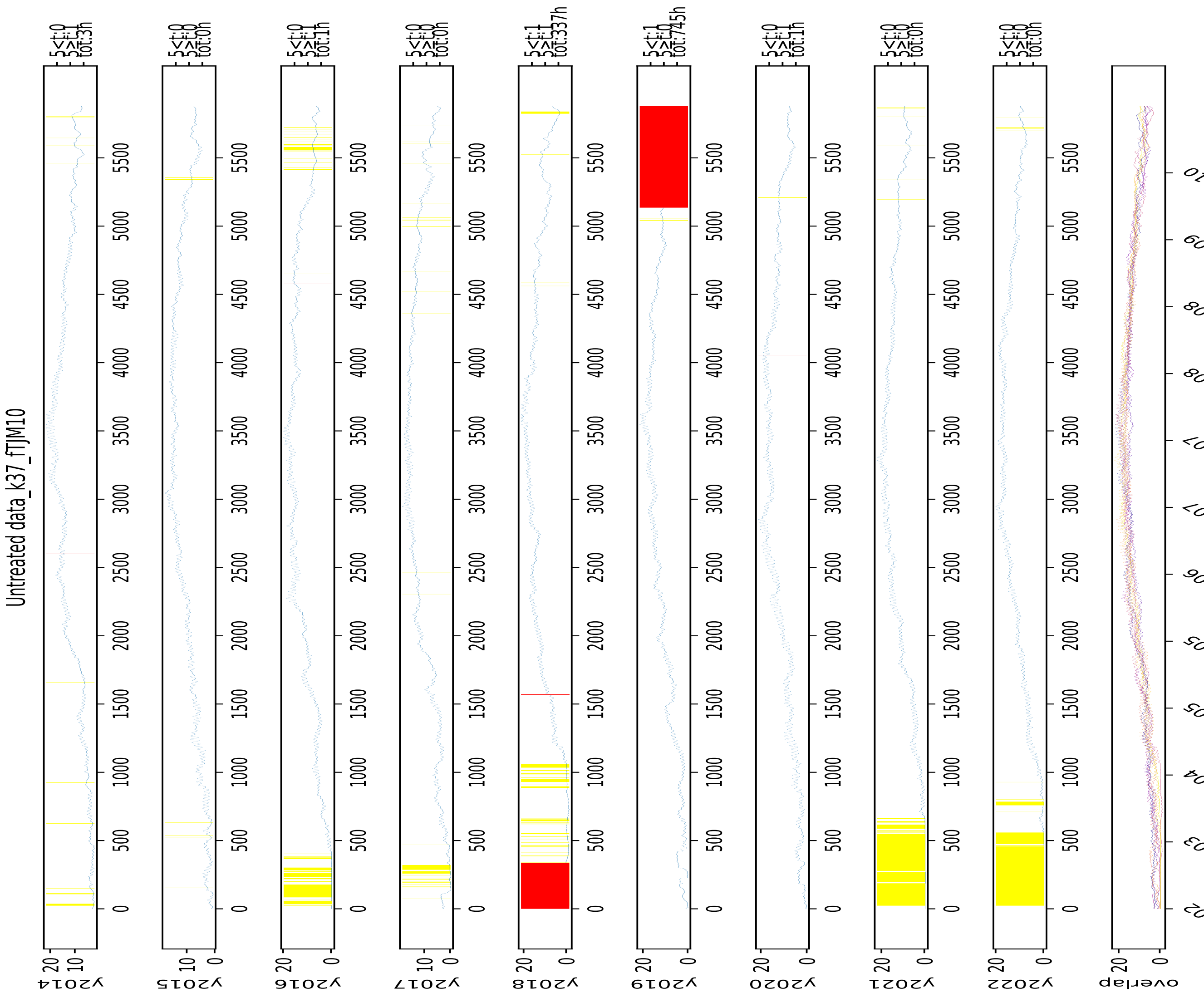


Figure 144: Visual representation of missing values at station 37 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

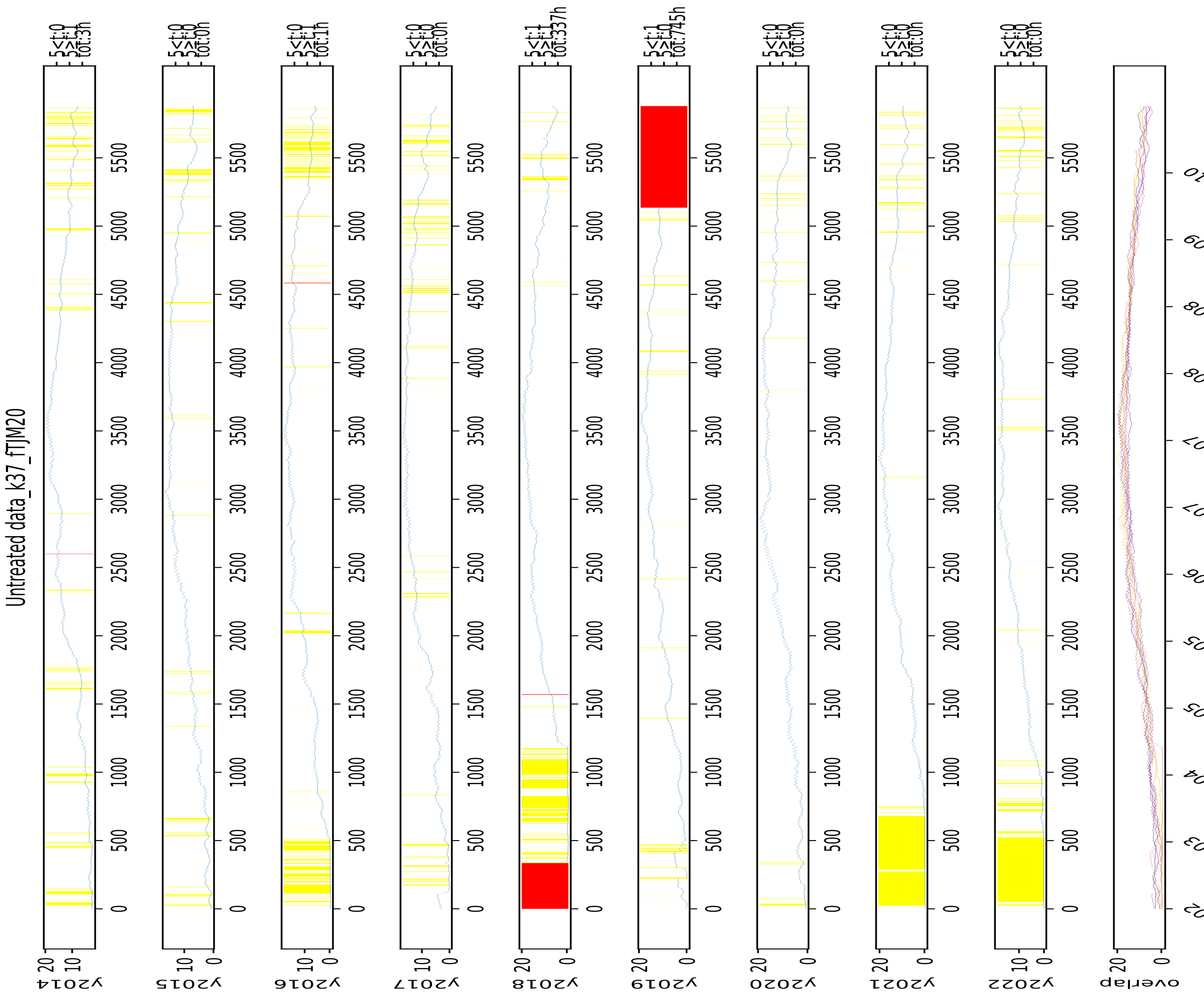


Figure 145: Visual representation of missing values at station 37 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

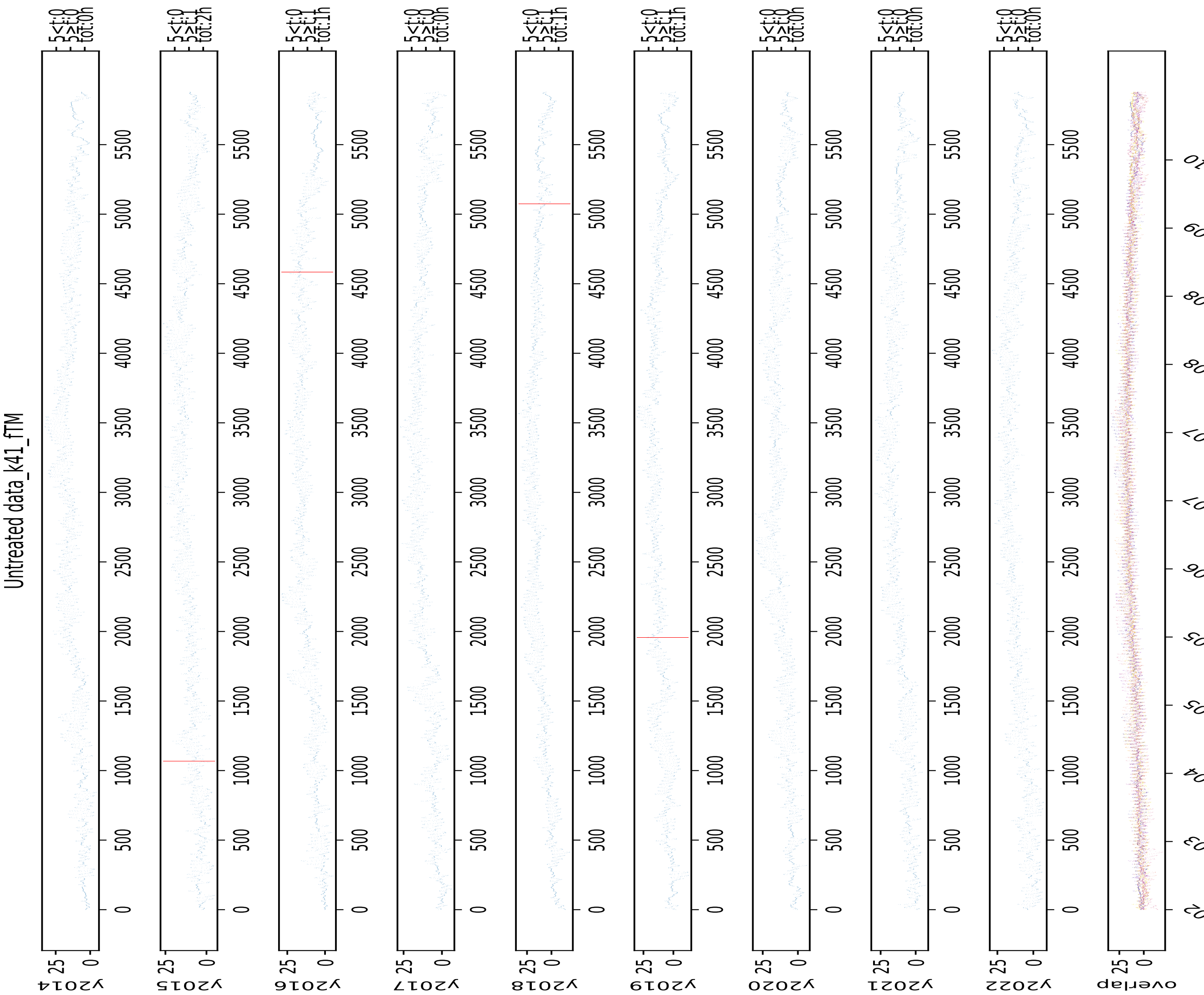


Figure 146: Visual representation of missing values at station 41 from 2014 to 2022 at the parameter ”TM” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

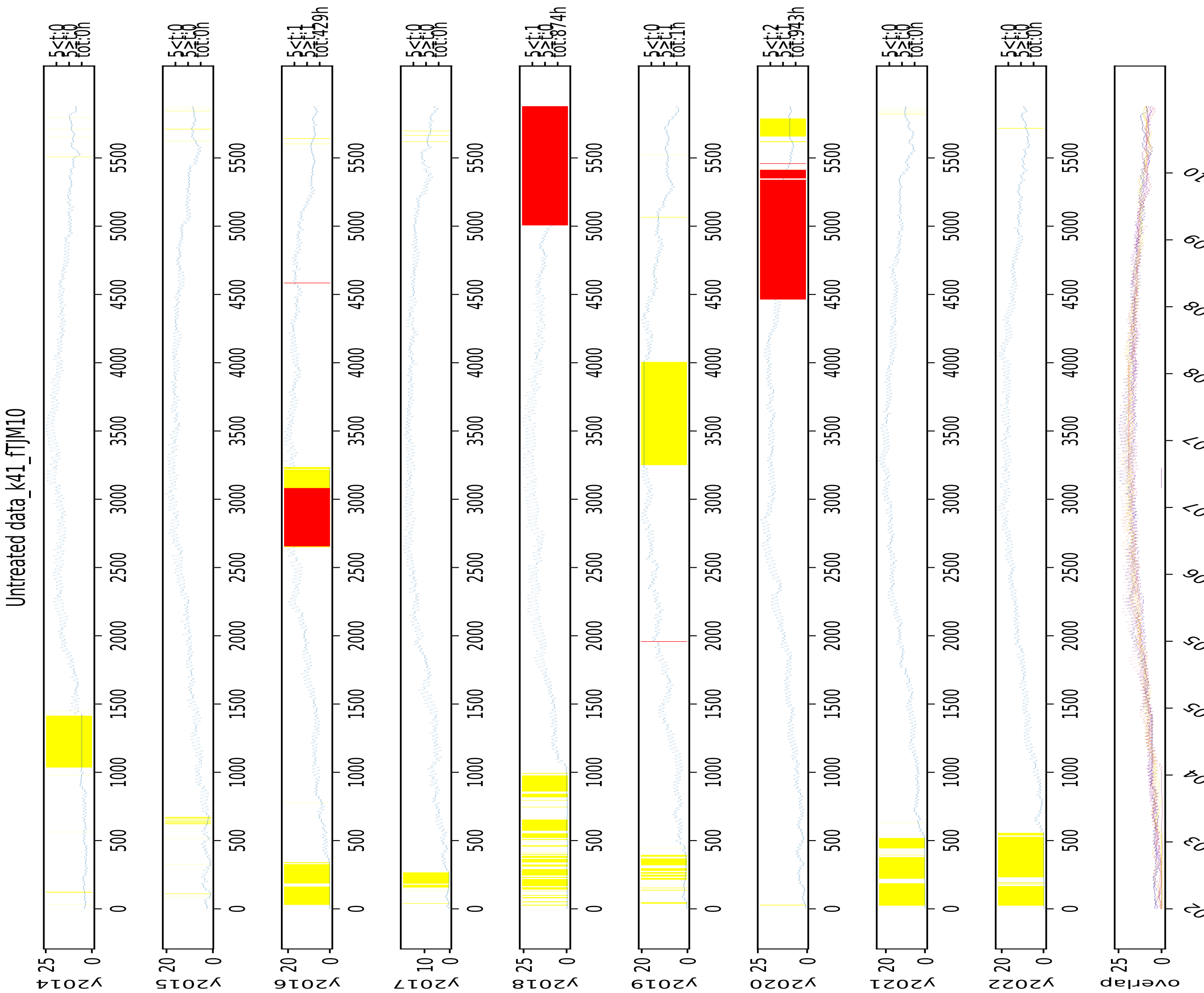


Figure 147: Visual representation of missing values at station 41 from 2014 to 2022 at the parameter ”TJM10” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

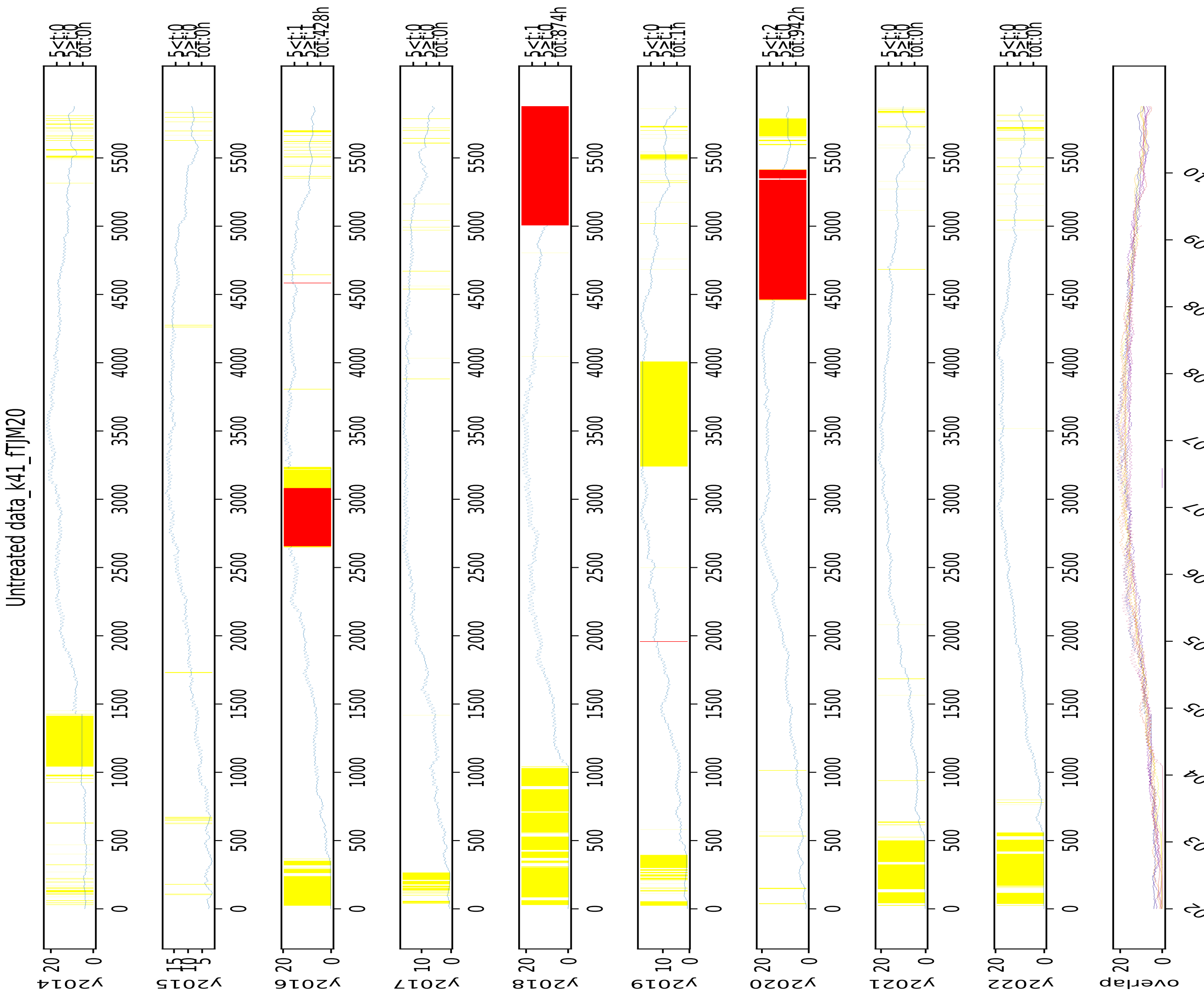


Figure 148: Visual representation of missing values at station 41 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

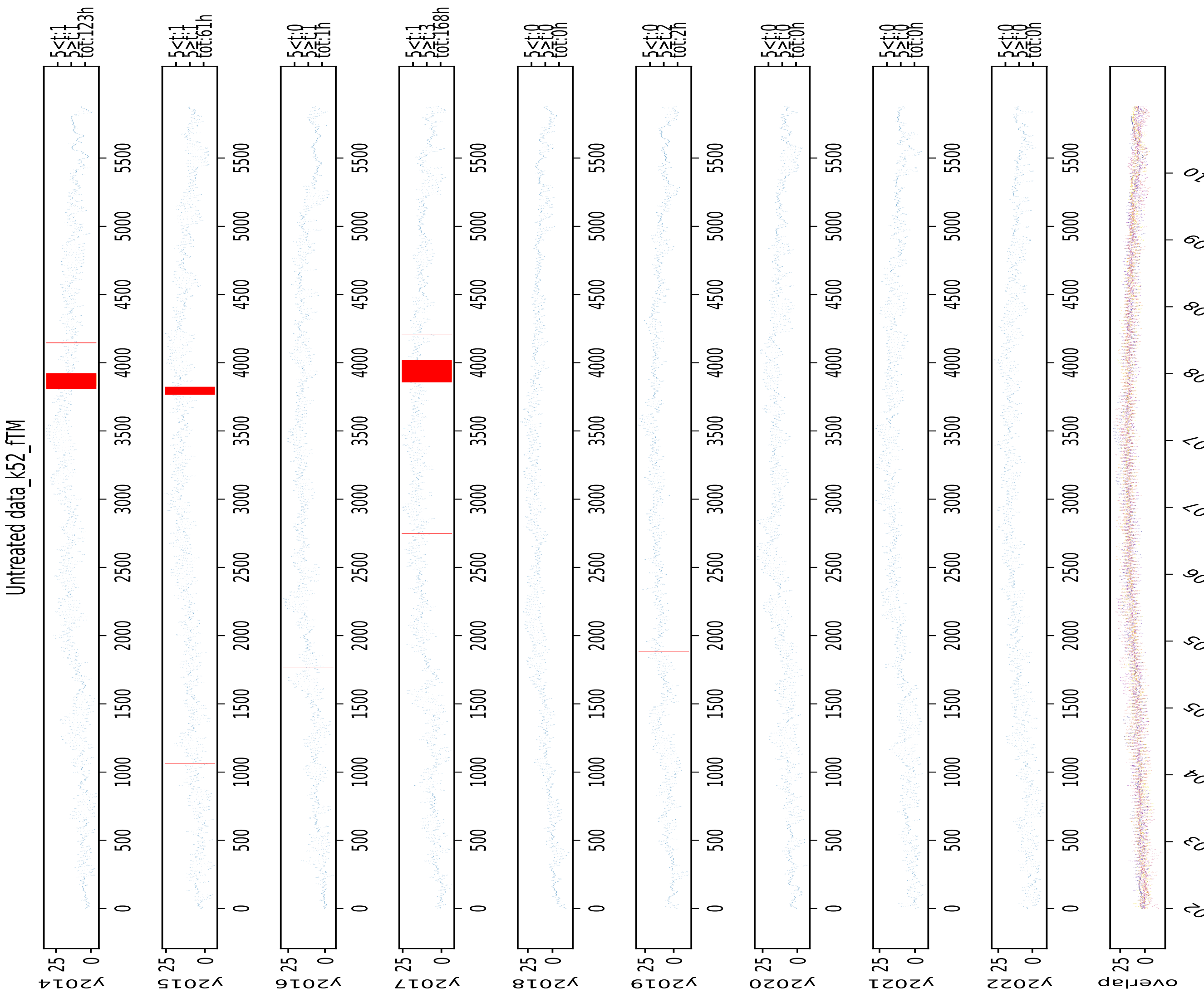


Figure 149: Visual representation of missing values at station 52 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

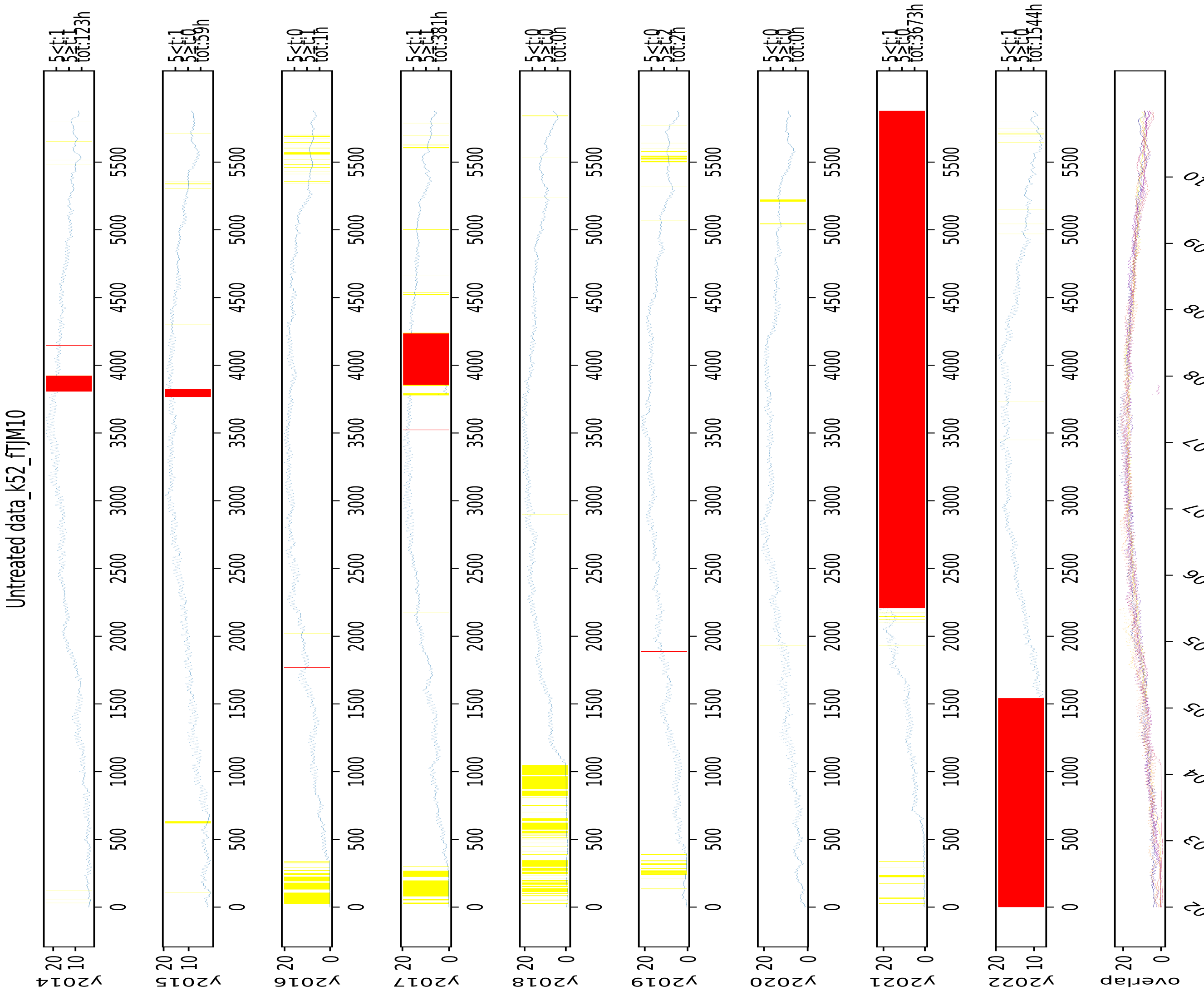


Figure 150: Visual representation of missing values at station 52 from 2014 to 2022 at the parameter ”TJM10” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

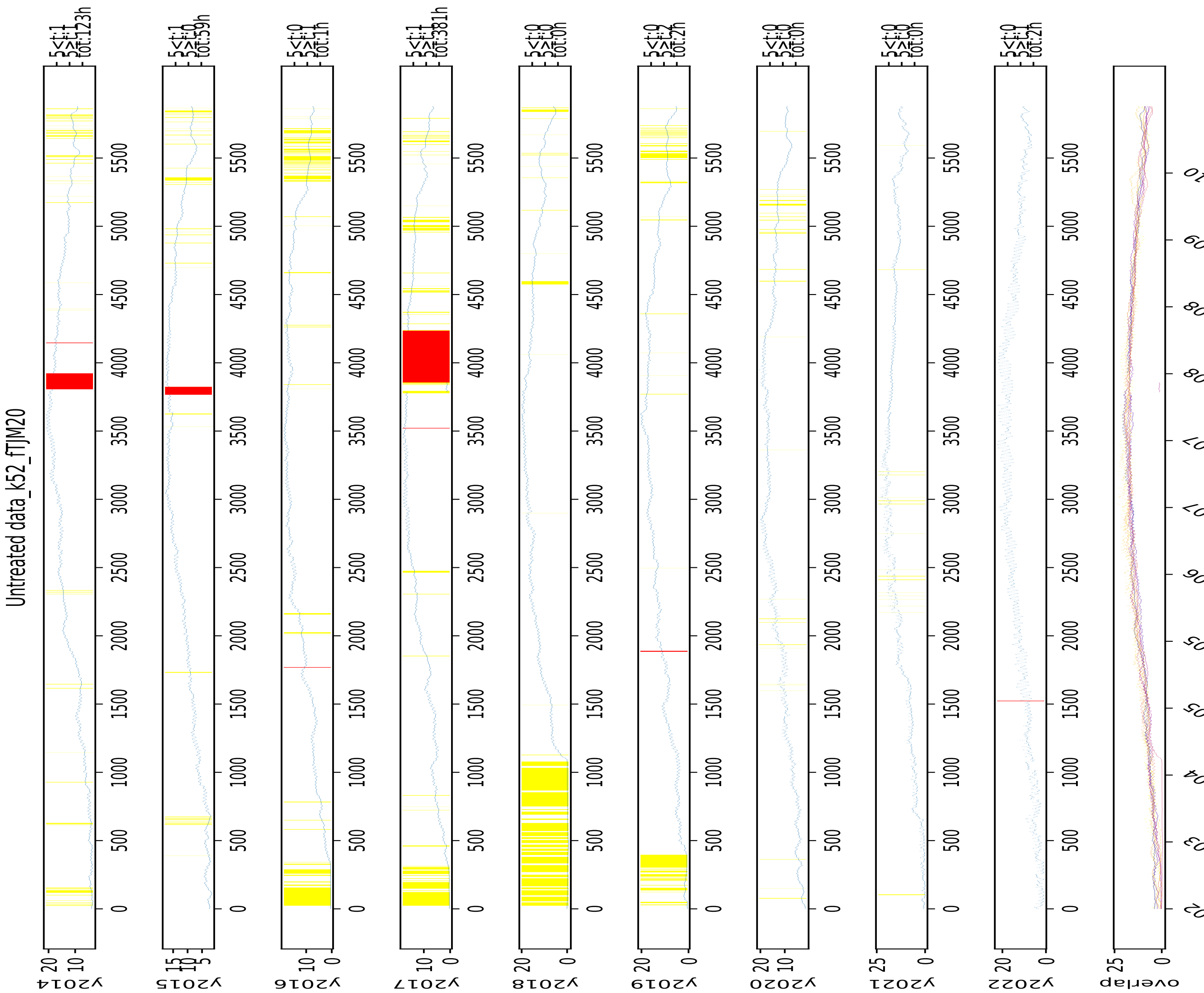


Figure 151: Visual representation of missing values at station 52 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

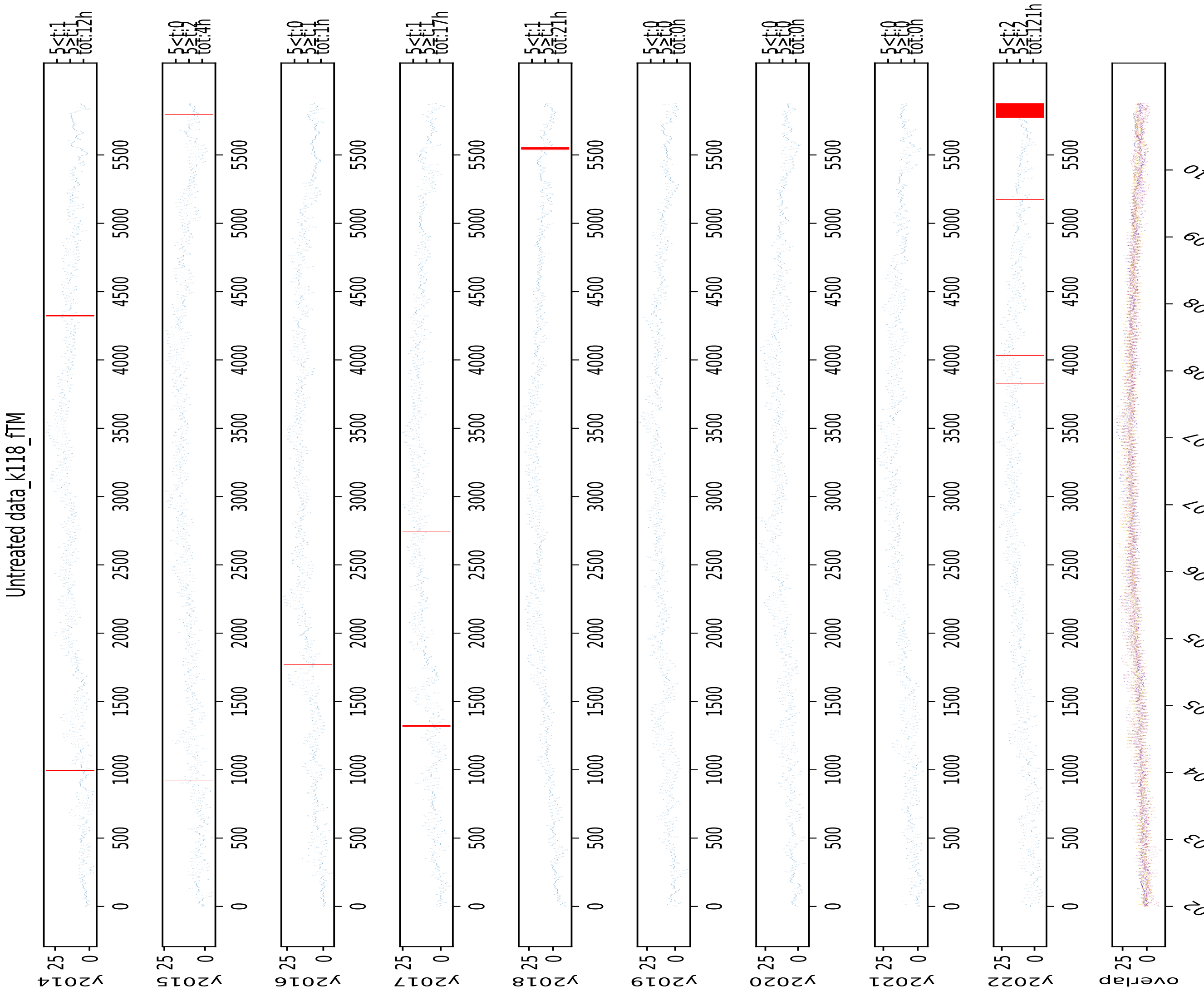


Figure 152: Visual representation of missing values at station 118 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

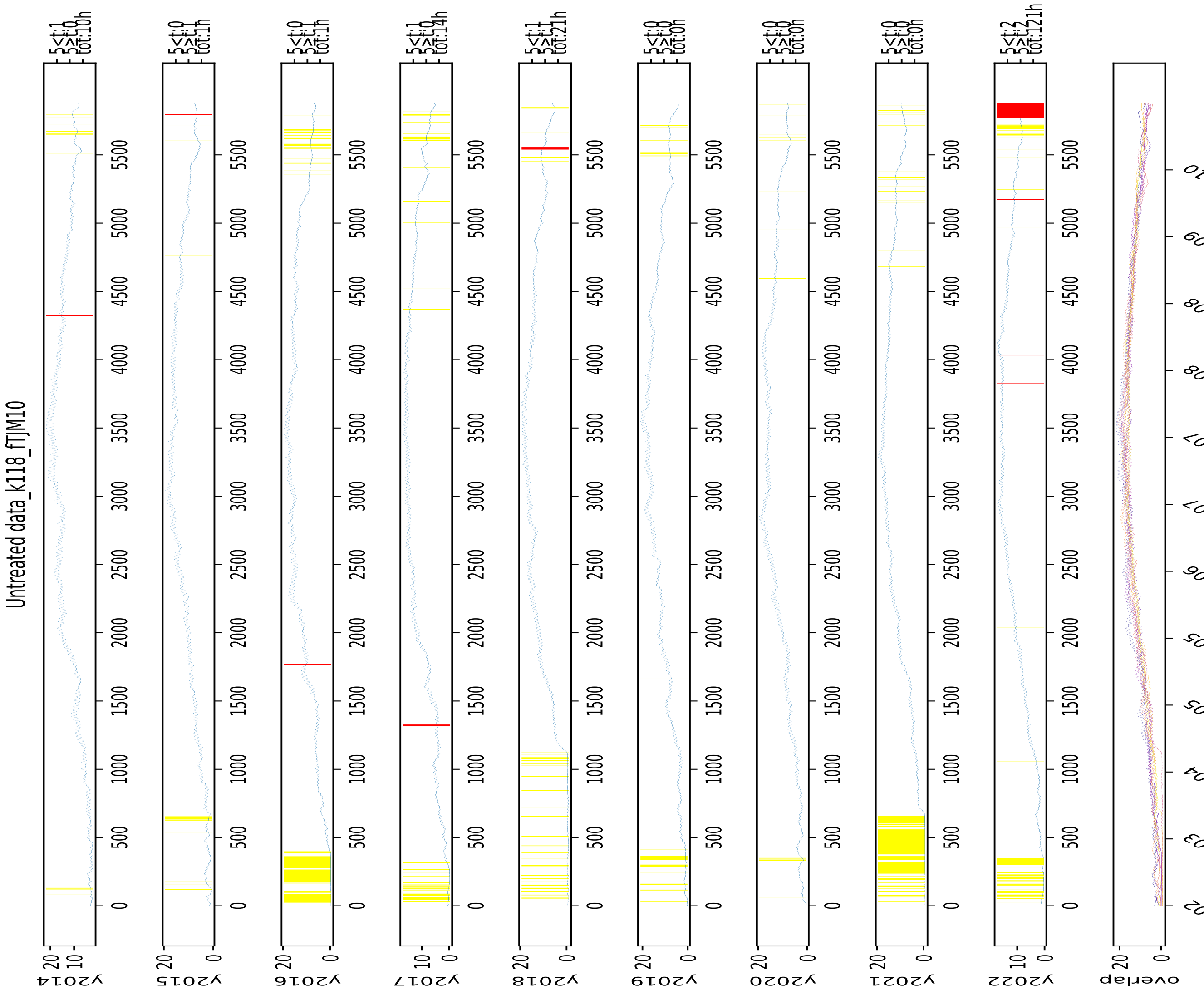


Figure 153: Visual representation of missing values at station 118 from 2014 to 2022 at the parameter ”TJM10” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

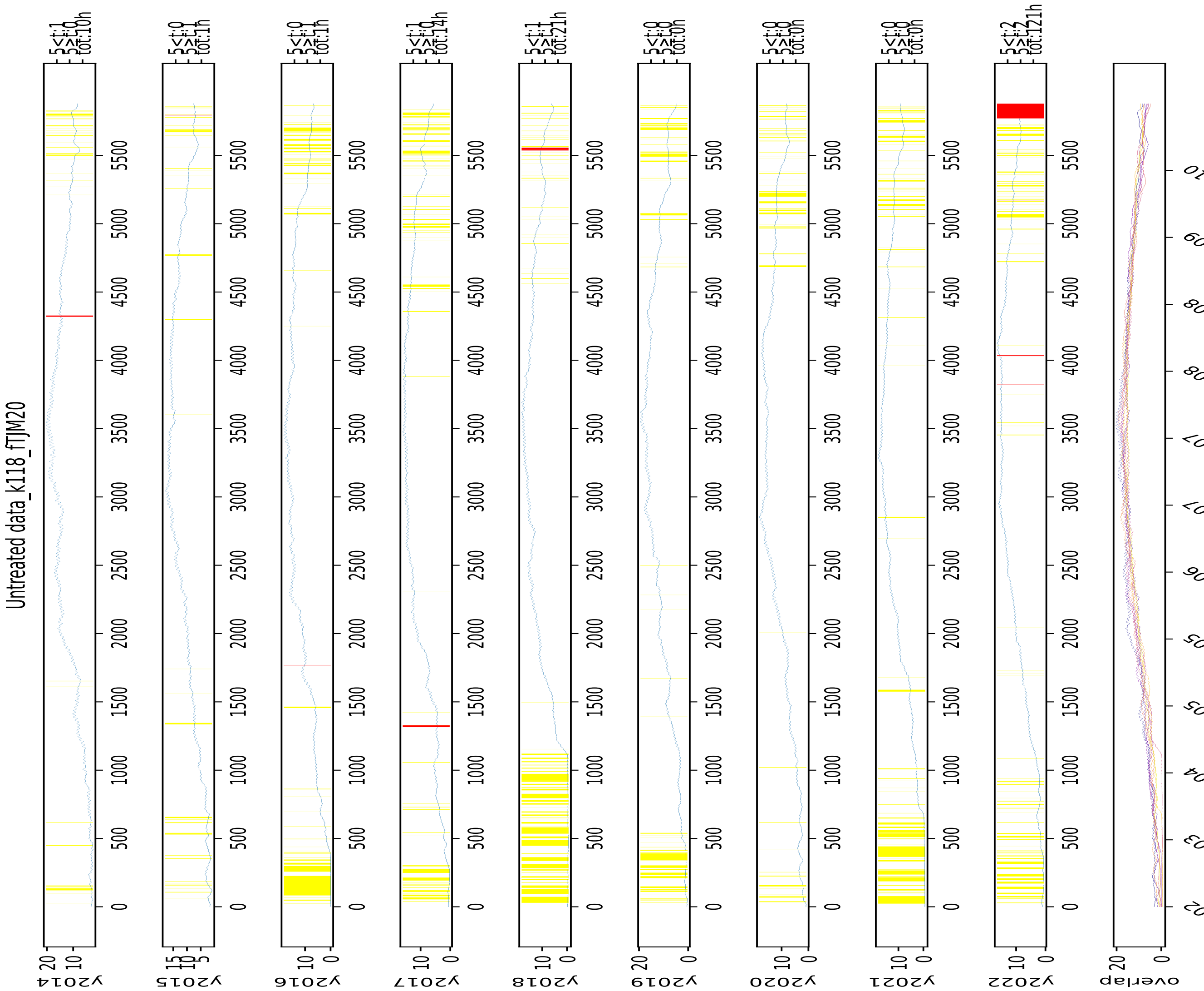


Figure 154: Visual representation of missing values at station 118 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

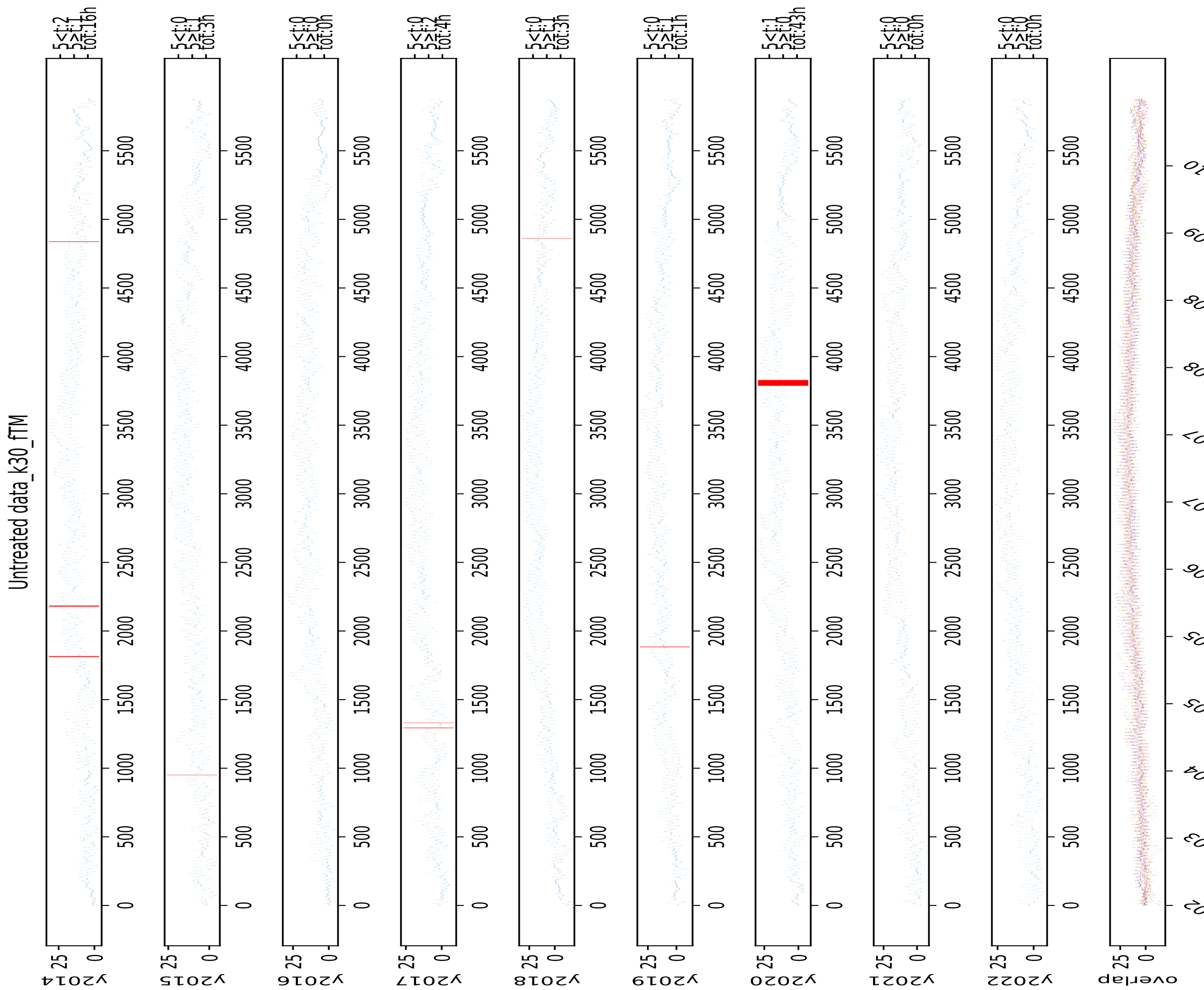


Figure 155: Visual representation of missing values at station 30 from 2014 to 2022 at the parameter ”TM” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

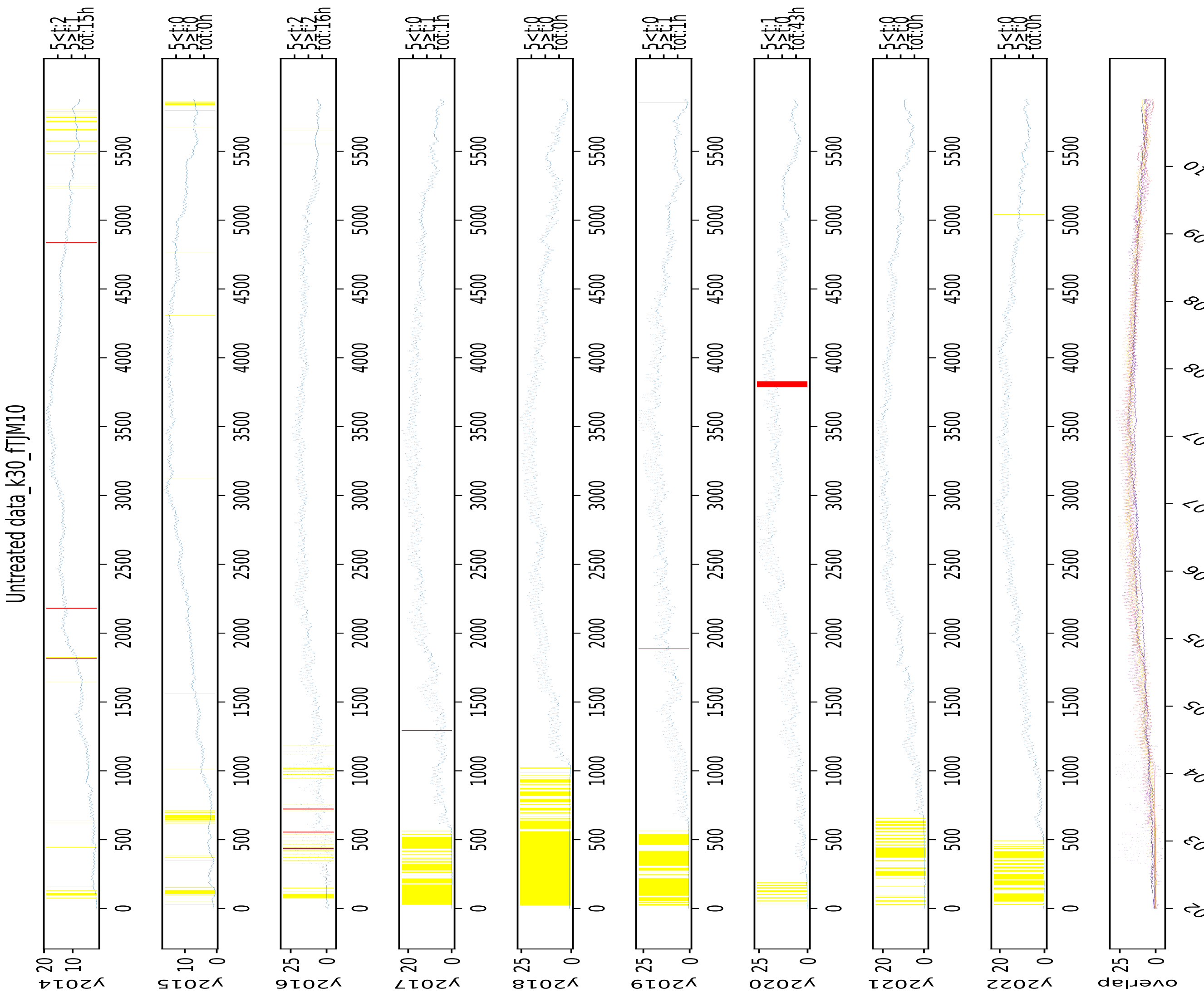


Figure 156: Visual representation of missing values at station 30 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

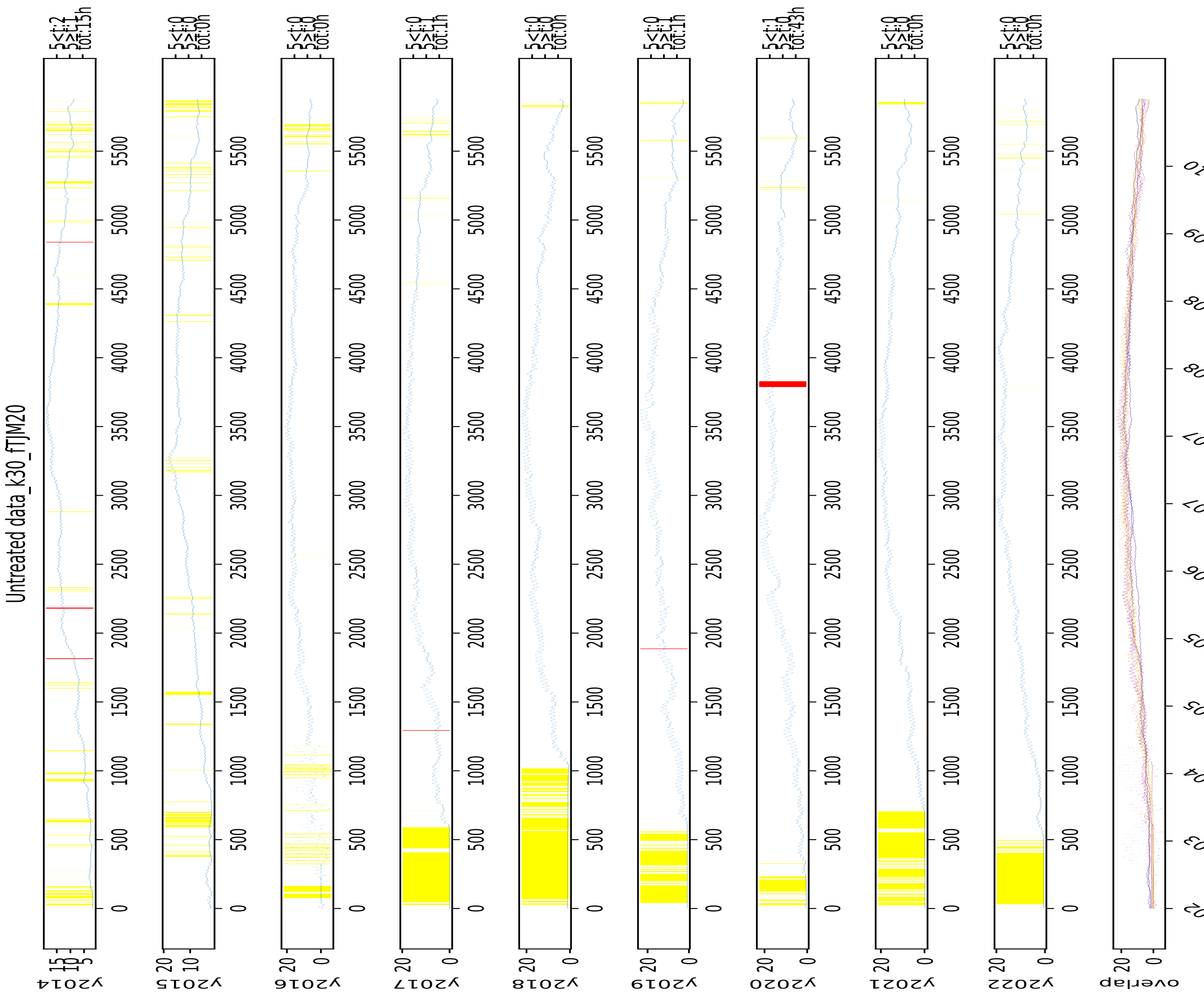


Figure 157: Visual representation of missing values at station 30 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

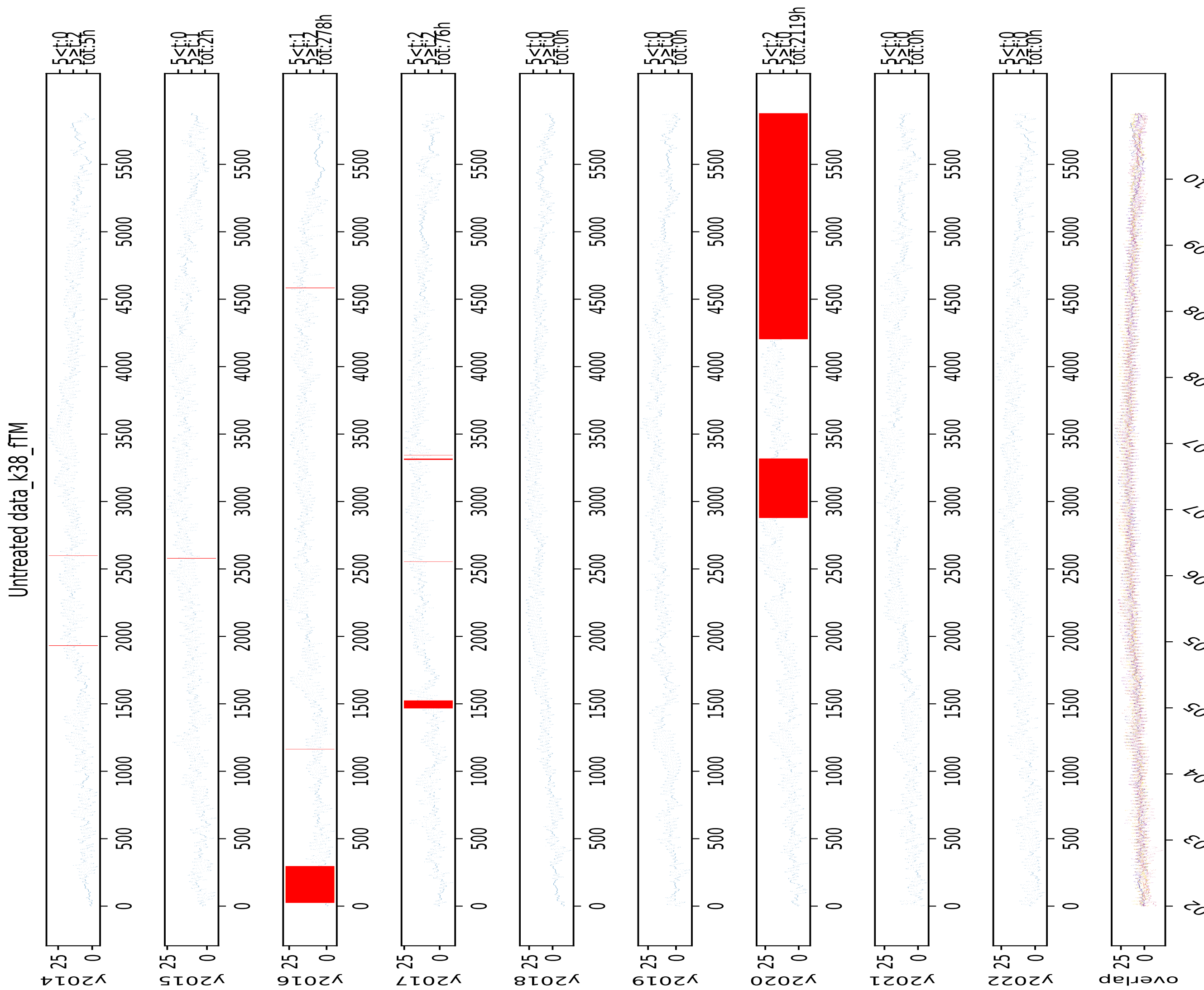


Figure 158: Visual representation of missing values at station 38 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

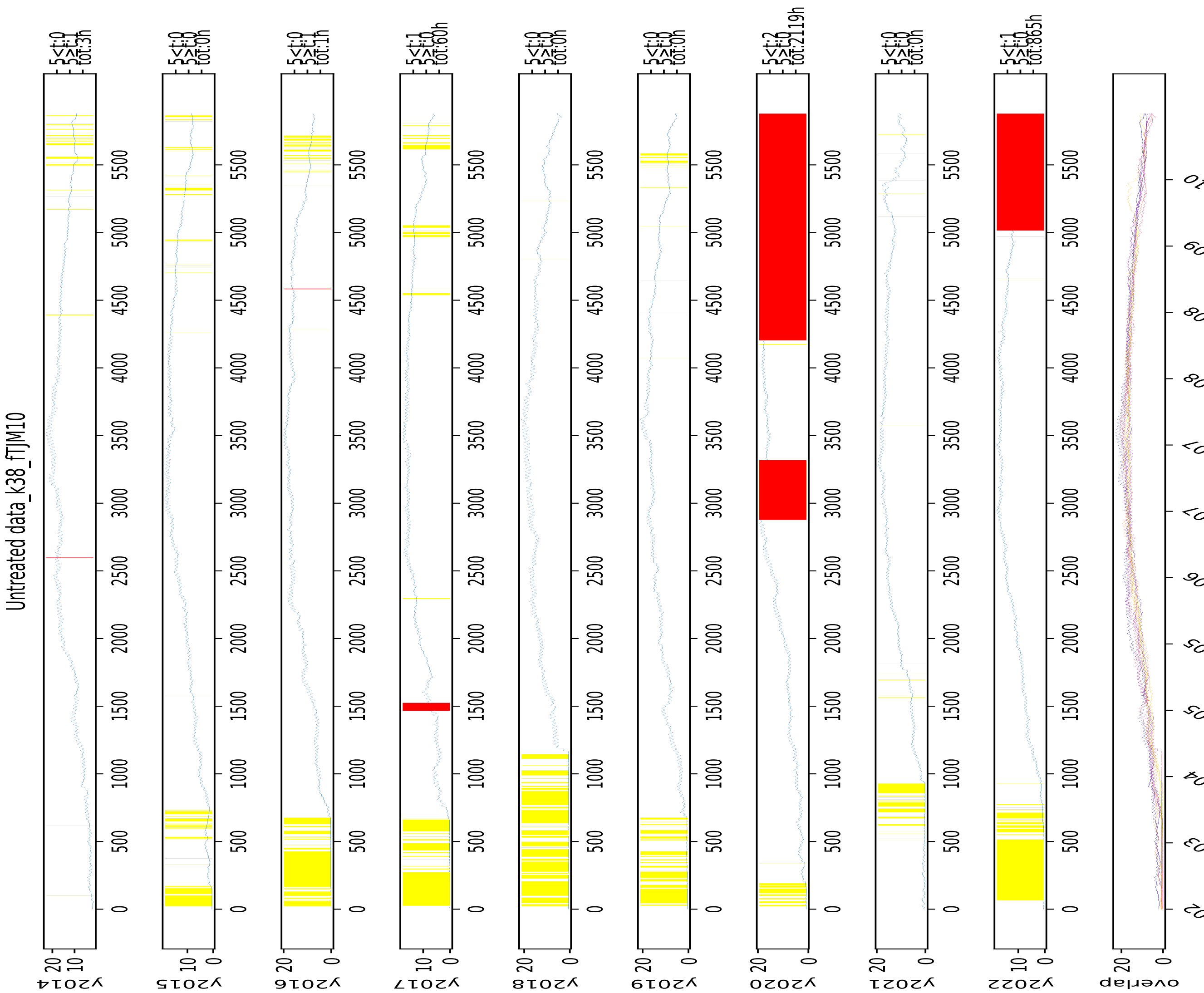


Figure 159: Visual representation of missing values at station 38 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

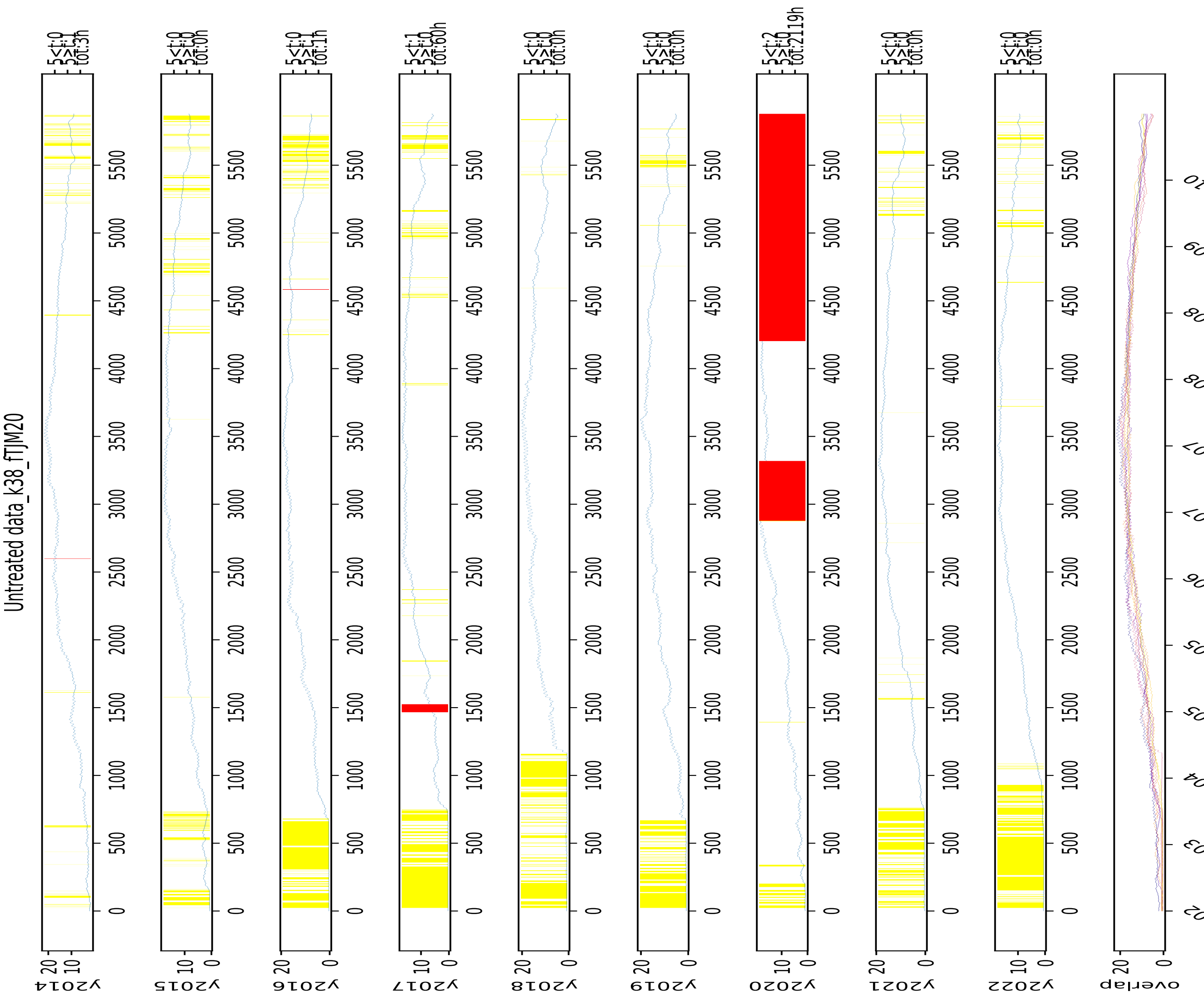


Figure 160: Visual representation of missing values at station 38 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

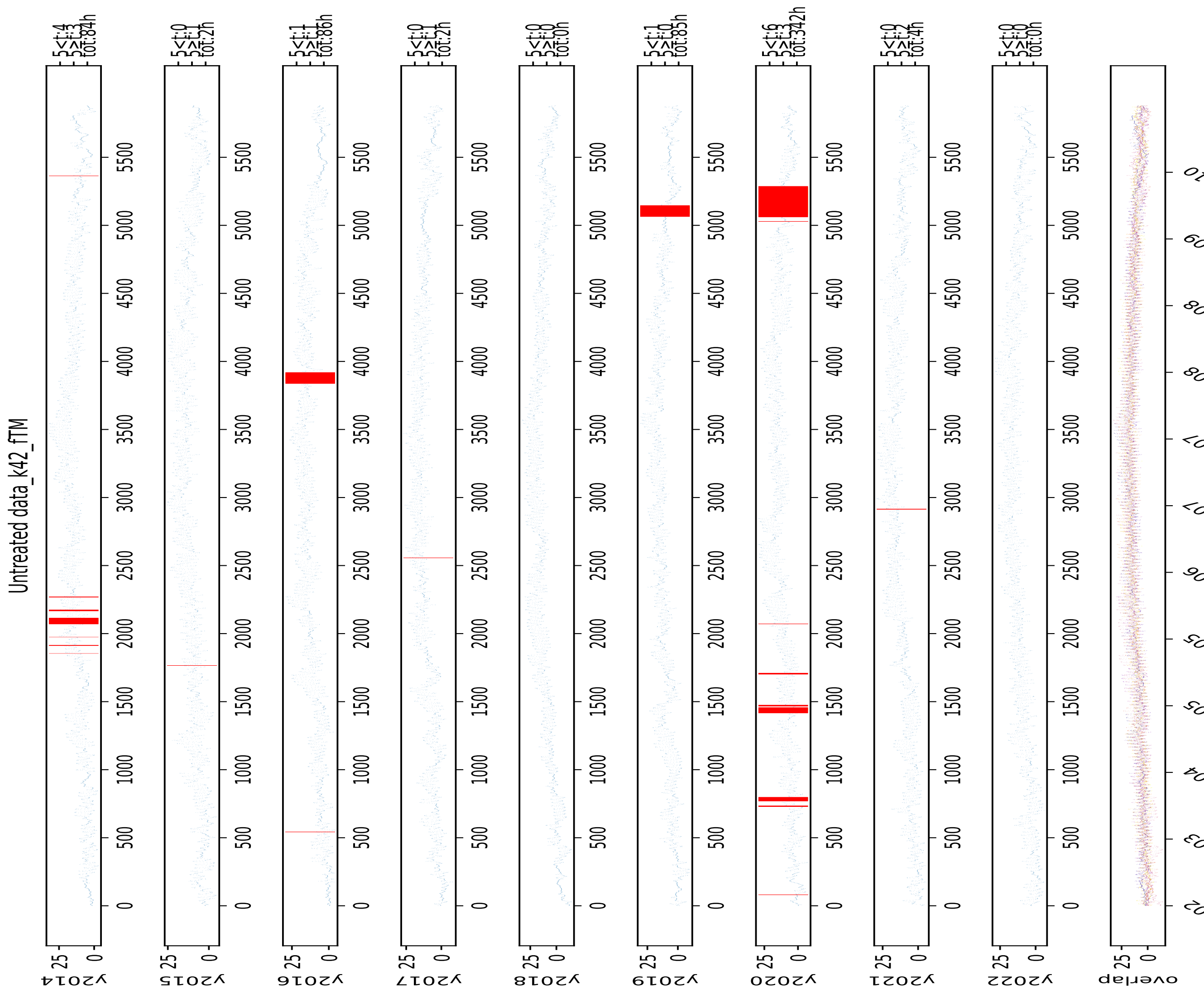


Figure 161: Visual representation of missing values at station 42 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

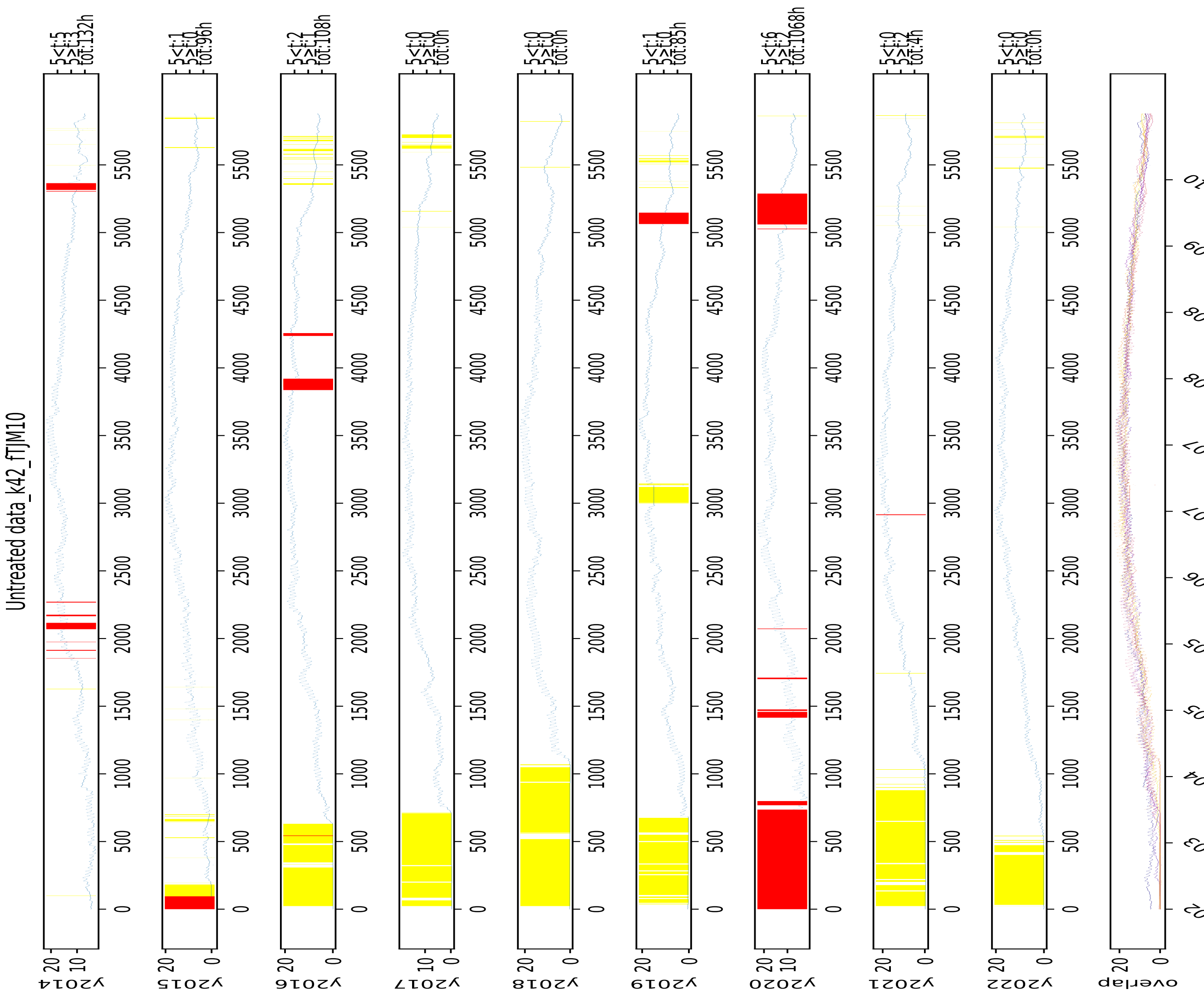


Figure 162: Visual representation of missing values at station 42 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

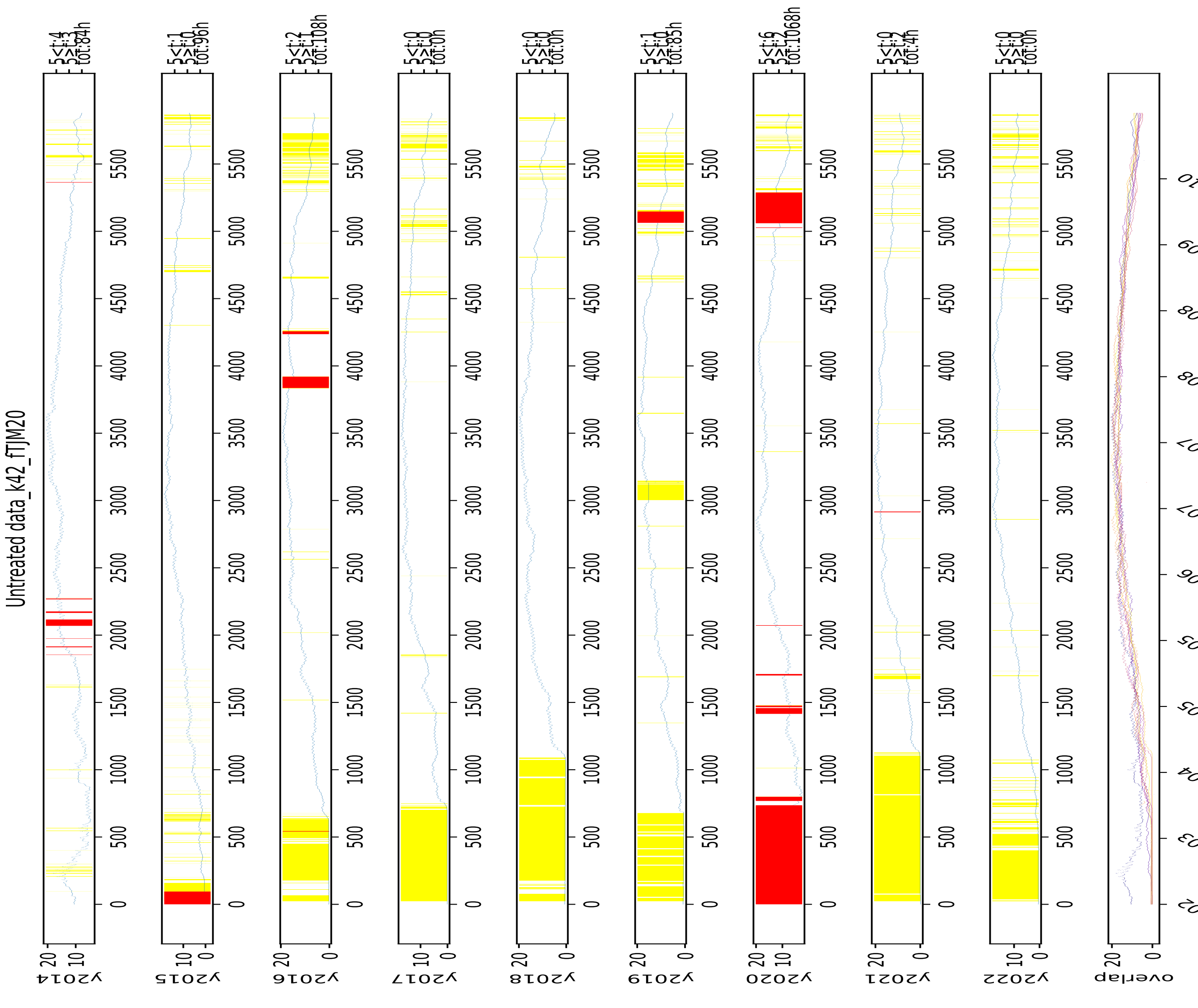


Figure 163: Visual representation of missing values at station 42 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

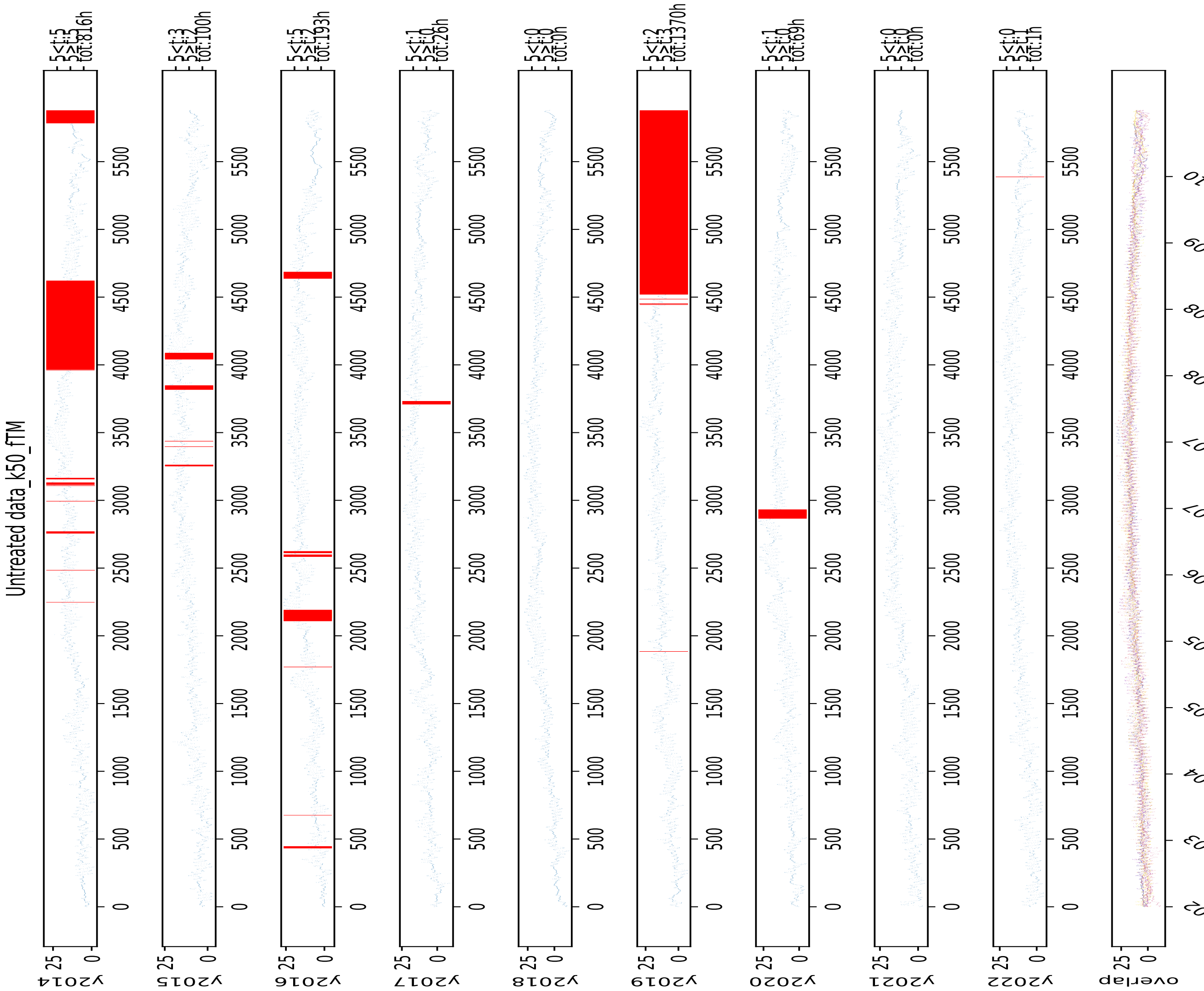


Figure 164: Visual representation of missing values at station 50 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

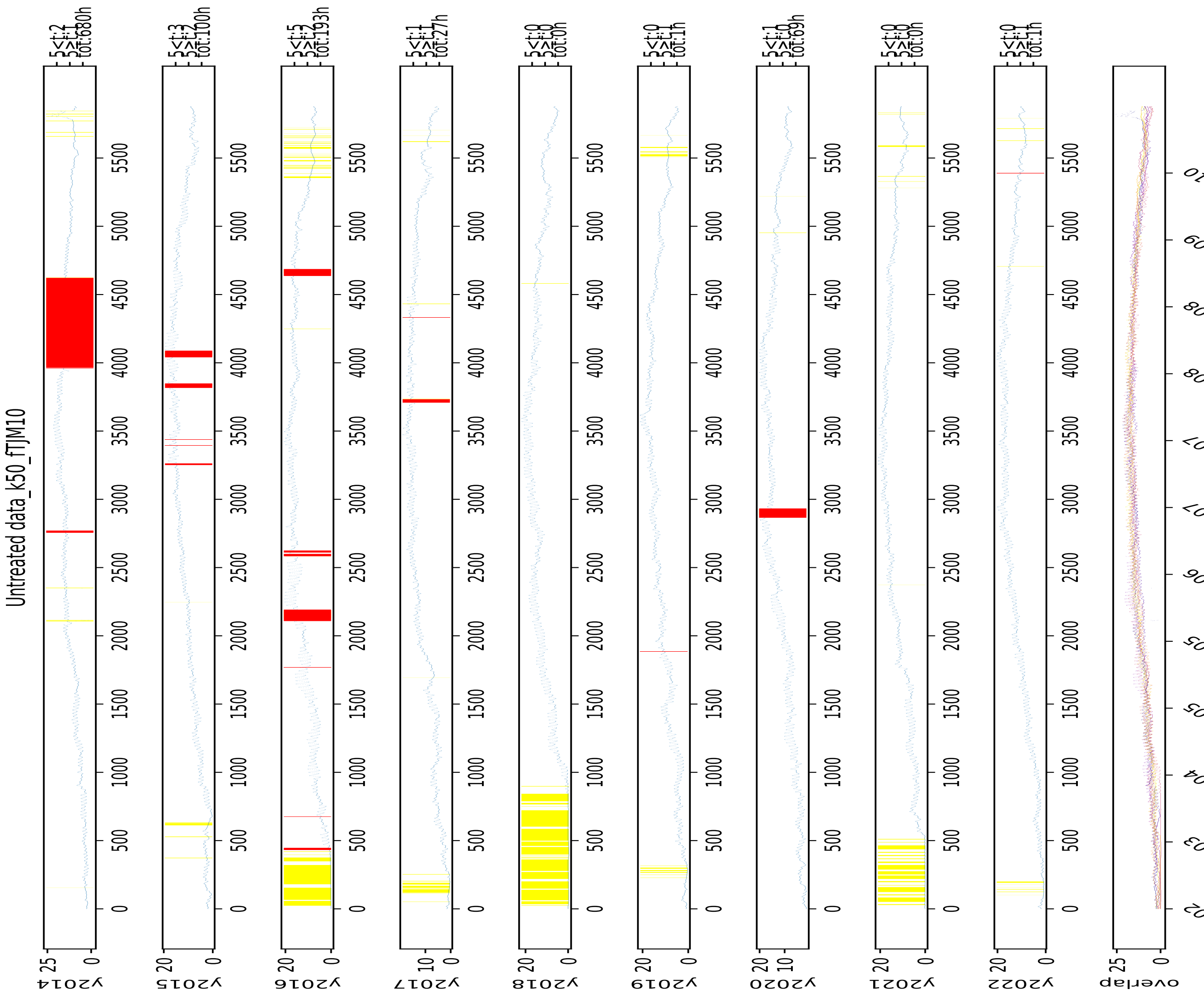


Figure 165: Visual representation of missing values at station 50 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

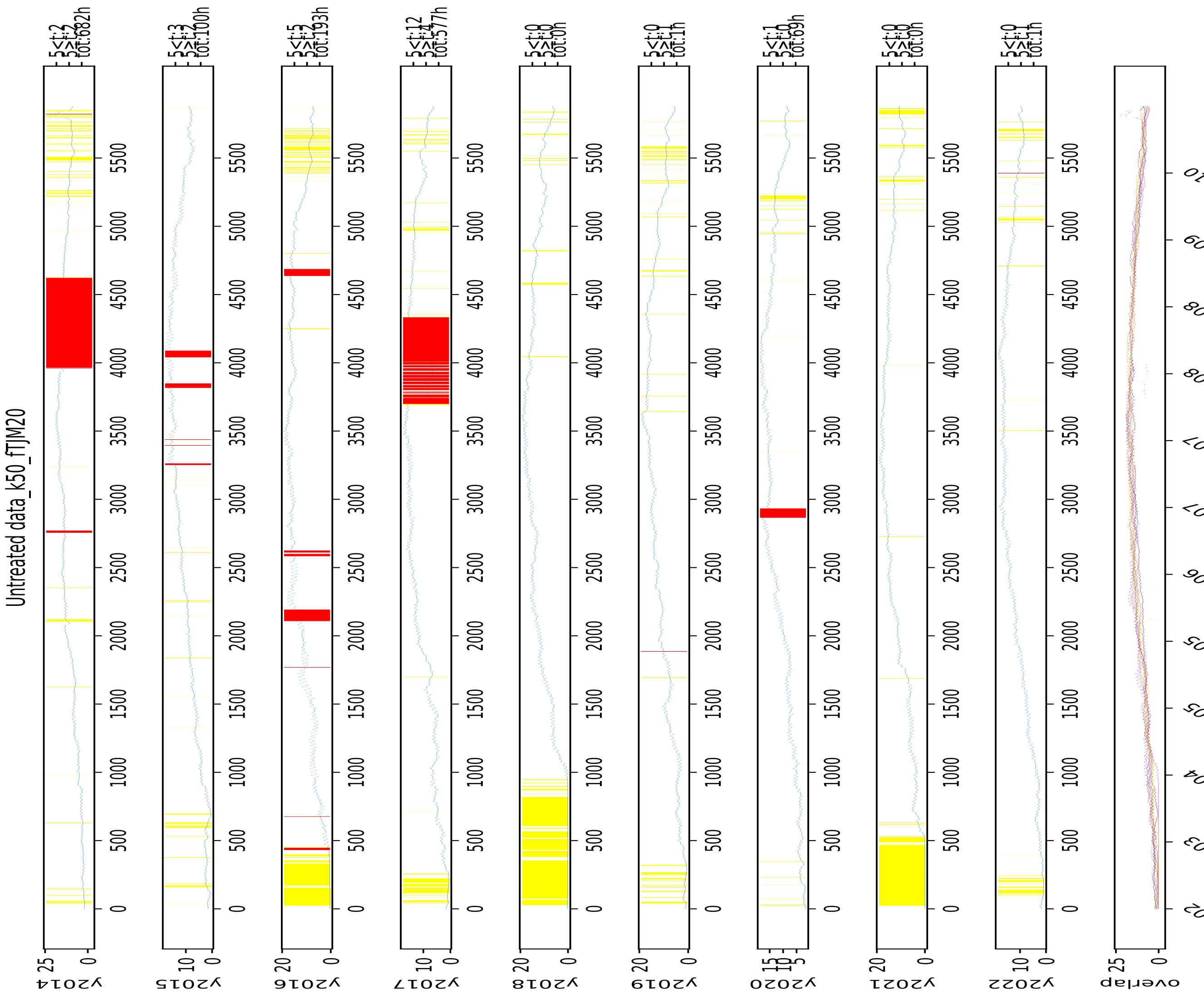


Figure 166: Visual representation of missing values at station 50 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

A.3 Data visiulation of data after treatment

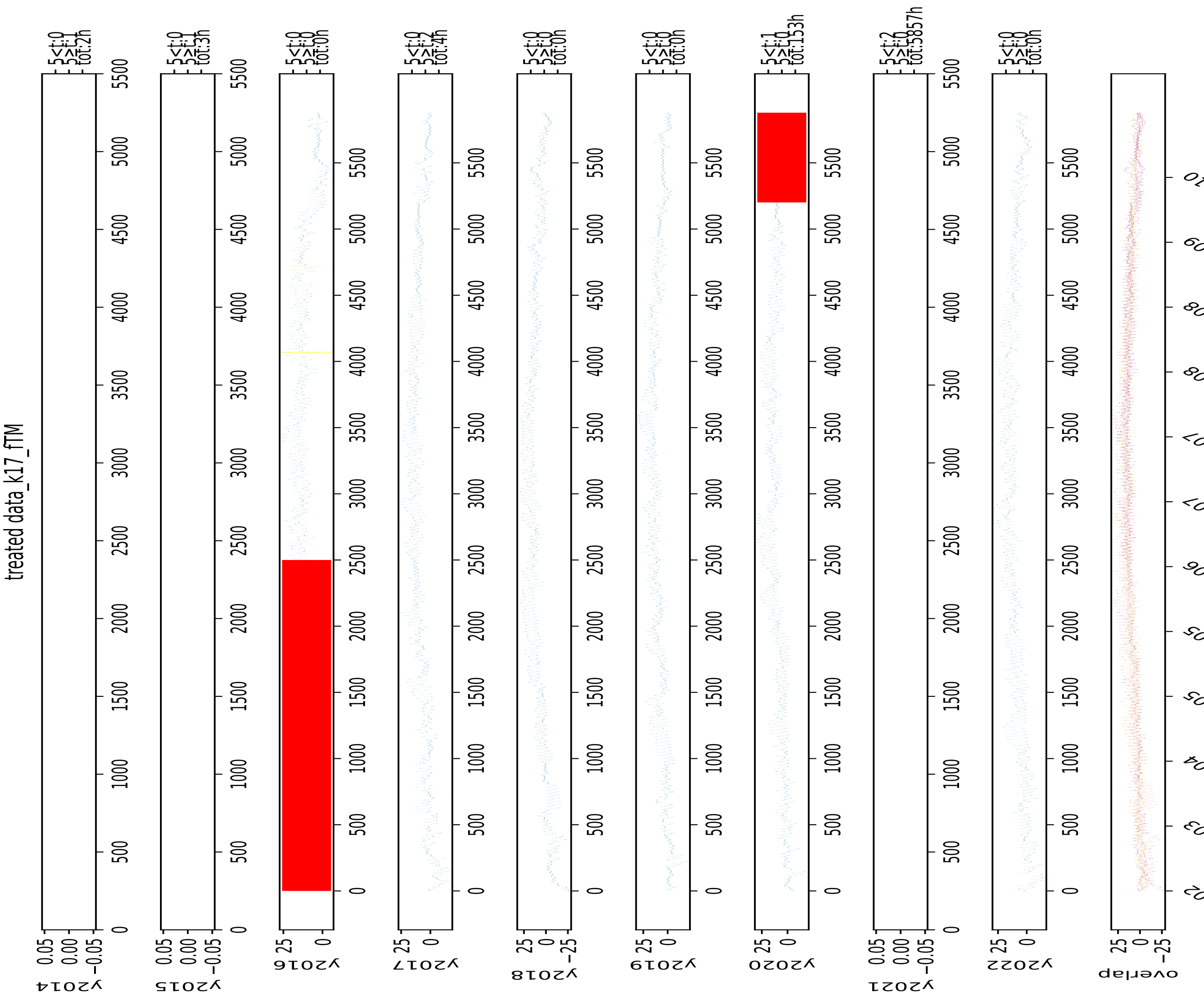


Figure 167: Visual representation of missing values at station 17 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

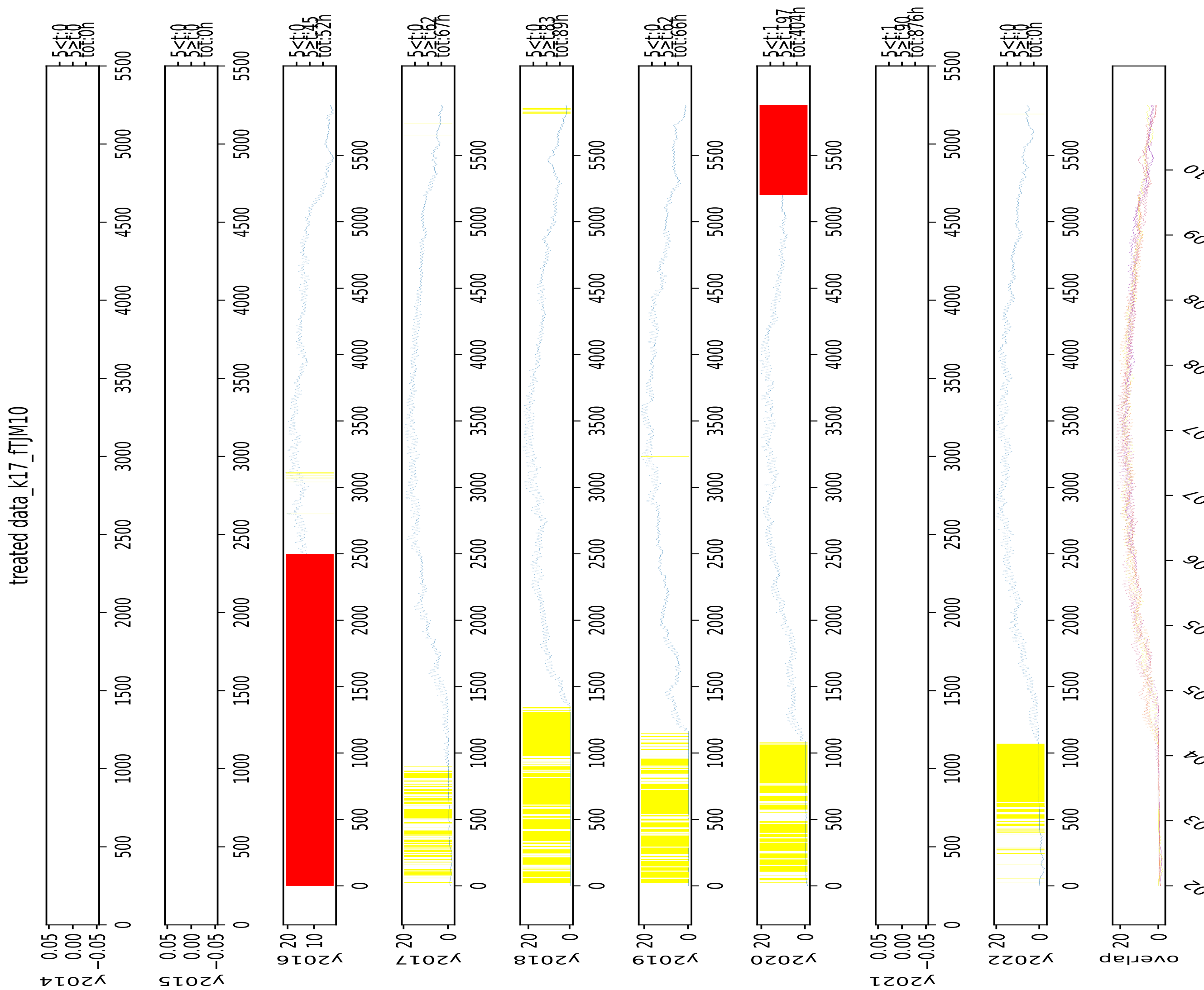


Figure 168: Visual representation of missing values at station 17 from 2014 to 2022 at the parameter ”TJM10” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

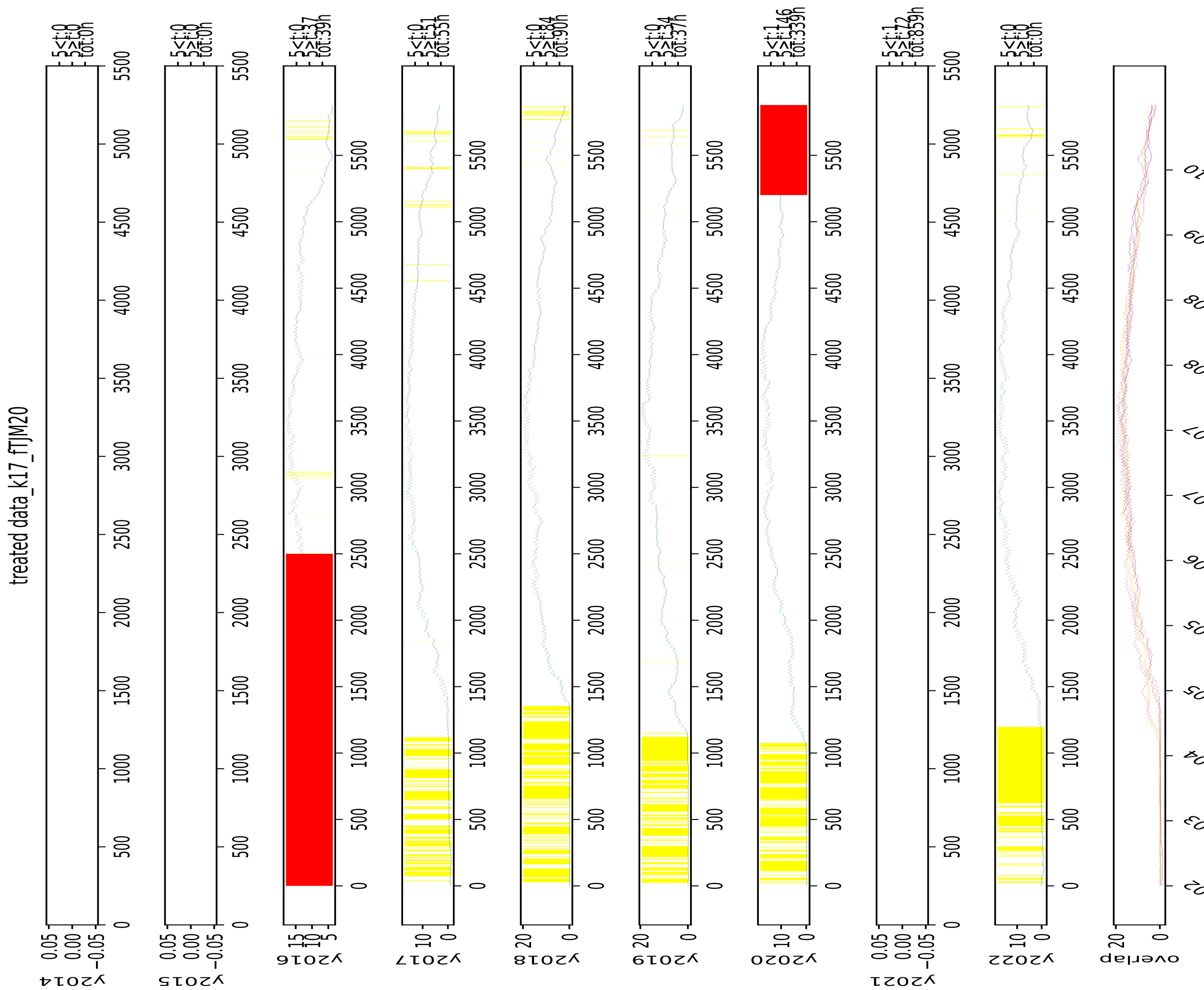


Figure 169: Visual representation of missing values at station 17 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

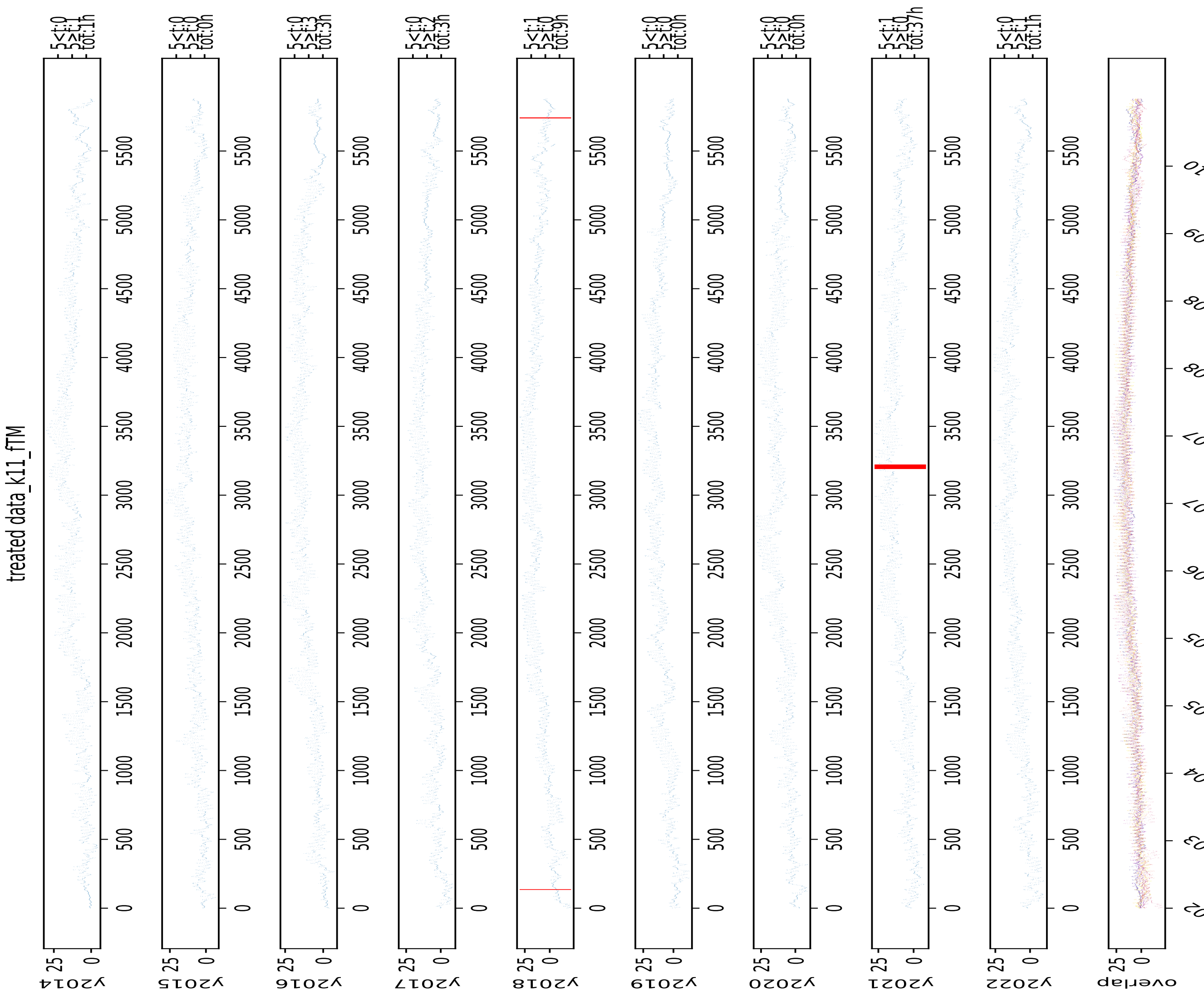


Figure 170: Visual representation of missing values at station 11 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

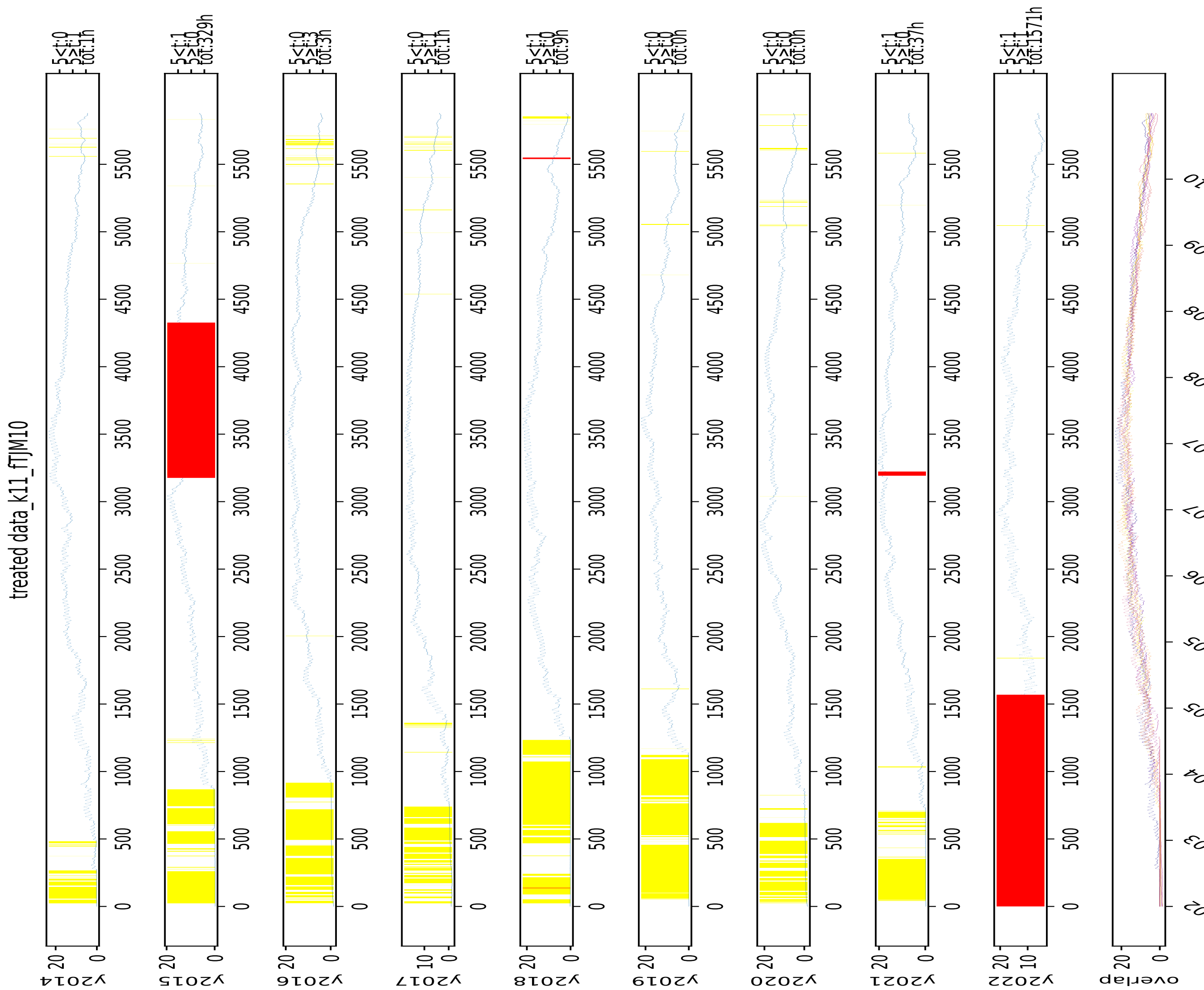


Figure 171: Visual representation of missing values at station 11 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

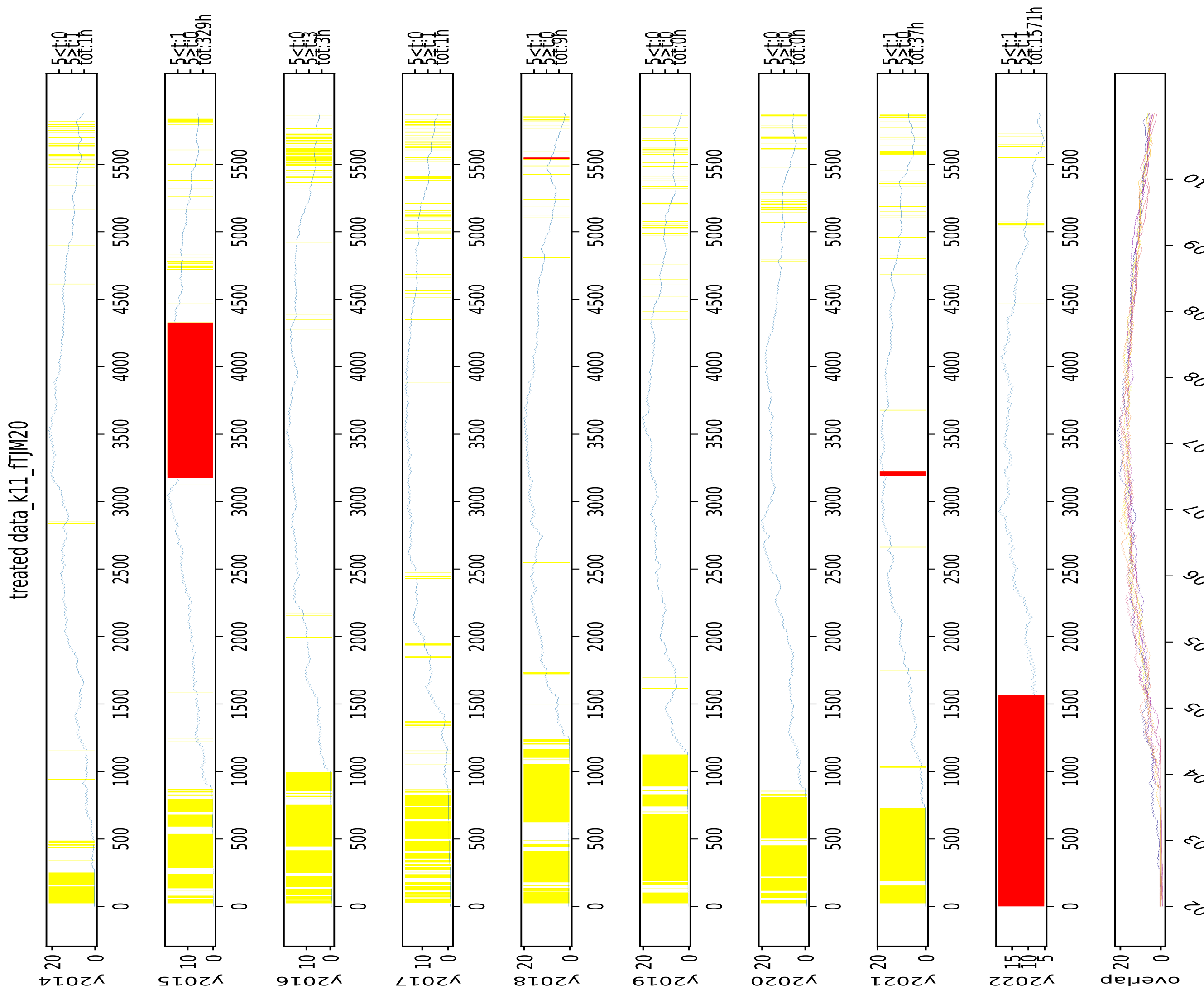


Figure 172: Visual representation of missing values at station 11 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

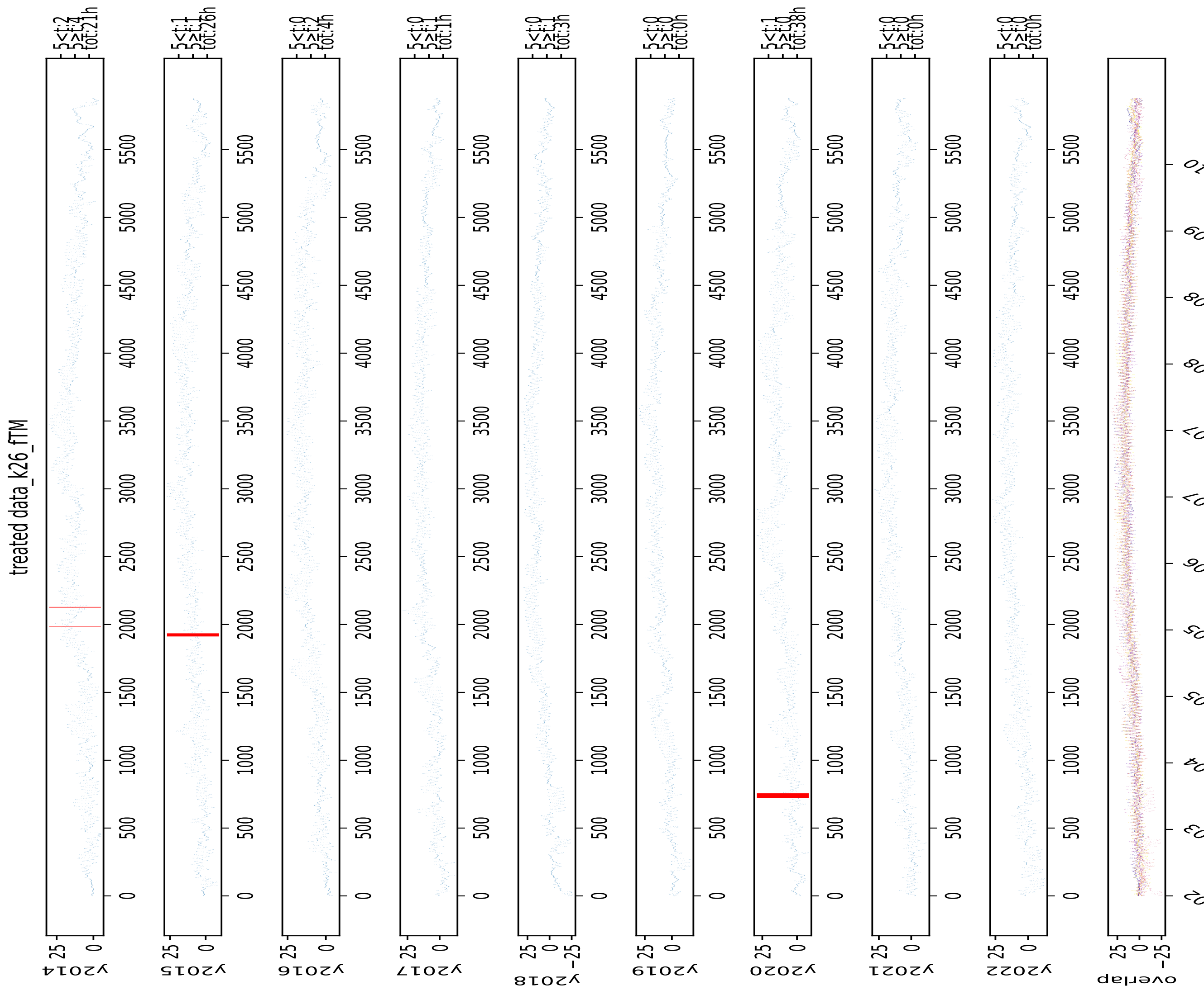


Figure 173: Visual representation of missing values at station 26 from 2014 to 2022 at the parameter ”TM” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

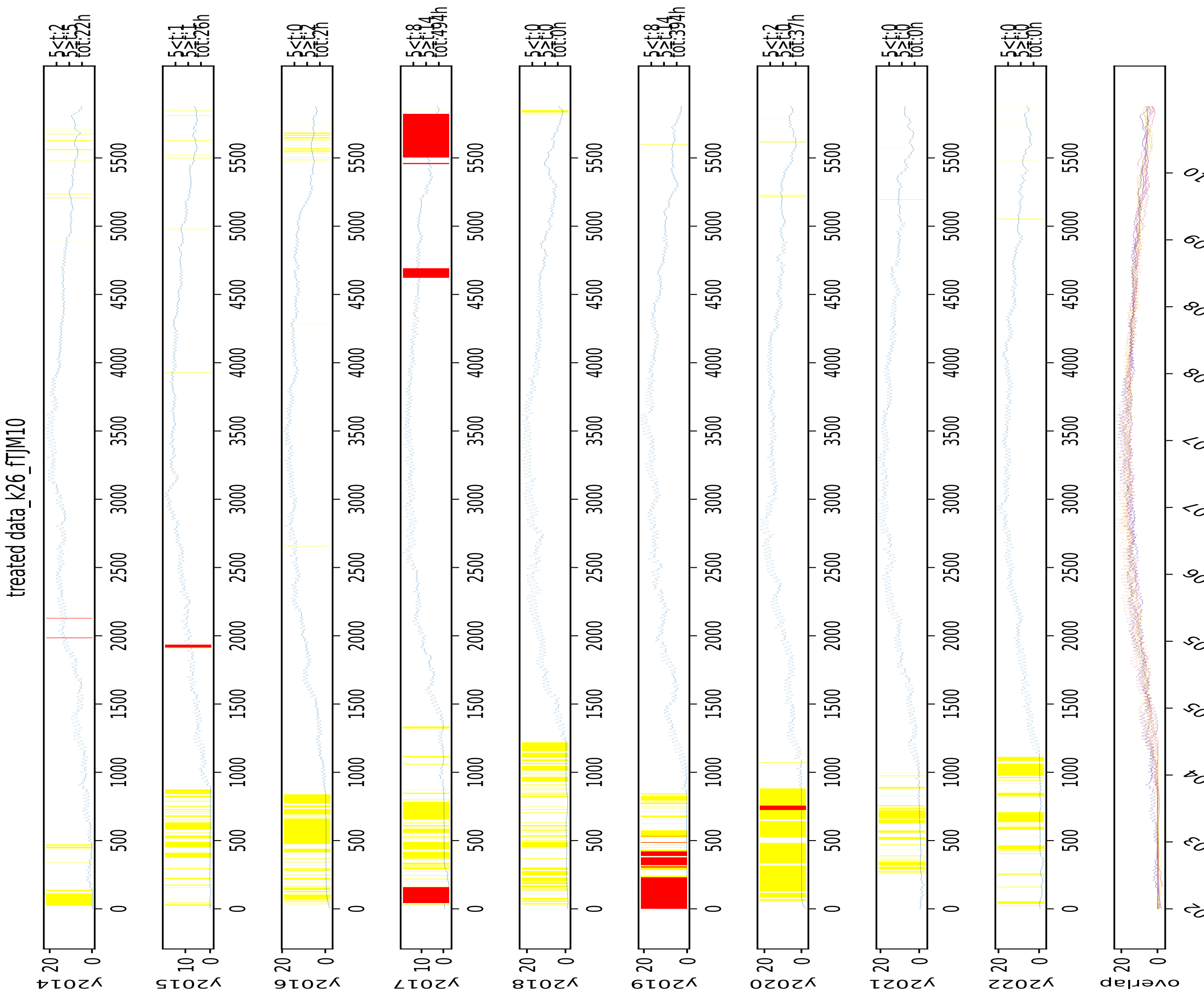


Figure 174: Visual representation of missing values at station 26 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

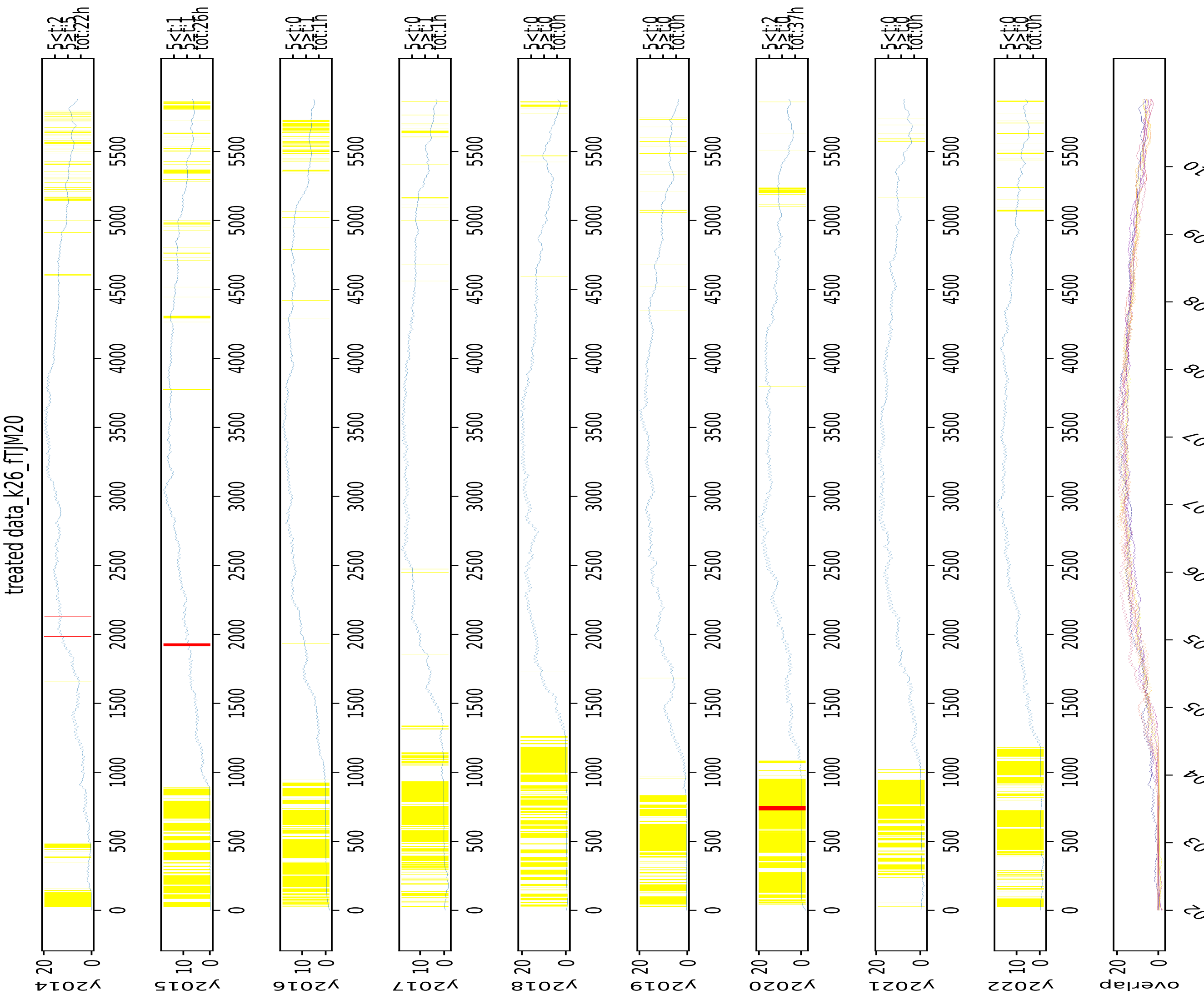


Figure 175: Visual representation of missing values at station 26 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

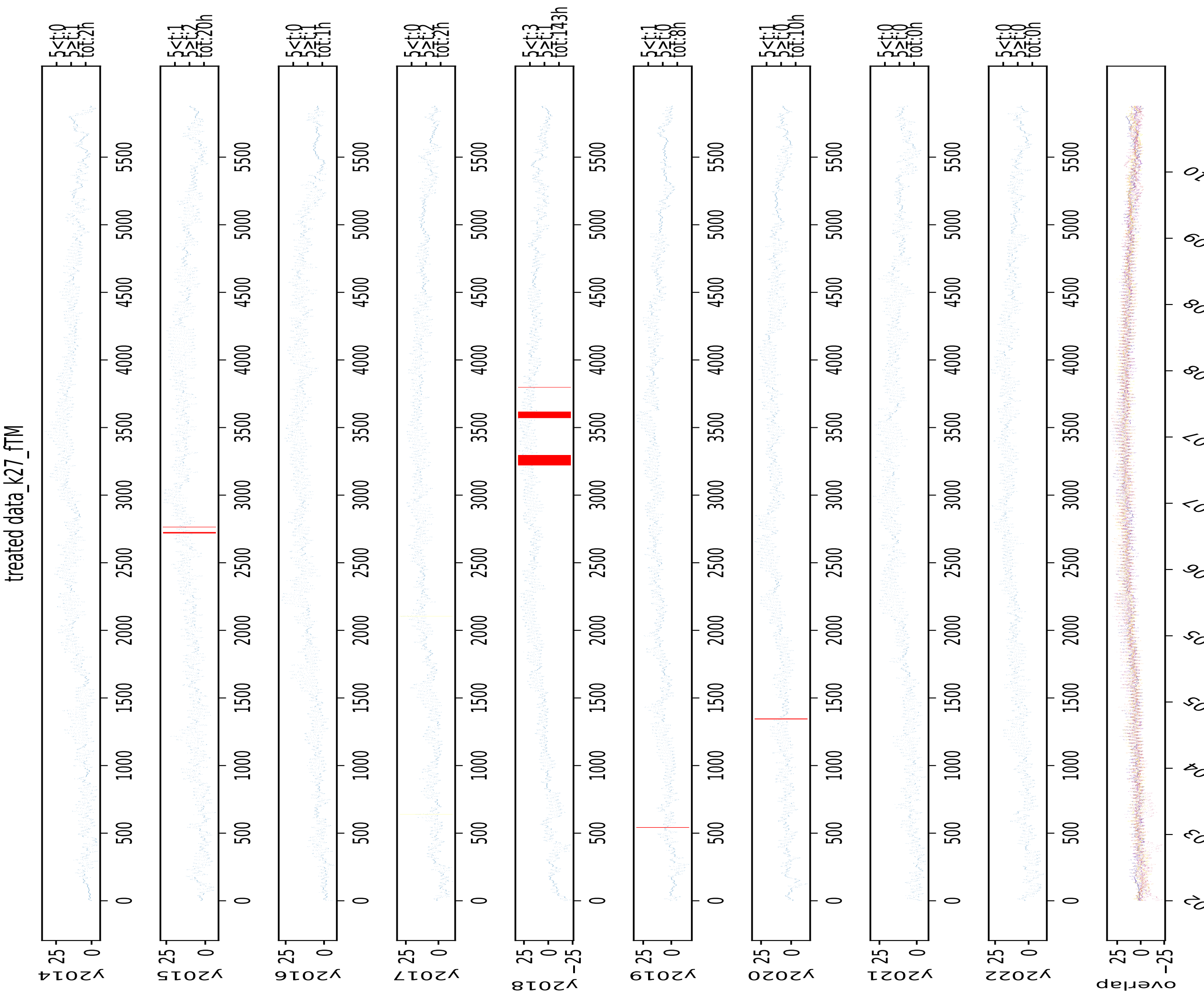


Figure 176: Visual representation of missing values at station 27 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

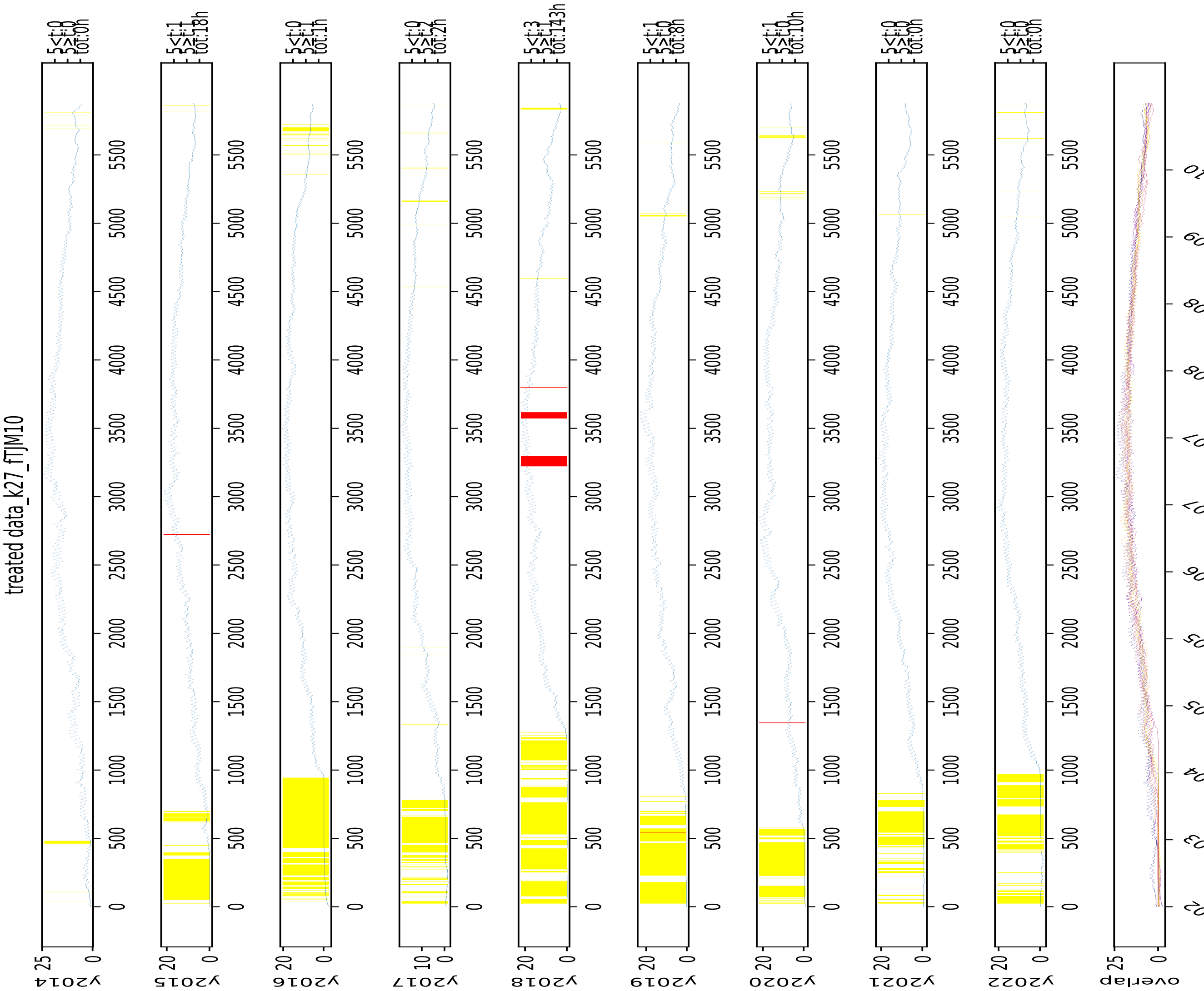


Figure 177: Visual representation of missing values at station 27 from 2014 to 2022 at the parameter ”TJM10” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

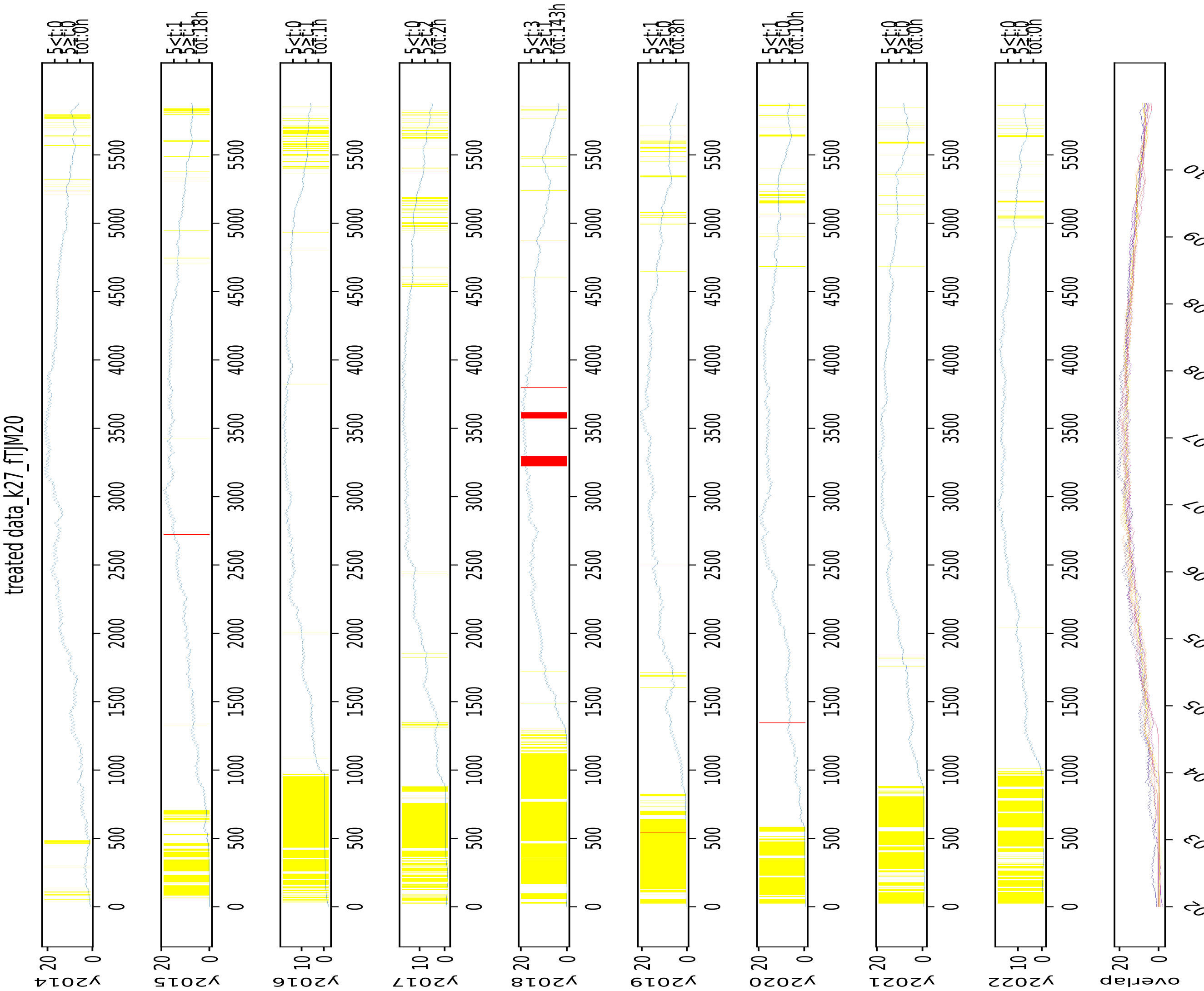


Figure 178: Visual representation of missing values at station 27 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

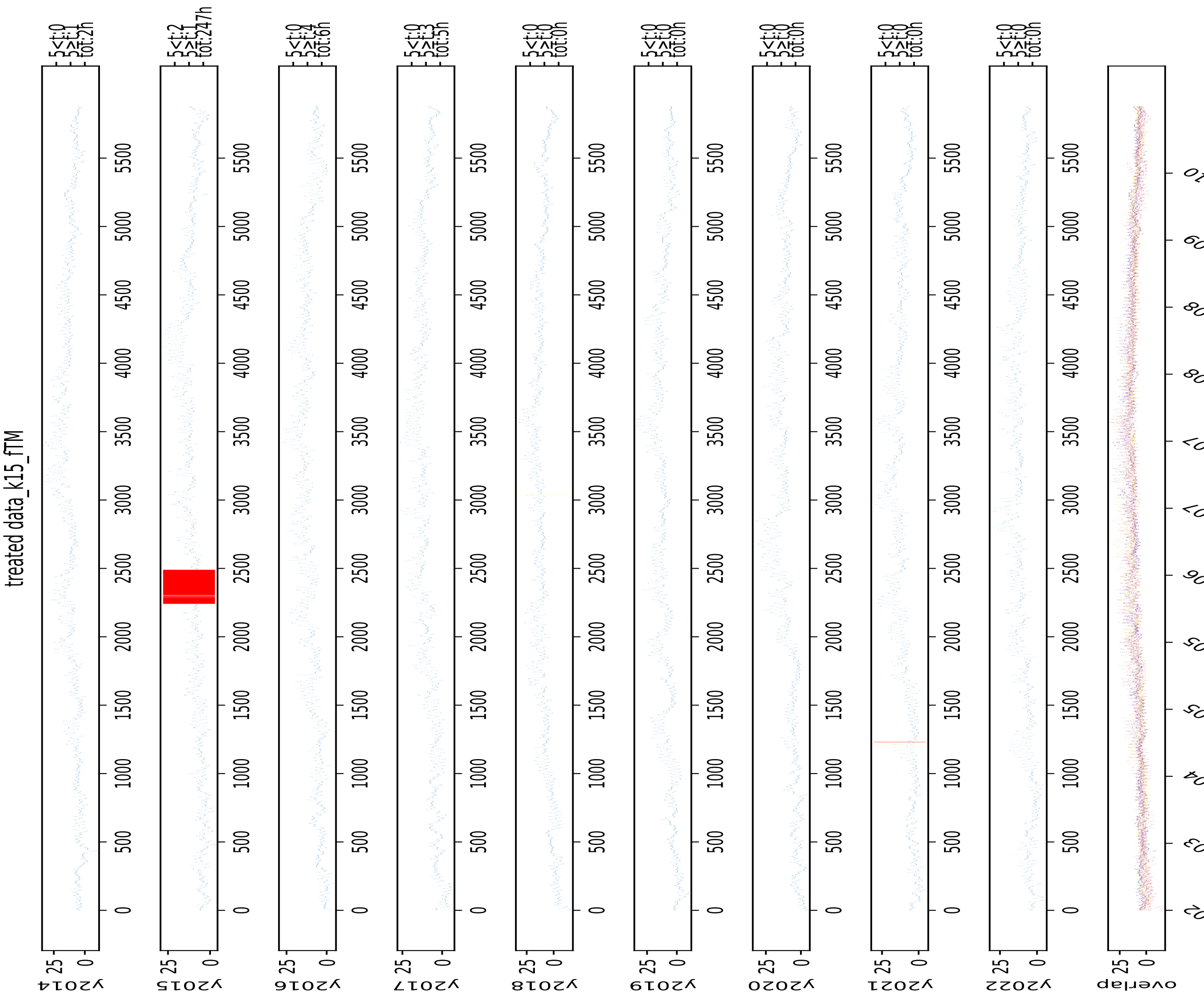


Figure 179: Visual representation of missing values at station 15 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

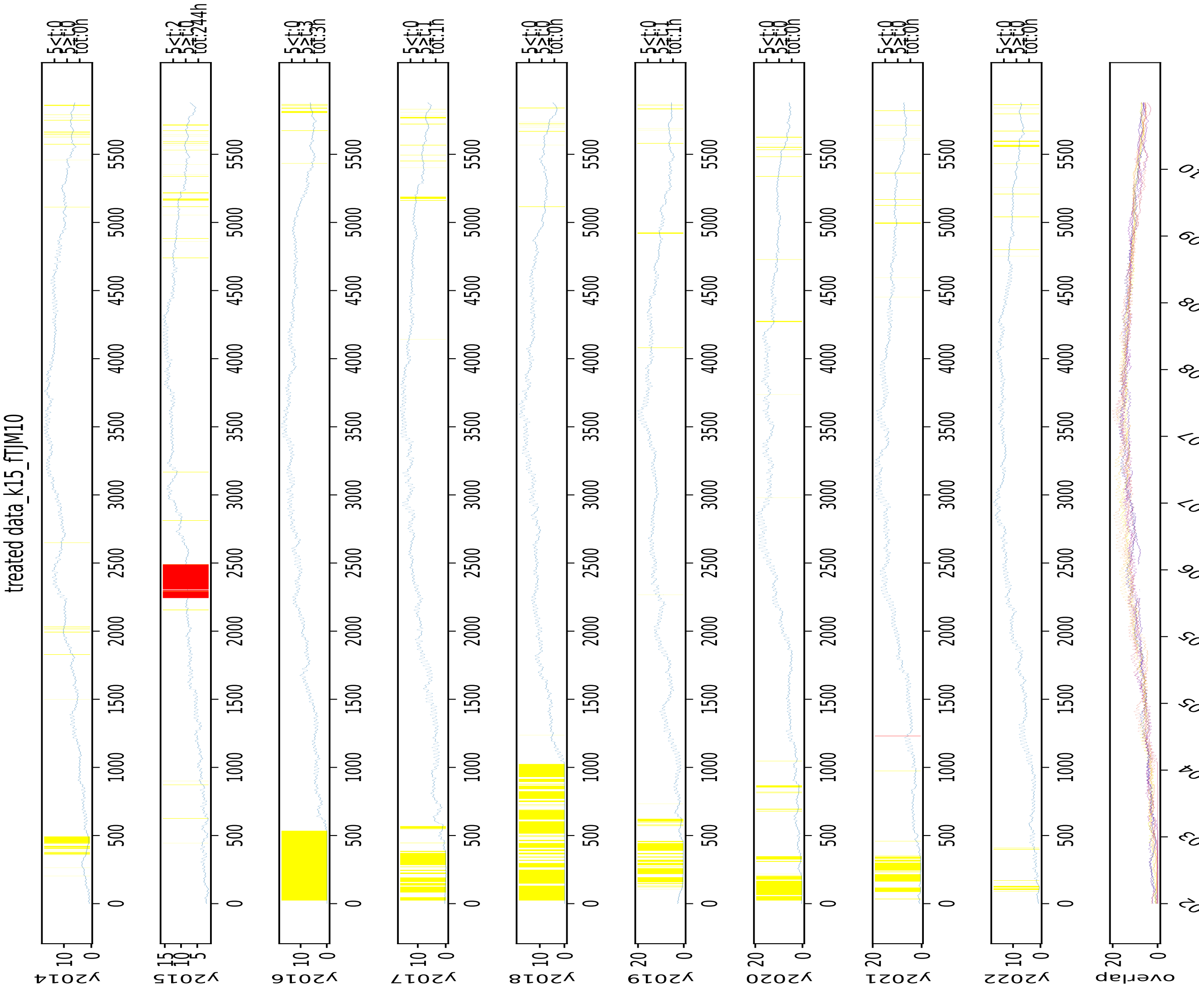


Figure 180: Visual representation of missing values at station 15 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

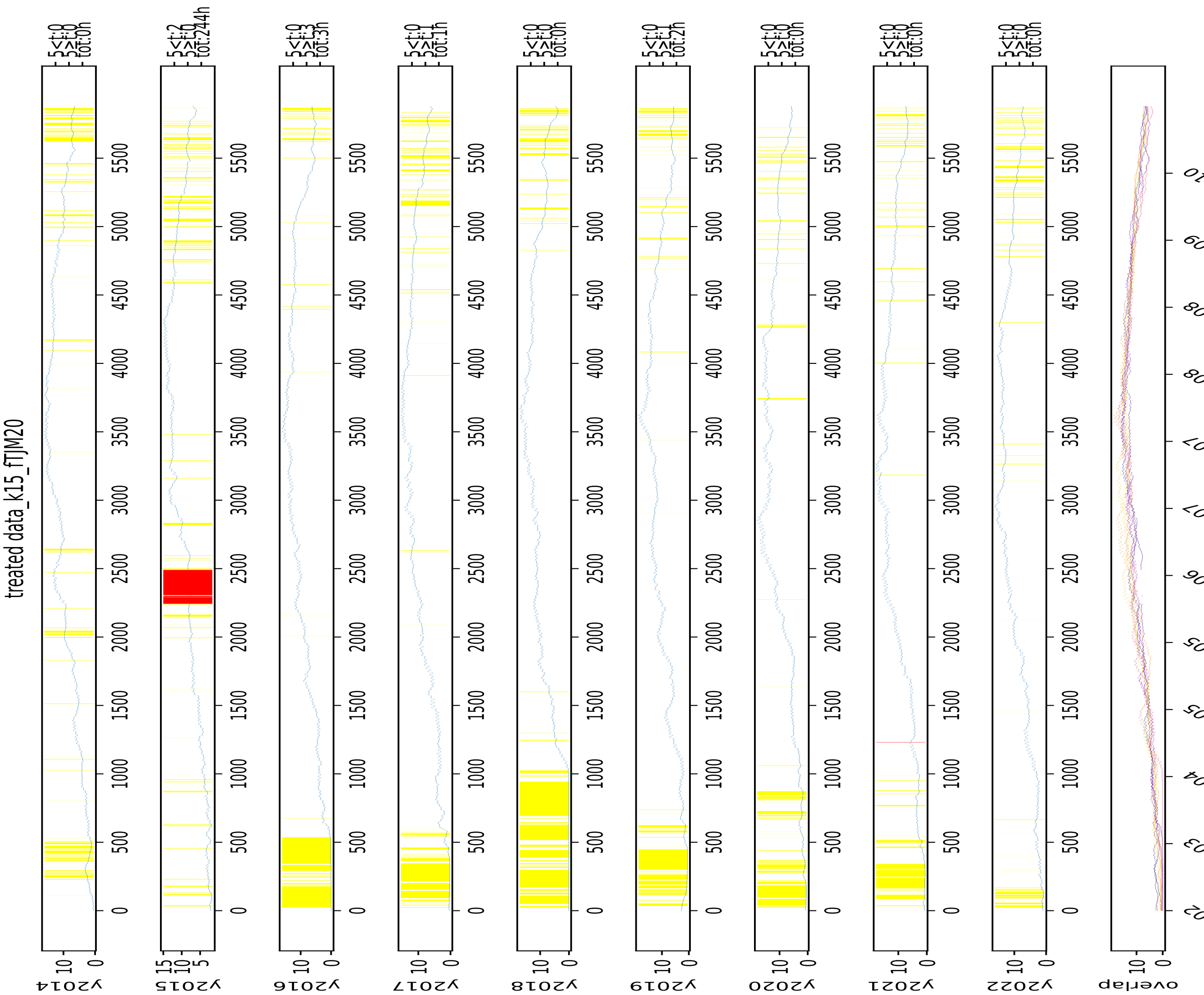


Figure 181: Visual representation of missing values at station 15 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

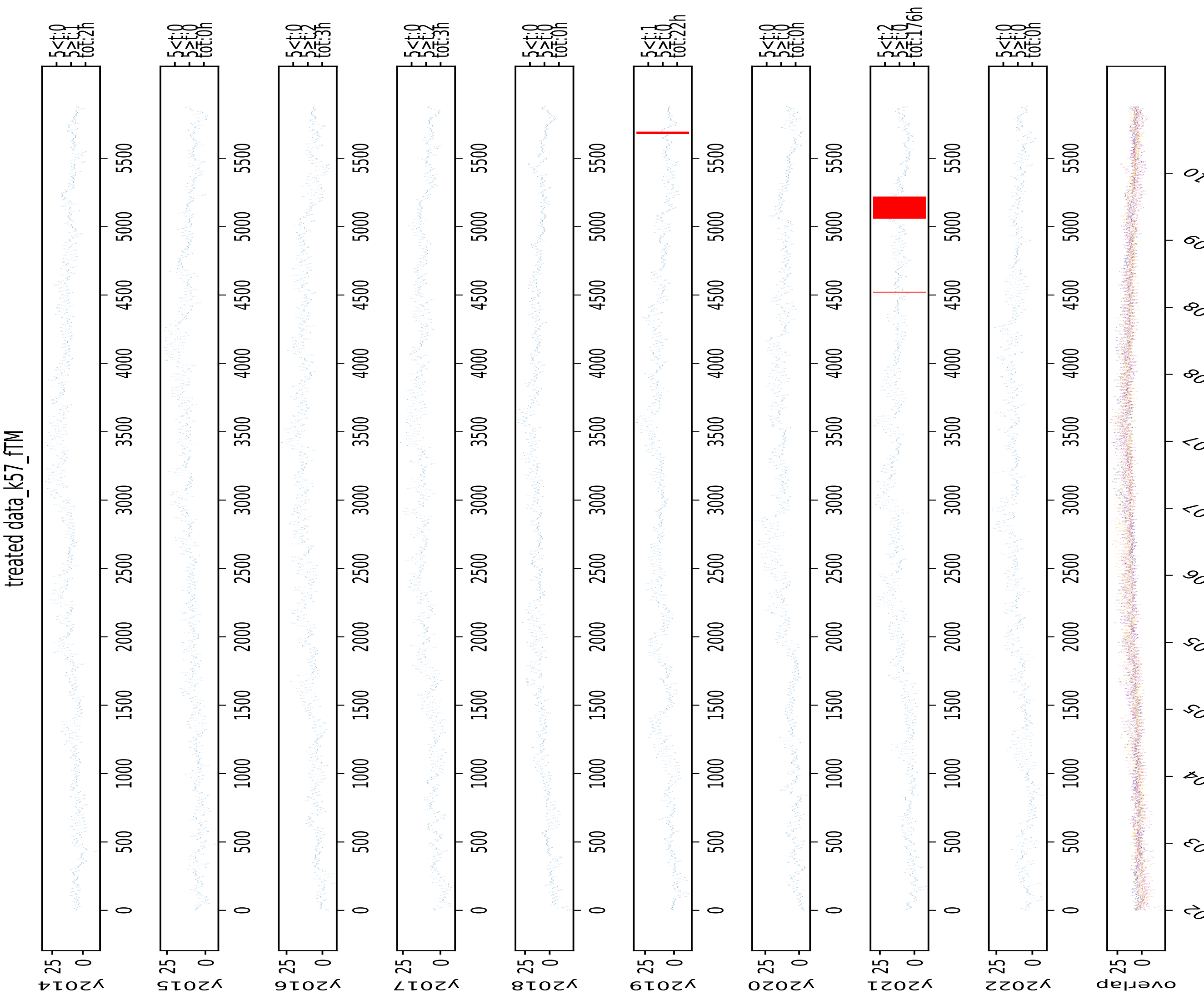


Figure 182: Visual representation of missing values at station 57 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

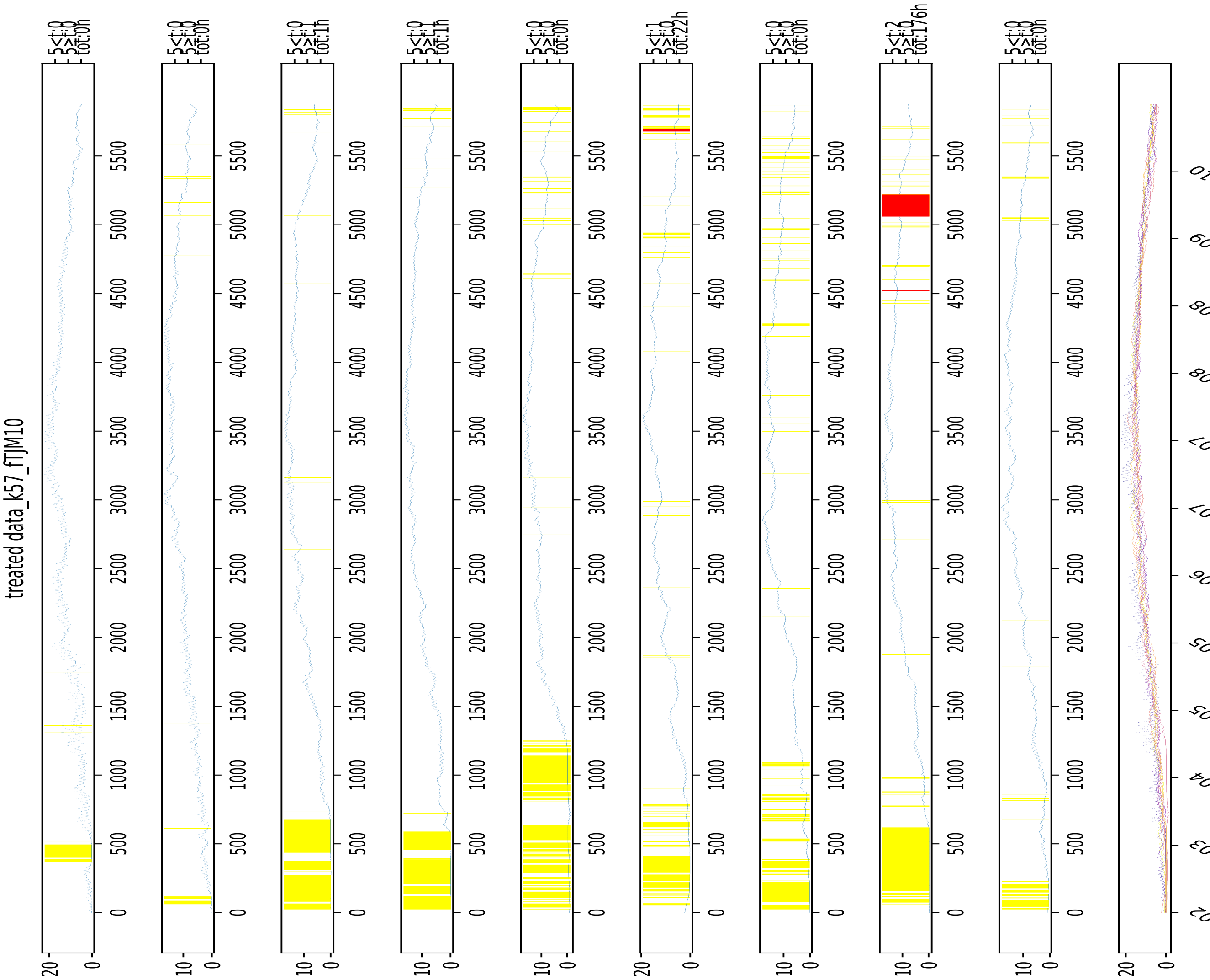


Figure 183: Visual representation of missing values at station 57 from 2014 to 2022 at the parameter ”TJM10” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

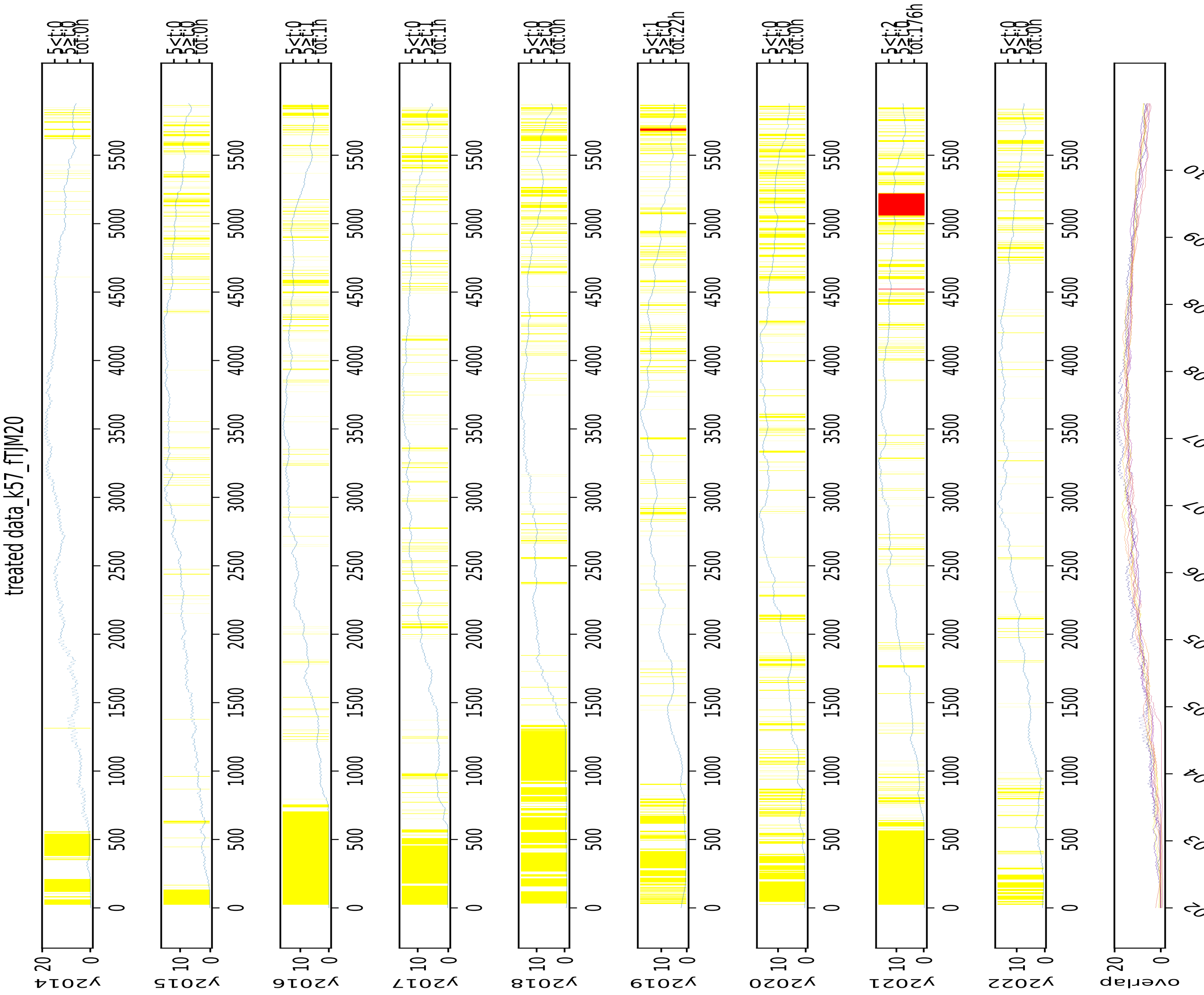


Figure 184: Visual representation of missing values at station 57 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

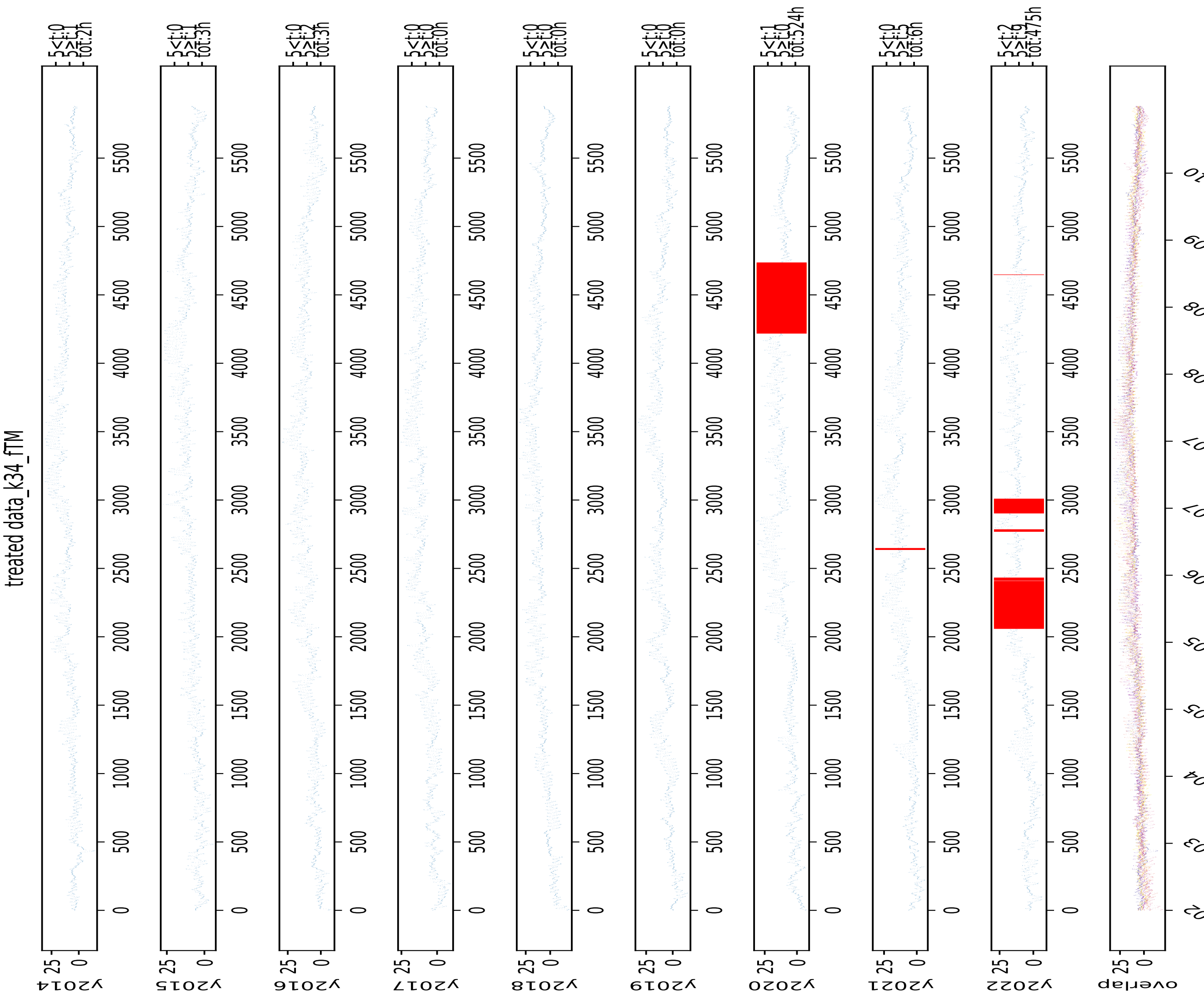


Figure 185: Visual representation of missing values at station 34 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

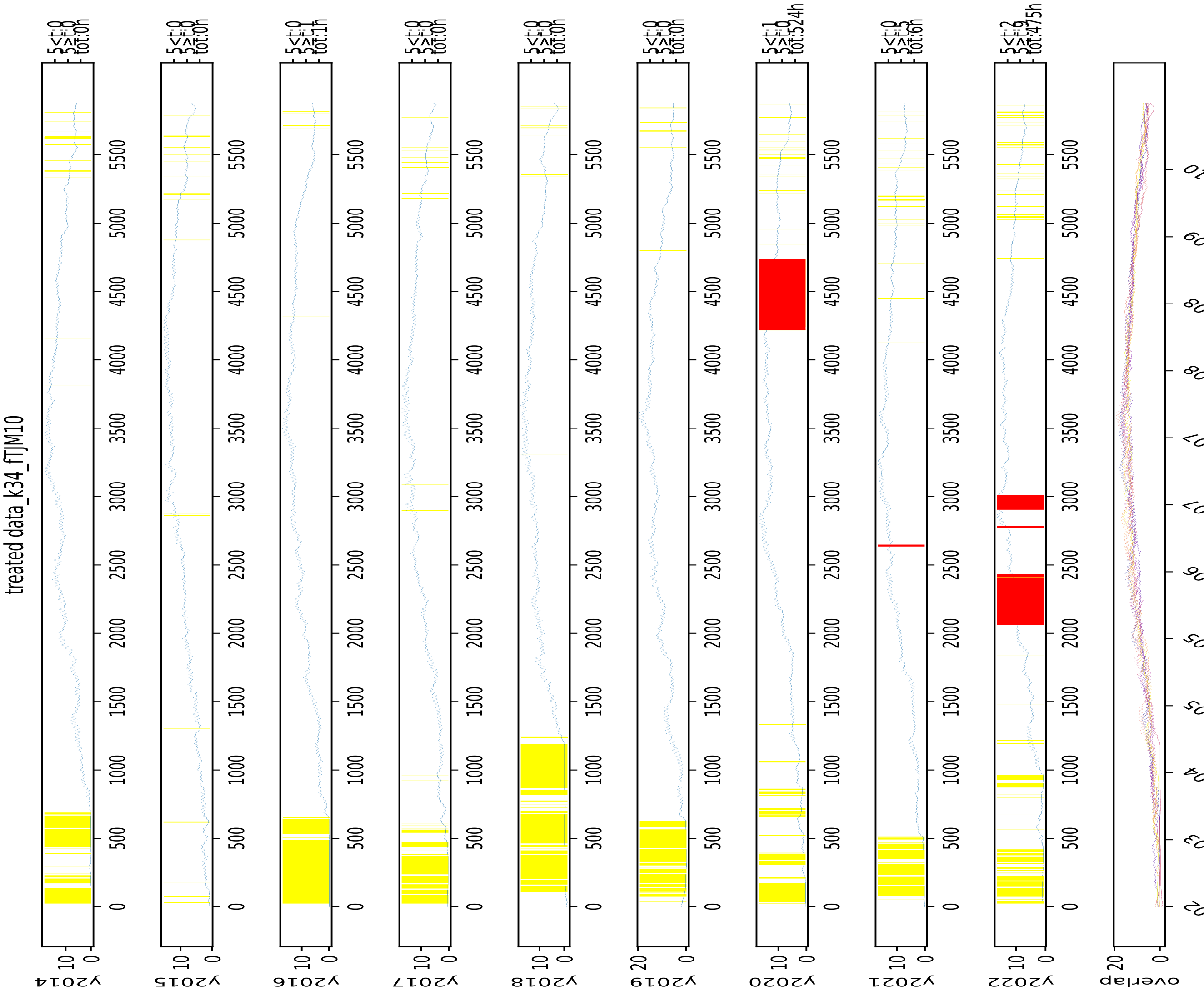


Figure 186: Visual representation of missing values at station 34 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

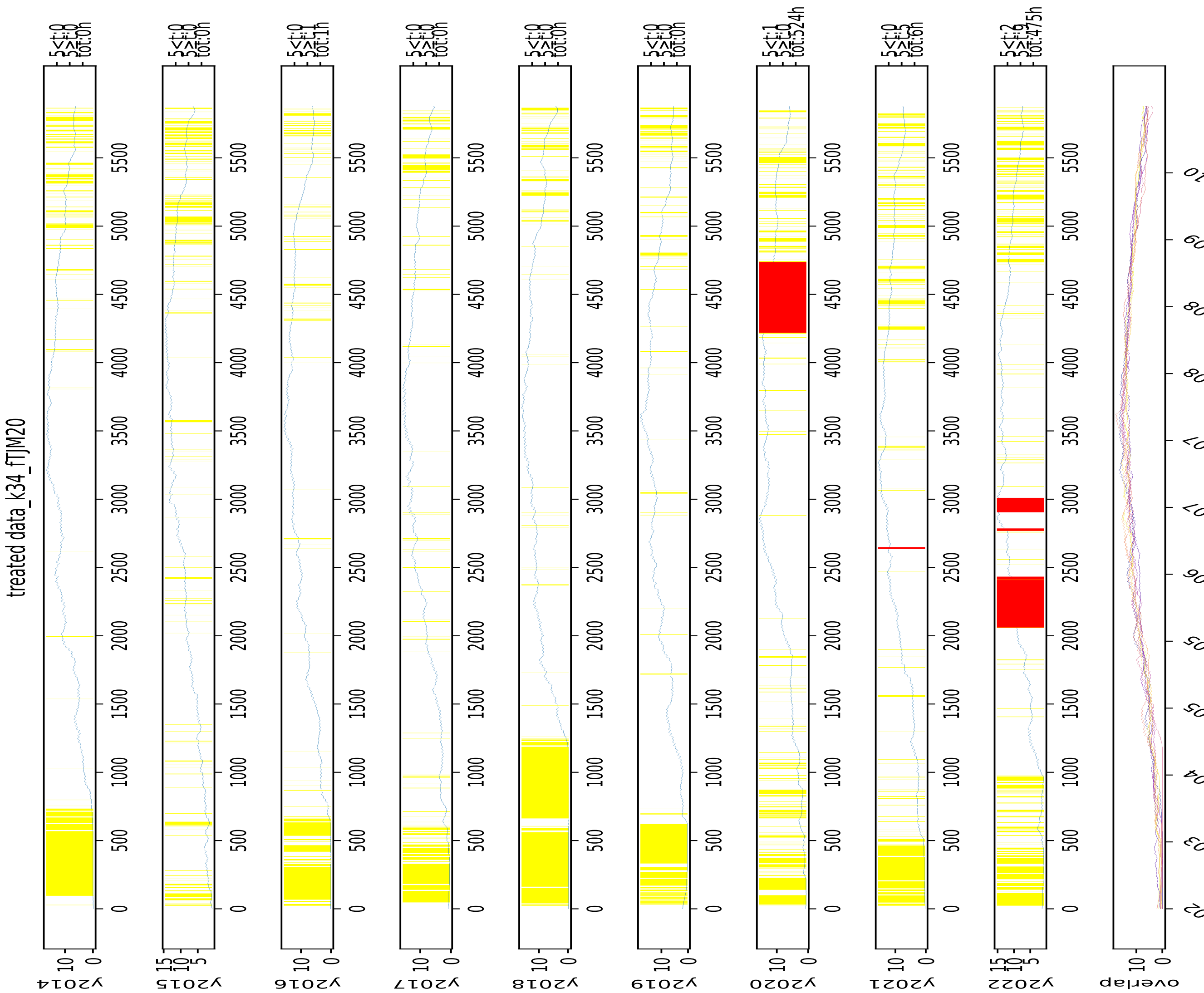


Figure 187: Visual representation of missing values at station 34 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

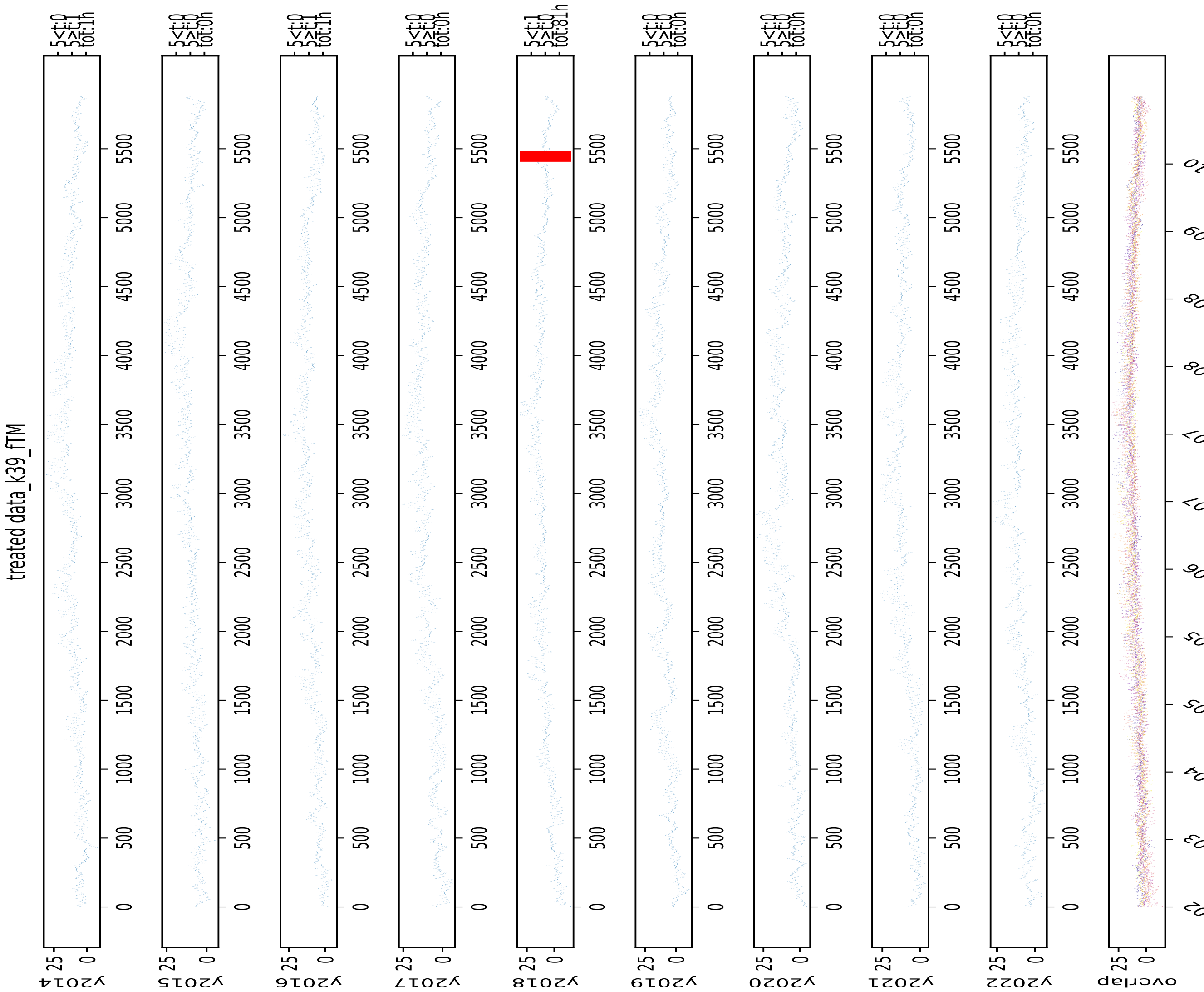


Figure 188: Visual representation of missing values at station 39 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

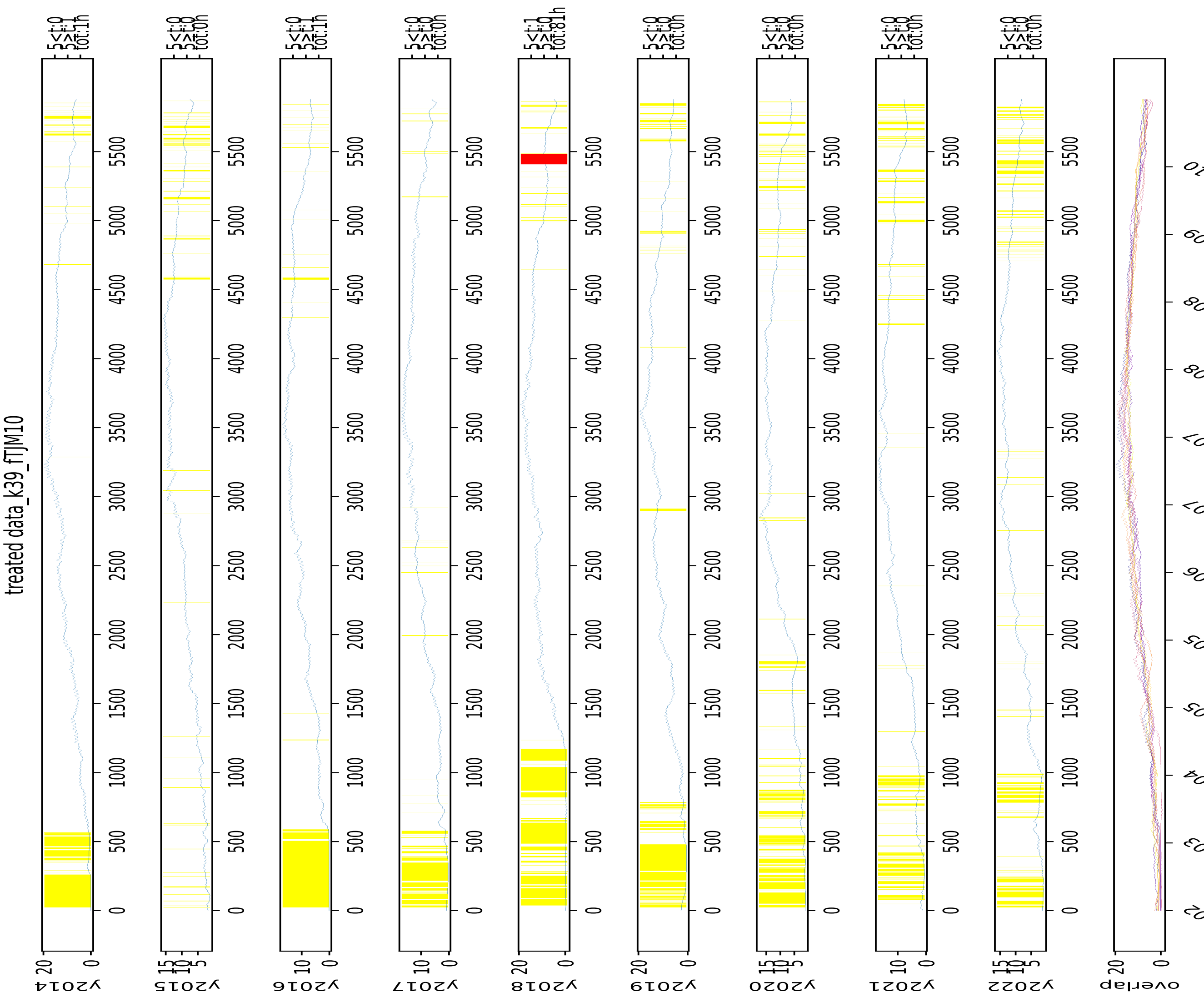


Figure 189: Visual representation of missing values at station 39 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

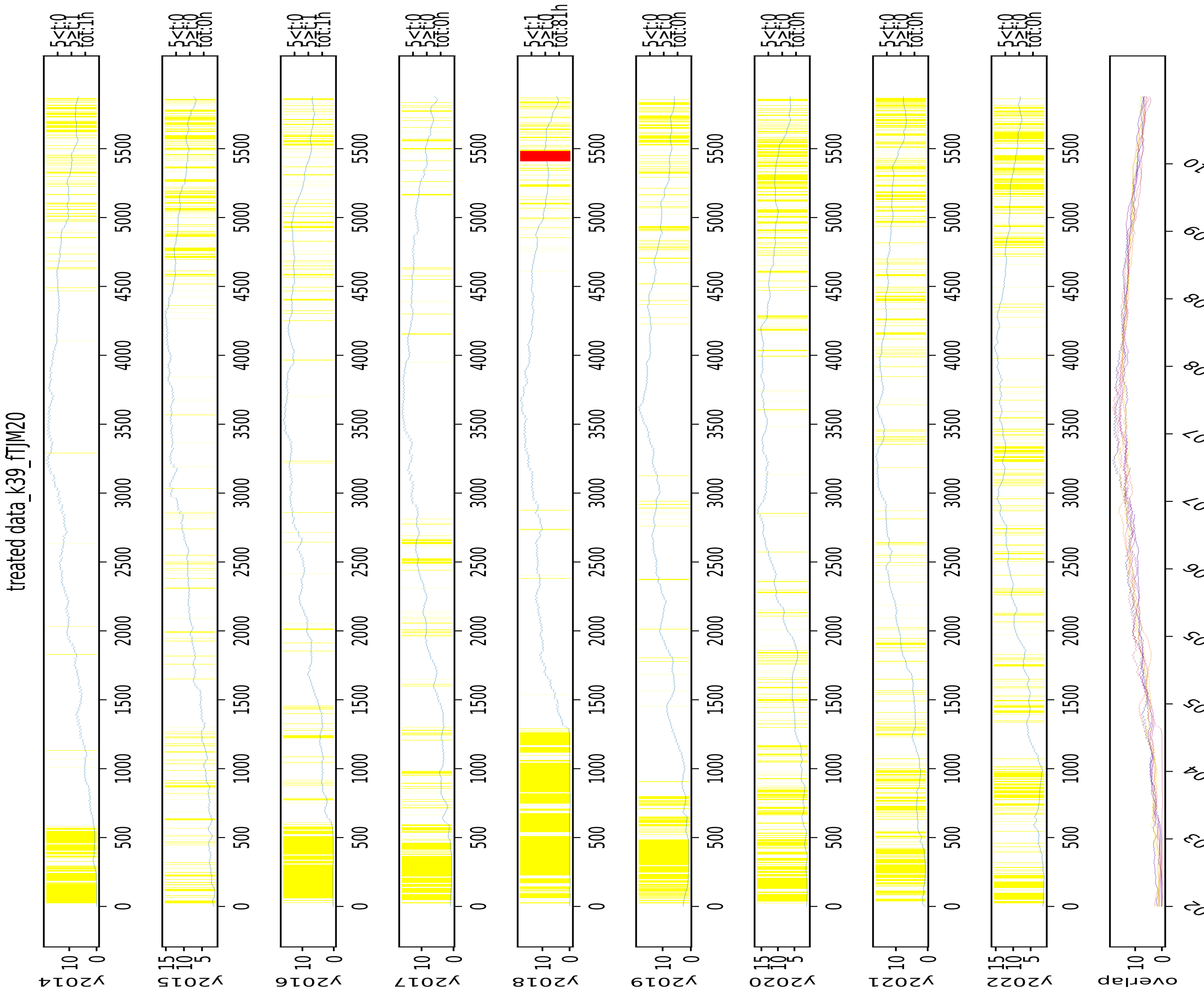


Figure 190: Visual representation of missing values at station 39 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

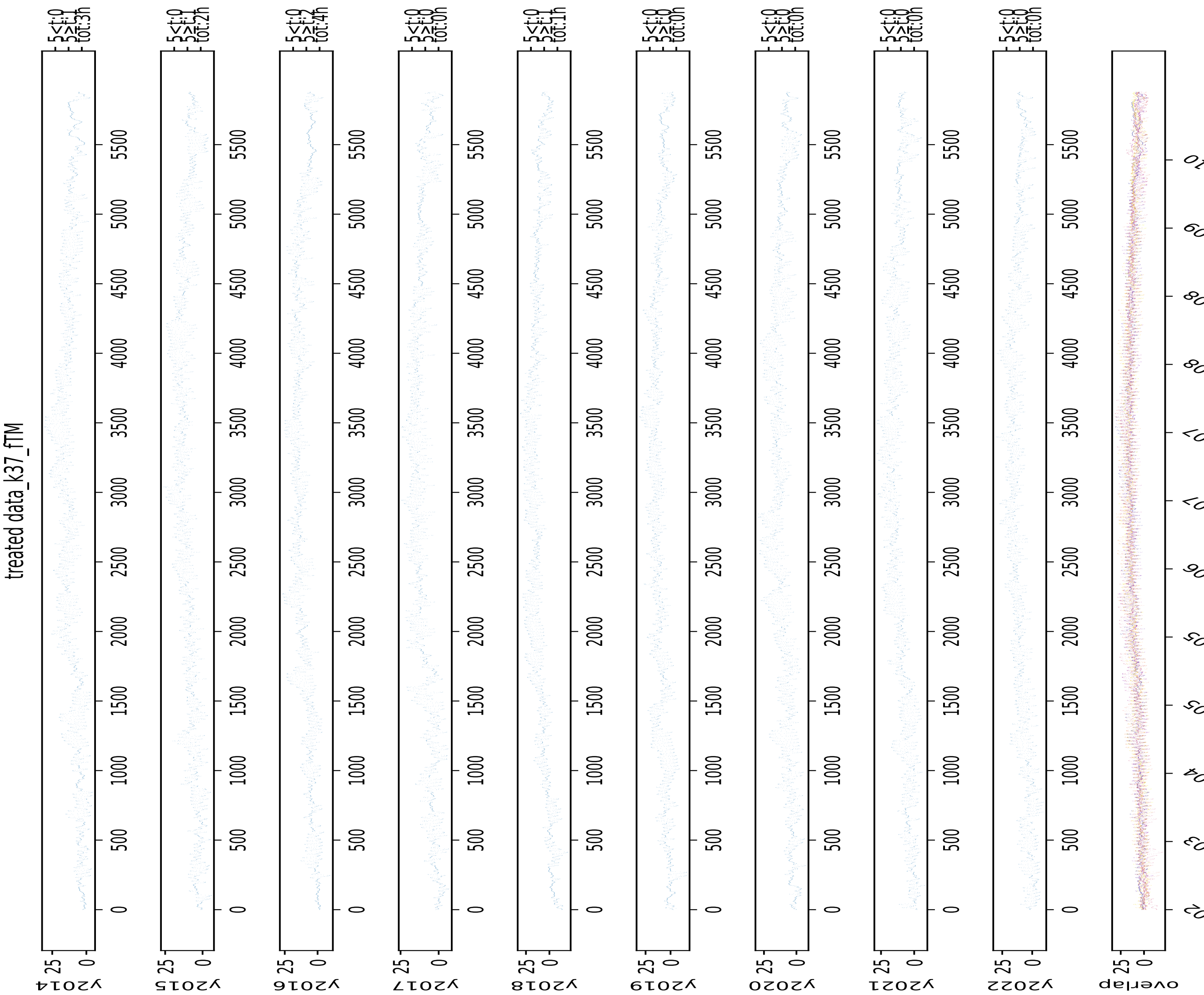


Figure 191: Visual representation of missing values at station 37 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

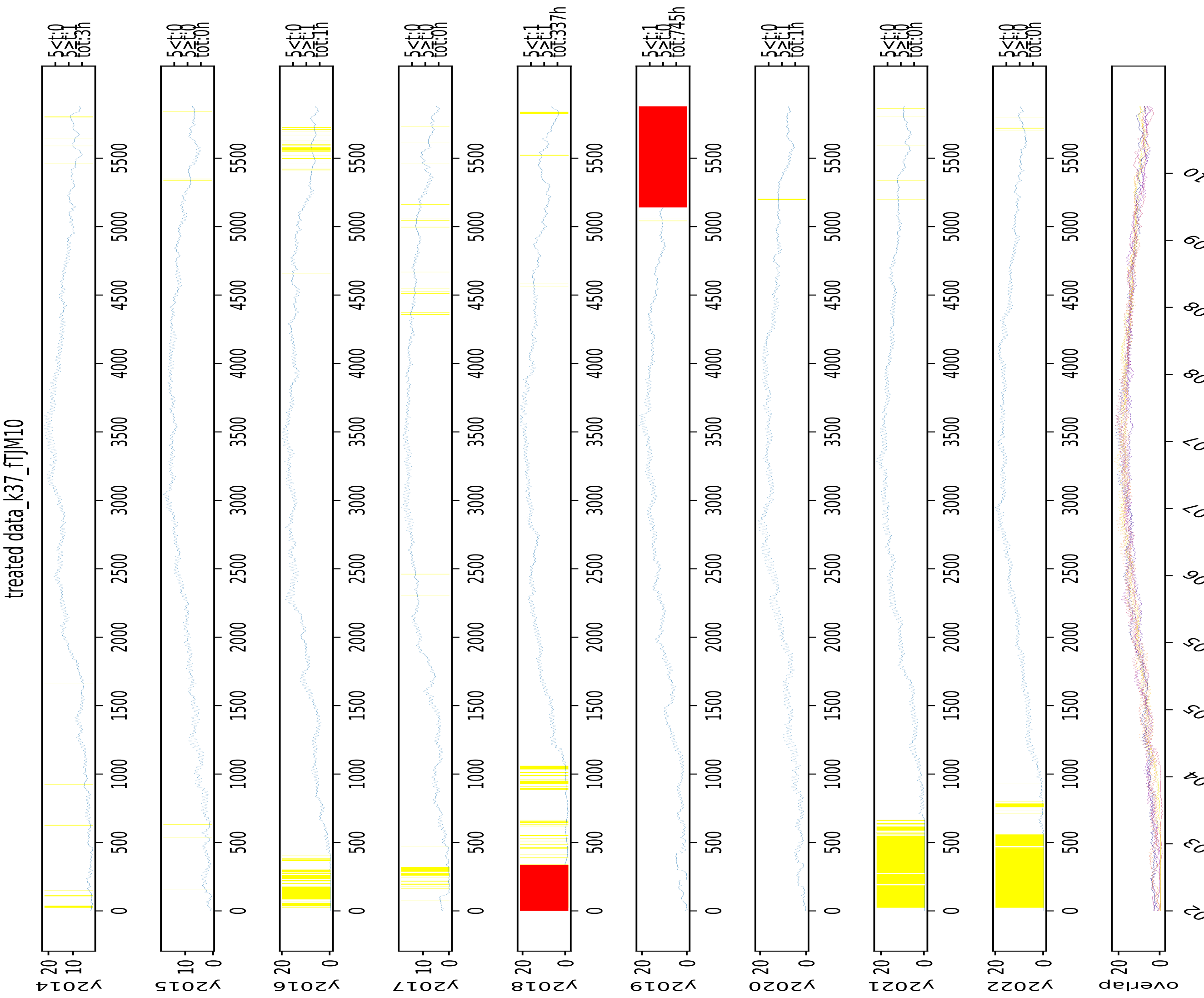


Figure 192: Visual representation of missing values at station 37 from 2014 to 2022 at the parameter ”TJM10” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

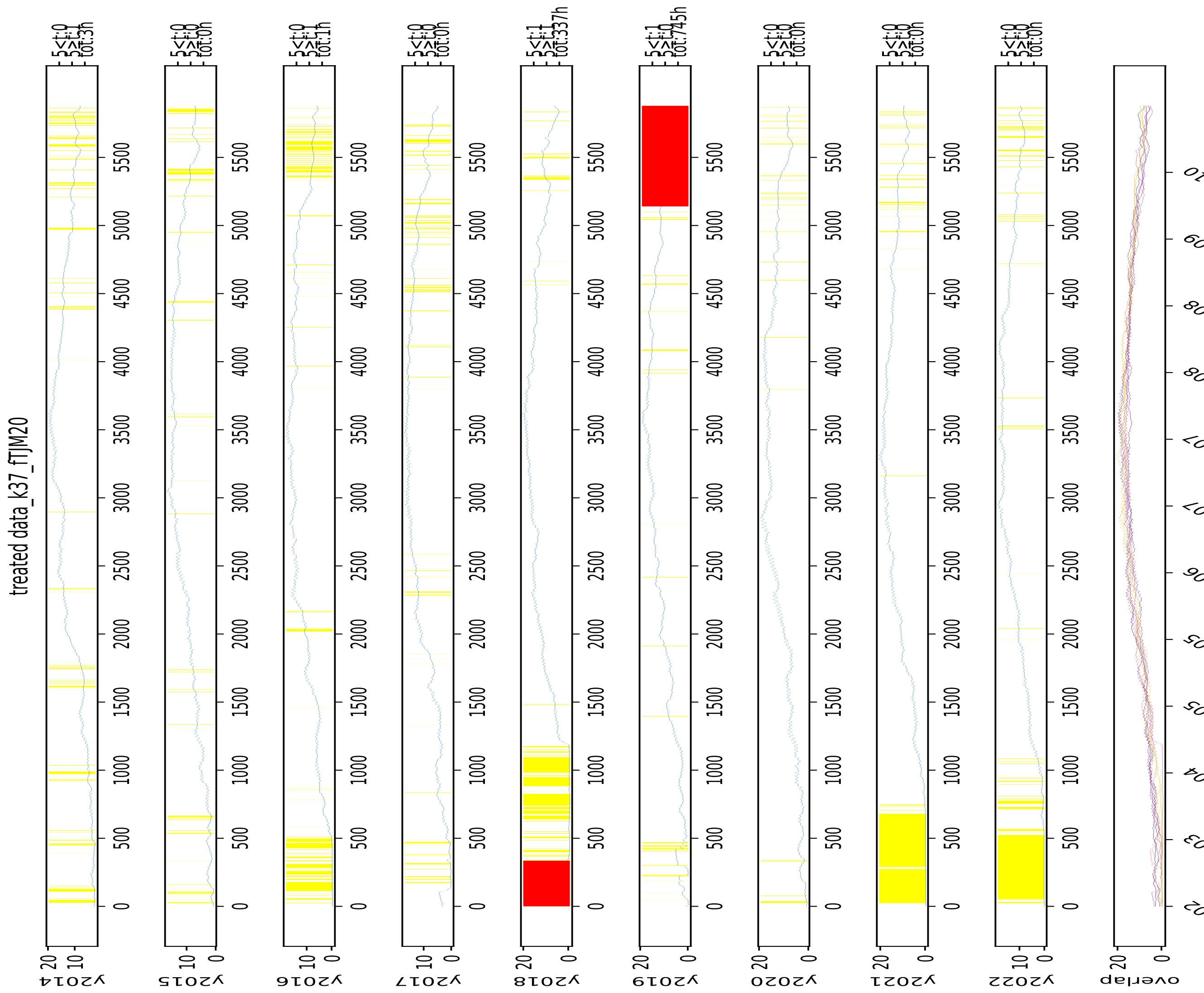


Figure 193: Visual representation of missing values at station 37 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

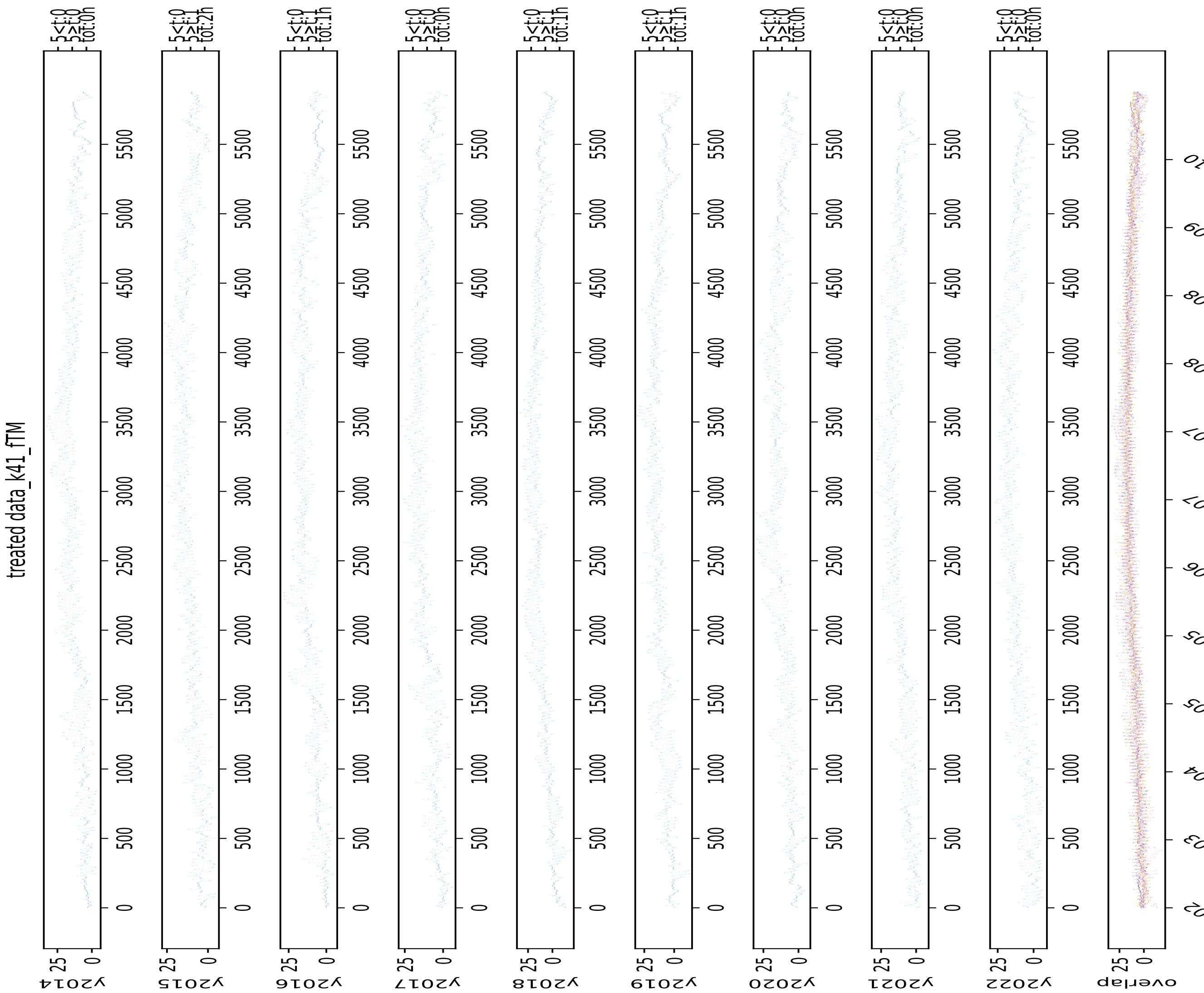


Figure 194: Visual representation of missing values at station 41 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

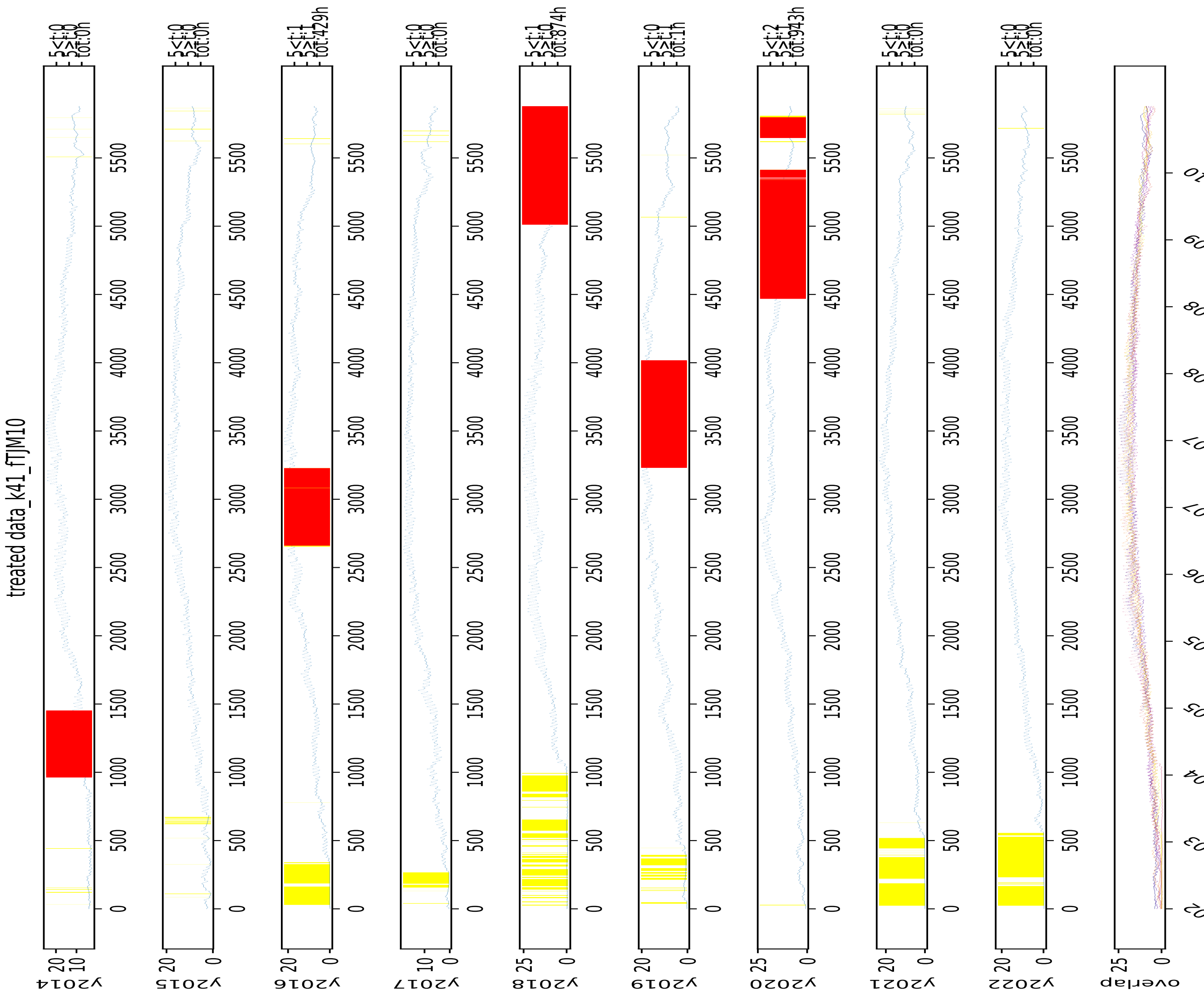


Figure 195: Visual representation of missing values at station 41 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

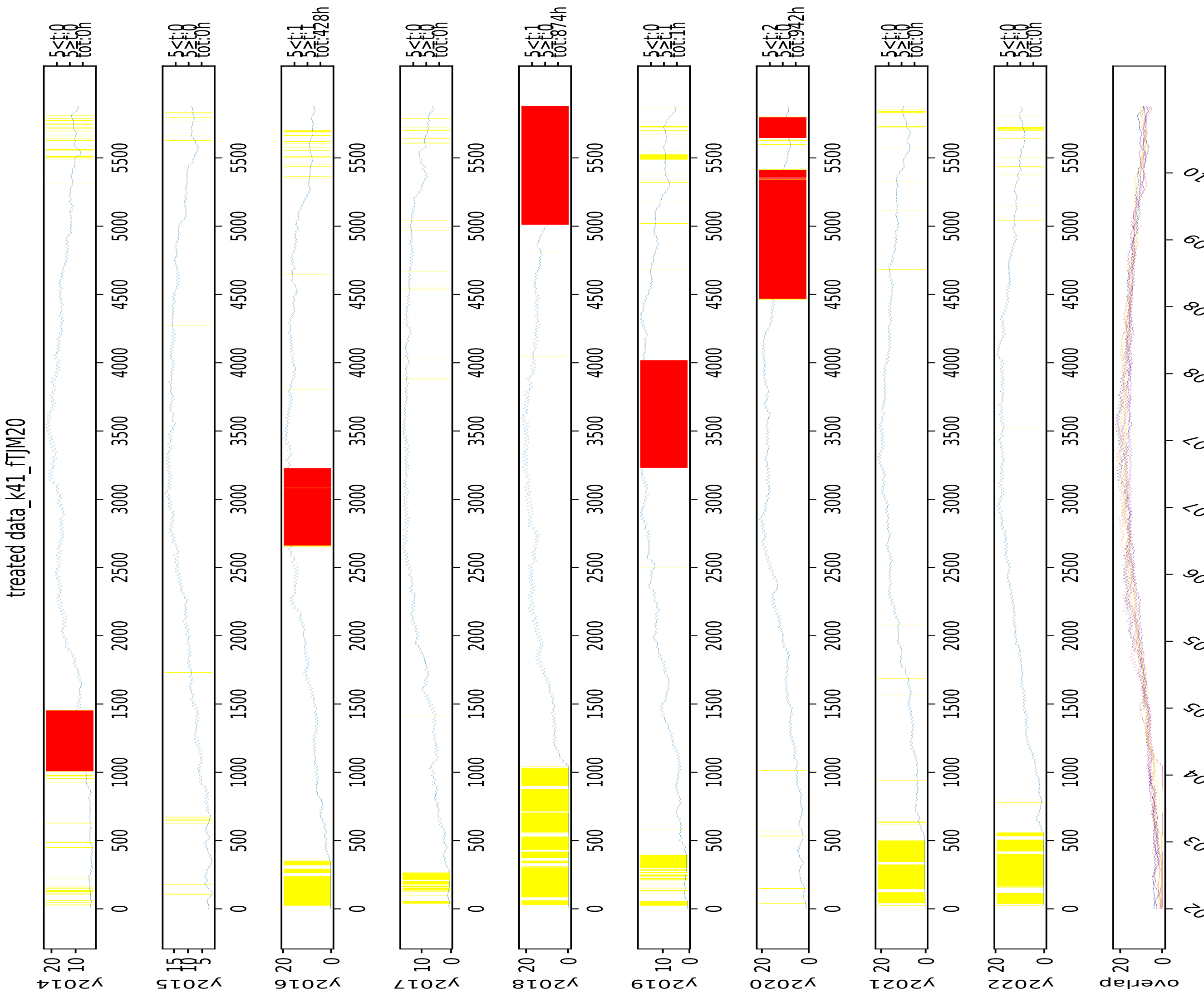


Figure 196: Visual representation of missing values at station 41 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

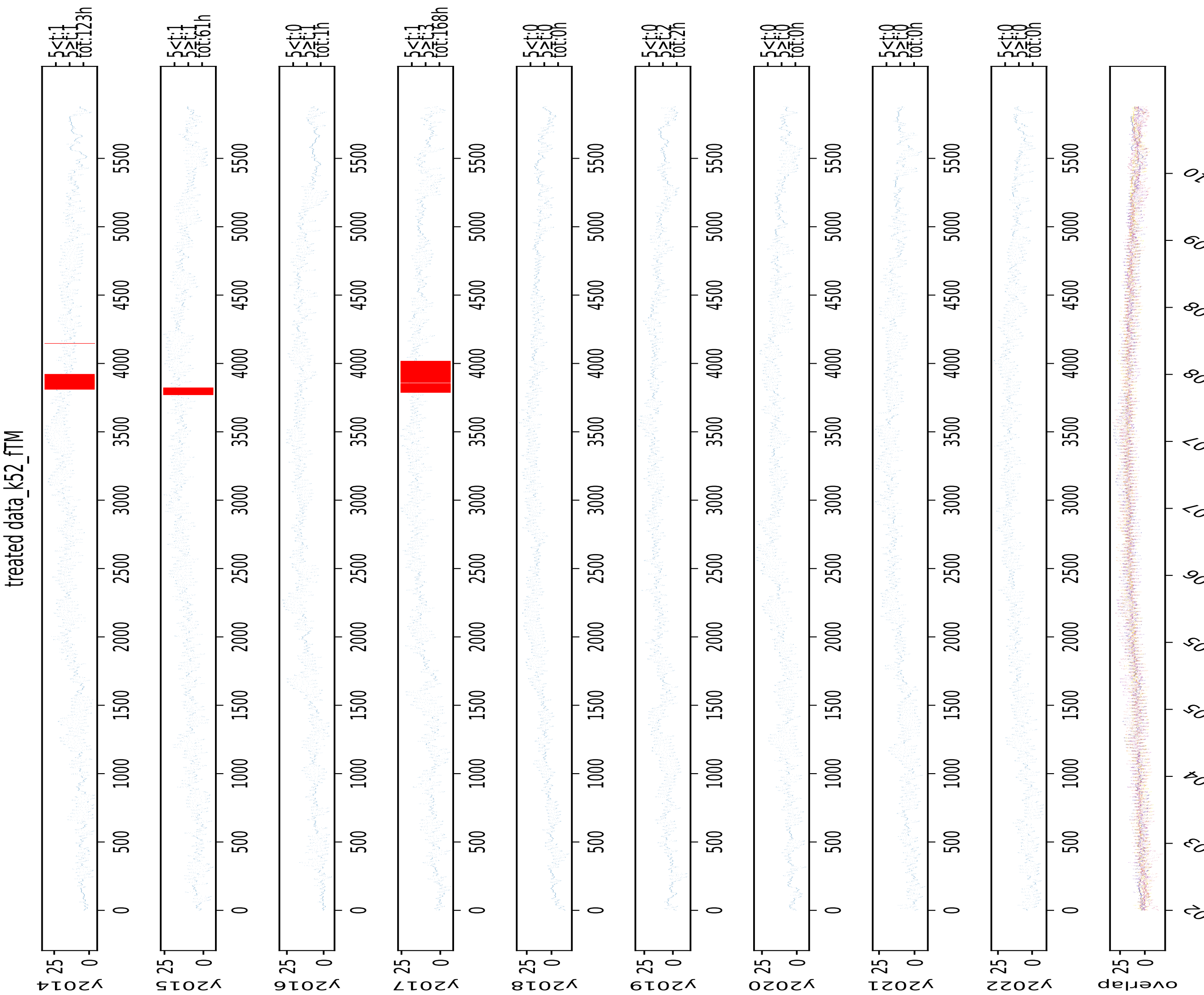


Figure 197: Visual representation of missing values at station 52 from 2014 to 2022 at the parameter ”TM” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

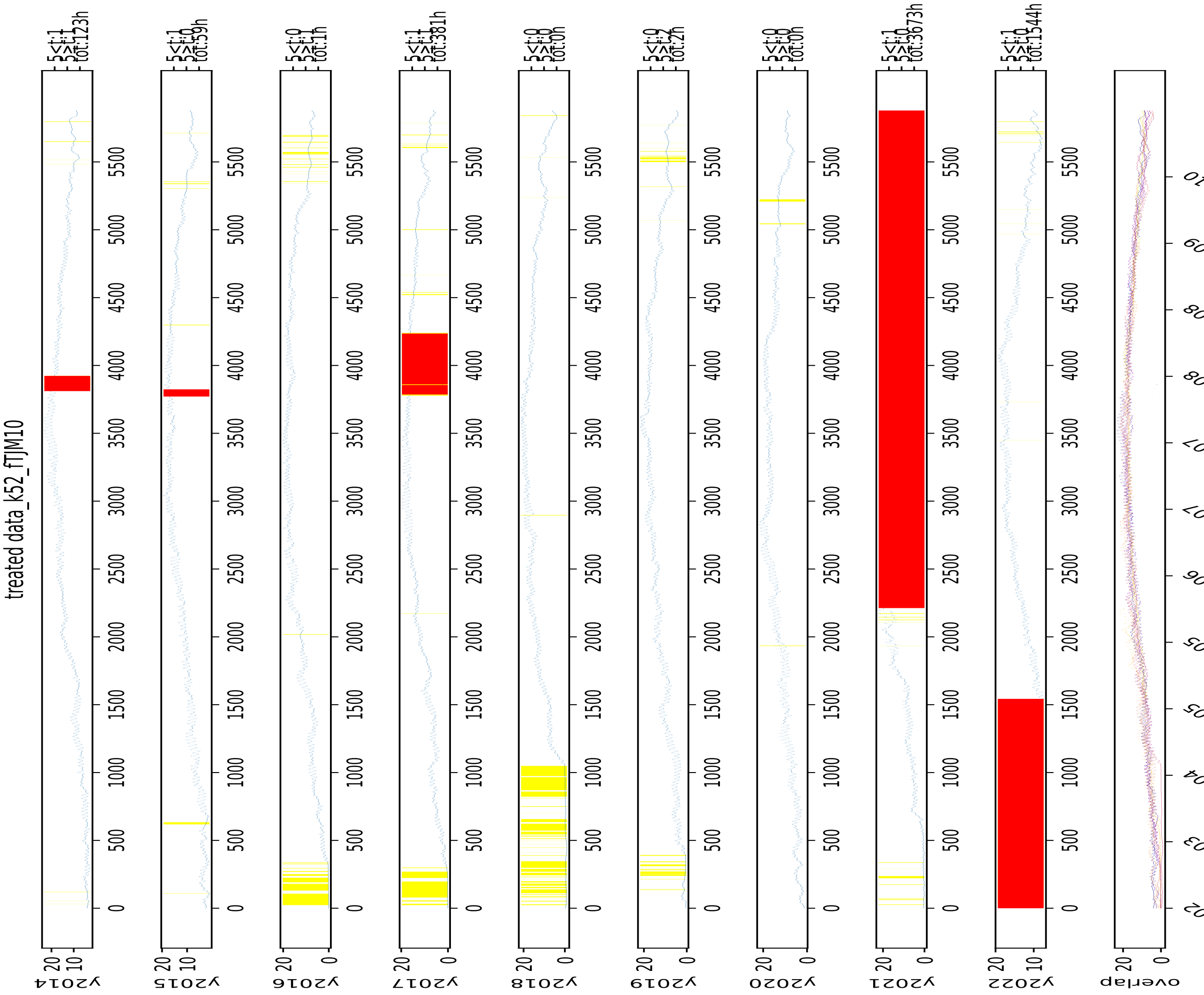


Figure 198: Visual representation of missing values at station 52 from 2014 to 2022 at the parameter ”TJM10” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

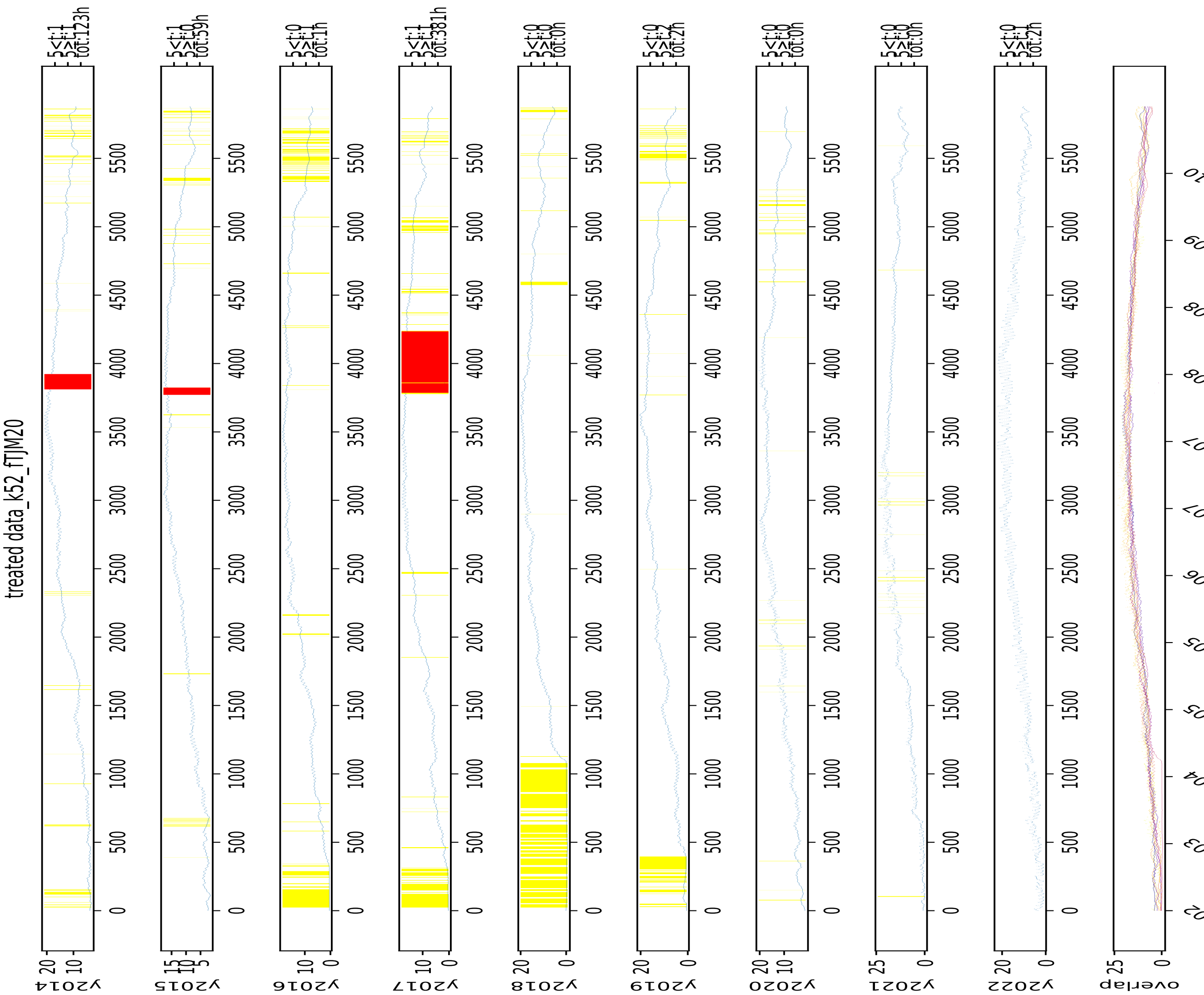


Figure 199: Visual representation of missing values at station 52 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

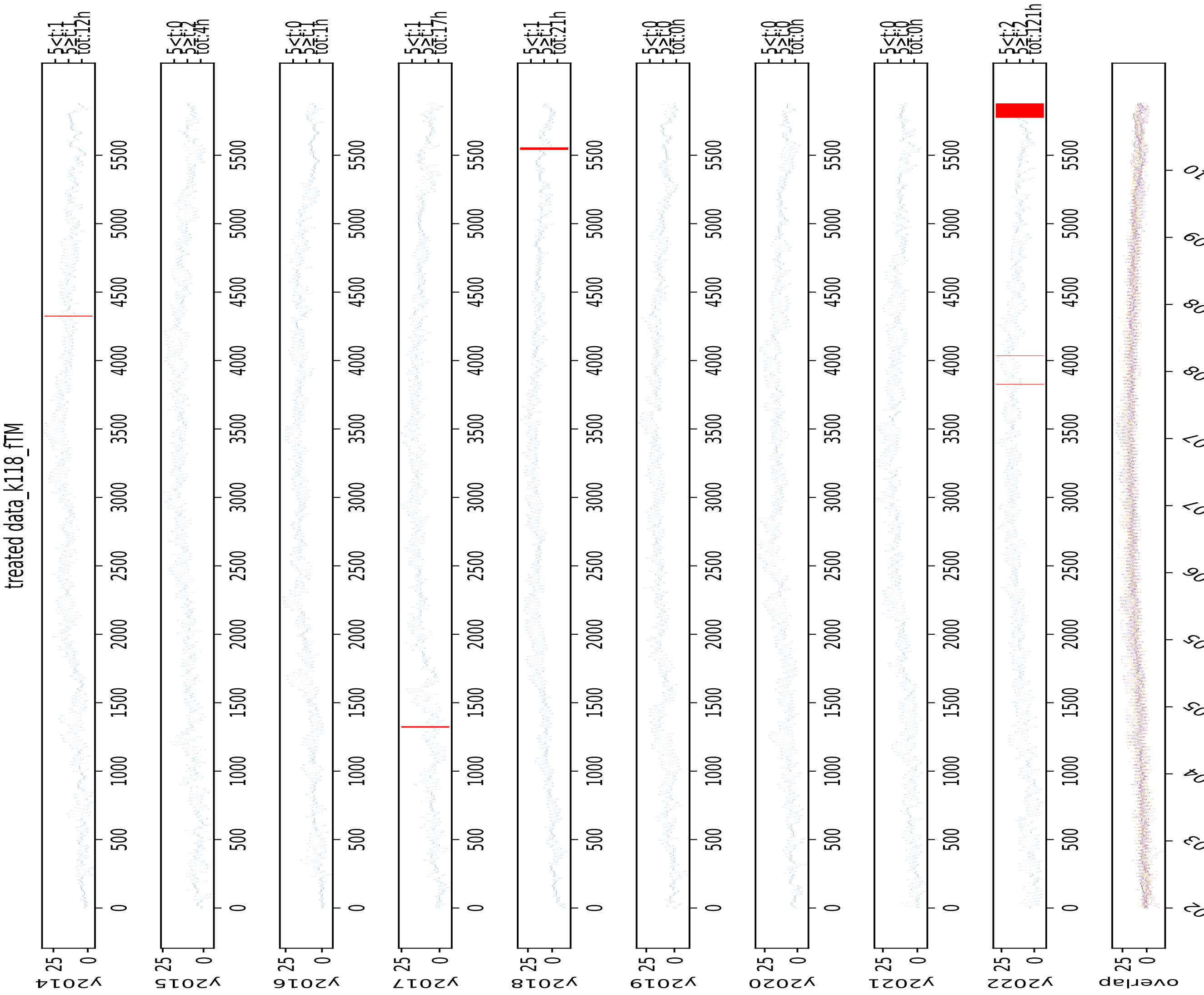


Figure 200: Visual representation of missing values at station 118 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

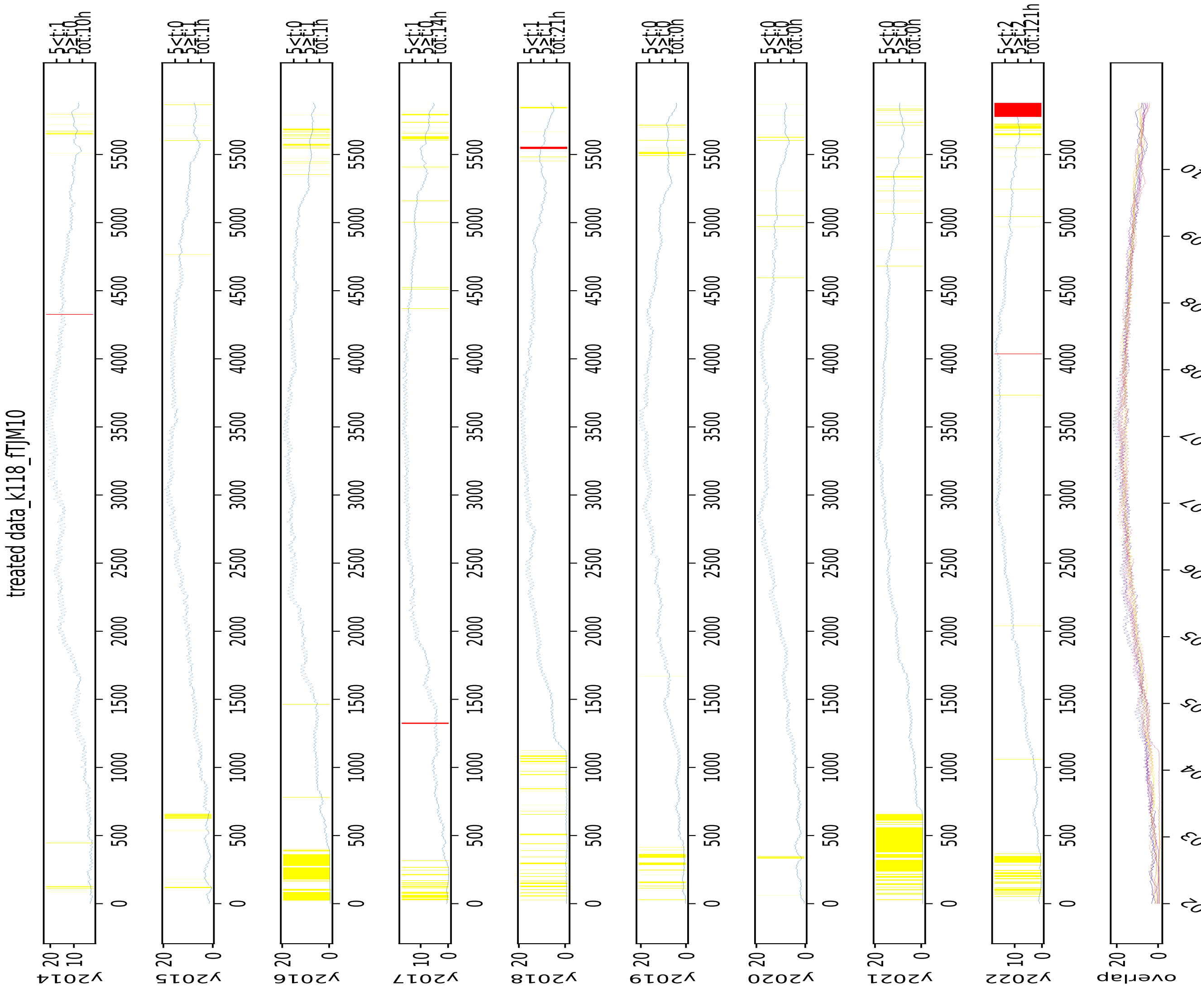


Figure 201: Visual representation of missing values at station 118 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

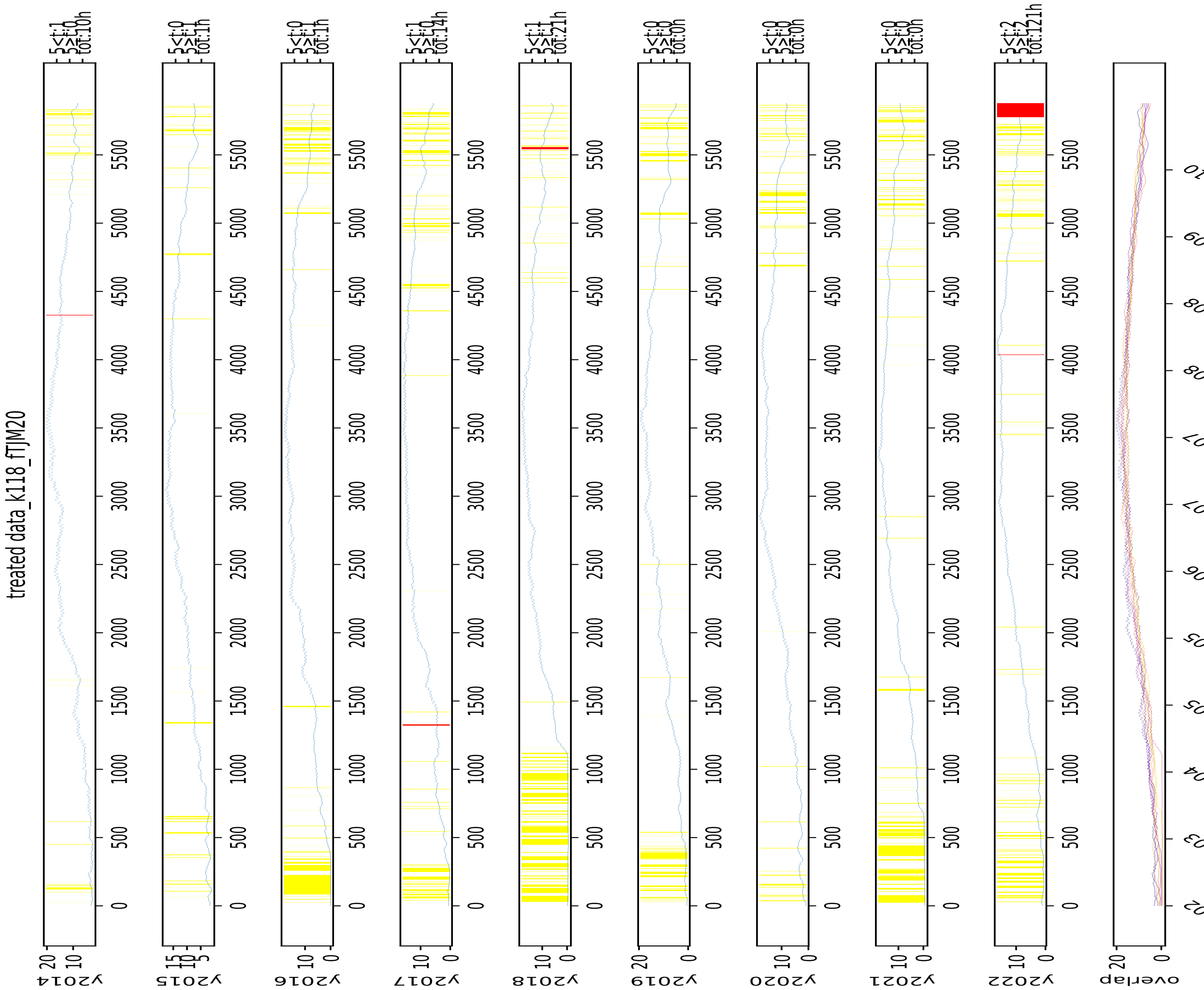


Figure 202: Visual representation of missing values at station 118 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

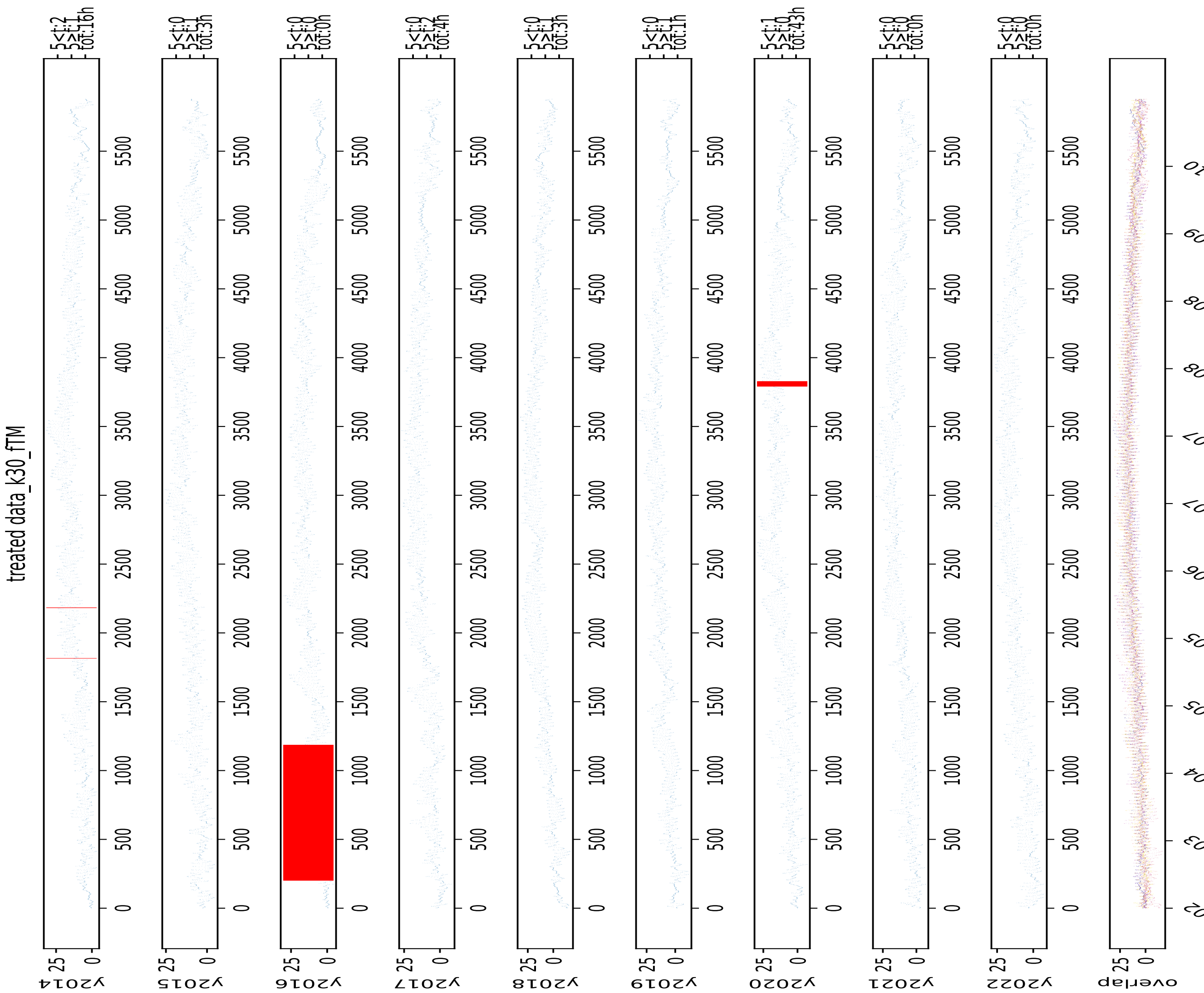


Figure 203: Visual representation of missing values at station 30 from 2014 to 2022 at the parameter ”TM” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

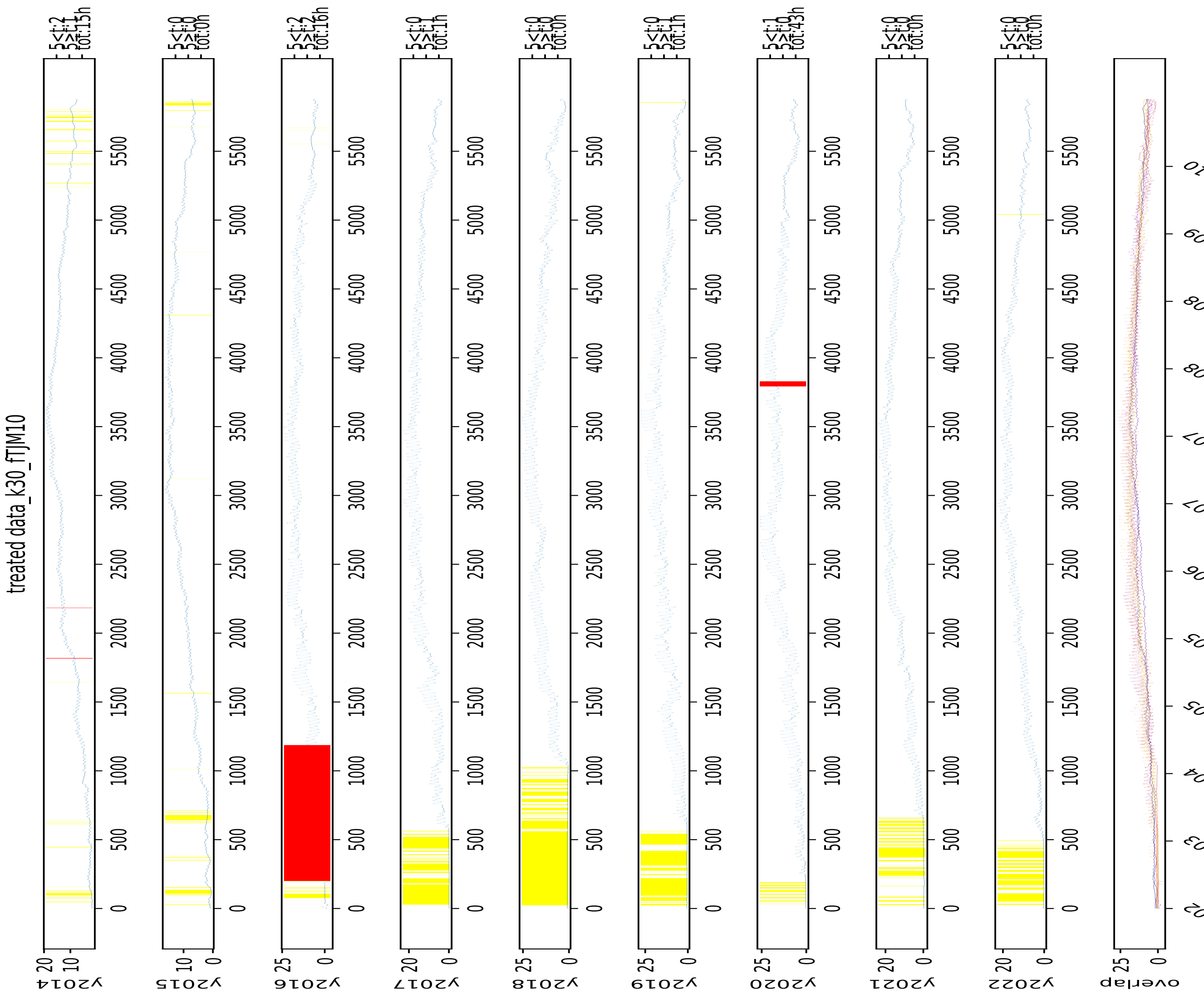


Figure 204: Visual representation of missing values at station 30 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

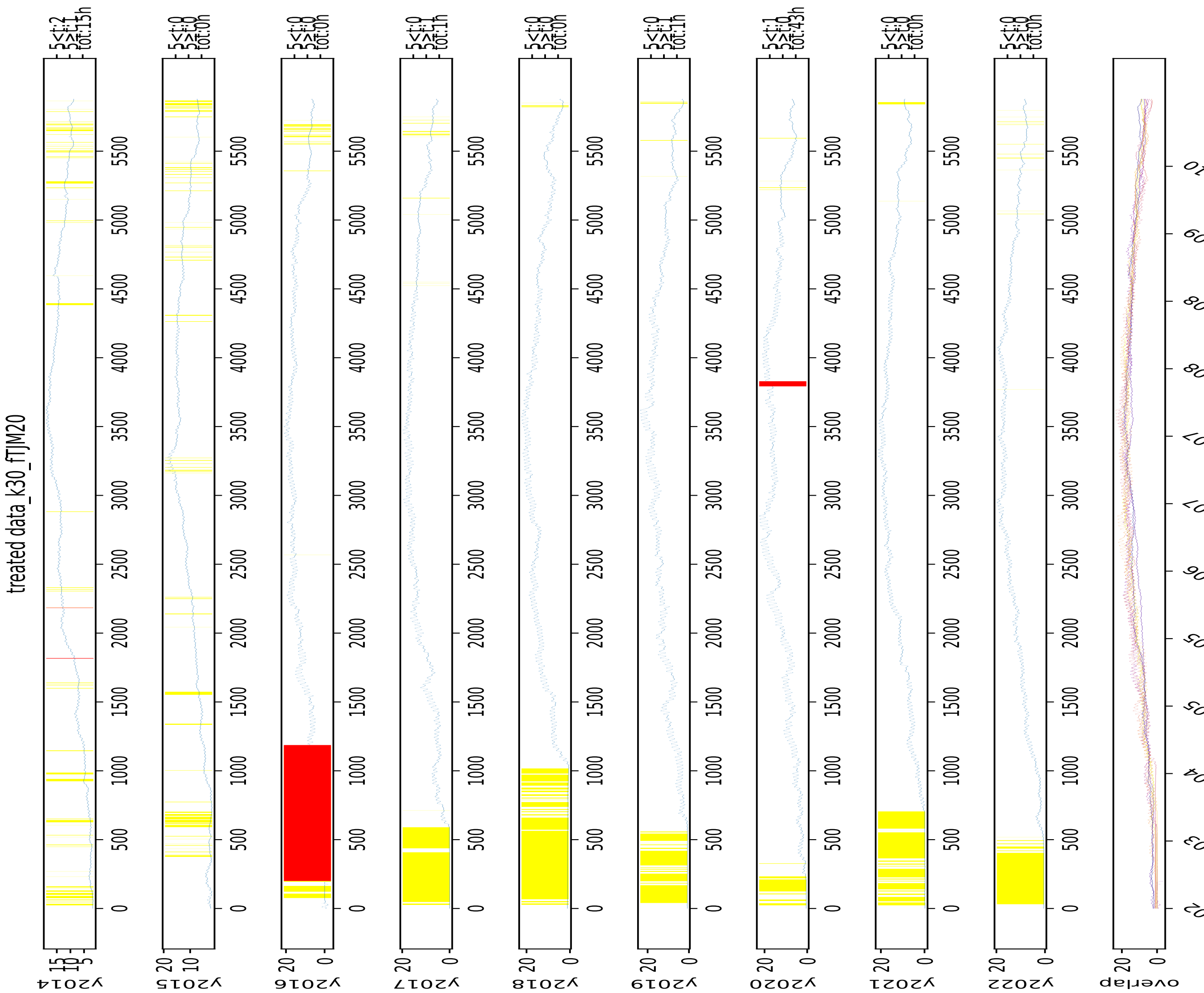


Figure 205: Visual representation of missing values at station 30 from 2014 to 2022 at the parameter ”TJM20” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

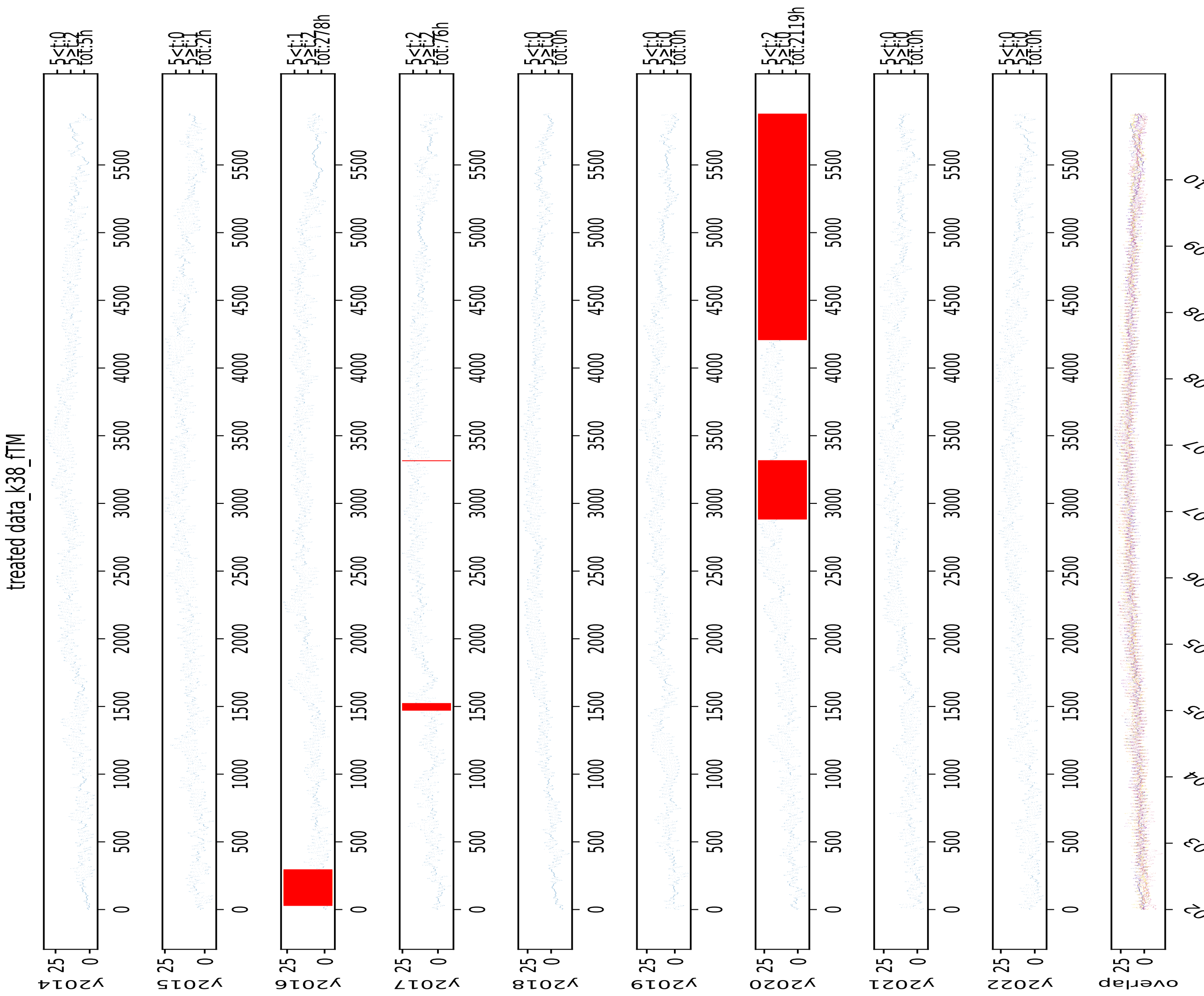


Figure 206: Visual representation of missing values at station 38 from 2014 to 2022 at the parameter ”TM” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

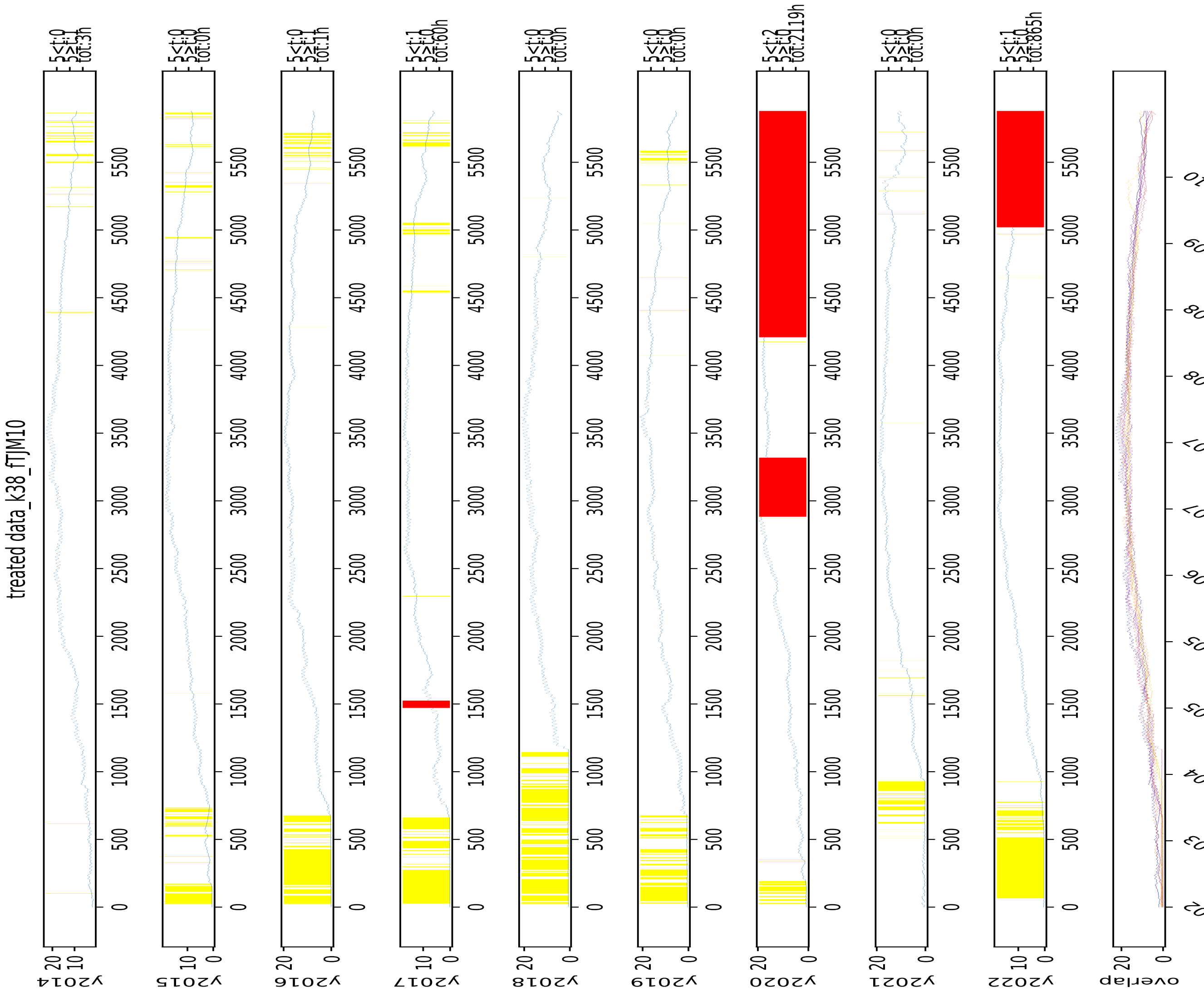


Figure 207: Visual representation of missing values at station 38 from 2014 to 2022 at the parameter ”TJM10” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

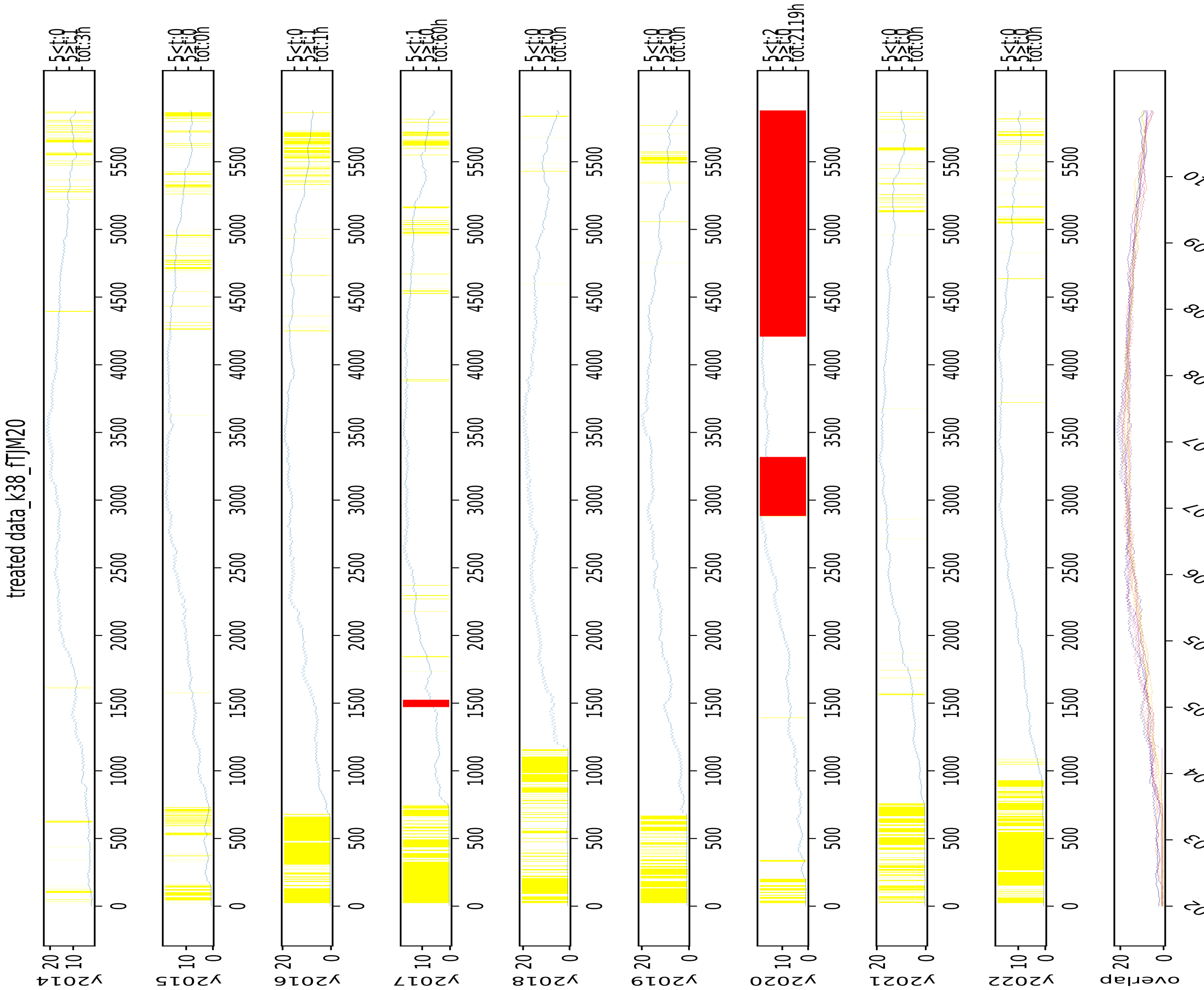


Figure 208: Visual representation of missing values at station 38 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

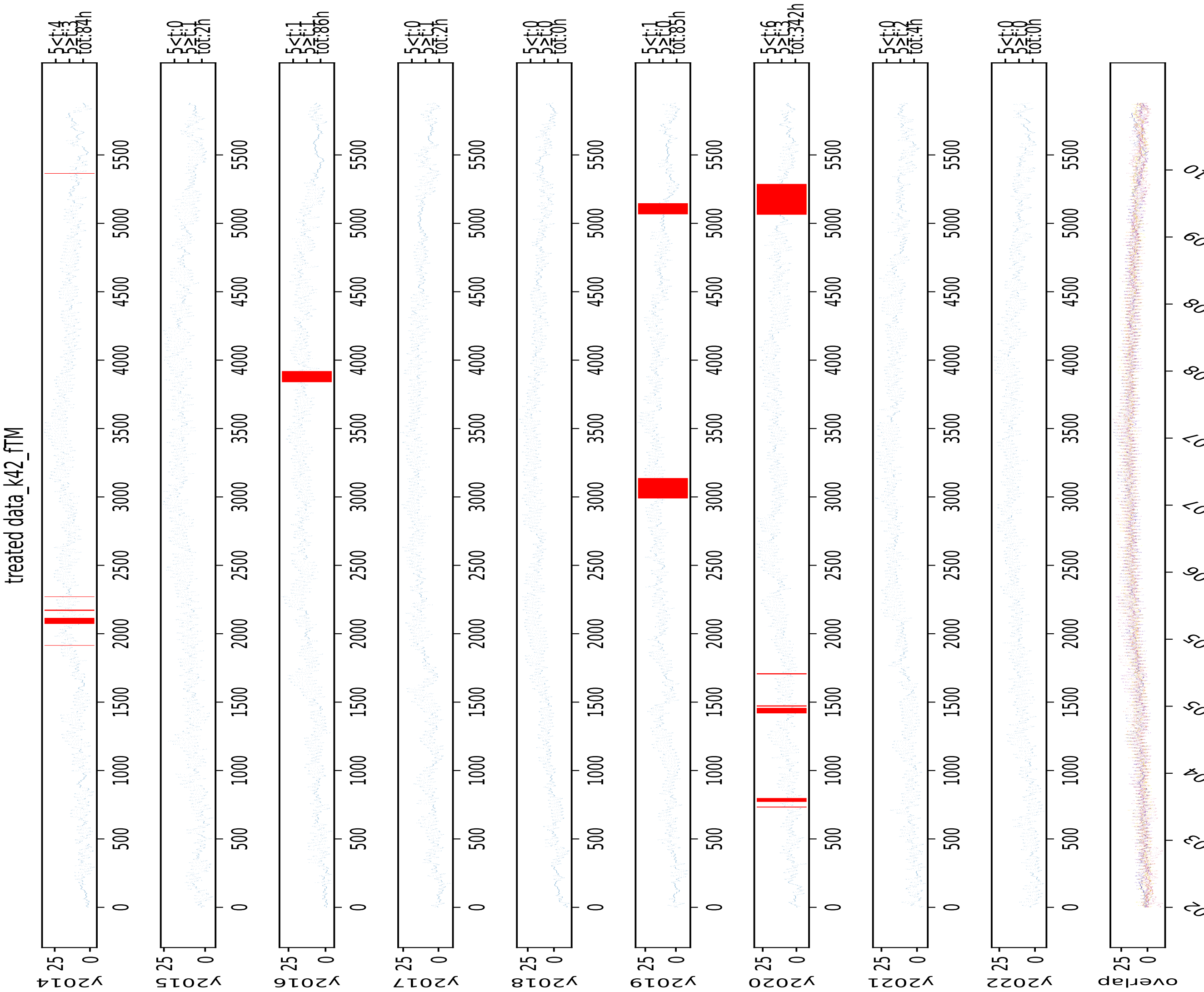


Figure 209: Visual representation of missing values at station 42 from 2014 to 2022 at the parameter "TM" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

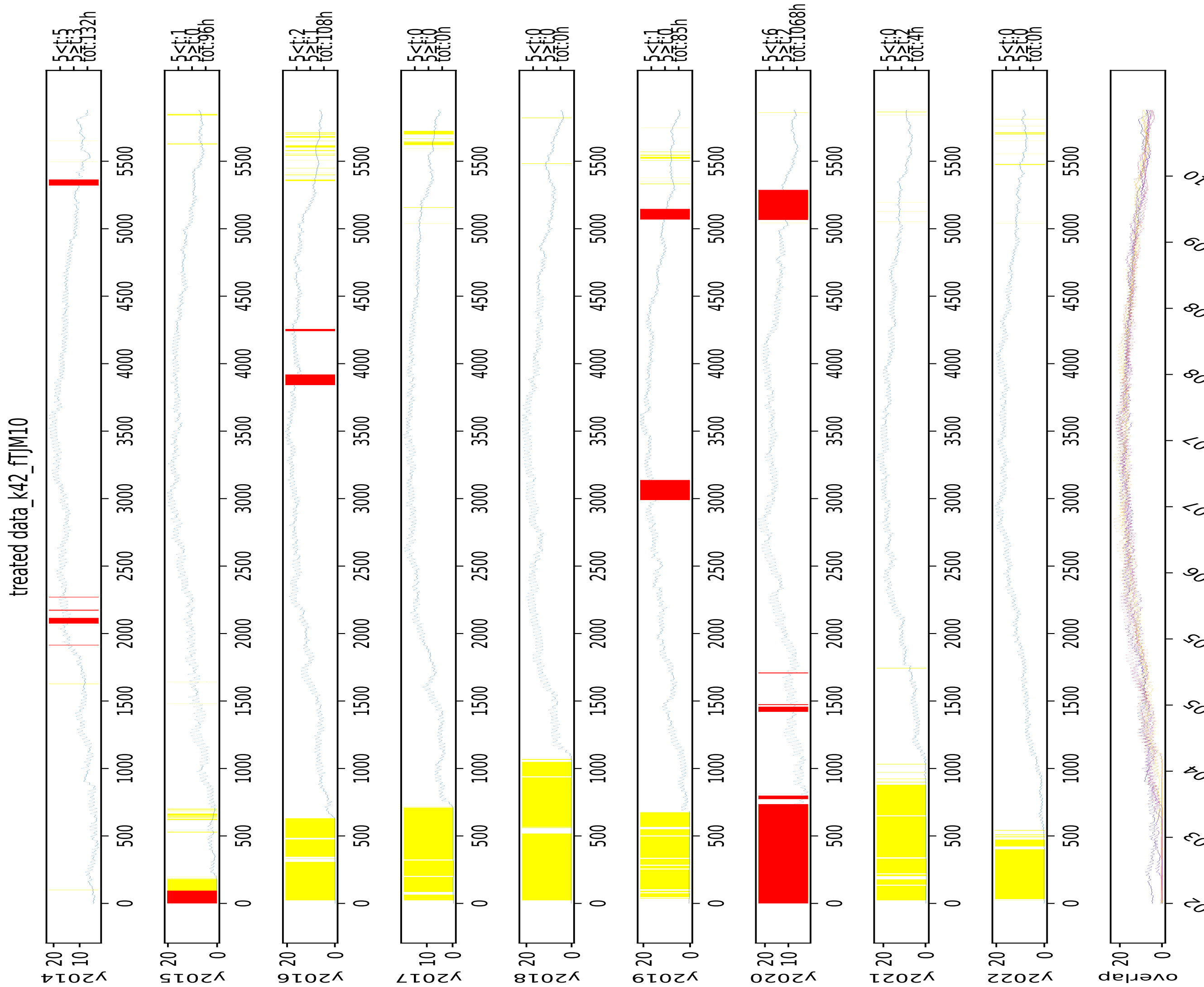


Figure 210: Visual representation of missing values at station 42 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

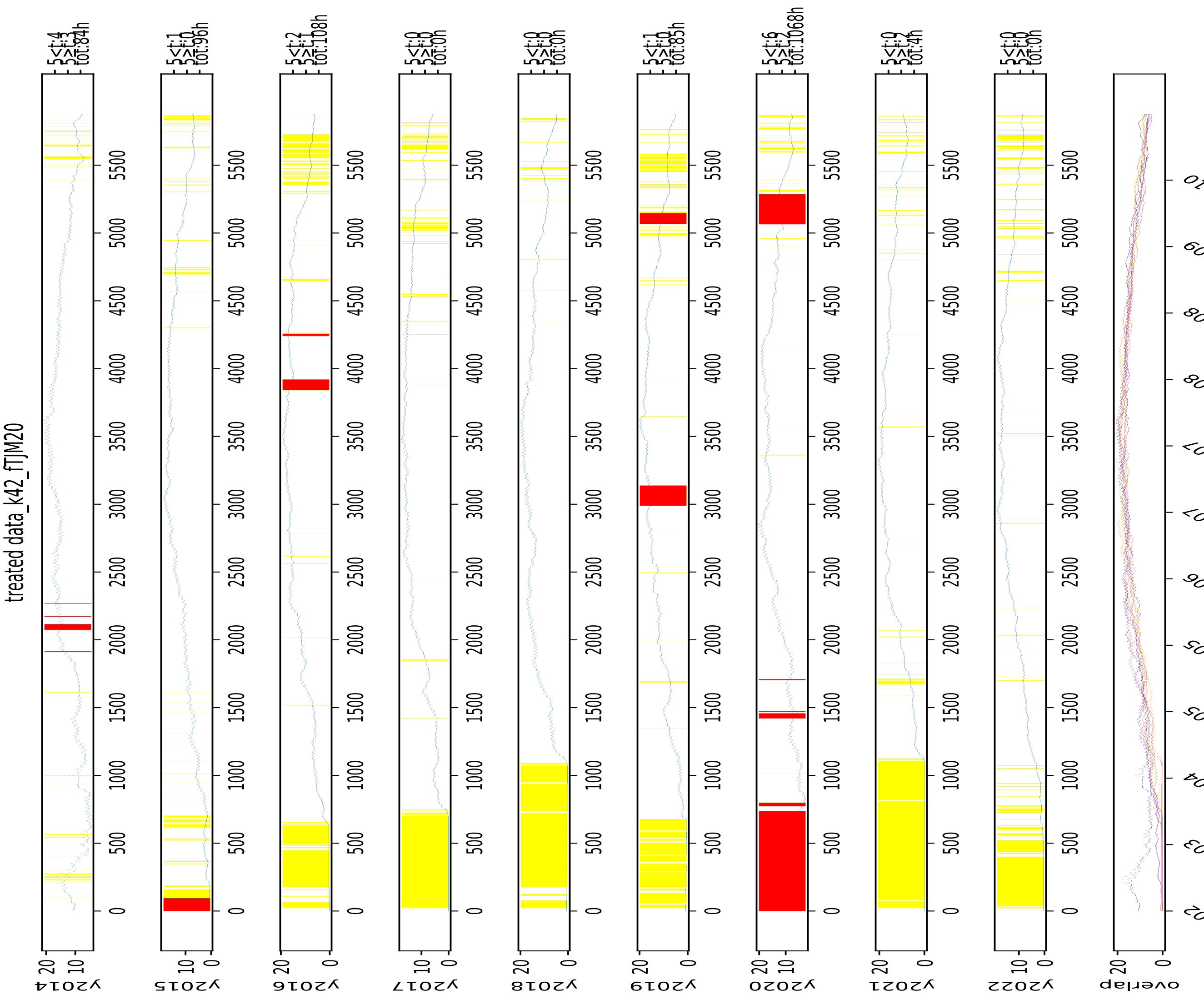


Figure 211: Visual representation of missing values at station 42 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

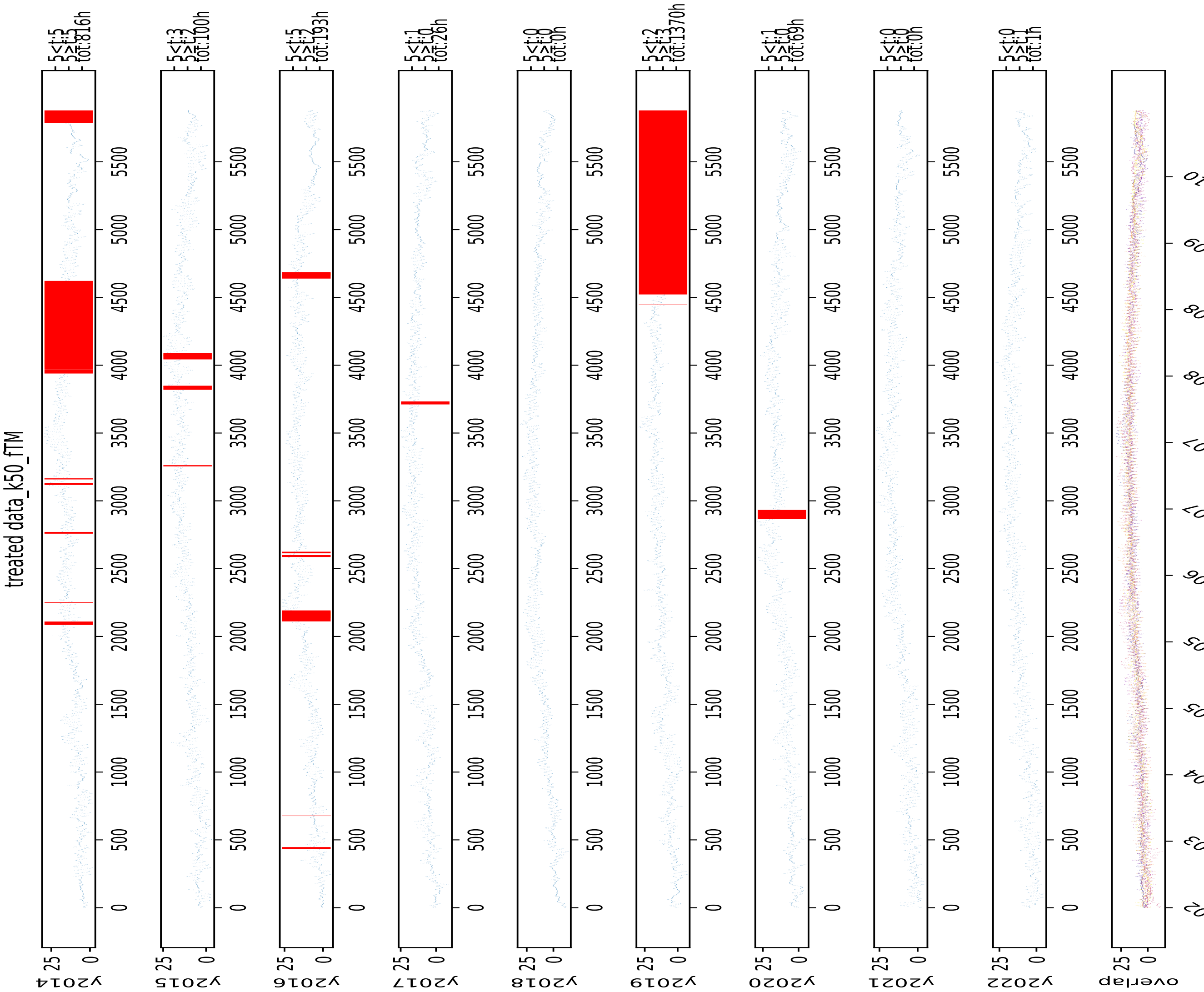


Figure 212: Visual representation of missing values at station 50 from 2014 to 2022 at the parameter ”TM” after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code ”NULL”).

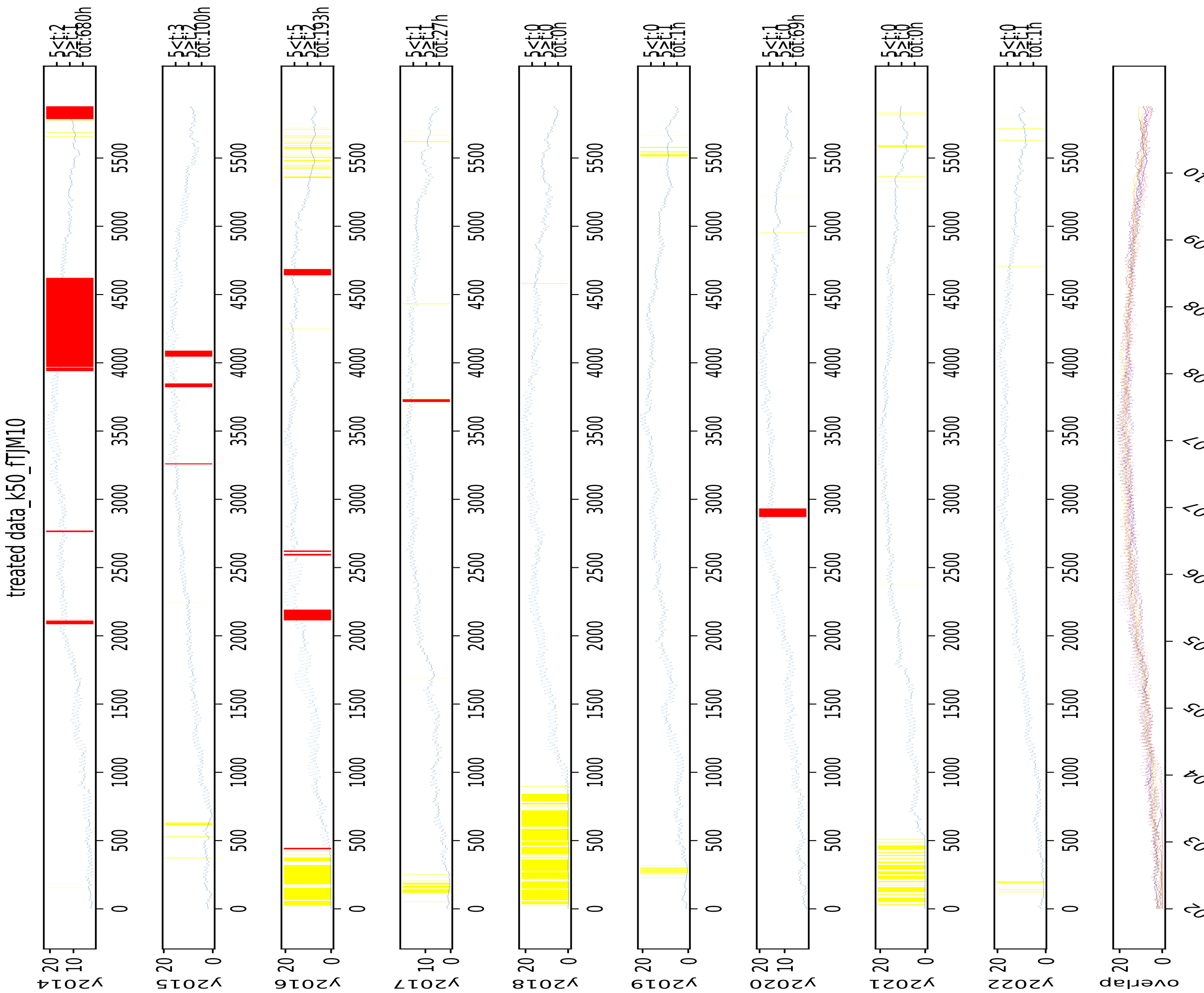


Figure 213: Visual representation of missing values at station 50 from 2014 to 2022 at the parameter "TJM10" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

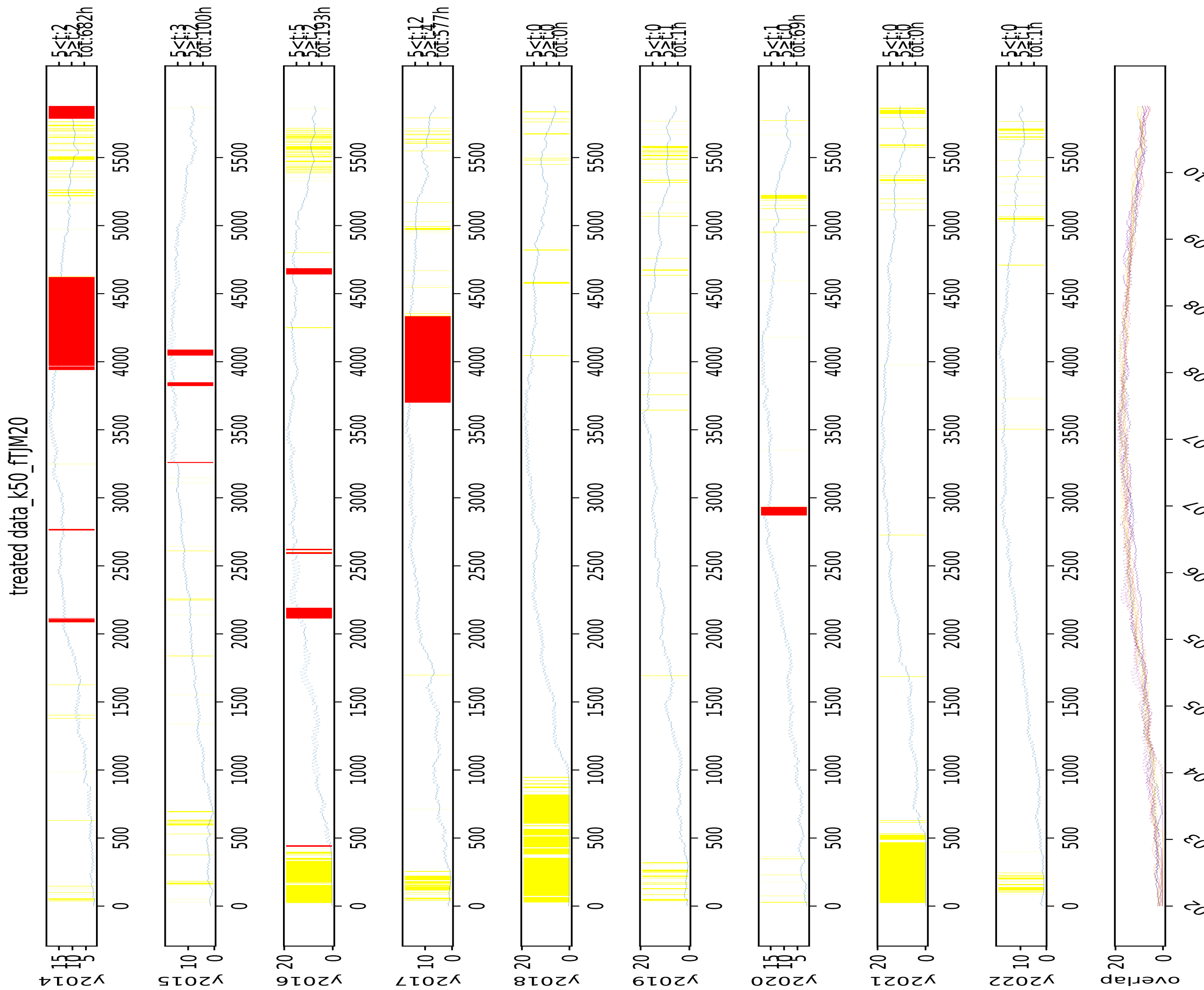


Figure 214: Visual representation of missing values at station 50 from 2014 to 2022 at the parameter "TJM20" after treating for outliers. The left numbers indicated how many hours that are missing and how many of them are shorter than or longer than 5 hours, however for this visualisation they indicate the untreated version of the data. The yellow markings indicate possible outliers based on the given year, all markings was checked if they were actual outliers. The red colouring indicate missing values in the data (represented in the data with code "NULL").

B Tables

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	2.676	2.06	0.528	-0.328	-1	0.756
region	Østfold	2.564	2	0.176	-0.324	-1	0.8
region	Vestfold	2.565	1.958	0.785	-0.317	-1	0.81
region	Trøndelag	2.938	2.279	0.75	-0.321	-1	0.477
region	Innlandet	2.612	1.997	0.379	-0.32	-1	0.799
local	52	2.504	1.976	-1.2	-0.327	-1	0.803
local	41	2.135	1.665	0.13	-0.322	-1	0.872
local	37	2.513	1.938	0.067	-0.322	-1	0.83
local	118	3.029	2.422	1.722	-0.323	-1	0.656
local	50	2.176	1.7	0.815	-0.327	-1	0.836
local	42	2.739	2.099	0.746	-0.323	-1	0.807
local	38	2.983	2.323	1.149	-0.333	-1	0.736
local	30	2.276	1.708	0.428	-0.328	-1	0.859
local	57	3.079	2.419	0.744	-0.33	-1	0.617
local	39	2.79	2.186	0.633	-0.322	-1	0.615
local	34	3.163	2.427	0.797	-0.329	-1	-0.547
local	15	2.706	2.094	0.827	-0.329	-1	0.484
local	27	2.455	1.885	0.335	-0.328	-1	0.839
local	26	2.757	2.105	0.892	-0.324	-1	0.801
local	17	3.023	2.265	0.137	-0.33	-1	0.762
local	11	2.346	1.847	-0.023	-0.33	-1	0.755

Table 18: Results from hourly version of the Plauborg model for 20cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	2.529	1.926	0.597	-0.449	-1	0.794
region	Østfold	2.448	1.894	0.512	-0.45	-1	0.816
region	Vestfold	2.412	1.81	0.733	-0.43	-1	0.846
region	Trøndelag	2.822	2.176	0.781	-0.439	-1	0.547
region	Innlandet	2.382	1.805	0.312	-0.448	-1	0.847
local	52	2.514	1.964	-0.349	-0.446	-1	0.636
local	41	1.938	1.519	0.151	-0.445	-1	0.903
local	37	2.344	1.804	0.237	-0.446	-1	0.857
local	118	2.928	2.322	1.639	-0.442	-1	0.706
local	50	1.908	1.472	0.558	-0.442	-1	0.884
local	42	2.501	1.885	0.703	-0.447	-1	0.852
local	38	3.055	2.368	1.363	-0.447	-1	0.754
local	30	2.072	1.555	0.354	-0.441	-1	0.892
local	57	2.906	2.263	0.677	-0.443	-1	0.677
local	39	2.77	2.151	0.701	-0.44	-1	0.633
local	34	3.013	2.306	0.845	-0.446	-1	-0.193
local	15	2.589	1.991	0.903	-0.444	-1	0.562
local	27	2.277	1.724	0.163	-0.444	-1	0.872
local	26	2.532	1.918	0.821	-0.442	-1	0.843
local	17	2.649	1.979	0.065	-0.44	-1	0.828
local	11	2.146	1.666	0.038	-0.443	-1	0.823

Table 19: Results from hourly version of the Plauborg model for 10cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.91	1.536	0.644	-1.918	-2	0.876
region	Østfold	1.94	1.541	-0.073	-1.917	-2	0.885
region	Vestfold	1.71	1.341	0.236	-1.899	-2	0.915
region	Trøndelag	1.843	1.56	1.461	-1.913	-2	0.794
region	Innlandet	2.16	1.735	1.02	-1.909	-2	0.863
local	52	2.33	1.873	-1.402	-1.908	-2	0.83
local	41	1.748	1.409	-0.371	-1.902	-2	0.914
local	37	1.877	1.496	0.353	-1.913	-2	0.905
local	118	1.742	1.384	1.139	-1.917	-2	0.886
local	50	1.251	0.985	0.096	-1.912	-2	0.946
local	42	1.966	1.54	0.346	-1.917	-2	0.901
local	38	1.721	1.367	0.515	-1.906	-2	0.912
local	30	1.817	1.471	-0.014	-1.918	-2	0.91
local	57	1.841	1.538	1.427	-1.919	-2	0.863
local	39	1.729	1.468	1.402	-1.912	-2	0.852
local	34	2.104	1.836	1.816	-1.91	-2	0.316
local	15	1.681	1.411	1.215	-1.903	-2	0.801
local	27	1.924	1.534	0.753	-1.905	-2	0.901
local	26	2.528	2.101	1.578	-1.912	-2	0.833
local	17	2.448	1.911	1.401	-1.912	-2	0.844
local	11	1.735	1.443	0.463	-1.91	-2	0.866

Table 20: Results from daily version of the Plauborg model for 20cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	2.074	1.621	0.608	-1.261	-2	0.861
region	Østfold	2.168	1.704	0.24	-1.27	-2	0.856
region	Vestfold	2.022	1.564	0.219	-1.263	-2	0.892
region	Trøndelag	1.957	1.528	1.235	-1.257	-2	0.782
region	Innlandet	2.165	1.71	0.714	-1.269	-2	0.873
local	52	2.418	1.837	-0.636	-1.265	-2	0.664
local	41	1.975	1.587	-0.293	-1.266	-2	0.9
local	37	2.206	1.755	0.373	-1.26	-2	0.873
local	118	2.165	1.697	1.137	-1.263	-2	0.839
local	50	1.395	1.105	-0.046	-1.265	-2	0.938
local	42	2.239	1.75	0.333	-1.266	-2	0.881
local	38	2.42	1.908	0.667	-1.261	-2	0.845
local	30	1.914	1.519	-0.046	-1.271	-2	0.908
local	57	1.978	1.547	1.108	-1.266	-2	0.85
local	39	1.896	1.455	1.193	-1.266	-2	0.828
local	34	2.143	1.687	1.535	-1.261	-2	0.397
local	15	1.806	1.428	1.114	-1.262	-2	0.787
local	27	2.063	1.627	0.396	-1.266	-2	0.895
local	26	2.43	1.937	1.251	-1.267	-2	0.855
local	17	2.26	1.78	0.921	-1.263	-2	0.875
local	11	1.879	1.504	0.339	-1.262	-2	0.864

Table 21: Results from daily version of the Plauborg model for 10cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	4.504	3.474	2.487	-0.796	-1	0.308
region	Østfold	4.348	3.363	1.901	-0.796	-1	0.424
region	Vestfold	4.564	3.47	2.297	-0.796	-1	0.397
region	Trøndelag	4.438	3.508	3.175	-0.796	-1	-0.194
region	Innlandet	4.688	3.568	2.601	-0.796	-1	0.353
local	52	3.556	2.841	0.559	-0.796	-1	0.604
local	41	4.248	3.286	1.677	-0.796	-1	0.491
local	37	4.754	3.675	2.174	-0.796	-1	0.391
local	118	4.726	3.654	3.208	-0.796	-1	0.162
local	50	4.048	3.025	2.207	-0.796	-1	0.434
local	42	4.863	3.741	2.364	-0.796	-1	0.393
local	38	4.832	3.682	2.601	-0.796	-1	0.308
local	30	4.465	3.433	2.015	-0.796	-1	0.456
local	57	4.655	3.636	3.153	-0.796	-1	0.125
local	39	4.31	3.39	3.083	-0.796	-1	0.081
local	34	4.583	3.675	3.471	-0.796	-1	-2.248
local	15	4.198	3.342	3.006	-0.796	-1	-0.241
local	27	4.672	3.547	2.535	-0.796	-1	0.415
local	26	5.17	4.009	3.282	-0.796	-1	0.302
local	17	5.049	3.84	2.939	-0.796	-1	0.336
local	11	3.821	2.924	1.692	-0.796	-1	0.35

Table 22: Results from the linear regression model for 20cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	4.231	3.267	2.303	-0.638	-1	0.423
region	Østfold	4.236	3.28	2.015	-0.638	-1	0.45
region	Vestfold	4.277	3.26	2.019	-0.638	-1	0.517
region	Trøndelag	4.133	3.274	2.893	-0.638	-1	0.028
region	Innlandet	4.282	3.254	2.246	-0.638	-1	0.504
local	52	3.679	2.889	1.226	-0.638	-1	0.221
local	41	3.976	3.07	1.494	-0.638	-1	0.593
local	37	4.501	3.503	2.07	-0.638	-1	0.473
local	118	4.5	3.486	2.93	-0.638	-1	0.306
local	50	3.611	2.702	1.766	-0.638	-1	0.584
local	42	4.571	3.525	2.109	-0.638	-1	0.506
local	38	4.815	3.741	2.502	-0.638	-1	0.388
local	30	4.053	3.106	1.733	-0.638	-1	0.588
local	57	4.293	3.356	2.775	-0.638	-1	0.295
local	39	4.057	3.193	2.835	-0.638	-1	0.213
local	34	4.262	3.415	3.176	-0.638	-1	-1.386
local	15	3.918	3.141	2.799	-0.638	-1	-0.003
local	27	4.272	3.236	2.078	-0.638	-1	0.551
local	26	4.714	3.651	2.902	-0.638	-1	0.456
local	17	4.518	3.432	2.506	-0.638	-1	0.501
local	11	3.567	2.713	1.529	-0.638	-1	0.51

Table 23: Results from the linear regression model for 10cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.7	1.294	0.244	-2.662	-3	0.901
region	Østfold	1.778	1.405	-0.482	-2.682	-3	0.903
region	Vestfold	1.417	1.118	-0.152	-2.646	-3	0.941
region	Trøndelag	1.622	1.255	0.993	-2.65	-3	0.841
region	Innlandet	1.979	1.419	0.706	-2.654	-3	0.884
local	52	2.517	2.148	-1.806	-2.633	-3	0.8
local	41	1.638	1.358	-0.779	-2.665	-3	0.924
local	37	1.414	1.146	-0.036	-2.639	-3	0.946
local	118	1.258	0.955	0.727	-2.629	-3	0.941
local	50	1.06	0.854	-0.355	-2.641	-3	0.961
local	42	1.628	1.307	-0.032	-2.655	-3	0.931
local	38	1.224	0.985	0.145	-2.664	-3	0.955
local	30	1.66	1.325	-0.368	-2.657	-3	0.924
local	57	1.575	1.194	0.982	-2.642	-3	0.899
local	39	1.563	1.217	0.917	-2.66	-3	0.878
local	34	1.831	1.489	1.339	-2.684	-3	0.502
local	15	1.518	1.138	0.763	-2.63	-3	0.836
local	27	1.626	1.2	0.362	-2.652	-3	0.929
local	26	2.187	1.595	1.246	-2.667	-3	0.874
local	17	2.335	1.634	1.106	-2.663	-3	0.857
local	11	1.869	1.344	0.244	-2.627	-3	0.844

Table 24: Results from the GRU model for 20cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.807	1.418	0.027	-0.013	-1	0.894
region	Østfold	1.917	1.507	-0.377	-0.042	-1	0.887
region	Vestfold	1.861	1.474	-0.435	-0.044	-1	0.908
region	Trøndelag	1.537	1.199	0.699	0.014	-1	0.866
region	Innlandet	1.911	1.513	0.212	-0.03	-1	0.901
local	52	2.289	1.686	-1.195	-0.034	-1	0.696
local	41	1.983	1.6	-0.937	0.008	-1	0.898
local	37	1.89	1.528	-0.155	-0.027	-1	0.906
local	118	1.619	1.288	0.439	-0.002	-1	0.911
local	50	1.442	1.183	-0.719	-0.036	-1	0.933
local	42	2.062	1.645	-0.309	-0.032	-1	0.899
local	38	2.005	1.569	0.018	-0.013	-1	0.893
local	30	1.881	1.507	-0.695	-0.029	-1	0.91
local	57	1.5	1.19	0.589	-0.045	-1	0.914
local	39	1.553	1.214	0.671	-0.014	-1	0.884
local	34	1.69	1.315	1.031	-0.033	-1	0.639
local	15	1.404	1.086	0.533	-0.028	-1	0.87
local	27	1.853	1.486	-0.139	-0.021	-1	0.915
local	26	2.081	1.654	0.765	-0.011	-1	0.893
local	17	1.976	1.54	0.499	-0.002	-1	0.904
local	11	1.722	1.367	-0.189	-0.027	-1	0.885

Table 25: Results from the GRU model for 10cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.575	1.236	0.012	-2.01	-3	0.915
region	Østfold	1.828	1.452	-0.583	-1.981	-2	0.898
region	Vestfold	1.409	1.112	-0.149	-1.981	-2	0.942
region	Trøndelag	1.456	1.141	0.584	-1.953	-2	0.872
region	Innlandet	1.572	1.236	0.247	-1.997	-2	0.927
local	52	2.565	2.131	-1.905	-2.004	-3	0.792
local	41	1.574	1.319	-0.749	-1.96	-2	0.93
local	37	1.562	1.283	-0.397	-1.996	-2	0.934
local	118	1.351	1.063	0.756	-1.943	-2	0.932
local	50	1.051	0.825	-0.151	-1.97	-2	0.961
local	42	1.585	1.259	-0.096	-2	-2	0.935
local	38	1.381	1.088	0.113	-1.996	-2	0.943
local	30	1.553	1.275	-0.464	-1.981	-2	0.934
local	57	1.527	1.198	0.549	-1.965	-2	0.905
local	39	1.404	1.095	0.514	-1.952	-2	0.902
local	34	1.578	1.228	0.807	-1.969	-2	0.63
local	15	1.311	1.052	0.486	-2.003	-3	0.878
local	27	1.52	1.205	0.037	-1.971	-2	0.938
local	26	1.677	1.314	0.763	-1.99	-2	0.926
local	17	1.629	1.208	0.434	-1.962	-2	0.931
local	11	1.467	1.199	-0.219	-1.975	-2	0.904

Table 26: Results from the BiGRU model for 20cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.722	1.36	-0.037	-1.734	-2	0.904
region	Østfold	1.793	1.408	-0.277	-1.717	-2	0.901
region	Vestfold	1.673	1.32	-0.211	-1.685	-2	0.925
region	Trøndelag	1.672	1.313	0.345	-1.7	-2	0.842
region	Innlandet	1.758	1.411	-0.023	-1.68	-2	0.916
local	52	2.31	1.687	-1.149	-1.7	-2	0.69
local	41	1.81	1.498	-0.716	-1.719	-2	0.915
local	37	1.659	1.369	-0.313	-1.705	-2	0.928
local	118	1.552	1.196	0.709	-1.681	-2	0.918
local	50	1.288	1.049	-0.397	-1.696	-2	0.947
local	42	1.818	1.455	-0.139	-1.709	-2	0.921
local	38	1.777	1.34	0.229	-1.731	-2	0.916
local	30	1.76	1.436	-0.507	-1.702	-2	0.922
local	57	1.736	1.377	0.226	-1.719	-2	0.884
local	39	1.659	1.286	0.283	-1.735	-2	0.867
local	34	1.732	1.336	0.554	-1.709	-2	0.621
local	15	1.562	1.256	0.334	-1.74	-2	0.839
local	27	1.774	1.438	-0.304	-1.716	-2	0.922
local	26	1.739	1.386	0.51	-1.658	-2	0.925
local	17	1.693	1.311	0.077	-1.705	-2	0.93
local	11	1.796	1.467	-0.375	-1.704	-2	0.875

Table 27: Results from the BiGRU model for 10cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.695	1.349	0.068	-1.797	-2	0.901
region	Østfold	1.902	1.532	-0.479	-1.831	-2	0.889
region	Vestfold	1.563	1.268	-0.093	-1.796	-2	0.929
region	Trøndelag	1.538	1.166	0.557	-1.787	-2	0.856
region	Innlandet	1.76	1.444	0.339	-1.786	-2	0.908
local	52	2.495	2.067	-1.788	-1.819	-2	0.803
local	41	1.745	1.455	-0.675	-1.755	-2	0.913
local	37	1.716	1.439	-0.255	-1.813	-2	0.92
local	118	1.497	1.162	0.82	-1.814	-2	0.915
local	50	1.201	0.986	-0.131	-1.832	-2	0.95
local	42	1.781	1.463	-0.034	-1.814	-2	0.918
local	38	1.54	1.244	0.189	-1.808	-2	0.929
local	30	1.67	1.379	-0.396	-1.818	-2	0.923
local	57	1.6	1.233	0.534	-1.823	-2	0.896
local	39	1.504	1.123	0.464	-1.83	-2	0.887
local	34	1.64	1.233	0.808	-1.819	-2	0.589
local	15	1.405	1.08	0.435	-1.807	-2	0.86
local	27	1.664	1.385	0.13	-1.822	-2	0.925
local	26	1.91	1.567	0.894	-1.801	-2	0.904
local	17	1.884	1.508	0.589	-1.835	-2	0.907
local	11	1.607	1.333	-0.207	-1.805	-2	0.885

Table 28: Results from the BiLSTM model for 20cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.423	1.111	0.06	-1.858	-2	0.934
region	Østfold	1.483	1.154	-0.252	-1.869	-2	0.932
region	Vestfold	1.341	1.03	-0.264	-1.832	-2	0.952
region	Trøndelag	1.467	1.133	0.524	-1.864	-2	0.877
region	Innlandet	1.4	1.135	0.225	-1.896	-2	0.947
local	52	2.08	1.479	-1.078	-1.838	-2	0.749
local	41	1.473	1.203	-0.757	-1.891	-2	0.944
local	37	1.221	1.006	-0.167	-1.88	-2	0.961
local	118	1.324	1.071	0.64	-1.893	-2	0.94
local	50	1.061	0.858	-0.475	-1.859	-2	0.964
local	42	1.403	1.103	-0.169	-1.861	-2	0.953
local	38	1.429	1.026	0.178	-1.878	-2	0.946
local	30	1.443	1.132	-0.559	-1.87	-2	0.947
local	57	1.405	1.106	0.419	-1.813	-2	0.924
local	39	1.592	1.239	0.496	-1.84	-2	0.878
local	34	1.557	1.171	0.749	-1.829	-2	0.685
local	15	1.3	1.018	0.445	-1.836	-2	0.888
local	27	1.356	1.126	-0.115	-1.836	-2	0.954
local	26	1.487	1.2	0.744	-1.848	-2	0.946
local	17	1.401	1.112	0.459	-1.875	-2	0.952
local	11	1.344	1.084	-0.118	-1.869	-2	0.93

Table 29: Results from the BiLSTM model for 10cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.762	1.363	0.423	-2.168	-3	0.893
region	Østfold	1.781	1.432	-0.288	-2.108	-3	0.902
region	Vestfold	1.433	1.153	0.033	-2.169	-3	0.94
region	Trøndelag	1.758	1.329	1.116	-2.177	-3	0.812
region	Innlandet	2.073	1.572	0.925	-2.165	-3	0.873
local	52	2.346	1.985	-1.611	-2.156	-3	0.826
local	41	1.609	1.339	-0.577	-2.15	-3	0.926
local	37	1.547	1.281	0.131	-2.132	-3	0.935
local	118	1.48	1.119	0.921	-2.149	-3	0.917
local	50	1.092	0.88	-0.165	-2.167	-3	0.958
local	42	1.637	1.343	0.163	-2.173	-3	0.931
local	38	1.326	1.082	0.31	-2.173	-3	0.947
local	30	1.606	1.308	-0.176	-2.147	-3	0.929
local	57	1.766	1.334	1.113	-2.133	-3	0.873
local	39	1.678	1.253	1.006	-2.158	-3	0.859
local	34	1.959	1.538	1.48	-2.165	-3	0.413
local	15	1.624	1.203	0.887	-2.173	-3	0.812
local	27	1.772	1.363	0.597	-2.137	-3	0.915
local	26	2.318	1.798	1.491	-2.153	-3	0.859
local	17	2.374	1.797	1.304	-2.103	-3	0.853
local	11	1.903	1.419	0.428	-2.164	-3	0.838

Table 30: Results from the LSTM model for 20cm depth.

scope	spesific scope	RMSE °C	MAE°C	bias °C	log(κ(model))	digit sensitivity	R^2
global	—	1.871	1.472	0.302	-1.544	-2	0.886
region	Østfold	1.904	1.522	0.067	-1.571	-2	0.888
region	Vestfold	1.834	1.422	0.119	-1.565	-2	0.91
region	Trøndelag	1.859	1.423	0.675	-1.57	-2	0.803
region	Innlandet	1.893	1.534	0.328	-1.552	-2	0.903
local	52	2.254	1.725	-0.807	-1.544	-2	0.705
local	41	1.85	1.544	-0.372	-1.581	-2	0.911
local	37	1.773	1.495	0.023	-1.565	-2	0.917
local	118	1.872	1.412	1.047	-1.6	-2	0.879
local	50	1.388	1.106	-0.047	-1.593	-2	0.938
local	42	2.002	1.591	0.199	-1.517	-2	0.905
local	38	2.014	1.477	0.537	-1.559	-2	0.892
local	30	1.876	1.518	-0.182	-1.581	-2	0.911
local	57	1.906	1.463	0.549	-1.592	-2	0.86
local	39	1.855	1.397	0.616	-1.596	-2	0.834
local	34	1.923	1.466	0.877	-1.609	-2	0.52
local	15	1.754	1.371	0.669	-1.581	-2	0.797
local	27	1.894	1.561	0.043	-1.568	-2	0.911
local	26	1.964	1.585	0.869	-1.568	-2	0.905
local	17	1.809	1.433	0.464	-1.535	-2	0.92
local	11	1.854	1.503	-0.048	-1.553	-2	0.867

Table 31: Results from the LSTM model for 10cm depth.

	27		15		30		26	
2014	μ:10.988	max:33.0	μ:11.166	max:32.4	μ:11.486	max:31.8	μ:10.327	max:30.2
	σ:6.564	min:-4.2	σ:6.305	min:-9.4	σ:6.473	min:-3.3	σ:6.829	min:-5.1
2015	μ:9.416	max:27.1	μ:9.858	max:27.9	μ:10.151	max:25.7	μ:8.993	max:27.1
	σ:5.906	min:-6.8	σ:5.173	min:-2.9	σ:5.807	min:-4.9	σ:6.135	min:-9.3
2016	μ:10.104	max:27.8	μ:9.98	max:28.1	μ:12.434	max:30.4	μ:9.74	max:26.7
	σ:6.508	min:-6.6	σ:5.581	min:-4.2	σ:6.282	min:-4.4	σ:6.727	min:-7.2
2017	μ:9.322	max:29.7	μ:9.675	max:26.5	μ:10.693	max:26.7	μ:8.927	max:28.5
	σ:6.339	min:-10.8	σ:5.598	min:-6.1	σ:6.371	min:-8.4	σ:6.477	min:-11.8
2018	μ:9.694	max:31.1	μ:9.885	max:32.6	μ:11.321	max:32.5	μ:9.949	max:32.4
	σ:9.471	min:-23.2	σ:7.042	min:-15.4	σ:9.113	min:-16.8	σ:9.999	min:-26.3
2019	μ:9.556	max:31.1	μ:9.7	max:33.5	μ:10.596	max:32.0	μ:9.242	max:30.4
	σ:7.083	min:-15.9	σ:6.673	min:-10.0	σ:6.99	min:-11.9	σ:7.476	min:-17.4
2020	μ:10.328	max:30.3	μ:9.938	max:32.5	μ:11.019	max:29.6	μ:9.949	max:28.9
	σ:6.639	min:-13.6	σ:6.464	min:-4.5	σ:6.703	min:-7.2	σ:6.895	min:-8.5
2021	μ:10.459	max:29.3	μ:10.378	max:30.6	μ:11.417	max:30.5	μ:10.075	max:28.1
	σ:6.827	min:-6.1	σ:6.122	min:-4.8	σ:6.888	min:-7.0	σ:7.169	min:-12.0
2022	μ:9.869	max:27.8	μ:10.035	max:30.2	μ:11.135	max:27.2	μ:9.558	max:27.0
	σ:6.89	min:-11.6	σ:5.497	min:-4.8	σ:6.841	min:-7.6	σ:7.236	min:-12.8

Table 32: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

	27		15		30		26	
2014	μ:12.241	max:23.9	μ:9.378	max:17.2	μ:10.76	max:19.2	μ:10.523	max:21.7
	σ:5.92	min:0.9	σ:4.483	min:0.5	σ:4.841	min:1.6	σ:5.778	min:-0.2
2015	μ:10.498	max:21.5	μ:8.814	max:15.6	μ:9.371	max:16.3	μ:8.861	max:18.2
	σ:5.579	min:-0.2	σ:3.979	min:1.5	σ:4.52	min:0.5	σ:5.282	min:-0.5
2016	μ:10.576	max:20.3	μ:8.882	max:17.3	μ:13.312	max:23.9	μ:9.257	max:19.1
	σ:6.462	min:-2.7	σ:4.8	min:-0.2	σ:5.304	min:-3.3	σ:6.053	min:-2.3
2017	μ:9.842	max:18.8	μ:8.926	max:16.3	μ:11.175	max:23.0	μ:8.904	max:18.5
	σ:6.066	min:-1.6	σ:4.519	min:-0.1	σ:5.969	min:-0.2	σ:6.113	min:-2.7
2018	μ:10.03	max:22.0	μ:8.819	max:18.5	μ:11.53	max:25.6	μ:9.852	max:22.6
	σ:6.856	min:-0.2	σ:5.224	min:-0.1	σ:6.856	min:0.4	σ:7.191	min:-1.4
2019	μ:10.502	max:23.1	μ:9.239	max:20.2	μ:11.156	max:27.6	μ:10.168	max:21.4
	σ:6.145	min:0.2	σ:4.968	min:0.2	σ:6.38	min:0.4	σ:5.659	min:0.0
2020	μ:10.567	max:21.9	μ:9.292	max:19.6	μ:11.59	max:26.1	μ:9.519	max:22.5
	σ:6.098	min:-0.9	σ:5.106	min:0.2	σ:6.166	min:0.1	σ:6.256	min:-2.5
2021	μ:10.417	max:21.4	μ:9.904	max:19.1	μ:11.455	max:23.6	μ:9.592	max:20.9
	σ:6.443	min:-1.1	σ:5.095	min:0.4	σ:6.647	min:-0.9	σ:6.593	min:-3.0
2022	μ:10.305	max:20.6	μ:9.424	max:17.1	μ:11.188	max:22.6	μ:9.098	max:20.4
	σ:6.311	min:-1.8	σ:4.408	min:0.4	σ:5.964	min:0.1	σ:6.181	min:-2.4

Table 33: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

	27		15		30		26	
2014	μ:11.865 σ:5.593	max:21.4 min:0.9	μ:9.213 σ:4.348	max:16.0 min:0.5	μ:11.023 σ:4.889	max:18.9 min:1.6	μ:10.488 σ:5.59	max:19.7 min:0.1
2015	μ:10.197 σ:5.33	max:19.1 min:-0.1	μ:8.708 σ:3.876	max:15.0 min:1.7	μ:9.661 σ:4.49	max:19.5 min:1.7	μ:9.01 σ:5.332	max:17.5 min:-0.2
2016	μ:10.205 σ:6.219	max:18.5 min:-2.2	μ:8.748 σ:4.634	max:16.1 min:0.0	μ:13.167 σ:4.825	max:21.1 min:-3.3	μ:9.588 σ:6.094	max:18.4 min:-1.6
2017	μ:9.489 σ:5.871	max:17.1 min:-1.3	μ:8.799 σ:4.382	max:15.4 min:0.2	μ:10.919 σ:5.686	max:19.8 min:-0.1	μ:9.095 σ:6.179	max:17.7 min:-1.8
2018	μ:9.664 σ:6.439	max:19.9 min:-0.1	μ:8.699 σ:5.032	max:17.1 min:0.1	μ:11.36 σ:6.404	max:22.3 min:0.6	μ:10.143 σ:7.061	max:20.6 min:-0.8
2019	μ:10.186 σ:5.838	max:20.7 min:0.3	μ:9.122 σ:4.749	max:18.7 min:0.5	μ:11.026 σ:5.874	max:23.5 min:0.7	μ:9.872 σ:5.894	max:19.9 min:0.3
2020	μ:10.21 σ:5.826	max:19.7 min:-0.5	μ:9.174 σ:4.899	max:17.9 min:0.5	μ:11.365 σ:5.725	max:22.5 min:0.4	μ:9.235 σ:6.061	max:20.1 min:-2.2
2021	μ:10.032 σ:6.194	max:19.3 min:-0.9	μ:9.778 σ:4.934	max:17.8 min:0.6	μ:11.118 σ:6.41	max:21.3 min:-0.3	μ:9.346 σ:6.37	max:19.3 min:-2.0
2022	μ:9.988 σ:6.024	max:18.5 min:-0.8	μ:9.34 σ:4.288	max:16.1 min:0.7	μ:10.98 σ:5.679	max:19.8 min:0.4	μ:8.855 σ:6.003	max:18.5 min:-1.7

Table 34: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

	17		41		11		50	
2014	μ:nan σ:nan	max:nan min:nan	μ:12.203 σ:6.63	max:33.0 min:-4.5	μ:10.66 σ:6.472	max:30.0 min:-4.5	μ:11.866 σ:5.688	max:29.4 min:-1.9
2015	μ:nan σ:nan	max:nan min:nan	μ:10.608 σ:5.676	max:26.4 min:-5.1	μ:9.121 σ:5.764	max:28.6 min:-7.1	μ:10.583 σ:5.236	max:24.9 min:-3.3
2016	μ:11.336 σ:6.029	max:25.8 min:-5.5	μ:11.218 σ:6.406	max:28.8 min:-5.5	μ:9.842 σ:6.49	max:26.4 min:-7.4	μ:11.349 σ:6.023	max:26.5 min:-4.9
2017	μ:8.24 σ:6.655	max:25.2 min:-16.7	μ:10.708 σ:6.077	max:27.3 min:-6.4	μ:9.049 σ:6.171	max:28.5 min:-8.7	μ:11.009 σ:5.673	max:24.7 min:-9.5
2018	μ:9.073 σ:9.791	max:29.8 min:-25.9	μ:11.554 σ:8.628	max:33.7 min:-19.2	μ:10.166 σ:9.322	max:30.2 min:-21.4	μ:11.734 σ:7.937	max:31.3 min:-14.0
2019	μ:8.751 σ:7.464	max:29.9 min:-20.4	μ:10.843 σ:6.679	max:30.4 min:-12.8	μ:9.275 σ:7.081	max:29.4 min:-13.2	μ:11.728 σ:6.324	max:29.5 min:-7.3
2020	μ:9.642 σ:7.221	max:29.4 min:-18.3	μ:11.323 σ:6.324	max:30.2 min:-6.6	μ:9.918 σ:6.662	max:28.5 min:-9.0	μ:11.443 σ:5.966	max:28.6 min:-5.3
2021	μ:nan σ:nan	max:nan min:nan	μ:11.519 σ:6.553	max:29.8 min:-7.1	μ:10.181 σ:6.762	max:27.7 min:-8.8	μ:11.77 σ:6.102	max:26.9 min:-4.8
2022	μ:9.07 σ:7.161	max:27.7 min:-15.0	μ:11.192 σ:6.625	max:28.5 min:-7.5	μ:9.644 σ:6.652	max:26.8 min:-10.3	μ:11.504 σ:6.004	max:27.7 min:-5.0

Table 35: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

	17	41	11	50
2014	μ:nan max:nan σ:nan min:nan	μ:13.38 max:24.8 σ:5.721 min:2.4	μ:11.146 max:23.3 σ:5.983 min:0.1	μ:11.194 max:21.6 σ:5.036 min:1.5
2015	μ:nan max:nan σ:nan min:nan	μ:11.099 max:20.6 σ:4.949 min:0.8	μ:8.035 max:19.7 σ:5.148 min:-0.2	μ:10.266 max:19.5 σ:4.814 min:0.2
2016	μ:12.884 max:20.7 σ:4.552 min:2.5	μ:11.169 max:21.9 σ:5.782 min:-0.1	μ:10.028 max:20.0 σ:6.286 min:-1.3	μ:11.612 max:20.7 σ:5.562 min:0.0
2017	μ:8.66 max:19.7 σ:6.312 min:-1.9	μ:11.002 max:18.5 σ:4.957 min:0.3	μ:9.306 max:18.3 σ:6.13 min:-1.4	μ:10.901 max:19.0 σ:4.833 min:0.3
2018	μ:9.555 max:22.8 σ:6.985 min:-0.5	μ:12.439 max:25.9 σ:7.815 min:-0.8	μ:10.312 max:21.9 σ:6.95 min:0.0	μ:10.924 max:21.9 σ:6.257 min:-0.3
2019	μ:9.094 max:21.4 σ:6.317 min:-0.5	μ:10.491 max:20.3 σ:5.405 min:0.2	μ:9.727 max:22.0 σ:6.264 min:-0.1	μ:10.783 max:21.0 σ:5.084 min:0.3
2020	μ:9.709 max:20.8 σ:6.534 min:-1.2	μ:12.209 max:23.9 σ:6.329 min:0.6	μ:9.927 max:22.4 σ:6.387 min:-1.6	μ:11.329 max:20.2 σ:4.97 min:1.3
2021	μ:nan max:nan σ:nan min:nan	μ:11.891 max:23.5 σ:6.341 min:0.0	μ:10.106 max:20.4 σ:6.248 min:-0.3	μ:11.513 max:21.0 σ:5.808 min:0.0
2022	μ:9.327 max:19.9 σ:6.397 min:-2.4	μ:11.318 max:21.8 σ:6.12 min:-0.2	μ:13.182 max:21.3 σ:3.953 min:3.9	μ:11.466 max:20.3 σ:5.388 min:0.7

Table 36: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

	17	41	11	50
2014	μ:nan max:nan σ:nan min:nan	μ:12.91 max:22.2 σ:5.34 min:2.6	μ:10.922 max:21.6 σ:5.761 min:0.3	μ:10.507 max:18.8 σ:4.766 min:1.8
2015	μ:nan max:nan σ:nan min:nan	μ:10.826 max:18.2 σ:4.705 min:1.4	μ:7.92 max:17.9 σ:4.912 min:0.0	μ:9.871 max:17.9 σ:4.792 min:0.4
2016	μ:12.627 max:18.0 σ:4.0 min:3.6	μ:10.834 max:19.7 σ:5.511 min:0.0	μ:9.842 max:18.5 σ:6.064 min:-0.7	μ:11.44 max:19.2 σ:5.37 min:0.3
2017	μ:8.311 max:17.2 σ:6.081 min:-1.4	μ:10.626 max:16.4 σ:4.751 min:0.4	μ:9.102 max:16.8 σ:5.993 min:-1.1	μ:10.237 max:17.5 σ:4.668 min:0.5
2018	μ:9.221 max:19.9 σ:6.456 min:-0.3	μ:11.758 max:22.2 σ:7.332 min:-0.3	μ:10.149 max:20.3 σ:6.539 min:0.2	μ:10.593 max:19.3 σ:5.956 min:0.0
2019	μ:8.893 max:19.4 σ:5.966 min:-0.3	μ:10.366 max:18.5 σ:5.125 min:0.7	μ:9.672 max:20.4 σ:5.961 min:0.3	μ:10.568 max:19.3 σ:4.903 min:0.5
2020	μ:9.347 max:18.3 σ:6.146 min:-0.4	μ:11.804 max:21.2 σ:6.003 min:1.0	μ:9.817 max:20.5 σ:6.134 min:-0.7	μ:11.046 max:18.6 σ:4.776 min:1.5
2021	μ:nan max:nan σ:nan min:nan	μ:11.7 max:21.4 σ:6.083 min:0.2	μ:10.014 max:19.1 σ:5.995 min:0.1	μ:11.052 max:19.1 σ:5.604 min:0.0
2022	μ:9.095 max:18.0 σ:6.196 min:-1.1	μ:11.15 max:19.5 σ:5.83 min:0.1	μ:12.987 max:19.2 σ:3.605 min:5.1	μ:11.001 max:18.5 σ:5.144 min:0.8

Table 37: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

	42		57		37		38	
2014	μ:12.0 σ:6.459	max:32.2 min:-3.1	μ:10.922 σ:6.59	max:31.2 min:-11.9	μ:10.8 σ:6.437	max:30.7 min:-4.6	μ:11.711 σ:6.4	max:31.8 min:-4.0
2015	μ:10.417 σ:5.871	max:25.4 min:-4.8	μ:9.567 σ:5.366	max:27.5 min:-6.7	μ:9.361 σ:5.738	max:26.2 min:-6.0	μ:10.385 σ:5.693	max:25.0 min:-5.7
2016	μ:11.037 σ:6.476	max:29.4 min:-4.2	μ:9.54 σ:6.016	max:28.6 min:-8.1	μ:10.161 σ:6.444	max:26.9 min:-9.8	μ:11.343 σ:5.919	max:27.4 min:-5.2
2017	μ:10.558 σ:6.214	max:26.8 min:-8.5	μ:9.412 σ:5.896	max:27.1 min:-8.0	μ:9.48 σ:6.112	max:24.6 min:-7.0	μ:10.346 σ:6.095	max:25.2 min:-9.2
2018	μ:11.399 σ:8.932	max:33.0 min:-14.5	μ:9.588 σ:7.611	max:32.0 min:-21.4	μ:10.513 σ:8.762	max:31.3 min:-20.1	μ:11.026 σ:8.809	max:32.3 min:-19.7
2019	μ:10.47 σ:6.822	max:30.9 min:-9.8	μ:9.306 σ:6.983	max:32.4 min:-13.0	μ:9.97 σ:7.077	max:31.6 min:-16.6	μ:10.944 σ:6.832	max:31.6 min:-12.0
2020	μ:11.027 σ:6.673	max:29.4 min:-7.5	μ:9.625 σ:6.619	max:32.8 min:-6.2	μ:10.403 σ:6.697	max:29.7 min:-7.4	μ:11.281 σ:7.453	max:29.7 min:-6.8
2021	μ:11.351 σ:6.746	max:29.2 min:-5.8	μ:9.889 σ:6.375	max:29.9 min:-8.3	μ:10.51 σ:6.807	max:28.2 min:-8.1	μ:11.722 σ:6.793	max:30.3 min:-6.1
2022	μ:11.094 σ:6.682	max:28.6 min:-6.3	μ:9.722 σ:5.791	max:28.4 min:-8.2	μ:10.162 σ:6.92	max:30.5 min:-9.7	μ:11.327 σ:6.786	max:29.1 min:-6.9

Table 38: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

	42		57		37		38	
2014	μ:12.296 σ:4.887	max:21.8 min:2.7	μ:10.603 σ:5.696	max:22.2 min:0.0	μ:11.237 σ:5.153	max:21.6 min:2.0	μ:12.793 σ:5.559	max:22.8 min:1.5
2015	μ:10.957 σ:5.29	max:20.3 min:0.1	μ:9.554 σ:4.382	max:16.7 min:0.1	μ:9.761 σ:4.506	max:17.8 min:0.5	μ:10.99 σ:5.321	max:18.7 min:0.4
2016	μ:11.293 σ:6.044	max:20.7 min:-0.1	μ:9.152 σ:5.249	max:17.4 min:-0.2	μ:10.593 σ:5.565	max:19.8 min:-0.4	μ:11.667 σ:6.074	max:19.9 min:0.1
2017	μ:10.46 σ:5.741	max:19.0 min:-0.3	μ:9.038 σ:4.871	max:16.3 min:-0.1	μ:10.147 σ:5.147	max:18.5 min:-0.2	μ:10.545 σ:5.467	max:17.6 min:0.1
2018	μ:11.048 σ:6.59	max:22.0 min:0.0	μ:8.456 σ:5.599	max:17.1 min:-1.3	μ:11.104 σ:6.219	max:21.2 min:-1.5	μ:11.123 σ:6.316	max:21.2 min:0.5
2019	μ:10.672 σ:6.061	max:21.4 min:0.1	μ:9.204 σ:5.274	max:19.4 min:0.1	μ:10.772 σ:5.612	max:21.4 min:-0.2	μ:10.961 σ:5.668	max:21.0 min:0.4
2020	μ:13.184 σ:4.763	max:22.8 min:1.6	μ:9.551 σ:5.288	max:18.0 min:0.1	μ:10.809 σ:5.292	max:21.0 min:0.0	μ:9.965 σ:5.872	max:19.5 min:0.8
2021	μ:10.999 σ:6.997	max:23.2 min:-0.4	μ:9.214 σ:5.324	max:17.8 min:-0.3	μ:10.592 σ:6.4	max:22.0 min:-0.5	μ:11.013 σ:6.182	max:19.5 min:0.1
2022	μ:10.835 σ:5.968	max:20.6 min:-0.1	μ:9.711 σ:4.965	max:18.1 min:0.3	μ:10.336 σ:5.99	max:19.9 min:-0.3	μ:10.615 σ:6.127	max:18.1 min:0.3

Table 39: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

Norwegian University of Life Sciences

	42		57		37		38	
2014	μ:12.967 σ:4.05	max:20.6 min:4.6	μ:10.149 σ:5.364	max:19.1 min:-0.1	μ:11.153 σ:4.936	max:19.7 min:2.3	μ:12.524 σ:5.424	max:21.7 min:1.7
2015	μ:10.765 σ:5.038	max:18.6 min:0.4	μ:9.335 σ:4.32	max:15.5 min:0.2	μ:9.719 σ:4.527	max:16.6 min:0.6	μ:10.849 σ:5.136	max:17.9 min:0.6
2016	μ:11.107 σ:5.776	max:19.3 min:0.2	μ:8.946 σ:5.12	max:16.2 min:0.0	μ:10.385 σ:5.425	max:18.3 min:-0.4	μ:11.562 σ:5.923	max:19.4 min:0.3
2017	μ:10.317 σ:5.501	max:17.3 min:0.1	μ:8.805 σ:4.74	max:15.2 min:0.1	μ:10.118 σ:5.043	max:16.9 min:0.0	μ:10.371 σ:5.432	max:16.9 min:0.3
2018	μ:10.863 σ:6.187	max:19.9 min:0.3	μ:8.108 σ:5.415	max:15.6 min:-0.9	μ:10.889 σ:6.133	max:20.0 min:-0.7	μ:11.093 σ:6.146	max:20.5 min:0.6
2019	μ:10.517 σ:5.775	max:19.9 min:0.5	μ:9.017 σ:5.122	max:18.3 min:0.2	μ:10.458 σ:5.432	max:19.3 min:-0.3	μ:10.873 σ:5.554	max:20.4 min:0.8
2020	μ:12.791 σ:4.488	max:20.1 min:2.5	μ:9.368 σ:5.173	max:16.8 min:0.1	μ:10.827 σ:5.16	max:19.4 min:0.5	μ:9.764 σ:5.746	max:18.7 min:0.8
2021	μ:10.749 σ:6.743	max:20.6 min:-0.1	μ:8.963 σ:5.177	max:16.6 min:-0.3	μ:10.524 σ:6.311	max:20.5 min:-0.3	μ:10.773 σ:6.007	max:19.2 min:0.3
2022	μ:10.601 σ:5.692	max:18.5 min:0.3	μ:9.478 σ:4.844	max:16.8 min:0.4	μ:10.379 σ:5.867	max:18.7 min:0.0	μ:10.387 σ:5.6	max:17.6 min:0.5

Table 40: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

	52		118		34		39	
2014	μ:11.961 σ:6.684	max:31.9 min:-4.0	μ:11.792 σ:6.384	max:31.5 min:-3.5	μ:10.545 σ:6.889	max:31.4 min:-14.5	μ:10.76 σ:6.24	max:30.7 min:-8.0
2015	μ:10.447 σ:5.932	max:25.8 min:-6.4	μ:10.283 σ:5.477	max:24.6 min:-4.9	μ:9.209 σ:5.42	max:27.4 min:-5.9	μ:9.524 σ:5.072	max:26.2 min:-5.8
2016	μ:11.096 σ:6.457	max:28.4 min:-5.7	μ:10.993 σ:6.22	max:27.3 min:-5.8	μ:9.341 σ:6.1	max:29.4 min:-8.4	μ:9.512 σ:5.45	max:29.0 min:-6.4
2017	μ:10.373 σ:6.159	max:25.6 min:-6.7	μ:10.425 σ:5.9	max:25.2 min:-6.4	μ:8.883 σ:5.987	max:25.6 min:-9.5	μ:9.479 σ:5.693	max:26.2 min:-6.4
2018	μ:11.451 σ:8.545	max:32.6 min:-18.0	μ:11.338 σ:8.456	max:33.1 min:-16.6	μ:9.291 σ:7.729	max:31.6 min:-19.5	μ:9.246 σ:6.877	max:31.4 min:-15.1
2019	μ:10.705 σ:6.754	max:30.6 min:-13.2	μ:11.057 σ:6.716	max:31.5 min:-11.3	μ:8.945 σ:7.017	max:32.7 min:-14.3	μ:9.193 σ:6.446	max:30.8 min:-10.9
2020	μ:11.072 σ:6.496	max:28.9 min:-7.1	μ:11.439 σ:6.453	max:30.2 min:-6.6	μ:9.051 σ:6.998	max:31.3 min:-7.5	μ:9.44 σ:6.225	max:29.8 min:-5.3
2021	μ:11.059 σ:6.618	max:27.7 min:-7.6	μ:11.586 σ:6.595	max:29.4 min:-5.5	μ:9.72 σ:6.481	max:31.7 min:-6.5	μ:9.915 σ:5.8	max:30.0 min:-3.9
2022	μ:10.878 σ:6.85	max:28.0 min:-8.7	μ:11.405 σ:6.714	max:29.4 min:-7.3	μ:8.985 σ:5.867	max:29.2 min:-7.9	μ:9.505 σ:5.247	max:27.5 min:-6.0

Table 41: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

	52		118		34		39	
2014	μ:12.462	max:23.2	μ:11.901	max:21.8	μ:9.229	max:18.4	μ:10.05	max:19.6
	σ:5.157	min:2.6	σ:5.262	min:1.9	σ:5.267	min:-0.1	σ:5.34	min:0.0
2015	μ:11.307	max:19.4	μ:10.506	max:19.5	μ:8.779	max:15.4	μ:9.121	max:15.7
	σ:4.865	min:1.1	σ:5.022	min:0.5	σ:4.095	min:0.6	σ:4.107	min:1.2
2016	μ:11.951	max:20.2	μ:10.866	max:19.6	μ:8.687	max:17.0	μ:9.108	max:16.9
	σ:5.67	min:0.1	σ:5.599	min:-0.4	σ:4.853	min:-0.1	σ:4.865	min:0.1
2017	μ:10.887	max:19.7	μ:10.099	max:16.9	μ:8.706	max:17.1	μ:8.998	max:16.9
	σ:5.337	min:0.1	σ:4.974	min:-0.2	σ:4.776	min:0.1	σ:4.753	min:0.3
2018	μ:11.39	max:21.2	μ:10.154	max:19.4	μ:8.444	max:17.6	μ:9.08	max:19.2
	σ:6.804	min:-1.0	σ:6.354	min:-0.9	σ:5.487	min:-1.3	σ:5.861	min:-0.9
2019	μ:11.665	max:21.7	μ:10.666	max:20.9	μ:9.078	max:19.3	μ:9.394	max:19.4
	σ:5.495	min:0.3	σ:5.485	min:0.1	σ:5.111	min:-0.2	σ:4.909	min:0.5
2020	μ:12.331	max:21.7	μ:10.823	max:19.6	μ:8.534	max:17.0	μ:9.111	max:17.6
	σ:4.968	min:0.8	σ:5.071	min:0.4	σ:4.839	min:0.3	σ:4.848	min:1.0
2021	μ:6.954	max:22.1	μ:10.277	max:19.6	μ:8.944	max:16.7	μ:9.436	max:16.9
	σ:5.746	min:0.0	σ:5.911	min:-0.7	σ:4.722	min:0.1	σ:4.789	min:0.5
2022	μ:13.996	max:19.4	μ:10.22	max:17.4	μ:8.694	max:16.4	μ:9.271	max:15.9
	σ:3.069	min:7.4	σ:5.147	min:0.2	σ:4.759	min:0.6	σ:4.345	min:1.0

Table 42: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

	52		118		34		39	
2014	μ:12.146	max:20.9	μ:11.672	max:20.3	μ:9.039	max:16.7	μ:9.896	max:18.4
	σ:4.928	min:3.3	σ:5.003	min:2.4	σ:5.106	min:-0.1	σ:5.215	min:0.0
2015	μ:11.079	max:17.8	μ:10.395	max:18.0	μ:8.689	max:14.6	μ:9.021	max:15.3
	σ:4.704	min:1.6	σ:4.73	min:0.9	σ:3.982	min:0.8	σ:4.022	min:1.4
2016	μ:11.672	max:18.7	μ:10.772	max:18.3	μ:8.616	max:15.7	μ:9.005	max:16.1
	σ:5.442	min:0.4	σ:5.258	min:-0.1	σ:4.685	min:0.0	σ:4.757	min:0.1
2017	μ:10.64	max:18.0	μ:9.993	max:16.0	μ:8.627	max:15.7	μ:8.923	max:16.2
	σ:5.151	min:0.2	σ:4.768	min:0.2	σ:4.606	min:0.4	σ:4.638	min:0.5
2018	μ:11.079	max:20.0	μ:9.931	max:18.3	μ:8.307	max:16.0	μ:8.928	max:18.2
	σ:6.518	min:-0.6	σ:6.071	min:-0.6	σ:5.27	min:-0.2	σ:5.749	min:-0.3
2019	μ:11.375	max:20.4	μ:10.45	max:19.4	μ:9.027	max:17.9	μ:9.308	max:18.5
	σ:5.302	min:0.5	σ:5.272	min:0.3	σ:4.907	min:0.2	σ:4.771	min:0.7
2020	μ:12.258	max:20.4	μ:10.61	max:18.3	μ:8.451	max:15.6	μ:9.02	max:16.3
	σ:4.773	min:1.2	σ:4.837	min:0.8	σ:4.62	min:0.6	σ:4.729	min:1.1
2021	μ:12.573	max:24.2	μ:9.995	max:18.1	μ:8.789	max:15.7	μ:9.346	max:16.4
	σ:6.025	min:0.0	σ:5.67	min:-0.7	σ:4.563	min:0.2	σ:4.714	min:0.6
2022	μ:12.151	max:22.4	μ:9.926	max:16.5	μ:8.61	max:15.2	μ:9.154	max:15.5
	σ:5.246	min:0.9	σ:4.949	min:0.5	σ:4.622	min:0.8	σ:4.267	min:1.1

Table 43: Every cell of the table has the mean temperature, the standard deviation, and the minimum and maximum temperature on one station in one year

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway