Assimilation of train based ZTD observations in AROME-RUC

Florian Weidle, Florian Meier, Christoph Wittmann

Outline

- Introduction to the project **TRain**
  - Project triggered by PhD of Matthias Aichinger-Rosenberger at TU Vienna
  - One conclusion of PhD: Retrieval of ZTD observations based on GNSS receiver on moving trains works reasonably well
  - Start of project in September 2020

- ZTD observations from moving trains
  - Data set for a 4 weeks period in autumn 2021 was provided in summer 2022
  - How do data look like? How many trains per day, how is the coverage?
  - Usage of train data in AROME-RUC
  - Analysis of data/fg-departures to apply bias correction

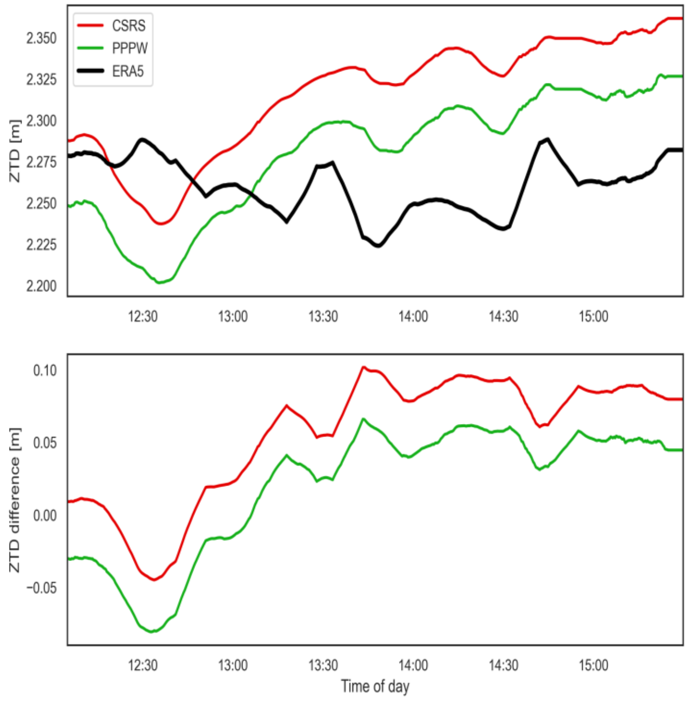
- Case study (on-going work)
  - Impact of train based observations on AROME-RUC forecasts
  - Evaluation of different setups

- Conclusions/Outlook

Motivation – Project history

- PhD at TU Vienna from Matthias Aichinger-Rosenberger dealt with the retrieval of ZTD data from GNSS-receivers on trains
- **Major challenges** have been an appropriate pre-processing and outlier detection of the data, and since only single-frequency observations are available a special treatment of the ionosphere has to be implemented. (PhD-thesis available online: *https://repositum.tuwien.at/bitstream/20.500.12708/17044/1/Aichinger-Rosenberger%20Matthias%20-%202021%20-%20Tropospheric%20parameter%20estimation%20based...pdf*)

Bad case

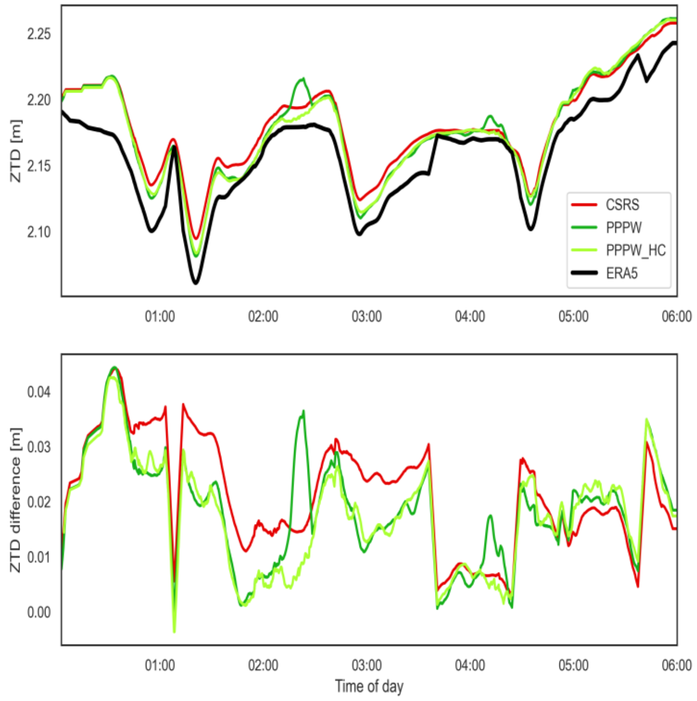


Courtesy of Matthias Aichinger-Rosenberger

3 out of 4 investigated cases showed high correlation with ERA5-ZTD data

In one case the observations were uncorrelated with ERA5 for parts of the track.
Could also be caused by errorenous ERA5 data….

Example of good cases



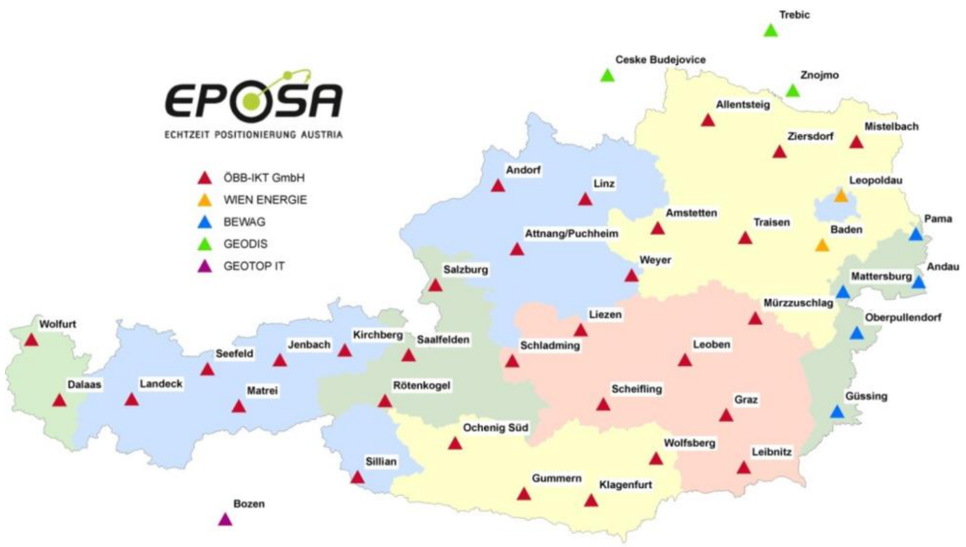
Courtesy of Matthias Aichinger-Rosenberger

GeoSphere Austria

Motivation – Project history

- Approx. 1500 Trains (freight and passenger trains) of federal railway company equipped with GNSS receivers and the number will even increase in the next years
- ZTD profiles with high resolution in space and time can be generated -> **Potential to increase the number of ZTD observations**

- Project started September 2020 and was prolongued to August 2023
- Delay in delivery of train based observations, data for one month period from Sep/Oct 2021 were provided in summer 2022

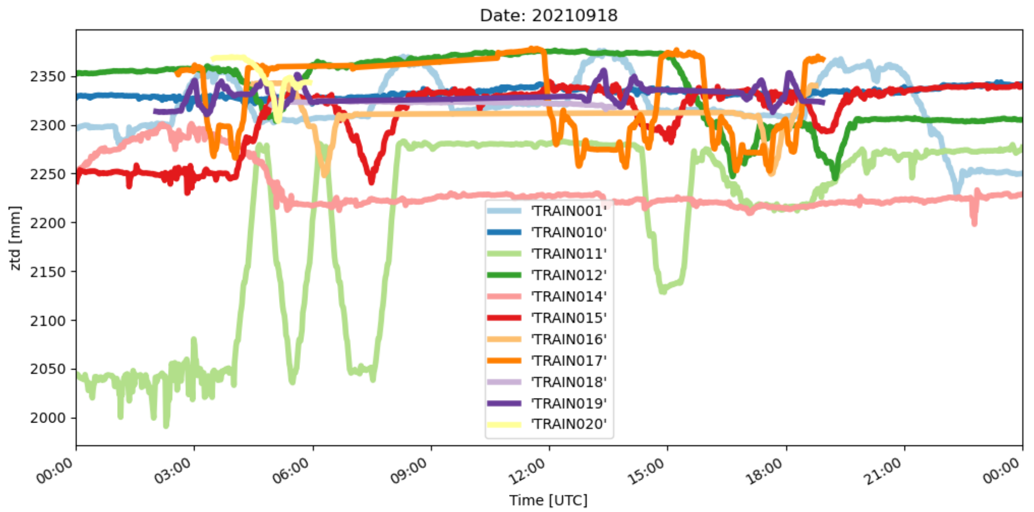


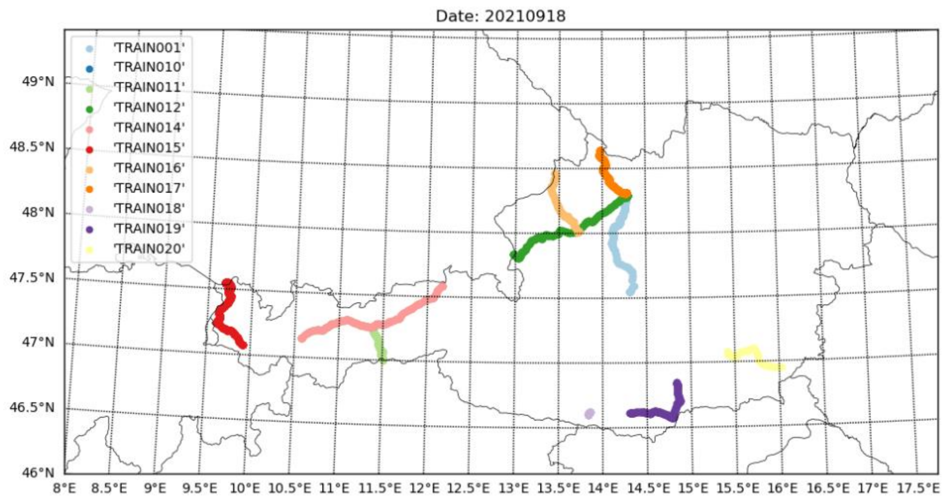
Source: OEBB 2017

Observation data set

- Data period from 9.9.2021 to 16.10.2021 with approx. 10 trains per day
- Raw data with 1 second time resolution, but processed data are averaged to 1 minute to reduce noise in the observations
- Trains are identified by a unique ID, trains are operating in the same area throughout the period but not necessarily every day on the exact same track
- Trains are not operating for the whole day, some have breaks or operate only for a limited time (train020)
- Some trains are circulating between two cities with remarkable height differences (train011)

- Question: How do bias behave if train operates in highly orographic terrain? How to apply bias correction?

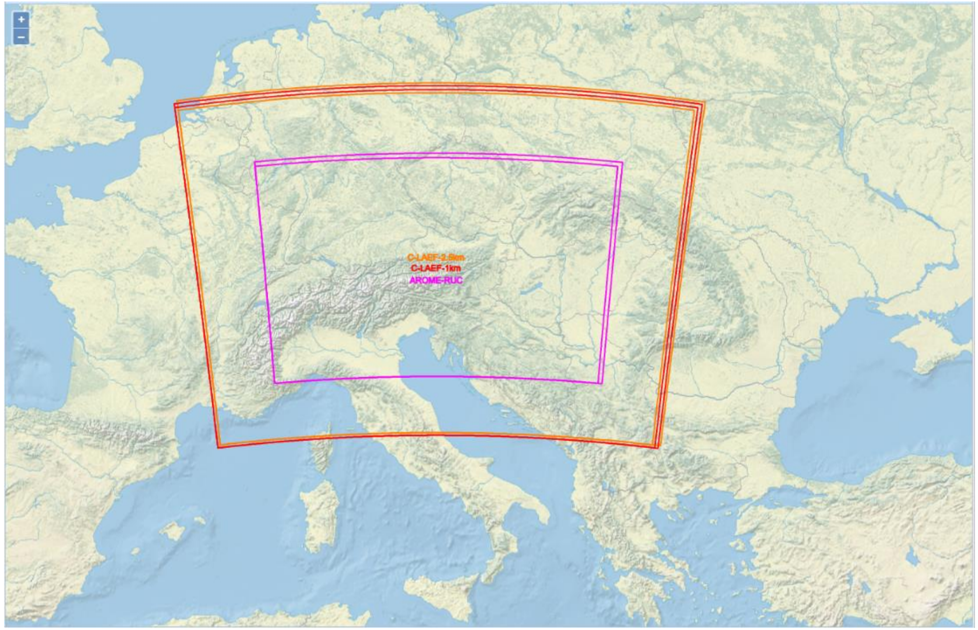




GeoSphere Austria

AROME-RUC setup and reference run

- Operational AROME-RUC setup with some limitations regarding availability of observations
  -> Radar, GPSRO, MODE-S MRAR, snow model are dismissed
- Station based GNSS observations with static bias correction
- Period 9.9. – 19.9.2021 serves as training period, for spin up and to calculate bias correction factors
- 20.9.2021 – 16.10.2021 is evaluation period

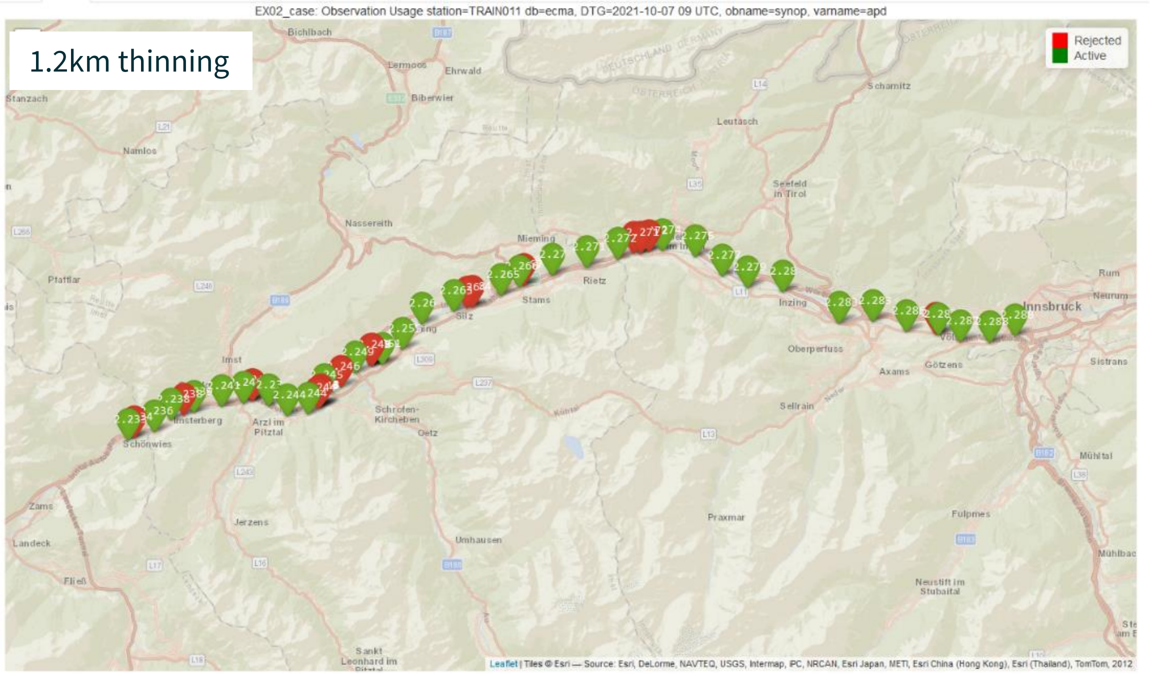


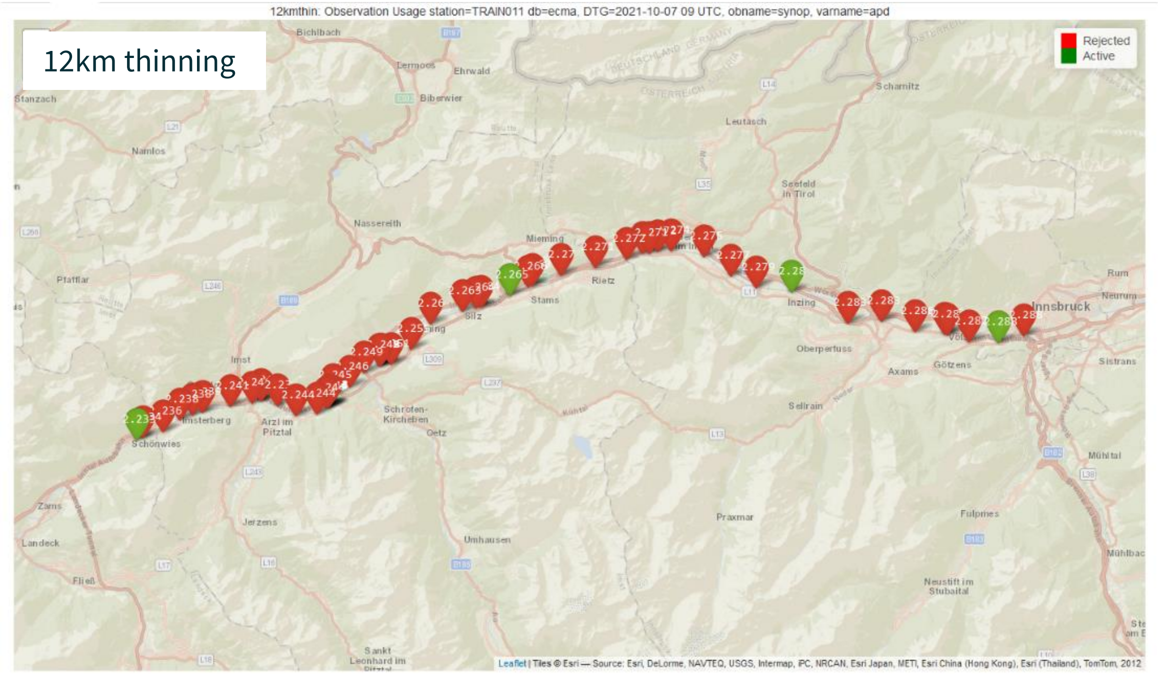
	AROME-RUC
Model version	cy43t2bf11
Resolution	1.2km
Area	Austrian area (900x576)
Members	1
Levels (lowest/highest)	90 (5m / 35km)
Starting times	00, 01,…,22,23 UTC
Forecast range	12 hours(05UTC+25h)
Time step	30s
Output Frequency	15min 2D/1h 3D
Orography / physiography	SRTM 90m ECOCLIMAP 1
LBC model	AROME-Aut
LBC update	1h
Surface scheme	SURFEX 8.0
Initial conditions (3D / Surf.)	3DVAR / OI +IAU+Nudging/LHN
Cycle interval	1 hour
Assimilation Window	-90min-+30min
B-Matrix	AROME-RUC EDA climatologic
Hardware	HPE Apollo 8600 (ZAMG)

Observation Type	Parameter assimilated	AROME-RUC	Experimental mode
SYNOP/TAWES	U10m, V10m, Z, T2m,RH2m	x	x
AMDAR, MODE-S (Flugzeuge)	U,V,T,Q	x	x
MODE-S MRAR (Flugzeuge)	U,V,T	x	
GEOWIND (SAT-Winde) MSG3	U,V	x	x
GEOWIND-HR (SAT-Winde) MSG3	U,V	x	x
TEMP (Radiosonde)	U,V,T,Q,Z	x (bufr)	x
PILOT	U,V	x	x
WINDPROFILER, SODAR	U,V	x	x
SCADA	U,V,T	x	x
MSG3-SEVIRI	WV-radiances	x	x
NOAA16/18/19/20/SNPP/MetOp-B-C AMSU-A, MHS, ATMS	radiances	x	x
MetOp-B-C IASI	radiances	x	x
ASCAT wind	U10m,V10m	x (12,5km)	x
RADAR	reflectivity / radial winds	x	
INCA-RR	RR 5min latent heat nudging	x	x
SNOWGRID Schneemodell	snowheight/density	x	
GNSS	ZTD (optional STD)	x	x (static bias corr.)
Radio Occultation	GPSRO bending angles	x	
ZTD on trains	ZTD		x
T-LAKE	T Water	x	

Include train observations in 3D-Var

- Necessary code adaptions to allow usage of ZTD observations from moving trains (pretty much hard coded at the moment):
  - Call spatial thinning method for ZTD obs from trains (reuse thinning code for ships); thinning threshold hard-coded
  - Adaptions in filter_gpssol/write_obsoul_gpssol to handle multiple entries with same statid: apply static bias correction, skip location check
  - Blacklisting in case trains should be used only in passive mode



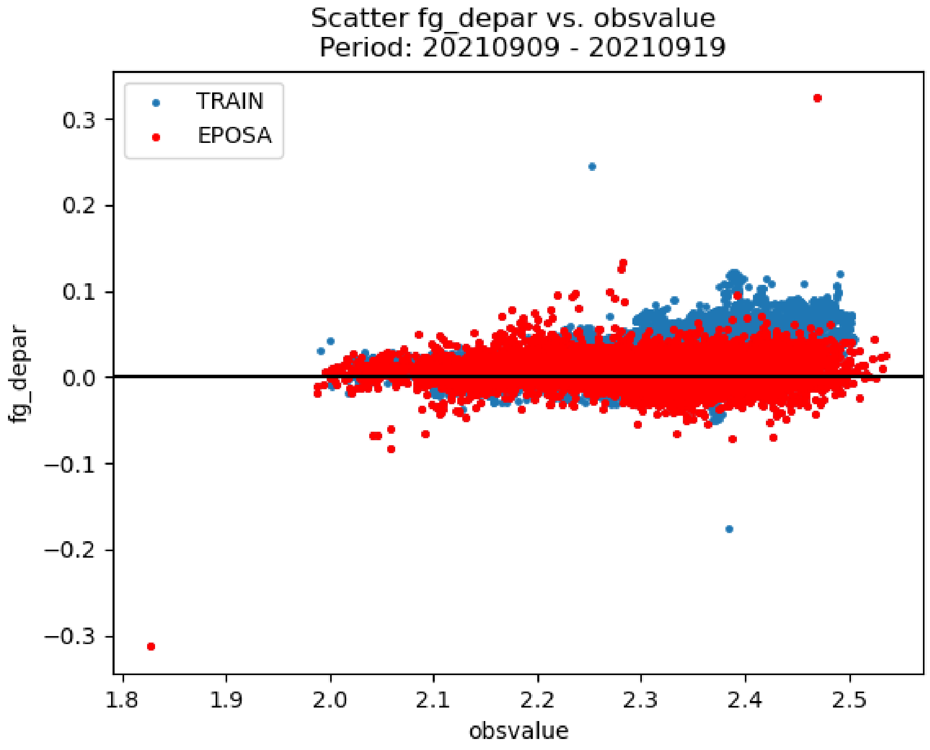


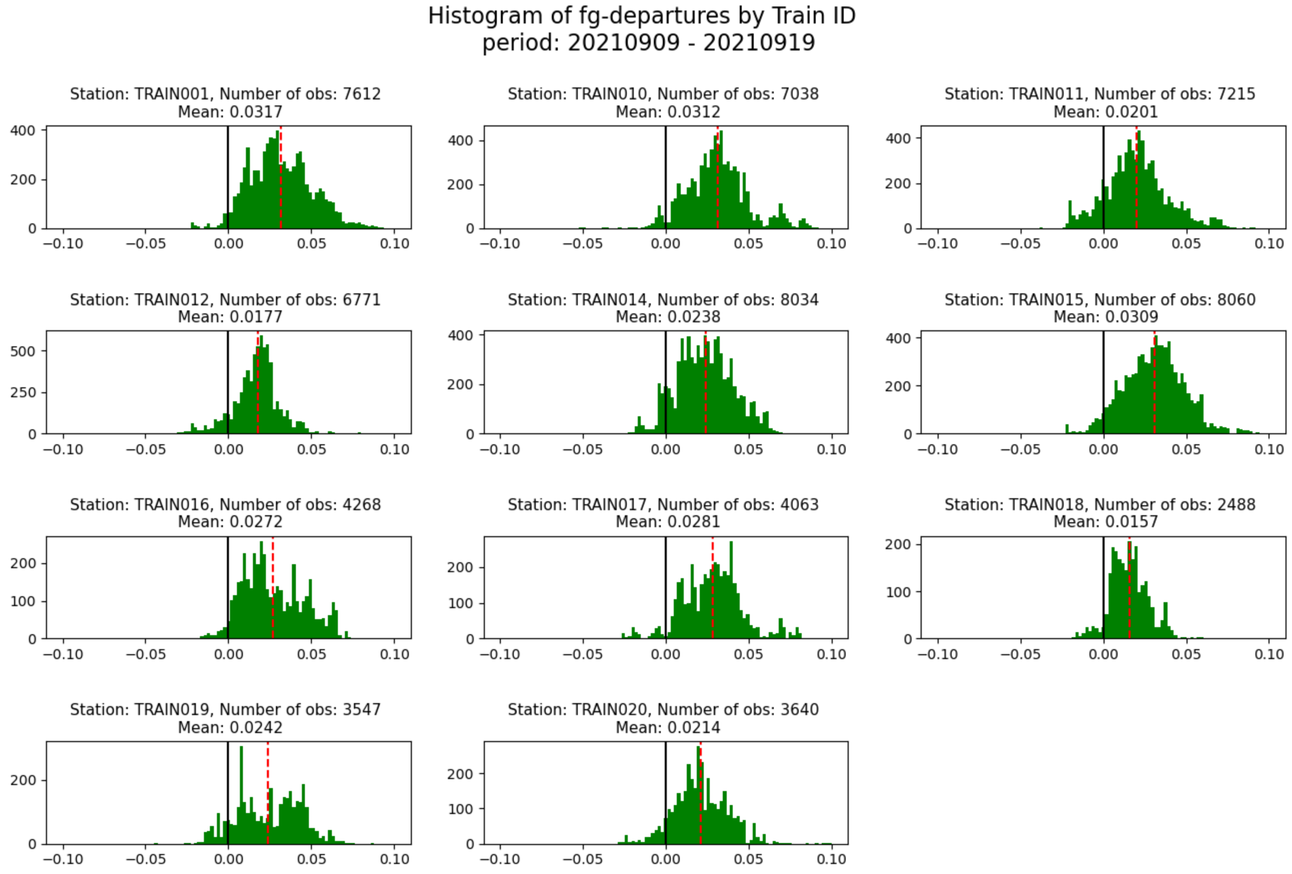
Demonstration that it technically works

Analyse first guess departures from trains

- Reference run with passive assimilation of train based ZTD for the first 10 days of period under study
- Spatial thinning of passive train observations 200m
- Observation error for train observations 0.02m

Both data set without bias correction

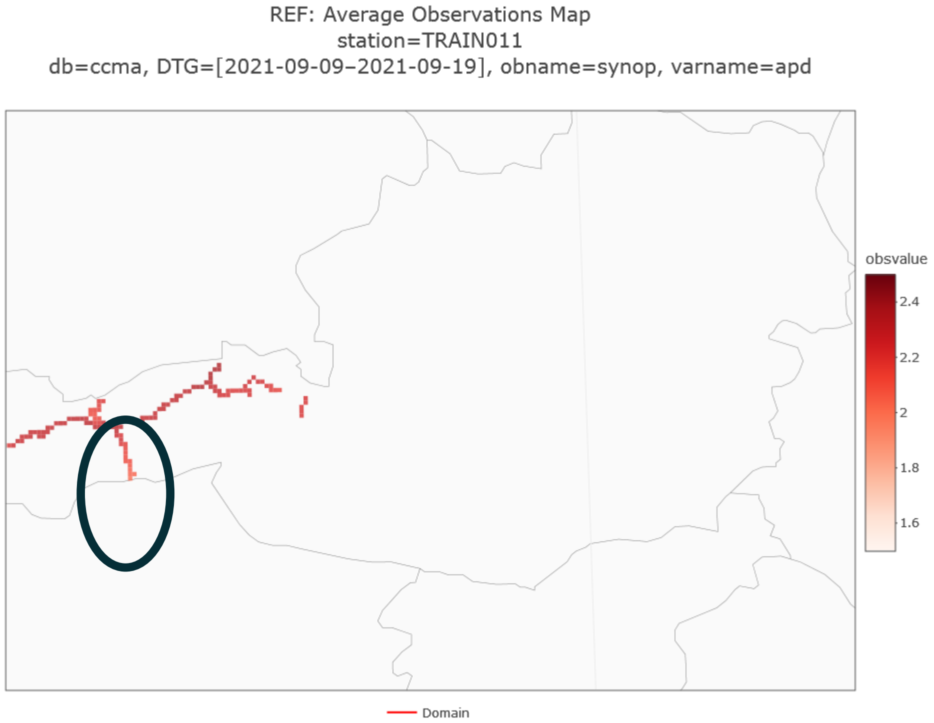




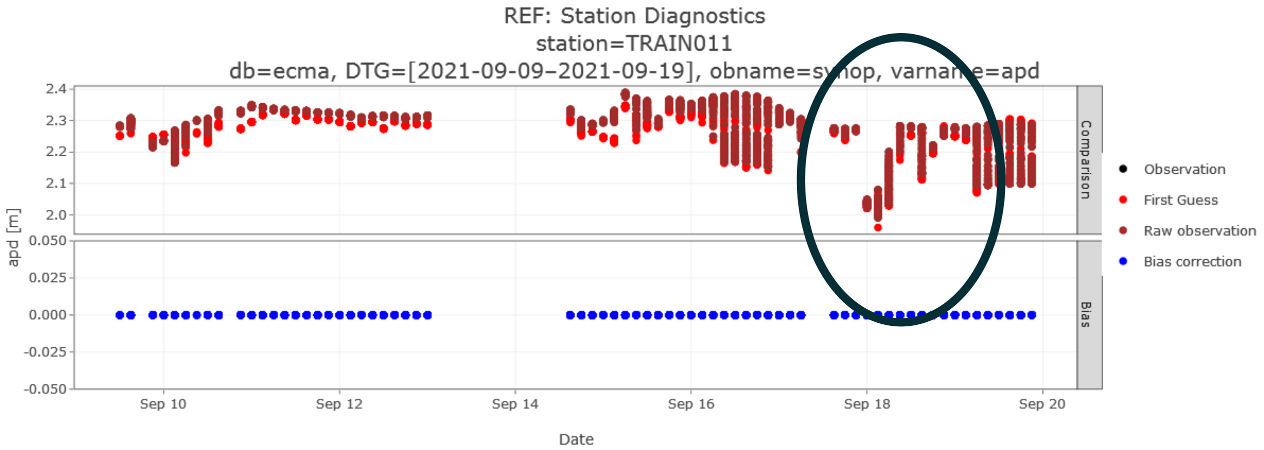
Used to populate LIST_GPSSOL for static bias correction

GeoSphere Austria

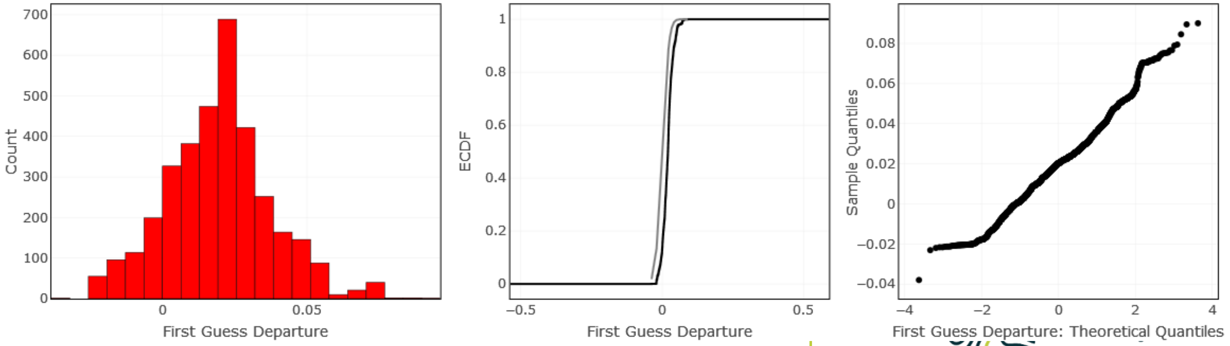
Analyse first guess departures from trains



- Data extraction for obsmon only every 3 hours
- Train operates in the area of Innsbruck
- 18/9/2021 train was moving between Innsbruck (570m) and Brenner Pass (1400m) in the morning

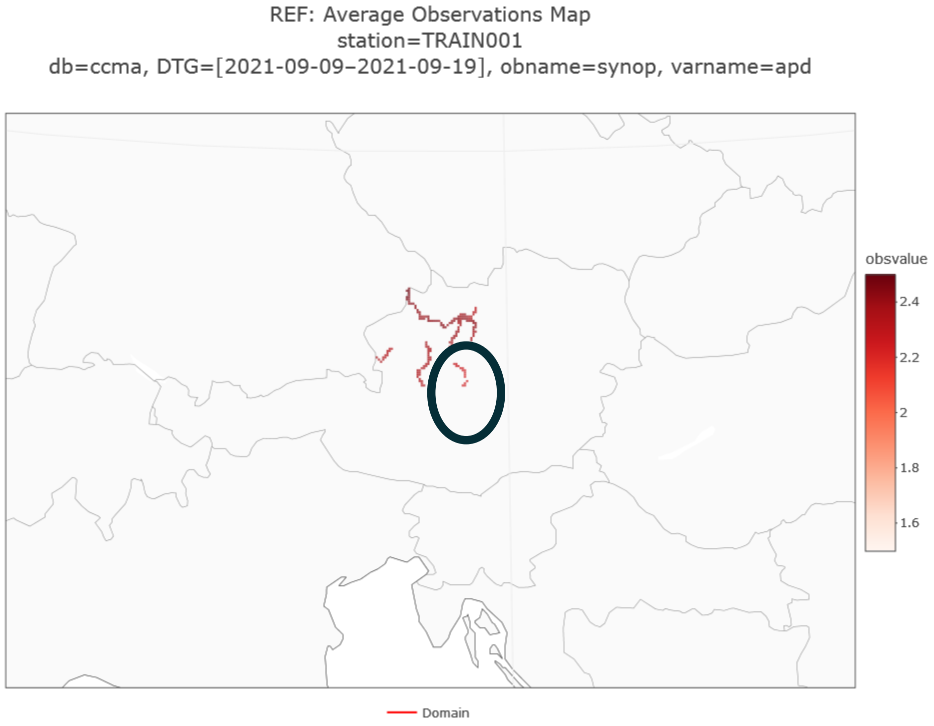


- Histogram of fg-departures indicate positive Bias
- Q-Q Plot confirms assumption of normal-distributed data set

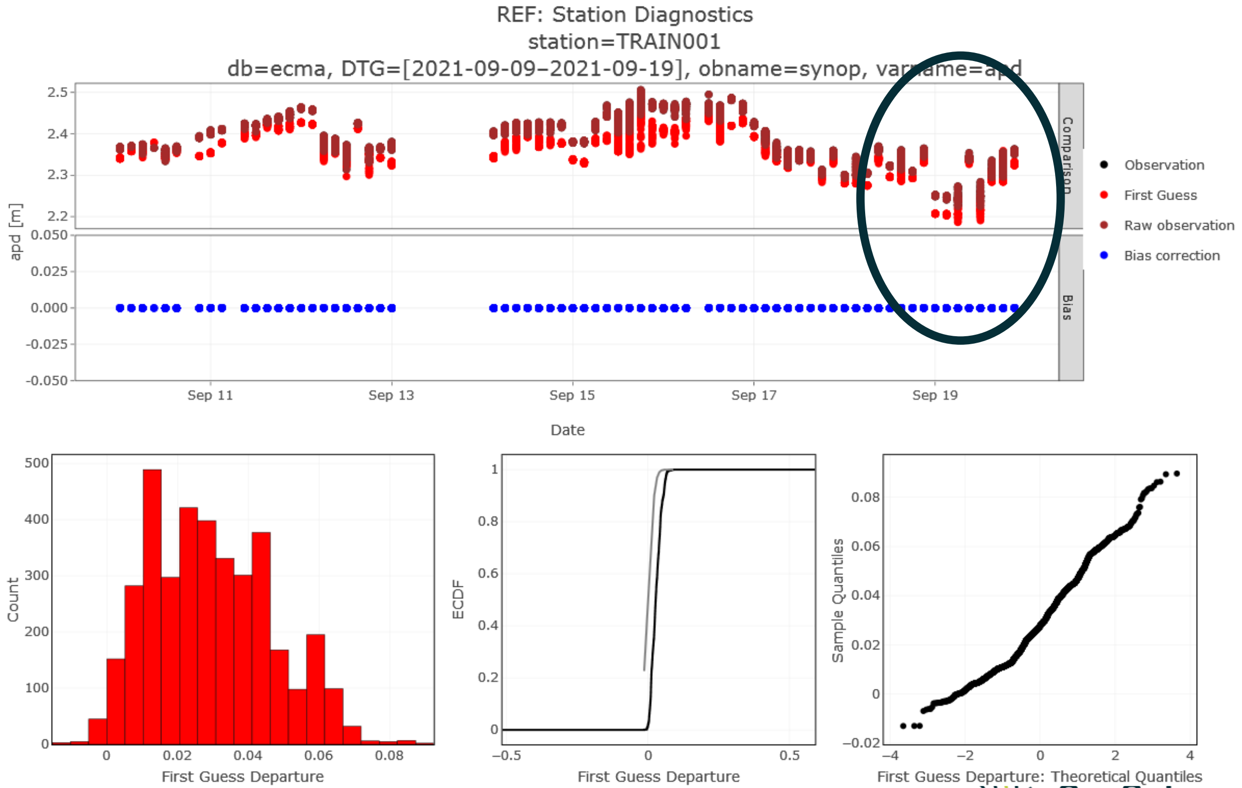


Austria

Analyse first guess departures from trains

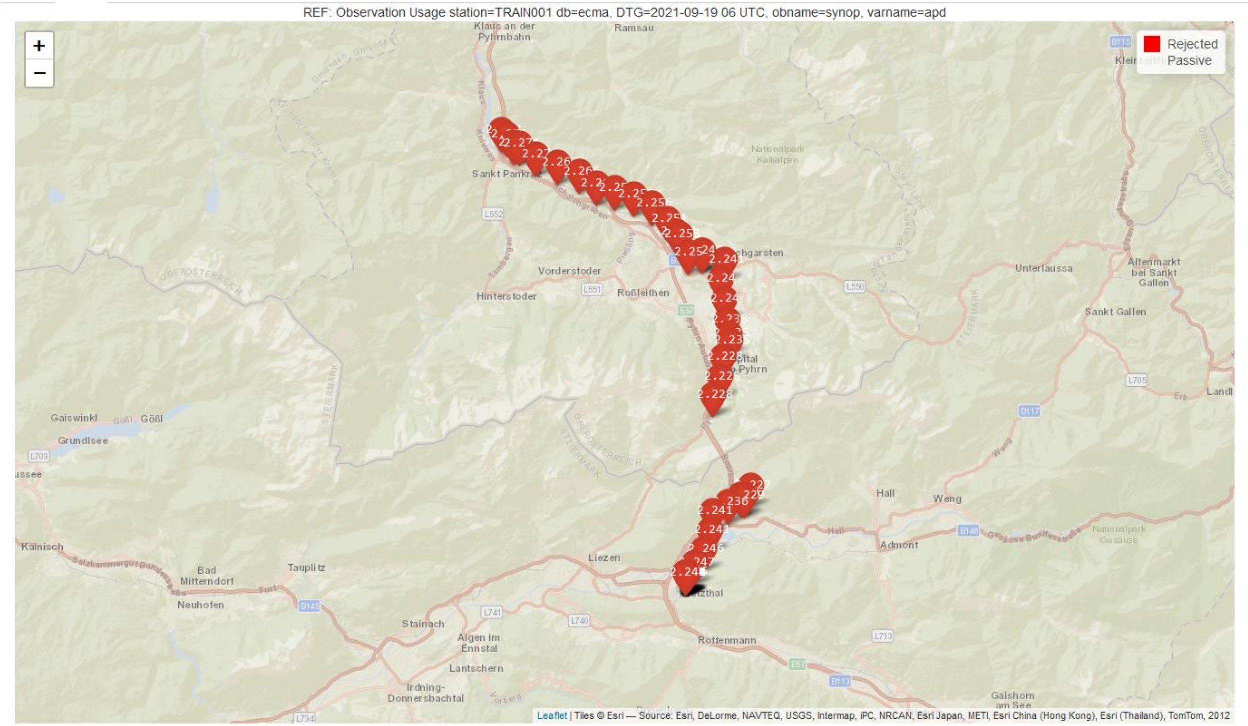


- Data extraction for obsmon only every 3 hours
- Train operates in upper Austria, both in relatively flat and orographic terrain
- 19/9/2021 moves along narrow valley

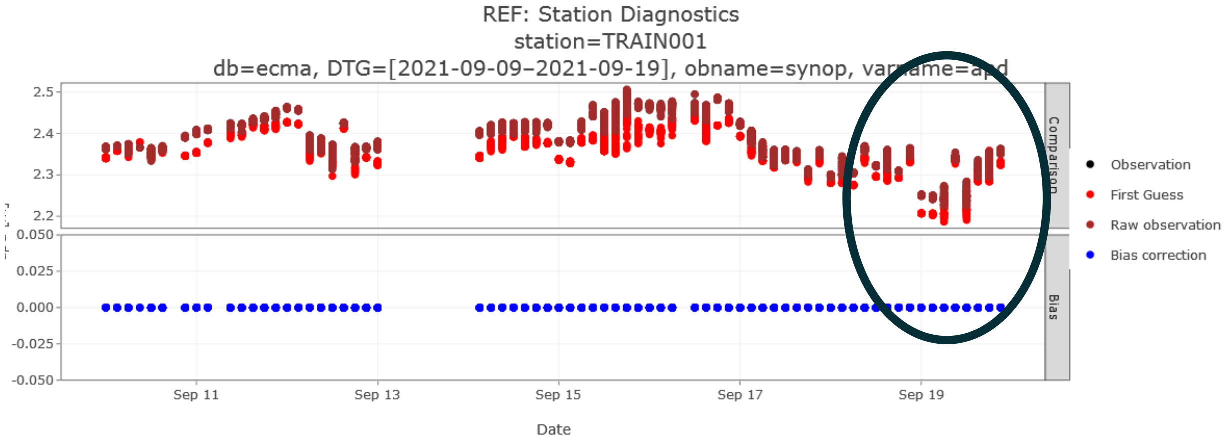


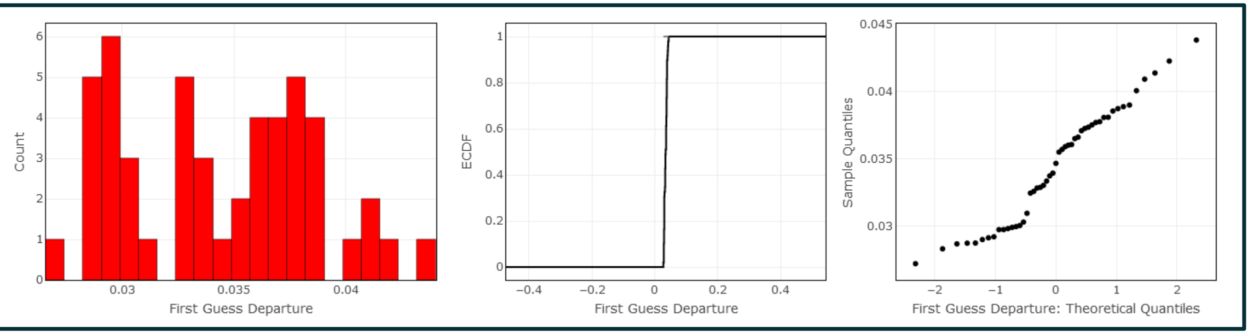
GeoSphere Austria

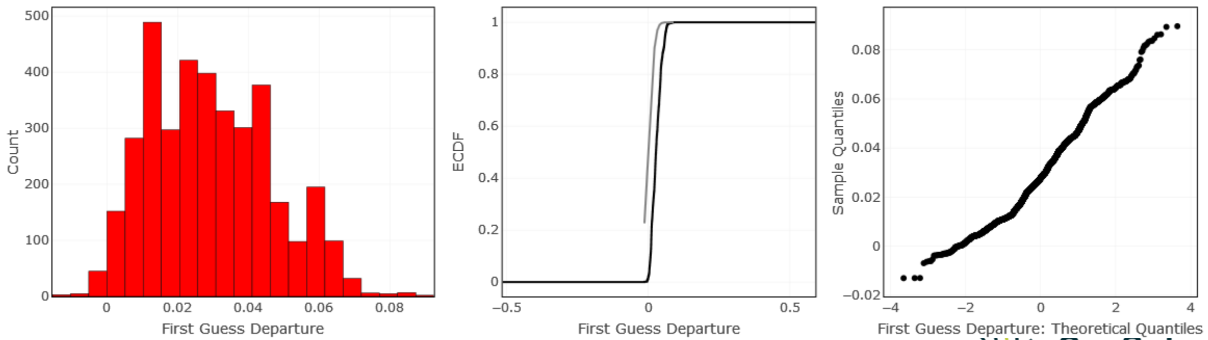
Analyse first guess departures from trains



- Even in complex terrain with tunnels data quality seems to be reasonable
- FG-departure distribution for 6 UTC run on 19/09/2021 is in the middle of 10-day distribution







06 UTC run on 19/09/2021

GeoSphere Austria

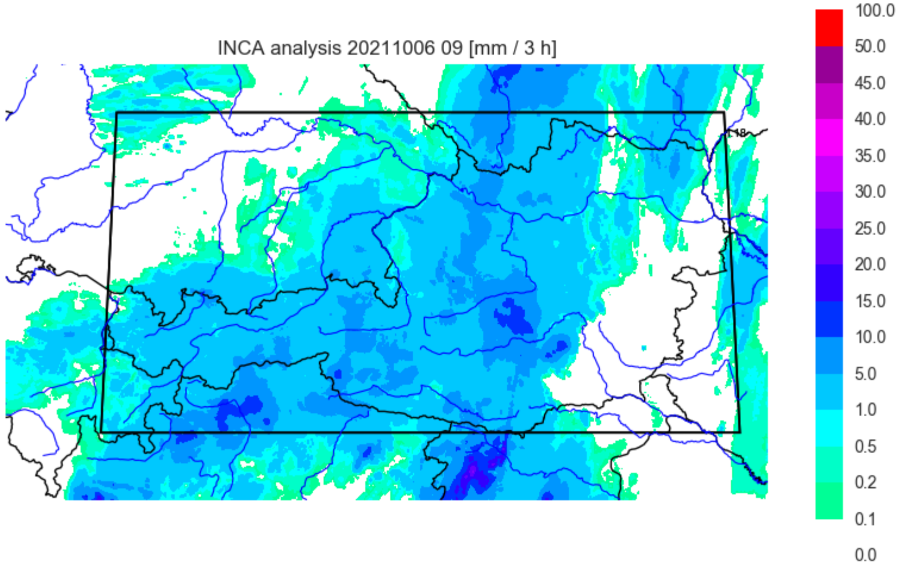
Case study 6.10.2021

Finding optimal setup for assimilation of train based ZTD observations:

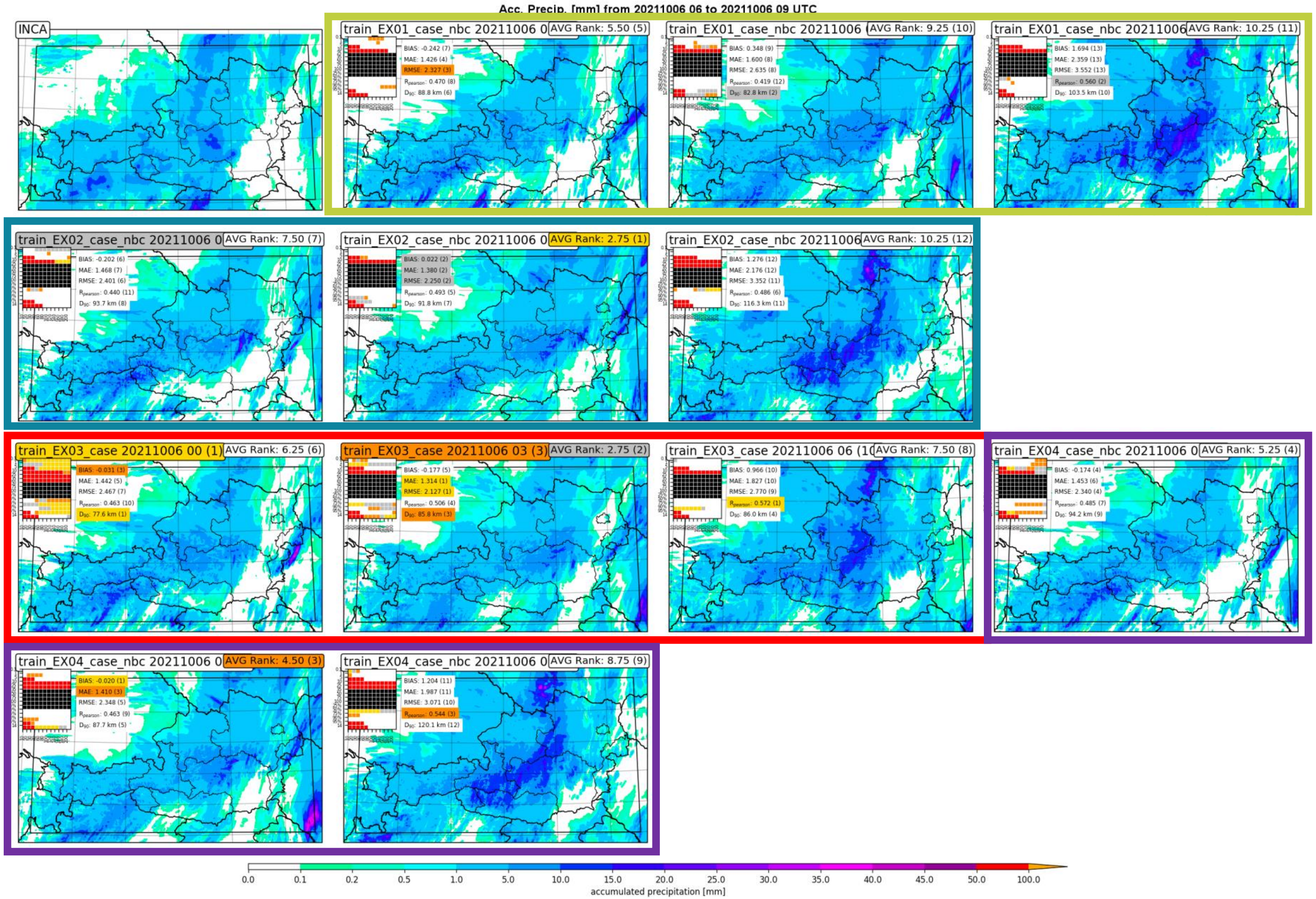
- Running case studies to evaluate impact of train observations
- Reference run (less observations than operational AROME-RUC) with static bias correction for stationary GNSS receivers
- Case selected as station based point verification revealed strong positive bias for 6 UTC precipitation forecast
- Experiments initialized on 5.10.2021 00 UTC with identical initial conditions

- Evaluation based on precipitation up to now

Experiment	Static GNSS	Train GNSS	thinning
REF	**X**		
incl_trains	**X**	**X**	1200m
Only_trains		**X**	1200m
12km_thin	**X**	**X**	12000m



Case study 6.10.2021



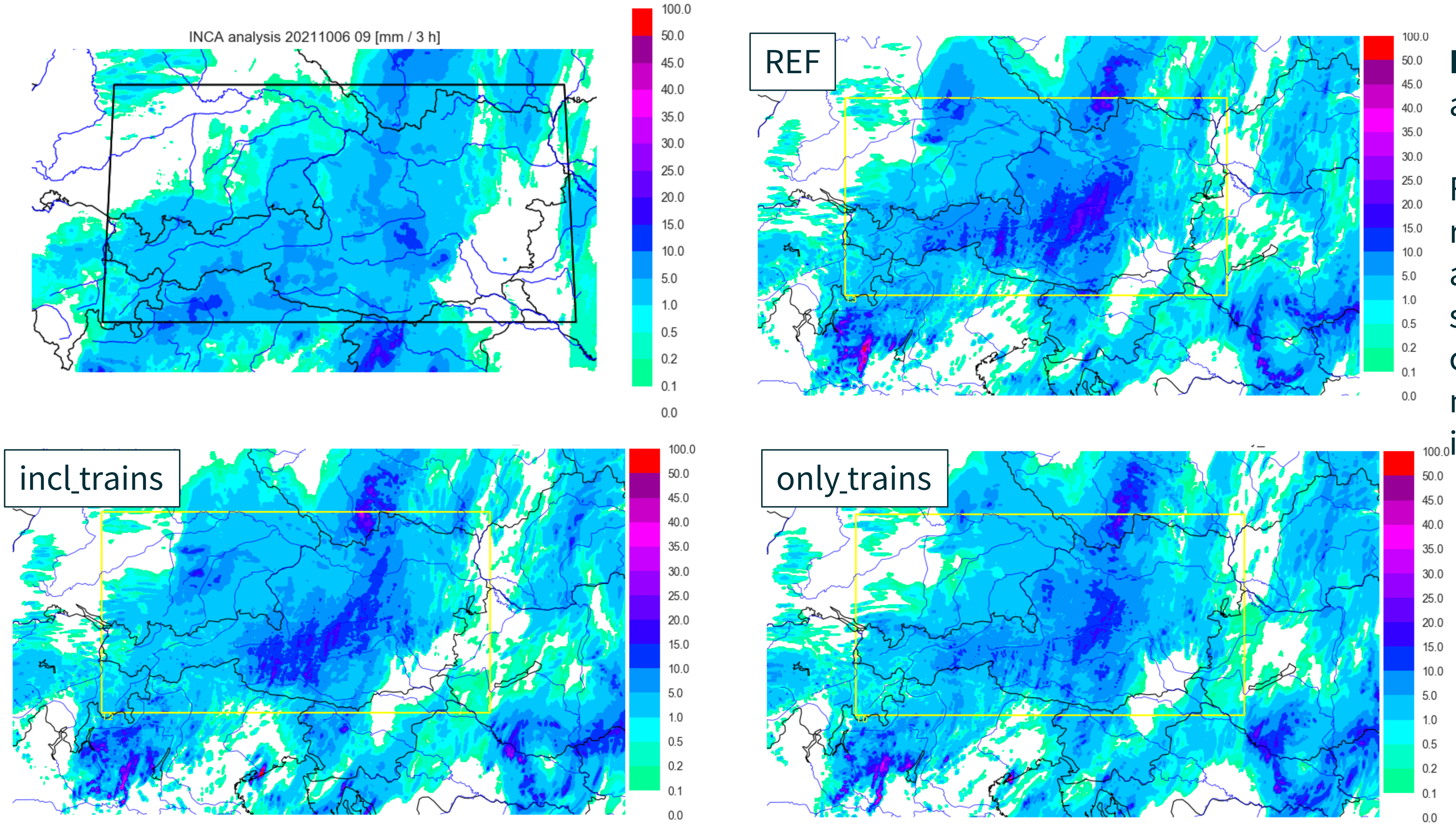
REFERENCE
Trains included
Only trains
12km tinning

Nice: Including train data is beneficial for that case study

Not so nice: Runs with longer lead times perform better than short lead times

GeoSphere Austria

Case study 6.10.2021, initialized at 06 UTC

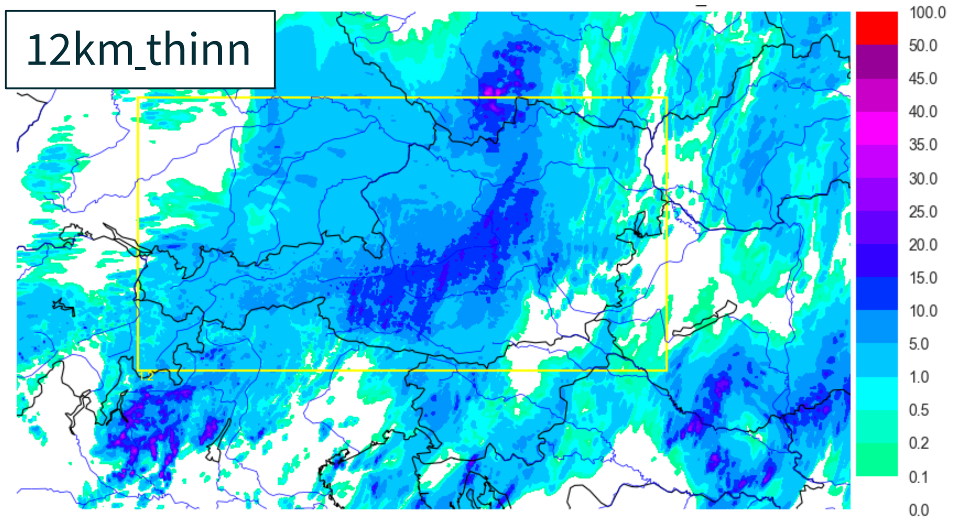


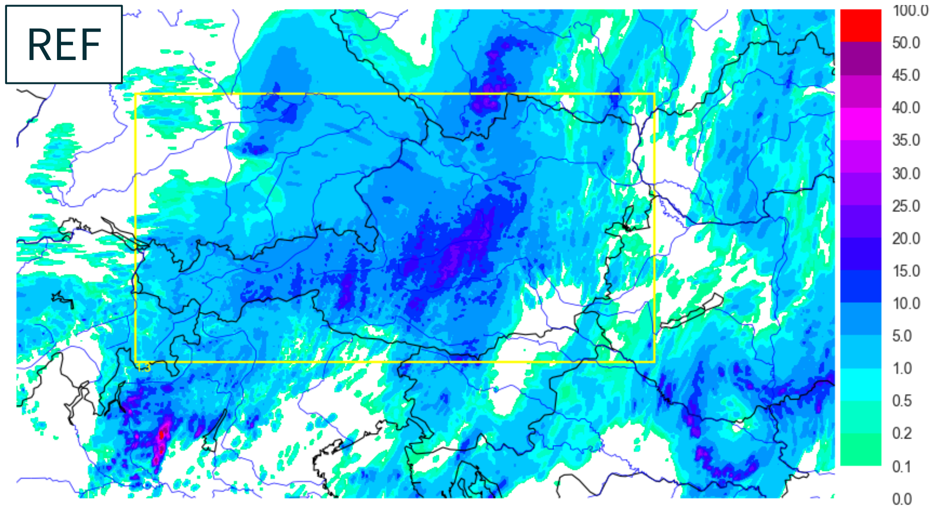
REF: overestimates prec amount

For this case, trains reduce humidity in the analysis, while some stations especially in central Alps depict negative fg-departures -> introduce moisture

GeoSphere Austria

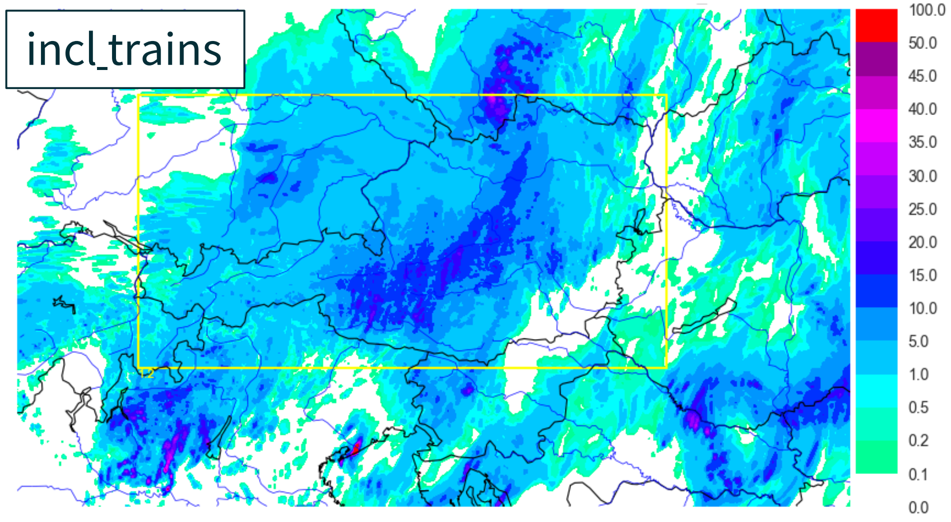
Case study 6.10.2021, initialized at 06 UTC

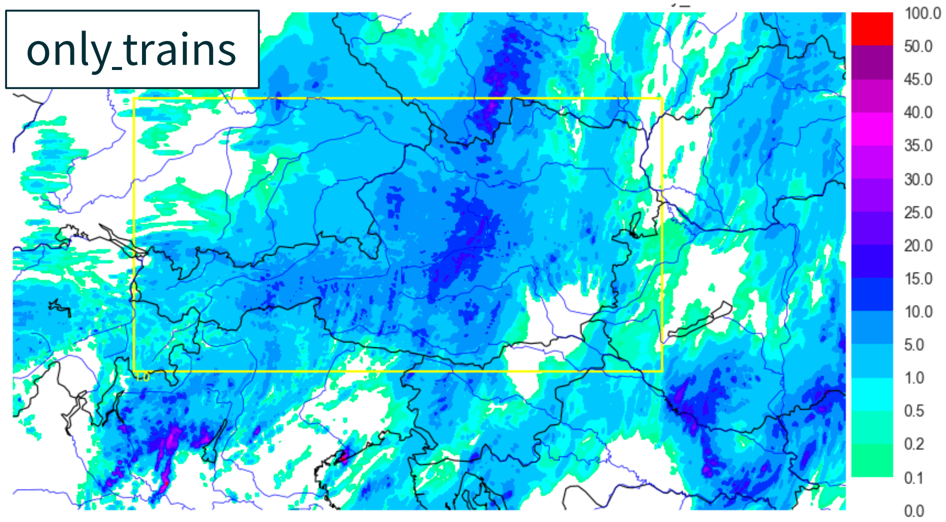




REF: overestimates prec amount

With stricter thinning impact of trains is reduced -> precipitation forecast is in the middle between REF and incl trains





GeoSphere Austria

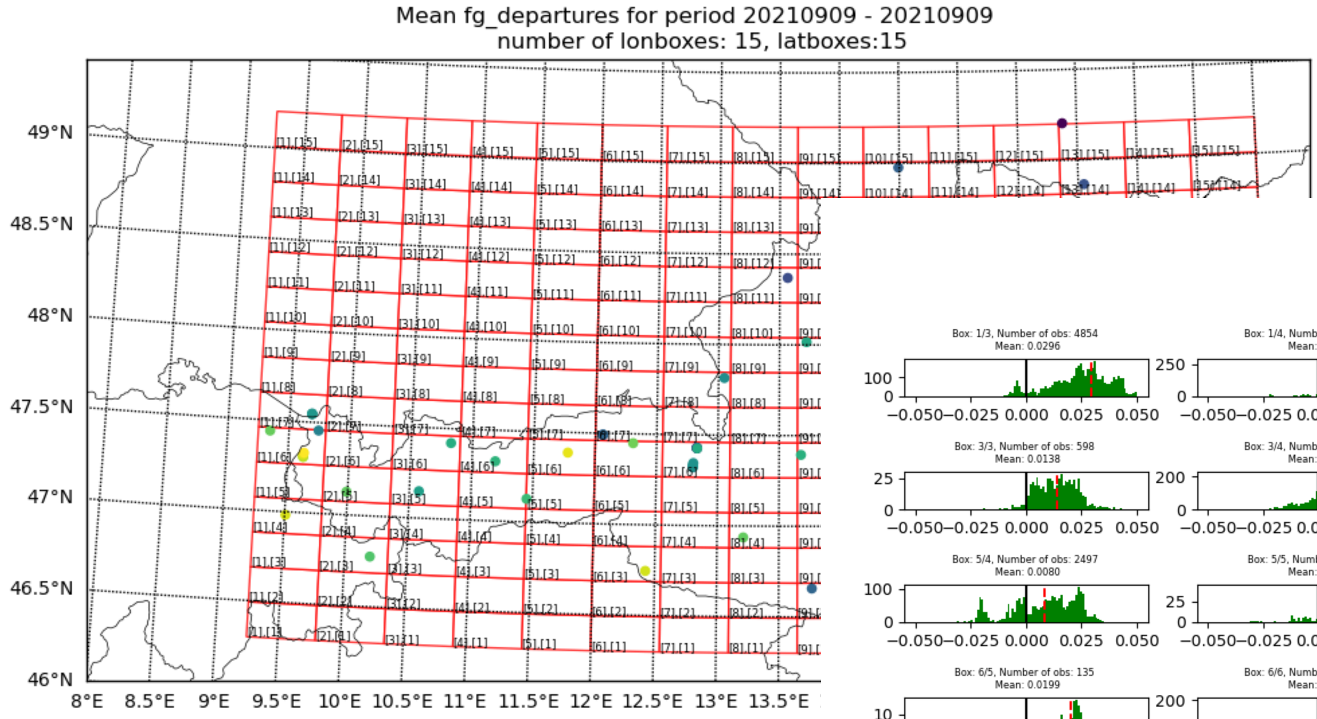
Summary and outlook

Summary:
- Approx. one month of data are available
- Train based ZTD observations look surprisingly good even in very complex terrain
- Simple bias correction by train-ID seems to do its job reasonably well
- Further testing of best setup required

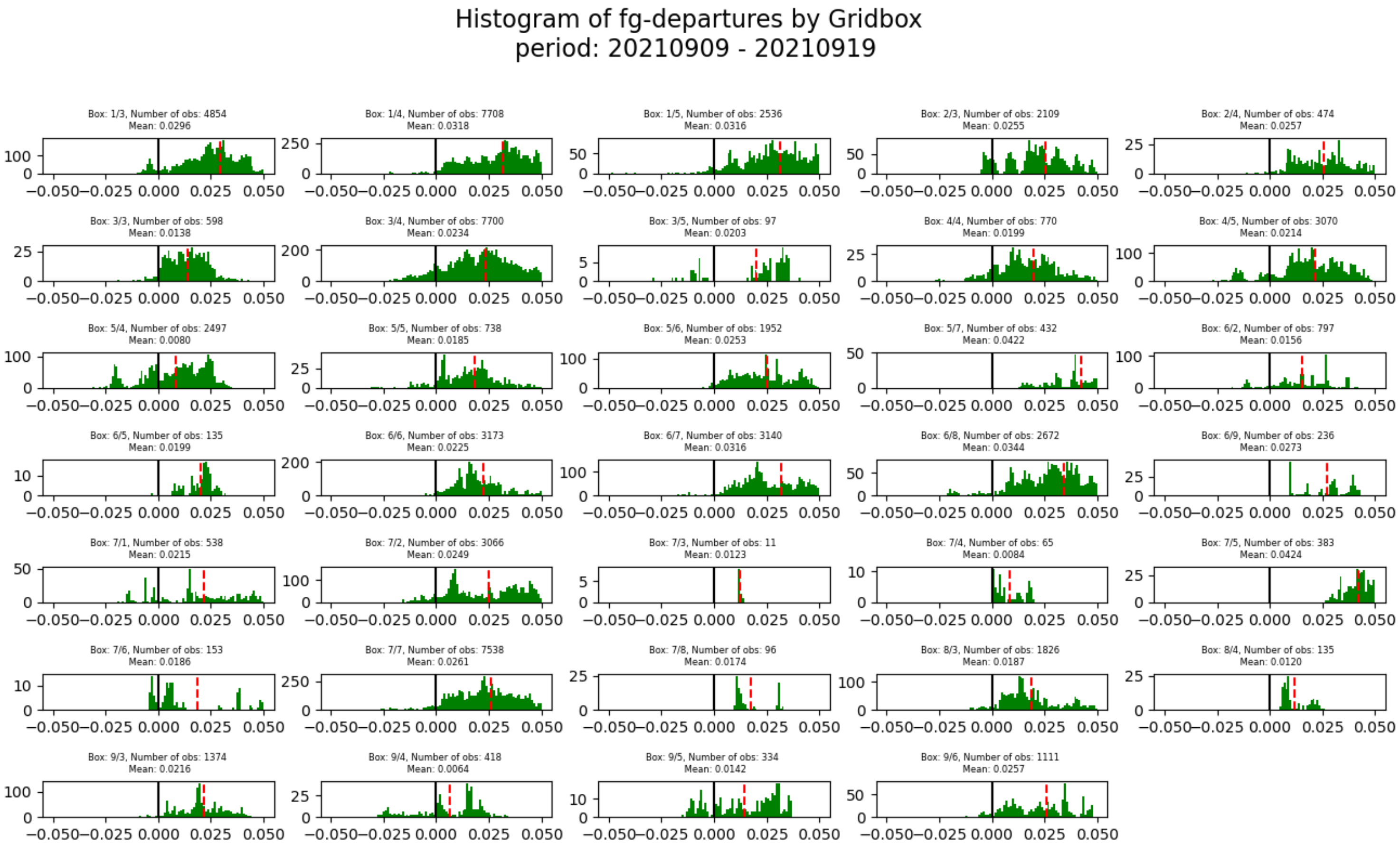
Outlook for the project:
- Can we apply VARBC? Height dependent predictors might require to treat observations seperated by trainID and terrain height.
- Further case studies planned as well as the usage in a sub-hourly AROME-RUC cycle
- At least one experiment over whole period planned

Can we go operational in a foreseenable future?
- Currently data retrieve is technically not possible on the fly, data have to be downloaded from on-board hard-disc manually
- Investigations at OEBB are on-going how to retrieve data on-the-fly, one possible solution would be a data transfer at larger train stations, or possibly with the usage of 5G
- Timely processing of large amount of data by TU-Wien is under investigation

Analyse first guess departures from trains

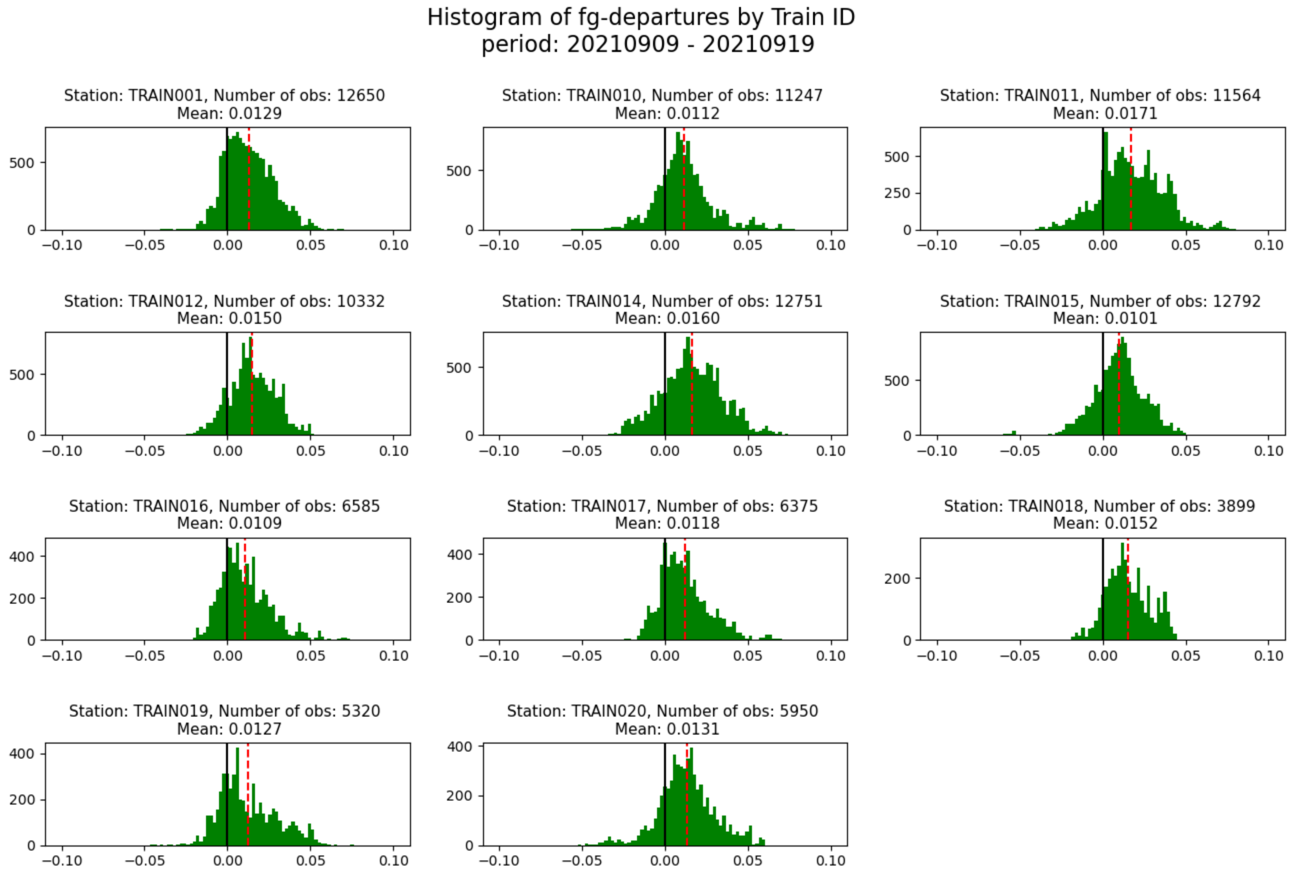


For completeness



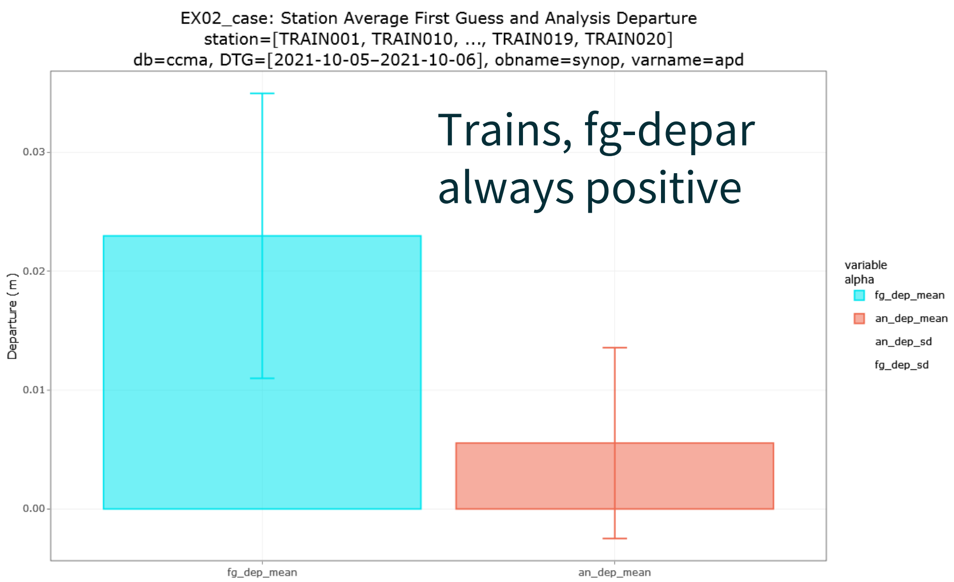
Apply static bias correction and rerun trainingsperiod

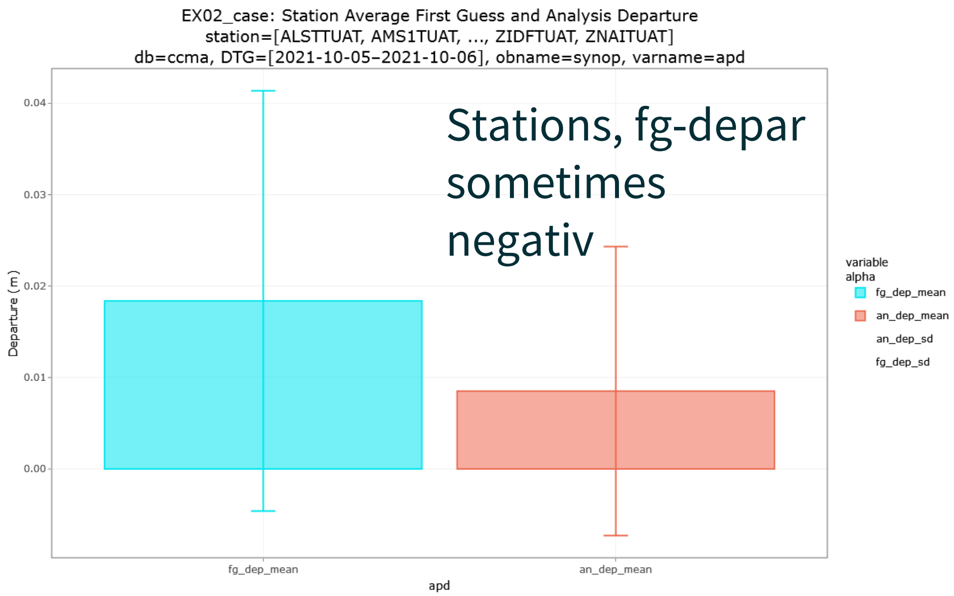
Bias reduced ….

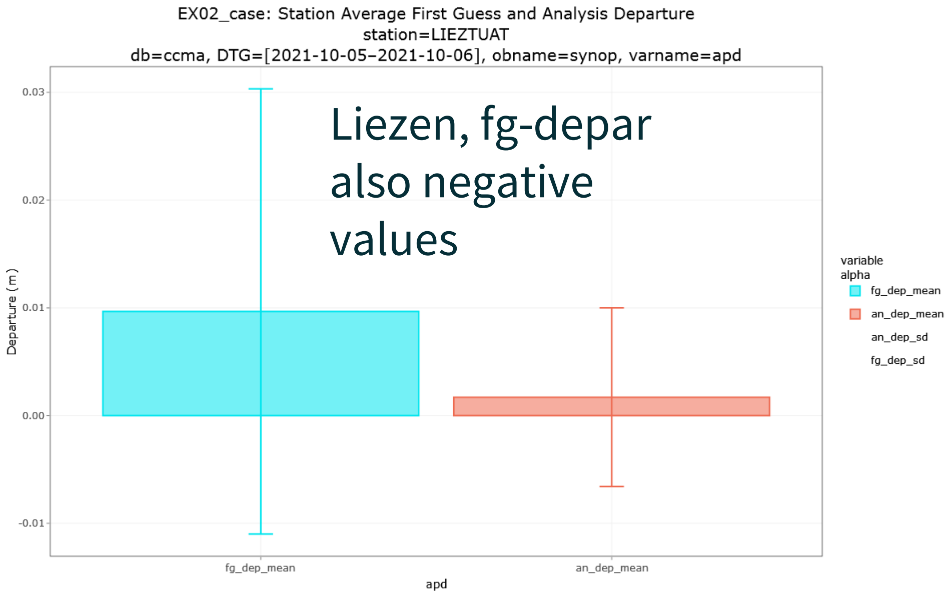


GeoSphere Austria

Analysis fg-departures case1







GeoSphere Austria