Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology MeteoSwiss

Nowcasting of thunderstorm severity with Machine Learning in the Alpine Region

Hamann U., Zeder J., Beusch L., Clementi L., Foresti L.,
Hering A., Nerini D., Nisi L., Sassi M., Germann U.

COALITION-3 – Faces

Ulrich Hamann
COALITION
project lead

Joel Zeder
Intern in 2018/2019
COALITION-3

Luca Nisi
COALITION-1
forecaster

Alessandro Hering
TRT developer

Elena Leonarduzzi
Intern in 2016
COALITION-2

Urs Germann,
head of MDR

Lorenzo Clementi
head of MDRD

Loris Foresti
precip. Attractor
pySTEPS

Daniele Nerini
pySTEPS

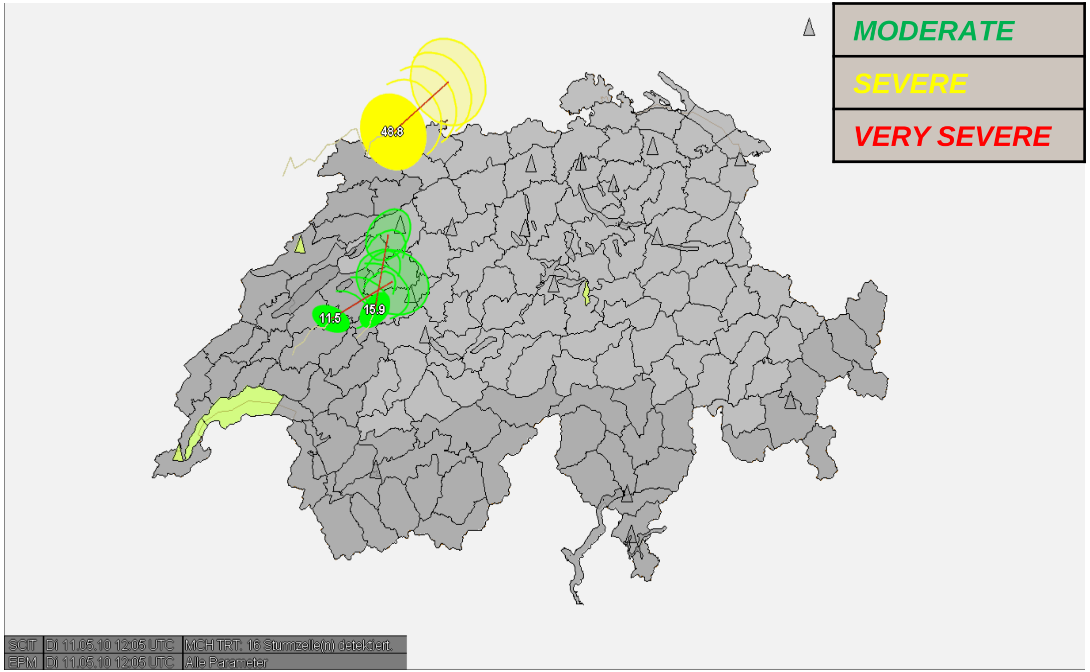
Marco Sassi
NWC-SAF,
deployment

Lea Beusch
Intern in 2017
Satellite Rainfall

Starting point before our study

- Thunderstorm cells are identified by radar with adaptive reflectivity thresholds.
- Thunderstorm intensity is expressed as heuristic TRT rank (in colours).
- Future position is extrapolated with the current motion.
- TRT rank is kept constant.
- Multi-sensor cell parameters are monitored.
- Automatic warning suggestions are generated for warning regions.
- Warning suggestions are modified by forecasters.

Example for Thunderstorm Radar Tracking (Hering et al)
11 May 2010 for Switzerland. TRT rank expressed in colours.

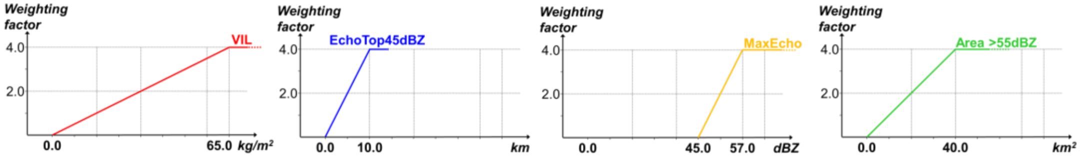


TRT rank (heuristic thunderstorm intensity)

Cell severity ranking: Single **numerical score** $[\mathbf{0.0}, \mathbf{4.0}]$ based on cell attributes integrated with a weighting scheme (**fuzzy logic** like approach):

- Vertical integrated liquid $\boldsymbol{VIL}$
- Median of 45 dBZ Echo Top altitude $\boldsymbol{ET45}$
- Maximum cell reflectivity $\boldsymbol{MaxEcho}$
- Area of cell reflectivity ≥ 57 dBZ $\boldsymbol{area57dBZ}$

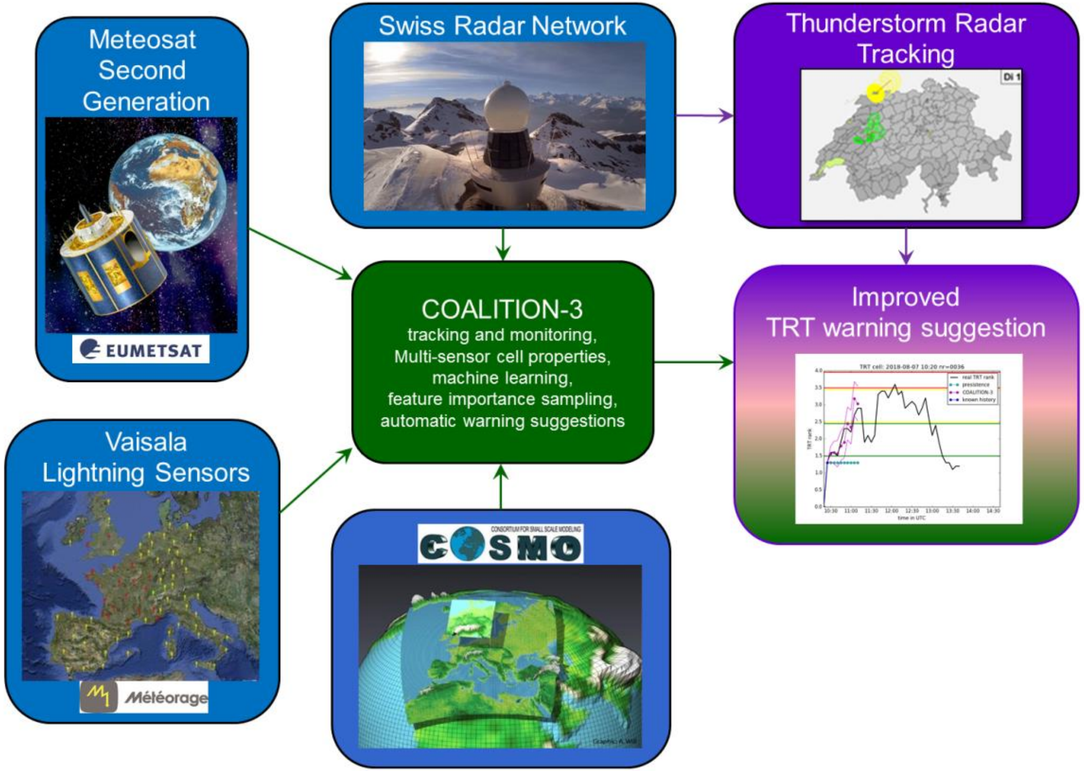
Severity	RANK
DEVELOPING	$RANK = [1.2 - 1.5[$
MODERATE	$RANK = [1.5 - 2.5[$
SEVERE	$RANK = [2.5 - 3.5[$
VERY SEVERE	$RANK = [3.5 - 4.0]$



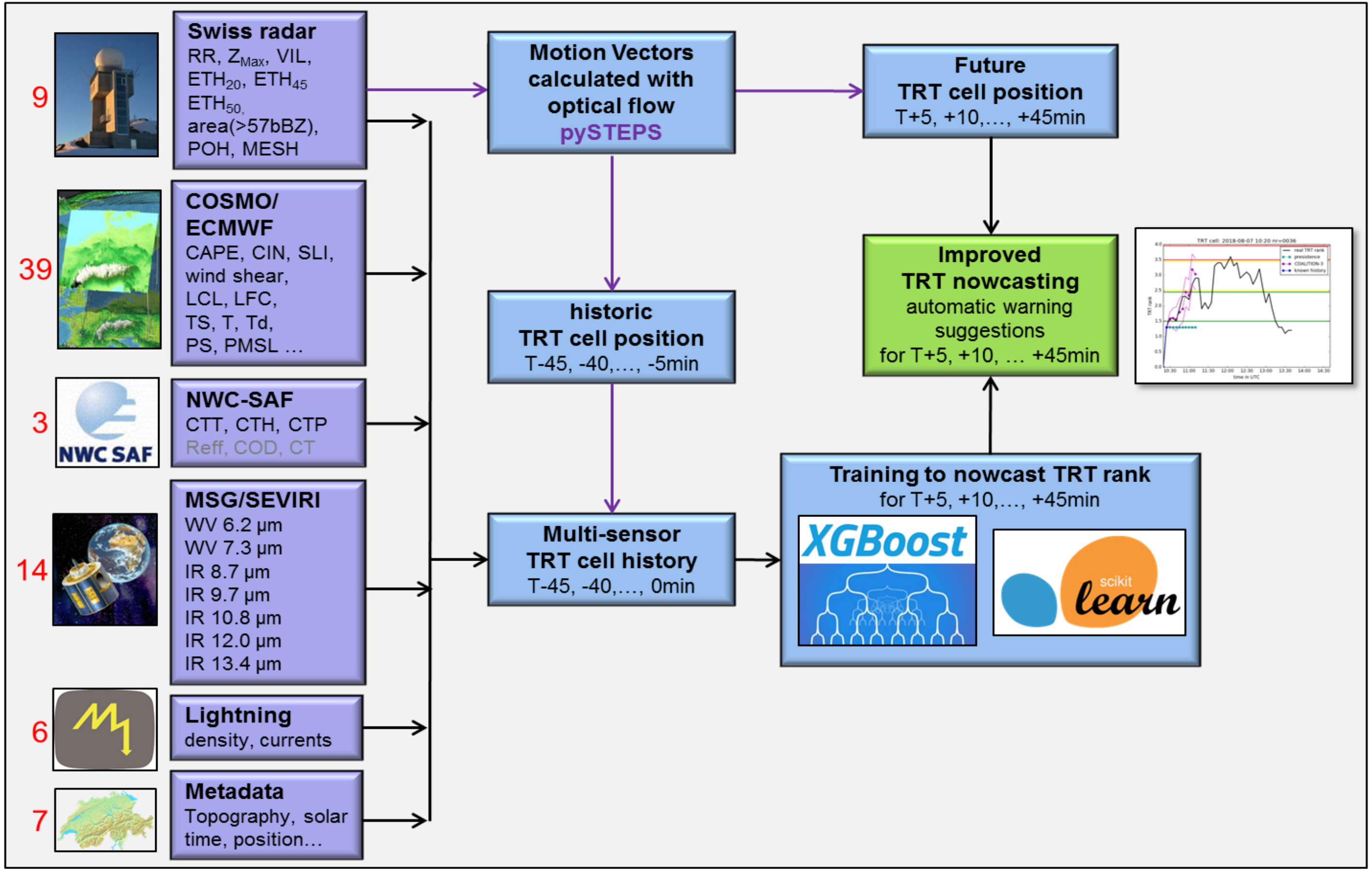
$$RANK = \frac{\mathbf{2.0} * \boldsymbol{g}(\boldsymbol{VIL}) + \mathbf{2.0} * \boldsymbol{g}(\boldsymbol{ET45med}) + \mathbf{1.0} * \boldsymbol{g}(\boldsymbol{dBZmax}) + \mathbf{2.0} * \boldsymbol{g}(\boldsymbol{area57dBZ})}{\mathbf{7}}$$

Goals for this study (COALITION-3)

- Improved automatic TRT warning suggestions
  a) thunderstorm position
  b) thunderstorm intensity
     (expressed as TRT rank)
- Multi-sensor retrieval (satellite, radar, lightning, COSMO, meta-data)
- Update cycle 5 min
- Nowcast up to 45 min
- Long warning lead times
- Quantitative, customer oriented output

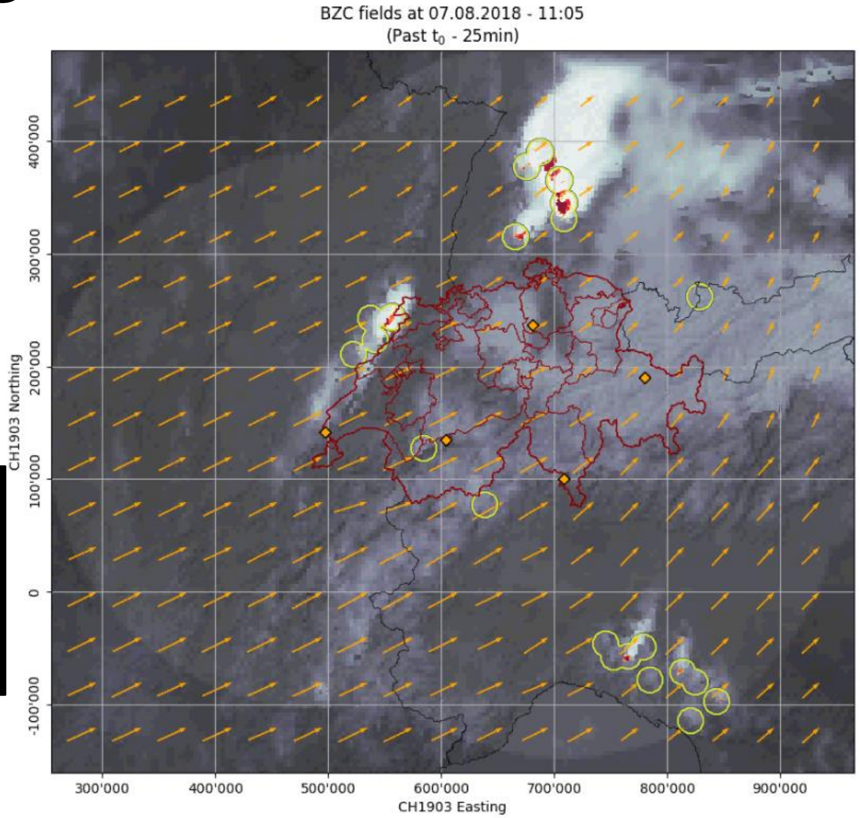


Flowchart of COALITION-3



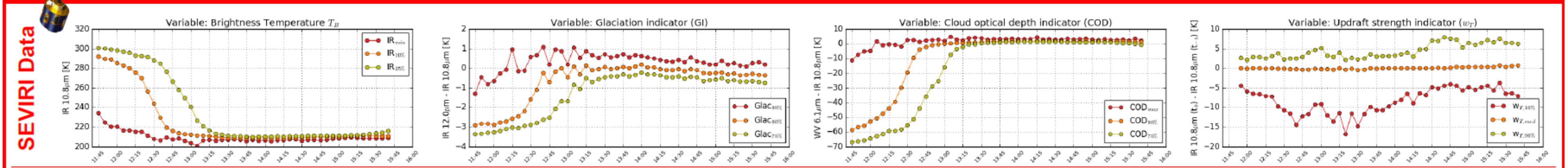
Retrieving cell histories

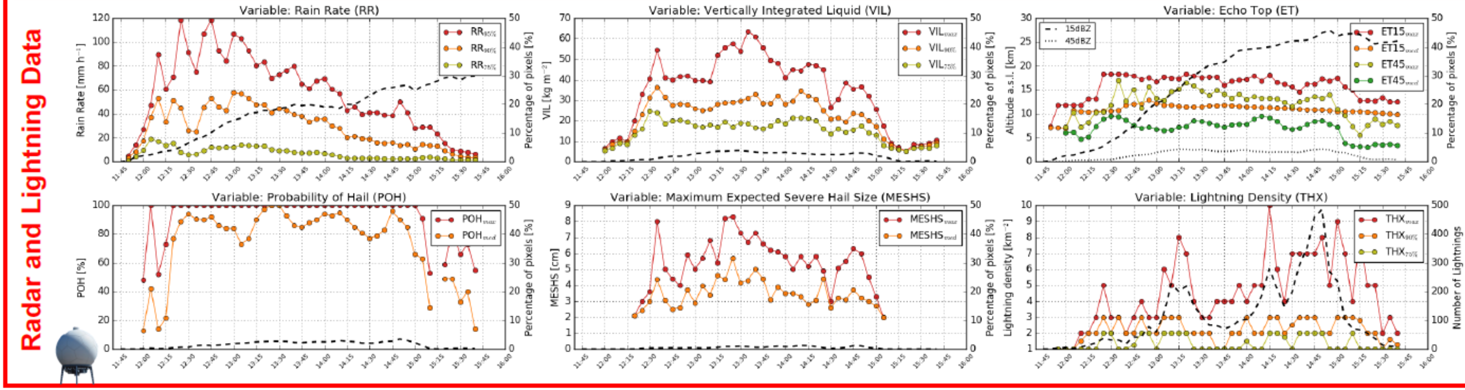
- Calculating motion vectors with pySTEPS. Feature selection with Shi and Tomasi (1994). Tracking with pyramidal implementation of Lucas-Kanade (1981) feature tracker algorithm
- Initial cell identification with TRT algorithm.
- Track position 45 min backwards in time to retrieve cell history (predictors).
- For each predictor following statistics are calculated in a 23 km diameter circle:
  - Mean, Sum, Standard Deviation
  - Minimum and maximum value
  - 1%, 5%, 25% 50%, 75%, 95%, 99% Quantiles
  - Number of pixels with certain properties, e.g. precipitation > 0mm/h
- Monitor the thunderstorm intensity (TRT rank) up to 45 min into the future (truth for training)



An example of typical cell histories

4. SEVIRI, Radar, and Lightning Data within Lagrangian Framework





- The **lagrangian framework** allows to derive **radar and satellite data for a specific non-stationary convective cell over time**.
- Depending on the variable, **specific quantiles** are derived for visualisation. Furthermore, the fraction **of covered pixels** is shown.
- This time series data can be further **exploited for automated nowcasting applications**, for example using machine learning algorithms.

Input data - Cardinality

Overall, the final count of features is composed of:

- **68** input variables (RZC, IR 10.8 μm, PV at 500 hPa, Slope …)
- **10** time steps (t_0, t_{-5min}, … , t_{-40min}, t_{-45min})
- **12** Statistics (`Sum, Mean, Quantiles, Pixel counts` …)
  plus some conditional statistics of Radar variables
- 68x10x12 = **8’160 possible input parameters**
- Training period summer 2018, about **10’000 observed cell histories**

TRT Cells identified by (unique) ID and Date/Time

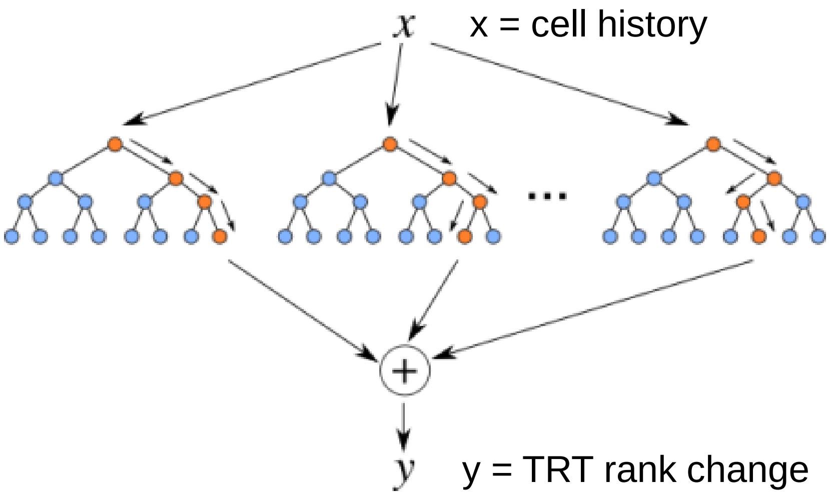
```
                                    RZC_stat|-15|PERC75  RZC_stat|-15|PERC95  RZC_stat|-15|PERC99  RZC_stat|-15|MAX  RZC_stat|-10|SUM  ...  TOPO_ALTITUDE_stat|-20|PERC01
DATE_TRT_ID
201804291235_2018042910550020                -22.248047             0.000000             0.000000          0.000000      -3424.883789  ...                     -11.875000
201804291235_2018042912000024                 -1.099609            44.785156            94.321289         87.427734       1311.467773  ...                     -88.125000
201804291235_2018042912000025                 -0.245117            -1.833008           -31.992966        -32.000977       -314.942383  ...                      -5.800003
201804291235_2018042912050022                 -0.861328             3.979492            14.365040         23.396484        103.207031  ...                     -16.274994
201804291235_2018042912100033                 -8.855469           -68.509766           -41.690041          0.000000      -1397.562500  ...                      -0.375000
201804291235_2018042912200024                  0.636719             0.706055             1.180468          0.000000        254.544922  ...                       0.000000
201804291235_2018042912250029                  0.804688             0.711914            -1.006836          1.375000        180.148438  ...                     -17.625000
...                                                 ...                  ...                  ...               ...               ...  ...                            ...
201809181900_2018091819000008                 -0.435547           -10.399414           -14.435349        -13.000000       -731.588867  ...                       8.199997
201809181900_2018091819000033                  0.289062             2.375000             6.605469         -2.296875        208.524414  ...                      57.875000

[34 rows x 10128 columns]
```

Machine learning nowcasting

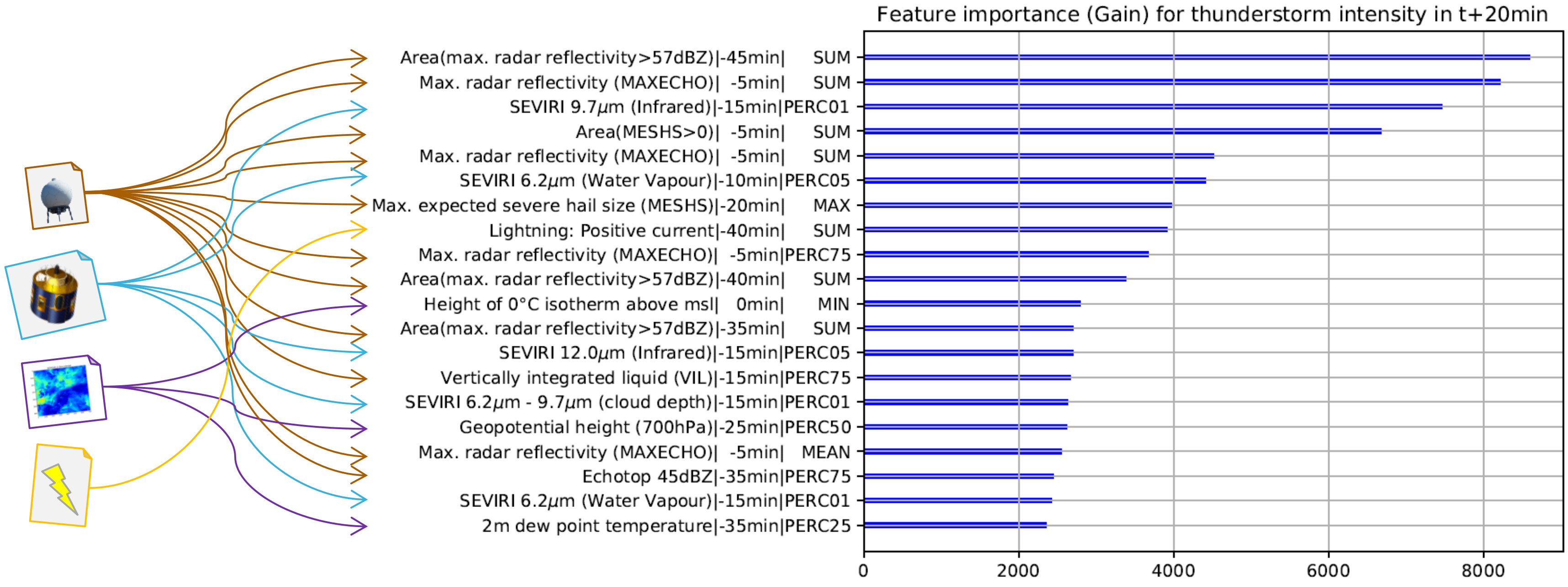
XGBoost model (gradient boosted trees)

- roughly **equivalent to random forest** with an ensemble of weak-learners (trees) as model
- but where during **training**, trees are added until the objective function converges
- where the mean squared error **MSE** is the **loss function**, and the **regularisation** term is the **sum of scores** at the leaves.
- XGBoost also produce an estimate how important an input variable is.
- Number of input parameter could be reduced to 750.



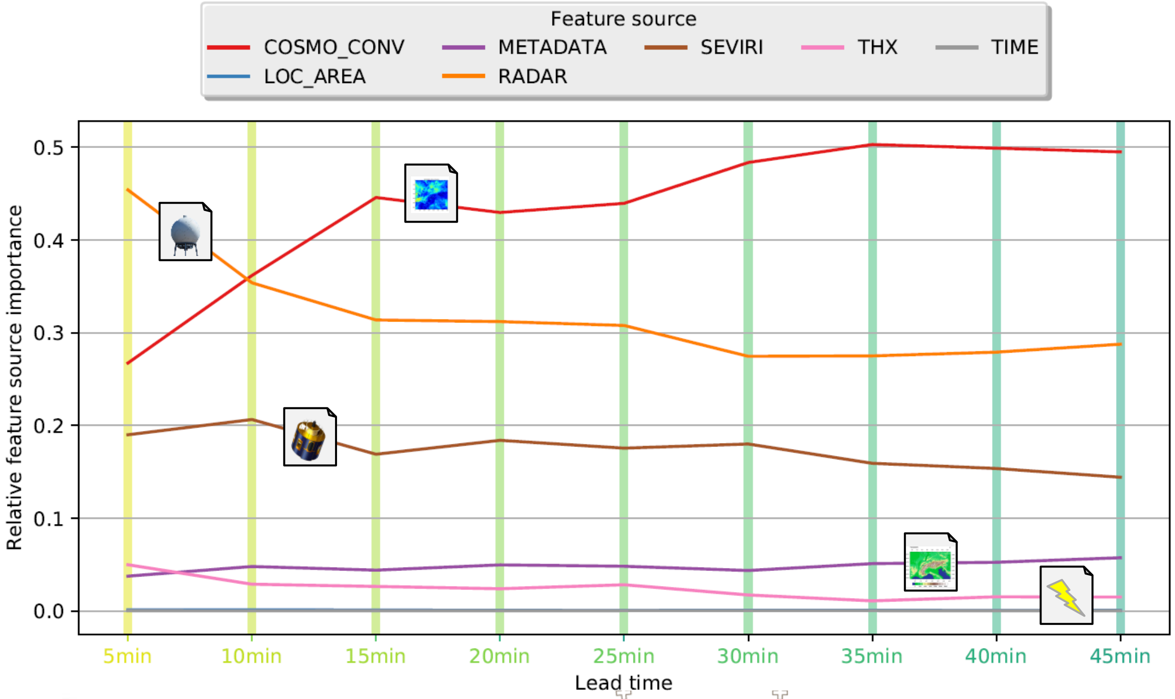
Statistical learning – Feature selection

Top 20 features according to XGB model for t_{+20min} lead time:



Statistical learning – Feature selection

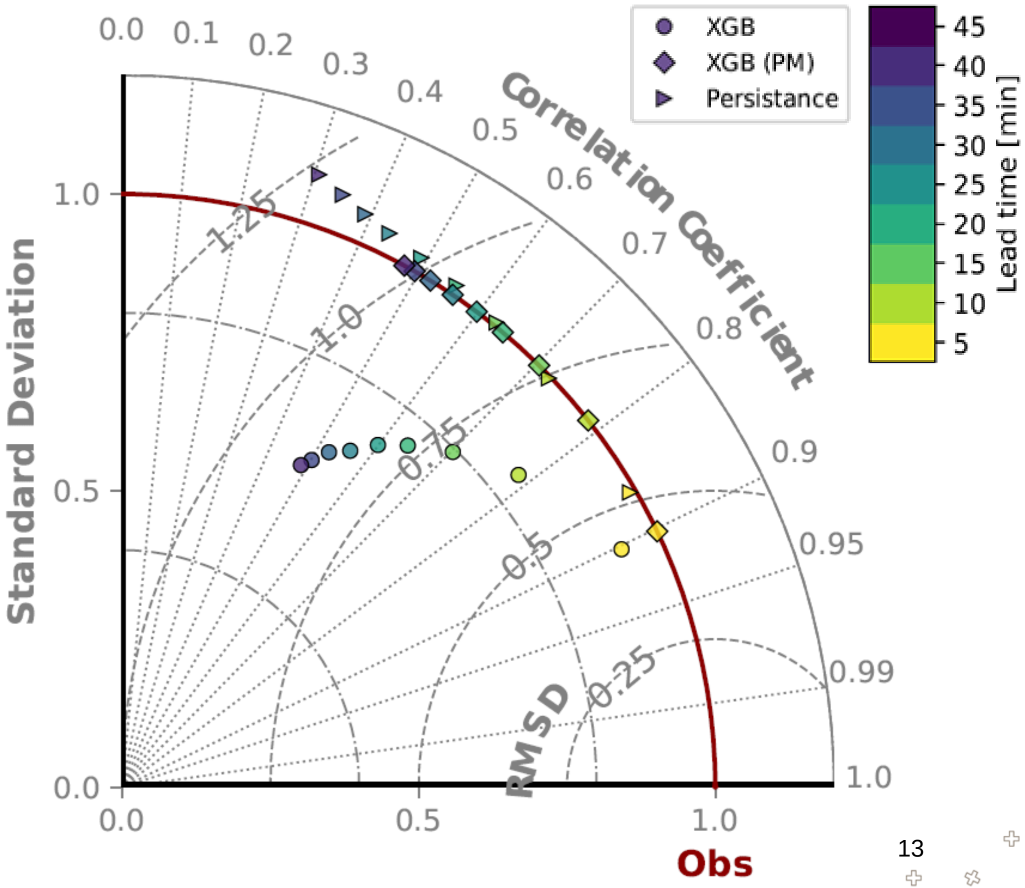
Relative importance of feature source:



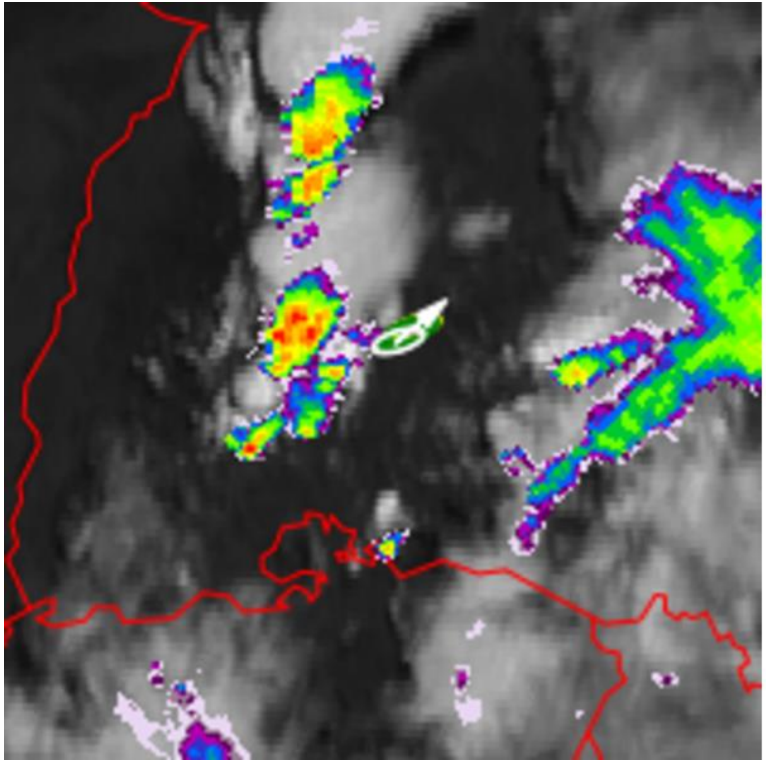
Model evaluation

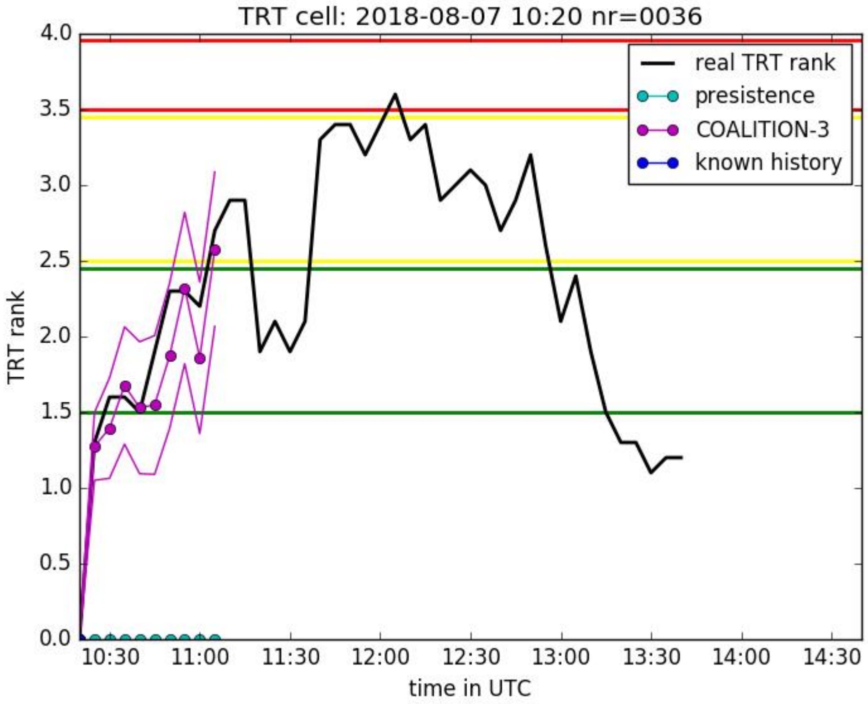
- XGBoost nowcasts (XGB, circles) always have the smallest RMSD (skilful up to t_{+45min})
- Probability matching (diamonds) is used to correct for the standard deviation (skilful up until t_{+35min})
- Probability matched results have a smaller RMSD as persistence (triangles) for all lead times.

Forecast times with same RMSD					
Persistence	5	10	15	20	25
XGBoost, PM	6	14	22	30	42
XGBoost	6	18	40	>45	>45

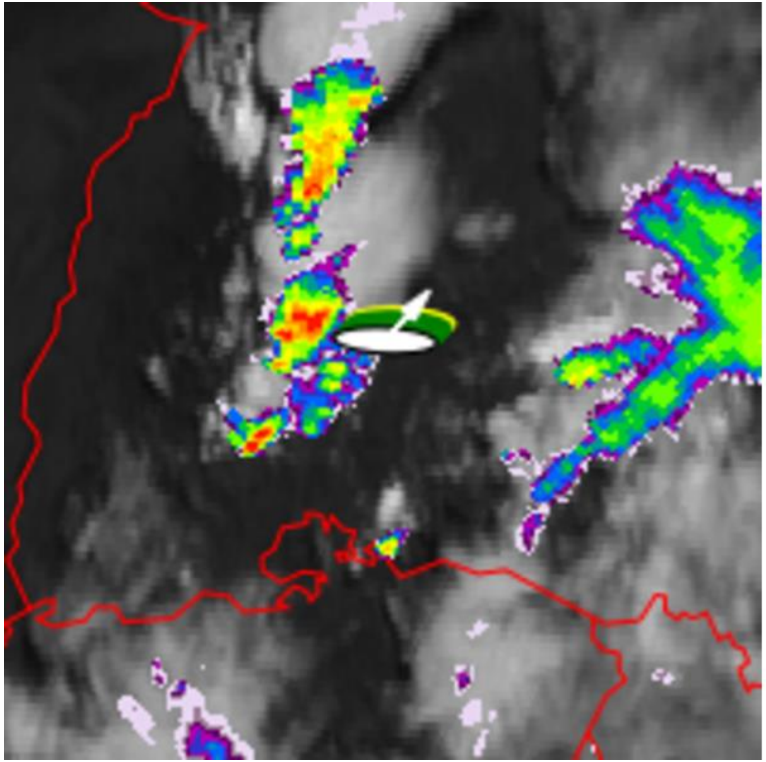


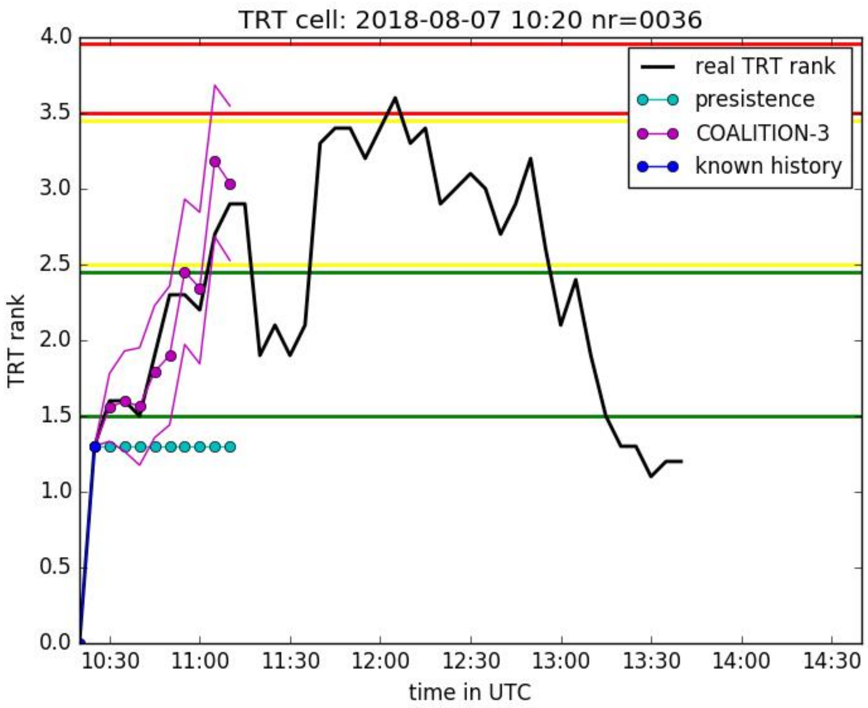
Visualization of the nowcasting results



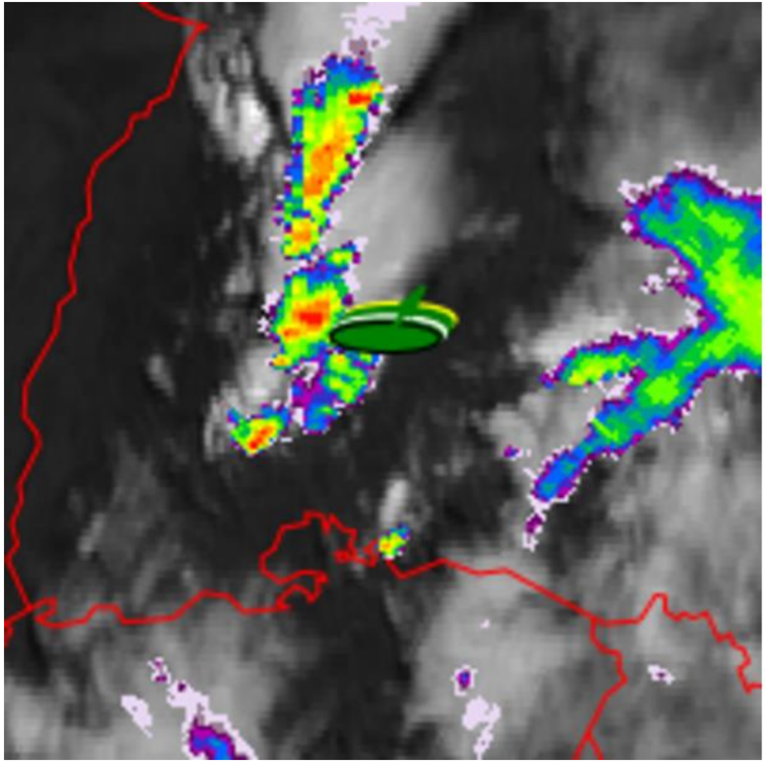


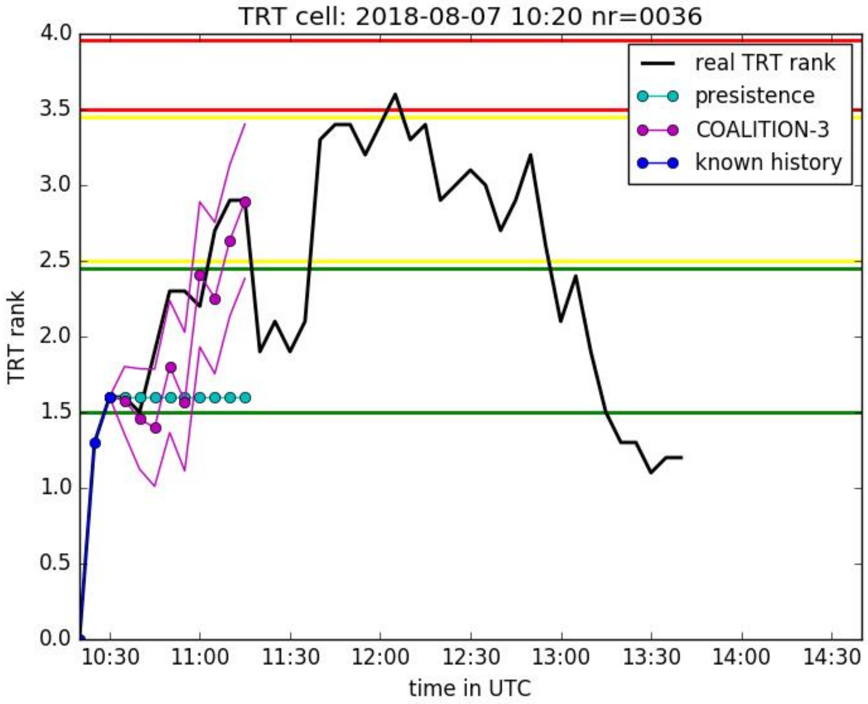
Visualization of the nowcasting results



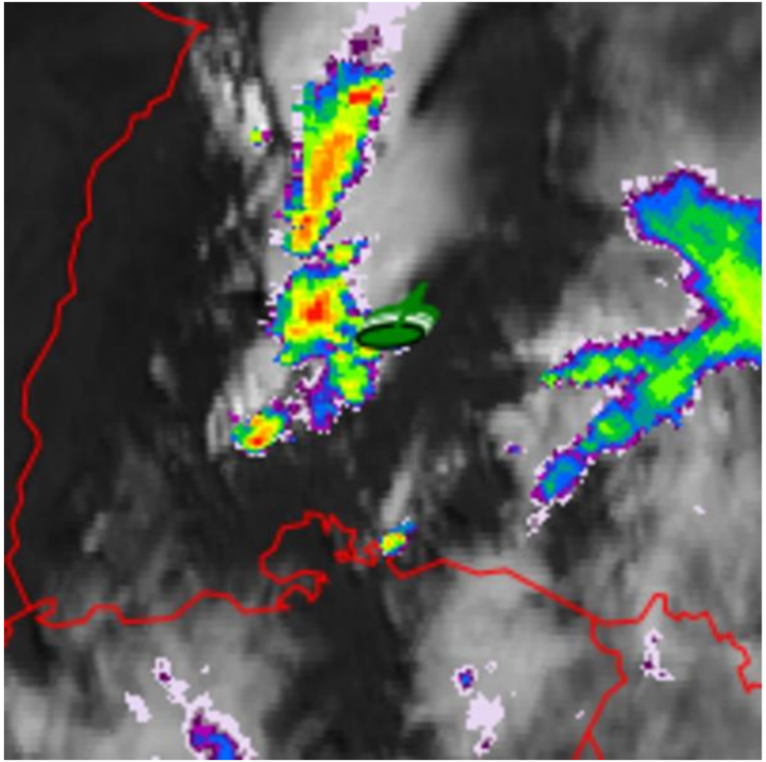


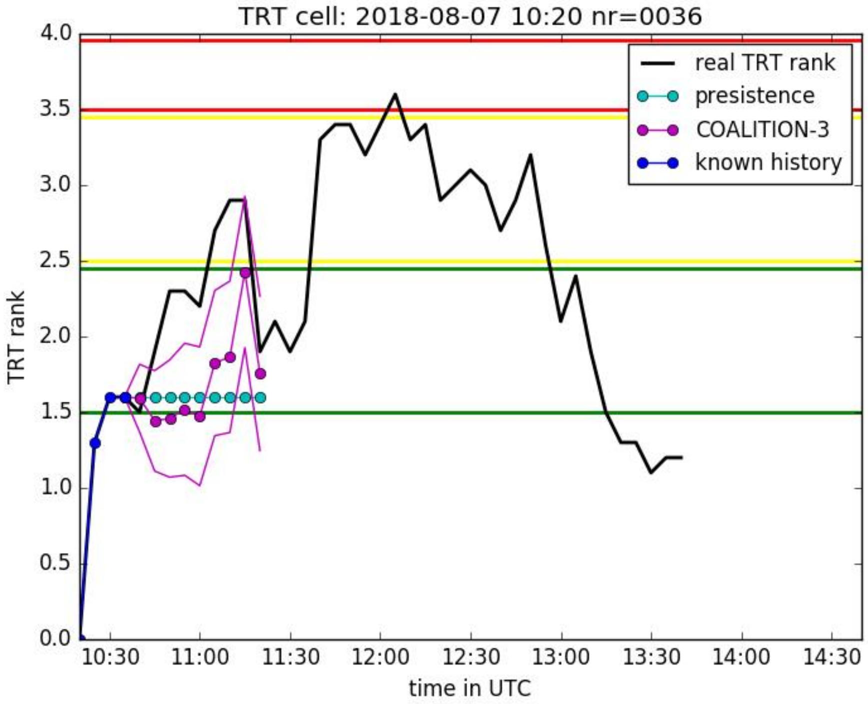
Visualization of the nowcasting results



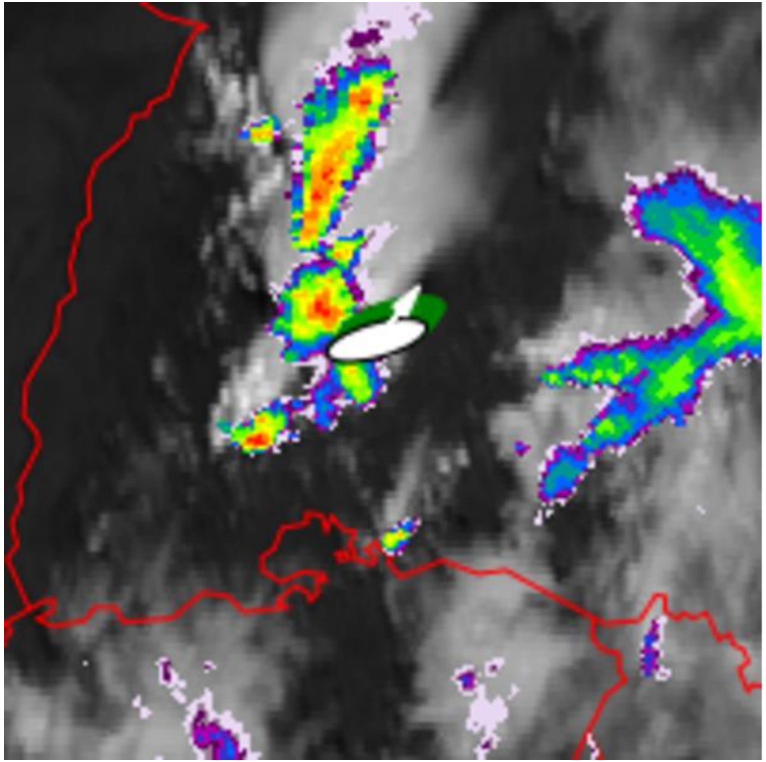


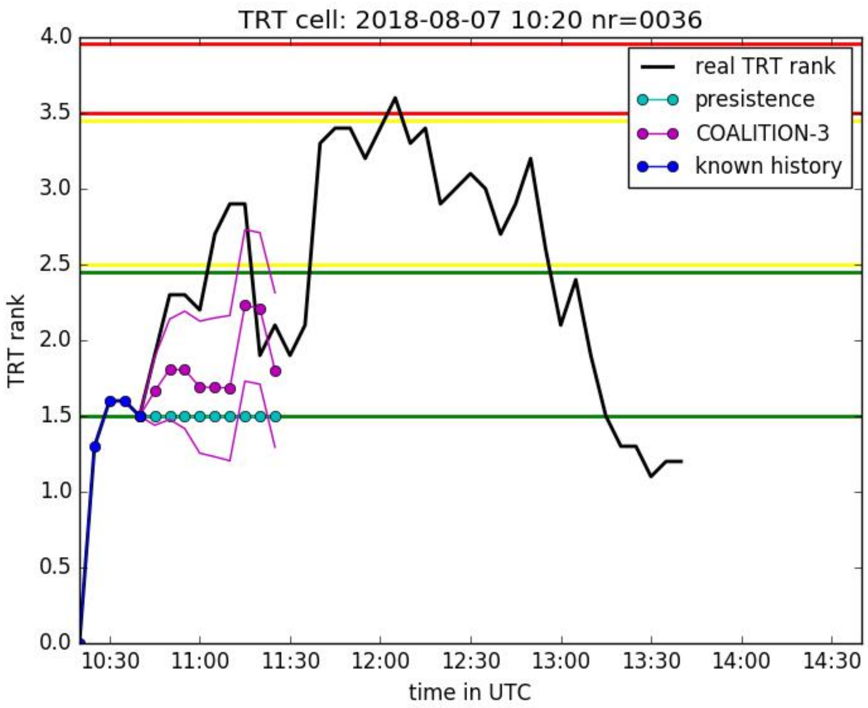
Visualization of the nowcasting results



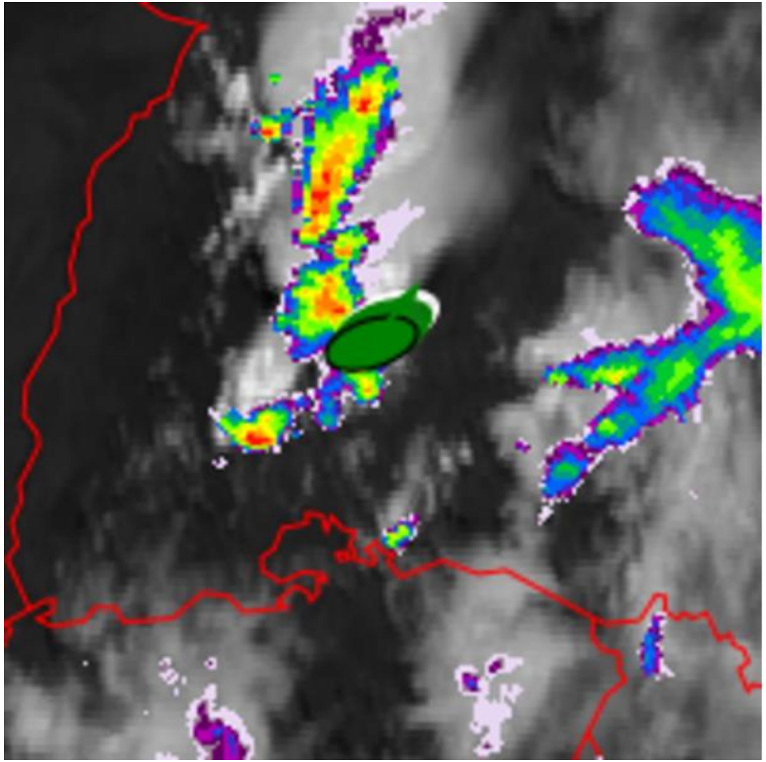


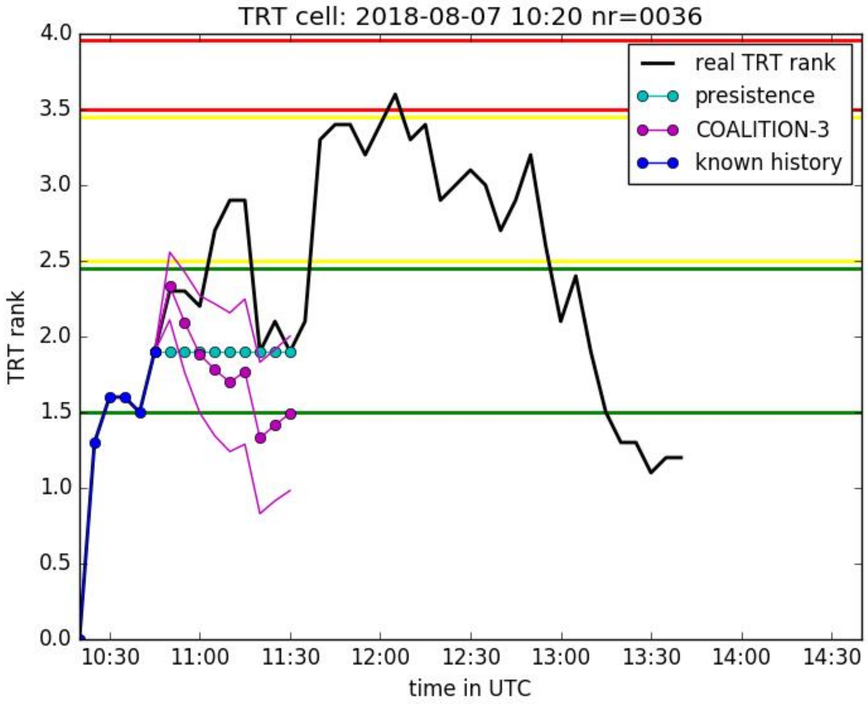
Visualization of the nowcasting results



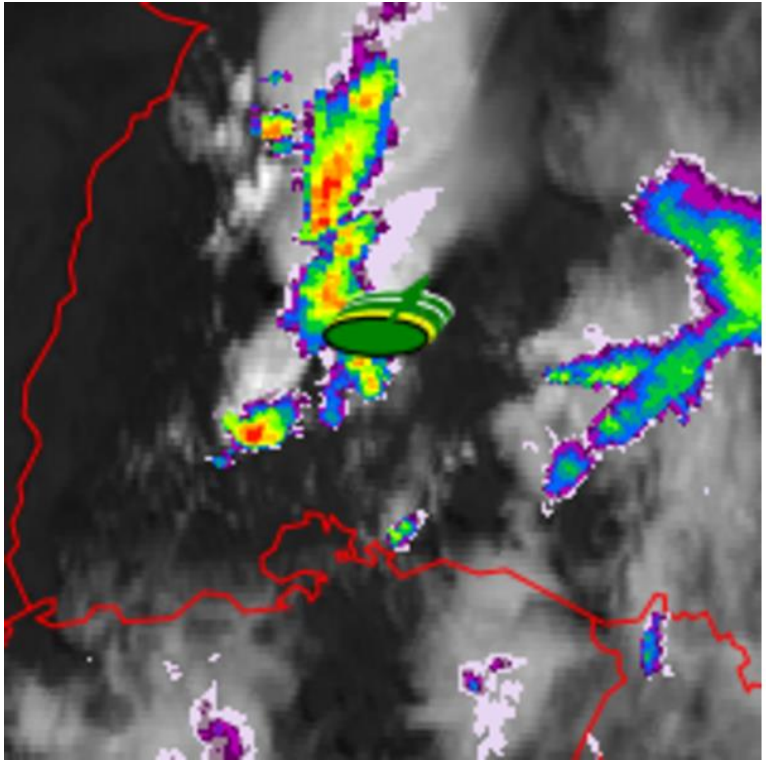


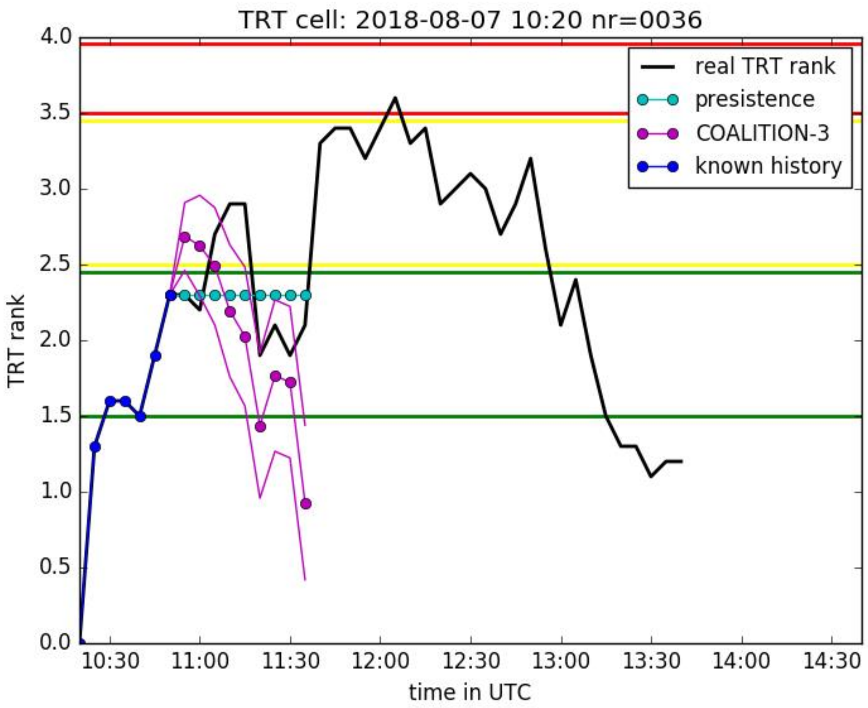
Visualization of the nowcasting results



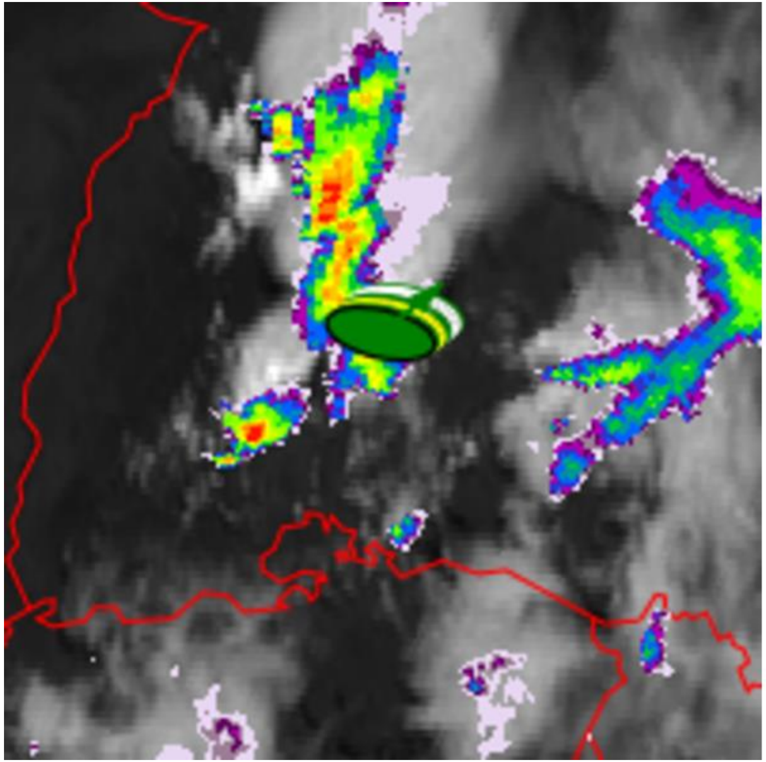


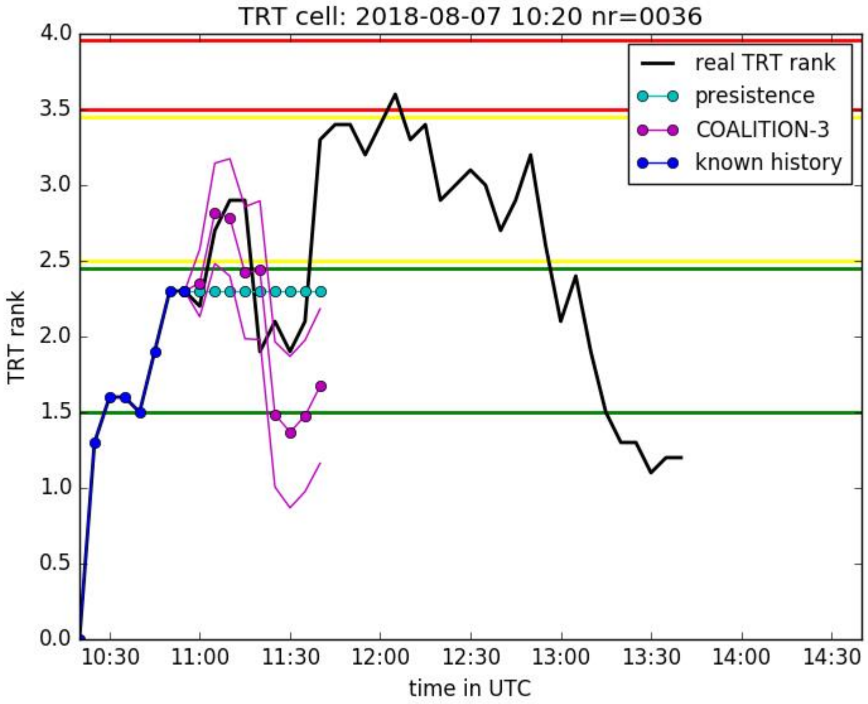
Visualization of the nowcasting results



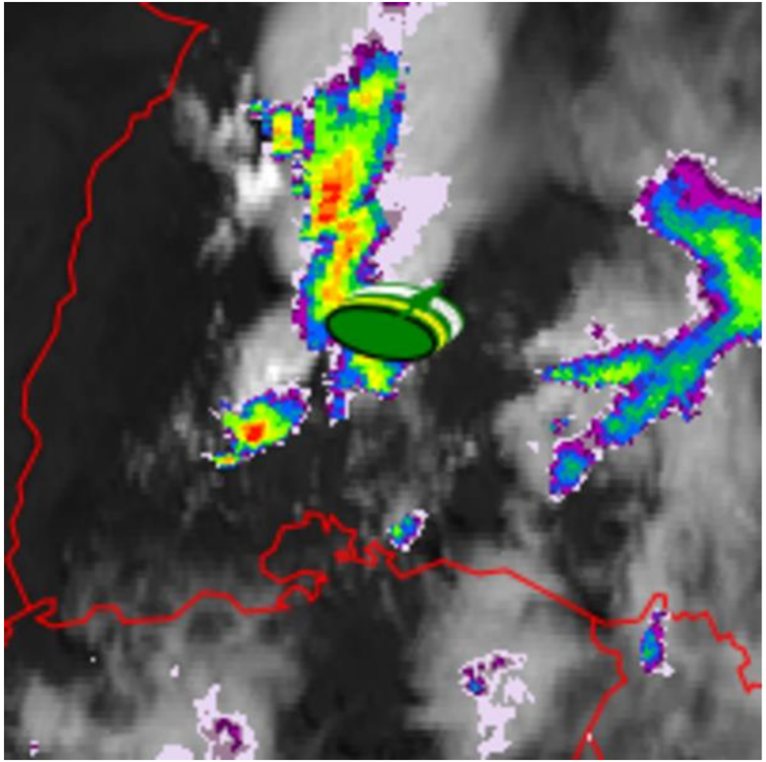


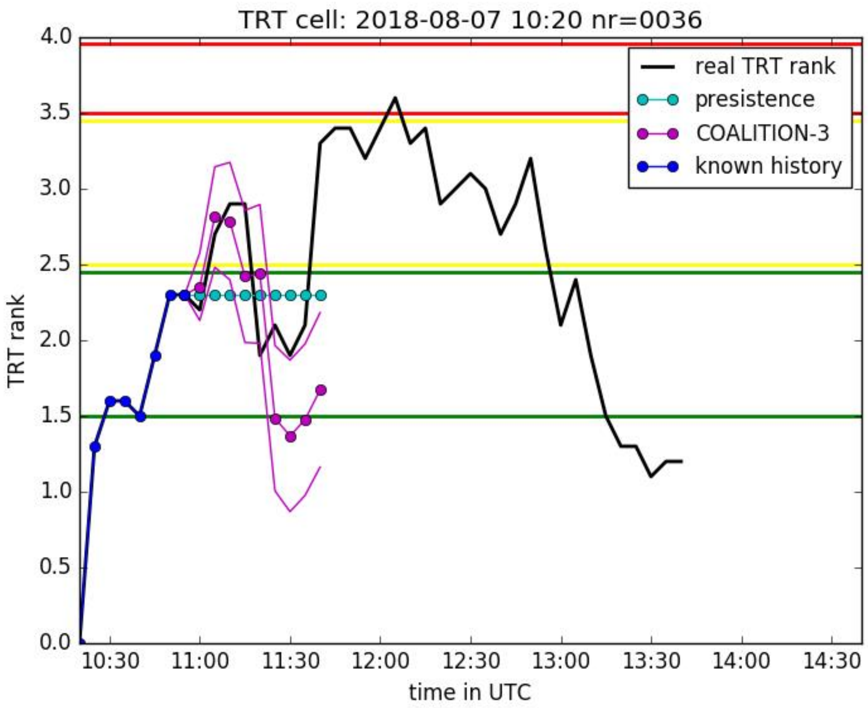
Visualization of the nowcasting results



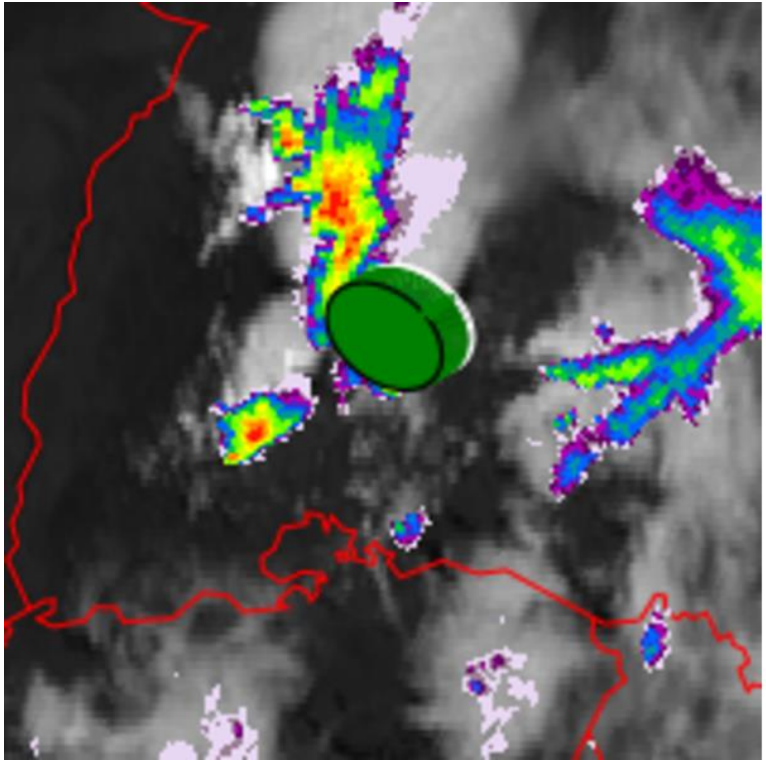


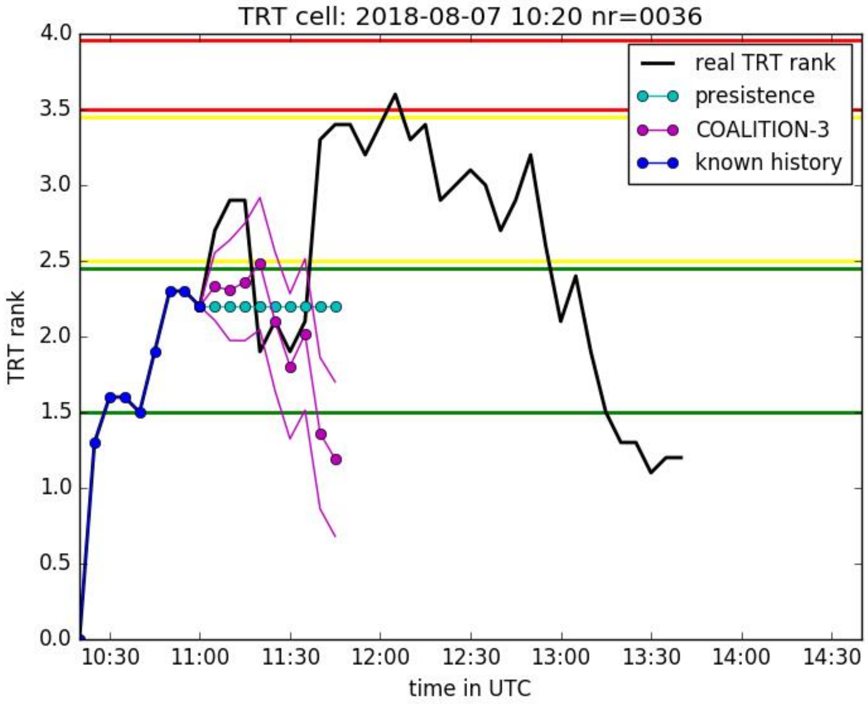
Visualization of the nowcasting results



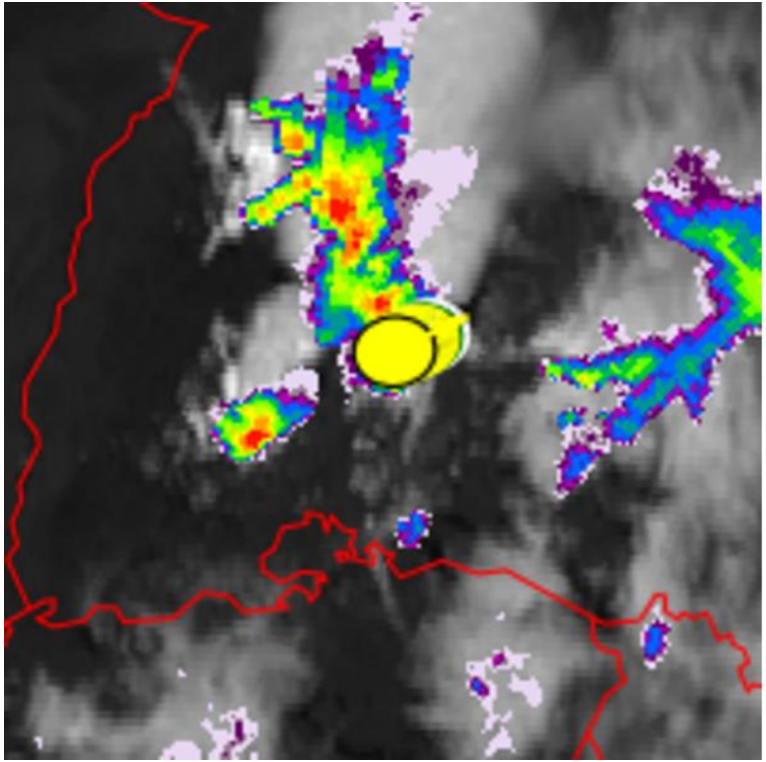


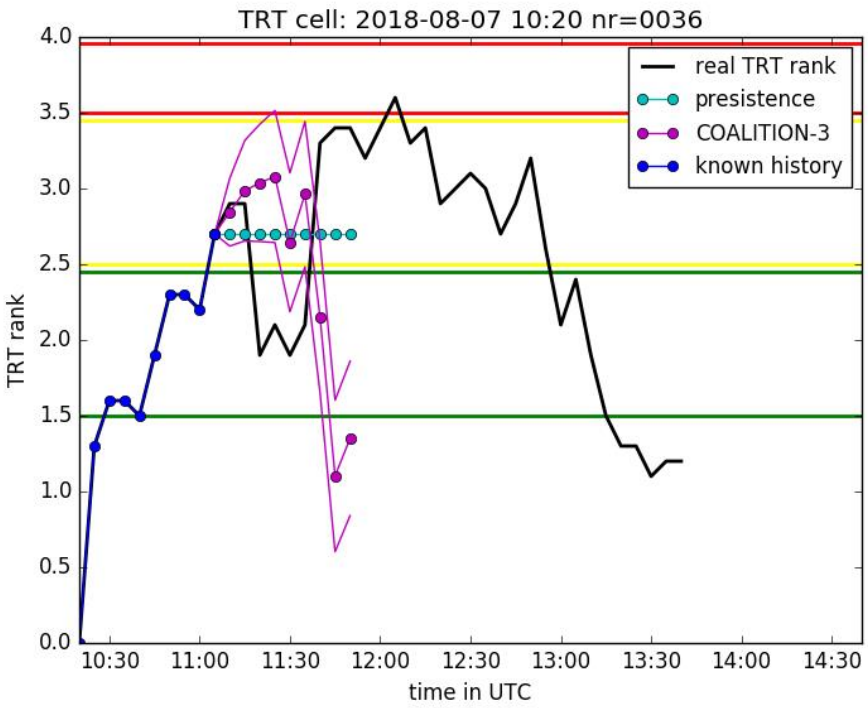
Visualization of the nowcasting results



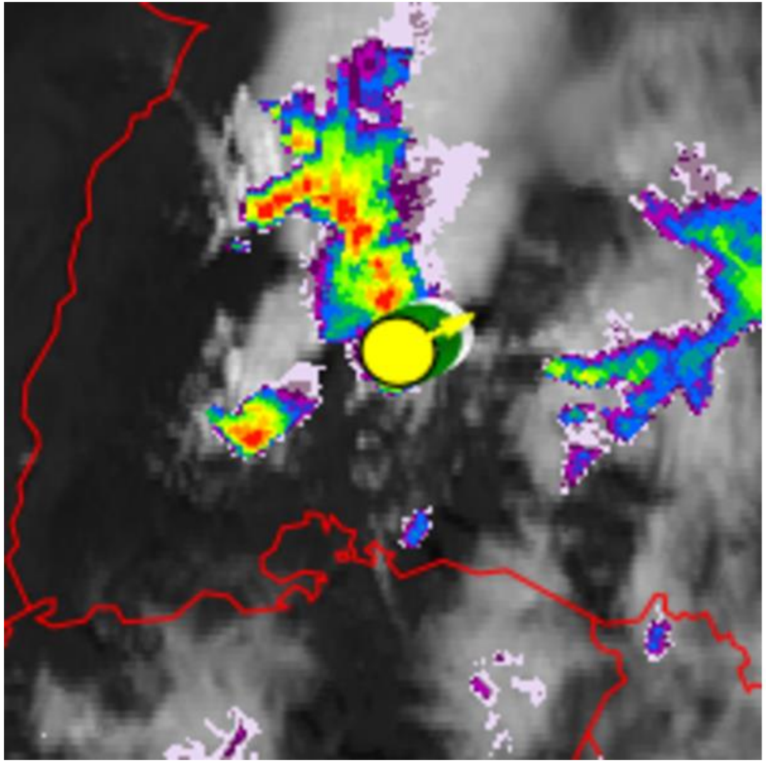


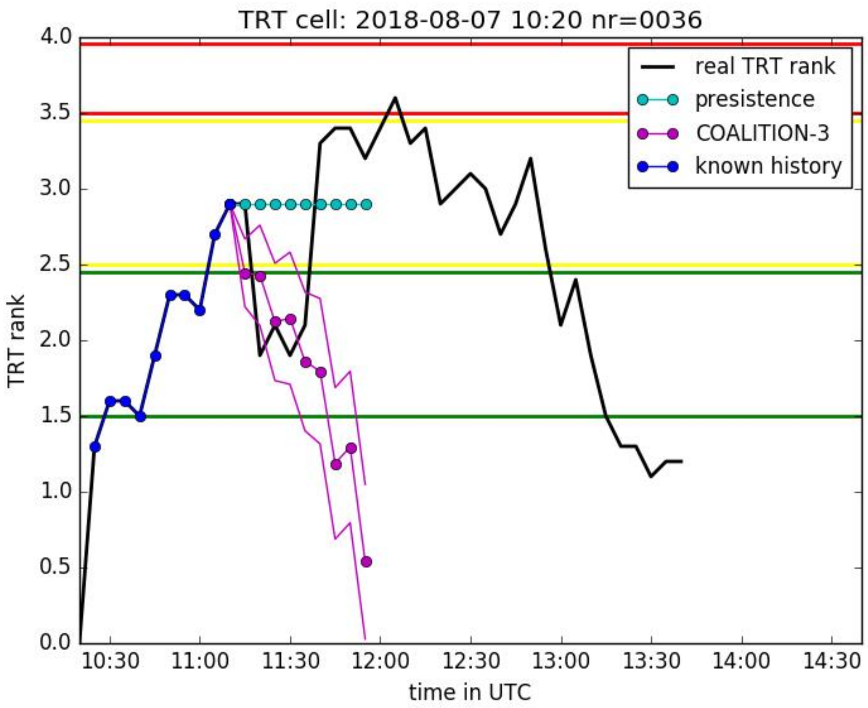
Visualization of the nowcasting results



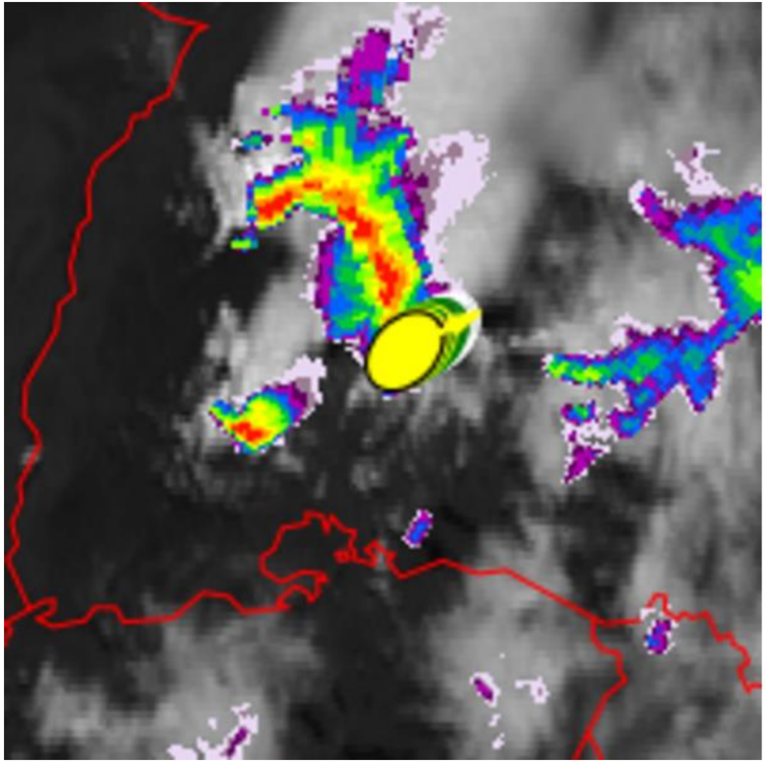


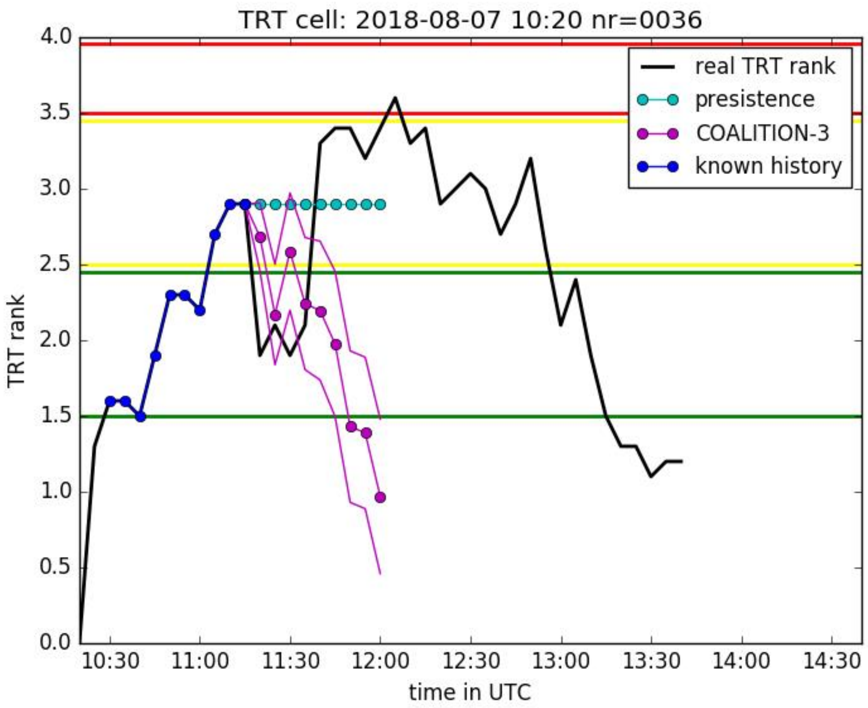
Visualization of the nowcasting results



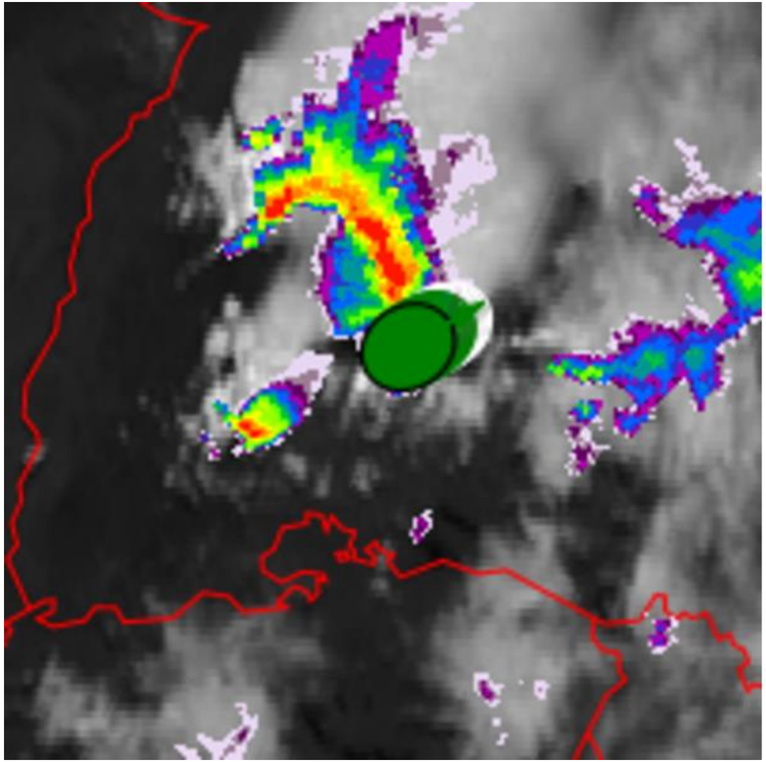


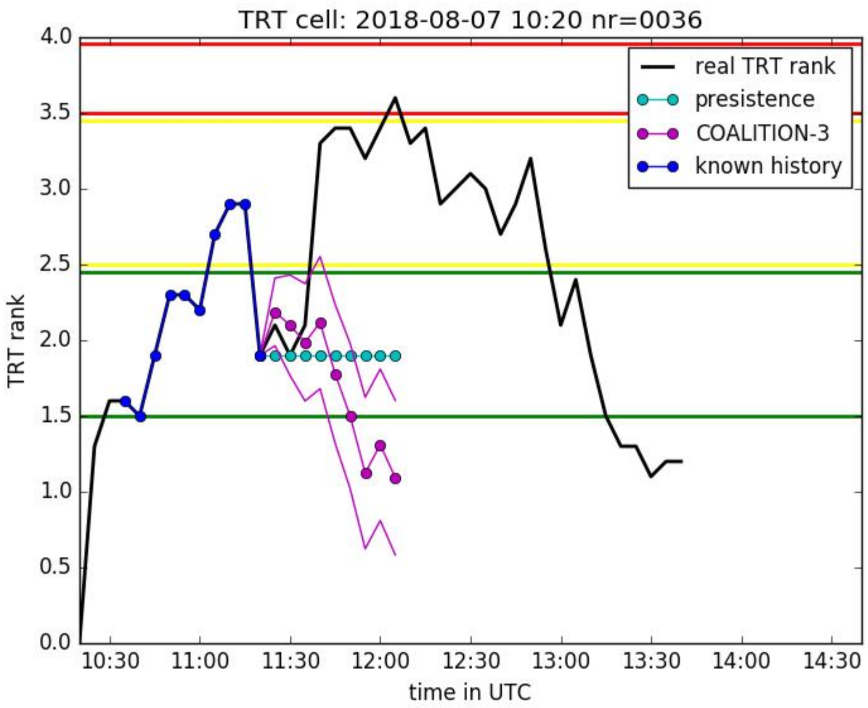
Visualization of the nowcasting results



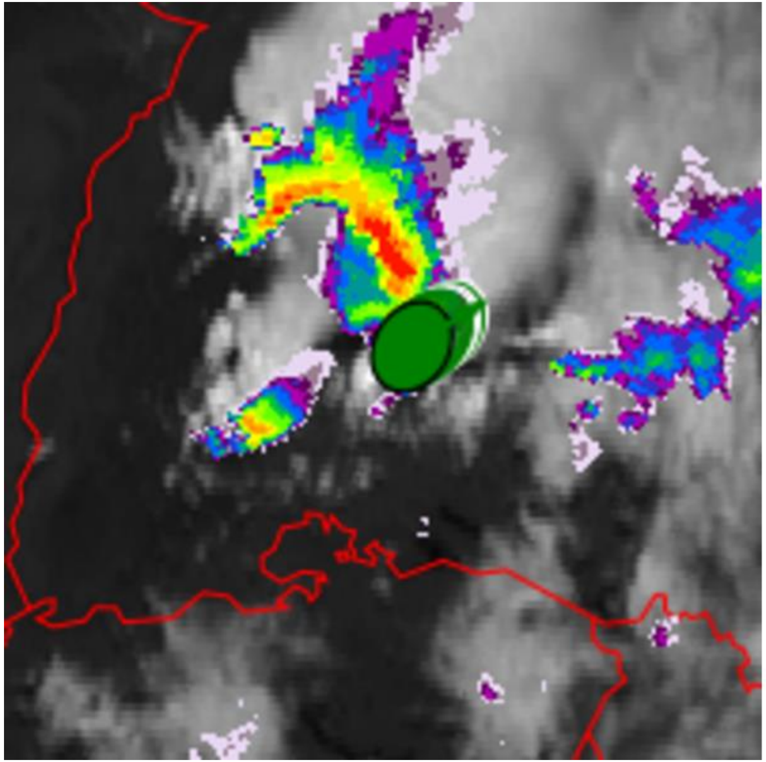


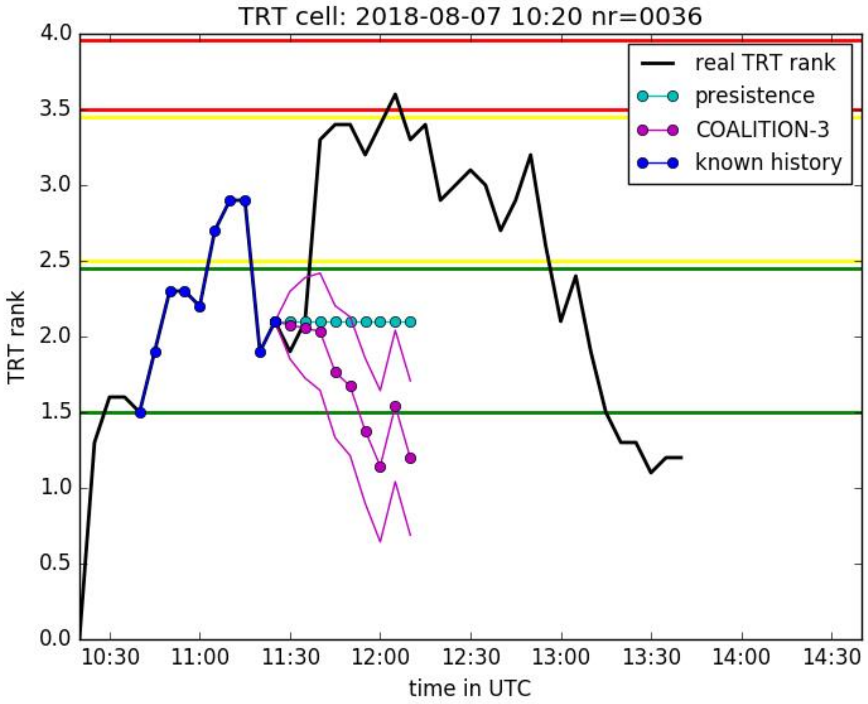
Visualization of the nowcasting results



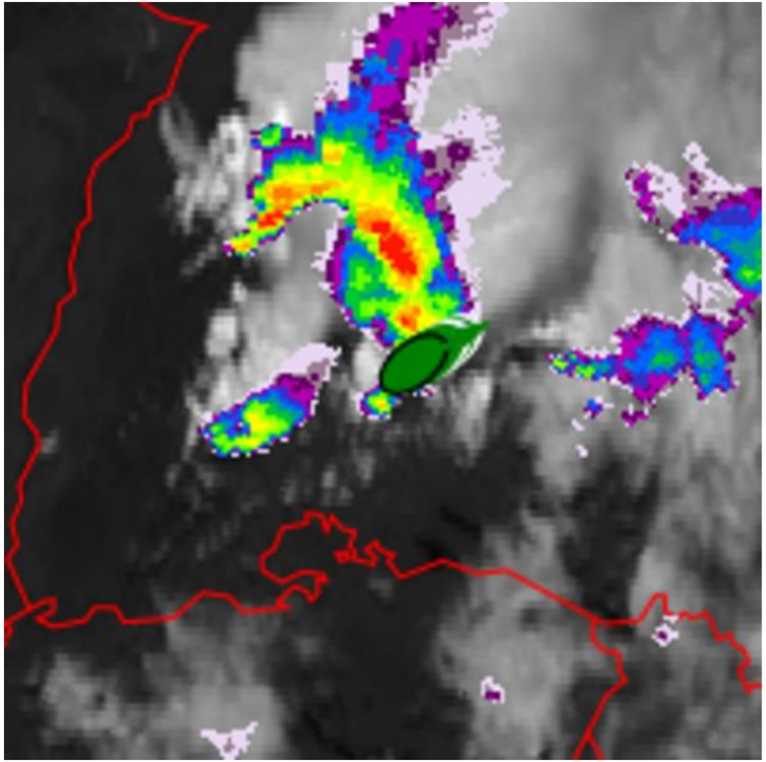


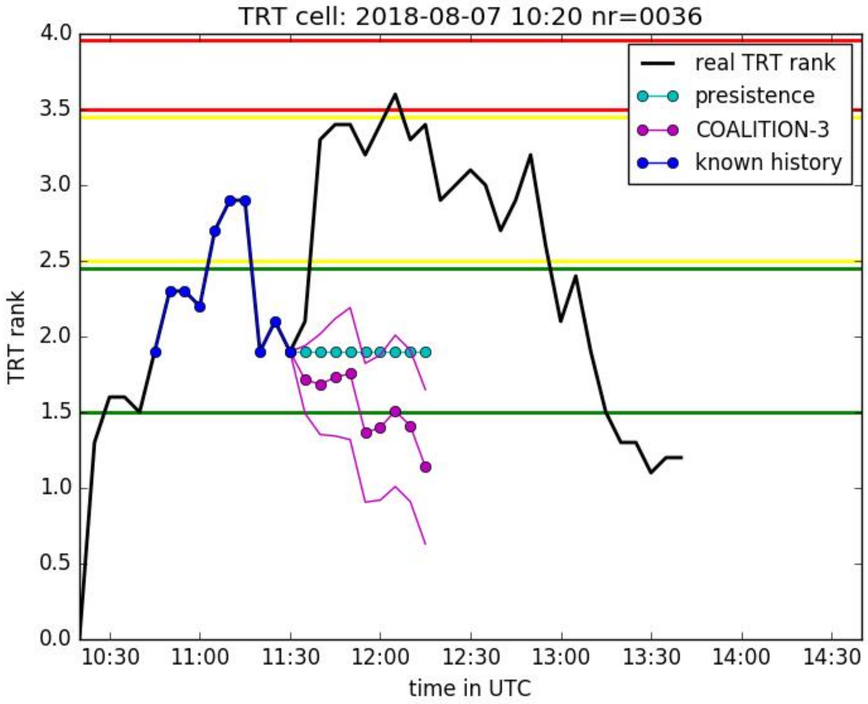
Visualization of the nowcasting results



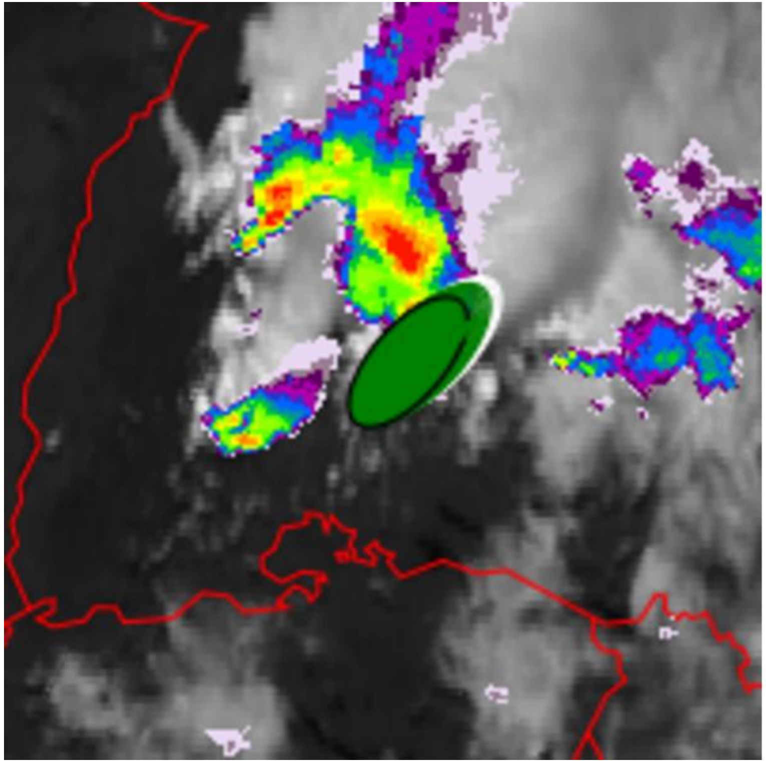


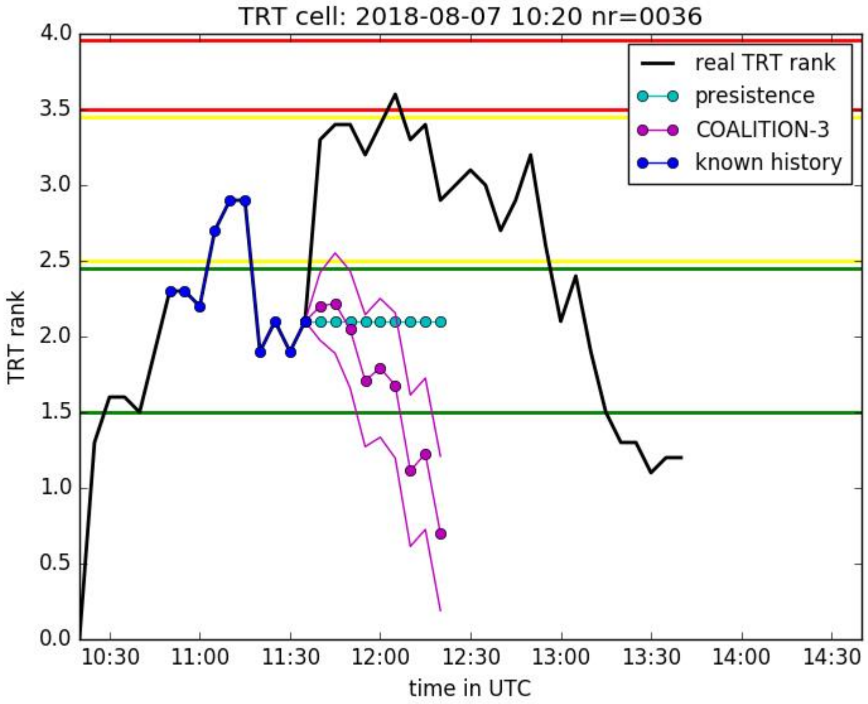
Visualization of the nowcasting results



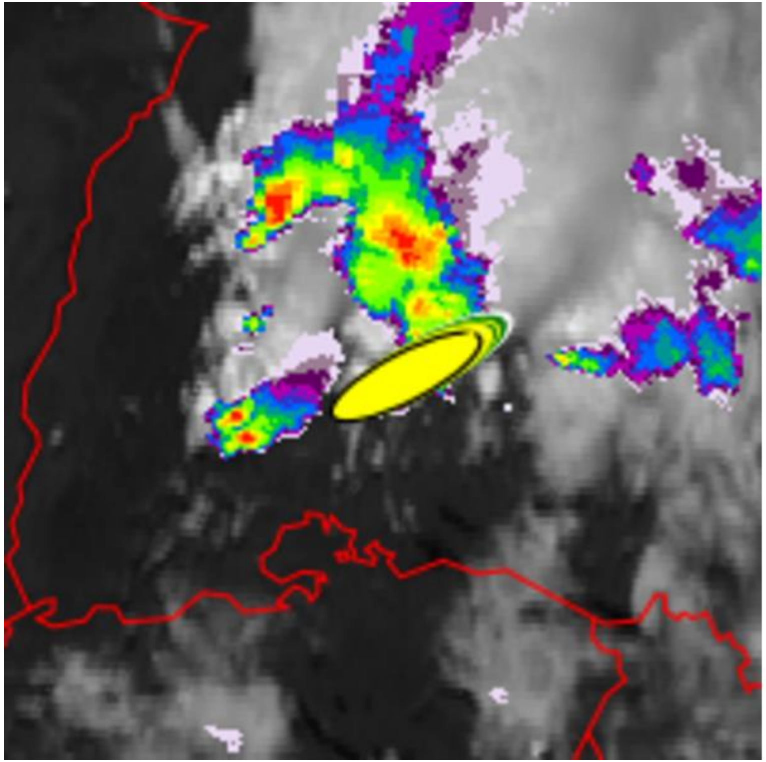


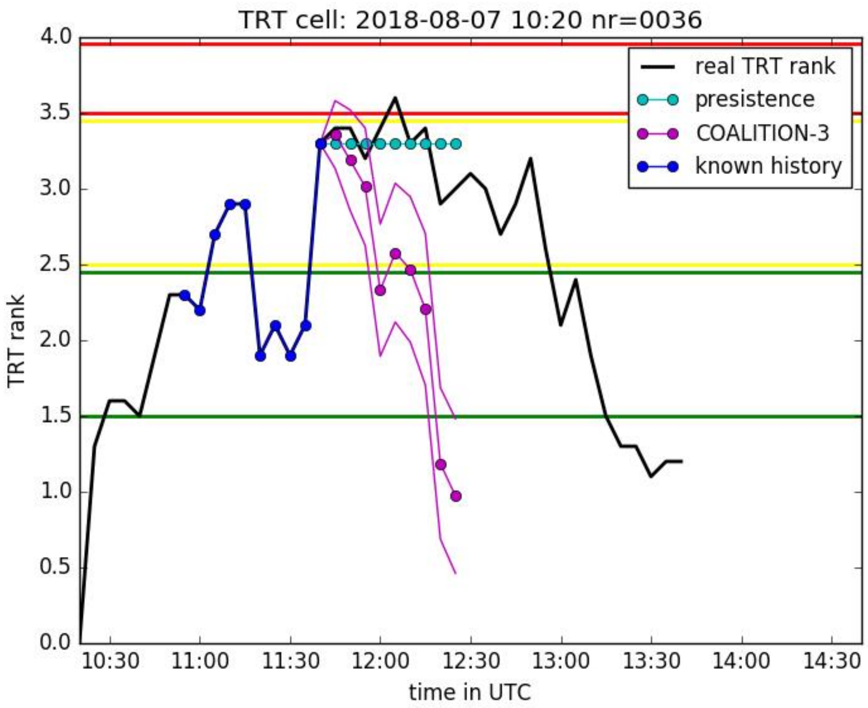
Visualization of the nowcasting results



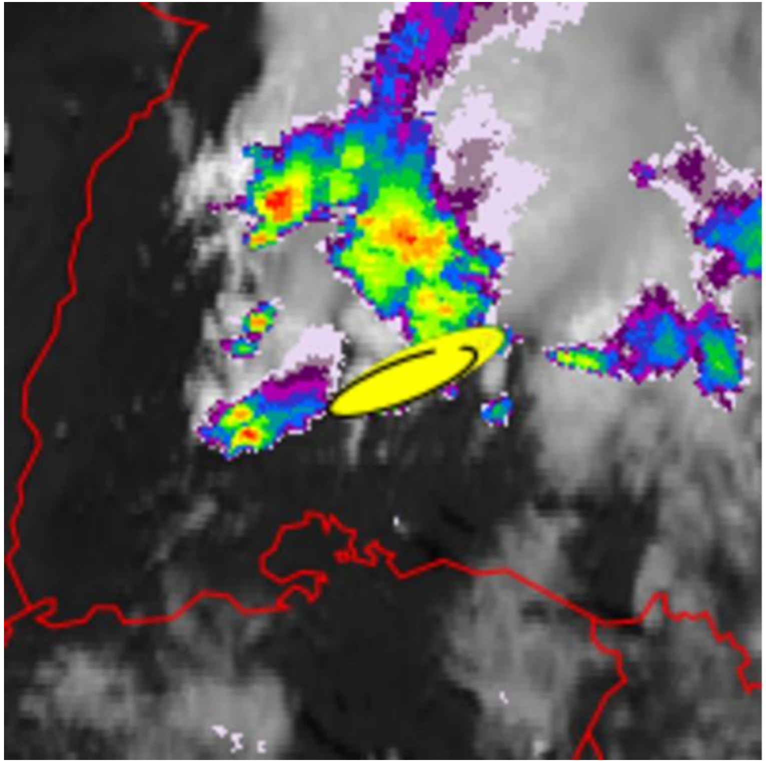


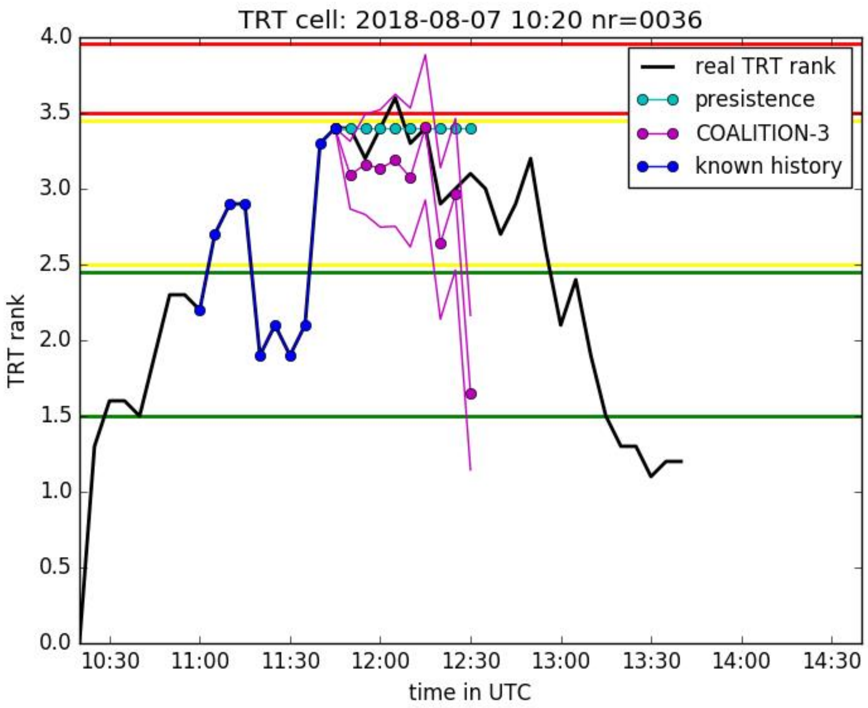
Visualization of the nowcasting results



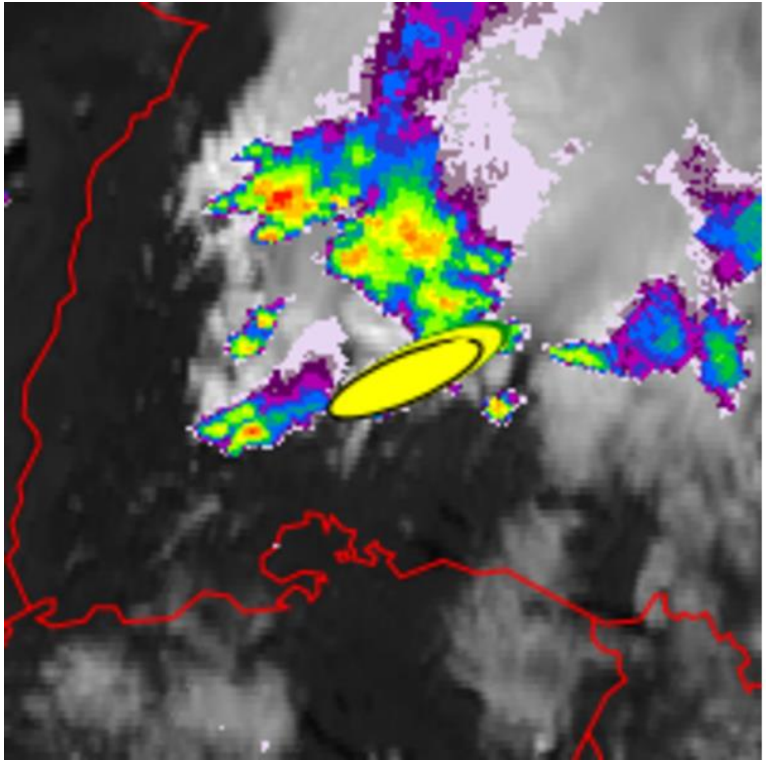


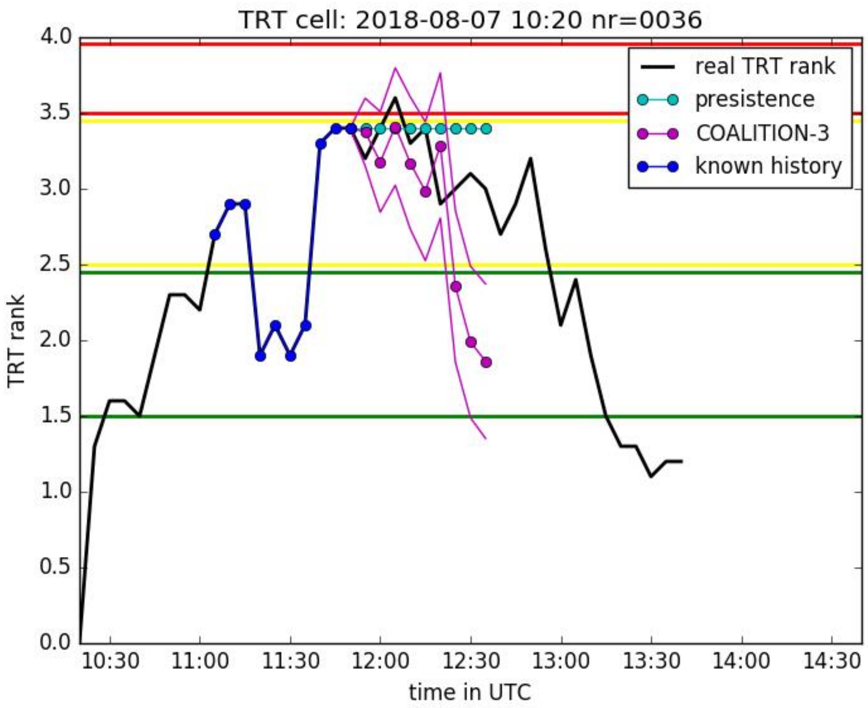
Visualization of the nowcasting results



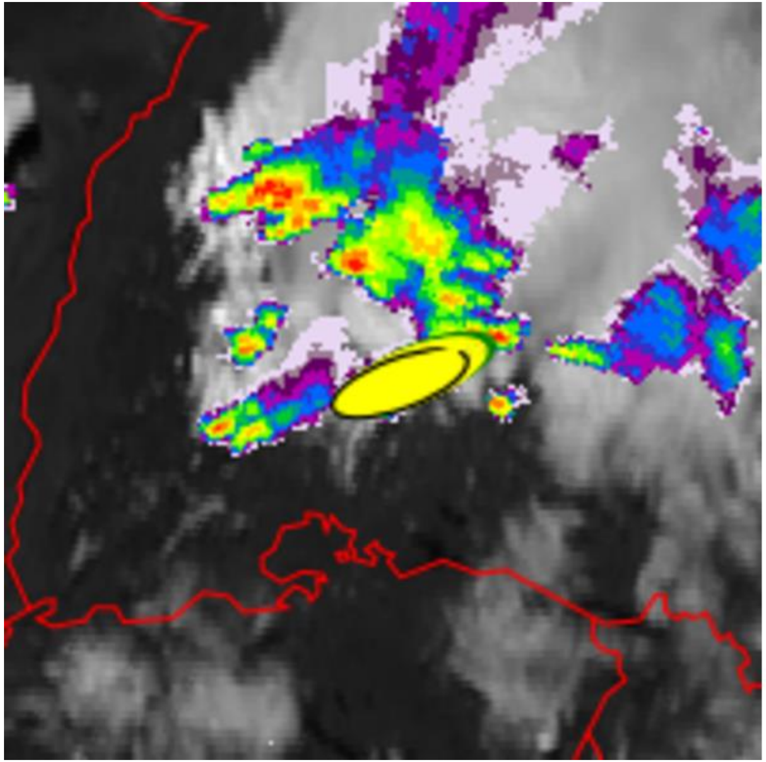


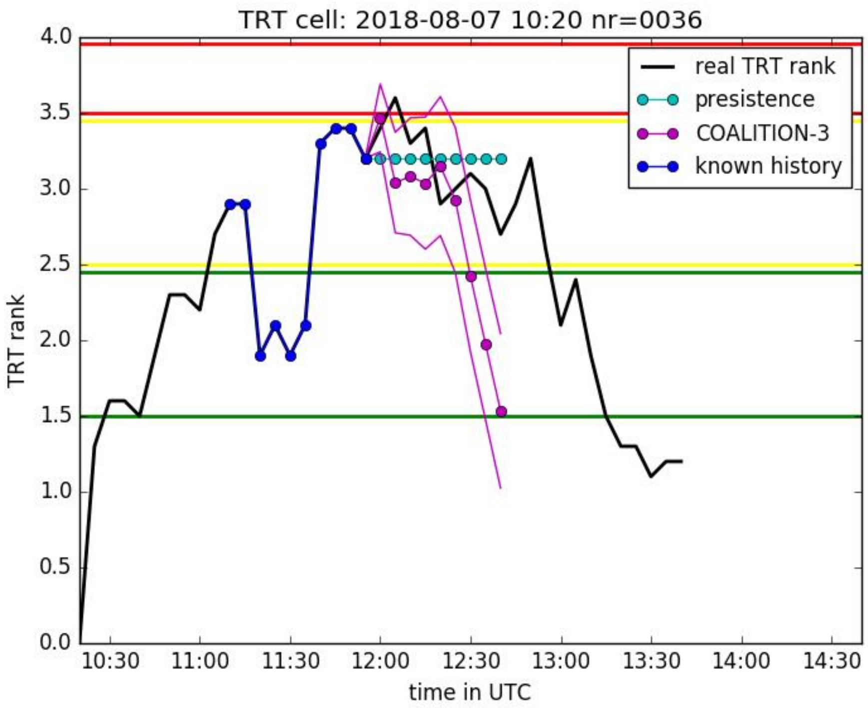
Visualization of the nowcasting results



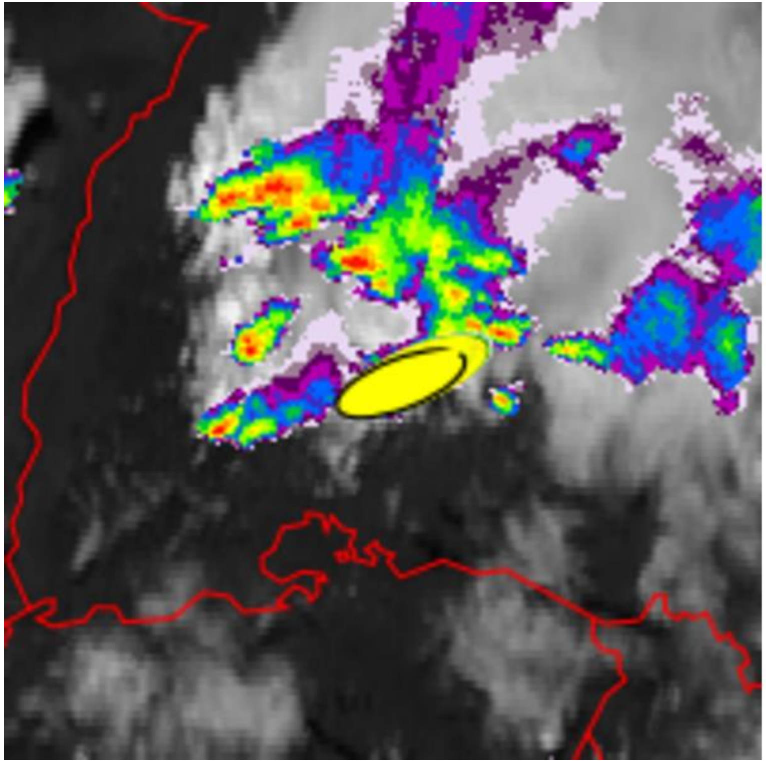


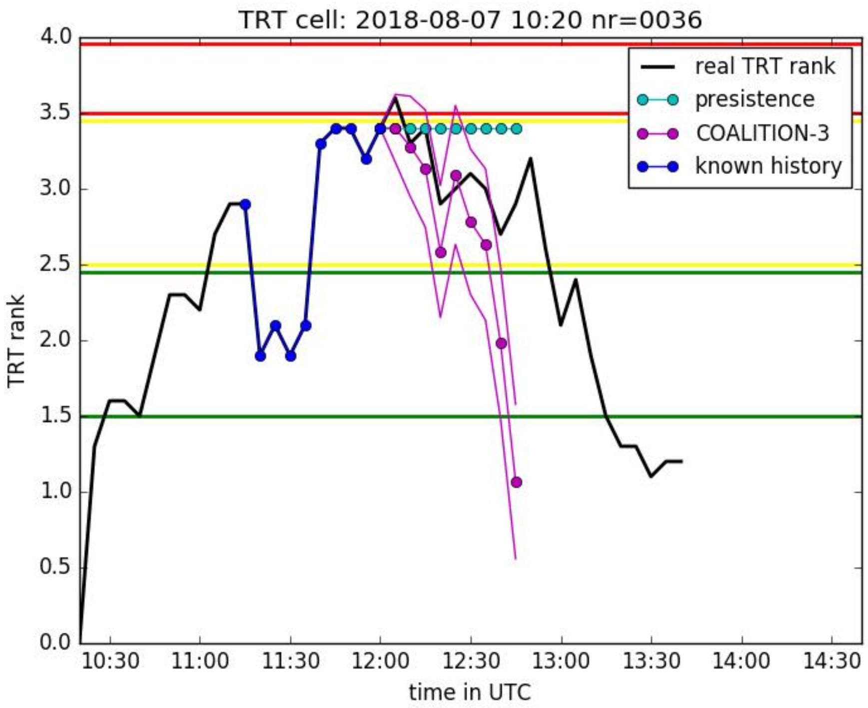
Visualization of the nowcasting results



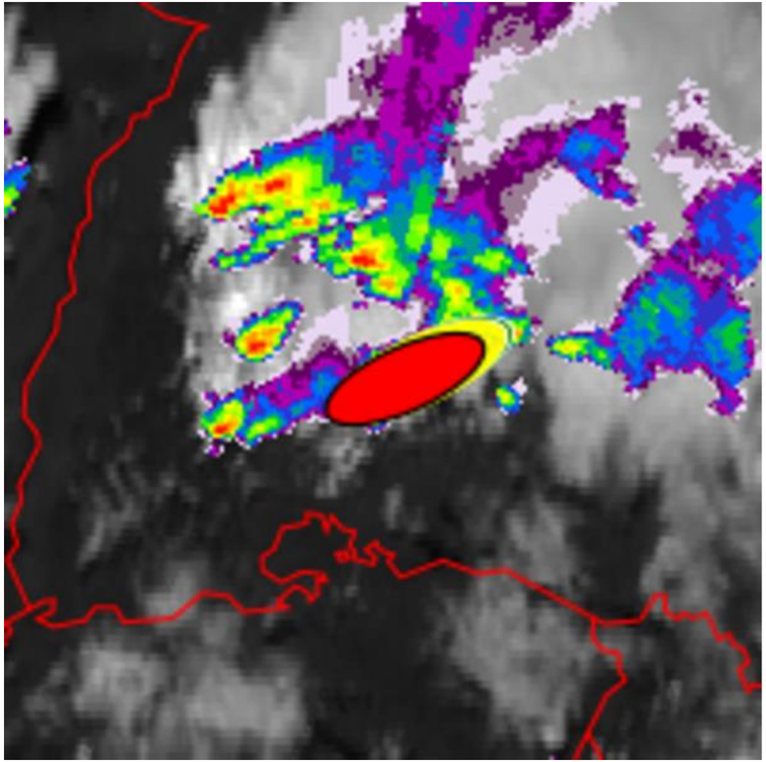


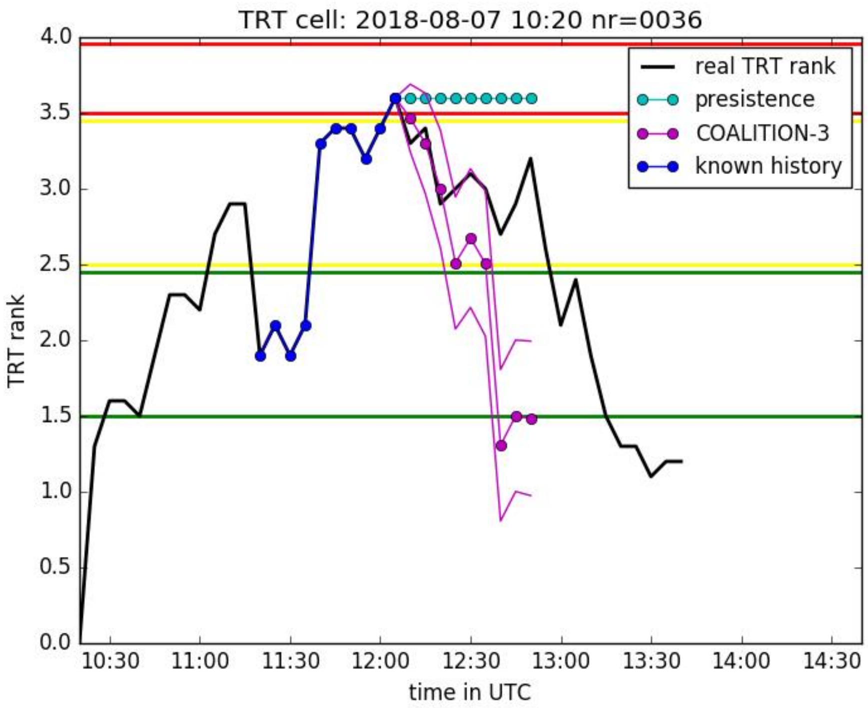
Visualization of the nowcasting results



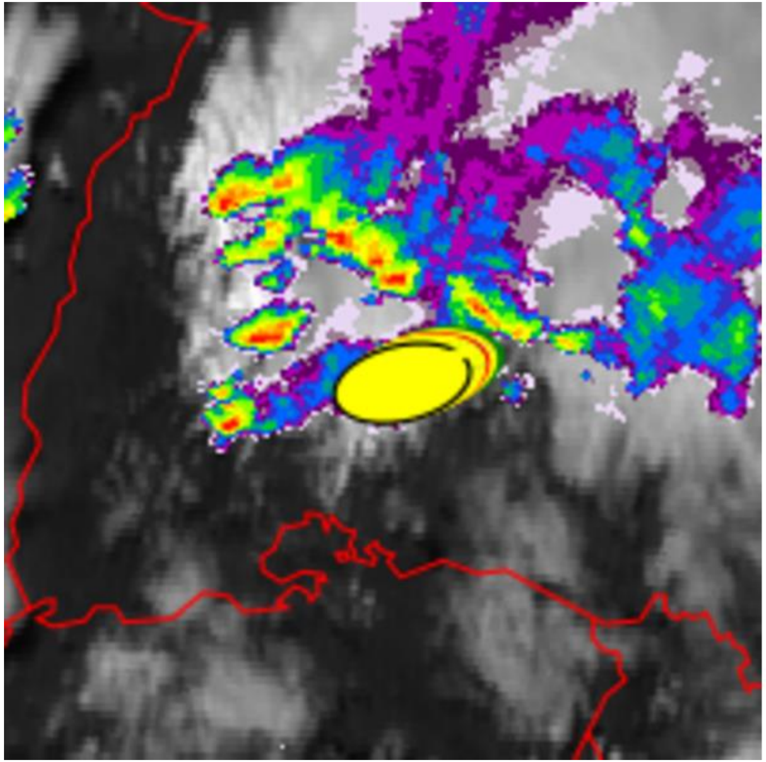


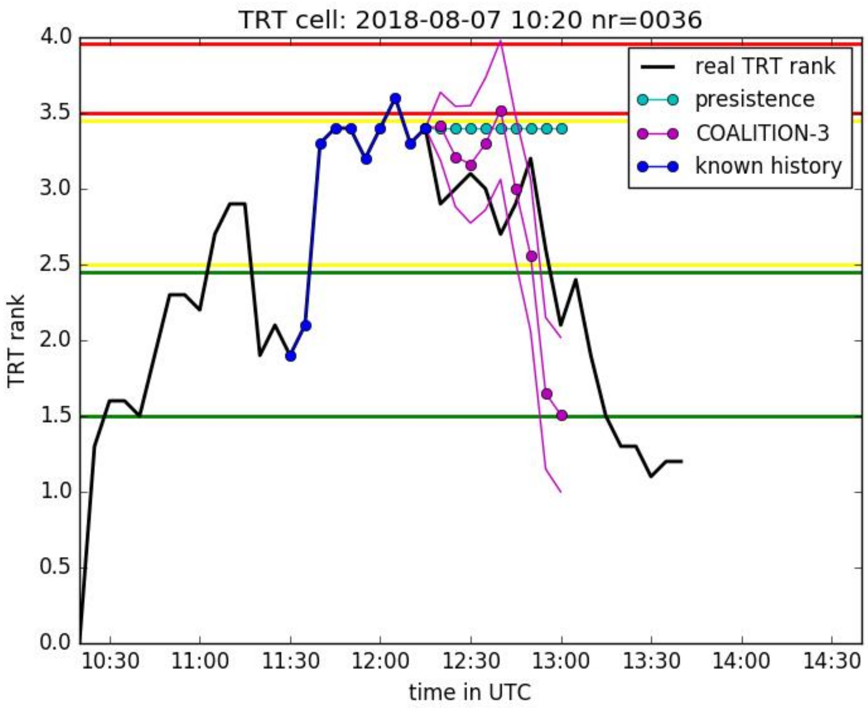
Visualization of the nowcasting results



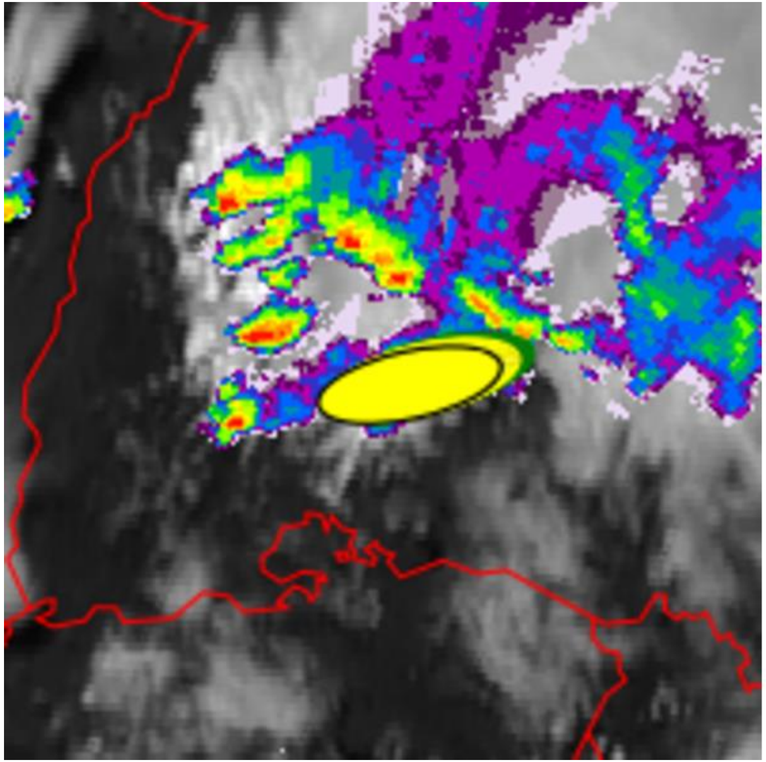


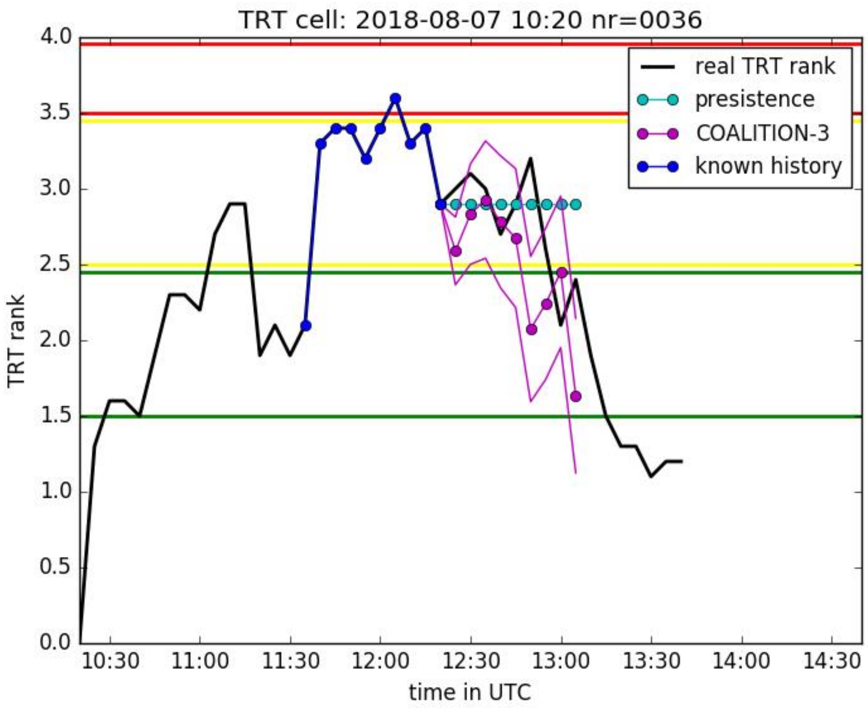
Visualization of the nowcasting results



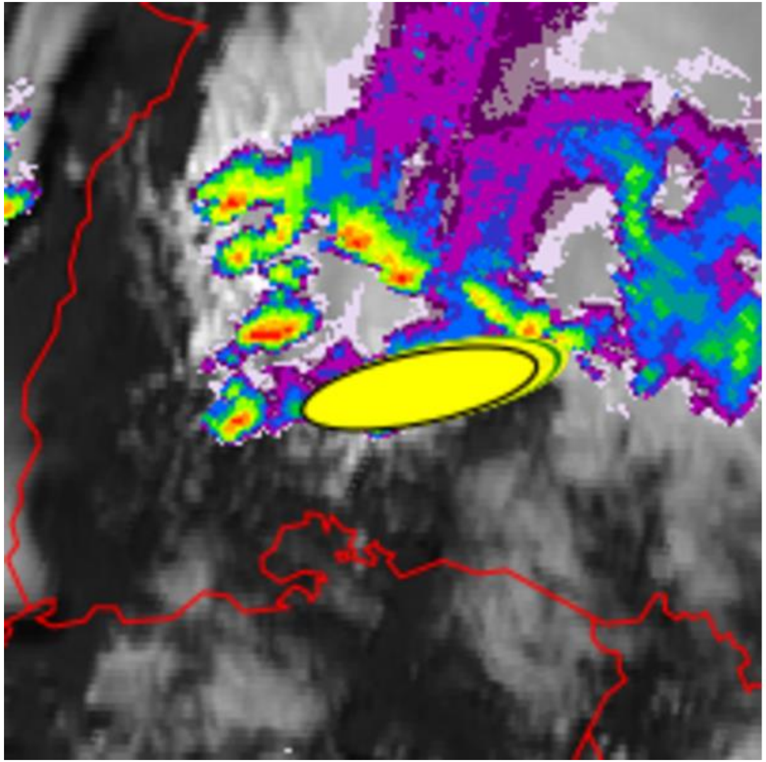


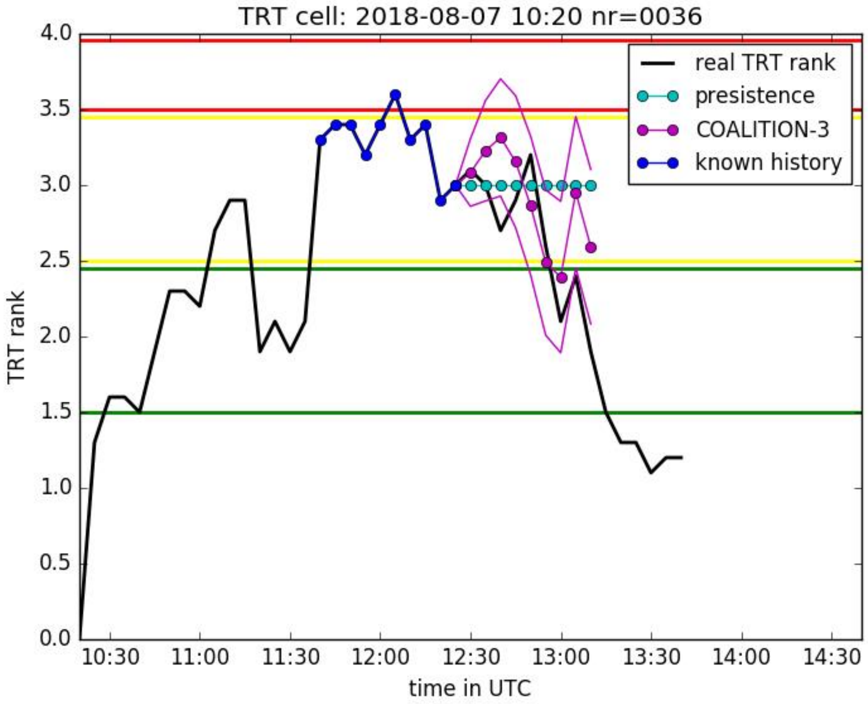
Visualization of the nowcasting results



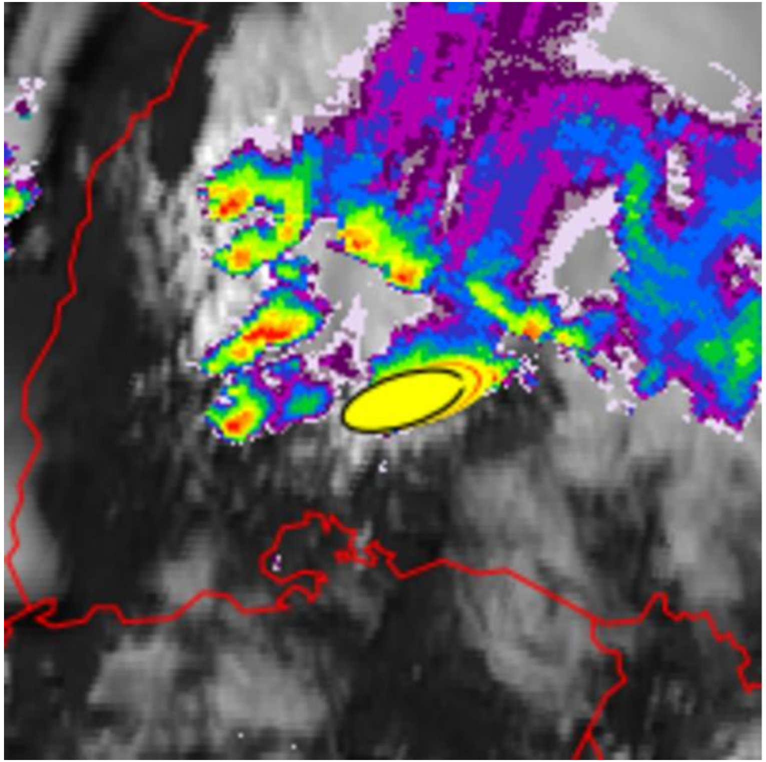


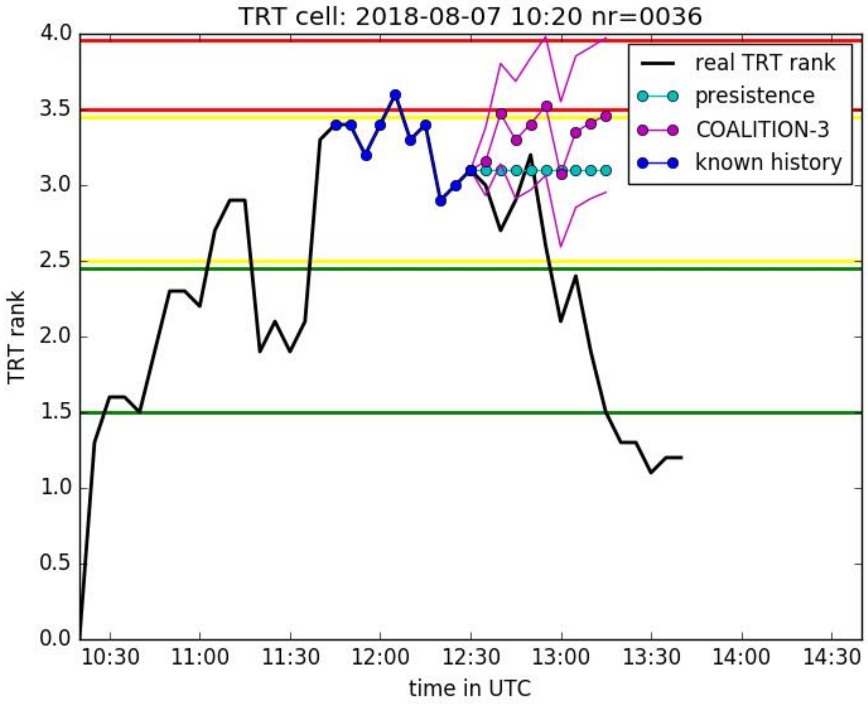
Visualization of the nowcasting results



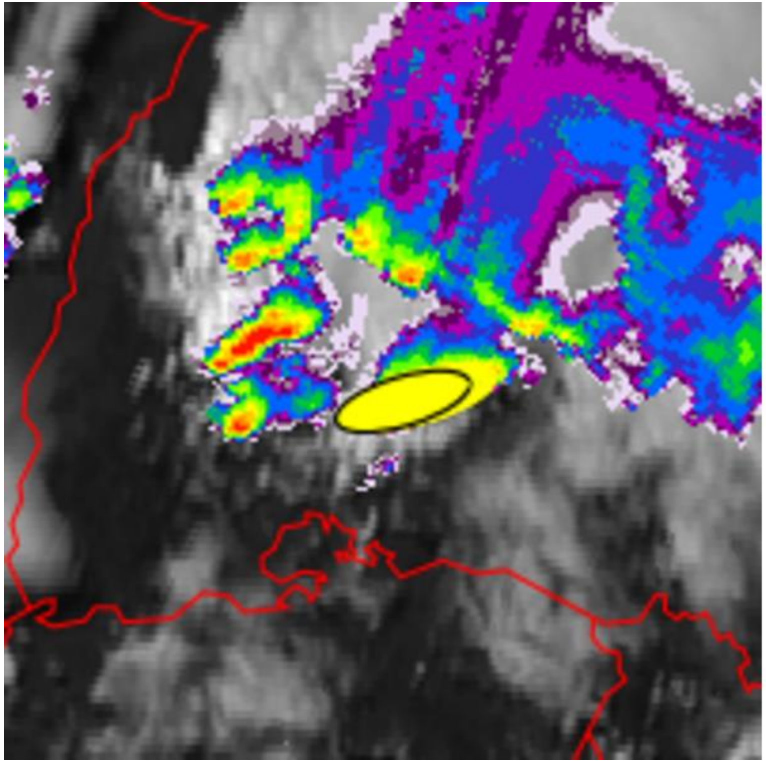


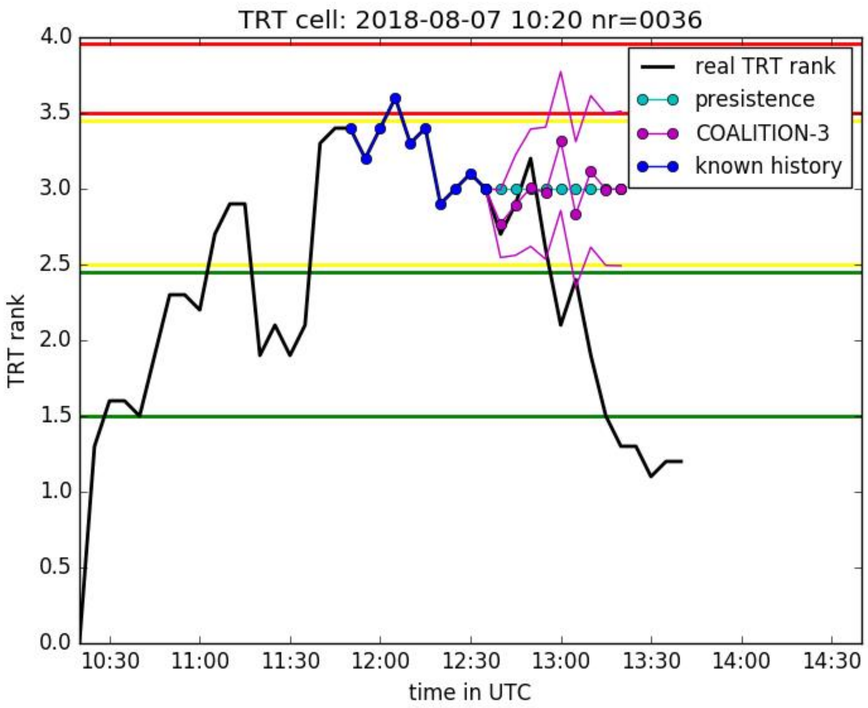
Visualization of the nowcasting results



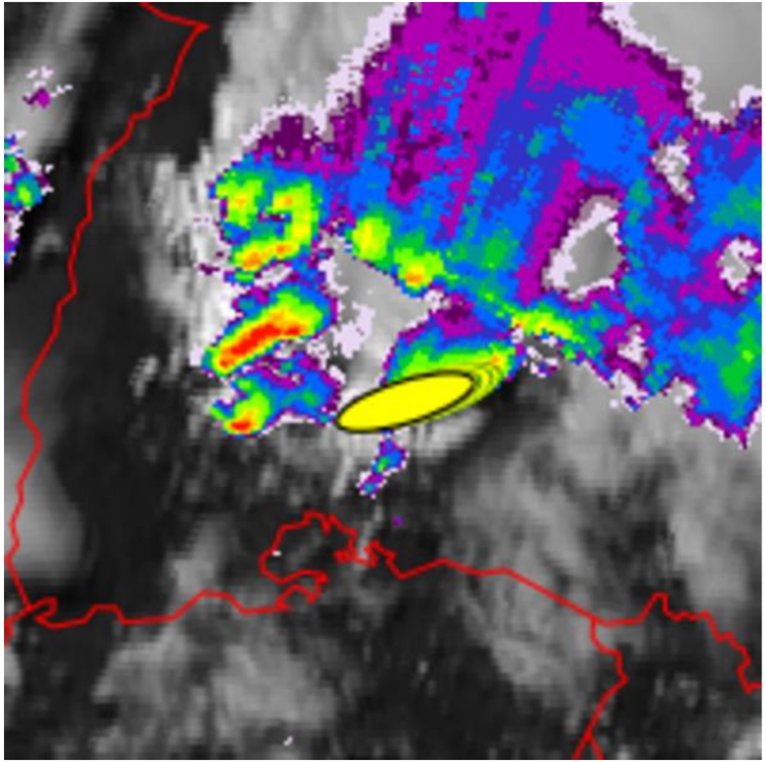


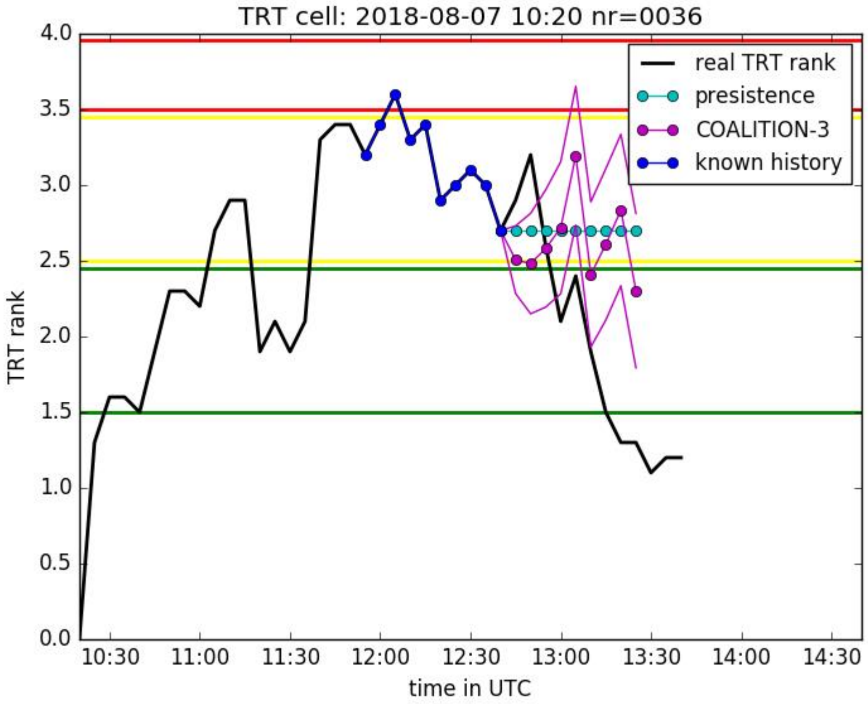
Visualization of the nowcasting results



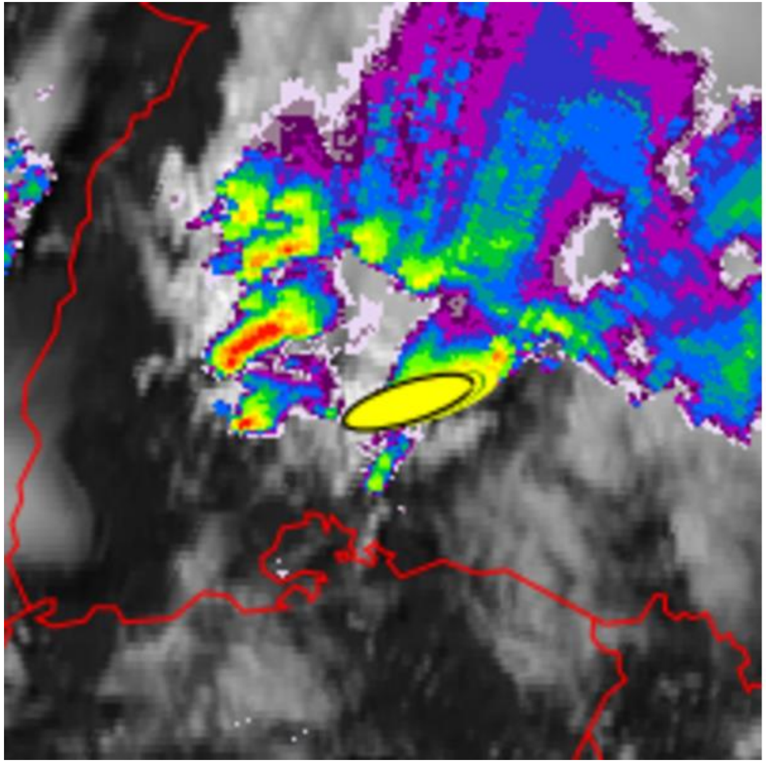


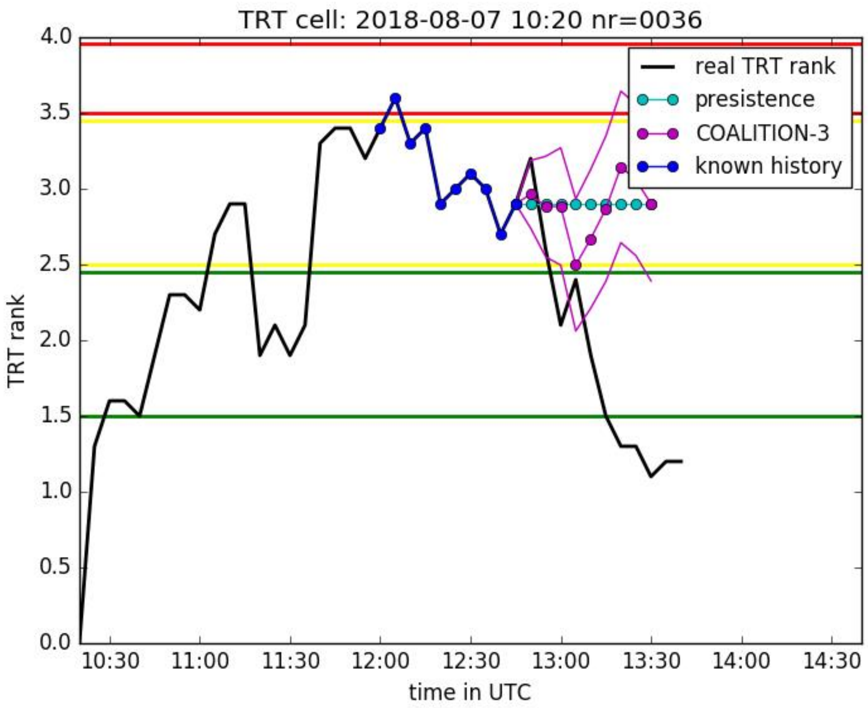
Visualization of the nowcasting results



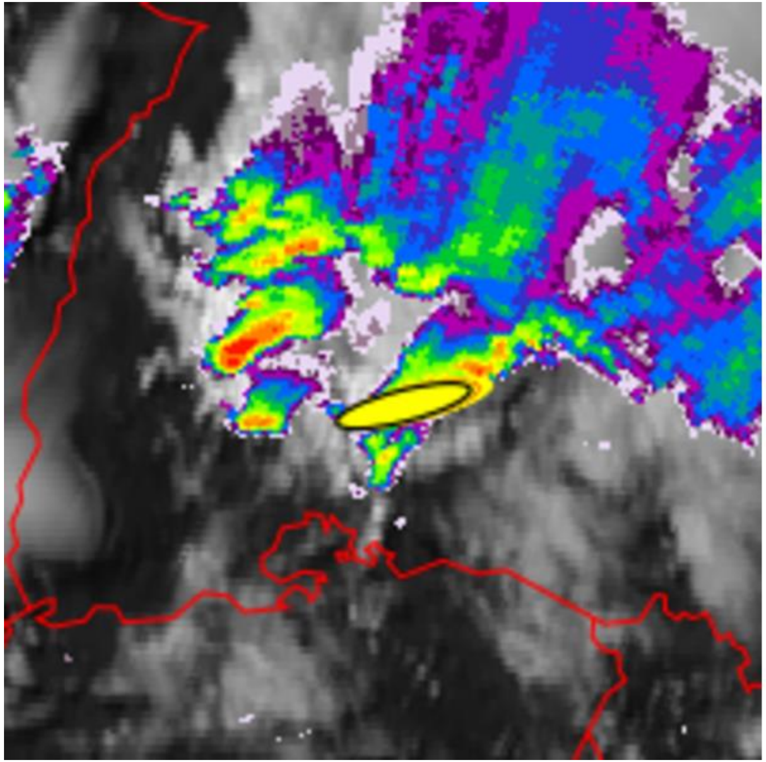


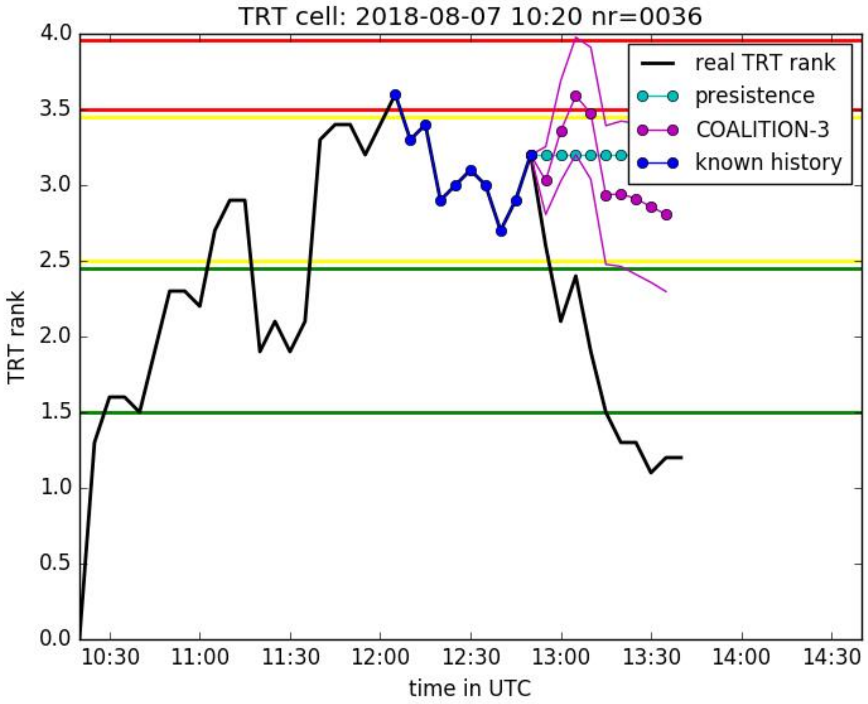
Visualization of the nowcasting results



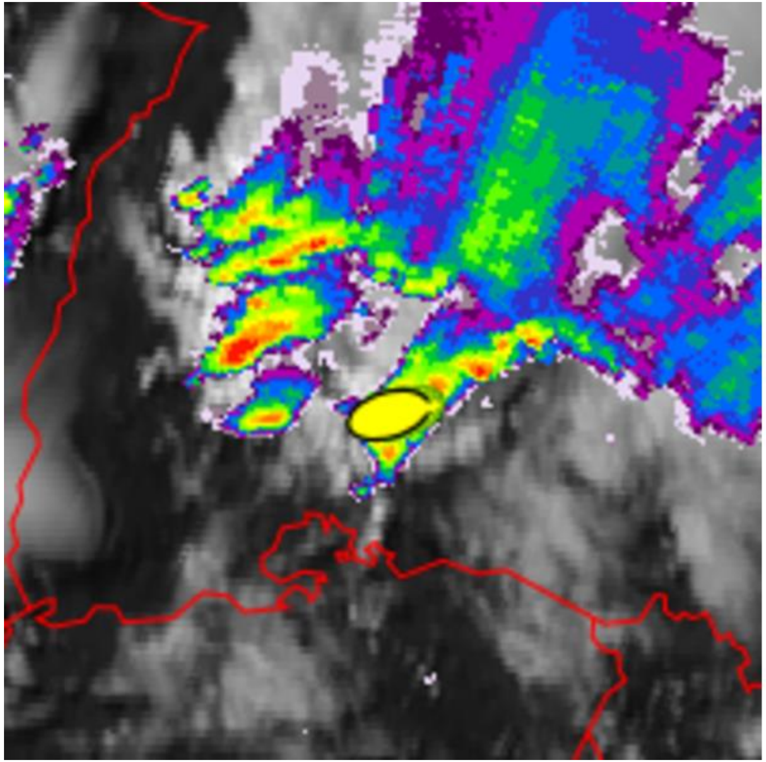


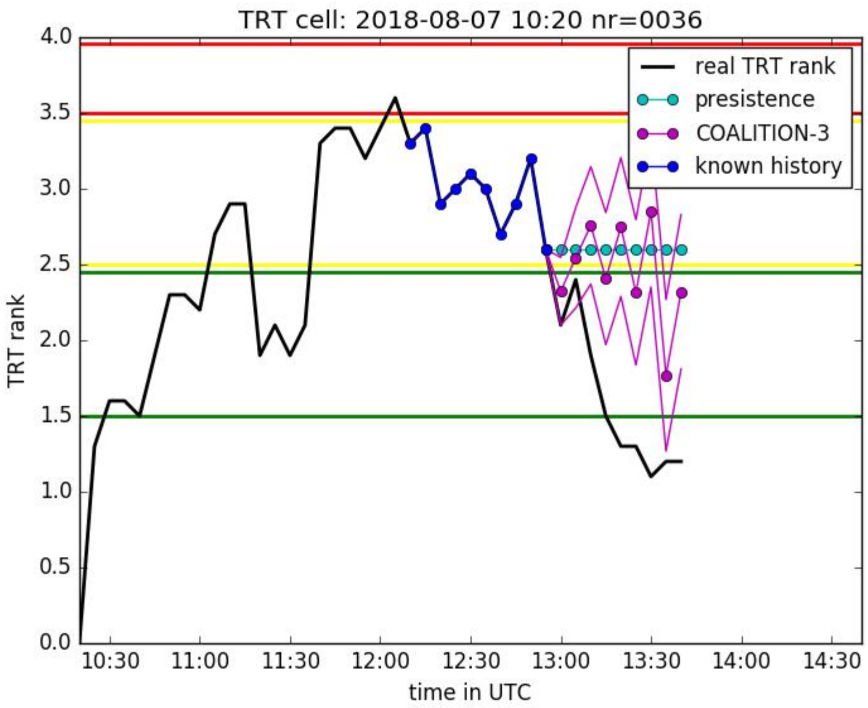
Visualization of the nowcasting results



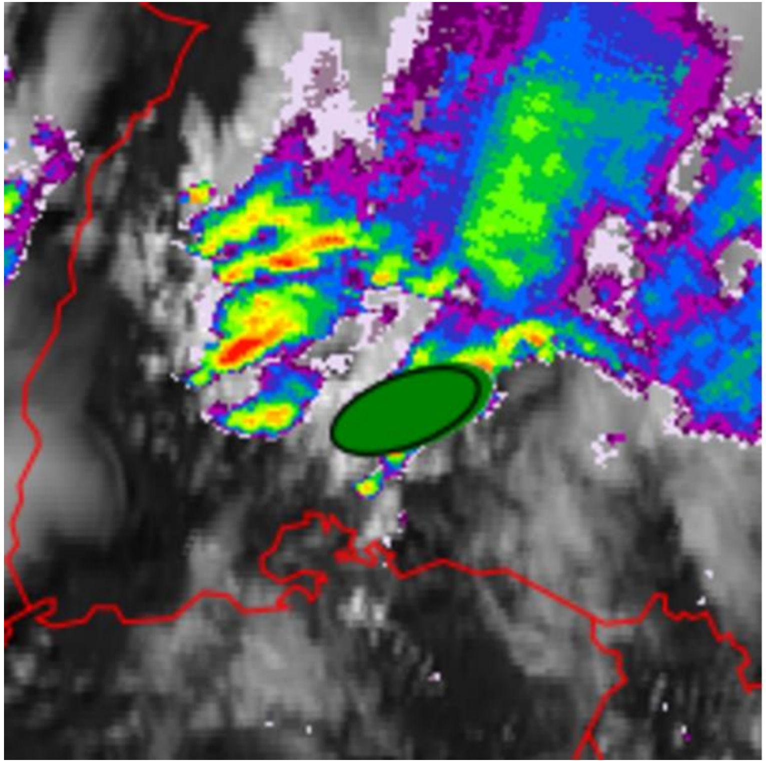


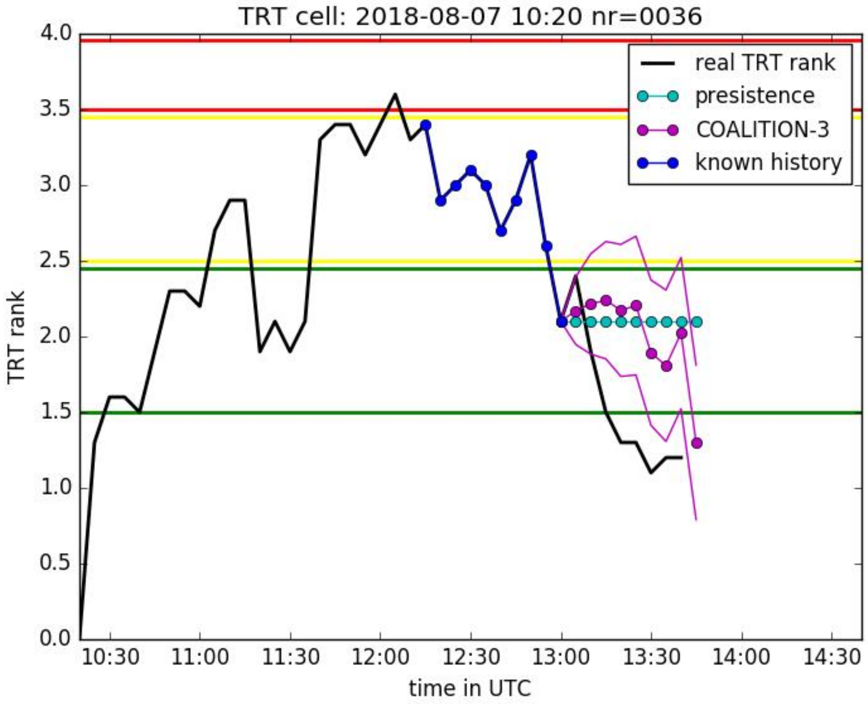
Visualization of the nowcasting results



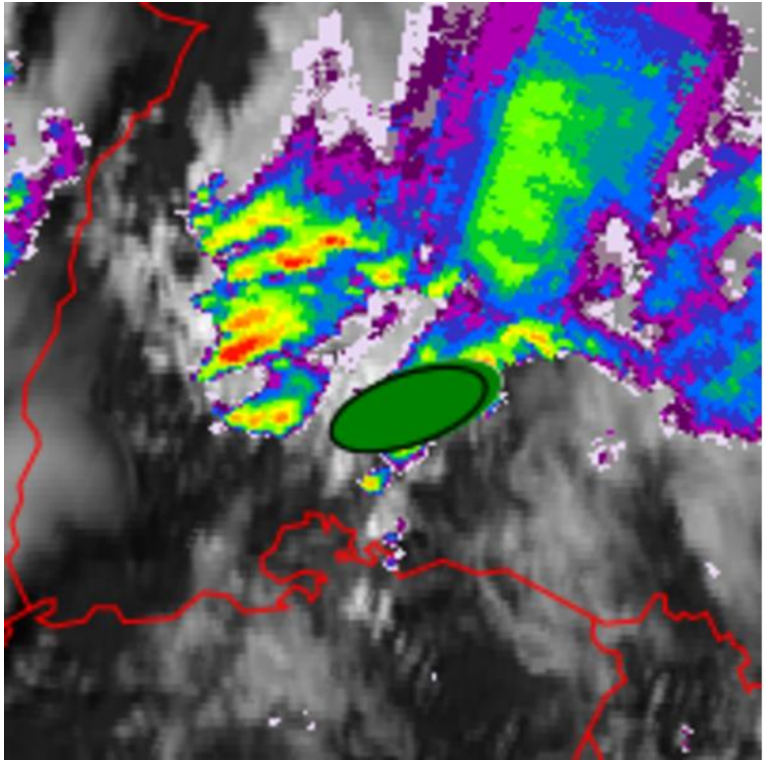


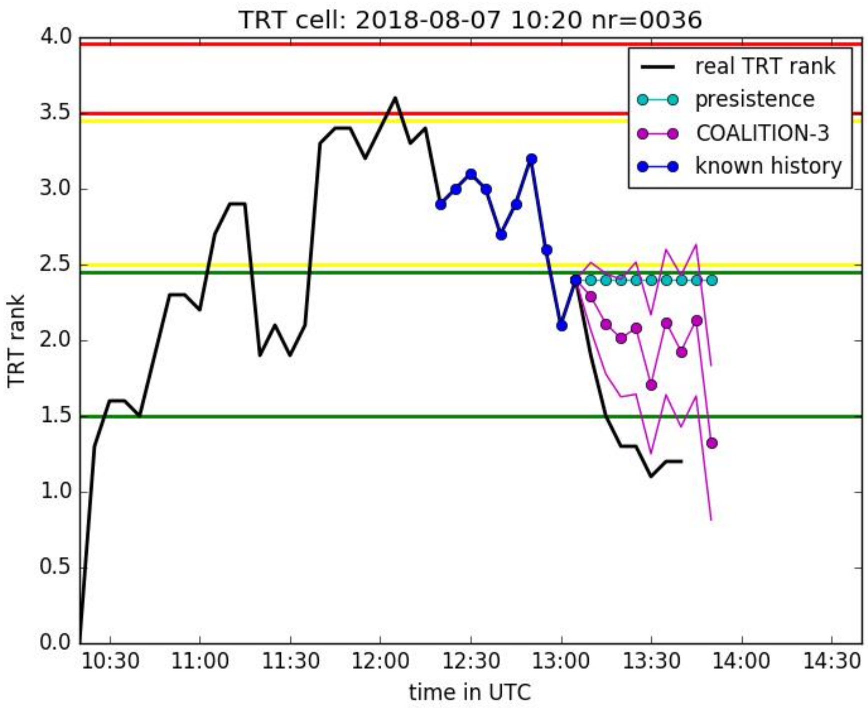
Visualization of the nowcasting results



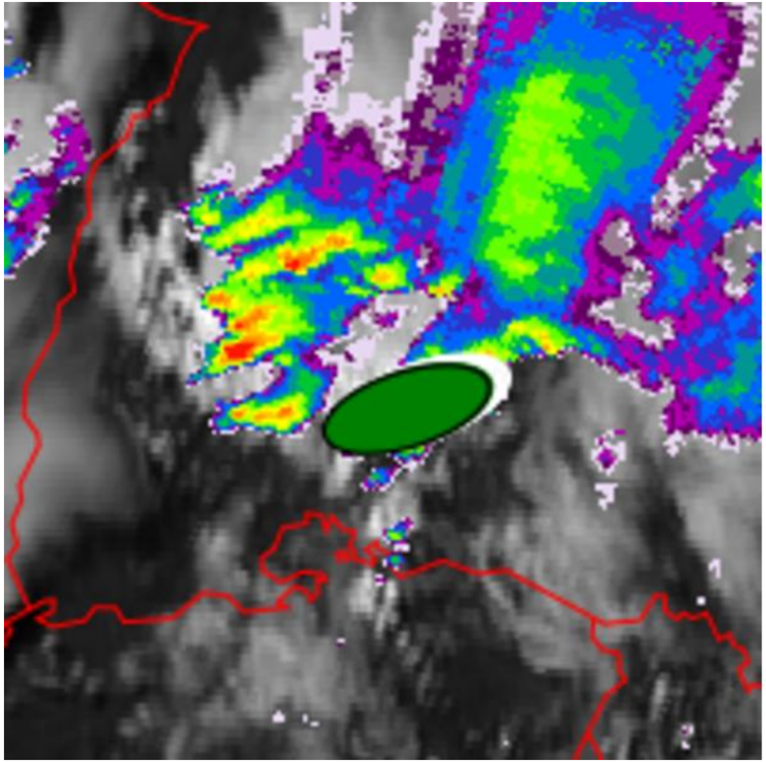


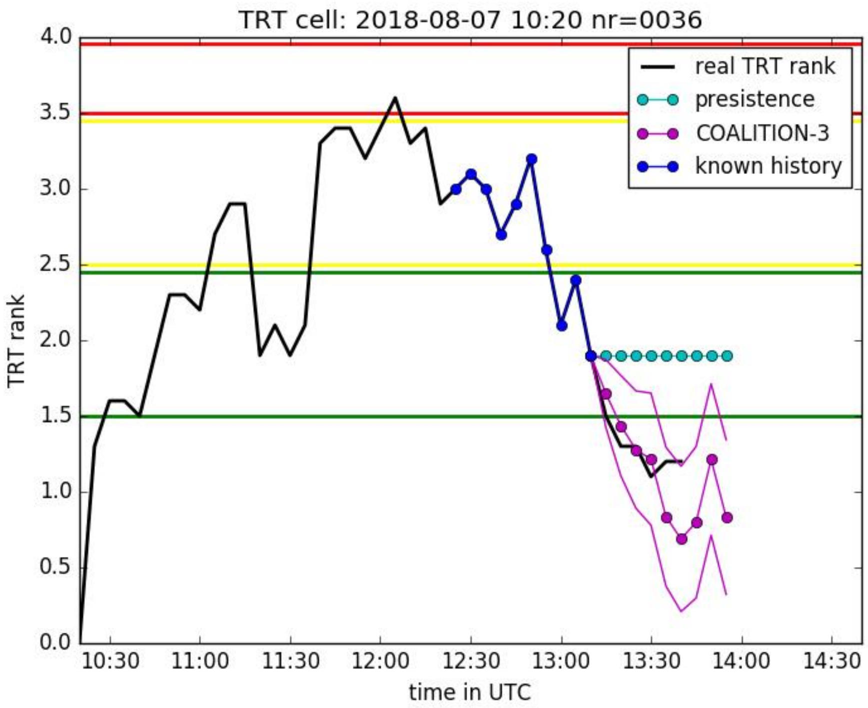
Visualization of the nowcasting results





Visualization of the nowcasting results





Summary

- A multi-sensor thunderstorm nowcasting was developed (COALITION-3)
- Thunderstorm cell are tracked with motion vectors
- Multi-sensor cell history is monitored
- Gradient boosted trees (XGBoost) is used to nowcast thunderstorm intensity
- Feature importance ranking enables reduction to 750 predictors
- Nowcasted TRT ranks for all forecast times better than persistence
- Nowcasted TRT ranks skilful up to 45 min into the future
- Probability matching is used to preserve standard deviation
- Easily expandable to more input variables
- Strait forward to train ML to nowcast other variables, e.g. lighting activity

Future Outlook

Future work COALITION-3

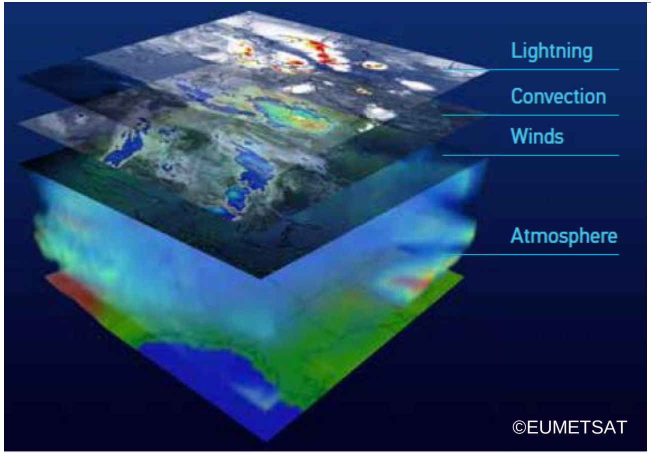
- Improved TRT rank **uncertainty estimates**
- **Feedback** from **forecasters**
- **Validation** of COALITION-3 in comparison to TRT and operational thunderstorm warnings (POD, FAR, SS for different warning levels)
- **Operationalisation** of COALITION-3 - Implement into automatic warning suggestion system of NinJo
- Explore **applicability for aviation** with European coverage

Future work COALITION-4

- Forecast of **specific thunderstorm hazards**: heavy precipitation, lightning, hail and wind gusts
- Improved **thunderstorm motion** (right movers, topographic steering)
- Exploiting **GPU-powered deep learning** technology,
  e.g. convolutional neural networks to improve further the nowcast quality and lead time
- **Prioritize warning** suggestions by risk & uncertainty

Future Outlook – Adaptation to MTG

1. MTG 4D Weather Cube

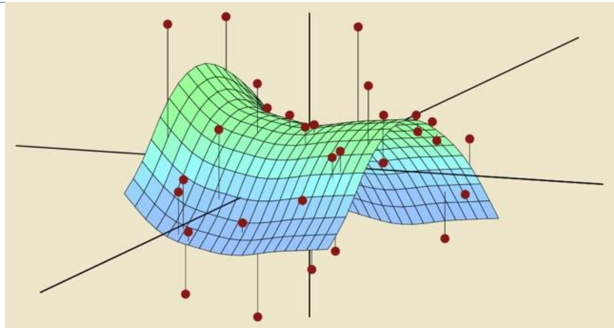


Radar

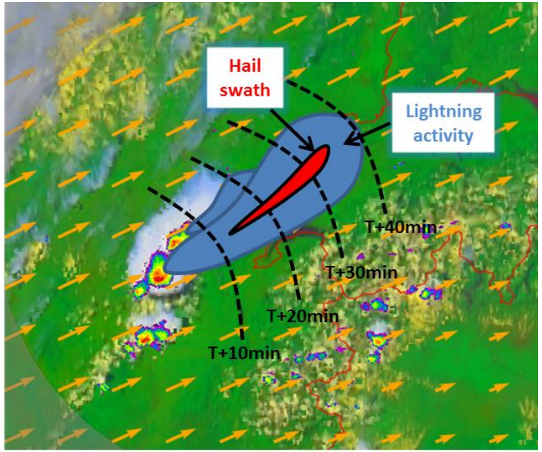
NWP model

Lightning

2. Advanced Machine Learning



3. Thunderstorm Nowcast



Rain

Hail

Lightning

Wind

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology MeteoSwiss

MeteoSwiss
Operation Center 1
CH-8058 Zurich-Airport
T +41 58 460 91 11
www.meteoswiss.ch

MeteoSvizzera
Via ai Monti 146
CH-6605 Locarno-Monti
T +41 58 460 92 22
www.meteosvizzera.ch

MétéoSuisse
7bis, av. de la Paix
CH-1211 Genève 2
T +41 58 460 98 88
www.meteosuisse.ch

MétéoSuisse
Chemin de l'Aérologie
CH-1530 Payerne
T +41 58 460 94 44
www.meteosuisse.ch