



Copernicus Emergency Management Service

What's New in EFAS Data Validation in Meteo DCC

EFAS Annual Meeting 2022

Presented by Zora Schirmeister and Dr. Christoph Schweim
27 09 2022



Emergency
Management

Data Validation in Meteorological Data Collection Center

- Meteo Data Collection Center checks incoming observation data using different quality checks, e.g.
 - Distance validation → time distance between two values, data availability
 - Min/max validation → values against min/max thresholds
 - Rate of change validation → rate of change between two values against max thresholds
- Resulting quality flags decide whether data are used for Meteo products or not

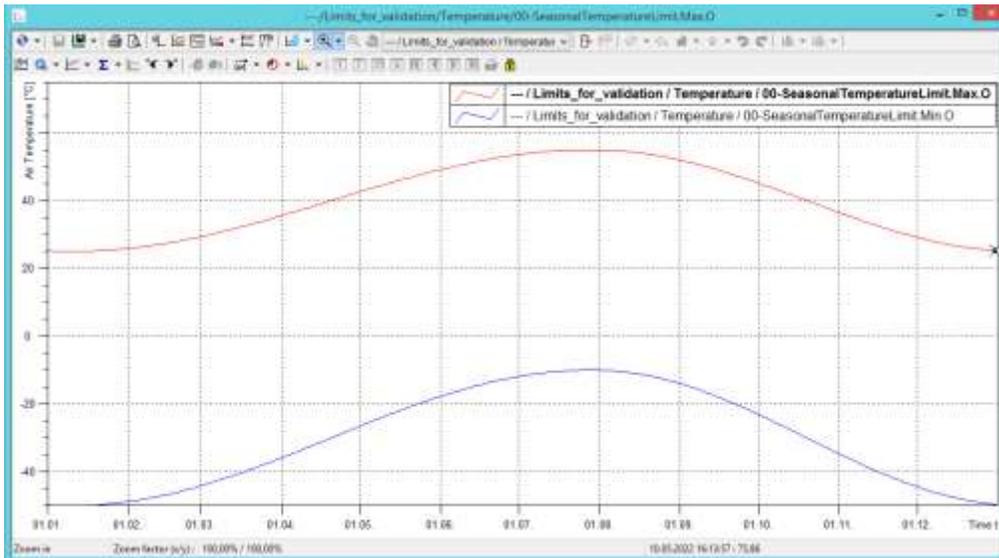
Parameter	Min threshold	Max threshold	Unit
Cloud cover	0	9	octas
Evaporation	0	2 3 15	mm/15 min mm/hour mm/day
Relative air humidity	5	100	%
Solar radiation	0	1360 cos(lat)	W/m ²
Sunshine duration	0	astronomic max	min
Vapour pressure	0	35	hPa
Wind direction	0	360	°
Wind speed	0	45	m/s

Example
min/max thresholds



Data Validation in Meteorological Data Collection Center

- Specific min/max validation rules
 - Seasonally varying thresholds for min/max temperature
 - Thresholds for precipitation totals
 - Thresholds are very tolerant to avoid excluding real extremes



Parameter	Min threshold	Max threshold	Unit
Precipitation	0	125	mm/15 min
Precipitation	0	200	mm/30 min
Precipitation	0	250	mm/60 min
Precipitation	0	350	mm/180 min
Precipitation	0	425	mm/360 min
Precipitation	0	475	mm/540 min
Precipitation	0	500	mm/720 min
Precipitation	0	525	mm/900 min
Precipitation	0	550	mm/1080 min
Precipitation	0	600	mm/1440 min



Emergency
Management

Data Validation in Meteorological Data Collection Center

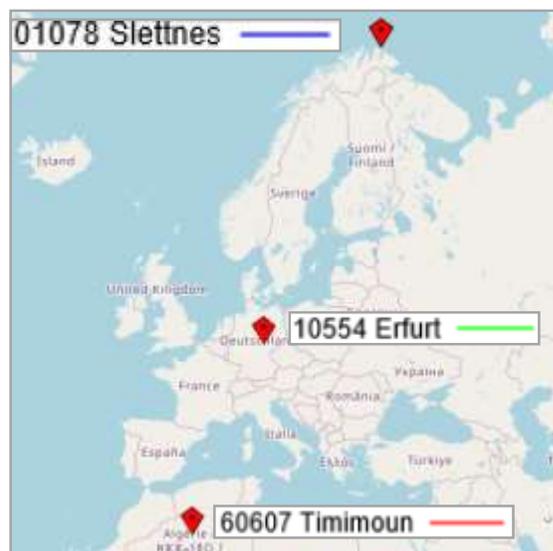
- Shortcomings
 - No geographical and few seasonal context:
 - Same thresholds for all stations in data base (e.g. North vs. South Europe)
 - No seasonal dependencies used for precipitation (e.g. ‚rain season‘ vs. ‚dry season‘)
 - Validation of individual time series
 - No information from neighboring stations
- Measures currently being setup
 - Introduction of climate zones and seasonally varying thresholds
 - Introduction of spatial comparison checks



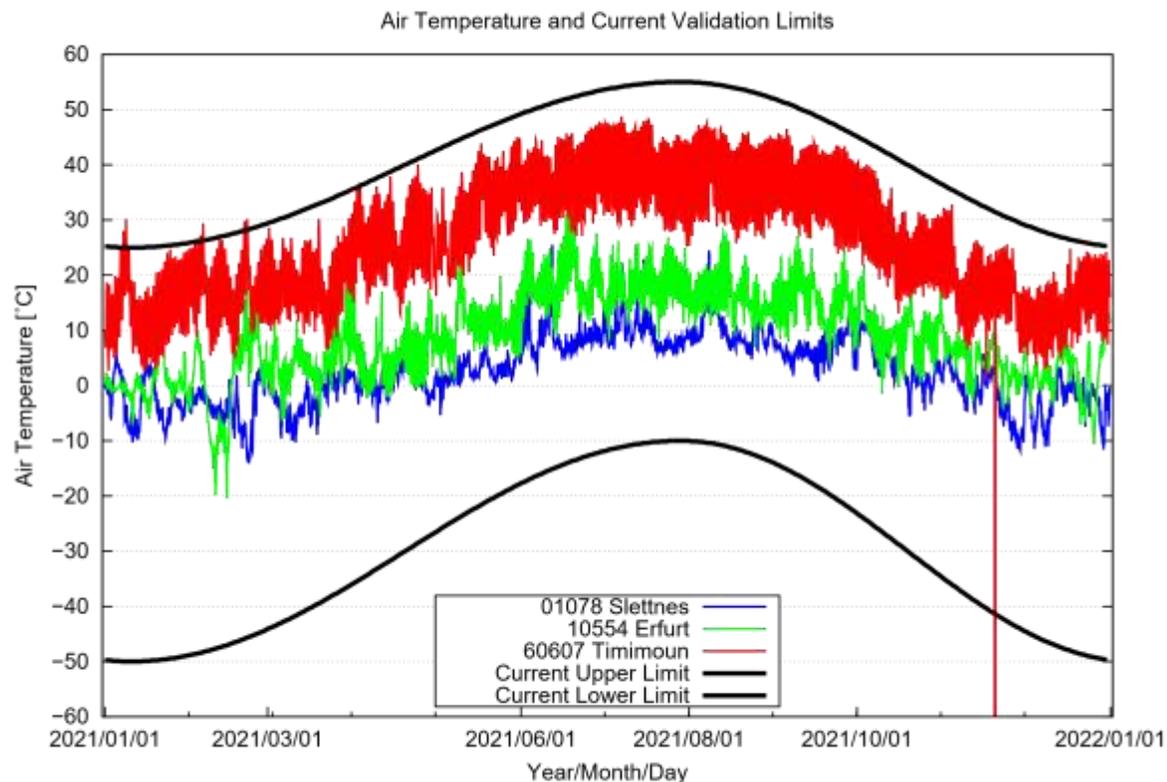
Emergency
Management

Meteo Data Validation: Thresholds by Climate Zone and Season

- Introduction of climate zones and seasonally varying thresholds
 - Example: Air temperature



openstreetmap.de/



- Goal:
 - Reduce range within limits
 - Don't flag real extremes

www.efas.eu



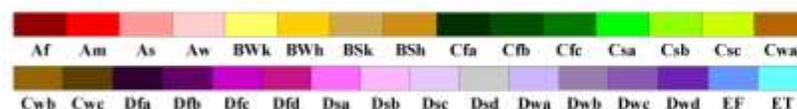


Meteo Data Validation: Thresholds by Climate Zone and Season

- Approach
 - Clever grouping of stations by similar geographical characteristics
→ using climate zones
 - Köppen-Geiger Climate Classification based on temperature and precipitation characteristics
 - Using data set provided by University of Veterinary Medicine, Vienna, based on CRU and GPCC data

World Map of Köppen–Geiger Climate Classification

updated with CRU TS 2.1 temperature and VASCLimO v1.1 precipitation data 1951 to 2000



Main climates

A: equatorial
B: arid
C: warm temperate
D: snow
E: polar

Precipitation

W: desert
S: steppe
f: fully humid
s: summer dry
w: winter dry
m: monsoonal

Temperature

h: hot arid
k: cold arid
a: hot summer
b: warm summer
c: cool summer
d: extremely continental
F: polar frost
T: polar tundra

Resolution: 0.5 deg lat/lon



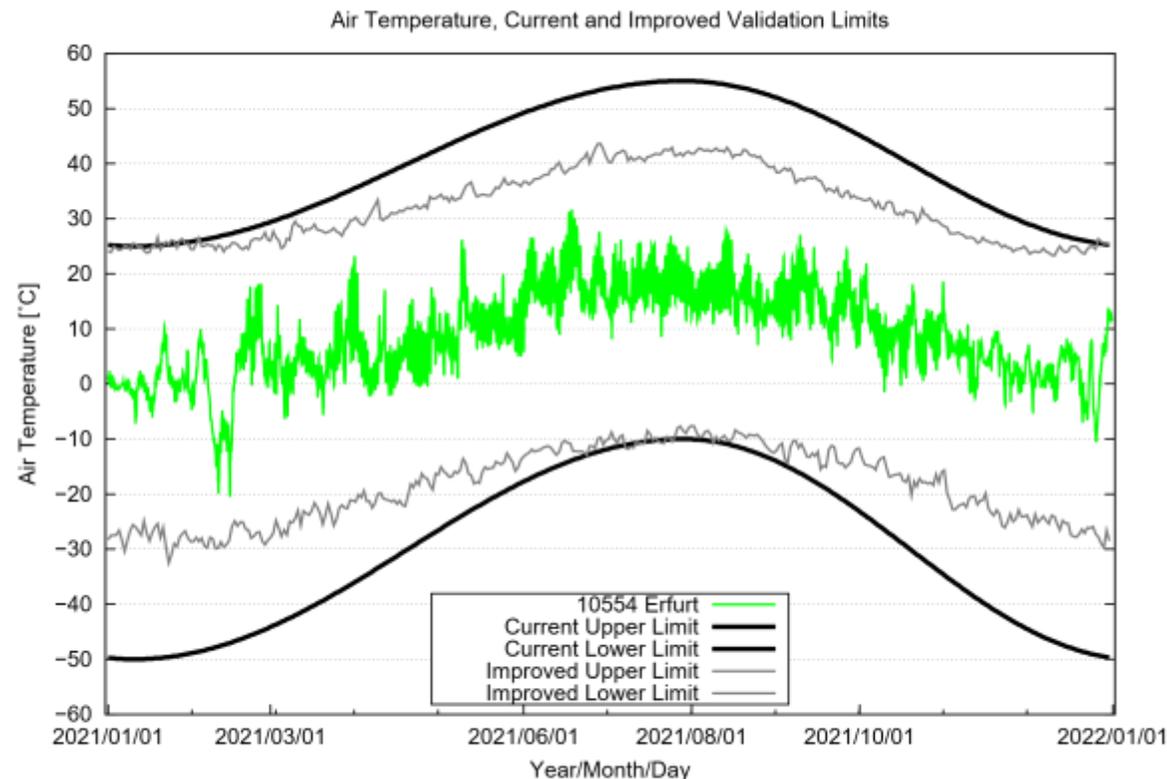
Version of April 2006



Emergency
Management

Meteo Data Validation: Thresholds by Climate Zone and Season

- Definition of new thresholds
 - Group stations by climate zones and calculate thresholds for each group
 - adds regional dependency
 - Separate percentile for each day → adds seasonal dependency
 - Example CfB (Central Europe)





Emergency
Management

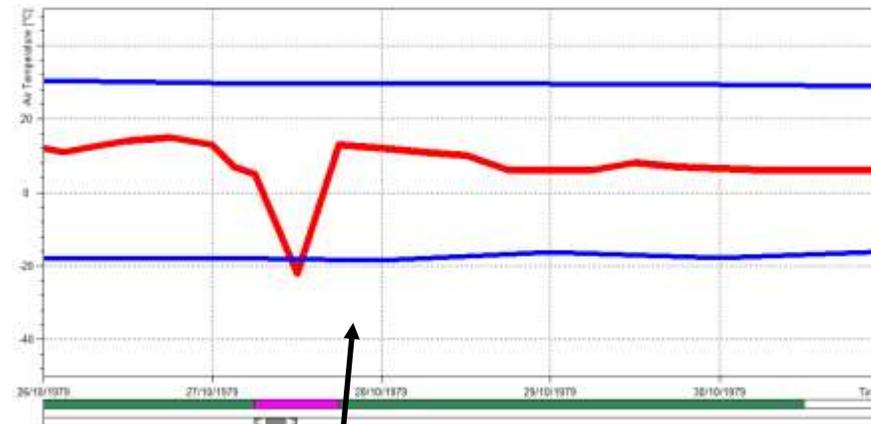
Meteo Data Validation: Thresholds by Climate Zone and Season

- Tests with real data

Current configuration



Improved configuration



Value now flagged as below improved threshold



Emergency
Management

Meteo Data Validation: Spatial Comparison

- Introduction of spatial comparison checks
- Objective
 - Use measured truth from neighboring stations to assess the quality of an isolated observation
- Approach
 - Identification of neighboring stations
 - Comparison with observed values from neighbors
 - Quality flag time series with high deviations („suspect“ or „rejected“)
 - Review or exclude from usage for Meteo products



- Methods for identification of neighboring stations

- Distance (nearest)
- Orientation (cardinal directions)
- Quadrant (coordinate system)
- Max / priority radius
- Max elevation difference

Type:

	Rule (Global)	Metadata Attributes
Absolute Scaling Factor:	<input type="text" value="3"/>	
Relative Tolerance Down:	<input type="text"/>	
Relative Tolerance Up:	<input type="text"/>	
Elevation:	<input type="text"/>	station.elevation
Elevation Lapse Rate:	<input type="text" value="0"/>	
Group size:	<input type="text" value="6"/>	
Latitude:	<input type="text"/>	station.position.latitude
Longitude:	<input type="text"/>	station.position.longitude
Max Elevation:	<input type="text" value="501"/>	
Max Horizontal:	<input type="text" value="50000"/>	
Max Horizontal Inw:	<input type="text" value="25000"/>	
Mode:	<input type="text" value="orientation"/>	
Value max:	<input type="text" value="60"/>	
Value min:	<input type="text" value="30"/>	

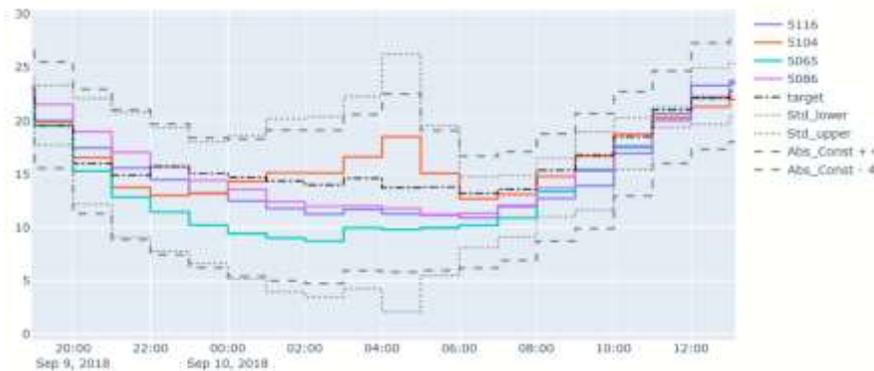
Cancel Save





Meteo Data Validation: Spatial Comparison

- Methods for definition of upper and lower limits of allowed values from relevant neighbors:
 - Standard deviation method σ multiplied by a factor
 - Mean multiplied with positive and negative factor
 - Constant offset value



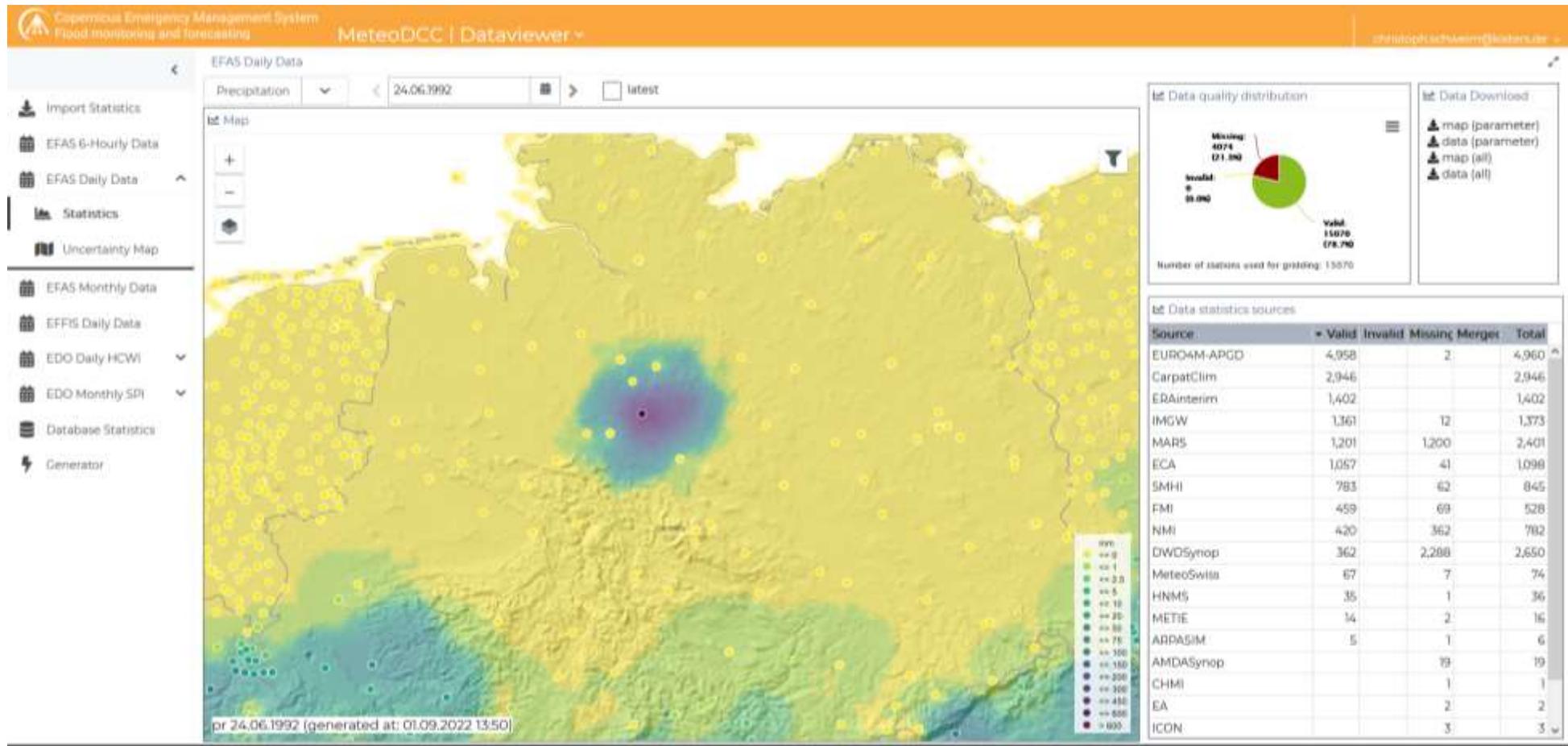
Field	Value	Metadata Attribute
Absolute Scaling Factor	3	
Relative Tolerance Down		
Relative Tolerance Up		
Elevation	501	station.elevation
Elevation Lapse Rate	0	
Group size	6	
Latitude		station.position.latitude
Longitude		station.position.longitude
Max Horizontal	50000	
Max Horizontal Inner	25000	
Mode	orientation	
Value max	60	
Value min	-30	



Emergency Management

Meteo Data Validation: Spatial Comparison

- Example for undetected issue and result in precipitation grid:





Emergency
Management

Meteo Data Validation: Spatial Comparison

- Identification of issue in Data Validation App:

KISTERS Applications | Data Validation

Configuration Results

Widget View

← Results | TEST-CASE #2

Score	Station Name	Station ID	Incident	Options
Fail	Celle	10343	Spatial comparis...	👁️ 📄
Good	Wunstorf	10334		👁️ 📄
Good	Soltau	10235		👁️ 📄
Good	Bergen	10238		👁️ 📄
Good	Fallberg	10246		👁️ 📄
Good	HANNOVER	476		👁️ 📄
Good	Braunschweig	10348		👁️ 📄
Good	Gardelegen	10359		👁️ 📄
No Data	HILDESHEIM	10337		👁️ 📄

Map: Celle (10343) Incident: Spatial comparison is exceeded (3)

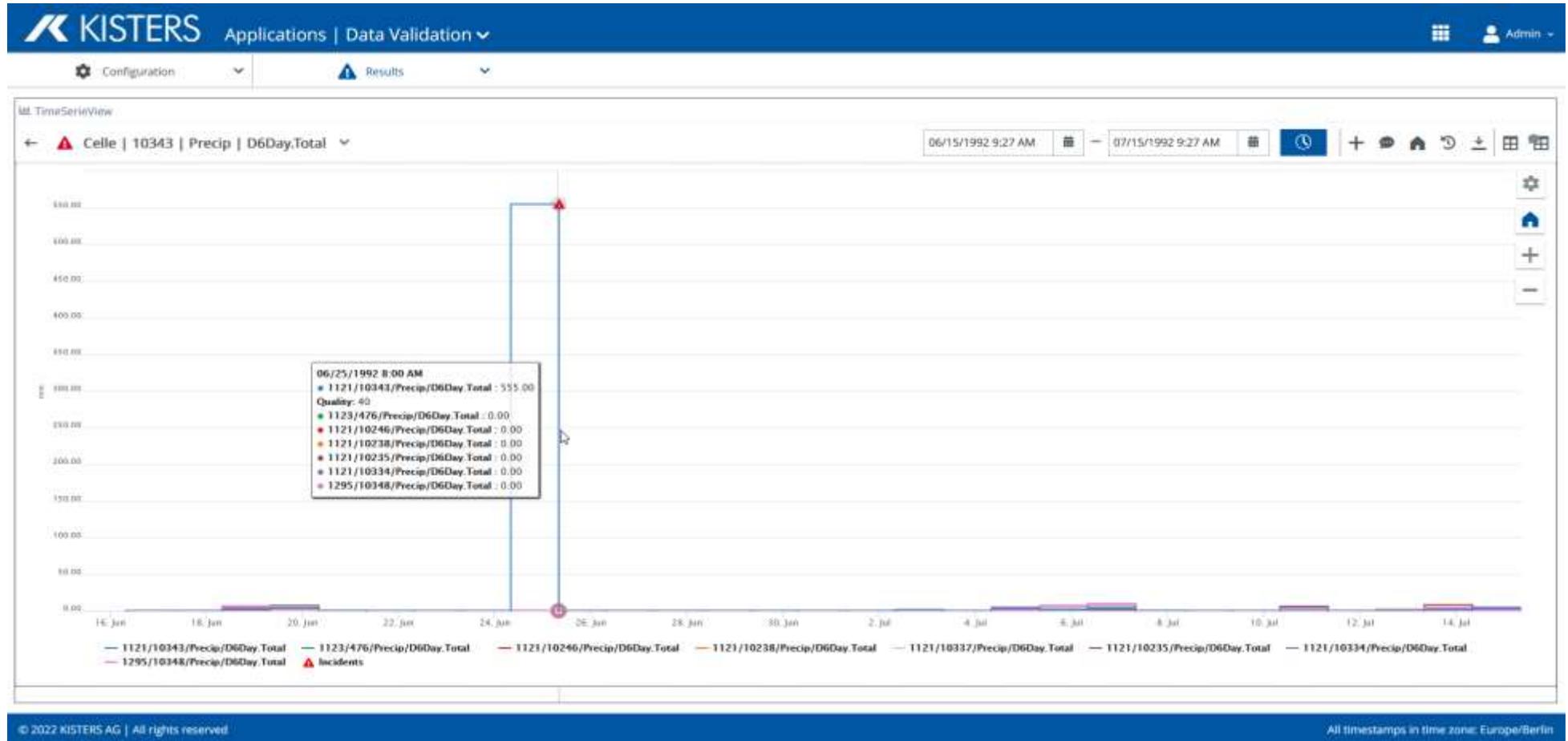
© 2022 KISTERS AG | All rights reserved. All timestamps in time zone: Europe/Berlin



Emergency
Management

Meteo Data Validation: Spatial Comparison

- Identification of issue in Data Validation App:

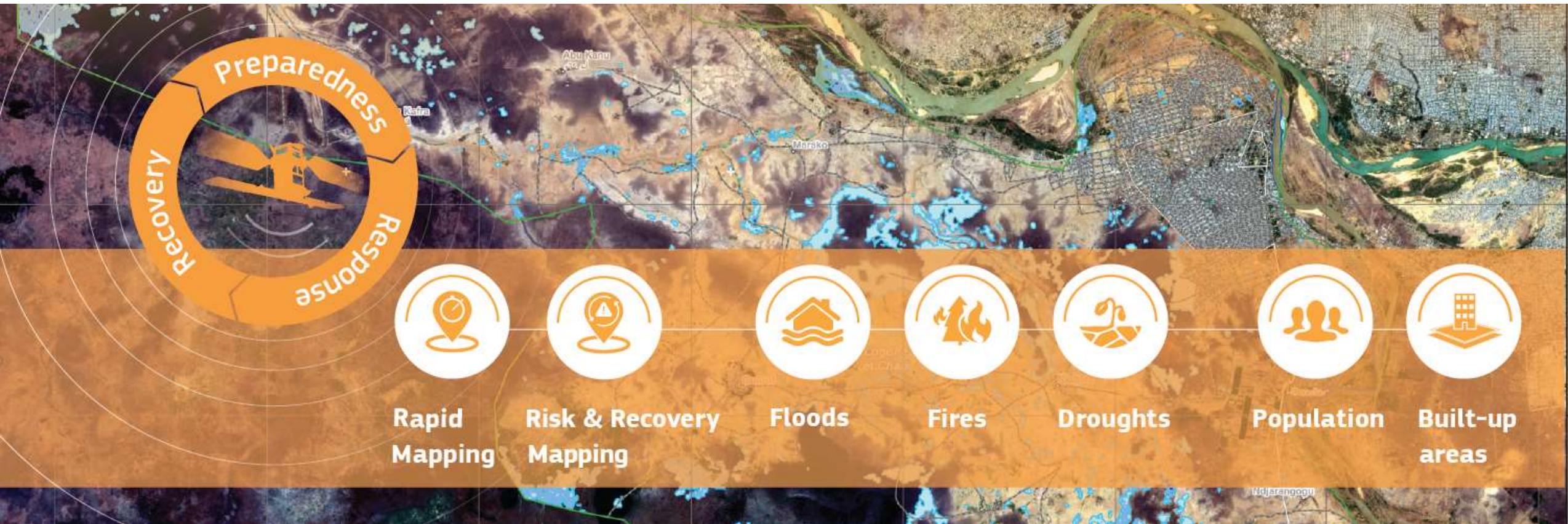




Emergency
Management

Meteo Data Validation: Status & Outlook

- Seasonally varying thresholds per climate zone
 - Statistical calculations of thresholds currently ongoing
 - Integration into production in near future
- Spatial comparison for precipitation
 - Software development under finalization
 - Integration into production in near future
- Expected Results
 - Better identification of issues in observation data
 - Better quality of data products created at Meteo DCC



Rapid Mapping



Risk & Recovery Mapping



Floods



Fires



Droughts



Population



Built-up areas