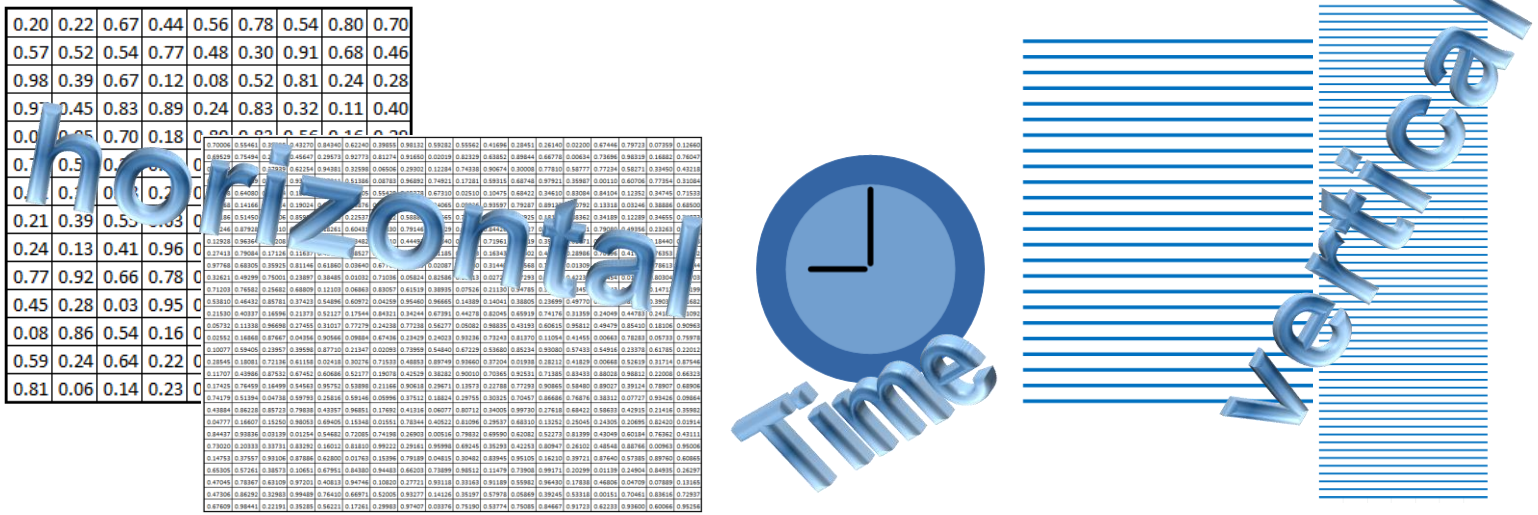
Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Significant increases in model data volumes

and other challenges for a meteorological visualisation platform

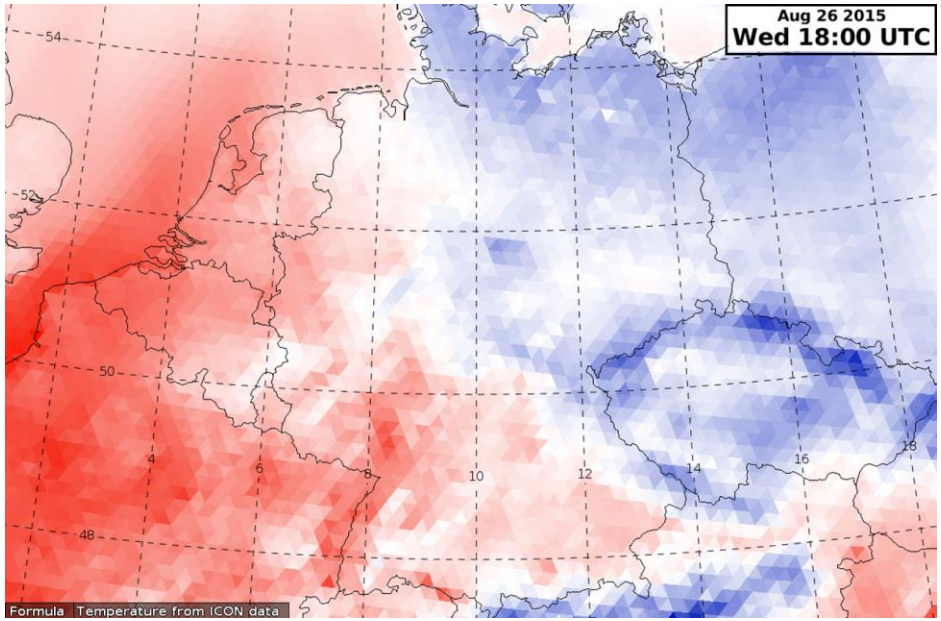


Visualisation in Meteorology Week

1st Oct 2015, Marcus Werner, DWD

Overview

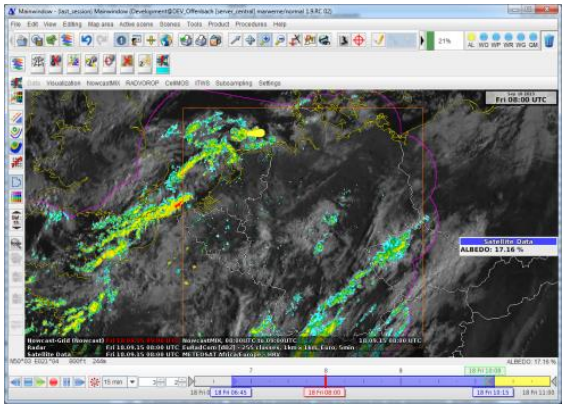
- NinJo at DWD
- Operational data volumes
- New model chain
- Distributed environment
- Generic unstructured grids
  - WMS experiment - ICON
  - Formula layer prototype
- Control volume growth
- Opening opportunities?
- Other meteorological data types



Bruno Zürcher, MeteoSwiss

DWD's meteorological workstation - NinJo

- NinJo is developed by an international Consortium

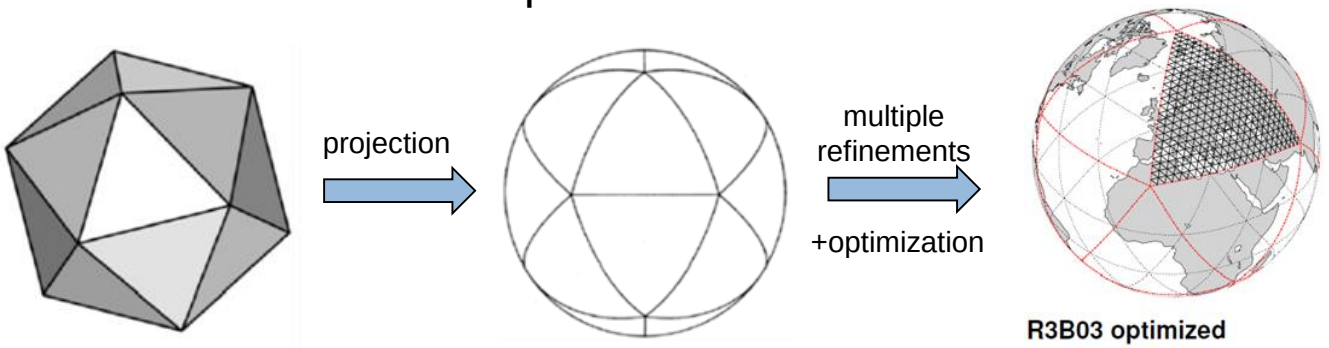


- Primary tool for our forecasters
- Short range forecast of DWD is based on
  - ICON Global, COSMO-EU, COSMO-DE, COSMO-DE EPS
  - In-house MOS systems
- ECMWF IFS, Euro4, GFS and other models are also available
- NWP counts for >50% cached data, followed by radar >30%

ICON Global & ICON-EU Nest

Florian Prill, DWD

- ICON joint development
  with Max-Planck-Institute for Meteorology
- ICON Global replaced GME (early 2015)
- ICON-EU Nest operational calculated
  - COSMO-EU is to be retired
- ICON 2-way nesting ensures consistency of global and regional run
- ICON is based on a generic unstructured grid
  - Optimized from a NWP calculation point of view



NWP volumes and new properties

- ➔ Users request high resolution NWP data (on all scales)
- ➔ NWP model updates provide 4x (horizontal)
- ➔ Vertical resolution, time resolution and forecast length increases
  - ➔ # of meteorological fields + field sizes
- ➔ NWP and NowCasting is going to merge (seamless prediction)
  - ➔ … acceleration … ?
- ➔ Unstructured generic grid (complexity++)
  - ➔ GRIBs do not contain all necessary information anymore
- ➔ Take advantage from consistency on different scales (glob./reg.)
  - ➔ GUI navigation + new visualisation algorithms

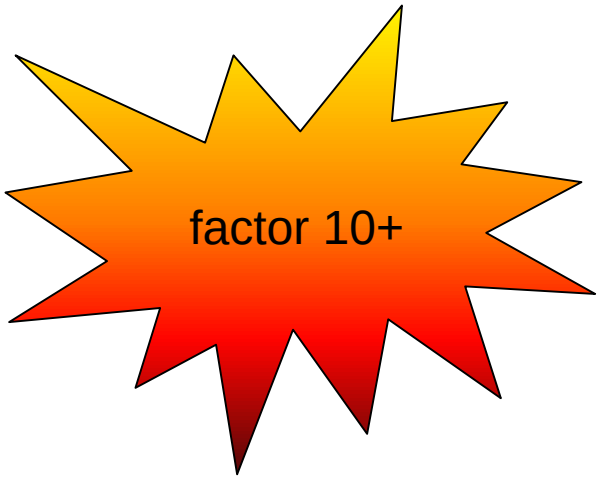
Some volume numbers around ICON

- Global ICON (R03B07)
  - ~3 mill. triangles, 4.5 mill. edges, ~13km
    - 90 vertical levels, +180h
  - Geometry file 1.8 GB netcdf + GRIB2 fields for model heights ~300 MB
    - Algorithmic grid generators do not work
- ICON EU-Nest (R03B08)
  - ~700k triangles, 1 mill. edges, ~6.5 km
    - 60 vertical levels, +120h
  - Geometry 400 MB netcdf

- Hence → dual output (native & regular)
  - Global ICON: regular lat/lon 0.25° grid (next step 0.125°)
  - ICON-EU Nest: regular lat/lon 0.0625° grid
  - Central and regional packages

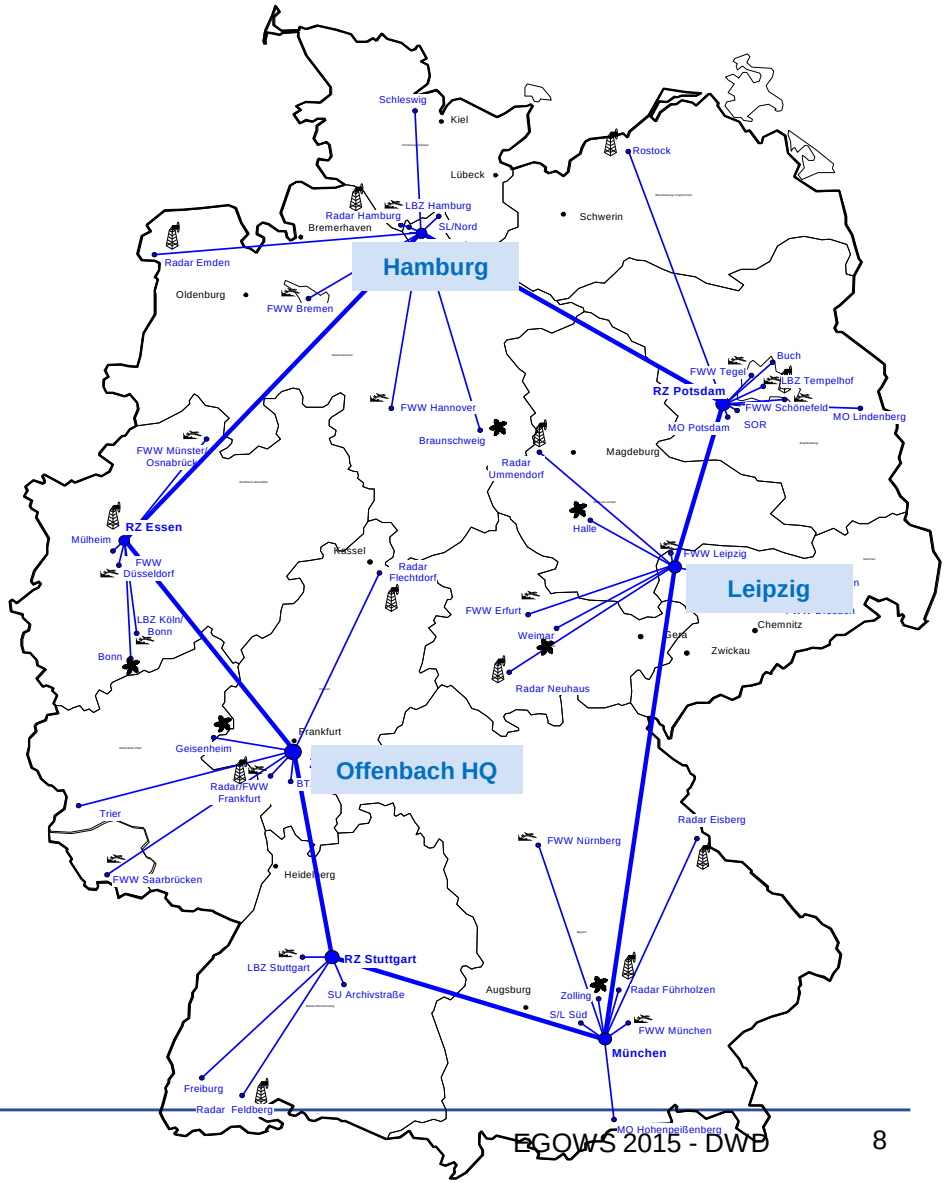
DWD NinJo - operational server cache sizes

- Hold time = 36h → ICON (last 6 model runs)
- Today (“measured”)
  - ICON Global 0.25°: 100 GB (~“old” GME data set)
  - COSMO-EU: 170 GB
- Incremental migration (extrapolated volumes)
  - ICON Global
    - New operational data set (0.25° grid)
      - ~1 TB
    - New operational data set (0.125° grid)
      - ~4 TB
    - Maximum extension up to ~7 TB
  - ICON-EU Nest (0.0625° grid)
    - New operational data set ~1.5 TB



NinJo deployment at DWD

- 3 central NinJo sites
  - Full operational data set
  - Backup each other
- Regional installations
  - Reduced data set
  - Regional clients can access full central data set
- Server and storage systems are up-to-date (16 TB ssd)
- HW is ready for new volumes
- Improved failover capabilities
- We are prepared for the moment…..



… planners, developers and budget responsibles face…

- Volumes seem to grow faster as hardware is replaced by renewal cycle
- Increasing NWP amounts effecting whole operational data chain
  (NWP DMO is processed and stored multiple times)
- Growing data volumes result in challenges for applications and libraries
  - Performance issues, technical limits → maintenance, test, cost ++
- Regional offices and backup/failover requires powerful WAN
- It is not easy to classify fields and get numbers for
  - Critical and heavy used data
  - Has to be there in case of need

Control of growth… ?

- DWD forecasters agreed new operational ICON data set

… important step in the right direction …

Recent discussions at DWD…

- Surface data, boundary layer & low heights are required in high resolution
- Top most layers are not required for operational usage (cap at ~20km)
- How much operational benefit provide high resolution data of the regional model for the free atmosphere in compare to the hi-res global model?
- Smooth fields vs. complex inherent structure
- What other options exist?
  - Well known: Precision and coding of numbers
  - Well known: Compression ←→ run times
  - Can we use problem specific resolutions and data? (→ general grids)
  - Technical: What is the best data format for our data caches, taking modern parallel CPUs, memory and storage systems into account?

Visualisation of generic grids (start point)

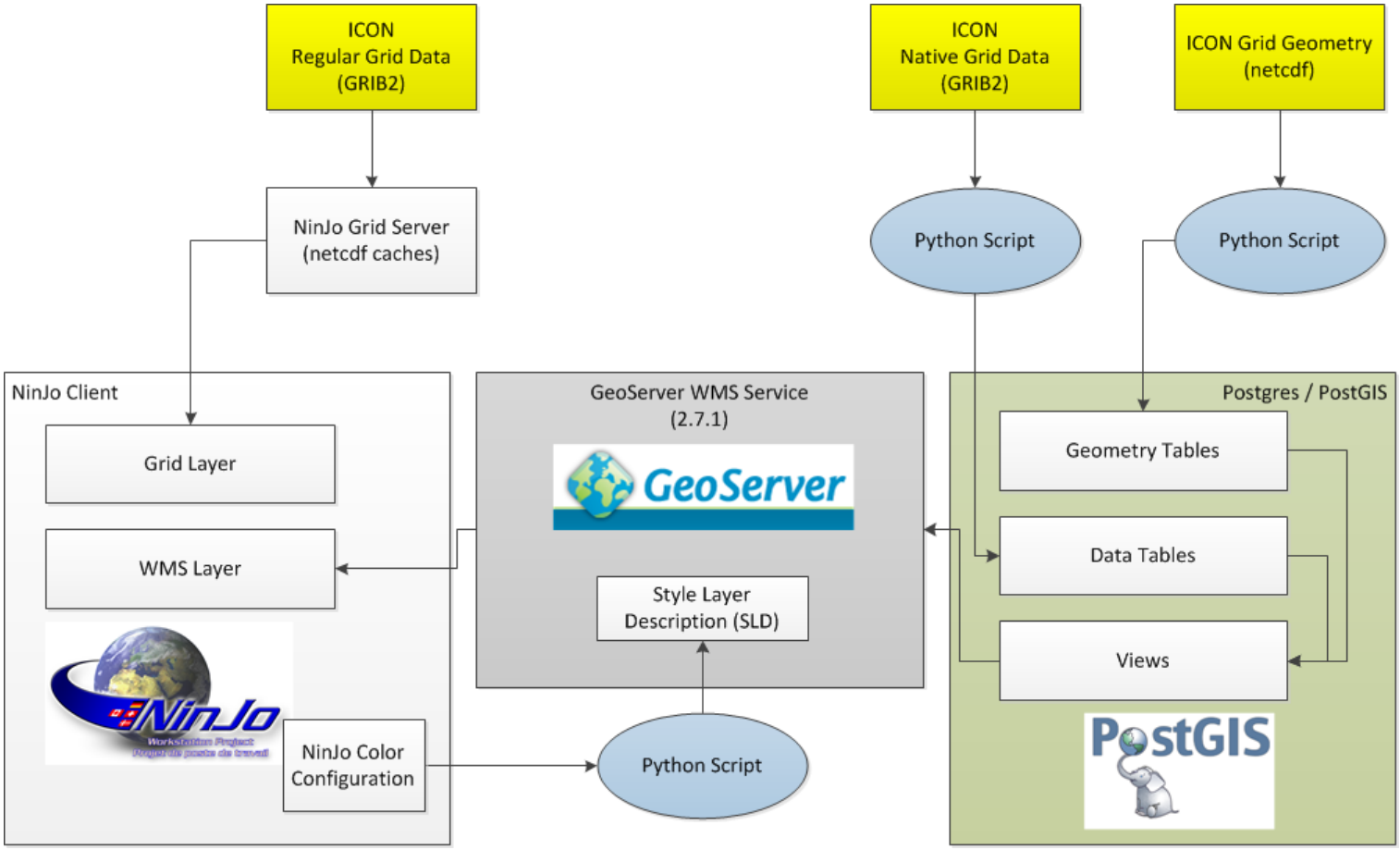
- Operational performance: max. 2 sec screen update

- Functional limit: NinJo Grid Layer cannot display generic grids
  - Internal data structures require extensions
  - Meteograms, cross-sections and model soundings need to work with generic grid data

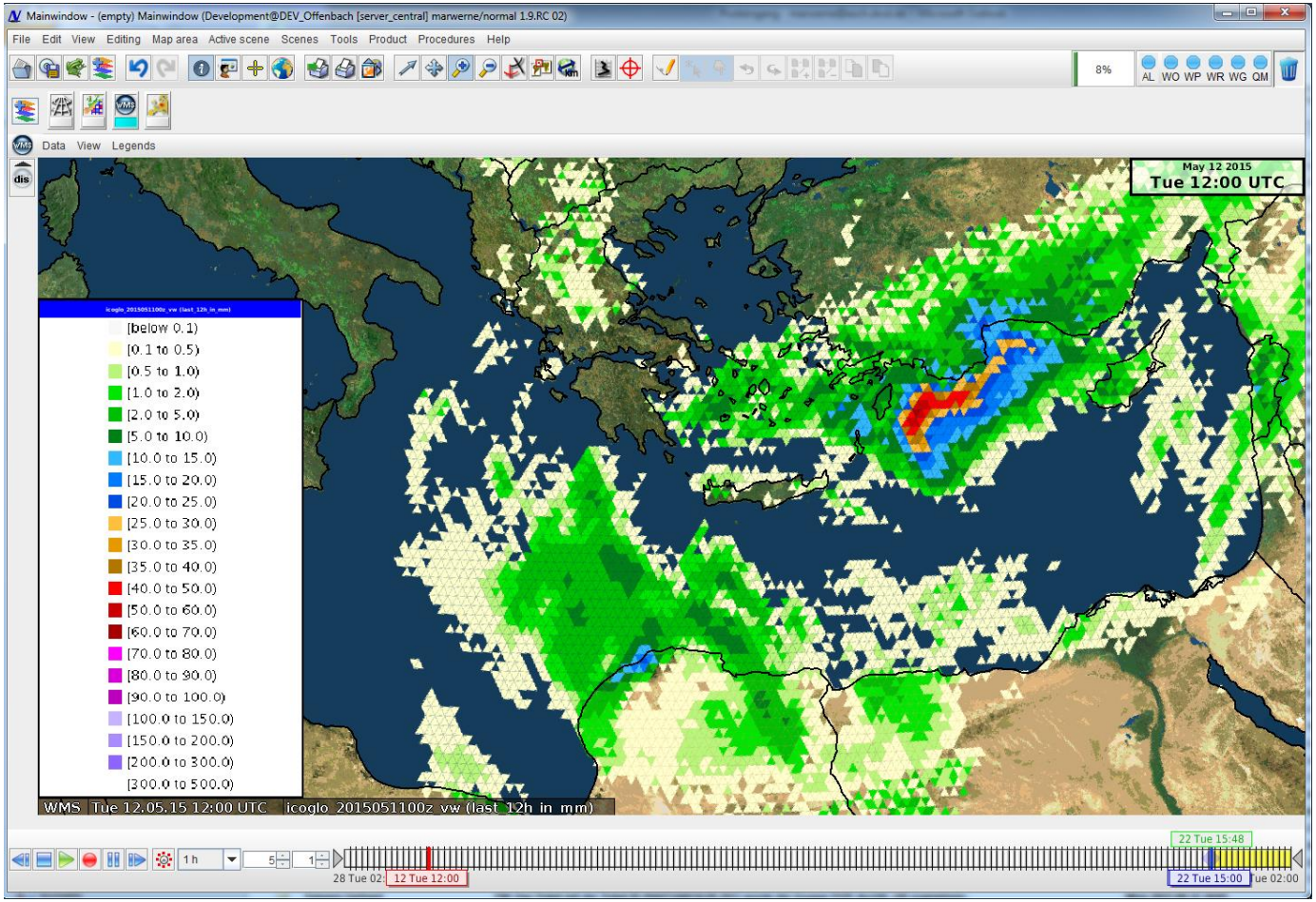
- Use forward mapping or backward mapping?
  - Forward: data values & geometry → pixels on the screen (requires cutting)
  - Backward: screen pixel → lat/lon → data values (requires “expensive” searches)

- Can we get a first display with low efforts?

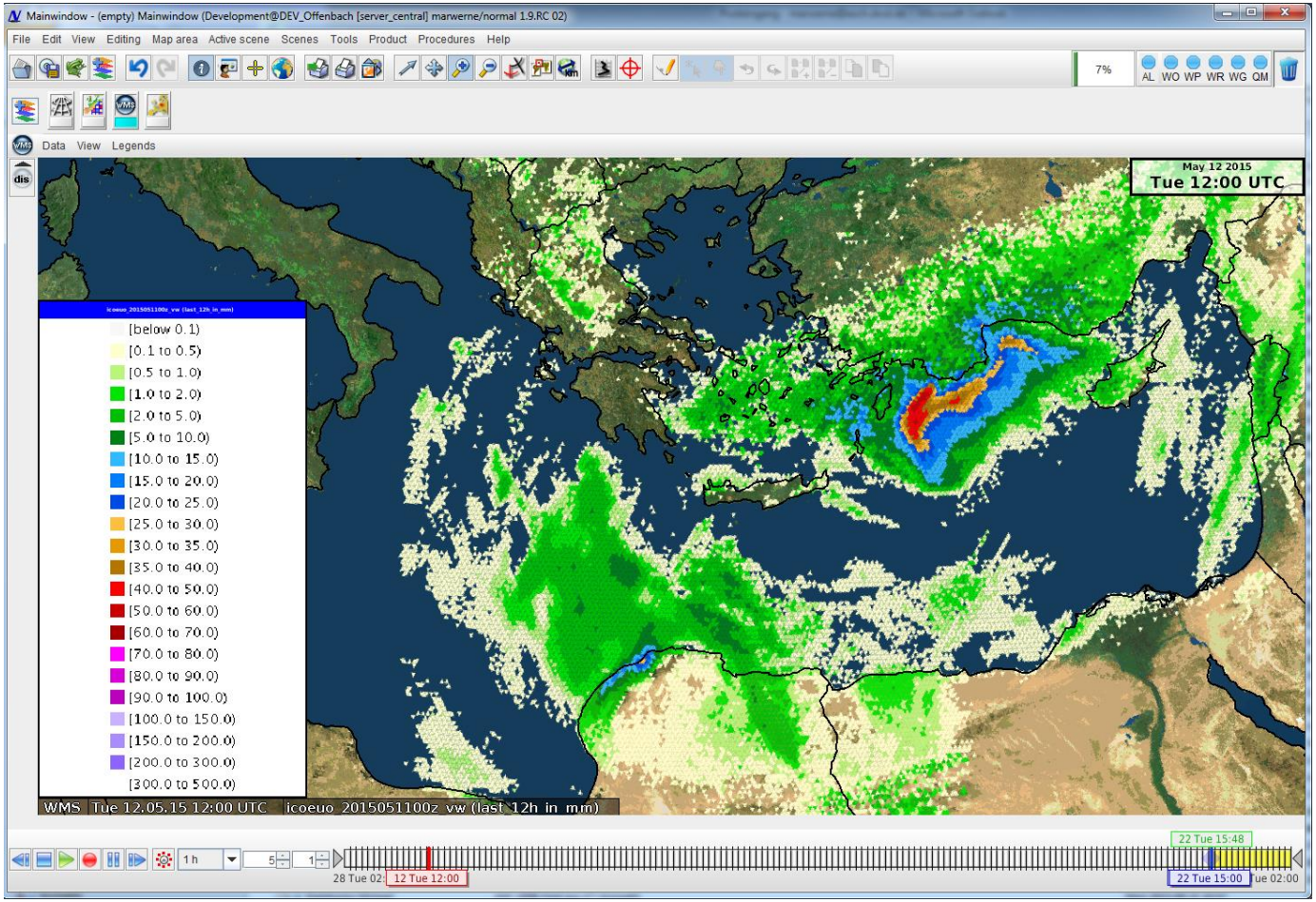
An ICON data experiment with GeoServer and NinJo (I)



… experiment … (II) – Global 13km grid – 12h precip.

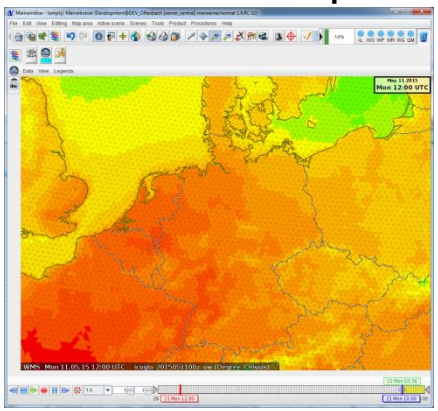


… experiment … (III) – Regional 6.5km grid – 12h precip.



An ICON data experiment with GeoServer and NinJo (IV)

- GeoServer 2.7.1 WMS and NinJo WMS Layer will display native ICON data
- Acceptable display times, require usage of tile cache
- 1 model run, 1 meteorological element (T2M) on surface level, 1 forecast step
  → 1 data field
  1 visualisation method (single SLD per met.-element),
  restricted to zoom level 0-8, 1 projection EPSG:4326:
  - ICON Global: ~25 MB tile cache size (png8)
  - ICON-EU Nest: ~40 MB tile cache size (png8)

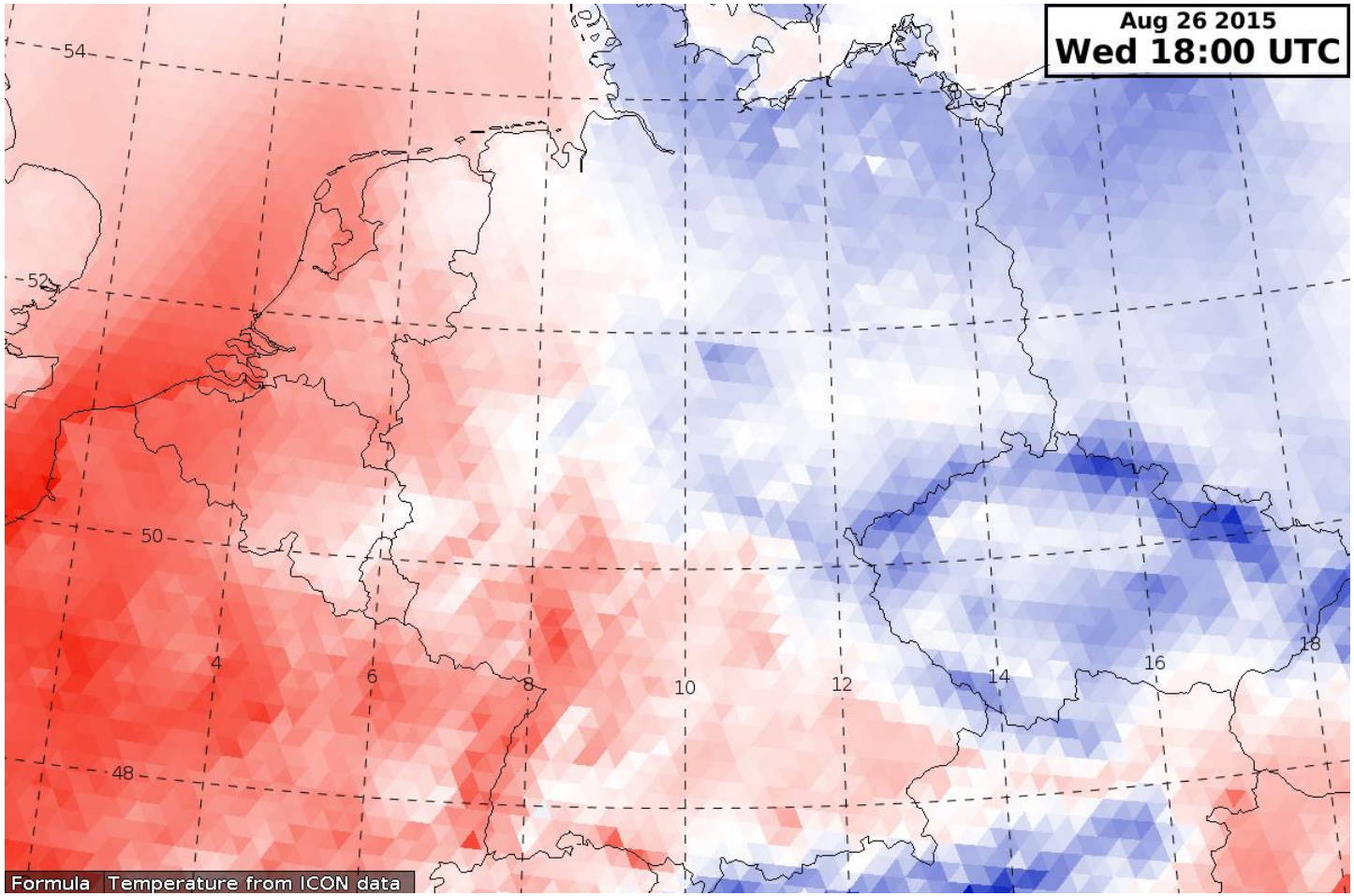


- New operational ICON data set:
  - ICON Global: ~45.000 fields per model run (→ ~1,2 TB)
  - ICON-EU Nest: ~68.000 fields per model run (→ ~2,7 TB)

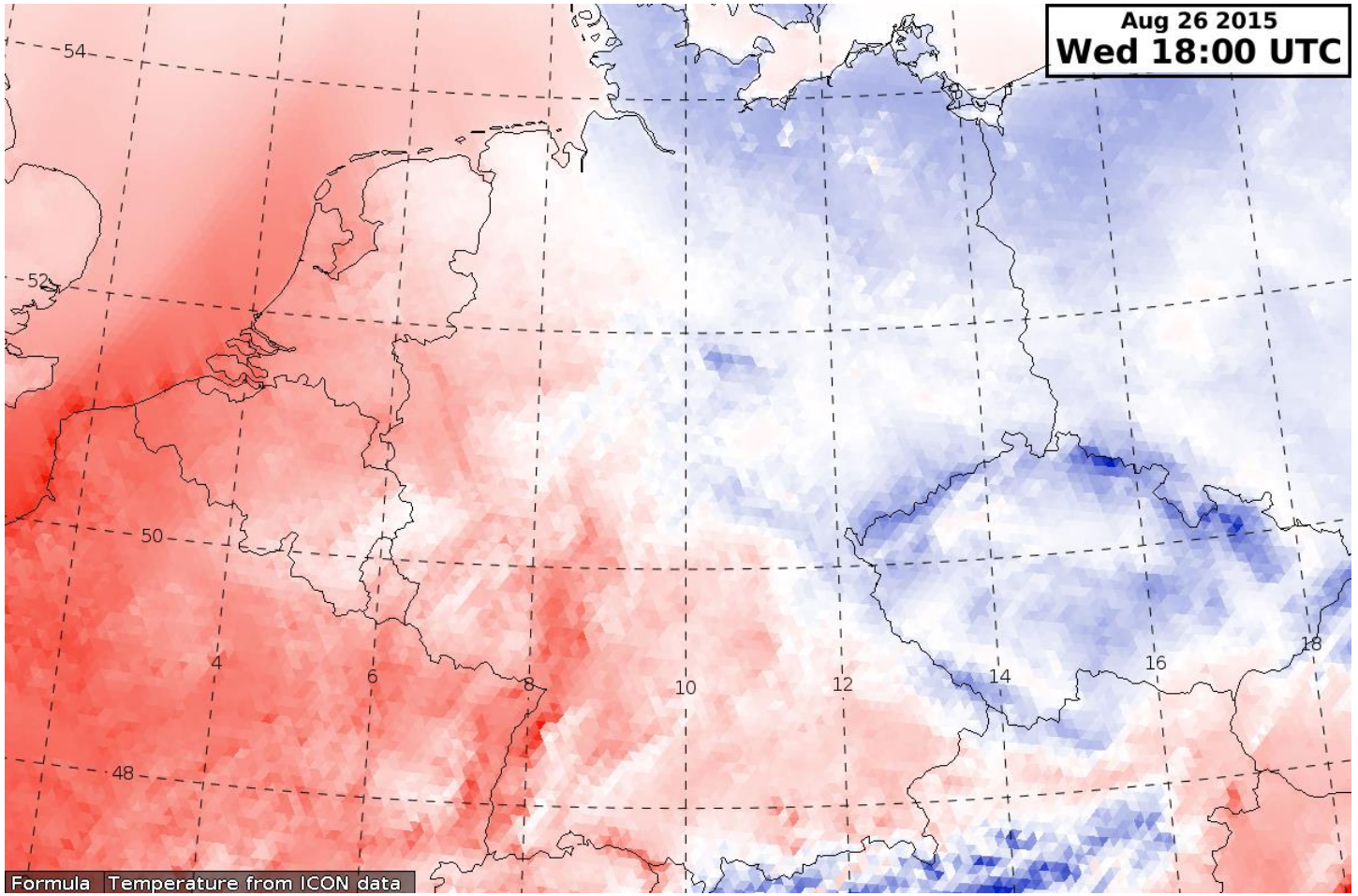
A native grid display prototype - NinJo formula layer

- Open development task to directly display native ICON data
- Can we get operational display performance for generic grid data?
- A generic display should help NWP developers and NWP model evaluators
  in this use case: visualisation performance is less critical
- Recently we got new prototype based on the NinJo formula layer
  First results look promising: display times of ~5-20 sec
- We are confident that parallel executions and pre-calculation (search tree) can improve timings
- Another result: Display of global grid and regional grid require different search algorithms and dedicated optimisations (<-> backward mapping case)

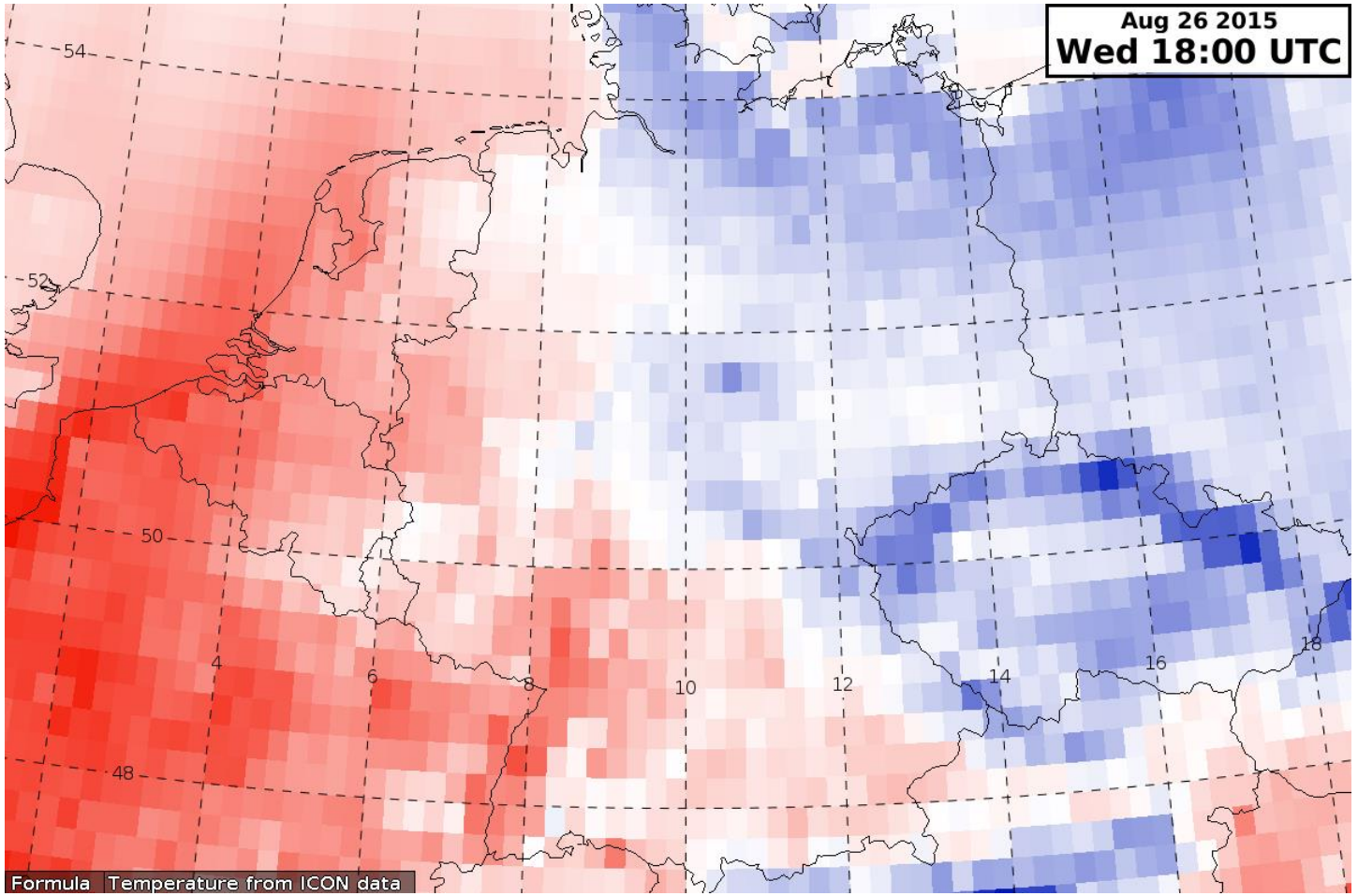
T2M: ICON Global (native 13km grid)



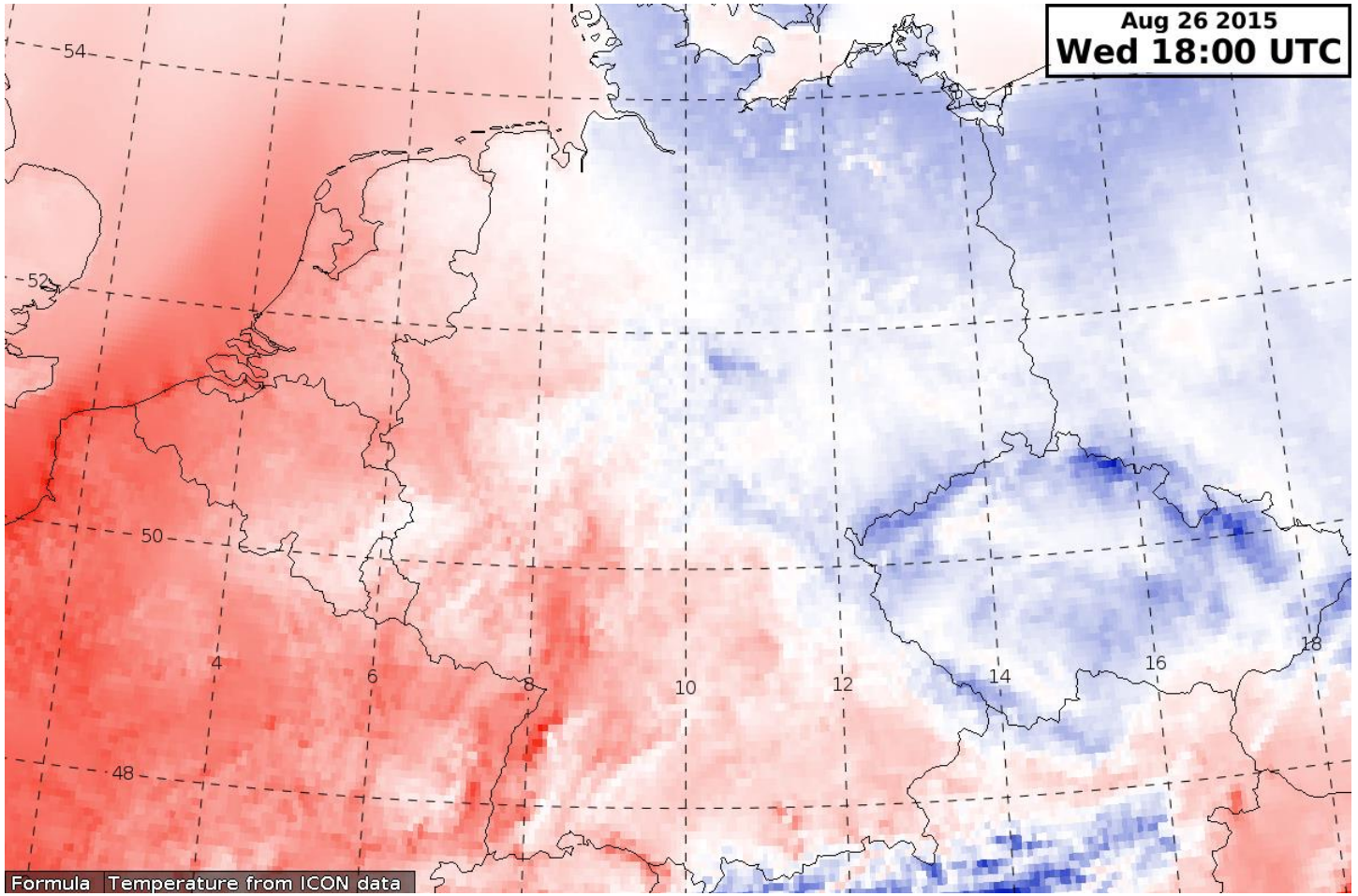
T2M: ICON-EU Nest (native 6.5 km grid)



T2M: ICON Global (regular 0.25° grid)



T2M: ICON-EU Nest (regular 0.0625° grid)

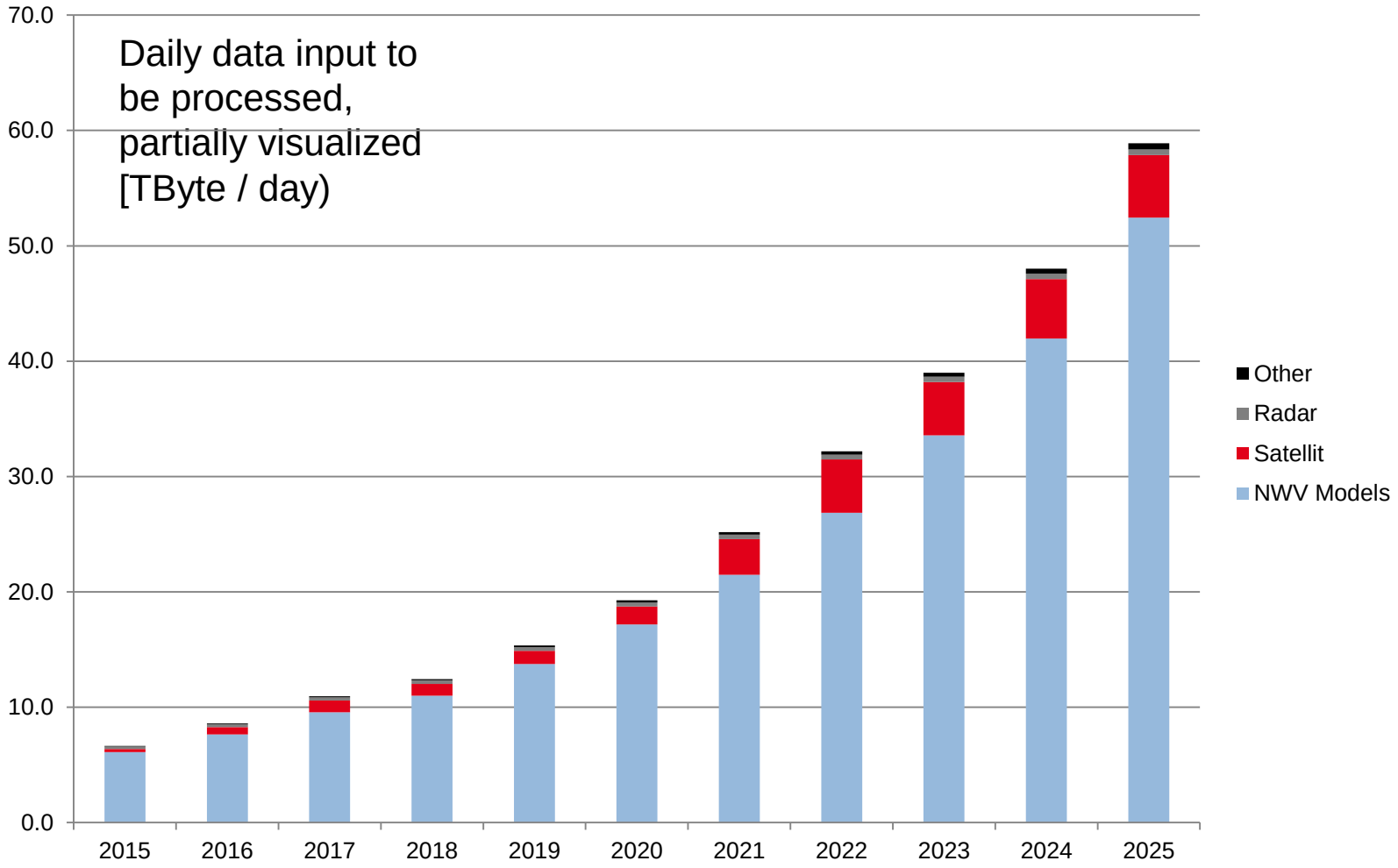


Available NWP data vs. operationally visualised data

- NWP models have to use: fine grids, high precision data types and use max. available resources on the super computer
- Operational visualized NWP data should not always take 1:1 NWP output
  - Forecast quality decreases by time
    - Available precision can pretend false truth (signal location & strength)
  - Short term FC +1h use same data representation as long term FC +180h
  - Forecasters anyway do not trust a deterministic “single pixels” in time
  - Increased use “Ensemble Prediction Systems” data
- Still deterministic high resolution NWP data uses significant amount of resources of our operational visualisation systems and processing chains
- Generic grids need to be supported by visualisation systems.

  Is here a hidden opportunity to reduce data volumes?

Expected increase of input data (NinJo & other applications)



Further estimates & rumours on volume increases

- Satellite data (MTG) → ~100x
- Radar data (volume scans, HR CAPPIs, 250m composites) → ~6x
- Strategic planning already includes next NWP refinement steps → ~6x
- NowCast data & seamless prediction (IVS) → unknown !!!
- Ensemble products → usage of some single ensemble members (~20-40x)
- DWD‘s central archive → Petabytes (needs dynamic load)
- Climate models → unknown !!!

Contact:

Marcus Werner
Referat FE ZE
Frankfurter Str. 135
63037 Offenbach

E-Mail: marcus.werner@dwd.de
Tel.: +49 (0) 69 / 8062 - 2076