Towards object-based verification in SINFONY

ICCARUS
28.02.2018

Michael Hoff
E-Mail: michael.hoff@dwd.de

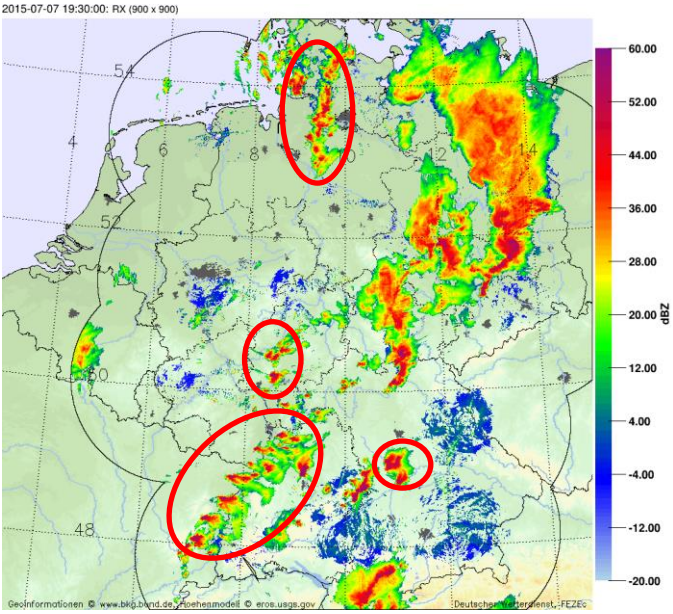
Project-team: **Ulrich Blahak, Kathrin Wapler, Roland Potthast, Kathleen Helmert, Alberto De Lozar, Axel Seifert, Elisabeth Bauernschubert, Christian Welzbacher, Manuel Werner, Robert Feger, Lilo Bach, Felix Fundel, Martin Rempel, Markus Junk, Lisa Neef, Kathrin Feige, Markus Schultze, Rafael Posada**

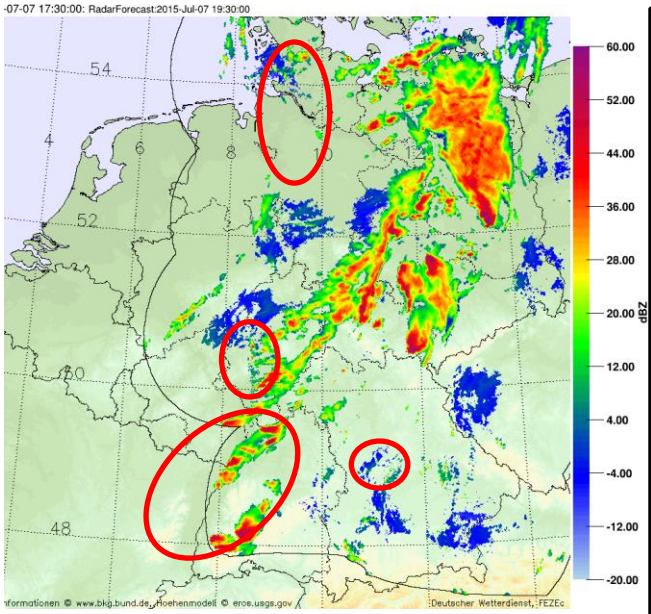
Motivation: improve the support of our forecasters for their weather warnings

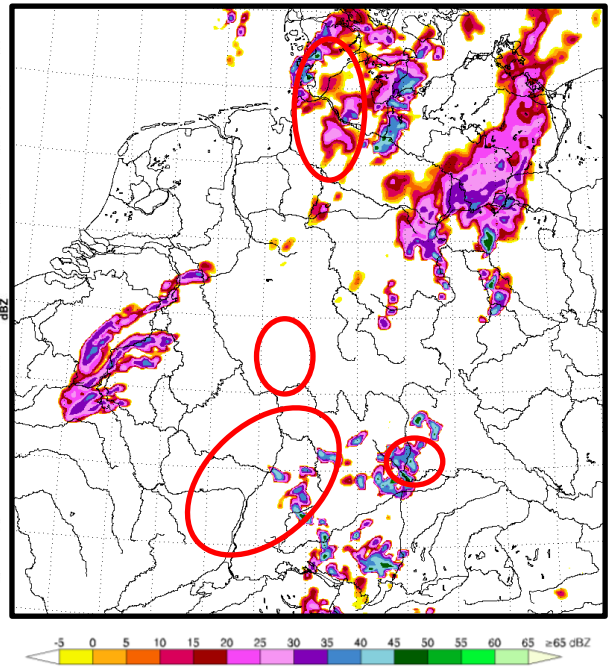
Measured Cells @ **19:30**

Nowcast **for 19:30 from 17:30**

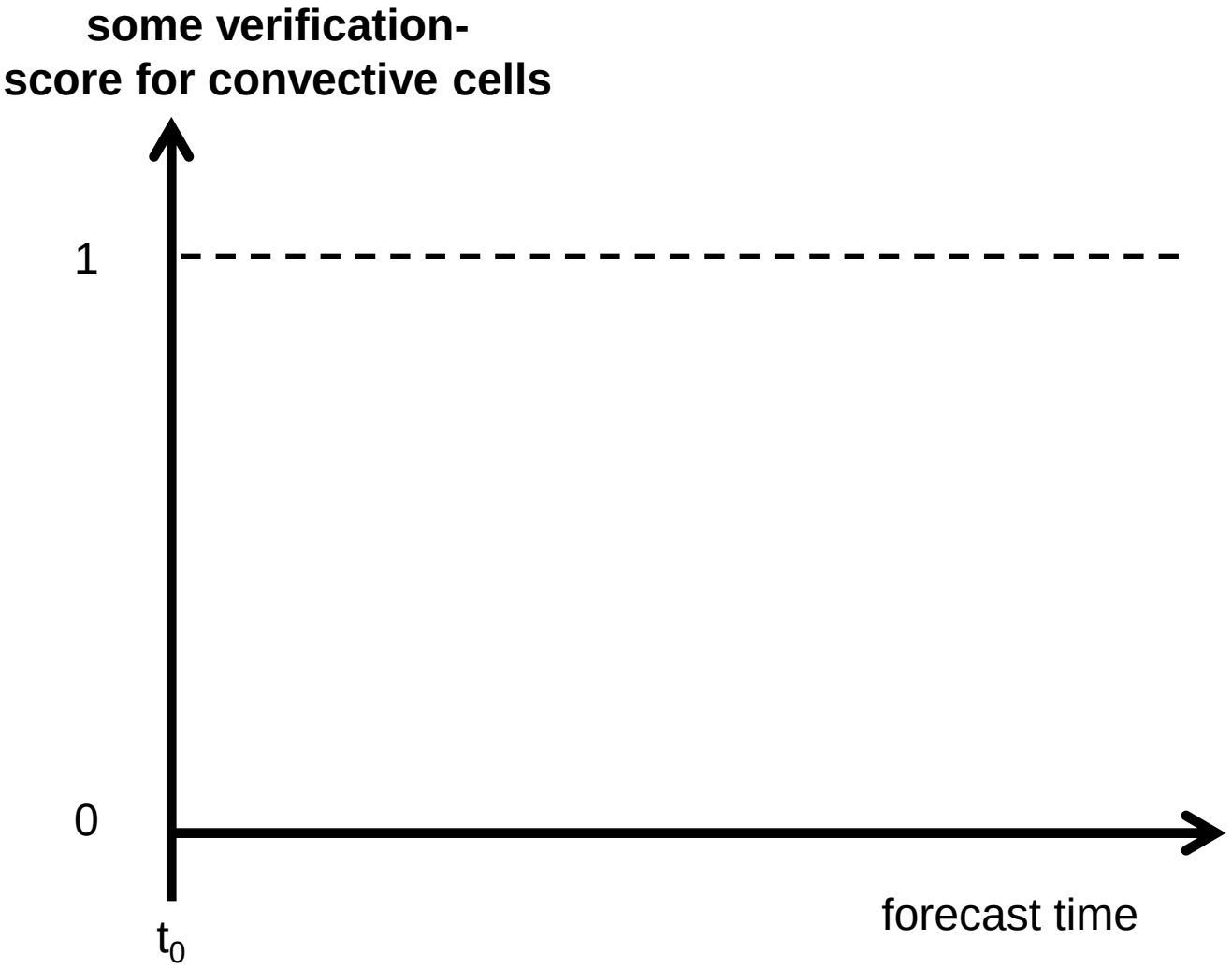
Model **for 19:30 from 15:00**



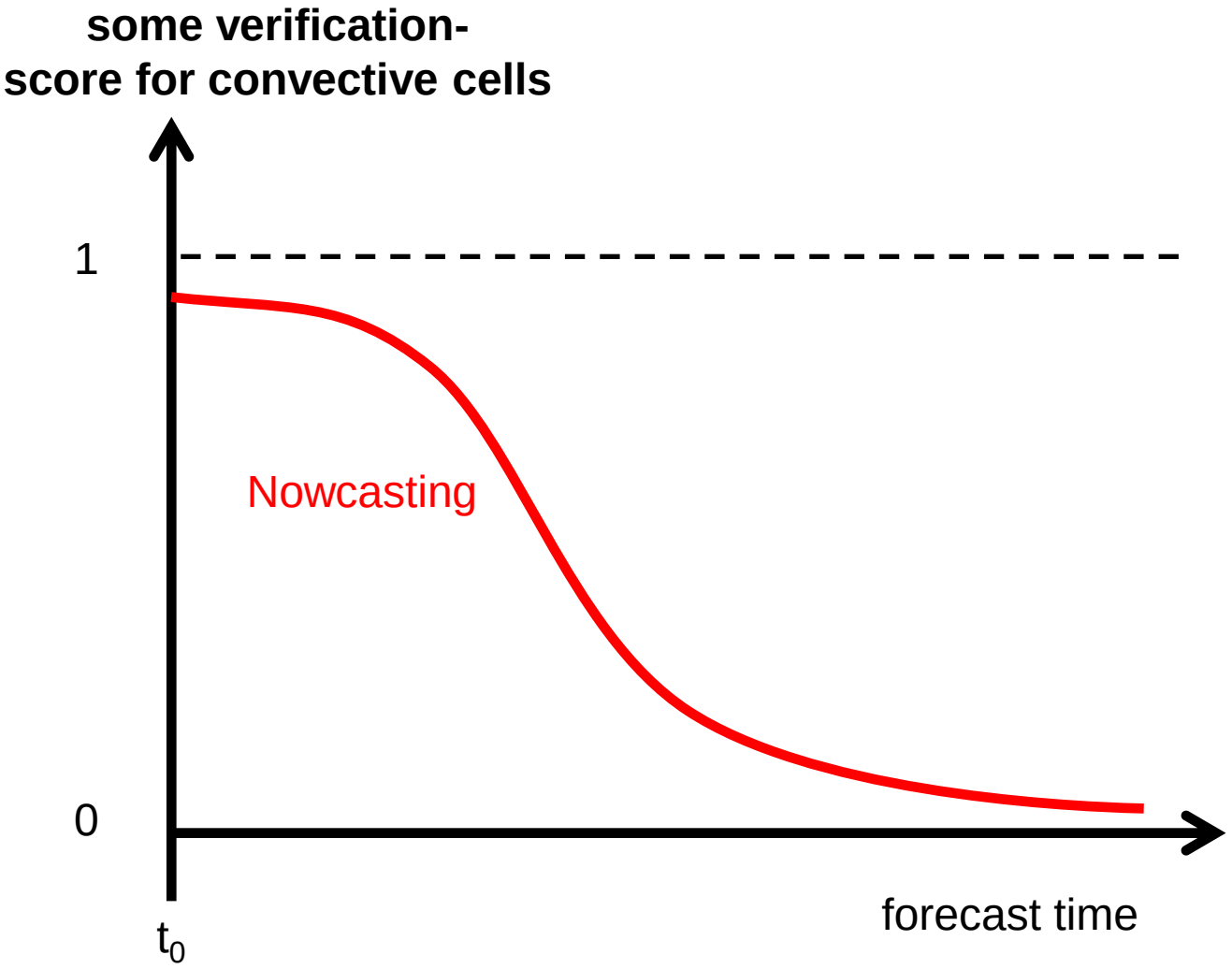




The basic idea of SINFONY

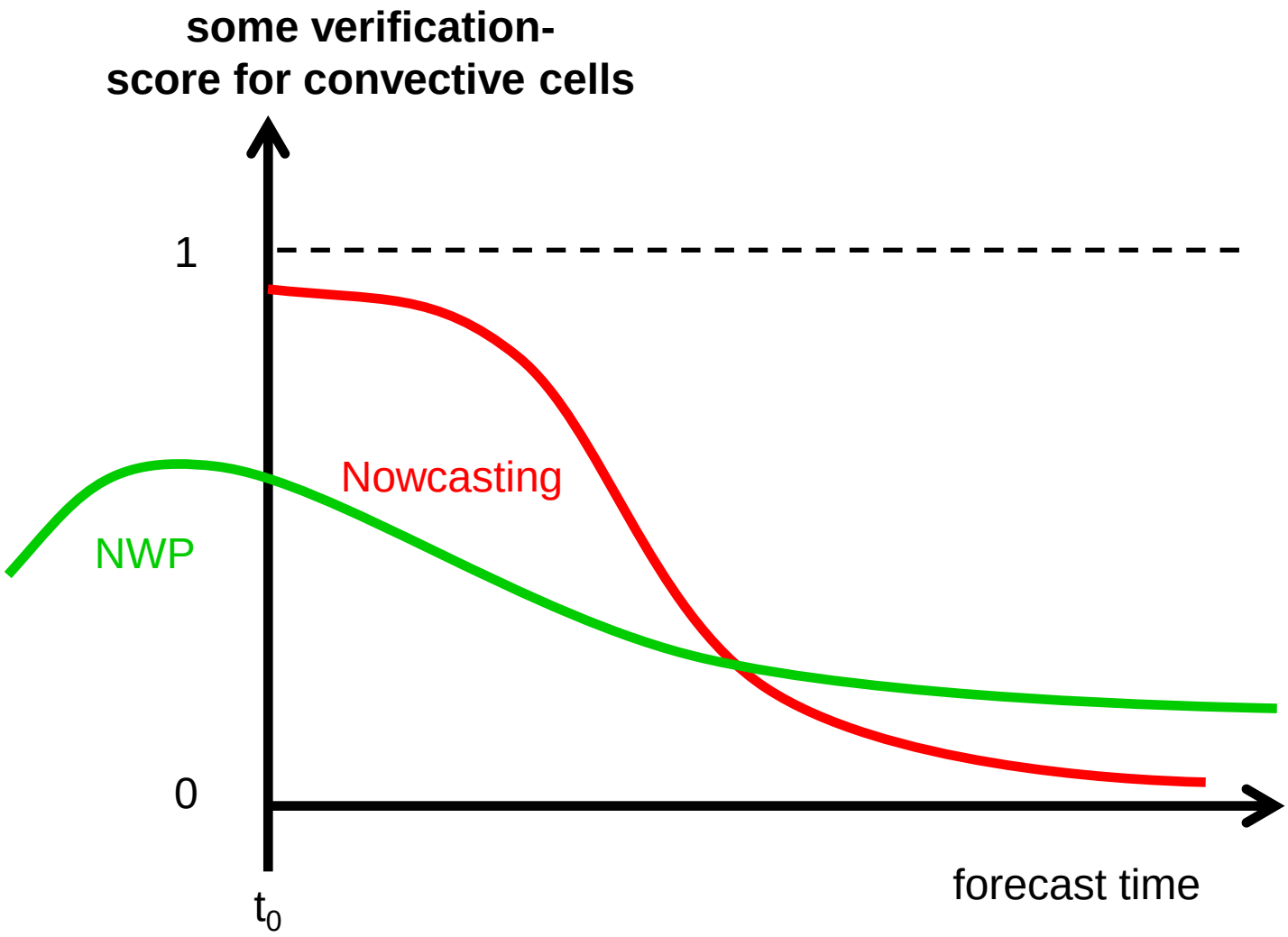


The basic idea of SINFONY

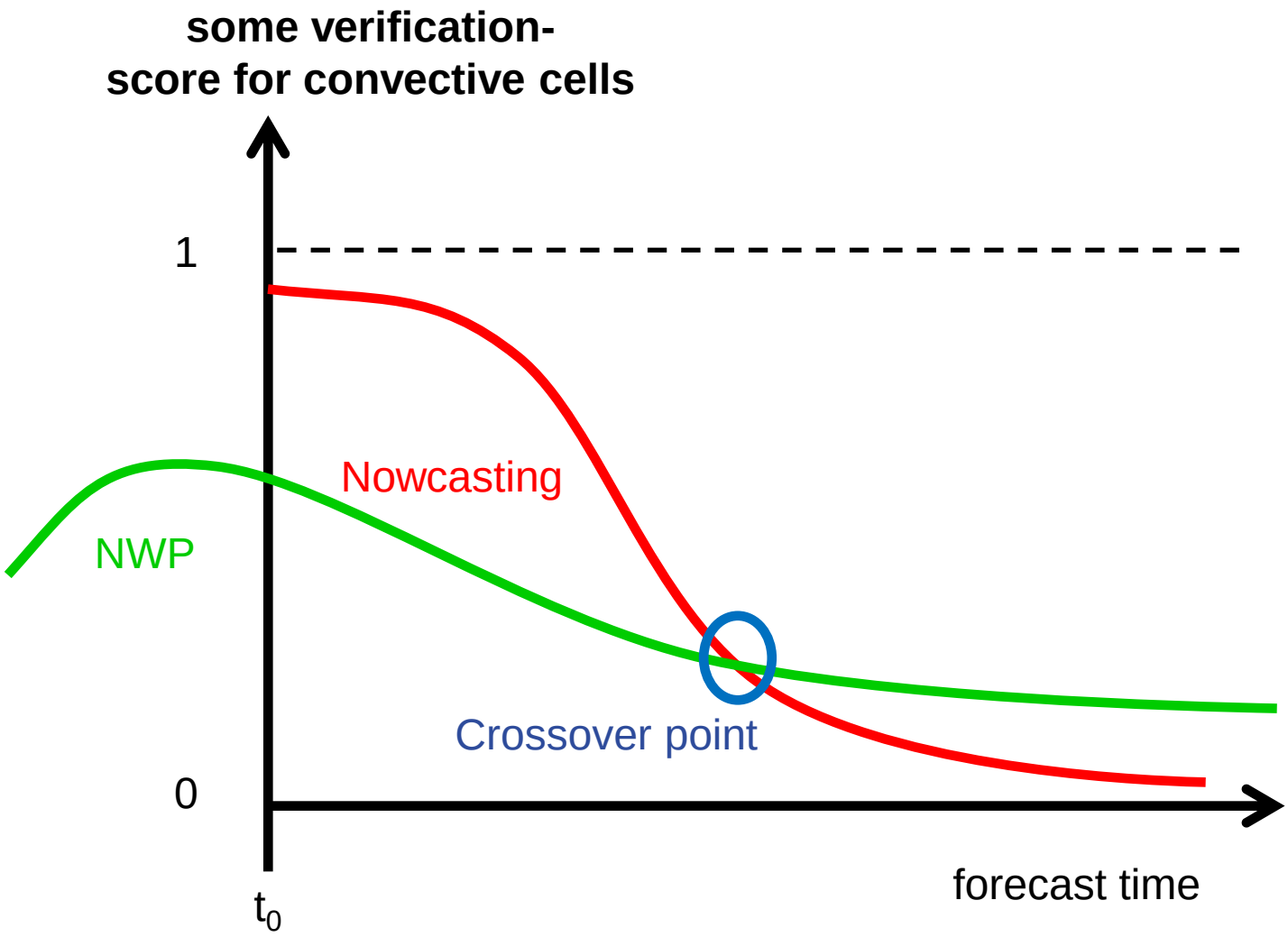


The basic idea of SINFONY

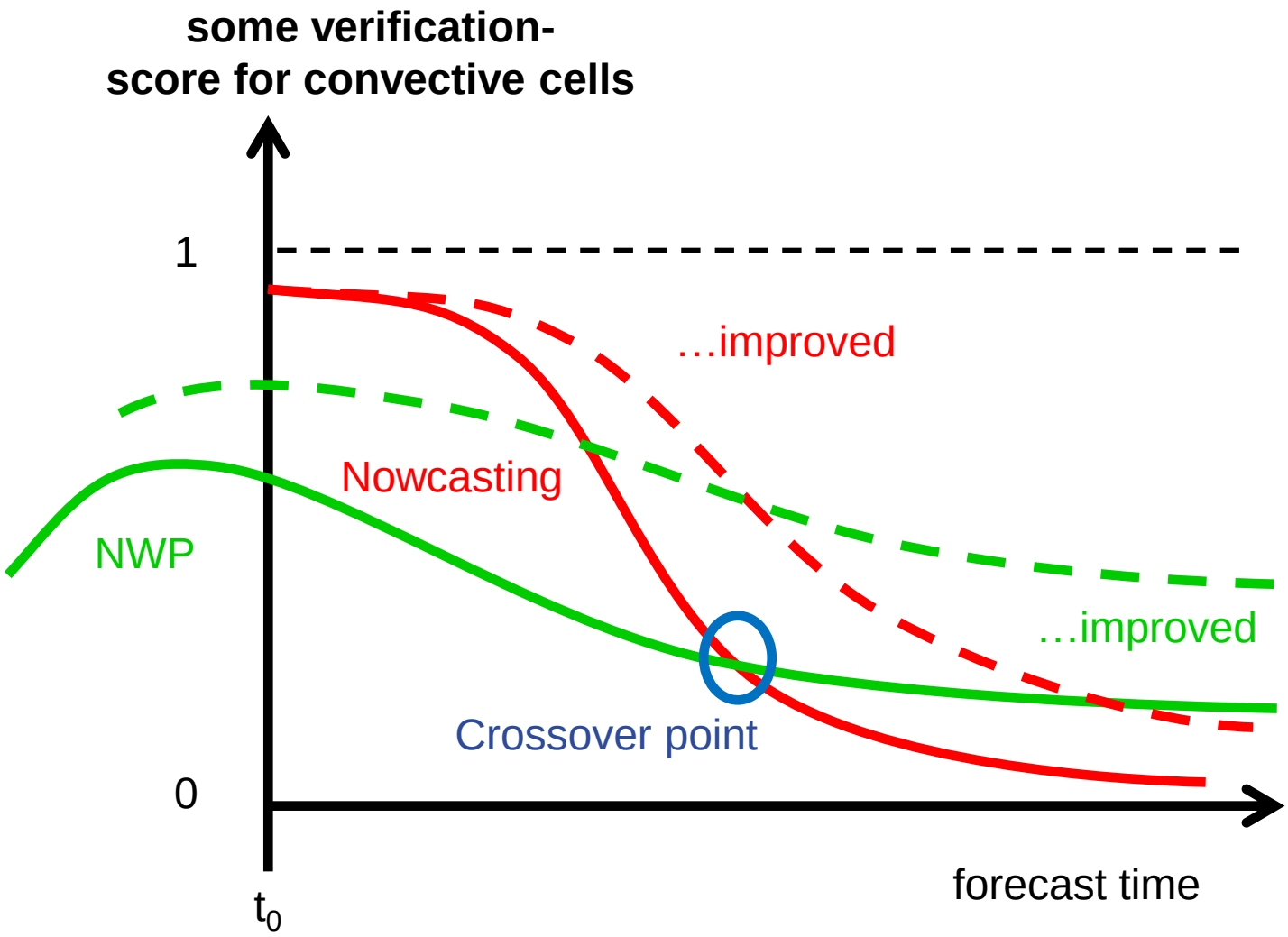
Deutscher Wetterdienst
Wetter und Klima aus einer Hand



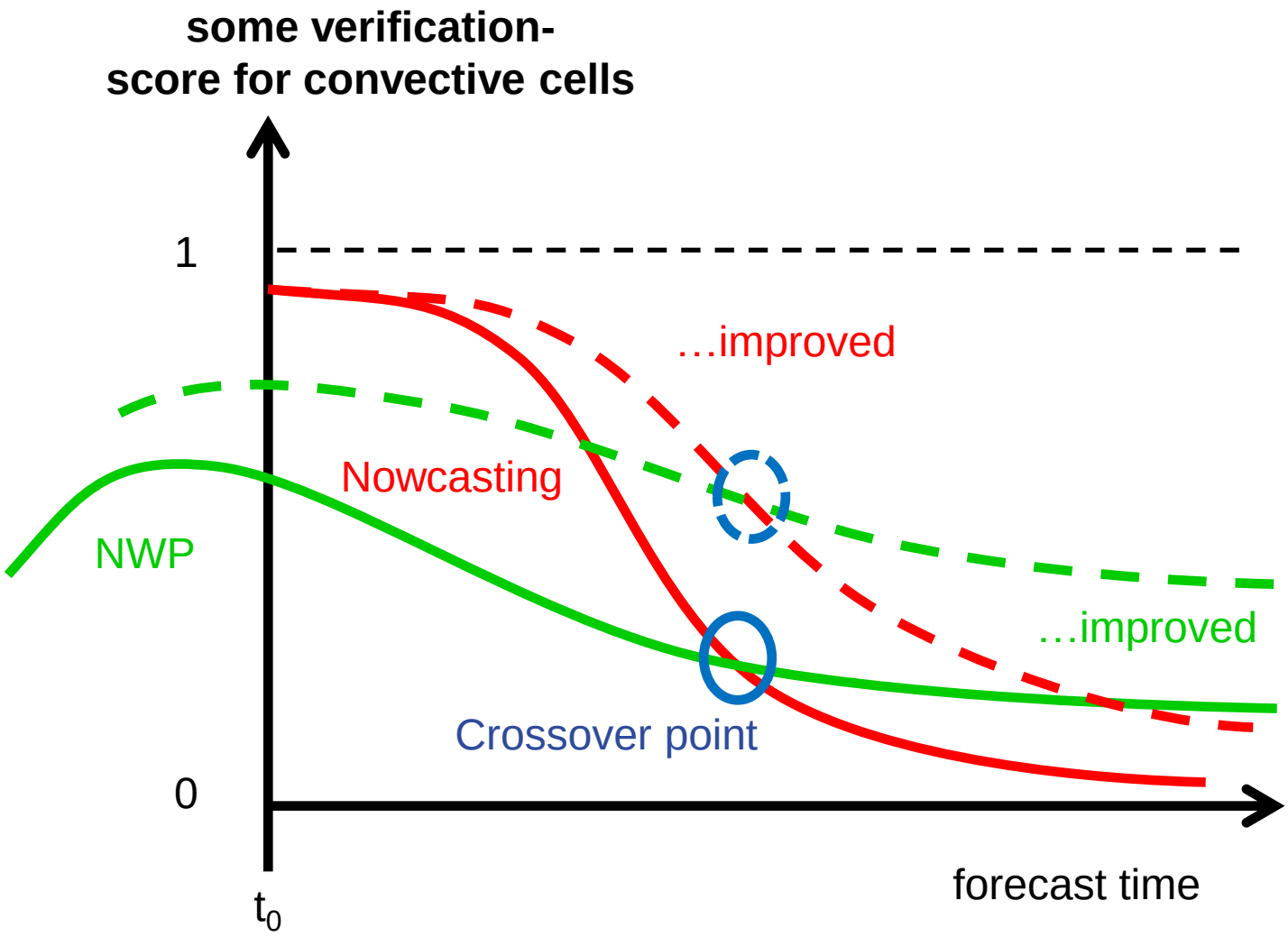
The basic idea of SINFONY



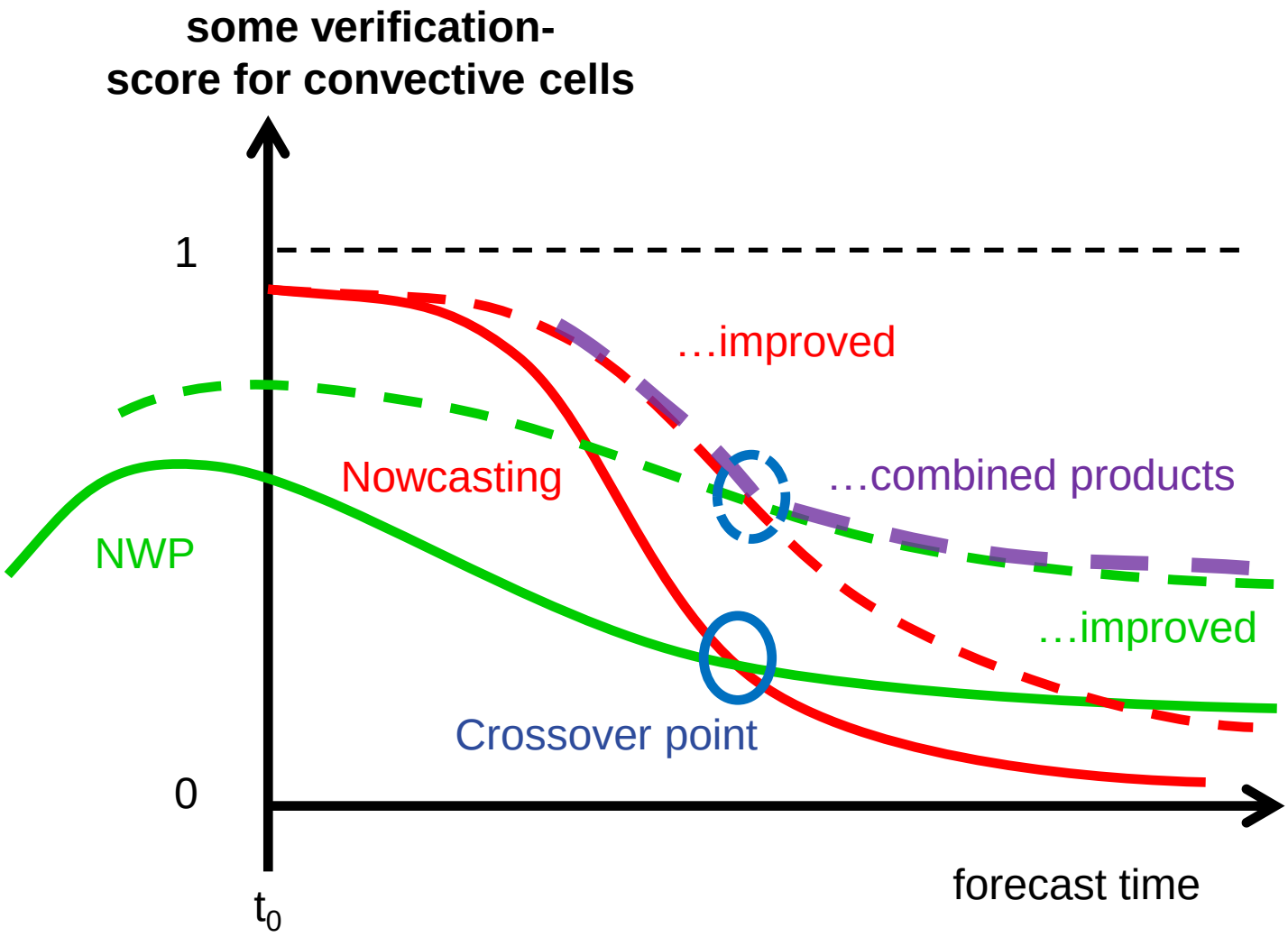
The basic idea of SINFONY



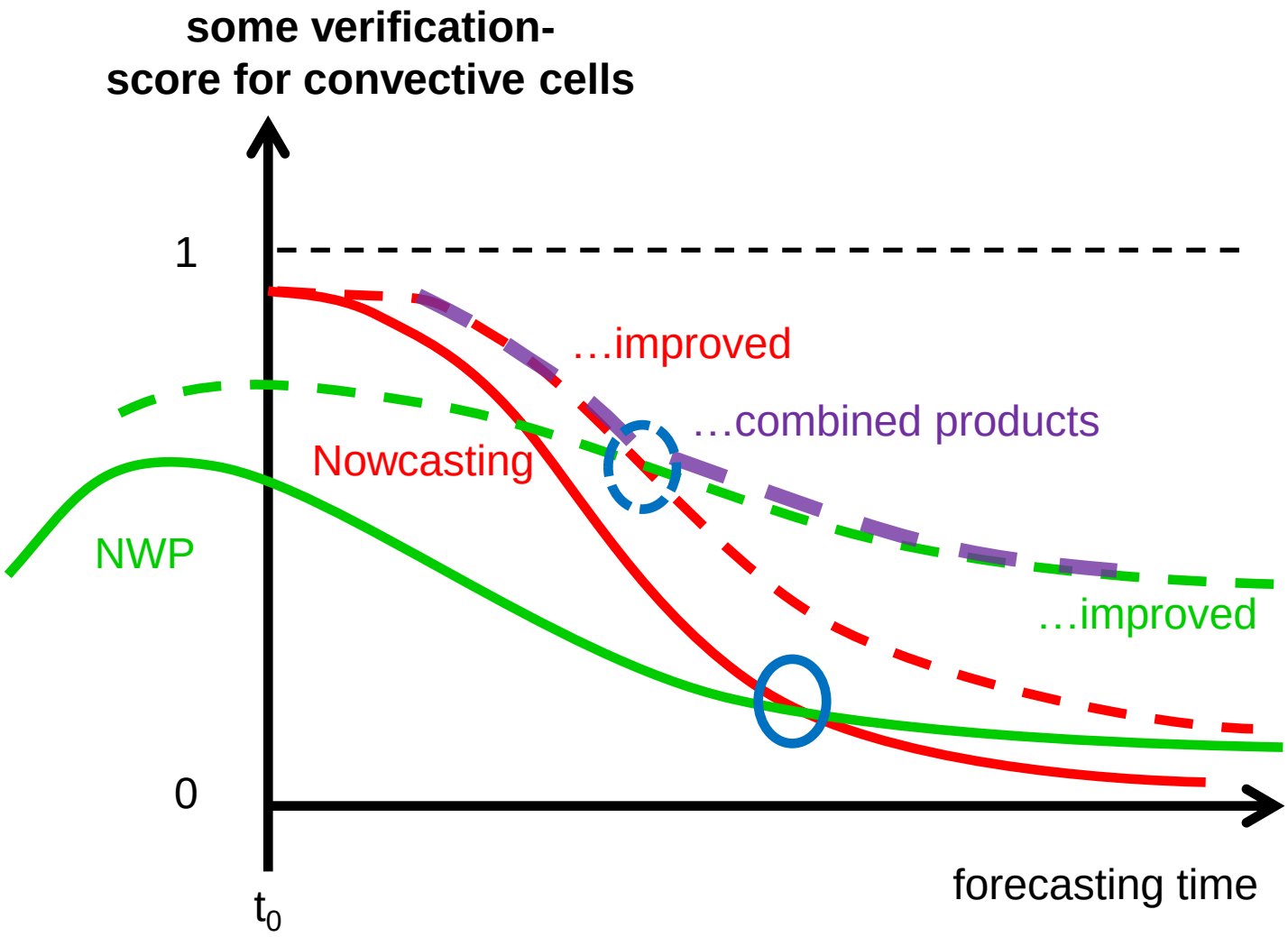
The basic idea of SINFONY



The basic idea of SINFONY

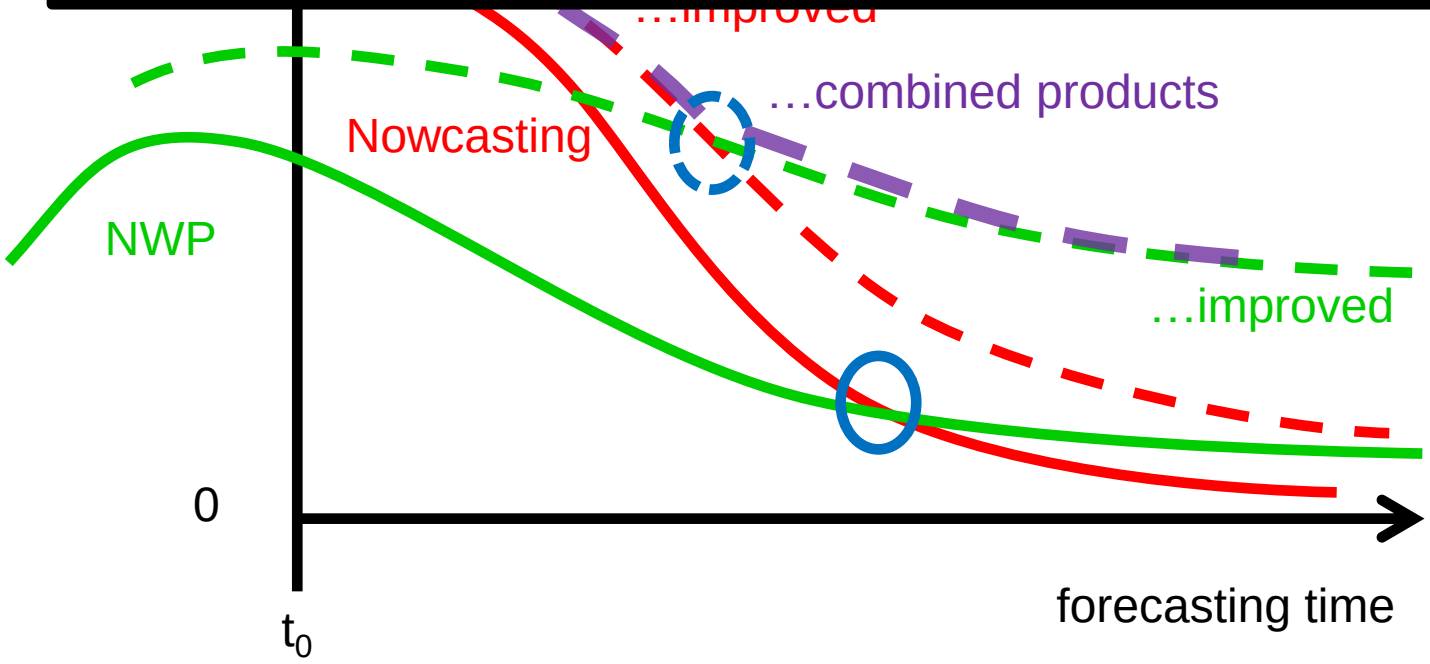


The basic idea of SINFONY



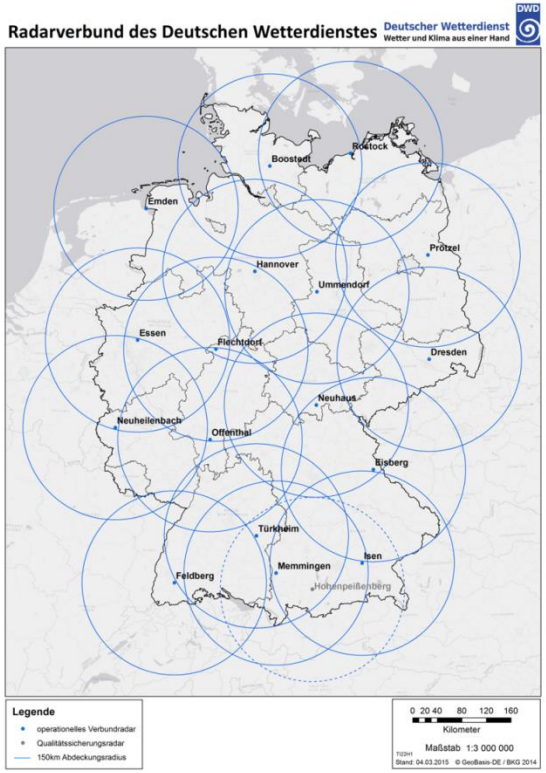
The basic idea of SINFONY

- It is an open issue at which leadtime the crossover point will be after improvements



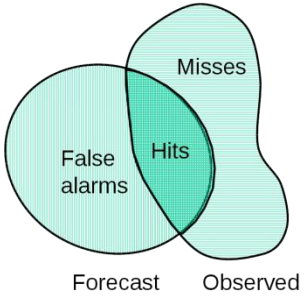
Only radar reflectivity composites for verification

- Testperiod: 26.05.2016 – 25.6.2016
- NWP model: COSMO-DE-EPS
  - quasi-operational setup (reference)
  - EMVORADO[1] to simulate reflectivities
  - ensemble + determinstic (output every 5 min)
- Nowcast
  - KONRAD3D using optical flow
  - only deterministic (output every 5 min)
- Observation (RX-composite)
- all data on the same grid (900x900; 1km resolution)
- Caution: reflectivities from COSMO (2.8km) interpolated to 1km composites
- set reflectivities ≤ 0 dBZ to zero (no precipitation)



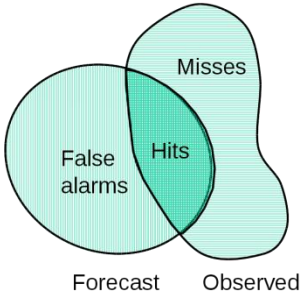
[1] Zeng *et al.*, 2016

Dichotomous verification: (rain: yes/no; thresholding)



forecast \ observed	yes	no	
yes	h (hits)	f (false alarms)	$h + f$
no	m (misses)	z (zero;corr. rejections)	$m + z$
	$h + m$	$f + z$	$N = h + f + m + z$

Dichotomous verification: (rain: yes/no; thresholding)

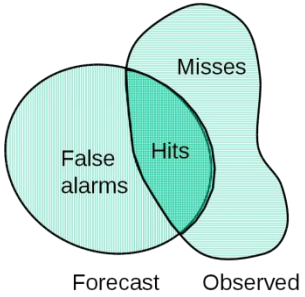


forecast \ observed	yes	no	
yes	h (hits)	f (false alarms)	$h+f$
no	m (misses)	z (zero;corr. rejections)	$m+z$
	$h+m$	$f+z$	$N=h+f+m+z$

obs events

$$s = \frac{h+m}{N} \qquad [0,1]$$

Dichotomous verification: (rain: yes/no; thresholding)



forecast \ observed	yes	no	
yes	h (hits)	f (false alarms)	$h+f$
no	m (misses)	z (zero;corr. rejections)	$m+z$
	$h+m$	$f+z$	$N=h+f+m+z$

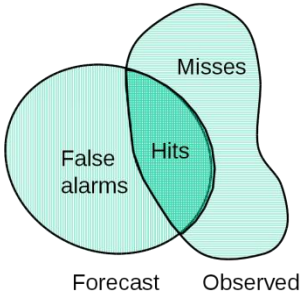
$$r = \frac{h+f}{N} \qquad [0,1]$$

fcst events

obs events

$$s = \frac{h+m}{N} \qquad [0,1]$$

Dichotomous verification: (rain: yes/no; thresholding)



forecast \ observed	yes	no	
yes	h (hits)	f (false alarms)	$h+f$
no	m (misses)	z (zero;corr. rejections)	$m+z$
	$h+m$	$f+z$	$N=h+f+m+z$

$$r = \frac{h+f}{N} \qquad [0,1]$$

fcst events

obs events

$$s = \frac{h+m}{N} \qquad [0,1]$$

$$BIAS = \frac{r}{s} = \frac{h+f}{h+m} \qquad [0,\infty]$$

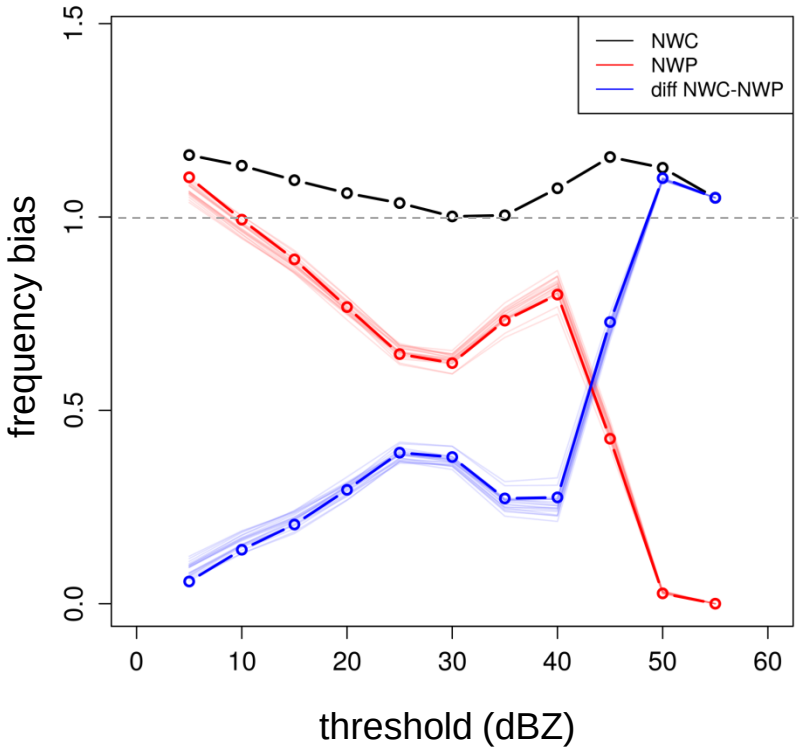
BIAS and ETS as a function of threshold

$$ETS = \frac{h - h_r}{h + f + m - h_r} \quad [-\frac{1}{3}, 1]$$

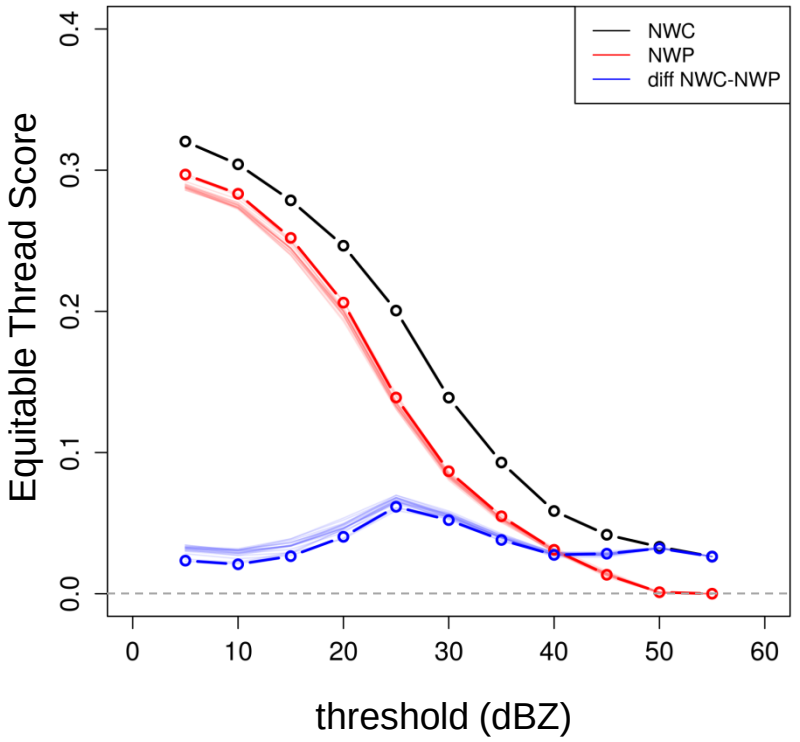
Summary scores over all days and all lead times

perfect = 1

BIAS (init: 15 UTC)

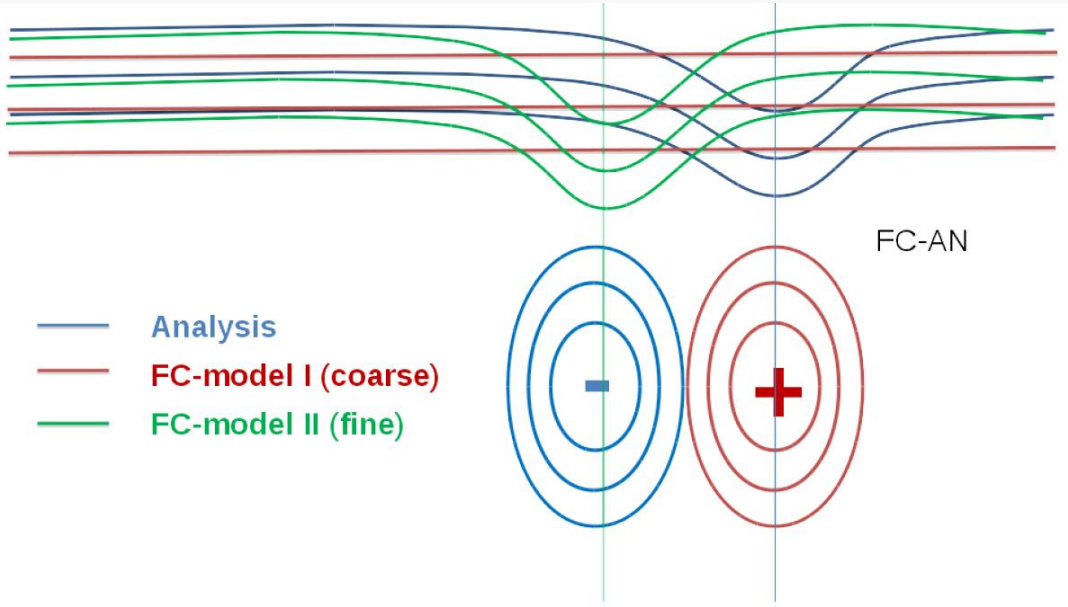


ETS (init: 15 UTC)



The problem of pixel-based verification:

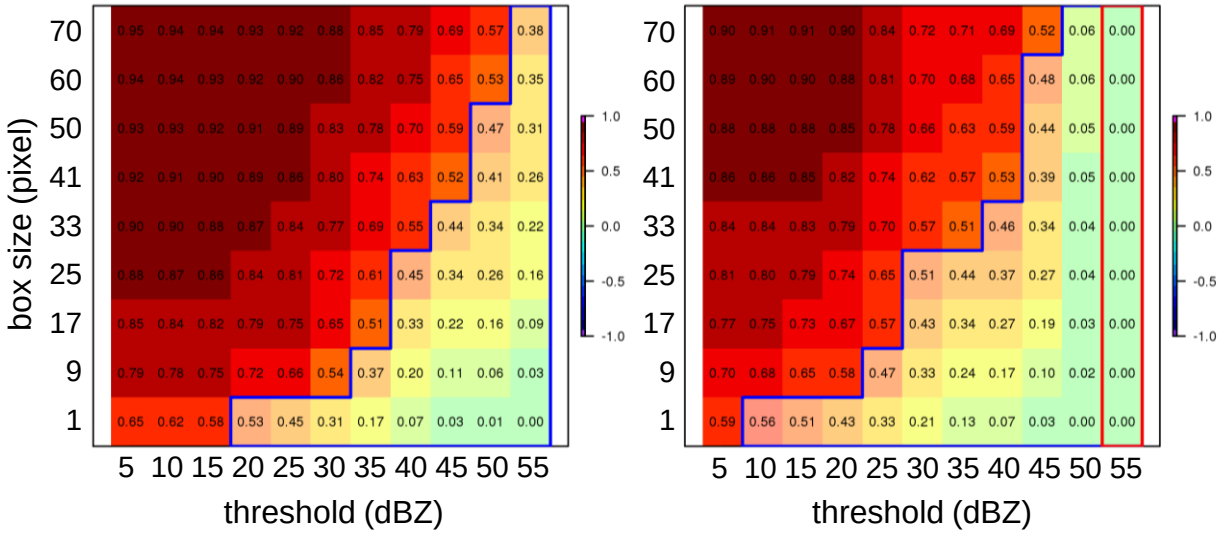
- small displacements are doubly penalized in traditional scores
- leads to unrealistic low scores
- effect larger, the finer the model grid resolution



Fraction Skill Score (Roberts & Lean, 2008)

Fraction Skill Score (Roberts & Lean, 2008)

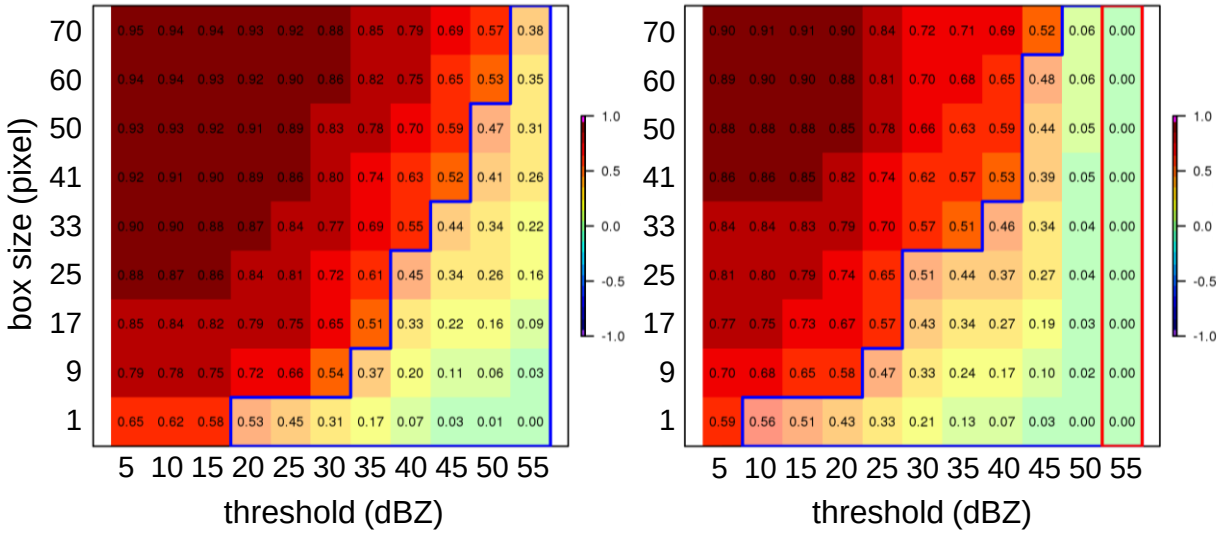
Nowcast:
15UTC + 1h



NWP:
15UTC + 1h

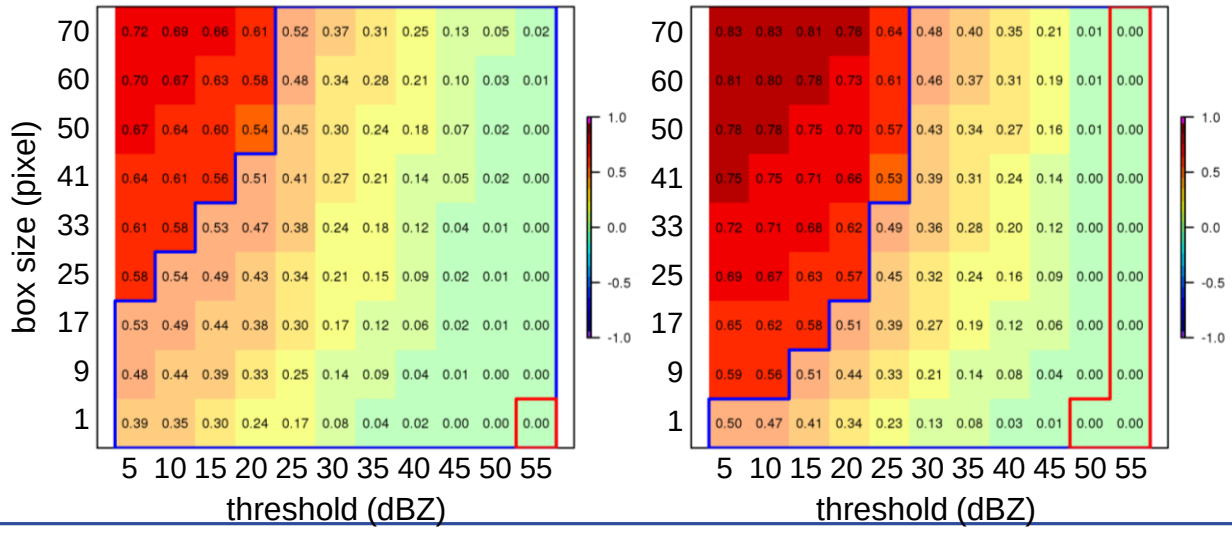
Fraction Skill Score (Roberts & Lean, 2008)

Nowcast:
15UTC + 1h



NWP:
15UTC + 1h

Nowcast:
15UTC + 4h

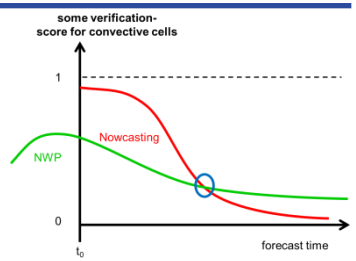


NWP:
15UTC + 4h

Nowcast vs. NWP – Crossover point

(init: 15 UTC, *threshold* = 25dBZ)

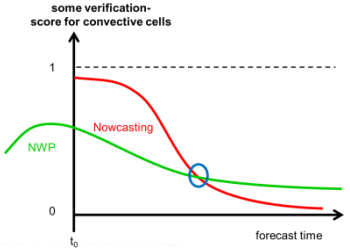
Summarized over all days

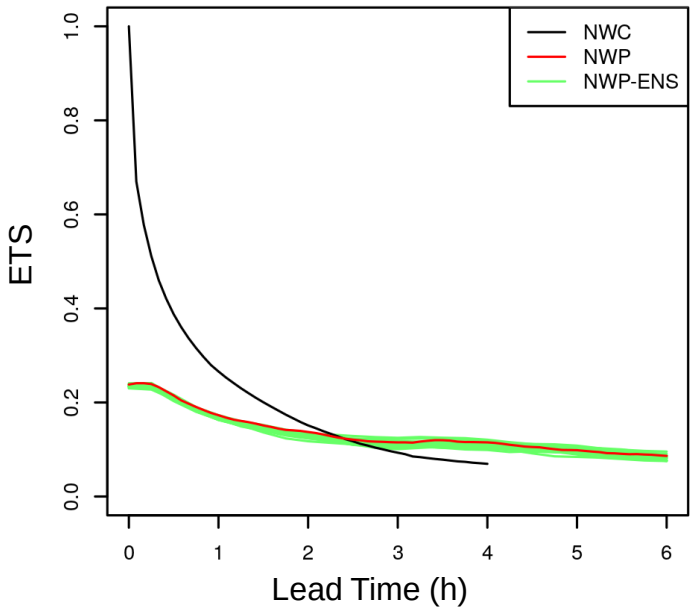


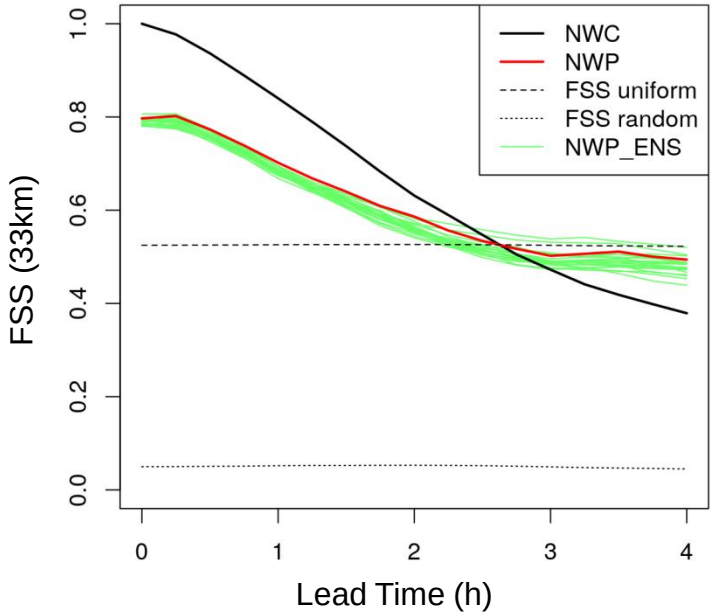
Nowcast vs. NWP – Crossover point

(init: 15 UTC, *threshold* = 25dBZ)

Summarized over all days





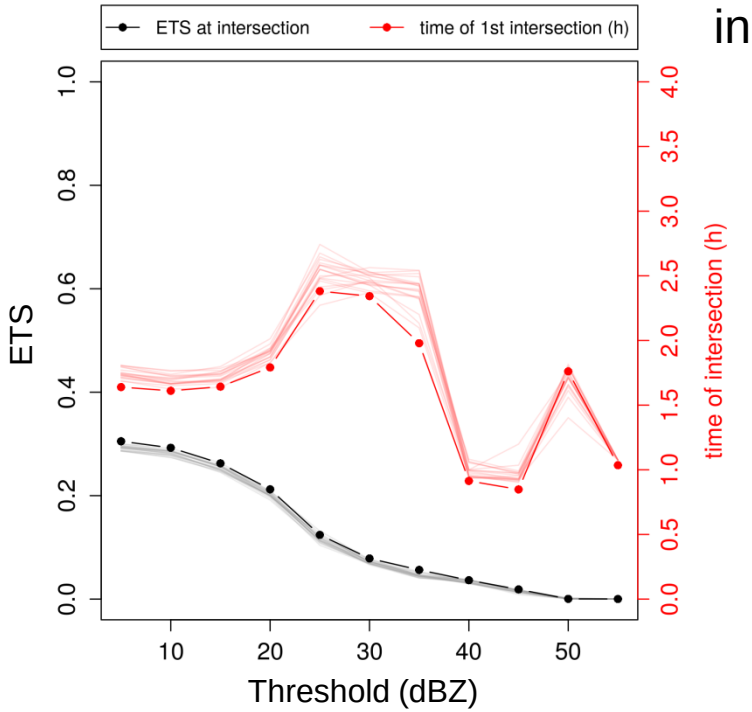


- ETS: pixel-based
- resolution: 1km
- ETS ~ 0.2 common value for precipitation forecasts

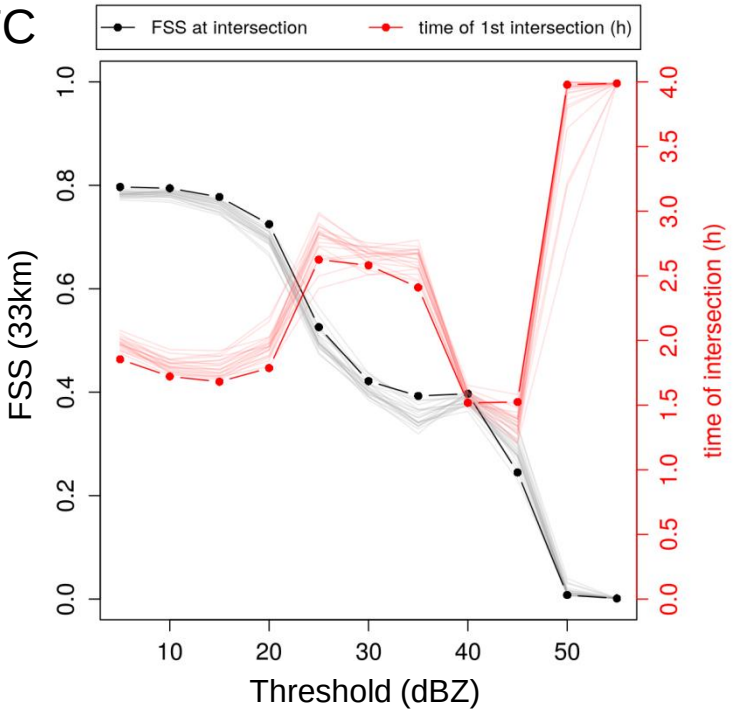
- FSS: neighbourhood-based
- smoothing with box size 33km
- FSS > ETS, because small displacements are not penalized

The crossover point as a function of threshold

→ Crossover point: lead time when NWP becomes better than nowcast

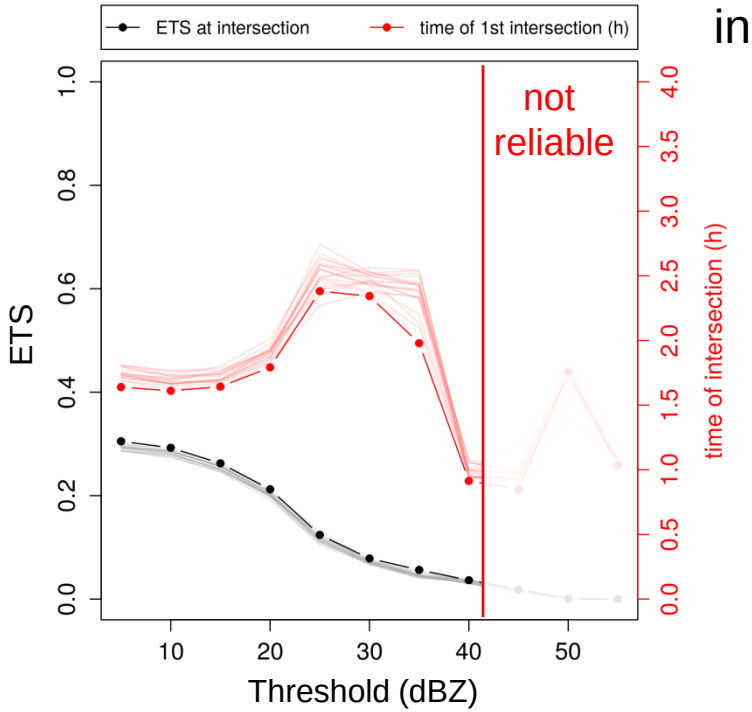


init: 15 UTC

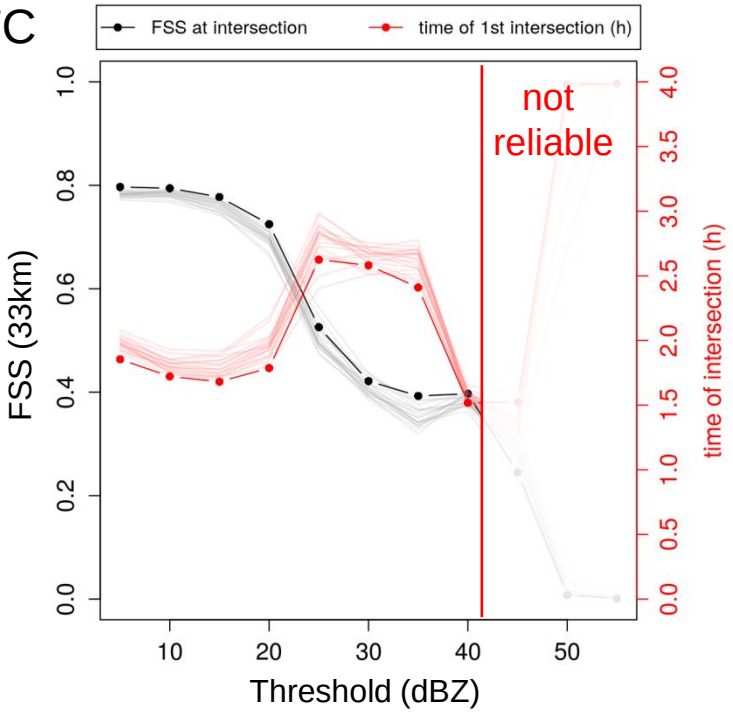


The crossover point as a function of threshold

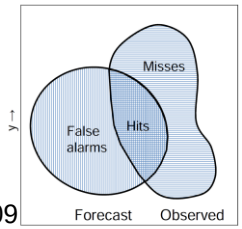
→ Crossover point: lead time when NWP becomes better than nowcast



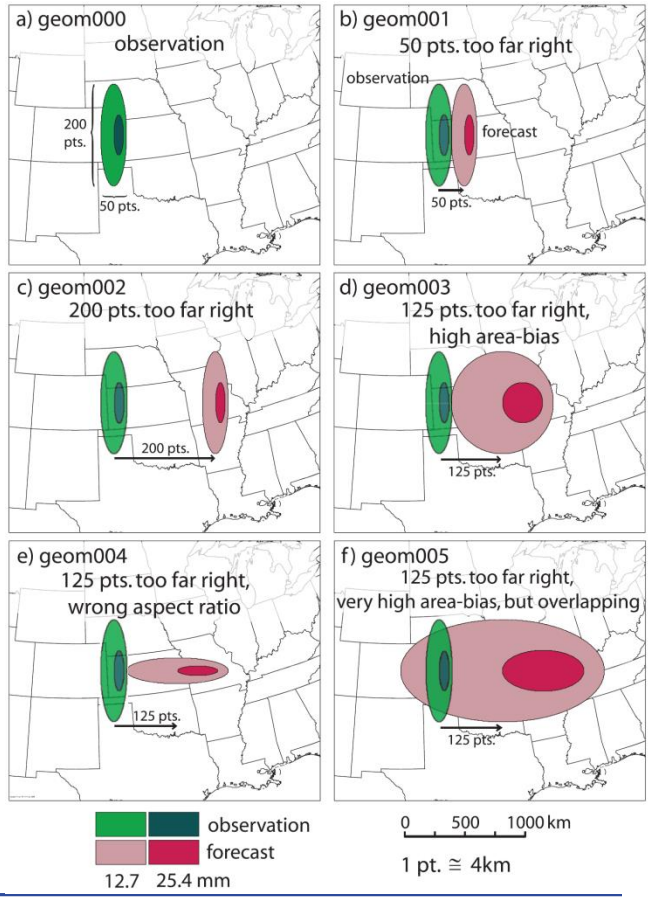
init: 15 UTC



Why do we want to use object-based verification?



Ahijevych *et al.* 2009



- so far scarce information about source of error
  - What went wrong? What right?
  - How close is the fcst to the observation?
  - Is the forecast realistic?
  - How can I improve the forecast?
- goals of object-based verification:
  - identify coherent spatial/temporal structures (closer to subjective perspective)
  - uncertainty in location and time
  - more information about errors in physical terms
  - reduction of data volume

The three main steps of object-based verification

- step 1: Define entity by thresholding (possibly smoothing)
- step 2: „pattern matching“ (most difficult part)
  - Minimum absolute error; max correlation; max overlap; centroid distance, …
- step 3: Analysing output (scatterplots, hits, misses, etc…)

Step 1: Two ways of how we plan to identify objects

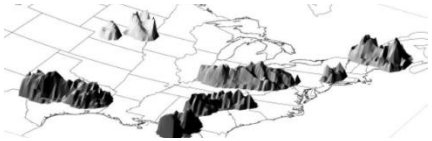
- KONRAD3D (developed by Manuel Werner (DWD))
  - based on strong convective cells (>36dBZ)
  - 2D cell identification (Johnson *et al*., 1998)
  - more flexible with adaptive thresholding scheme (Hering *et al*., 2004)
- MODE (Davies *et al*. 2006a,b;2009)
  - convolution filter via disk to obtain contiguous areas
  - thresholding to obtain a mask
  - keeping the original field for statistical analyses

Objects must be as simple as possible to get an unambiguous interpretation

Some attributes we want to take into account

Objects must be as simple as possible to get an unambiguous interpretation

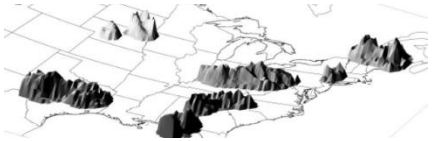
Some attributes we want to take into account

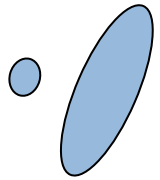


➔ intensity (stochastic)

Objects must be as simple as possible to get an unambiguous interpretation

Some attributes we want to take into account

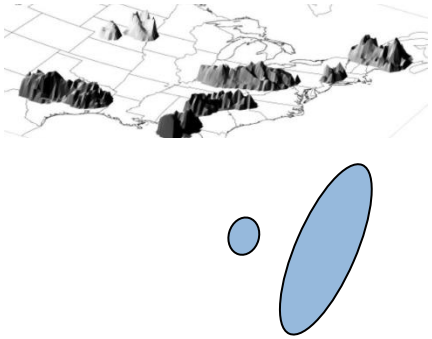




- ➔ intensity (stochastic)
- ➔ area (different scales, different dynamical processes and predictability)

Objects must be as simple as possible to get an unambiguous interpretation

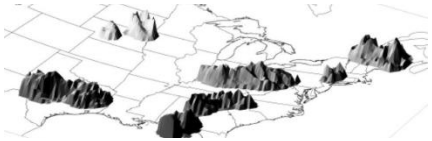
Some attributes we want to take into account

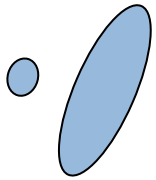


- intensity (stochastic)
- area (different scales, different dynamical processes and predictability)
- centroid (lat/lon, x/y) → systematic location bias in fcst

Objects must be as simple as possible to get an unambiguous interpretation

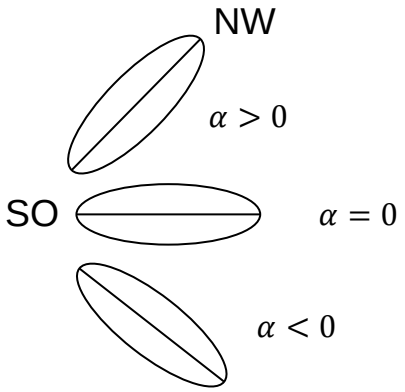
Some attributes we want to take into account





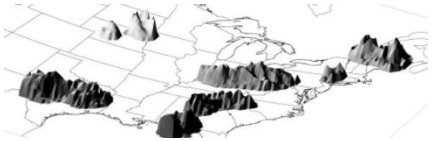
- intensity (stochastic)
- area (different scales, different dynamical processes and predictability)
- centroid (lat/lon, x/y) → systematic location bias in fcst
- axis angle → line through centroid for „best“

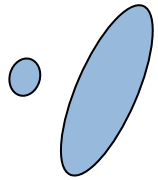
characterization of objects

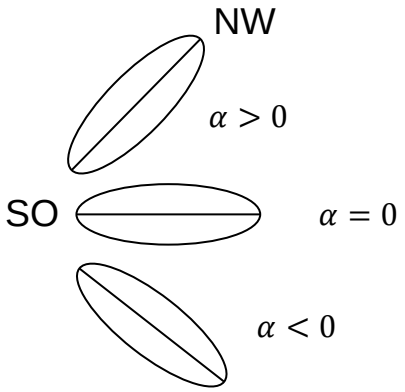


Objects must be as simple as possible to get an unambiguous interpretation

Some attributes we want to take into account





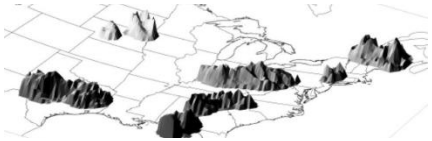


- ➔ intensity (stochastic)
- ➔ area (different scales, different dynamical processes and predictability)
- ➔ centroid (lat/lon, x/y) → systematic location bias in fcst
- ➔ axis angle→ line through centroid for „best“

  characterization of objects
- ➔ aspect ratio (ellipses)

Objects must be as simple as possible to get an unambiguous interpretation

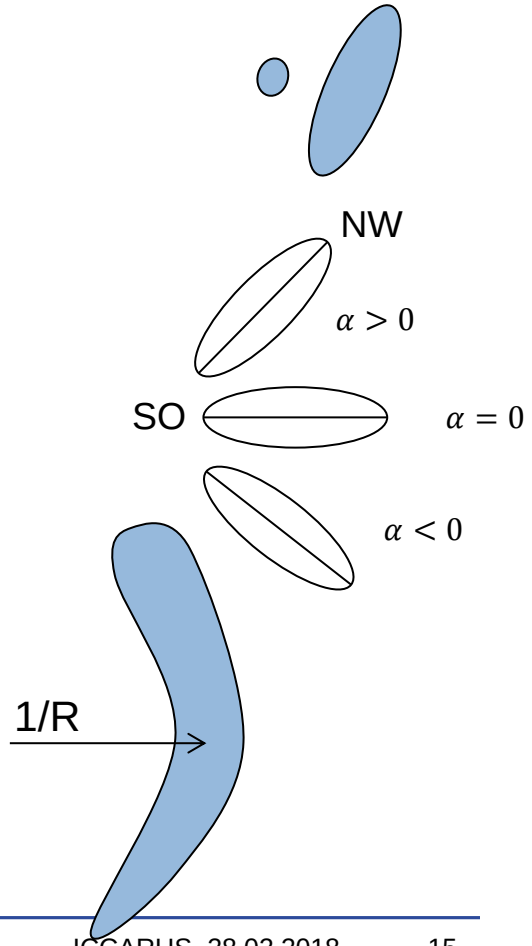
Some attributes we want to take into account



- intensity (stochastic)
- area (different scales, different dynamical processes and predictability)
- centroid (lat/lon, x/y) → systematic location bias in fcst
- axis angle → line through centroid for „best“

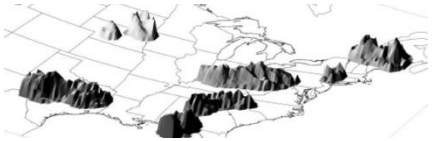
  characterization of objects
- aspect ratio (ellipses)
- curvature → fit circle, deviation from straightness.

  Helpful for characterisation of convective systems

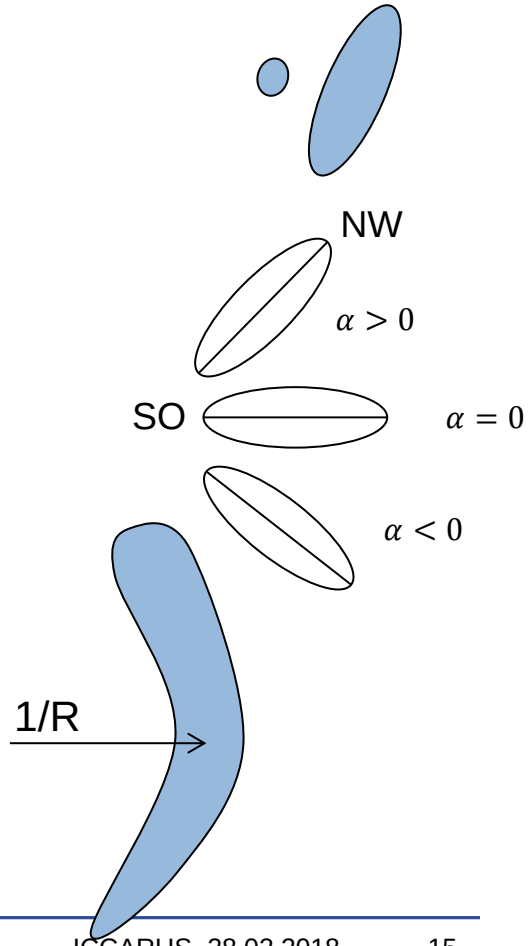


Objects must be as simple as possible to get an unambiguous interpretation

Some attributes we want to take into account



- intensity (stochastic)
- area (different scales, different dynamical processes and predictability)
- centroid (lat/lon, x/y) → systematic location bias in fcst
- axis angle → line through centroid for „best“ characterization of objects
- aspect ratio (ellipses)
- curvature → fit circle, deviation from straightness. Helpful for characterisation of convective systems
- (ingredients method, lightnings, etc.)



Step 2: Two ways of how we want to match objects

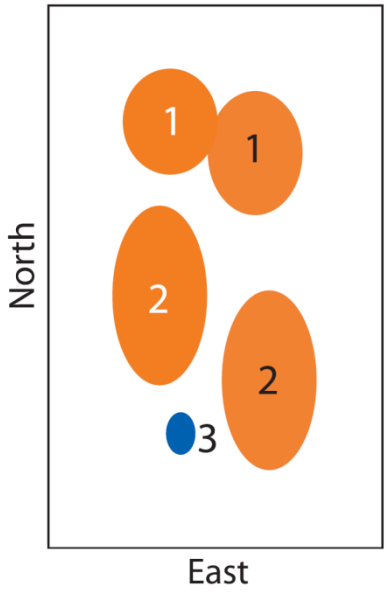
- KONRAD3D (Manuel Werner)
  - tracking (optical flow) and matching (max. overlap)
  - tracked nowcast objects get IDs and have counterparts in observation
  - Up to now: Only for nowcast available
- MODE (Davies *et al*. 2006a,b;2009)
  - „Total Interest“ for each pair of objects (Interest Matrix)
  - includes different weights for centroid distance, minimum boundary separation, orientation angle difference, area ratio, intersection area
  - matching if total interest above certain value (merging if multiple matches)

Step 3: Verification of identified matched objects

- Median of Maximum Interest (MMI) – MODE (Davies *et al*., 2009)
- applicable to both matched objects from MODE and matched objects from KONRAD3D (XML output file)

Step 3: Verification of identified matched objects

- Median of Maximum Interest (MMI) – MODE (Davies *et al*., 2009)
- applicable to both matched objects from MODE and matched objects from KONRAD3D (XML output file)



Interest Matrix

Forecast \ Observed	1	2
1	0.90	0.75
2	0.50	0.80
3	0.40	0.55

figure from Davies et al. (2009)

Step 3: Verification of identified matched objects

- Median of Maximum Interest (MMI) – MODE (Davies *et al.*, 2009)
- applicable to both matched objects from MODE and matched objects from KONRAD3D (XML output file)

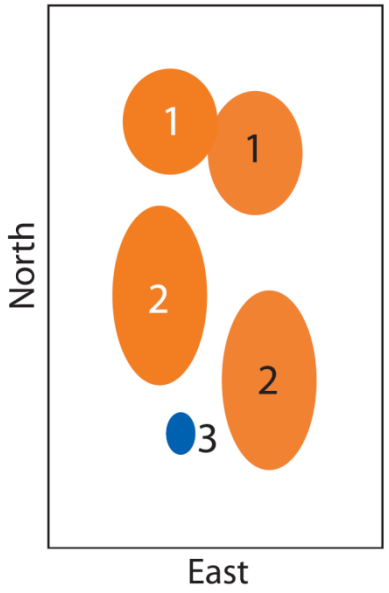


figure from Davies et al. (2009)

Interest Matrix

Forecast \ Observed	1	2	max:
1	0.90	0.75	0.90
2	0.50	0.80	0.80
3	0.40	0.55	0.55
max:	0.90	0.80	

0.80 (MMIF)

0.85 (MMIO)

Step 3: Verification of identified matched objects

- Median of Maximum Interest (MMI) – MODE (Davies *et al.*, 2009)
- applicable to both matched objects from MODE and matched objects from KONRAD3D (XML output file)

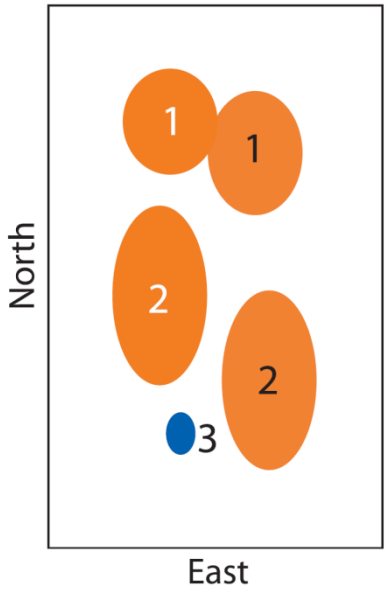


figure from Davies et al. (2009)

Interest Matrix

Forecast \ Observed	1	2	max:
1	0.90	0.75	0.90
2	0.50	0.80	0.80
3	0.40	0.55	0.55
max:	0.90	0.80	

0.80 (MMIF)

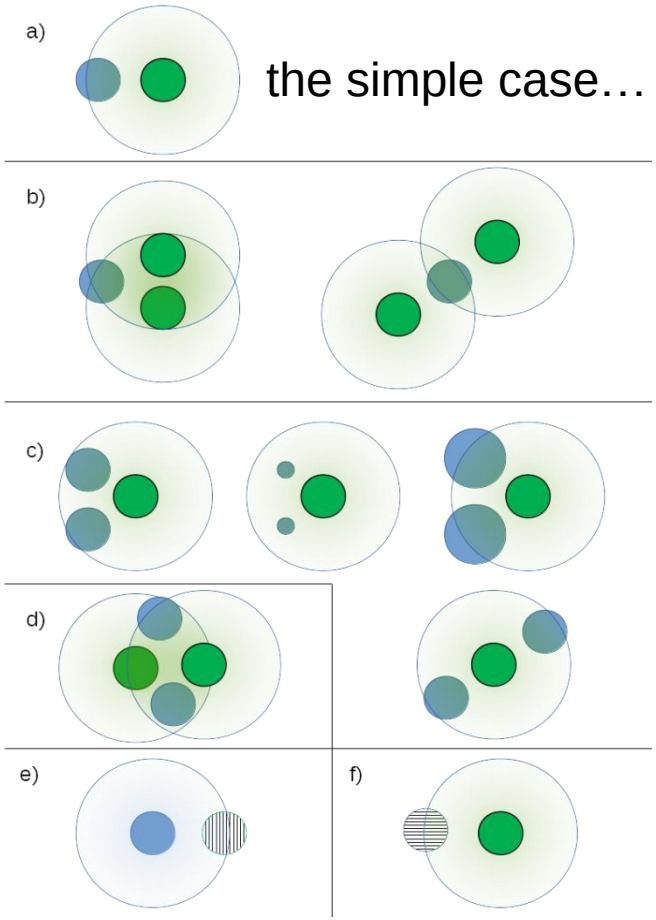
0.85 (MMIO)

MMI – Median of all maximum interest values

An alternative approach without matching…

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

For some situations, matching becomes a big problem:

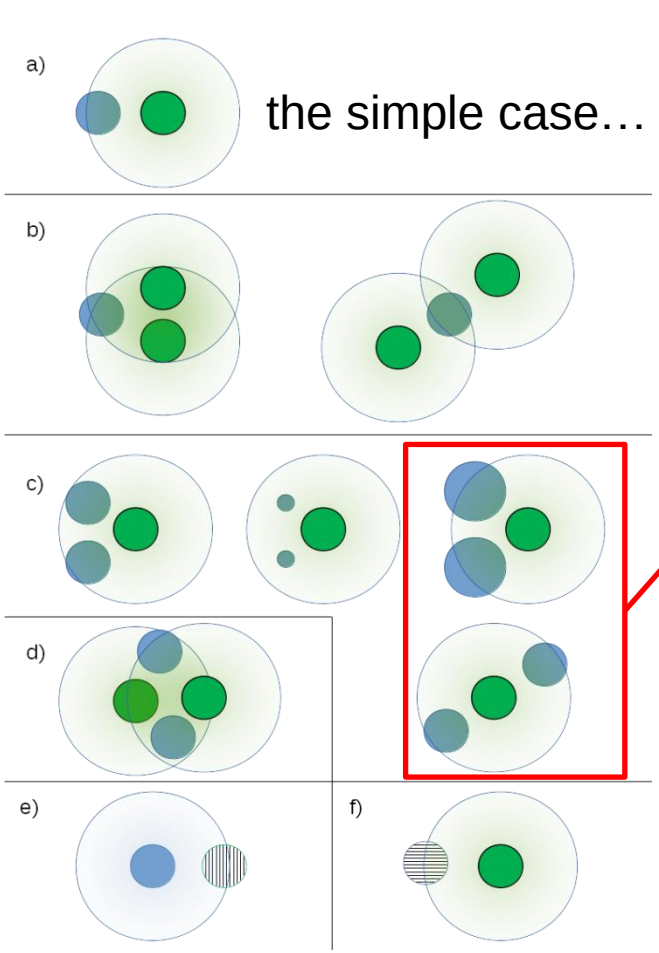


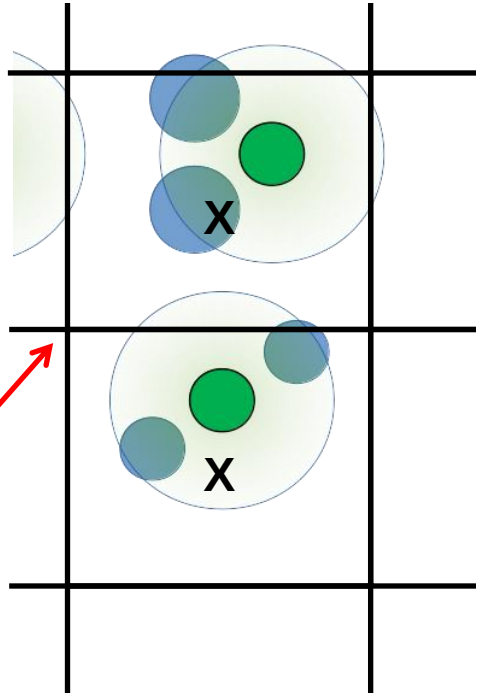
the simple case…

An alternative approach without matching…

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

For some situations, matching becomes a big problem:

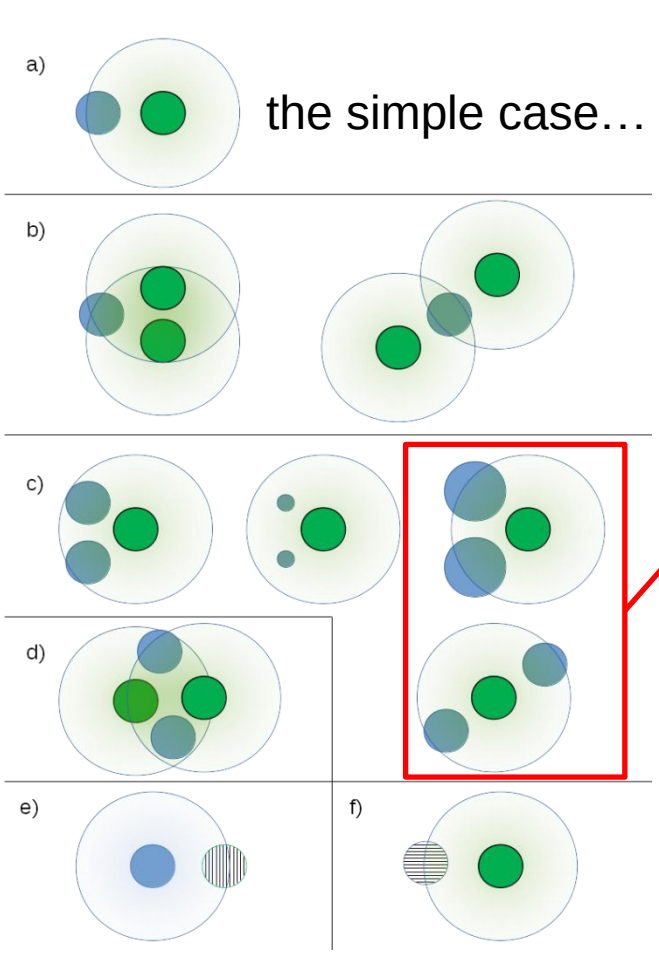


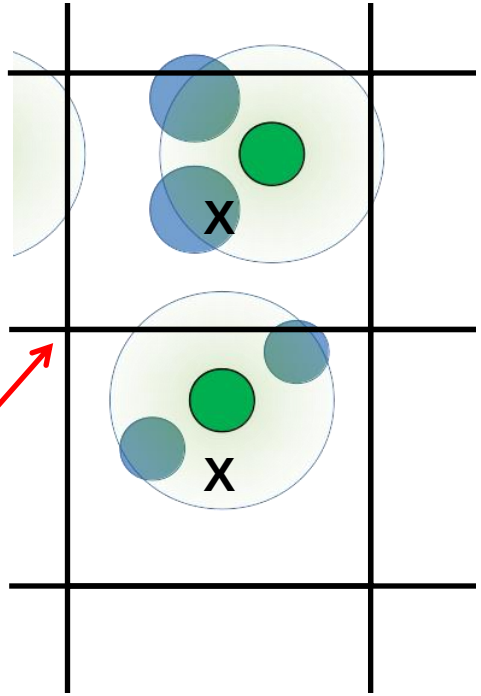


An alternative approach without matching…

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

For some situations, matching becomes a big problem:



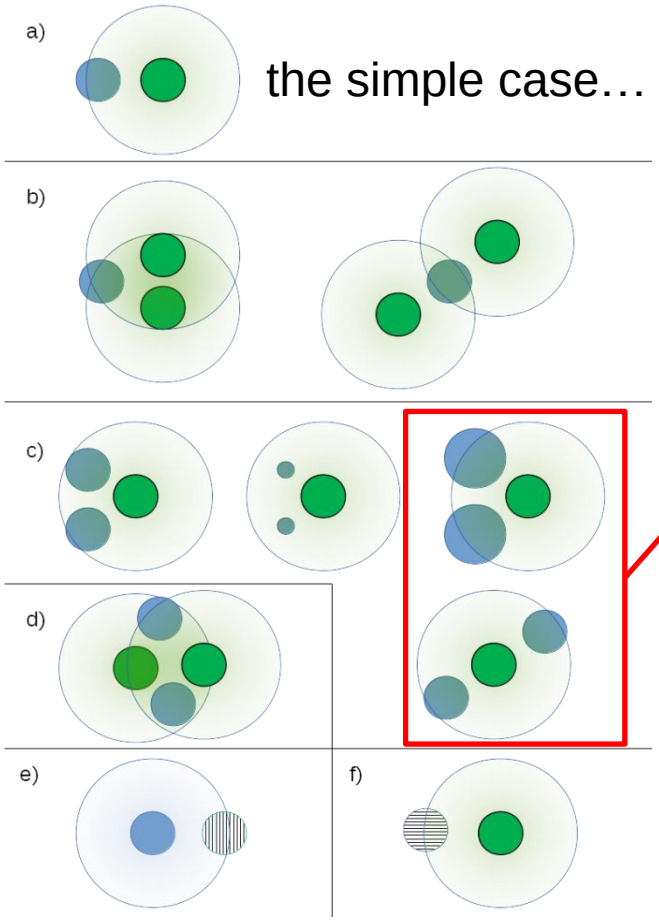


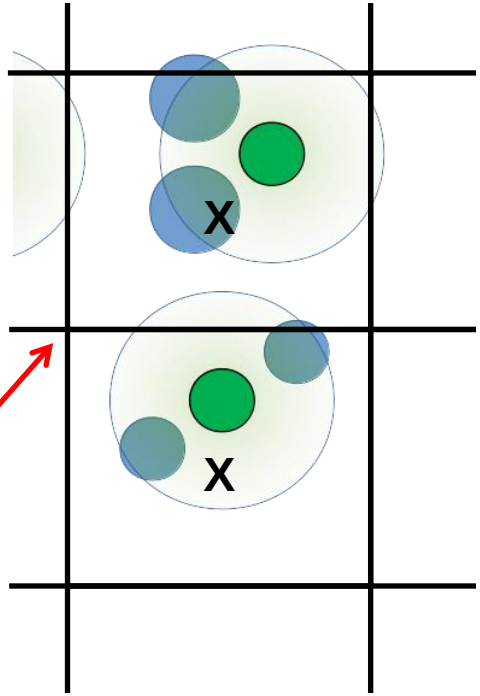
➔ Determine object-related properties as „averages“ in a local neighbourhood around a fixed location **X** in space

An alternative approach without matching…

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

For some situations, matching becomes a big problem:





- ➔ Determine object-related properties as „averages“ in a local neighbourhood around a fixed location **X** in space
- ➔ e.g. #objects > thresh,

  area > thresh

  mean distance betw. obs & sim objects

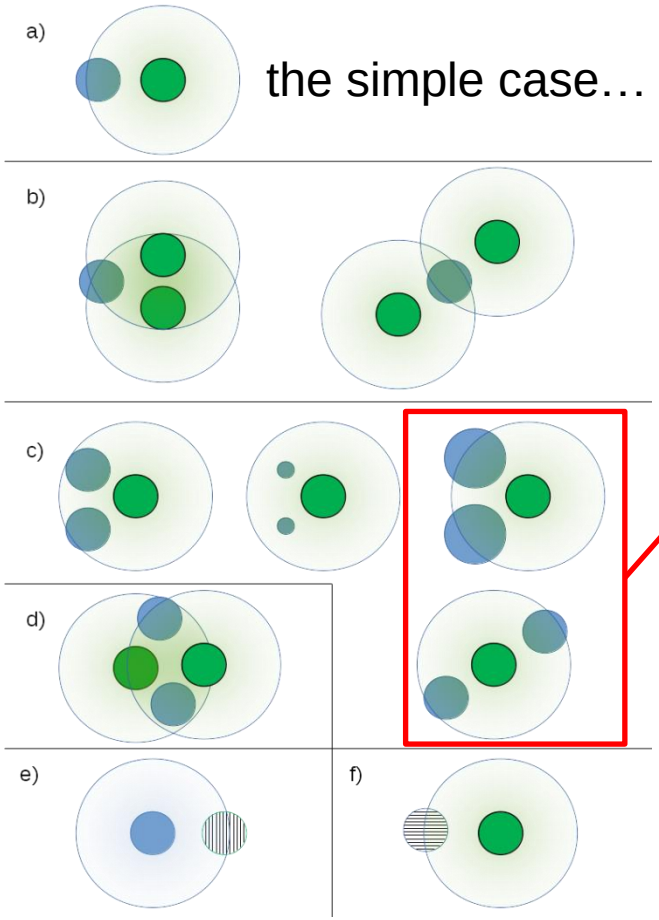
  lightning activity

  Echo tops/base

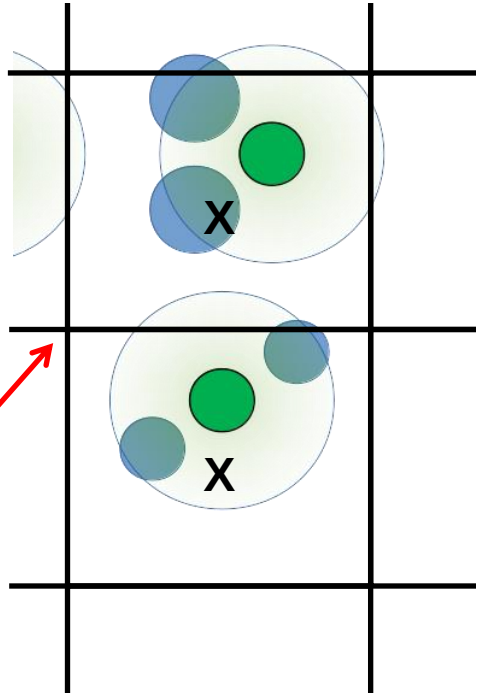
An alternative approach without matching…

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

For some situations, matching becomes a big problem:



the simple case…



- ➔ Determine object-related properties as „averages“ in a local neighbourhood around a fixed location **X** in space
- ➔ e.g. #objects > thresh,

  area > thresh

  mean distance betw. obs & sim objects

  lightning activity

  Echo tops/base

- ➔ Do this both for **obs** and **simulations**
- ➔ verify this „gridded“ information locally at **X**
- ➔ At the moment just an idea, we are currently starting with it.

Summary and Outlook

- (pixel-based) standard verification of recalculated forecast in a sampling period May/ June 2016 based on 5-minute radar-reflectivity composites
- comparison of Nowcast (+4h) and NWP (+6h)

Nowcast	NWP
• Mainly linear evolution • Bias strongly dependent on initial weather condition and its evolution	• Mostly slight under-forecast • Scarcely high reflectivities (1-mom. scheme)

Scores in the range of expectation

- crossover point in range +1.5h to +3h
- next steps: tests of existing object-based verification methods, parameter studies, including ensemble forecasts

Thank you for your attention!

Contact:

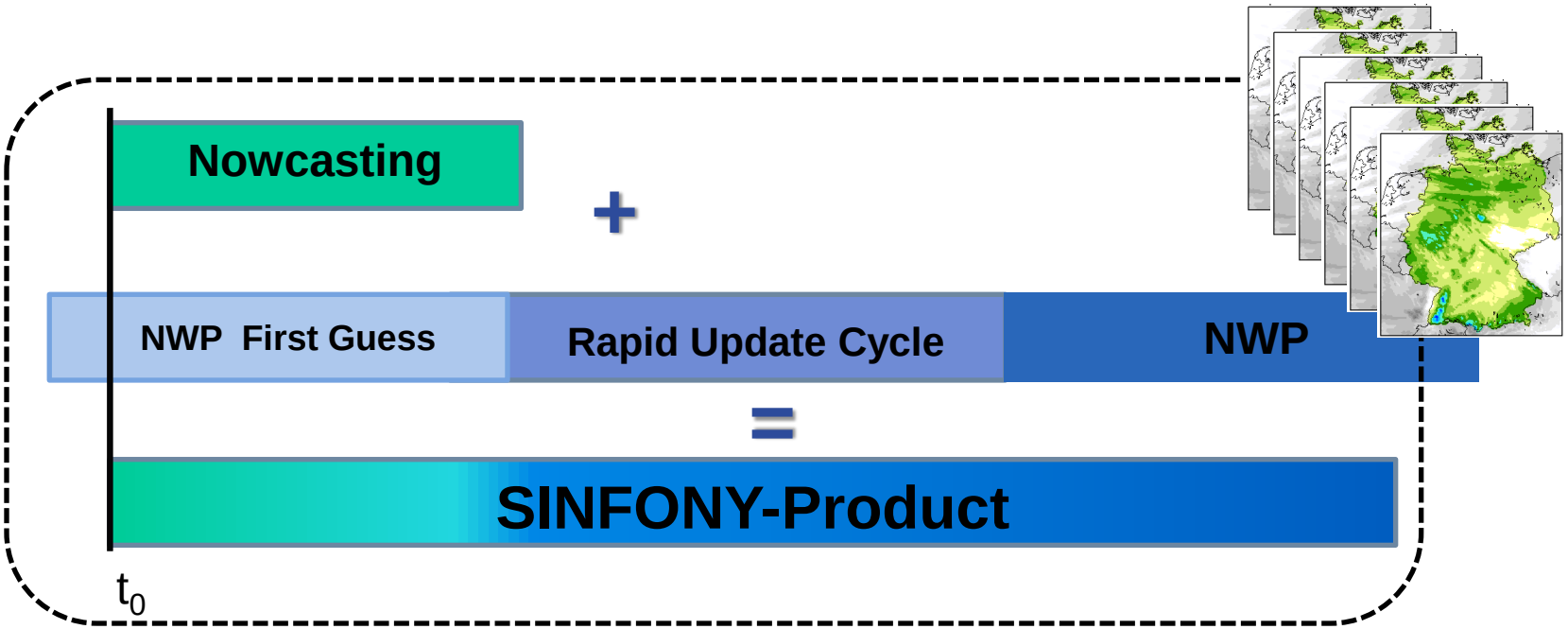
Michael Hoff
Deutscher Wetterdienst
Referat FE 15
Frankfurter Straße 135 / Strahlenberger Str. 17 (Omegahaus C)
63067 Offenbach am Main

E-Mail: michael.hoff@dwd.de
Tel.: +49 (0) 69 / 80 62 -3416
Büro: E8.C.05 (Omegahaus C)

Anhang

Was ist SINFONY?

Seamless Detection and Ensemble Assimilation / Forecasting



Combination of (advanced) Nowcasting and (improved) NWP

Seamless **IN**tegrated **FO**recasti**N**g s**Y**stem

Concept of SINFONY

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Cell indentifi-cation
Cell motion (tracking)
5‘
Cell indentifi-cation
Cell motion (tracking)
5‘ …

ENS-Nowcasting
Life cycle analysis
Verification

Obs. objects
Simul. objects
From NWV: Envir. cond.
From NWV: Envir. cond. ENS uncert.

Data:
3D Radar
MSG
Lightning

Combined products for forecasters

Silently:
- „conventional“ data
- GPS derived moisture
- other satellite data

KM-Scale ENS-NWV

RAD
SAT
Light ning
LETKF

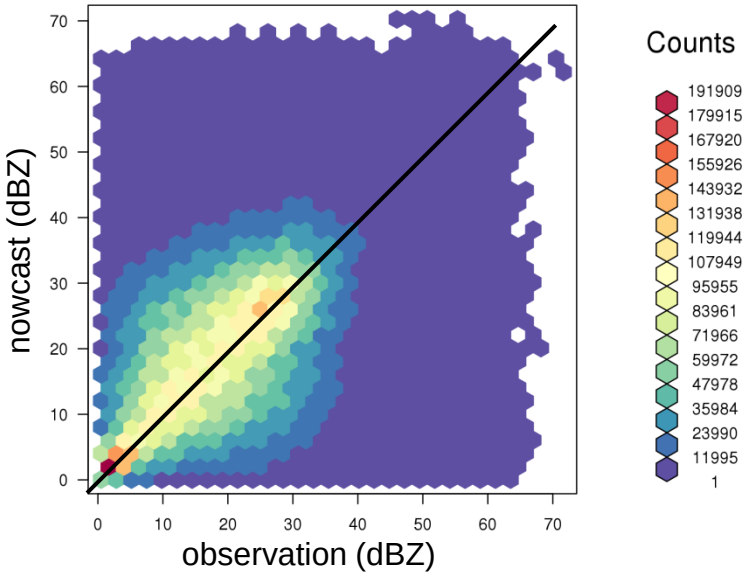
NWV ANA$_{-1}$
NWV First Guess
NWV ANA$_0$
NWV First Guess
NWV ANA$_1$

+1 h
+1 h

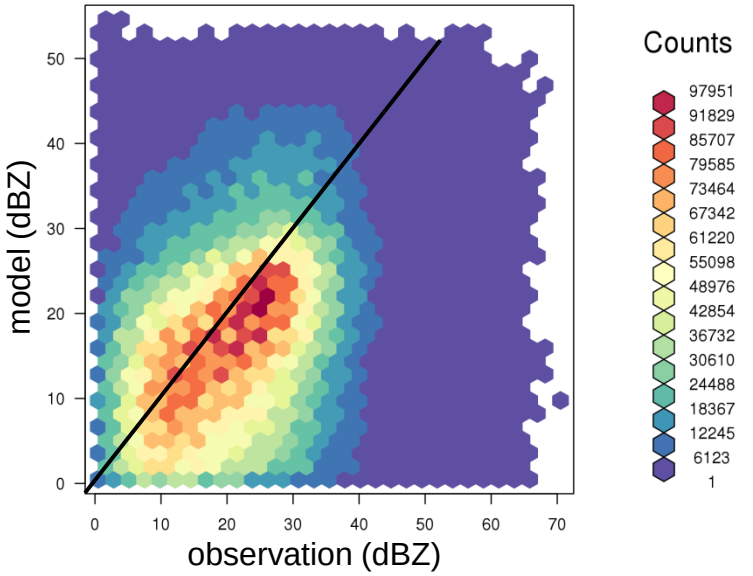
Forward Operators

Hexbinplots for the period 26.05.2016 – 25.06.2016

nowcast (init: 15UTC +0.25h to +4h)



NWP (init: 15UTC +0.25h to +4h)



- ➔ high reflectivities possible (if in observation)
- ➔ linear behaviour, tendency to over-fcst

- ➔ scarcely high reflektivities (>45dBZ)
- ➔ slight under-forecast

Fraction Skill Score (Roberts & Lean, 2008)

Was wollen wir wissen?

- Entspannt die Bedingung des exakten „Matching“ von Events zur Evaluierung der Vorhersage in einer lokalen Umgebung um die Beobachtung
- Wie variiert Vorhersageskill mit Fenstergröße (räuml. Skala) und mit Schwellwert (gibt es z.B. lokale Max./ Min.)
- Kleinste Fenstergröße für ausreichend akkurate Vorhersagen

  (z.B. 1km Gitter → 2km Phänomene auflösen, aber nicht vertrauenswürdig. Praktisch: 4-6km)
- Geben höher-aufgelöste NWV wirklich genauere Vorhersagen in interessierten Skalen?

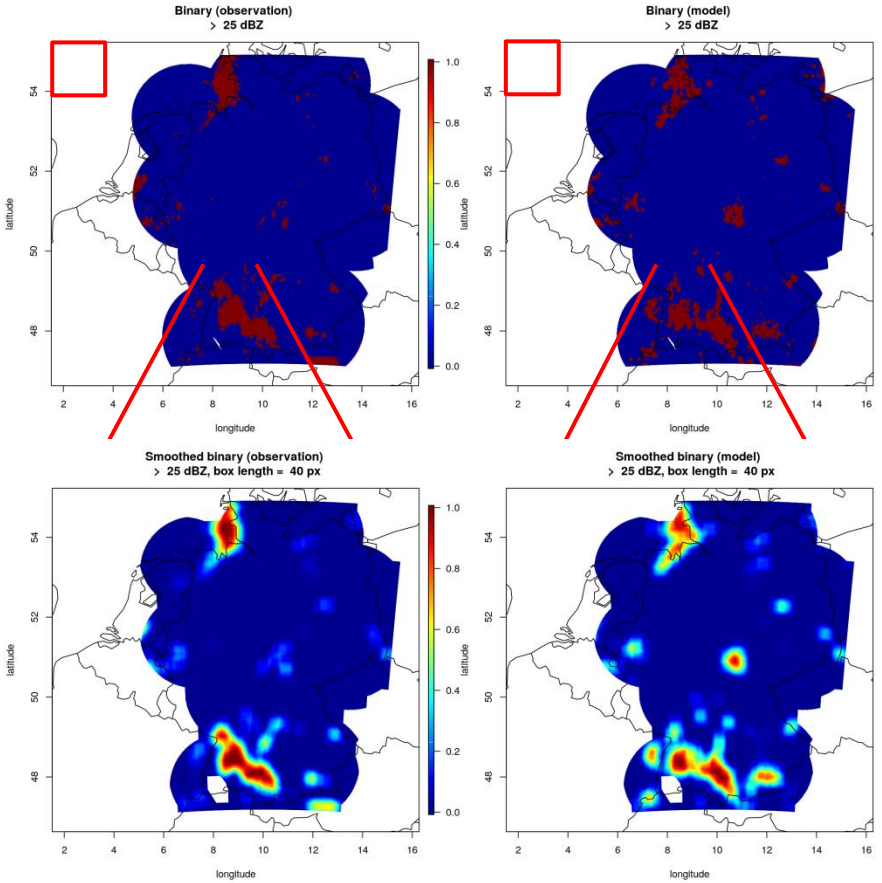
Fraction Skill Score (Roberts & Lean, 2008)

How to calculate FSS?

1. Fcst and obs on same grid
2. Choose proper thresholds
3. Convert fcst and obs to binary field for each threshold
4. Choose box length and scan the entire domain pixelwise
5. For each box:
   - ➔ count pixel > *threshold*
   - ➔ divide by number of pixel in box
   - ➔ get a smoothed picture with fractions
6. Calculate FSS:

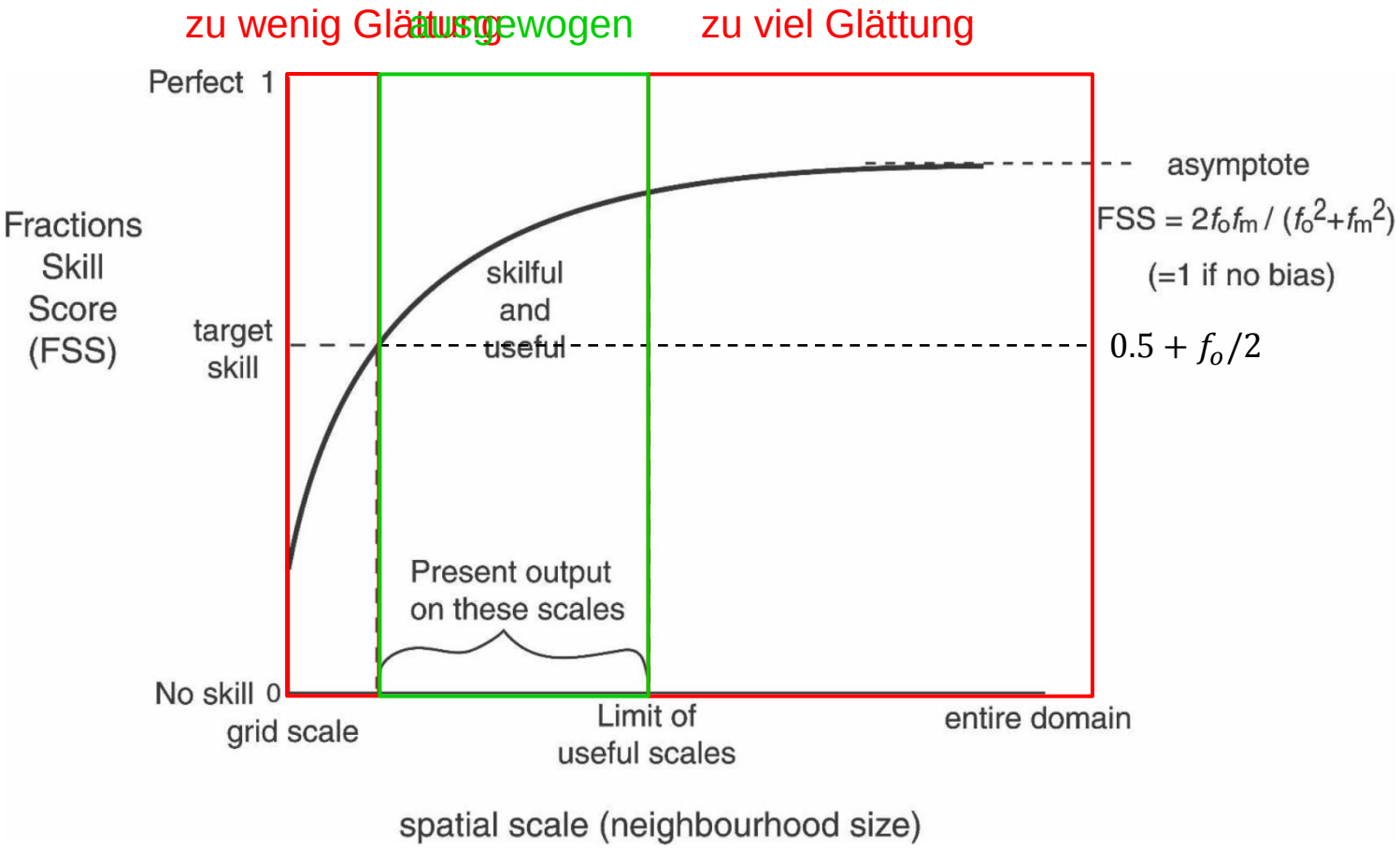
$$FSS = 1 - \frac{\sum \left(P_f - P_o\right)^2}{\sum P_f^2 + \sum P_o^2}$$

7. Repeat steps 2 to 6 for different thresholds and box lengths

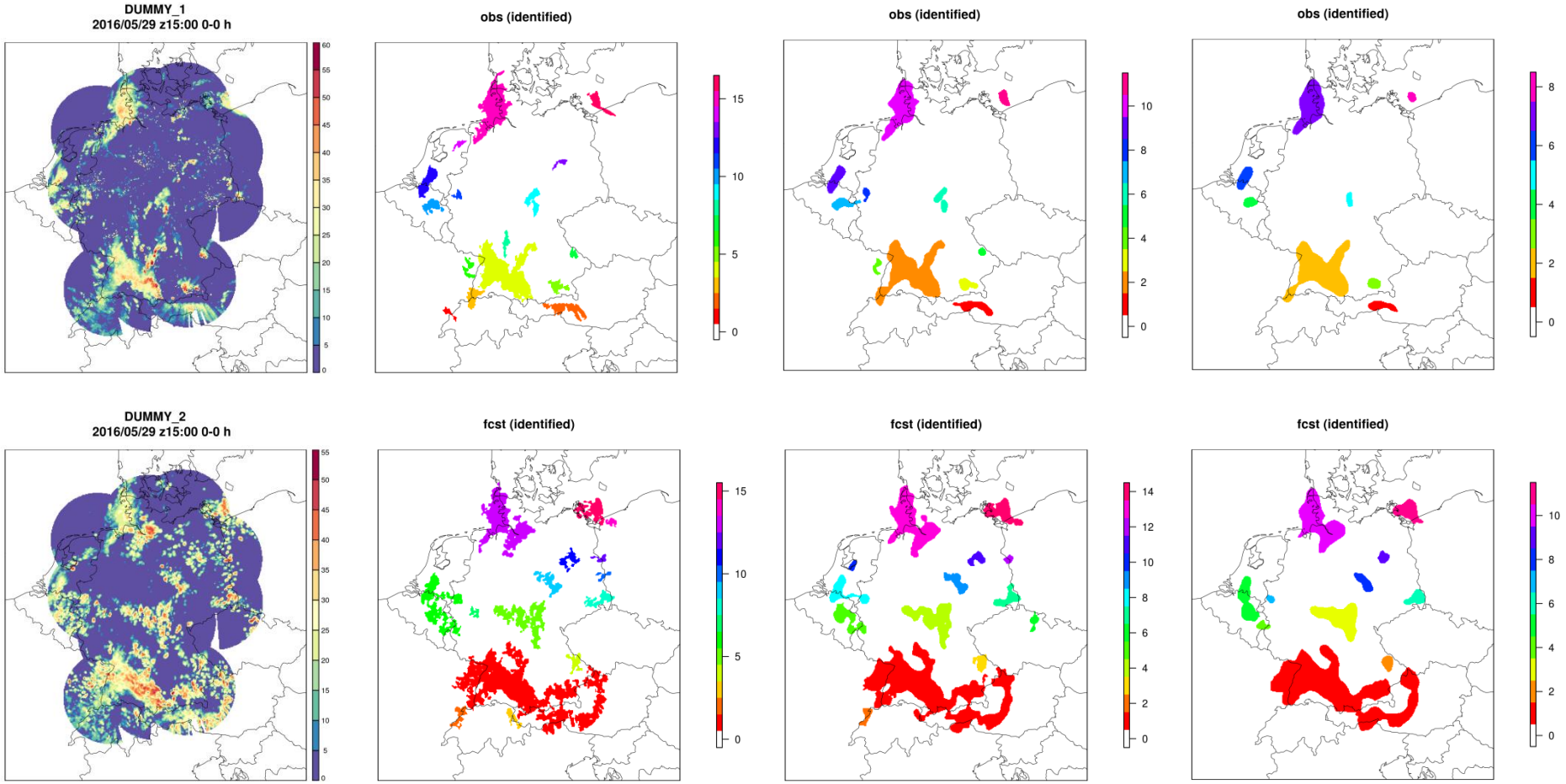


Fraction Skill Score (Roberts & Lean, 2008)

Zur Interpretation des FSS

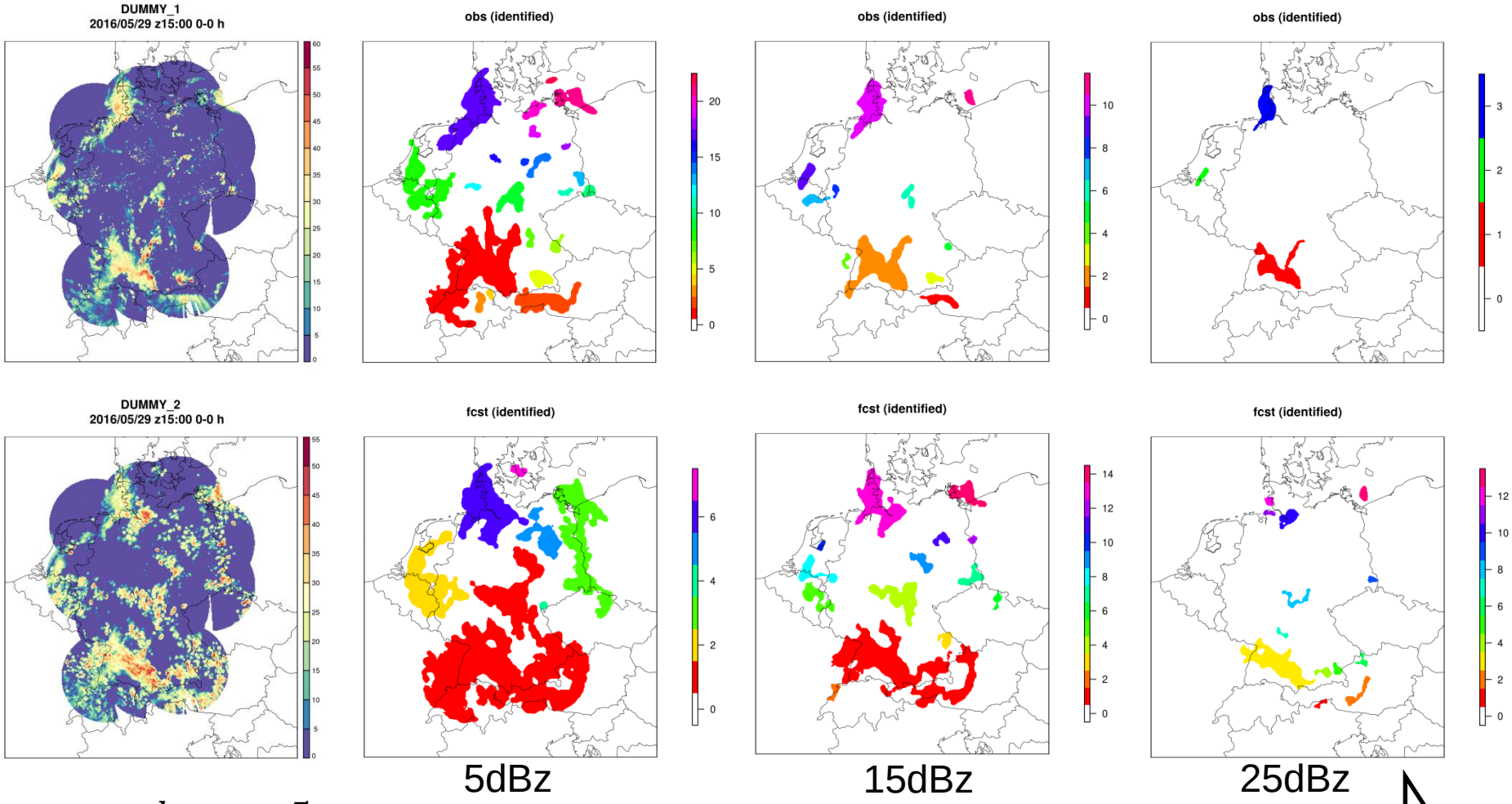


Step 1: Define features



$SW = 15dBz$
$minsize = 100px^2$

Step 1: Define features



$smoothpar = 5$

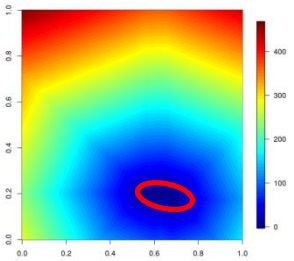
$minsize = 100px^2$

increasing threshold

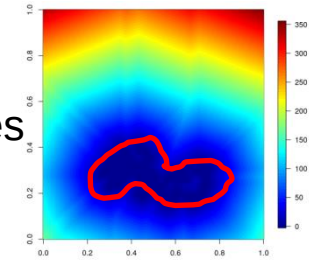
Bisher bewährte Matchingmethoden

Minimum Boundary Separation

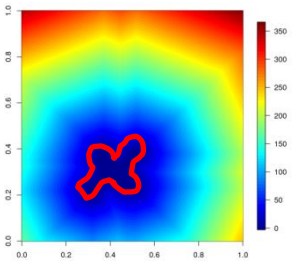
beobachtetes Feature 1



vorhergesagtes Feature

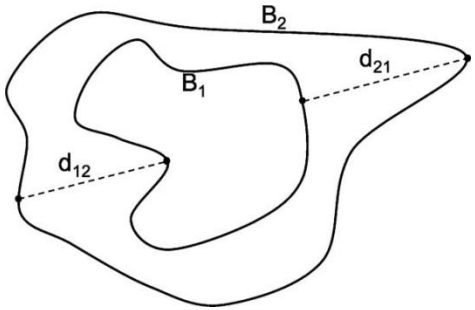


beobachtetes Feature 2



Delta Metric Methode (Haussdorffmetric)

- Eine Art Ähnlichkeitenvergleich zwischen Objekten



Centroid Distanz Methode

- Vergleiche Distanzen zw. Centroiden der Objekte
- Gültige Distanzen über Objektgröße definiert