User Oriented Verification

Cristina Primo
Deutscher Wetterdienst (DWD)

UEF- 8.Juni.2016

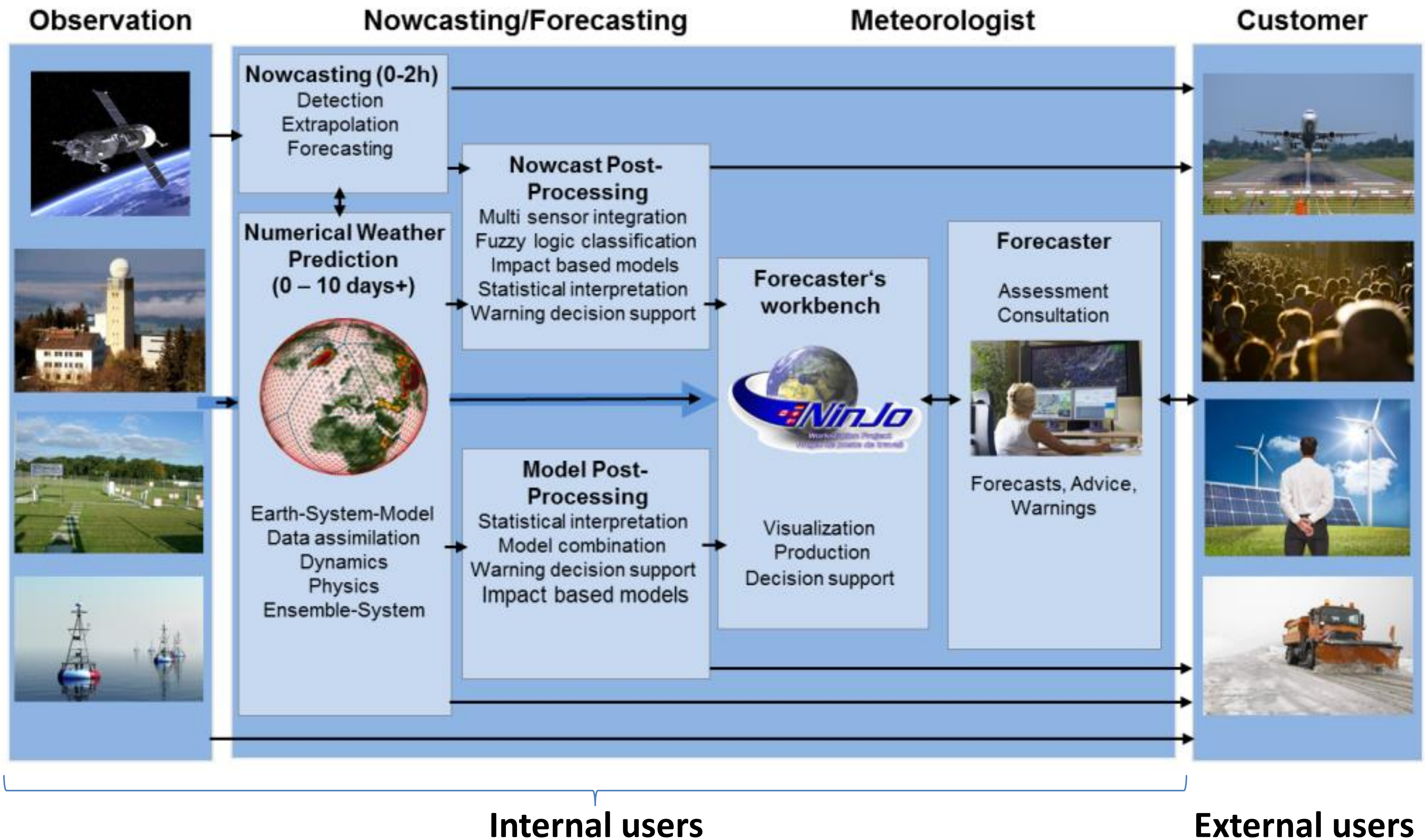
Outline

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- Forecast users
- What is a „good“ forecast
- Verification aspects to take into account
- DWD user oriented verification

Weather forecasting

Deutscher Wetterdienst
Wetter und Klima aus einer Hand



Forecast users

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- **Internal users:** researchers, model developers and forecasters.
  – Does the model represent well the physics?
  – Does the new version improve the previous one?
  – How is the performance during the last years to justify costs?

- **External users:** general public and customers.
  – Interested particularly in surface parameters.
  – Do they help me in the decission process to take action?
  – Can I benefit from using them? Do they help me to prevent losses?

The meaning of a „good“ forecast can differ among users

„good forecast“

What makes a forecast good? According to Murphy, 1993:

- QUALITY: correspondence between observations and forecasts
  - of the interest of forecast providers : Accuracy, reliability, skill, bias, etc..
- CONSISTENCY: forecasts agree with forecaster’s true belief about the future weather.
- VALUE: benefit as a result of using the forecasts (it does not have to be necessarily economic, saving lives has a lot of value!).
  - of the interest of users since value means that forecasts are useful in the decision problem.

„good forecast“

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

No magic number summarizes all the properties of a forecast.

External users are interested in forecasts because they DO have value. We should share our experience to help them to make the most of them. (E.g. Help them to interpret probabilities).

The relationship between value and quality is complex:

- In a non linear relationship, small increases of skill may lead to notable increases of value and vice versa.
- A totally wrong forecast system can still have a lot of value if the user knows that they have to choose exactly the opposite outcome.

Verification issues

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Matching forecast and observations:

- Forecasts are not perfect but we do not know the truth either:
  - Best estimate of the initial state for the models (field).
  - Observation measurements (points): Some users have more observations than the weather centres.
- How to match observations and forecasts:
  - Spatial representativeness problems
- Forecast and observations have to be defined in the same way (do not compare pears with apples!).
  - Feedback from users requirements is necessary.

Verification

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- **Identify verification purposes**: Choose the best score that gives information about the aspect of the forecast we are interested in.
- **The score has to be proper**: Forecasts have to be consistent with forecaster's belief. Some scores give better results when the forecasts are closer to climatology. The score must be sensitive to both reliability and resolution. That is, a forecast must sort the observed states of the system into groups that are different from each other.
- **Show the uncertainty around the score** (confidence intervals on the verification measures) to be confident that the estimate of forecast quality is not misleading.

If verification is run by non experts, it may lead to wrong conclusions!

Forecast types

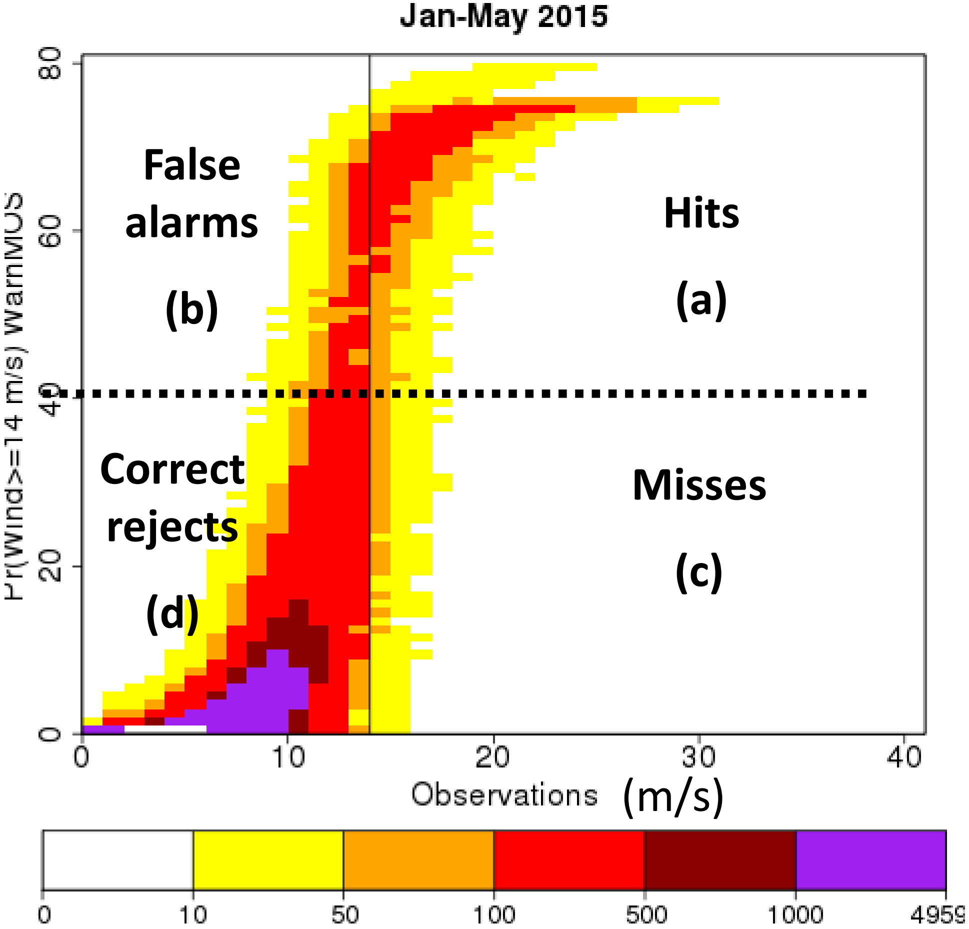
Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- Probabilistic forecasts include information about forecast uncertainty.
  - If we show the users that forecasts are able to discriminate successfully between the observations, users can take appropriate decisions based on those forecasts.
  - ROC curve gives information about discrimination and helps to make the YES/NO decision.

- Some users prefer to have a YES/NO forecast, so they loose information about uncertainty:
  - Contingency tables

Turning probs into Yes/No

Wind Gust, leadtime :1h

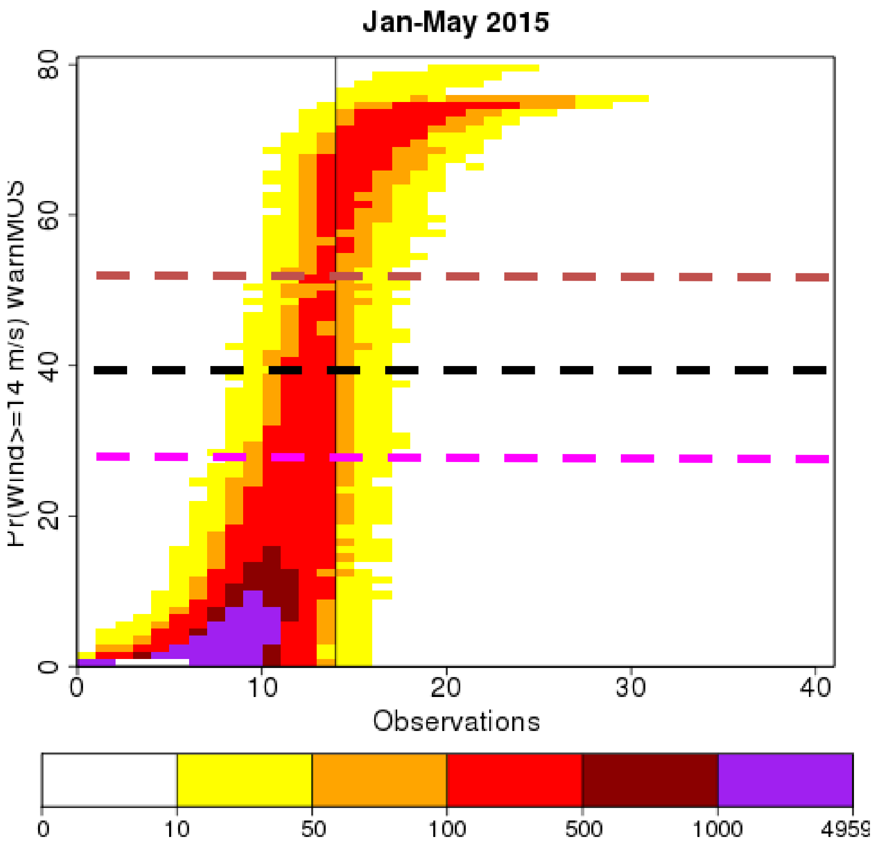


Observations

Forecasts

	Yes	No	Total
Yes	**Hits**	**False Alarms**	
No	**Misses**	**Correct rejects**	
Total			

Turning probs into Yes/No



Increase P_{th} → the model says NO more.
Less false alarms and less hits.

Decrease P_{th} → the model says YES more.
Less misses but less correct rejections

What „the best threshold“ is depends on user priorities.
Some users penalize more the false alarms while others the misses.

Value for a binary event

Roebber und Bosart, 1996

			Event: Frost in a citrus farm
Action: Spraying crops with water to protect against freezing		**Adverse weather**	**No adverse weather**
	Do not Protect	Net payoff for adverse weather and no protection (a)	Net payoff for non-adverse weather and no protection (b)
	Protect	net payoff for adverse weather and protection (c)	net payoff for non-adverse weather and protection (d)

expected return is optimized if :

$$P(event) \geq \frac{b-d}{(b-d)+(c-a)} \quad \Rightarrow \text{Take action}$$

Net payoff= expenses. Usually negatives for a, c, d.

Cost-loss model

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Observations

Forecasts

	YES	NO
YES	Mitigated Loss	Cost to prevent weather related damages
NO	Loss in case the user does not protect his operations	No Cost

- This decision model maximizes gain or loss-avoidance, but it requires cost information about the user, and it is difficult to generalize.
- Sometimes we cannot express this table in terms of costs. Which is the cost of saving a live or the loss of not saving it?
- Avoiding cost/loss function, we choose a threshold that allows a compromise between hit rate and false alarm ratio.

DWD examples:

- Warnings
- Aviation weather data
- Solar and Wind power

DWD key customers:

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- Civil protection, federal and regional authorities
- Provincial administration of Road construction, winter services
- Municipalities
- Energy supply companies
- Service Business, other important business companies
- Media, agencies
- Trading
- private customers / general public

Warnings

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

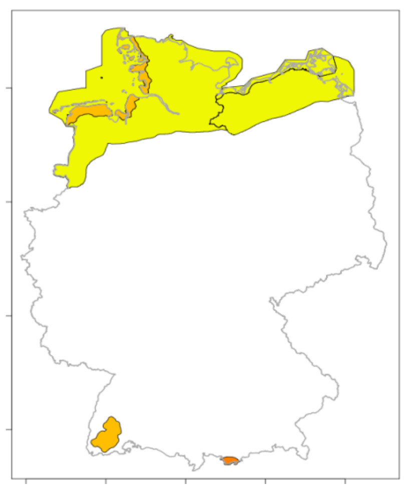
- Time
- Area
- Intensity

Lead time t1 - t0 | Duration t2 - t1

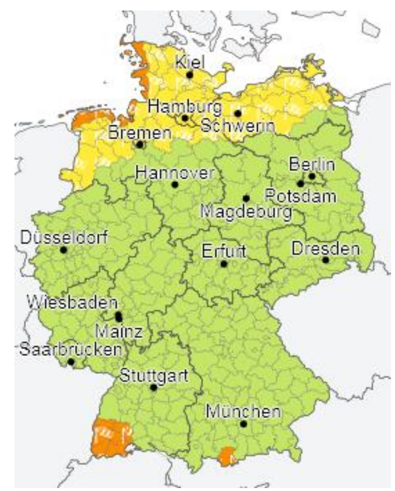
Issue Time t0 | start Time t1 | end Time t2

Model output + MOS
→ Warning Polygons

Polygons + Forecasters
→ Warnings in county



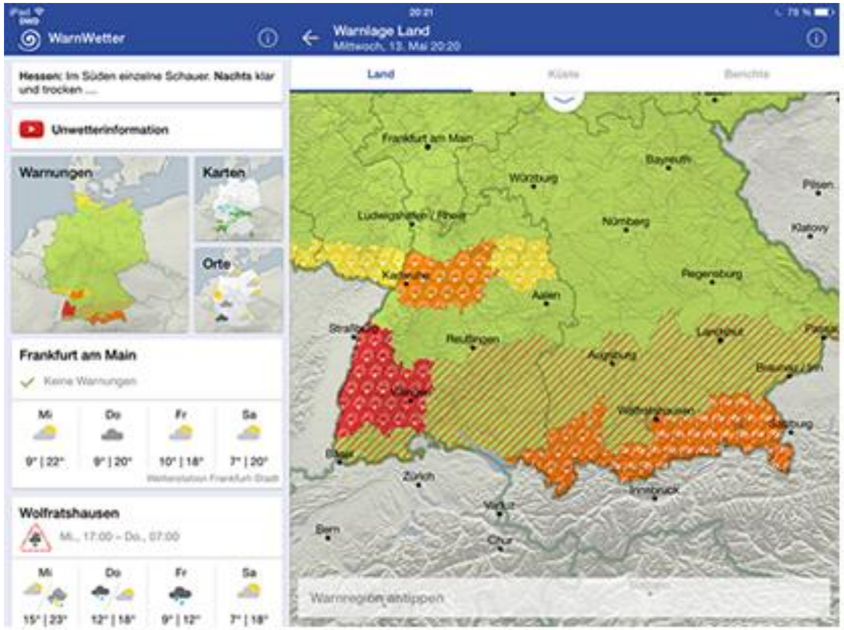
No warning
Warning level 1
Warning level 2

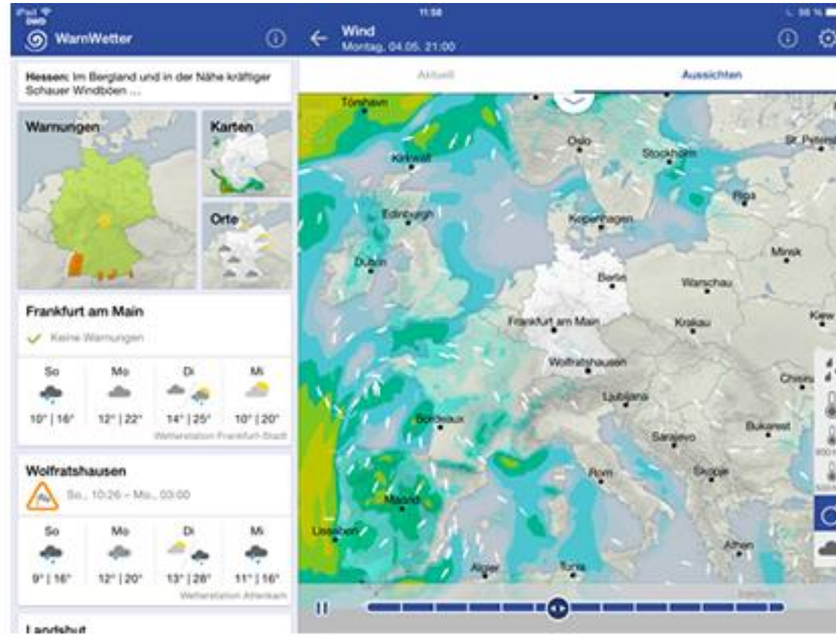


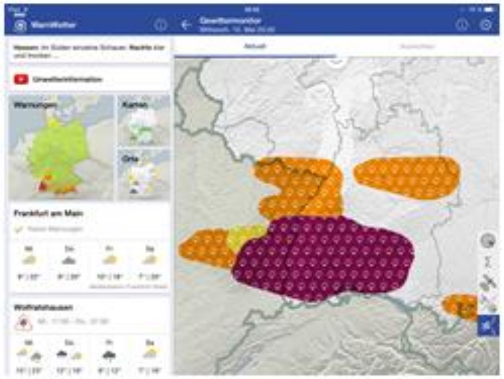
30.03.2016, 13 UTC

WarnWetter App

Deutscher Wetterdienst
Wetter und Klima aus einer Hand



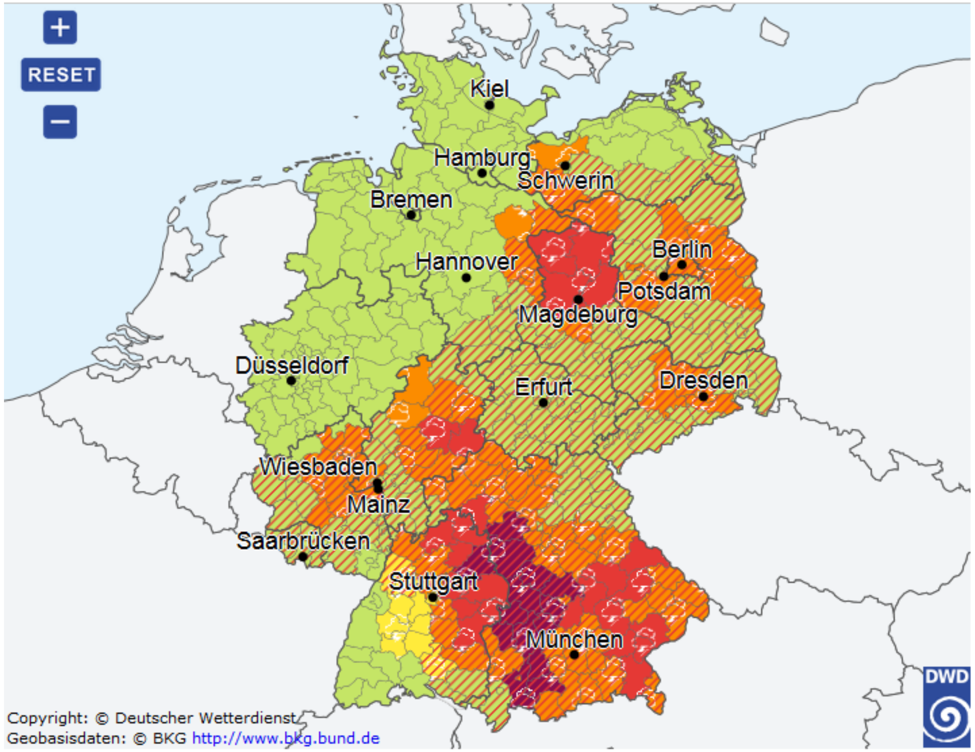




Warnings

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Floodings in Germany last week



Eifelkreis Bitburg-Prüm

VORABINFORMATION UNWETTER vor HEFTIGEM / ERGIEBIGEM REGEN

So, 29. Mai, 19:00 – Mo, 30. Mai 08:00 Uhr

Am Abend und in der Nacht zu Montag kommt von Südosten Starkregen auf, in den auch einzelne Gewitter eingelagert sein können. Dabei können Regenmengen zwischen 25 und 40 Liter pro Quadratmeter innerhalb weniger Stunden auftreten, lokal auch noch mehr.

Warning verification

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

We are provided with database of warnings and observations

Hourly verification

0 0 0 0 0 1 1 1 1 0 0 0 0 1 0 1 0 1 1 1 1 1 0

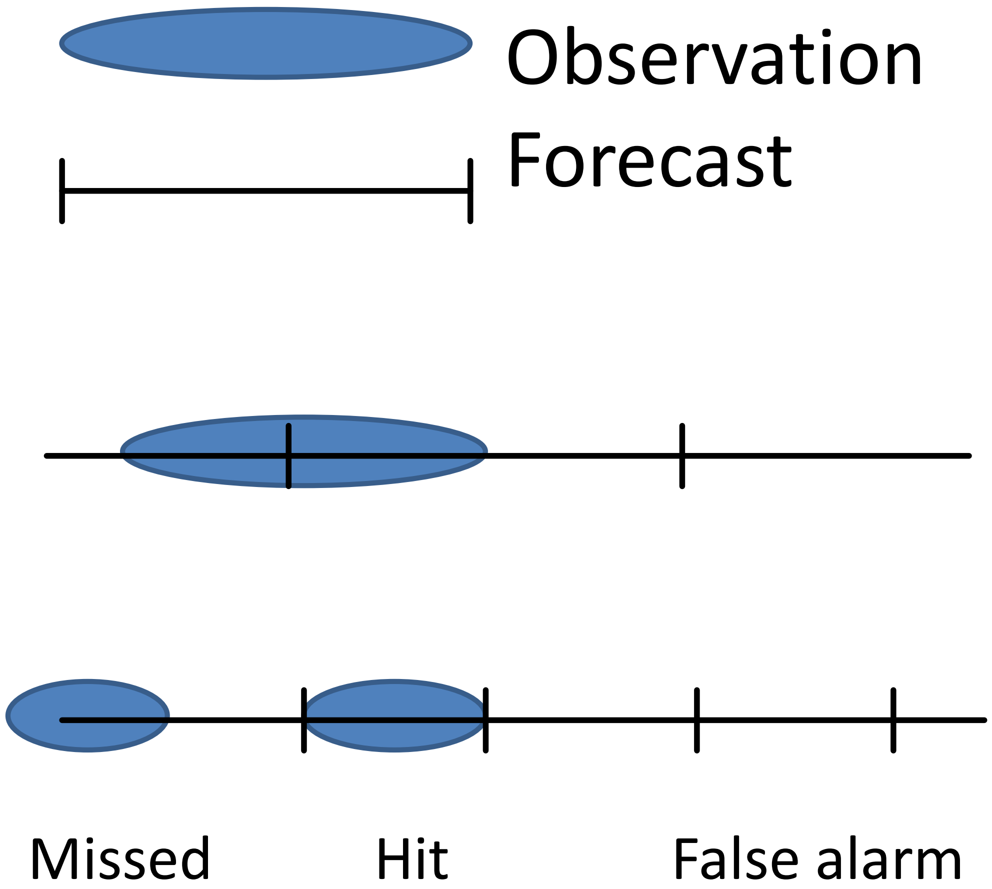
0 0 0 0 1 1 1 0 0 1 1 0 0 0 1 1 1 1 1 1 1 1 0

Contingency table

Hits, false alarms, misses, correct rejects

SCORES

Double penalty



Double negatives: For the same warning, what is counted as a miss, will be later counted as a false alarm, and vice versa.

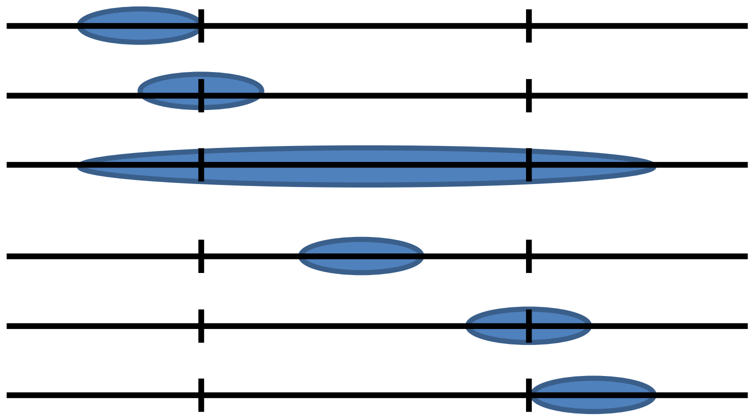
Move into an event-based verification

Questions

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- There is no one by one relationship between observed events and warnings → Verification of the observed events (more priority to misses) or the warnings (more priority to false alarms)?

- How do we deal with misplaced forecasts?



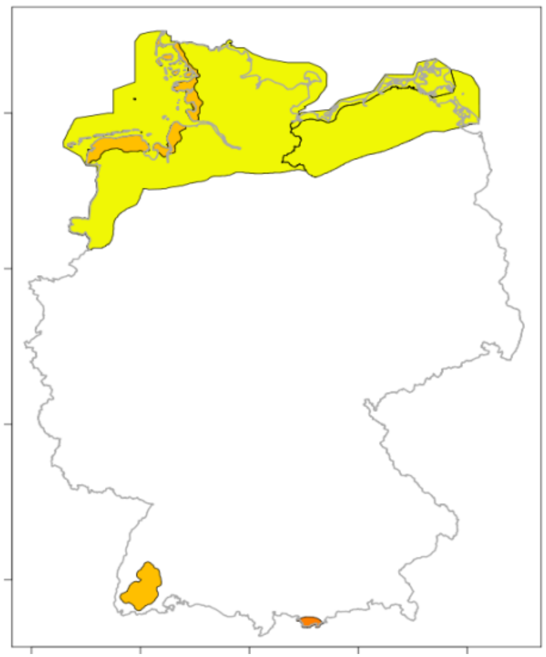
Choose a criterion to define what a hit is.
What do the users penalize more?

Intensity errors
Spatial misplaces
Temporal errors

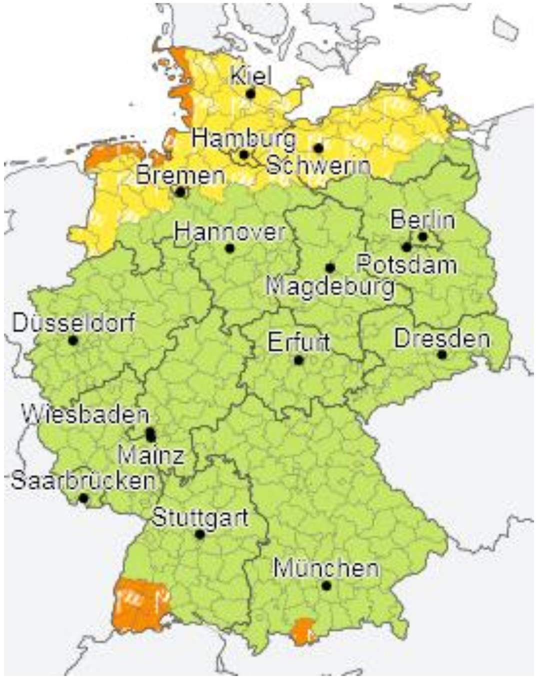
Other questions

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

Moving into a polygon verification



?



Aviation weather data

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

The Aeronautical Meteorology Department of Deutscher Wetterdienst co-operates with several providers of aviation services (such as navigation software, flight planning tools or reservation systems).

Verification

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- DWD has recently started a new project (new 15 permanent positions) to provide new products to the airlines and airports customers.

- Airlines industry requires specific products related to turbulence and icing. Attention has to be paid to get observations of the new forecast products. The USA is providing an observational dataset of eddy diffusivity rate (EDR), a measure of turbulence.

- No systematic verification is run yet, but DWD gets direct feedback from the user in a single-case situations.

EWeLiNE Project

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- Solar energy forecasts:
  - quantile forecasts are provided as decision variables for global radiation (main weather variable affecting solar energy forecasts).
  - Verification measures and tools for the assessment of quantile forecast discrimination (event based and user based) and value.

Bouallègue et al., 2015: „Quantile forecast discrimination ability and value“

EWeLiNE Project

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- Wind power:
  - Scenarios that describe spatio-temporal wind variability are relevant products for end-users of wind forecasts.
  - Calibrated quantile forecasts of wind at 100 meter height above the ground are provided to remove statistical inconsistencies and get reliable forecasts.
  - Different scenarios are generated with a dynamic ensemble copula coupling approach (more details in Bouallègue et al., 2016)
  - Product oriented verification is provided.
  - Applications that require temporal trajectories will fully benefit of this dynamic approach.

Bouallègue et al., 2016 : „Generation of scenarios from calibrated ensemble forecasts with a dynamic ensemble copula coupling approach“

Conclusions

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

- Users are more interested in verifying value of the forecasts.
- It is important to compare forecasts and observations representing the same, to use proper scores and to include confidence intervals.
- Verification should be run by experts to get the right conclusions.
- DWD runs user-oriented verification for warnings, aviation weather data and solar and wind power forecasts.

Thank you for your attention!