Data preparation for verification

L. Wilson
Associate Scientist Emeritus
Environment Canada

Outline

- Sources of observation data
- Sources of forecasts
- Types of variables
- Matching issues
  - Forecasts to the observations
  - Observations to the forecast
  - Examples

Observation data sources for verification

- Wouldn't it be nice if we had observations for every location and every point in time for the valid period of the forecast?
  - Then we could do complete verification of any forecast
- Observations represent a "Sample" of the true state of the atmosphere in space and time.
  - The "truth" will always be unknown
- Observations too may be valid at points or over an area
  - In situ observations or remotely sensed
- In situ observations – surface or upper air
  - Valid at points, in situ
  - High resolution, but drastically undersamples in space
  - Newer instruments can sample nearly continuously in time
  - Only important error is instrument error, usually small

Remotely sensed observations

- Satellite and radar most common
  - Radar
    - Measures backscatter from hydrometeors in a volume above the surface
    - Relationship to rain rate in the sensed volume is a complicated function but known
    - The link between the average rain rate in the sensed volume and rain rates (or total rainfall at the surface) is much more tenuous
    - Several sources of error: attenuation, anomalous propagation, bright band near the freezing level etc.
  - Satellite
    - Measures backscattered radiation in one or more frequency bands according to the instrument.
    - Usually low vertical resolution – may measure total column moisture for example
    - Transfer function needed to translate returns into estimates of the variable of interest.
    - Most useful for cloud, especially in combination with surface observations

Remotely sensed data (cont'd)

- Large data volumes
- Variable sensed is usually not the variable to be verified – transfer function required – one source of error
- Resolution dependent on the instrument, order of a few m for radar, 1km or so for satellite data.
- High coverage spatially, may be sporadic in time
- Beware of errors due to external influences on the signal

"I've looked at clouds from both sides now/ From up and down/

And still somehow/ it's clouds illusions I recall/ I really don't know clouds at all"/ --J. Mitchell

Summary of data characteristics

	In situ	Radar	Satellite
Resolution - space	High - point	Fairly high – radar volume avg	Depends on footprint 1 km or so
Resolution - time	high	high	high
Space sampling frequency	Low except for special networks	High – essentially continuous	High for geos within their domain Variable for polar orbit
Temporal sampling frequency	Can be high	High, typically 10 min or so	Medium for geos.; low for polar orbiting

Resolution: The distance in time or space over which an observation is defined
Sampling frequency (granularity): Frequency of observation in time or space

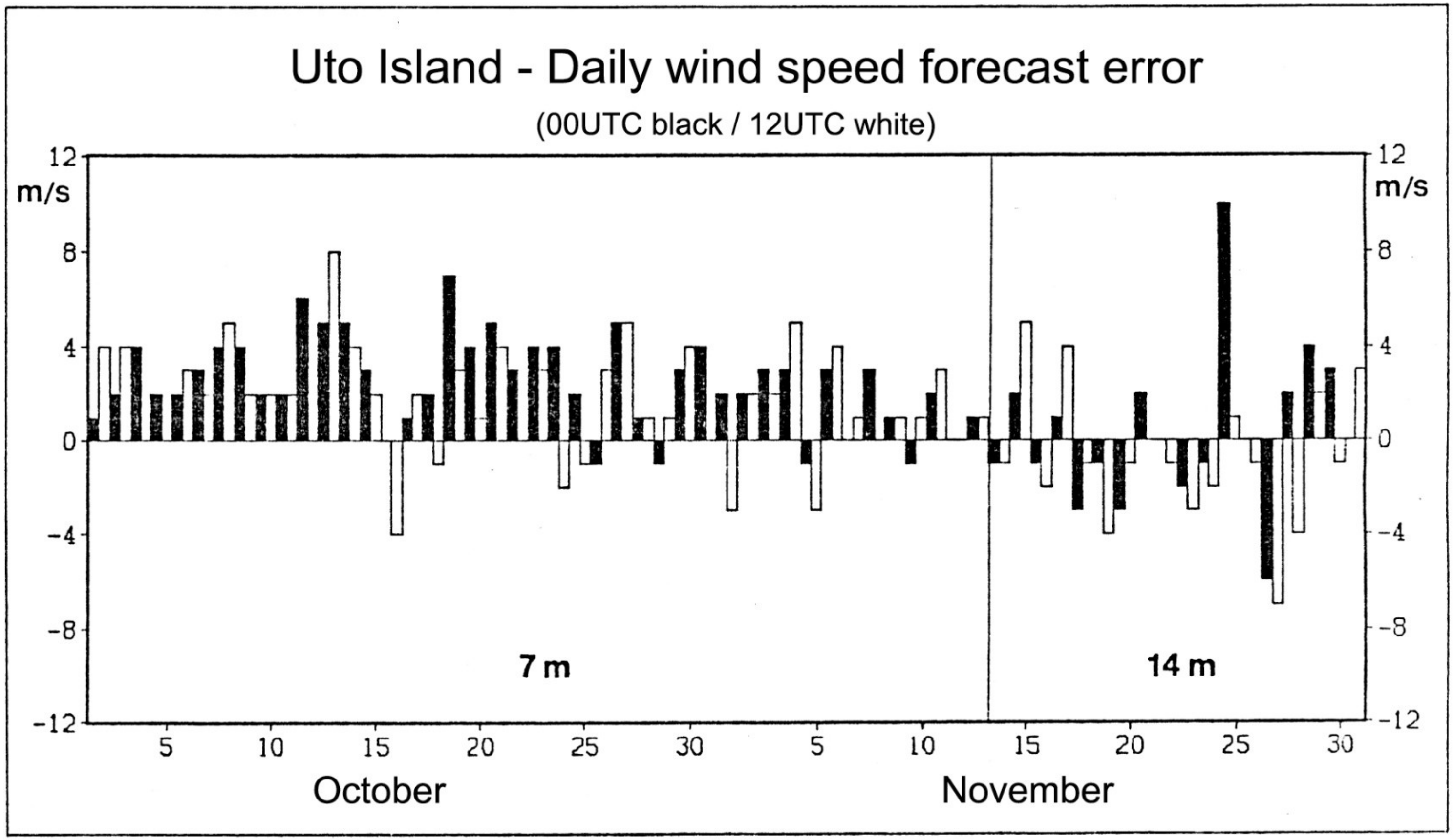
Sources of error and uncertainty

- Biases in frequency or value
- Instrument error
- Random error or noise
- Reporting errors
- Subjective obs
  - E.g. cloud cover
- Precision error
- Transfer function error
- Analysis error
  - When analysis is used
- Other?

Quality control of observations

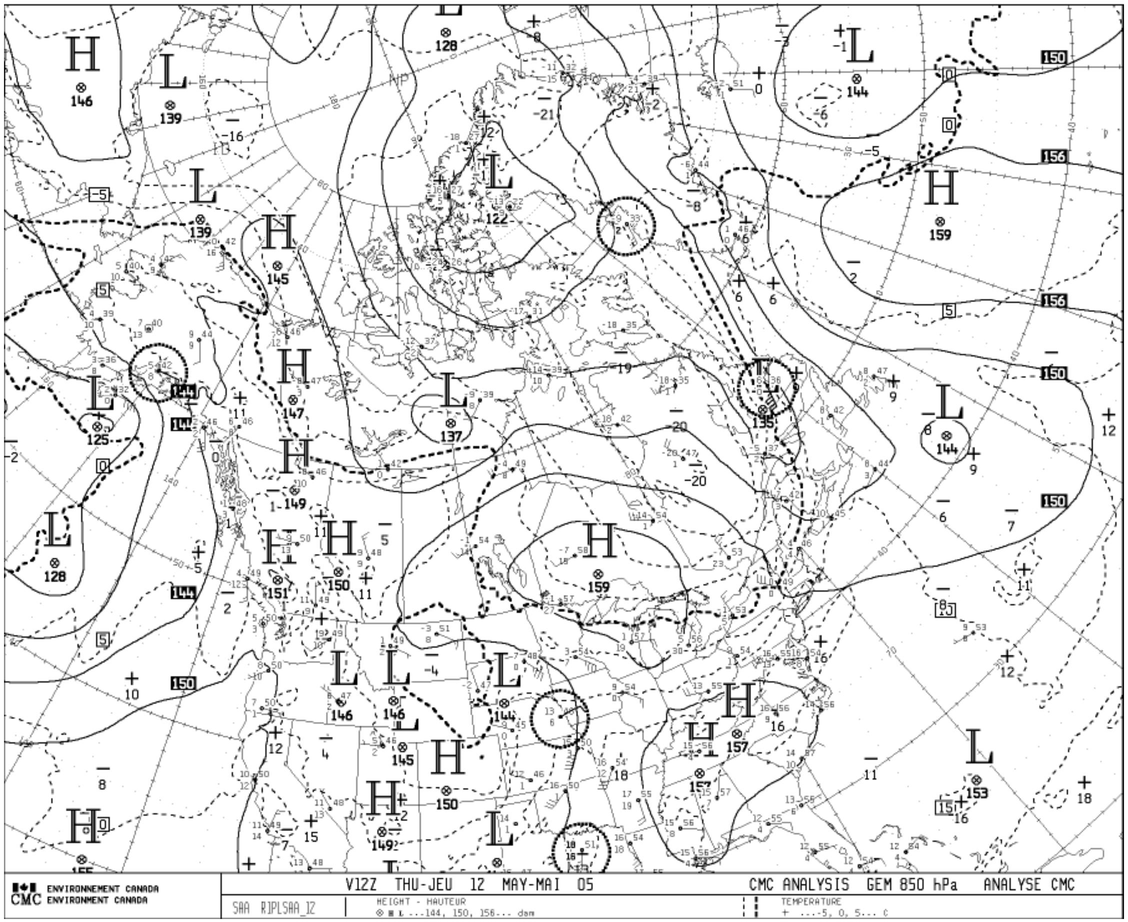
- Absolutely necessary to do it
- Basic methods: buddy checks, trend checks (checking with nearby independent obs in space and or time); absolute value checks etc.
- NOT a good idea to use a model as a standard of comparison for observations, acts as a filter to remove e.g. extremes that the model can't resolve
  - Makes the observation data model-dependent
  - Model used in the qc gets better verification results
- Important to know details about the instrument and its errors.

Importance of knowing measurement details



From P. Nurmi

Quality control of observations



- Quality control of observations:
  - Necessary, even for “good” stations
  - Buddy checks (space and time)
  - Simple range checks
  - Get rid of “bad” data without eliminating too many “good” cases
  - But NOT forecast-obs difference checks

Types of forecast validity

- For objective verification.....
- “Forecasts must be stated so they are verifiable”
- What is the meaning of a forecast? Exactly?
  - Needed for Objective verification
  - User understanding is important if the verification is to be user-oriented
  - All forecasts are valid for a point in space OR an area
  - At all points in the area?
- Similarly for time: A forecast may be
  - An instant in time
  - An instant in time, but “sometime” in a range
  - A total over a period of time e.g. 24h precip
  - An extreme during a period of time?

Forecast data sources for verification

- NWP models of all types
  - Deterministic forecasts of primary variables (P or Z, T, U, V, RH or Td), usually at grid points over the model's 3-d domain
  - Other derived variables: precip rate, precip totals, cloud amount and height etc, computed from model, may not be observed
  - Spatial and temporal representation considered to be continuous, but restricted set of scales can be resolved.
- Post-processed model output
  - Statistical methods e.g. MOS
  - Dynamic or empirical methods e.g. precip type
  - Dependent models e.g. ocean waves
- Operational forecasts
  - Format depends on the needs of the users
  - May be for points, may be a max or min or average over an area or over a period of time
- "Everything should be verified"

Types of Variables

- 1. Continuous
  - can take on any value (nearly) within its range
  - e.g. temperature, wind
  - forecast is for specific values
- 2. Categorical
  - can take on only a small set of specific values
  - may be observed that way e.g. precipitation, precipitation type, obstructions to vision
  - may be “categorized” from a continuous variable e.g. precipitation amount, ceiling, vis, cloud amount
  - Verified as categorical or probability of occurrence if available

Types of Variables (continued)

- 3. Probability distributions
  - Verified as a probability distribution function or cumulative distribution function
- 4. Transformed variables
  - values have been changed from the original observation
  - Examples:
    - Categorization of a quasi continuous variable e.g. cloud amount
    - To evaluate according to user needs:
      - “upscaling” to model grid boxes
      - Interpolation
    - Transforming the distribution of the observation:
      - E.g. subsetting to choose the extremes

Are continuous variables really continuous?

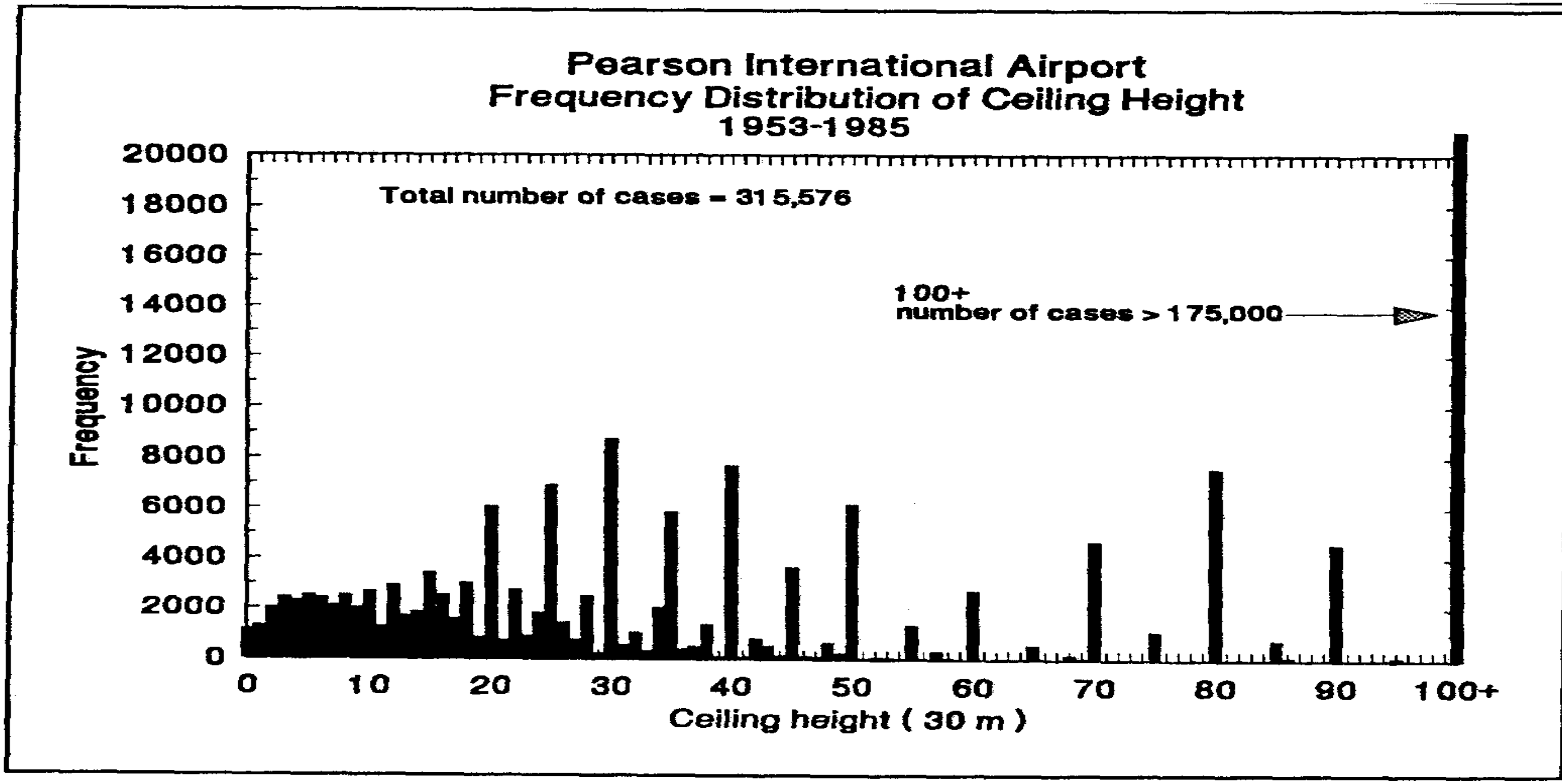


Figure VII-1. Frequency distribution of ceiling height for Pearson International Airport (Toronto, Ontario). Although it is nominally a quasi-continuous variable with 30 m resolution, many values of ceiling height are never reported.

Data Matching issues

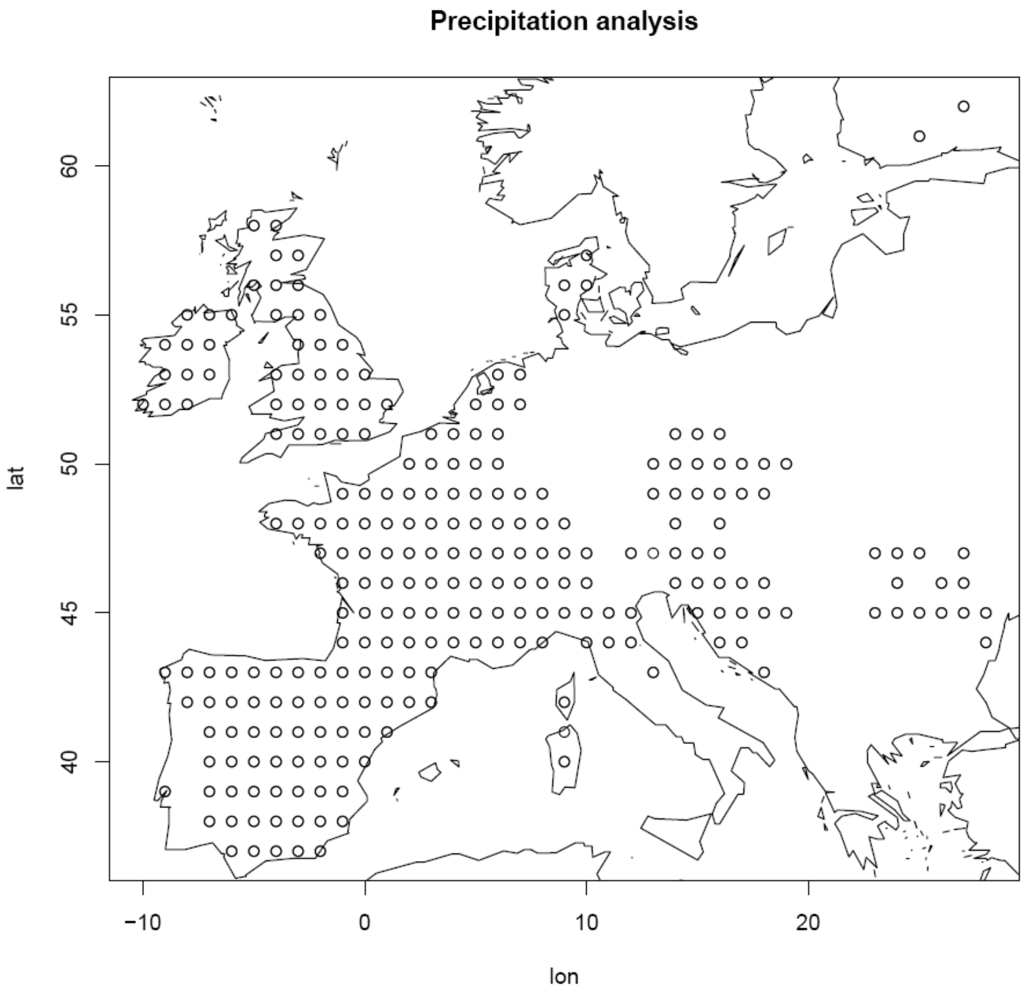
- Forecasts may be spatially defined as a “threat area” for example, or expressed on a grid (models)
  - Restricted set of scales
  - Correlated in space and time
- Observations come as scattered point values
  - All scales represented, but valid only at station
  - Undersampled as field
- Forecast to observation techniques:
  - Ask: What is the forecast at the verification location?
  - Recommended way to go for verification – Leave the observation value alone.
  - Interpolation to the observation location – for smooth variables
  - Nearest gridpoint – for “episodic” or spatially categorical variables
  - Observation is left as is except for QC
  - Sometimes verification is done with respect to remotely sensed data by transforming the model forecast into “what the satellite would see if that forecast were to be correct”

Data matching issues (2)

- Observation to forecast techniques (really for modelers):
  - Upscaling – averaging over gridboxes – only if that is truly the definition of the forecast (model) E.g. Cherubini et al 2002
    - Local verification
  - Verify only where there is data!

Precipitation verification project : methodology - Europe

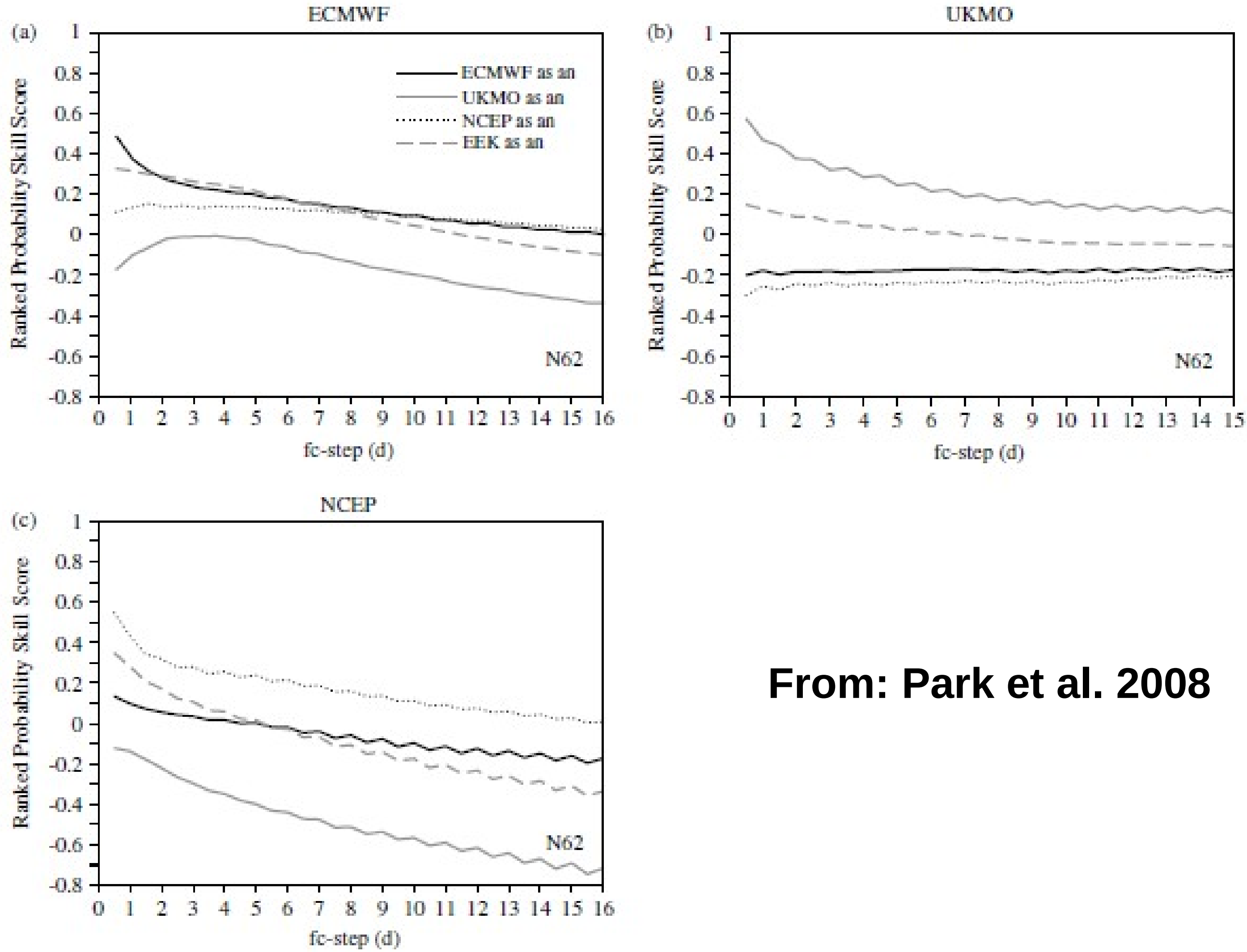
- Upscaling:
  - 1x1 gridboxes, limit of model resolution
  - Average obs over grid boxes, at least 9 stns per grid box (Europe data)
  - Verify only where enough data
  - Answers questions about the quality of the forecasts within the capabilities of the model
  - Most likely users are modelers.



Data matching issues (2)

- Observation to model techniques:
  - Upscaling – averaging over gridboxes – only if that is what the model predicts. E.g. Cherubini et al 2002
    - Local verification
  - Analysis of observation data onto model grid
    - Frequently done, but not a good idea for verification except for some kinds of model studies.
    - Analysis using model-independent method e.g. Barnes
    - Analysis using model-dependent method – data assimilation (bad idea for verification!) e.g. Park et al 2008

The effects of different “truths”

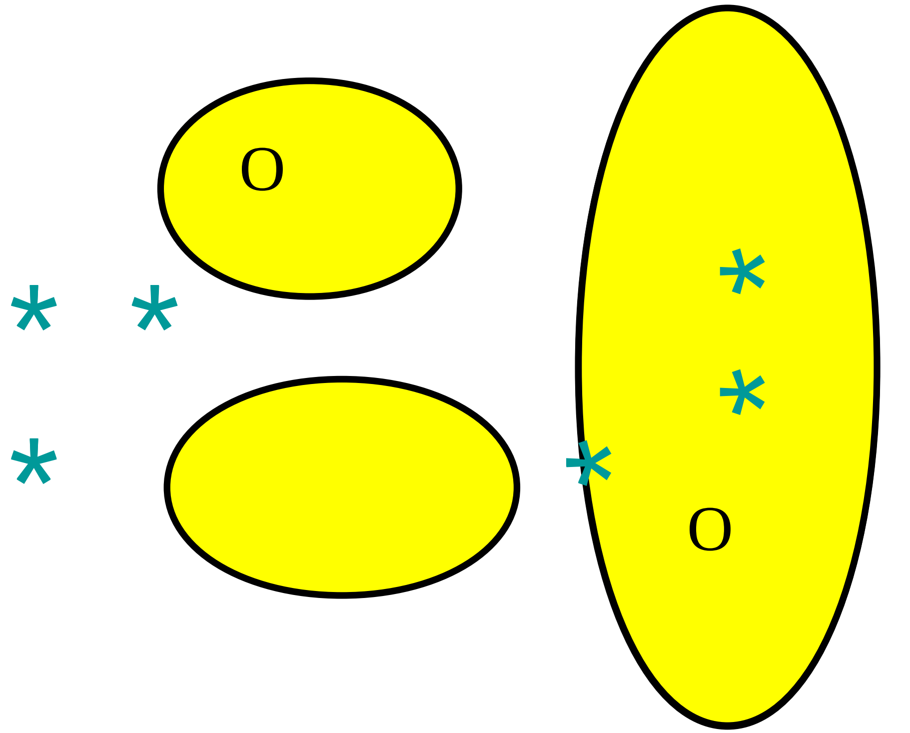


From: Park et al. 2008

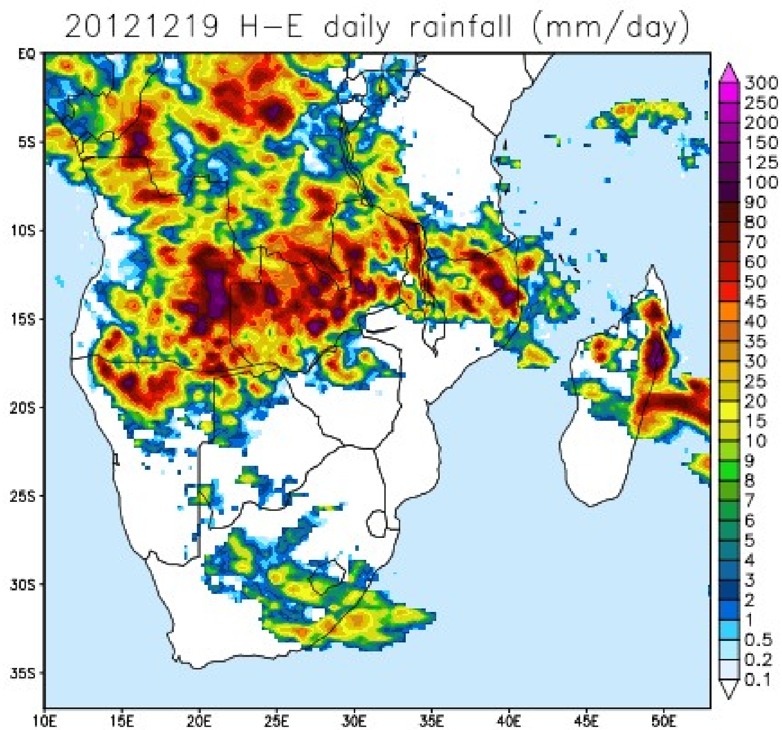
Das Ende – The End - Fini

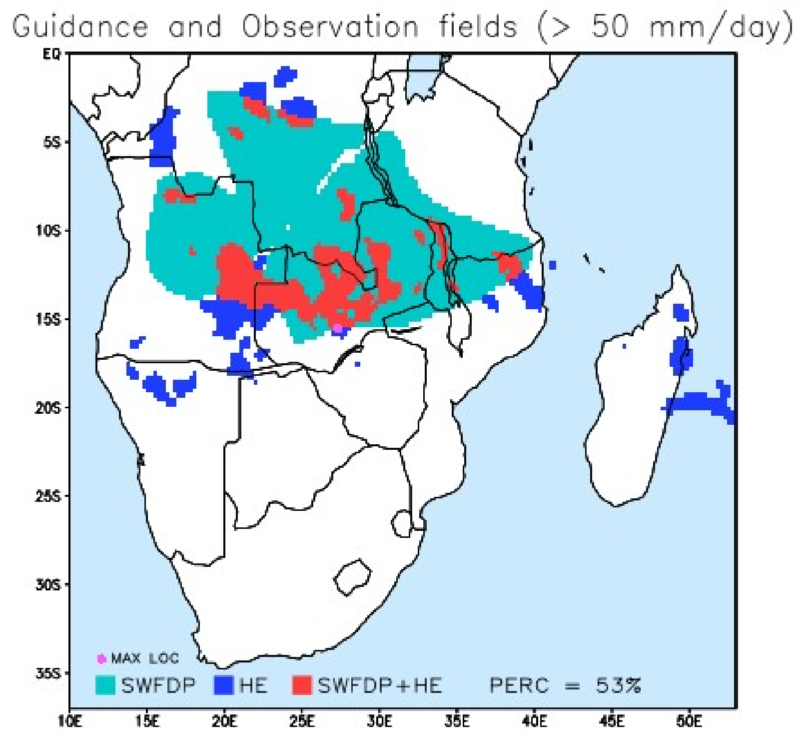
Matching point obs with areally defined forecasts: what is the Event?

- For categorical forecasts, one must be clear about the "event" being forecast
  - Location or area for which forecast is valid
  - Time range over which it is valid
  - Definition of category
- And now, what is defined as a correct forecast?
  - The event is forecast, and is observed – anywhere in the area? Over some percentage of the area?
  - Scaling considerations



Verification of regional forecast map using HE





Verification statistics for 20121219 : Grid Size = 0.25° : Units = mm/day : n = 25777

	Guidance	H–E
Number of gridpoints >= 50 mm	3294	1243
Average Rain over domain	~	19.7012
>= 50 mm Rain Area (km²*10⁶)	2.05875	0.776875
Maximum Rainfall Observed (mm)	~	151.124

Categorical Forecasts

Frequency Bias	2.65004
Probability of Detection	0.526146
False Alarm Ratio	0.801457
Hansen & Kuipers Score	0.418541
Equitable threat score	0.132959
Spatial Correlation	0.264835

GUIDANCE \ OBSERVATION	>=50	<50
>=50	654	2640
<50	589	21894

Extreme Events Verification

Extreme Dependency Score	0.650434
Symmetric Extreme Dependency Score	0.385181
Extremal Dependency Index	0.552717
Symmetric Extremal Dependency Index	0.59486

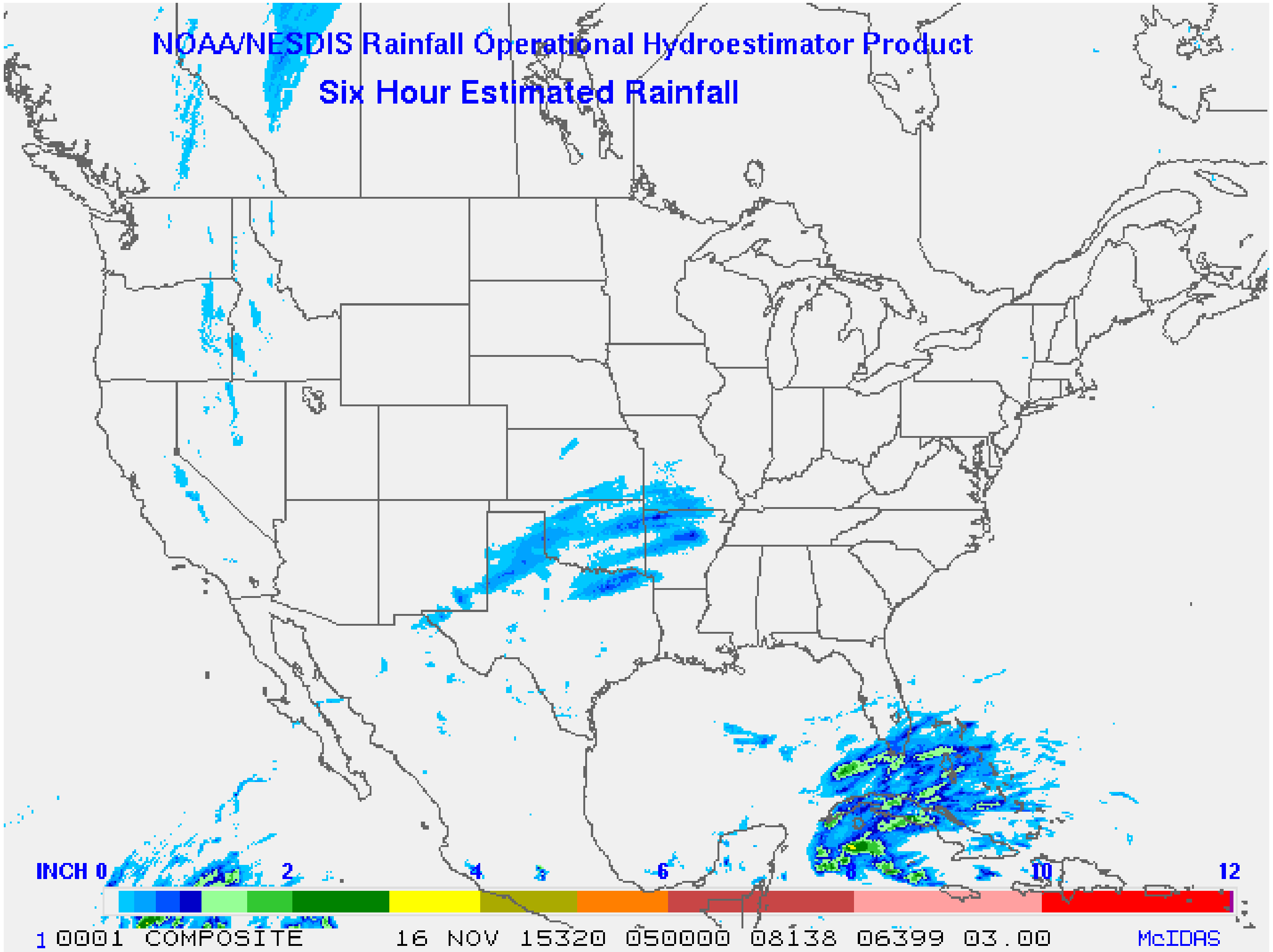
(**Ferro and Stephenson, 2011***)

http://rsmc.weathersa.co.za/RSMC/index.php
Format based on IPWG verification output

WMO

South African Weather Service
ISO 9001 Certified Organisation

US Precipitable water estimates



Collecting data for verification

- Archive forecasts AND observations
  - Your own: station observations AND corresponding forecasts
  - Most NWP centers archive their forecasts and observations; if you use their model, you can probably get them to give you relevant data for verification.
- Goal: Generate matched set of forecasts and observations