6th WMO International Verification Methods Workshop
Bayesian model verification –
Predictability of convective conditions
based on EPS forecasts and observations

Andreas Röpnack, Deutscher Wetterdienst
Andreas Hense, University of Bonn
Motivation

Convection

- **Unstable stratification**
  vertical structure of temperature and moisture is significant for convective initiation
  (Crook, 1996), (Keil et al., 2008)

- **Trigger**

(Doswell, 1993)
Observations

Radiosondes

- Measurement error
- Representativeness error
- Position error

Observation uncertainty
- data-assimilation of DWD

\[
\begin{align*}
\bar{f}_{avg} & = \frac{1}{3} \left( \bar{f}_1 + \bar{f}_2 + \bar{f}_3 \right) \\
\bar{f}_{cor} & = \begin{bmatrix} \bar{f}_1^T \\ \bar{f}_2^T \\ \bar{f}_3^T \end{bmatrix}^T
\end{align*}
\]
**Forecasts**

**Ensemble prediction systems (EPS)**

- Representation of **forecast uncertainty** via Initial-conditions, Boundary-conditions, Model-physics.

<table>
<thead>
<tr>
<th></th>
<th>COSMO-DE-EPS</th>
<th>COSMO-SREPS</th>
<th>COSMO-LEPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member</td>
<td>20</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Grid size</td>
<td>2.8km</td>
<td>10km</td>
<td>10km</td>
</tr>
<tr>
<td>Vertical levels</td>
<td>50</td>
<td>40</td>
<td>40</td>
</tr>
</tbody>
</table>

*Consortium for small-scale modelling*

*COSMO short-range EPS*  
*COSMO limited-area EPS*  
*(COSMO; Non-hydrostatic Limited-Area model of DWD)*  
*(COSMO-SREPS)*  
*(COSMO-LEPS)*
Aim

Probabilistic method \textit{(Min and Hense, 2006)}

\rightarrow \textbf{Verification} and \textbf{comparison} of EPS forecasts

\rightarrow \text{Consideration of forecasts- and observation uncertainty}

\textbf{Application 1:} \textbf{Comparison}
COSMO-DE-EPS vs. COSMO-SREPS

\textbf{Application 2:} \textbf{Verification}
COSMO-SREPS and COSMO-LEPS vs. COSMO-EU analysis
1. Motivation and aim
2. Bayesian theory
3. Comparison COSMO-DE-EPS vs. COSMO-SREPS
4. Verification SREPS / LEPS with COSMO-EU analysis
5. Summary
Bayesian approach

- Probability of a **hypothesis** for a given **observation** (*Malakoff, 1999*)

- Extension of the classical frequentist statistics, here degree of personal conviction as a probability (*Jaynes, 2003*)

\[
P(H | D) = \frac{\ell(D | H) P(H)}{P(D)}
\]

- **Posterior probability**
- **Likelihood**
- **Prior probability**

\(H:\) hypothesis → **model-forecast** \(m\)
\(D:\) data → **observation** \(o\)
Conditional probability

\[ P(m \mid o) \propto \ell(o \mid m) P(m) \]

- **Posterior**
- **Prior**
  - Personal believe
  - Prior – variation

- **Likelihood**
  - Data

Probability density functions (PDFs)
- \( p_m \) Model
- \( p_o \) Observation
- \( p_o \cdot p_m \)
**Likelihood**

Multivariate normal distribution

\[
p_o = N(o, \Sigma_o)
\]

\[
p_m = \frac{1}{K_m} \sum_{k=1}^{K_m} N_k(f_k, \Sigma_m)
\]

*(Röpnack et al., 2013)*

*Kernel Dressing (Bröcker, 2007), (Schölzel and Hense, 2010)*
Bayes factor

\[ B_{ir} = \frac{P(m_i \mid o)}{P(m_{ref} \mid o)} \]

- \( B_{ir} > 1 \) meaning \( m_i \) is more likely than \( m_{ref} \)
- \( B_{ir} < 1 \) meaning \( m_{ref} \) is more likely than \( m_i \)

\( \Rightarrow \) Classification depends on the application (comparison or verification)!
1. Motivation und aim
2. Bayesian Theory
3. Comparison of COSMO-DE-EPS vs. COSMO-SREPS
4. Analysis
5. August 2007
Comparison of ensembles

<table>
<thead>
<tr>
<th>log B_{ir}</th>
<th>Evidence for ensemble</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;5</td>
<td>Decisive for DE-EPS</td>
</tr>
<tr>
<td>2.5 – 5.0</td>
<td>Strong for DE-EPS</td>
</tr>
<tr>
<td>1.0 – 2.5</td>
<td>Substantial for DE-EPS</td>
</tr>
<tr>
<td>-1.0 – 1.0</td>
<td>Neutral</td>
</tr>
<tr>
<td>-2.5 – -1.0</td>
<td>Substantial for SREPS</td>
</tr>
<tr>
<td>-5.0 – -2.5</td>
<td>Strong for SREPS</td>
</tr>
<tr>
<td>&lt; -5</td>
<td>Decisive for SREPS</td>
</tr>
</tbody>
</table>

*after Kass and Raftery (1995)*
DE-EPS vs. SREPS, 00UTC

Univariate – temperature@850hPa (1500m)

No distinction possible

→ Neutral evidence
DE-EPS vs. SREPS, 00UTC

Multivariate – temperature@8Levels

Temperature-profiles

DE-EPS more likely as SREPS

⇒ Decisive evidence
for DE-EPS if three stations
unaveraged threatened together
DE-EPS vs. SREPS, 00UTC

Variation of observation-covariance matrix

Observation errors are important

\[ \Sigma_0 = \gamma \Sigma_o \]

\[ \Rightarrow \text{The evidence of DE-EPS is weaker for larger observation errors} \]
DE-EPS vs. SREPS, 00UTC

Prior – Variation

\[
P(m | o) = \frac{\ell(o | m)P(m)}{r(o)}
\]

Strong evident for DE-EPS
Weak evident for DE-EPS

Robust result
1. Motivation and aim
2. Bayesian Theory
3. Comparison of COSMO-DE-EPS vs. COSMO-SREPS
4. Verification SREPS / LEPS with COSMO-EU analysis
5. June, July, August 2007
## Verification of ensembles

<table>
<thead>
<tr>
<th>log B_ir</th>
<th>Validation against analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;5</td>
<td>-</td>
</tr>
<tr>
<td>2.5 – 5.0</td>
<td>-</td>
</tr>
<tr>
<td>1.0 – 2.5</td>
<td>-</td>
</tr>
<tr>
<td>-1.0 – 1.0</td>
<td>High level of confidence</td>
</tr>
<tr>
<td>-2.5 – -1.0</td>
<td>Medium level</td>
</tr>
<tr>
<td>-5.0 – -2.5</td>
<td>Low level</td>
</tr>
<tr>
<td>&lt; -5</td>
<td>Very low level</td>
</tr>
</tbody>
</table>

_after Kass and Raftery (1995)_
SREPS vs. COSMO-EU analysis

SREPS vs. COSMO-EU analysis

COSMO-SREPS

Temperature

High

Medium

Low

Level of confidences

Eq.-pot temperature

COSMO-LEPS

18. March 2014

Bayesian model verification
Summary

Bayesian model verification

⇒ explicit consideration of the observation uncertainty
⇒ multivariate application

Application 1: DE-EPS vs. SREPS

- COSMO-DE-EPS temperature-profiles are more likely
- Observation errors are important
- Robust results (Prior-variation)

Application 2: SREPS/LEPS vs. COSMO-EU analysis

- Verification of case studies as well as of longer periods are possible
- Linear decrease of the forecast quality after vv=72h transition from „medium“ to „low“ forecast quality

⇒ Good judgment of forecasted vertical profiles with one score
⇒ Method is appropriated to compare as well as to verify ensembles
sensorimotor learning...


Thank you!
Questions?

Appendix – COSMO-DE-EPS
Appendix – Multivariate aspect

Verticale ensemble-profile

\[ \begin{pmatrix}
  f_{200\text{hPa},m1} & \cdots & f_{200\text{hPa},m16} \\
  \vdots & \ddots & \vdots \\
  f_{1000\text{hPa},m1} & \cdots & f_{1000\text{hPa},m16}
\end{pmatrix} \]

Model-levels

Ensemble-Member

\[
\bar{f}_{\text{avg}} = \frac{1}{3} (\bar{f}_1 + \bar{f}_2 + \bar{f}_3)
\]

\[
\bar{f}_{\text{cor}} = \begin{bmatrix} \bar{f}_1^T & \bar{f}_2^T & \bar{f}_3^T \end{bmatrix}^T
\]
Appendix – DE-EPS vs. SREPS, 00UTC

Temperature

Eq.-pot.-temperature

vv time / h

log B_{ir}

three stations together

arithmetic mean of three stations

vv time / h

three stations together

arithmetic mean of three stations
Appendix – DE-EPS vs. SREPS, 00UTC

Prior – Variation

\[
P(m \mid o) = \frac{\ell(o \mid m)P(m)}{r(o)}
\]
DE-EPS vs. SREPS, 00UTC

Bayes factor

Ignorance skill score

\[
\log B_{ir}
\]

\[
\text{temperature@850hPa}
\]

\[
\text{temperature@850hPa}
\]
Continuous Ranked Probability Skill Score

CRPSS = 1 - \frac{CRPS_{DE-EPS}}{CRPS_{SREPS}}

CRPS = \int_{-\infty}^{\infty} [F(t) - H(t-o)]^2 \, dt

Integral of the Brier Score over all possible thresholds

(Gneiting et al., 2004)
Continuous Ranked Probability Skill Score

August 2007

-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0

temperature@850hPa

without SKD

Standard

with SKD

Result is similar to Bayes factor, however without consideration of the observation errors