Multi-Model Ensemble Wake Vortex Prediction

Stephan Körner *, Frank Holzäpfel *, Nash'at Ahmad+

* German Aerospace Center (DLR)
  Institute of Atmospheric Physics

+ NASA Langley Research Center

WakeNet 2015, Amsterdam
22.04.2015
Multi-Model Ensemble Wake Vortex Prediction

- Vortex position
- Decay
- Circulation
- APA3.2
- D2P
- NASA
- TDP2.1
- APA3.4
- DLR
- Ground effect

Reliability
Ensemble Averaging

Deterministic forecast

Proabilistic forecast

Bayesian Model Averaging

Outperform individual models

En-route

Airports
Revision of P2P - Motivation

wake vortex descent

\[ \Gamma^* = \Gamma / \Gamma_0, \quad z^* = z / b_0, \quad y^* = y / b_0, \quad t^* = t / t_0 \]

- Initial vortex spacing: \( b_0 \)
- Initial vortex descent speed: \( w_0 \)

88 landings

70 landings

\[ \Gamma^* = \Gamma / \Gamma_0, \quad z^* = z / b_0, \quad y^* = y / b_0, \quad t^* = t / t_0 \]
Revision of P2P

wake vortex descent

- secondary vortices weakened by 30 % after first orbit (0.28 * $\Gamma_0$)

- tertiary vortices weakened by 30 % from the beginning (0.28 * $\Gamma_0$)

- vortex-ground interaction above $b_0$: not yet further investigated

- vortex ground interaction not only distance but also time dependent?
Revision of P2P

bias = model - observation
Revision of P2P

![Graph showing revision of P2P metrics for different locations (MUC, FRA, OP). The graphs compare various versions of P2P, including 'no rev', 'img vtx rev', 'tert vtx rev', 'sec vtx rev', and the average of 'img vtx rev' and 'tert vtx rev'. The metrics evaluated are rmse \( y^*_\text{luf t} \) and \( z^*_\text{luf t} \).]
Multi Model Ensemble

Sugar

How to mix several good ingredients?

Water

Lemon Juice

Lemonade
Why not use the best ensemble member exclusively?

- which is the best member?
- in average best performing member can sometimes be the worst one

Can an ensemble outperform its best member?

- success of ensemble appr.: any model can be the best sometimes
- consistently low performing models → no increase of skill

Yes!

Hagedorn et al., 2005
Ensemble Members

NASA-DLR cooperation

**D2P**
- deterministic output of P2P
- based on decaying potential vortex, adapted to LES results (DLR)

**TDP 2.1**
- considers effect of crosswind shear on vortex descent (NASA)

**APA 3.2**
- decay and transport model according to Sarpkaya (NASA)

**APA 3.4**
- reduced effect of stratification (NASA)

Probability that one of the models delivers the best forecast
(in ground-effect, on the basis of rmse for 99 example cases)
Multi-Model Ensemble

Reliability Ensemble Averaging (REA)

\[ \tilde{f} = \tilde{A}(f) = \frac{\sum_i R_i f_i}{\sum_i R_i} \]

\[ R_i = [(R_{B,i})^m \cdot (R_{D,i})^n]^{1/(m \cdot n)} = \left\{ \left[ \frac{nv}{abs(B_i)} \right]^m \left[ \frac{nv}{abs(D_i)} \right]^n \right\}^{1/(m \cdot n)} \]

- \( f_i \): forecast of model i
- \( \tilde{A}(f) \): REA-forecast
- \( R_i \): reliability factor of model i
- \( nv \): natural variability
- \( B_i \): absolute bias of model i
- \( D_i \): absolute difference between forecast of model i and ensemble mean

Giorgi and Mears, 2002
Multi-Model Ensemble

Reliability Ensemble Averaging

\[ R_{D,i} \text{ depends on distance to ensemble mean:} \]

\[ z \]

\[ t \]

\[ \text{ensemble mean} \]

natural variability

if bias or distance to ensemble mean < \( nv \) → model reliable (\( R_{B,i} \) or \( R_{D,i} = 1 \))

\[ \text{natural variability} \]

\[ \rightarrow nv = \text{model resolution limit} \]
Multi-Model Ensemble

Reliability Ensemble Averaging

uncertainty bounds depend on ensemble spread

\[ \tilde{\delta}_f = \left[ \bar{A}(f_i - \bar{f})^2 \right]^{1/2} = \left[ \frac{\sum_i R_i (f_i - \bar{f})^2}{\sum_i R_i} \right]^{1/2} \]

\[ f_+ = \bar{f} + \tilde{\delta}_f \]

\[ f_- = \bar{f} - \tilde{\delta}_f \]

according to Giorgi and Mearns, 2002
Application to Wake Vortex Models

Reliability Ensemble Averaging

Training
- mixture of landings from WakeFRA, WakeMUC and WakeOP
- 95 selected cases

$R_{B,i}$ and $R_{D,i}$
- $R_{B,z,i}(t)$, $R_{B,y,i}(t)$, $R_{B,\Gamma,i}(t)$, $R_{D,z,i}(t)$, $R_{D,y,i}(t)$, $R_{D,\Gamma,i}(t)$
- $\Delta t^* = 2 t_0$
- separately for luff and lee vortices
- weights for reliability factors: $R_{B,z,i} : m=1.0$, $R_{D,z,i} : n=0.3$

Uncertainty envelope
- initial condition uncertainty added (not considered in original approach):

<table>
<thead>
<tr>
<th>variable</th>
<th>unit</th>
<th>$\sigma$ (standard deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>true airspeed</td>
<td>[m/s]</td>
<td>4</td>
</tr>
<tr>
<td>air density</td>
<td>[kg/m³]</td>
<td>0.0048</td>
</tr>
<tr>
<td>weight</td>
<td>[kg]</td>
<td>1300</td>
</tr>
<tr>
<td>$z_0$</td>
<td>[m]</td>
<td>7</td>
</tr>
<tr>
<td>$y_0$</td>
<td>[m]</td>
<td>25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>variable</th>
<th>unit</th>
<th>$\sigma$ (standard deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>if initial conditions derived from lidar:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_0$</td>
<td>[m]</td>
<td>9</td>
</tr>
<tr>
<td>$y_0$</td>
<td>[m]</td>
<td>13</td>
</tr>
<tr>
<td>$\Gamma_0$</td>
<td>[m²/s]</td>
<td>13</td>
</tr>
</tbody>
</table>
**Application to Wake Vortex Models**

### REA natural variability, $\Gamma^*$

<table>
<thead>
<tr>
<th>N* = N*t₀</th>
<th>$\epsilon^<em>$ = ($\epsilon</em>b₀$)$^{1/3}/w₀</th>
<th>$v^*$ = $v/w₀$</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>$N^* &lt; 0.3, \epsilon^* &gt; 0.25$</th>
<th>$N^* &lt; 0.3, \epsilon^* &lt; 0.25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{obs,\Gamma_{luff}}$</td>
<td>0.097 (1073)</td>
</tr>
<tr>
<td>$\sigma_{obs,\Gamma_{lee}}$</td>
<td>0.081</td>
</tr>
</tbody>
</table>

- $\epsilon^* > 0.2$, $N^* < 0.3$
- $\epsilon^* < 0.2$, $N^* < 0.3$
Application to Wake Vortex Models

**REA natural variability, $z^*$**

\[
\sigma_{obs} = \sqrt{\sigma_{err}^2 + \sigma_{nv}^2}
\]

| $N^* < 0.3$, $|v^*| > 0.5$ | $N^* < 0.3$, $|v^*| < 0.5$ |
|--------------------------|--------------------------|
| $\sigma_{obs,zz_{uff}}$  | 0.35 (1224)              |
| $\sigma_{obs,zz_{ece}}$  | 0.35                     |
|                          | 0.200 (805)              |
|                          | 0.262                    |

| $|v^*| > 1.0$, $N^* < 0.3$ | $|v^*| < 1.0$, $N^* < 0.$ |
|---------------------------|---------------------------|
| $z^* - z_0$               | $z^* - z_0$               |
| 0.5                       | 0.5                       |
| 1.0                       | 1.0                       |
| 1.5                       | 1.5                       |
| -0.5                      | -0.5                      |
| -1.0                      | -1.0                      |
| -1.5                      | -1.5                      |

\[t^*\]
Results

REA forecast
(one single landing)

enhancement:

\[
\begin{align*}
\text{rmse}_{z^*,\text{TDP}} &= 0.158 \\
\text{rmse}_{z^*,\text{REA}} &= 0.148 \\
\text{rmse}_{\Gamma^*,\text{D2P}} &= 0.085 \\
\text{rmse}_{\Gamma^*,\text{REA}} &= 0.072
\end{align*}
\]

probability levels according to
- 99 testcases
- WakeFRA & WakeOP
Results

REA reliability factors (one single landing)

no correlation between $R_D$ and $R_B$ can be found!
## Results

### REA scoring

- 99 randomly chosen cases
- Skill factor $s_i$:

$$s_i = \frac{\sum_{p=1}^{n} \frac{rmse_{e,p}}{rmse_{i,p}}}{n} - 1$$

<table>
<thead>
<tr>
<th>median</th>
<th>rms $\Gamma_{luff}^*$</th>
<th>rms $\Gamma_{lee}^*$</th>
<th>rms $y_{luff}$</th>
<th>rms $y_{lee}^*$</th>
<th>rms $\bar{z}_{luff}^*$</th>
<th>rms $\bar{z}_{lee}^*$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>REA</td>
<td>0.116</td>
<td>0.099</td>
<td>0.433</td>
<td>0.409</td>
<td>0.163</td>
<td>0.149</td>
<td>0.000</td>
</tr>
<tr>
<td>TDP 2.1</td>
<td>0.127</td>
<td>0.106</td>
<td>0.590</td>
<td>0.415</td>
<td>0.237</td>
<td>0.147</td>
<td>-0.122</td>
</tr>
<tr>
<td>APA 3.4</td>
<td>0.160</td>
<td>0.107</td>
<td>0.602</td>
<td>0.413</td>
<td>0.179</td>
<td>0.154</td>
<td>-0.127</td>
</tr>
<tr>
<td>APA 3.2</td>
<td>0.204</td>
<td>0.140</td>
<td>0.612</td>
<td>0.410</td>
<td>0.179</td>
<td>0.154</td>
<td>-0.190</td>
</tr>
<tr>
<td>D2P</td>
<td>0.122</td>
<td>0.120</td>
<td>0.406</td>
<td>0.408</td>
<td>0.140</td>
<td>0.166</td>
<td>-0.016</td>
</tr>
</tbody>
</table>
### Results

#### REA scoring

- 99 randomly chosen cases
- Skill factor $s$:

\[
s_i = \frac{\sum_{p=1}^{n} \text{rmse}_{e,p}/\text{rmse}_{i,p}}{n} - 1
\]

<table>
<thead>
<tr>
<th>median</th>
<th>$\text{rms} , \Gamma_{lu}^*$</th>
<th>$\text{rms} , \Gamma_{le}^*$</th>
<th>$\text{rms} , y_{lu}^*$</th>
<th>$\text{rms} , y_{le}^*$</th>
<th>$\text{rms} , z_{lu}^*$</th>
<th>$\text{rms} , z_{le}^*$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>REA</td>
<td>0.116</td>
<td>0.099</td>
<td>0.433</td>
<td>0.409</td>
<td>0.163</td>
<td>0.149</td>
<td>0.000</td>
</tr>
<tr>
<td>DEA</td>
<td>0.131</td>
<td>0.093</td>
<td>0.525</td>
<td>0.411</td>
<td>0.173</td>
<td>0.149</td>
<td>-0.048</td>
</tr>
</tbody>
</table>

Advanced MME approach outperforms Direct Ensemble Average (DEA)
PDD of models and ensemble

**overconfident ensemble:**
ensemble spread too low

**well-dispersed ensemble:**
coverage of full spectrum of possible solutions

Weigel et al., 2008, Hagedorn et al., 2004

well-dispersed model forecasts → rmse improvement

overconfident ensemble → small or no rmse improvement
Conclusion

- ensemble **can improve quality** of wake vortex forecasts on average
- however **only 1.6 %** improvement compared to best model
  
  **reasons:** - ensemble is **overconfident** for \( z^* \) and \( y^* \)
  - uncertainties from env. and natural variability dominate model uncertainty
- **but:** models might behave differently in particular ambient weather conditions and out-of-ground → investigation with pdds
Further Development

- How does a good training data set look like?
- Can the results be further improved by distinguishing various weather conditions?
- How does the Bayesian Model Averaging (BMA) perform?

[source: Raftery et al., 2005]
Backup
Results

REA forecast reliability (one single landing)

\[ \tilde{\rho} = \frac{\sum_i R_i^2}{\sum_i R_i} \]

- low reliability for \( y \) - forecast
- high reliability for \( z \) - forecast
- medium reliability for \( \Gamma \) - forecast

forecast reliability \( \rho \)
Wake Vortex Predictions

Motivation

1. - optimization of tactical separation at airports
   - hazard warning system

2. - Wake Encounter Avoidance & Advisory System (WEAA)
   - “Free Flight”
Ensemble Methoden
Bayesian Model Averaging

P(B) = Wahrscheinlichkeit des Eintretens von B
P(B|A) = Wahrscheinlichkeit für B, unter Vorraussetzung A
PDF = Probability Density Function (Wahrscheinlichkeitsdichtefunktion)

Law of total probability:

\[ P(B) = \sum_{n} P(B \cap A_n) = \sum_{n} P(A_n) P(B|A_n) \]

Beispiel:
Wir befinden uns auf einem Schiff:
- wir wollen die Position B bestimmen
- 3 Crew-Mitglieder (A1,A2,A3) wissen wie es geht, haben aber unterschiedliche Methoden

according to Grimmett and Welsh., 1986
Ensemble Methoden
Bayesian Model Averaging

Law of total probability:

\[ P(B) = \sum_{n} P(B \cap A_n) = \sum_{n} P(A_n) P(B|A_n) \]

<table>
<thead>
<tr>
<th>Methode</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
</tr>
</thead>
<tbody>
<tr>
<td>individuelle Wahrscheinlichkeit, dass die Methode Erfolg hat: P(B</td>
<td>A_n)</td>
<td>0.6</td>
<td>0.9</td>
</tr>
<tr>
<td>Wahrscheinlichkeit, dass wir A1, A2 or A3 fragen: P(A_n)</td>
<td>0.2</td>
<td>0.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>

\[ s = \sum_{n} v_n \cdot t_n \]

\[ P(B) = 0.78 \]
Ensemble Methoden
Bayesian Model Averaging

Law of total probability:

\[ P(B) = \sum_n P(B \cap A_n) = \sum_n P(A_n) P(B|A_n) \]

Methode

<table>
<thead>
<tr>
<th>Methode</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDF der Methode (Modell-Unsicherheiten): P(B</td>
<td>A_n)</td>
<td>![Diagram A1]</td>
<td>![Diagram A2]</td>
</tr>
<tr>
<td>Wahrscheinlichkeit, dass wir A1, A2 oder A3 fragen: P(A_n)</td>
<td>0.1</td>
<td>0.6</td>
<td>0.3</td>
</tr>
</tbody>
</table>
Ensemble Methoden
Bayesian Model Averaging

Law of total probability:
\[ P(B) = \sum_n P(B \cap A_n) = \sum_n P(A_n) P(B \mid A_n) \]

angewandt auf Vorhersage-Modelle:
Annahme: es gibt immer ein bestes Ensemble-Glied

\[ A_n = \text{Modell } n \]
\[ B = \text{vorherzusagende Größe} \]
\[ B^T = \text{Trainings-Daten} \]
\[ P(A_n) = \text{Wahrscheinlichkeit, dass } A_n \text{ das beste Modell ist} \]
\[ (\text{Gewichtungsfaktor}, \text{basierend auf } B^T) \]
\[ P(B \mid A_n) = \text{PDF of } A_n \text{ alone (Gaussian distribution, given that } A_n \text{ is the best forecast)} \]

\[ \triangleq \text{gewichtete Summe von Wahrscheinlichkeitsdichtefunktionen (PDFs)} \]

according to Raftery et al., 2005
Ensemble Methoden
Bayesian Model Averaging

BMA applied on 48-h surface temperature forecast (bias corrected)

- ensemble forecast
- individual model PDF
- individual model forecast
- 90% interval
- verification

source: Raftery et al., 2005
Multi-Model Ensemble

benefit

- increase deterministic skill
- predict forecast skill
- provide probabilistic forecast

individual model forecasts

average

probabilistic envelope
Multi-Model Ensemble

- Lidar port and starboard data plotted against time (t [s]) with labels for different models (APA 3.2, APA 3.4, TDAWP 2.1, D2P).
- Y-axis for z [m] and y [m] showing vortex development over time.
- Gamma (Γ [m²/s]) plotted against time (t [s]) for various models.