Long evaluations of cloud in SAM at the SGP

Peter Henderson   &   Robert Pincus

University of Colorado–CIRES
NOAA–ESRL–PSD

Fort Collins, July 2008

Project outline

- ▶ Goal: to evaluate cloud in SAM

- ▶ Using 3-year runs

- ▶ Evaluating POC (aka cloud fraction) against ARSCL

  - ▶ with and without temporal averaging

- ▶ Using observations to define cloud-occurrence in model

  - ▶ via radar-simulator & lidar-proxy

Project outline

- ▶ Goal: to evaluate cloud in SAM

- ▶ Using 3-year runs

- ▷ Evaluating POC (aka cloud fraction) against ARSCL

  - ▶ with and without temporal averaging

- ▷ Using observations to define cloud-occurrence in model

  - ▶ via radar-simulator & lidar-proxy

Project outline

- ▶ Goal: to evaluate cloud in SAM
- ▶ Using 3-year runs
- ▶ Evaluating POC (aka cloud fraction) against ARSCL
    - ▶ with and without temporal averaging
- ▶ Using observations to define cloud-occurrence in model
    - ▶ via radar-simulator & lidar-proxy

Project outline

- ► Goal: to evaluate cloud in SAM
- ► Using 3-year runs
- ► Evaluating POC (aka cloud fraction) against ARSCL
  - ► with and without temporal averaging
- ► Using observations to define cloud-occurrence in model
  - ► via radar-simulator & lidar-proxy

Project outline

- ▶ Goal: to evaluate cloud in SAM
- ▶ Using 3-year runs
- ▶ Evaluating POC (aka cloud fraction) against ARSCL
  - ▶ with and without temporal averaging
- ▶ Using observations to define cloud-occurrence in model
  - ▶ via radar-simulator & lidar-proxy

Project outline

- ▶ Goal: to evaluate cloud in SAM
- ▶ Using 3-year runs
- ▶ Evaluating POC (aka cloud fraction) against ARSCL
    - ▶ with and without temporal averaging
- ▶ Using observations to define cloud-occurrence in model
   - ▶ via radar-simulator & lidar-proxy

Model runs

▶ 3-year runs of SAM (1999–2002) at SGP, forced with ARM data (M. Zhang, J. Lin, S. Xie, et al)

▶ Comparing results of 2D, 2D+IPHOC and 3D configurations

    ▶ 2D: 32 x 1 x 28; $\Delta x = 4$ km (MMF-like)

    ▶ 3D: 128 x 128 x 64; $\Delta x = \Delta y = 500$ m

Model runs

- ▶ 3-year runs of SAM (1999–2002) at SGP, forced with ARM data (M. Zhang, J. Lin, S. Xie, et al)
- ▶ Comparing results of 2D, 2D+IPHOC and 3D configurations
  - ▶ 2D: 32 x 1 x 28; $\Delta x = 4$ km (MMF-like)
  - ▶ 3D: 128 x 128 x 64; $\Delta x = \Delta y = 500$ m

Model runs

- 3-year runs of SAM (1999–2002) at SGP, forced with ARM data (M. Zhang, J. Lin, S. Xie, et al)
- Comparing results of 2D, 2D+IPHOC and 3D configurations
  - 2D: 32 x 1 x 28; $\Delta x = 4$ km (MMF-like)
  - 3D: 128 x 128 x 64; $\Delta x = \Delta y = 500$ m

Performance metrics

- $MSE = bias^2 + random^2$

- $b = reliability - resolution + uncertainty$

2D cloud-errors: $\tau_T \in \{2, 4, 8, 16, \ldots, 512, 1024, inf\}$
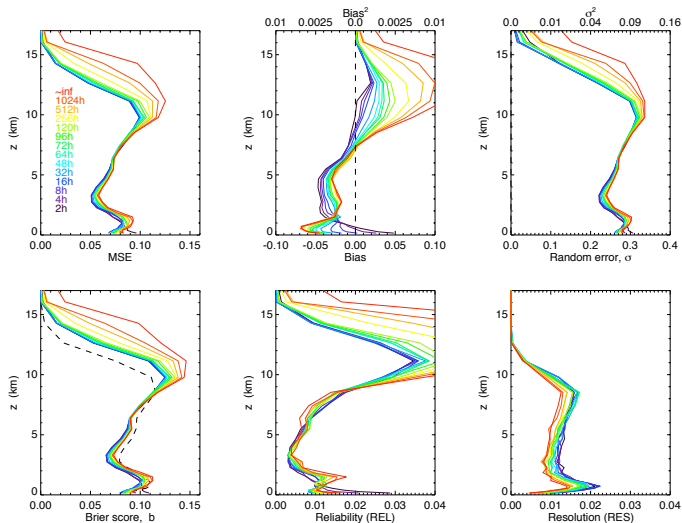
Cloud errors in 2D, 2D+IPHOC & 3D runs; ($\tau_T = 24$)
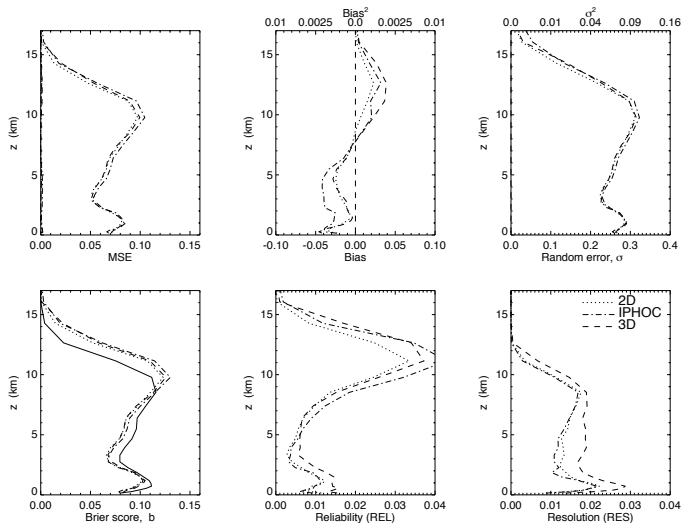
Brier Score, $b = $ limiting case of MSE
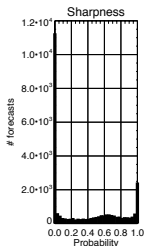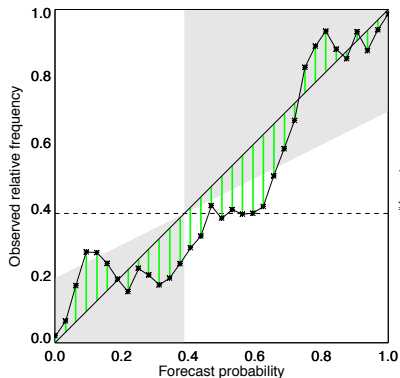
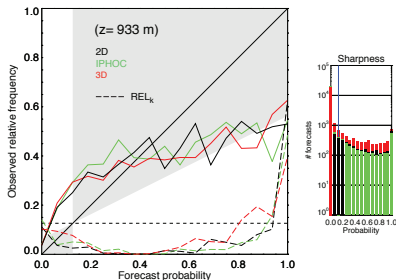$$b = \frac{1}{N} \sum_{i=1}^{N} (p_i - o_i)^2$$

$$b = \textit{reliability} - \textit{resolution} + \textit{uncertainty},$$

$$= \frac{1}{N} \sum_{k=1}^{K} n_k (p_k - \bar{o}_k)^2 - \frac{1}{N} \sum_{k=1}^{K} n_k (\bar{o}_k - \bar{o})^2 + \bar{o}(1 - \bar{o})$$

Brier Score, b = limiting case of MSE
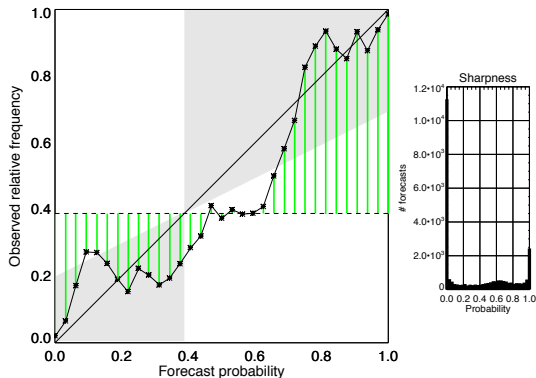
$$b = \frac{1}{N} \sum_{i=1}^{N} (p_i - o_i)^2$$

$$b = reliability - resolution + uncertainty,$$

$$= \frac{1}{N} \sum_{k=1}^{K} n_k (p_k - \bar{o}_k)^2 - \frac{1}{N} \sum_{k=1}^{K} n_k (\bar{o}_k - \bar{o})^2 + \bar{o}(1 - \bar{o})$$

Brier Score, b = limiting case of MSE
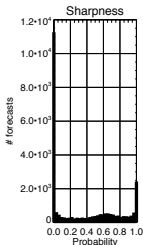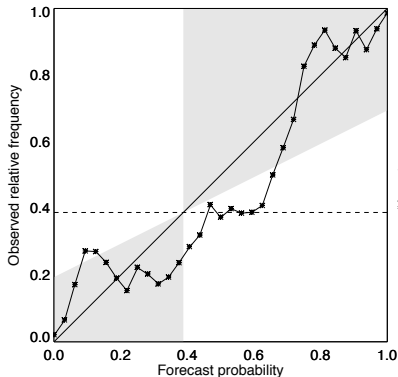
$$b = \frac{1}{N} \sum_{i=1}^{N} (p_i - o_i)^2$$

Measures variability
of observations only;
max at $\bar{o} = \frac{1}{2}$.

$$b = reliability - resolution + uncertainty,$$

$$= \frac{1}{N} \sum_{k=1}^{K} n_k (p_k - \bar{o}_k)^2 - \frac{1}{N} \sum_{k=1}^{K} n_k (\bar{o}_k - \bar{o})^2 + \bar{o}(1 - \bar{o})$$

Weightings of conditional biases (reliability)

- ▶ All over-predict most p
- ▶ 3D has least bias due to extreme p & most for mid–high p
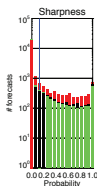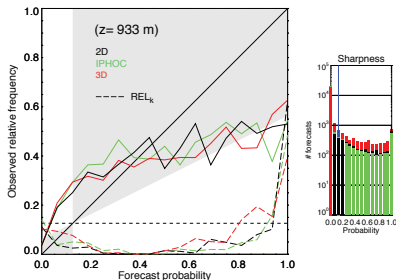- ▶ IPHOC most biased ∴ $p = 1$ contributions

- ▶ All under-predict low p
- ▶ All over-predict high p
- ▶ 3D most biased ∴ non-extreme contributions
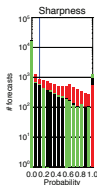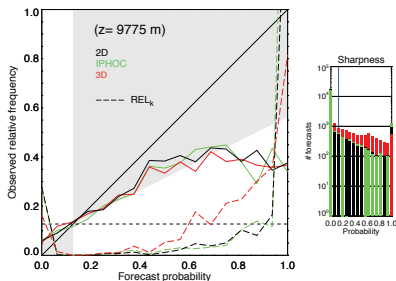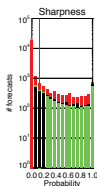
Weightings of conditional biases (reliability)

- All over-predict most p
- 3D has least bias due to extreme p & most for mid–high p
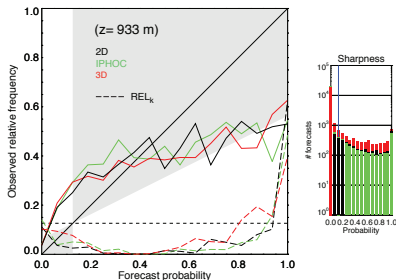- IPHOC most biased ∵ $p = 1$ contributions

- All under-predict low p
- All over-predict high p
- 3D most biased ∵ non-extreme contributions
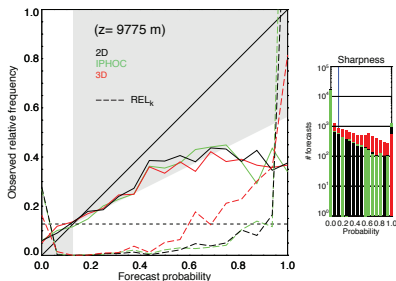
Weightings of conditional biases (reliability)

- ▶ All over-predict most p
- ▶ 3D has least bias due to extreme p & most for mid–high p
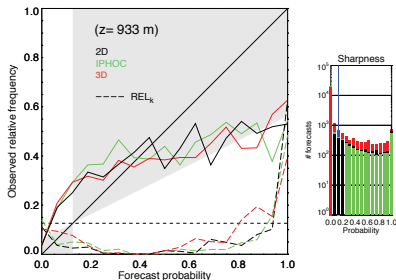- ▶ IPHOC most biased ∵ $p = 1$ contributions

- ▶ All under-predict low p
- ▶ All over-predict high p
- ▶ 3D most biased ∵ non-extreme contributions

Weightings of conditional biases (reliability)

- All over-predict most p
- 3D has least bias due to extreme p & most for mid–high p
- IPHOC most biased ∵ $p = 1$ contributions

- All under-predict low p
- All over-predict high p
- 3D most biased ∵ non-extreme contributions
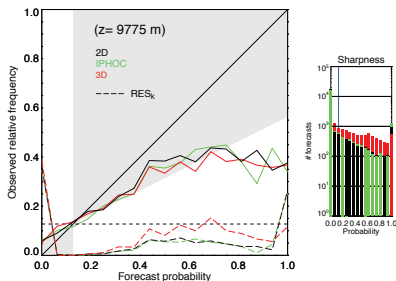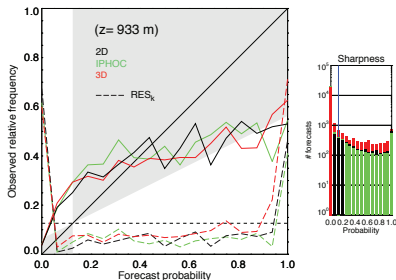
Weightings of conditional biases (reliability)

- All over-predict most p
- 3D has least bias due to extreme p & most for mid–high p
- IPHOC most biased $\because$ $p = 1$ contributions

- All under-predict low p
- All over-predict high p
- 3D most biased $\because$ non-extreme contributions
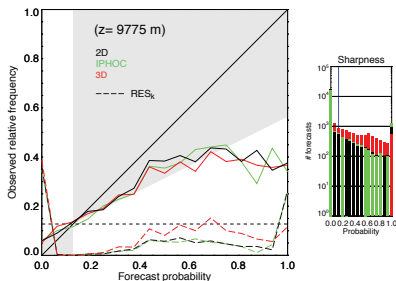
Weightings of timing-of-events (resolution)

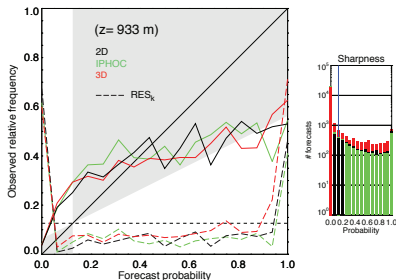- Most contributions for $p = 0, 1$
- 3D best ∵ mid–high p

- Even more $p = 0, 1$
- Clear-sky ($p = 0$) contributions similar in all runs
- 3D best ∵ most high p
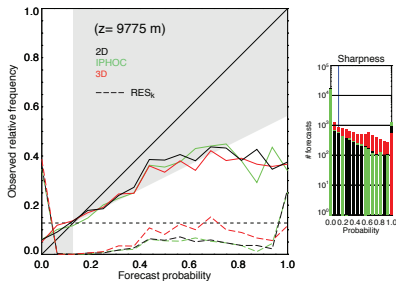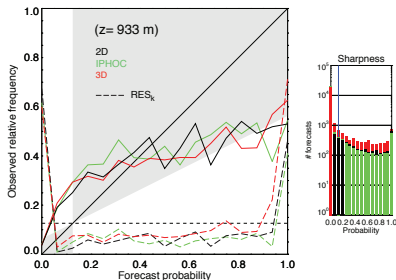
Weightings of timing-of-events (resolution)

▶ Most contributions for
  $p = 0, 1$

▶ 3D best ∵ mid–high p

▶ Even more $p = 0, 1$

▶ Clear-sky ($p = 0$)
  contributions similar in all
  runs

▶ 3D best ∵ most high p
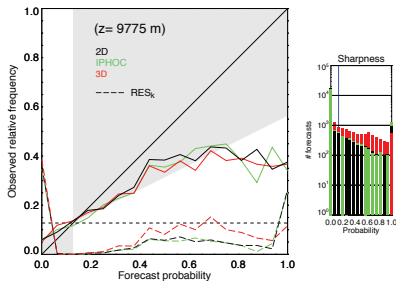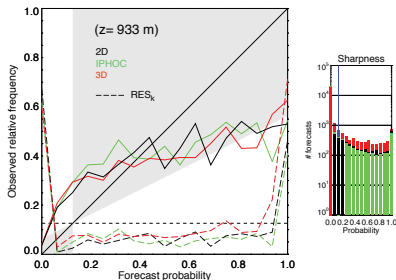
Weightings of timing-of-events (resolution)

- Most contributions for $p = 0, 1$
- 3D best $\therefore$ mid–high p

- Even more $p = 0, 1$
- Clear-sky ($p = 0$) contributions similar in all runs
- 3D best $\therefore$ most high p

Weightings of timing-of-events (resolution)

- Most contributions for $p = 0, 1$
- 3D best ∴ mid–high p

- Even more $p = 0, 1$
- Clear-sky ($p = 0$) contributions similar in all runs
- 3D best ∴ most high p
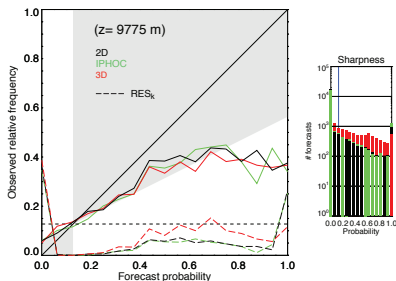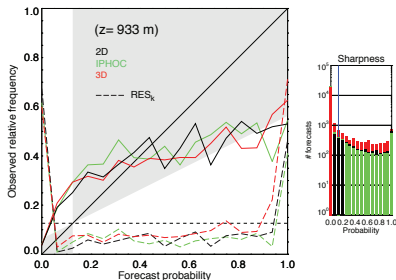
Weightings of timing-of-events (resolution)

- Most contributions for $p = 0, 1$
- 3D best $\therefore$ mid–high p

- Even more $p = 0, 1$
- Clear-sky ($p = 0$) contributions similar in all runs
- 3D best $\therefore$ most high p

Some conclusions

▶ $b \sim \mathrm{MSE} \implies$ temporal stats $\sim$ spatial stats

▶ Small bias $\implies$ model well calibrated (overall)

▶ Range in {RES, RES} for 2D-3D $\sim$ range seen over τ_T
  $\implies$ significant differences between configurations

▶ Conditional bias & timing best for low cloud

▶ 3D has best timing at all z, but most conditional bias in PBL

Some conclusions

- $b \sim \mathrm{MSE} \implies$ temporal stats $\sim$ spatial stats

- Small bias $\implies$ model well calibrated (overall)

- Range in {RES, RES} for 2D-3D $\sim$ range seen over τ_T
  $\implies$ significant differences between configurations

- Conditional bias & timing best for low cloud

- 3D has best timing at all z, but most conditional bias in PBL

Some conclusions

- $b \sim \mathrm{MSE} \implies$ temporal stats $\sim$ spatial stats

- Small bias $\implies$ model well calibrated (overall)

- Range in $\{\mathrm{RES, RES}\}$ for 2D-3D $\sim$ range seen over τ_T
  $\implies$ significant differences between configurations

- Conditional bias & timing best for low cloud

- 3D has best timing at all z, but most conditional bias in PBL

Some conclusions

- $b \sim \mathrm{MSE} \implies$ temporal stats $\sim$ spatial stats

- Small bias $\implies$ model well calibrated (overall)

- Range in {RES, RES} for 2D-3D $\sim$ range seen over τ_T
  $\implies$ significant differences between configurations

- Conditional bias & timing best for low cloud

- 3D has best timing at all z, but most conditional bias in PBL

Some conclusions

- $b \sim \mathrm{MSE} \implies$ temporal stats $\sim$ spatial stats

- Small bias $\implies$ model well calibrated (overall)

- Range in {RES, RES} for 2D-3D $\sim$ range seen over τ_T
  $\implies$ significant differences between configurations

- Conditional bias & timing best for low cloud

- 3D has best timing at all z, but most conditional bias in PBL
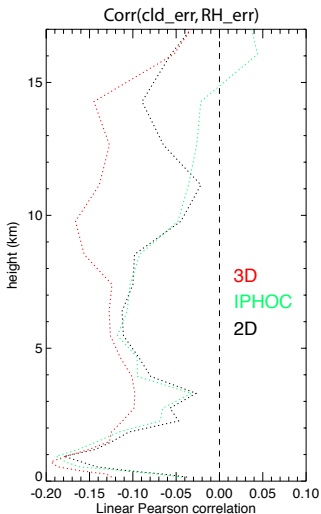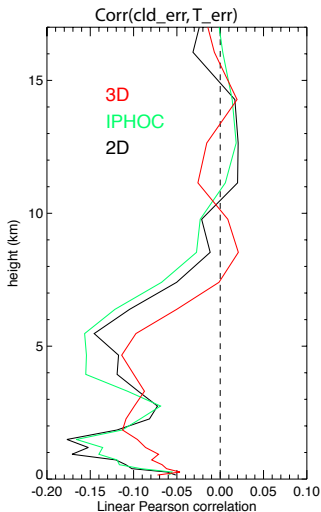
When do differences in performance occur?
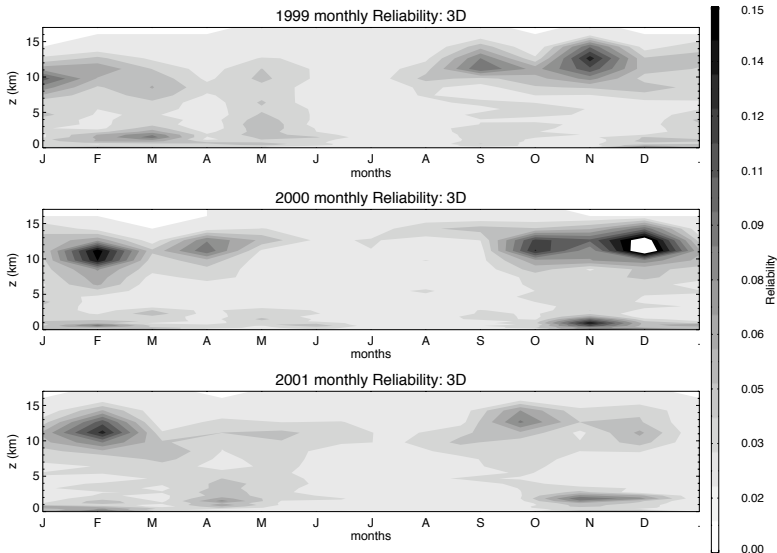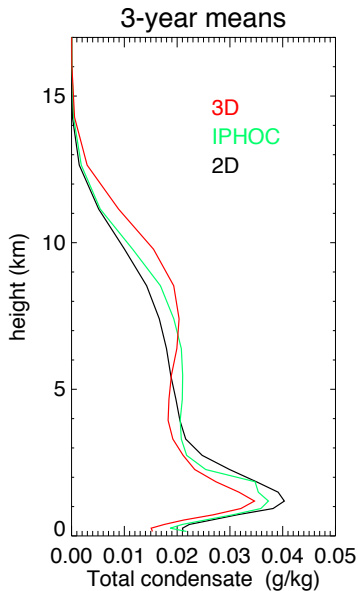
Extra slide: ARSCL monthly cloud fraction
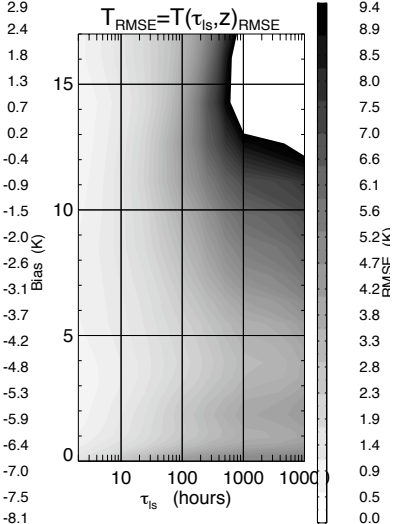
Extra slide:
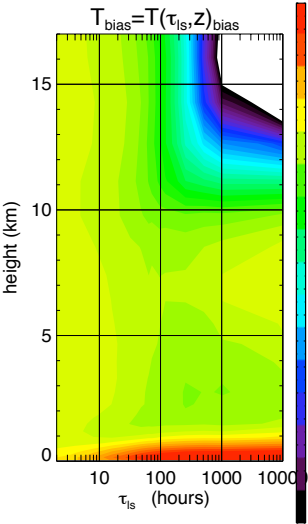

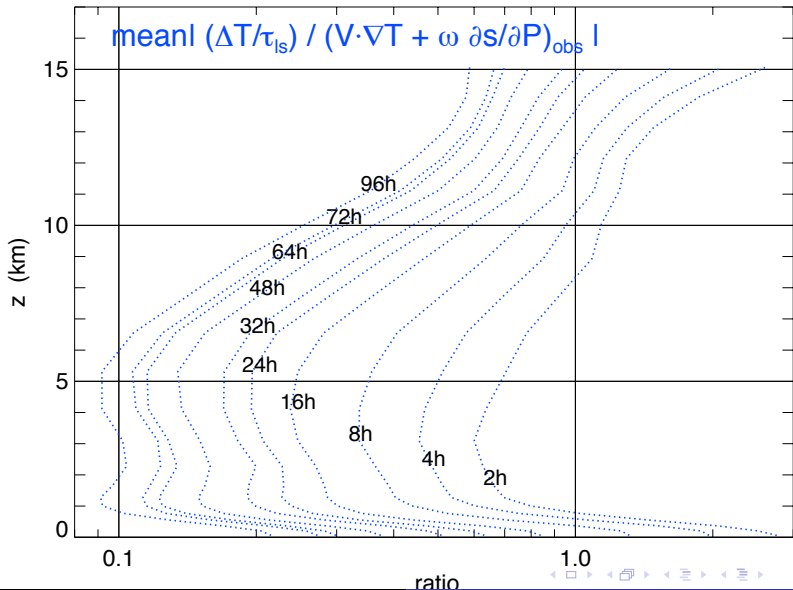Ratios of nudging & LS-advective tendency terms, for T
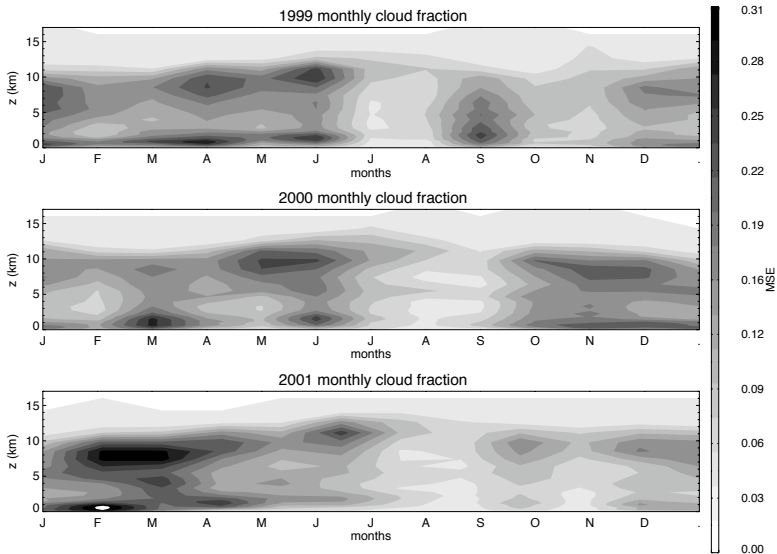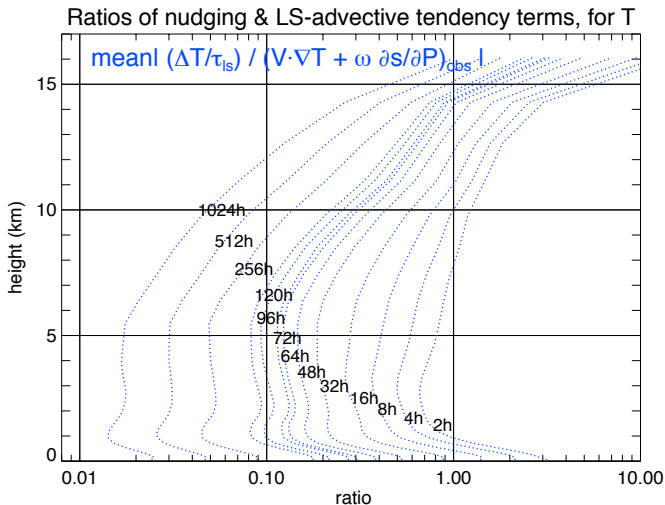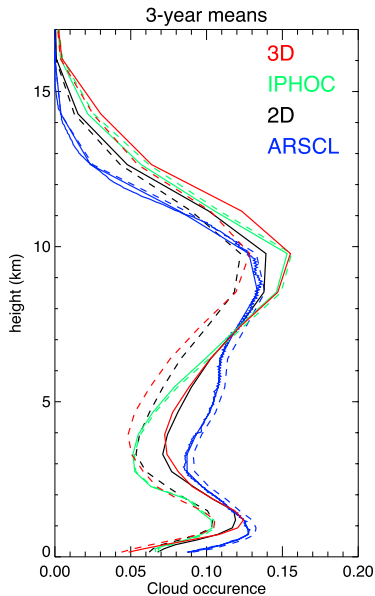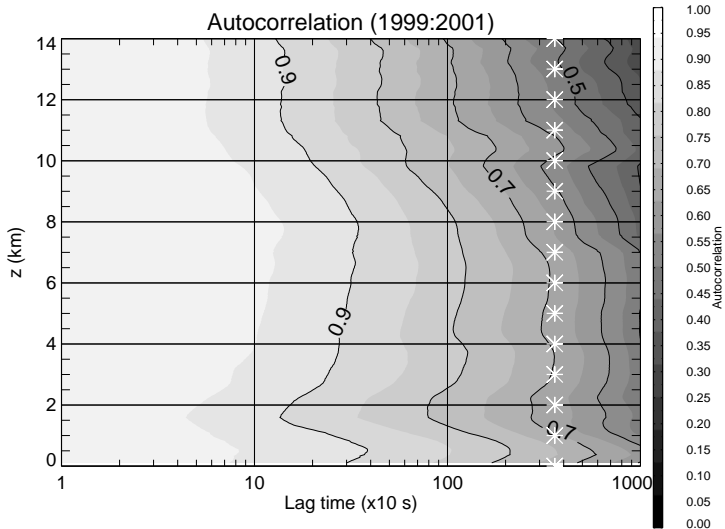
Extra slide:

Extra slide:

Extra slide:

Extra slide: Ratios of nudging & advective tendency of T

Autocorrelation (1999:2001)

Extra slide: model grid

Model configuration: 32 x 1 x 28 (4 km, 20 s)