# Pitfalls when comparing ICON output with ERA5 reanalysis data

Henning Franke

Max-Planck-Institut
für Meteorologie

# Background

- **Aim:** Finding an ICON configuration (vertical diffusion, Rayleigh damping) which simulates a realistic stratospheric circulation
- Known reference state needed for comparison: ERA5 reanalysis
- ICON simulations initialized from ERA5 state → Basically identical at first timestep
- Assumption: Zonal mean weather is predictable for ~10 days → ICON should approximately reproduce ERA5 for up to 10 days
- Changes in ICON configuration which reduce the bias to ERA5 are interpreted as model improvement

# First pitfall: Vertical coordinate transformation

Hydrostatic equation:

$$\frac{dp}{dz} = -\rho g$$

Using ideal gas law:

$$\frac{dp}{dz} = -\frac{p}{T_v R_l} g$$

Rearranging:

$$\frac{1}{p} dp = -\frac{g}{T_v R_l} dz$$

Integration:

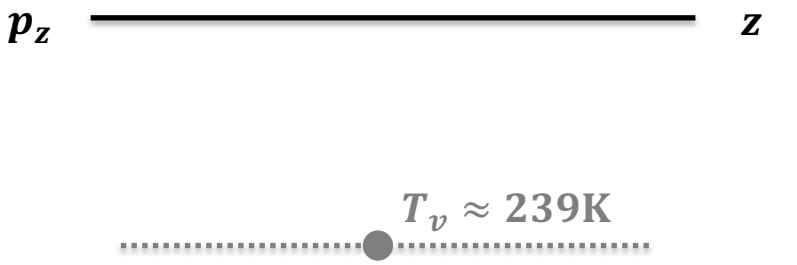
$$\ln\left(\frac{p_{\text{up}}}{p_{\text{low}}}\right) = -\frac{g}{T_v R_l}\left(z_{\text{up}} - z_{\text{low}}\right)$$

Rearranging:

$$z_{\text{up}} = z_{\text{low}} + \frac{T_v R_l}{g} \cdot \ln\left(\frac{p_{\text{low}}}{p_{\text{up}}}\right)$$

# Vertical coordinate transformation: Approximation

$$z_{\text{up}} = z_{\text{low}} + \frac{T_v R_l}{g} \cdot \ln\left(\frac{p_{\text{low}}}{p_{\text{up}}}\right)$$




Isothermal atmosphere ($T_v \approx 239K$):

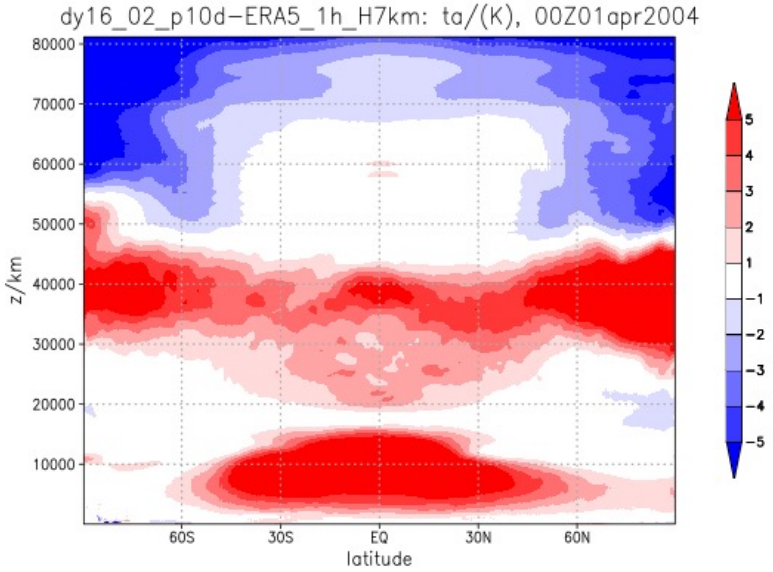
$$\frac{T_v R_l}{g} = H = 7000\text{m}$$

$$z = z_{\text{sfc}} + H \cdot \ln\left(\frac{p_{\text{sfc}}}{p(z)}\right)$$

```
cdo ml2hl,hlevels infile outfile
```

# Approximation of fixed scale height does not hold

Temperature bias between ICON and ERA5 at model initialization

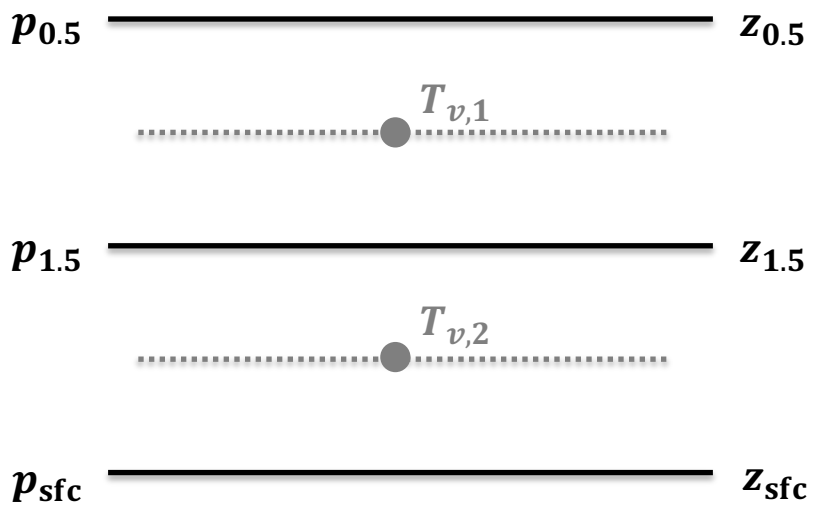



Plot: Marco Giorgetta

# Vertical coordinate transformation: Second approach

$$z_{\mathrm{up}} = z_{\mathrm{low}} + \frac{T_v R_l}{g} \cdot \ln\left(\frac{p_{\mathrm{low}}}{p_{\mathrm{up}}}\right)$$

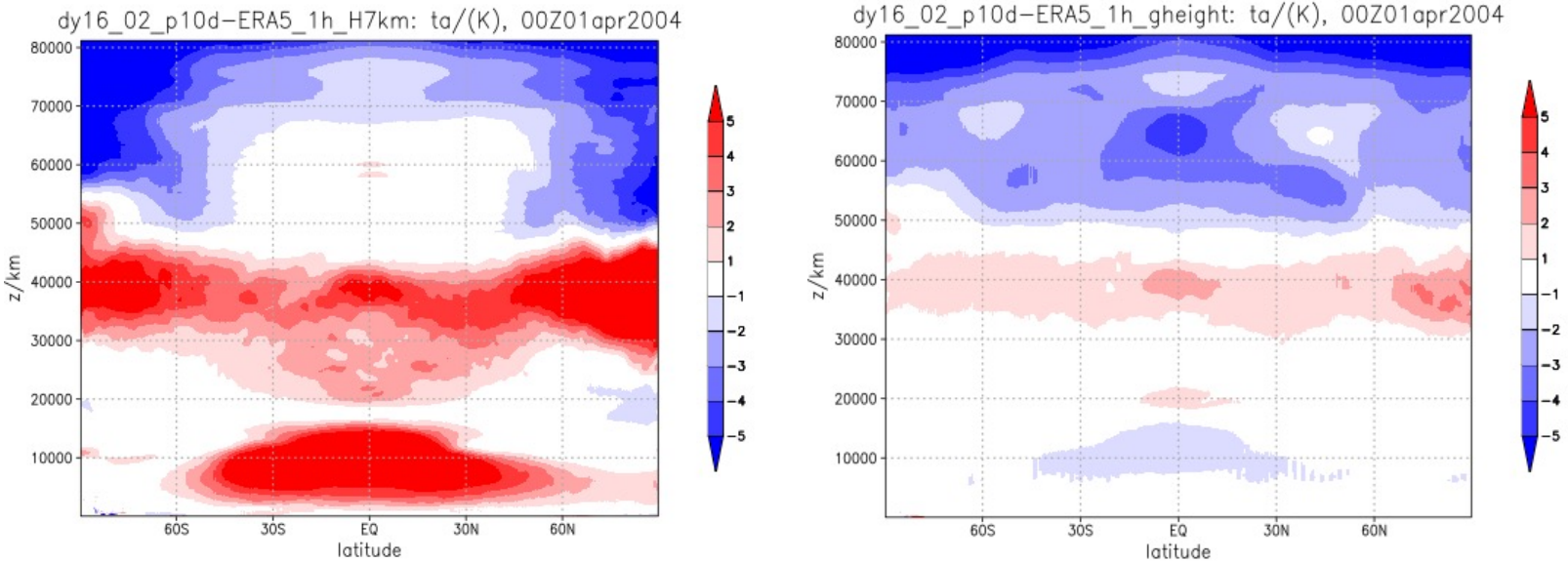
→ Sum up layer by layer




```
cdo gh2hl,hlevels -gheight infile outfile
```

# Much smaller but still significant biases

Temperature bias between ICON and ERA5 at model initialization

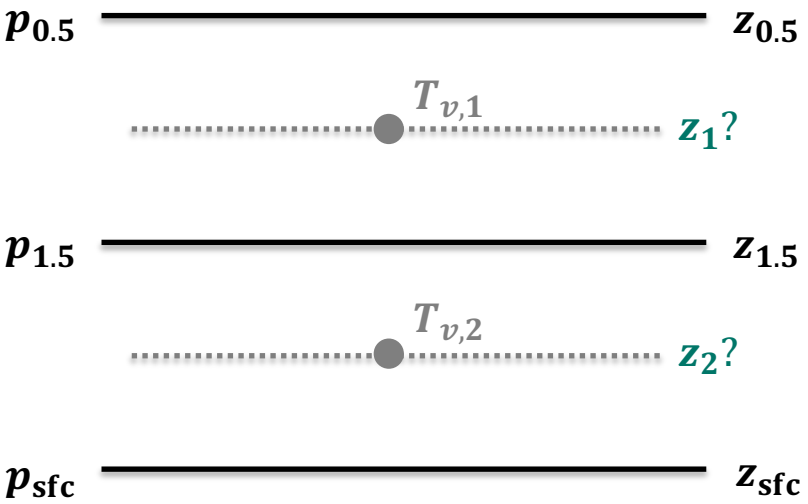



Plots: Marco Giorgetta

# A more detailed look into the calculation of geopotential height

$$z_{\text{up}} = z_{\text{low}} + \frac{T_v R_l}{g} \cdot \ln\left(\frac{p_{\text{low}}}{p_{\text{up}}}\right)$$

Geopotential height on half levels is calculated exactly




**How should $z$ interpolated to full (model) levels?**

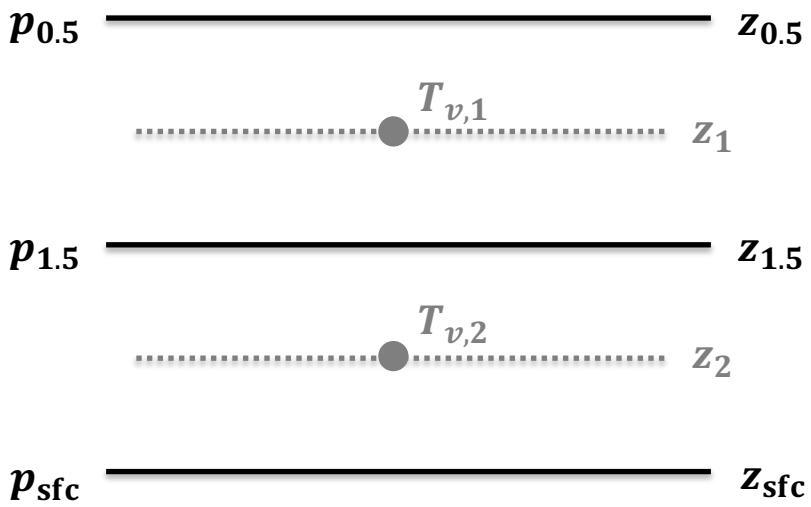
# Vertical coordinate transformation

$$z_{\mathrm{up}} = z_{\mathrm{low}} + \frac{T_v R_l}{g} \cdot \ln\left(\frac{p_{\mathrm{low}}}{p_{\mathrm{up}}}\right)$$

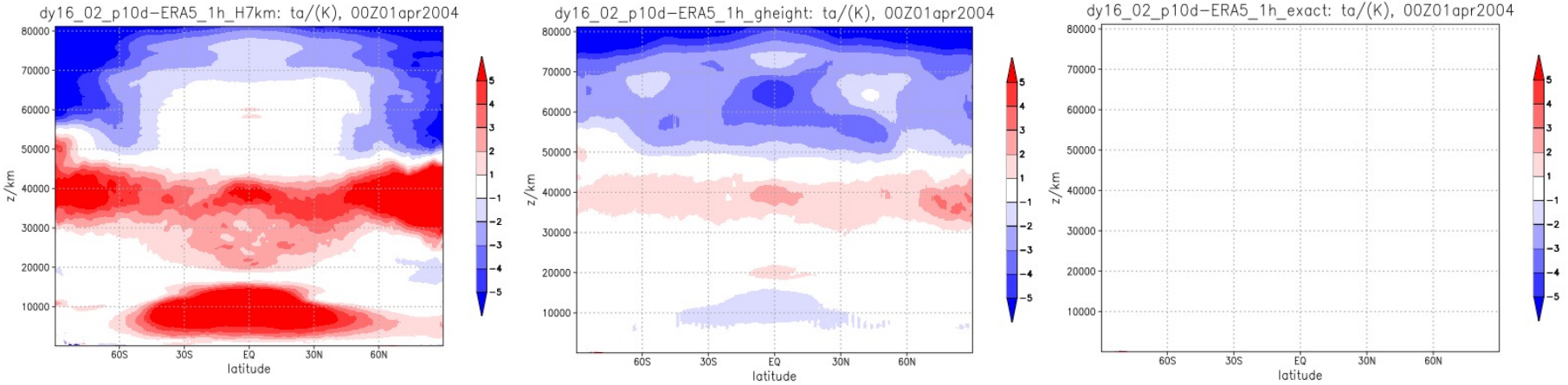
$$z_{\mathrm{fl}} = z_{\mathrm{hl,\,low}} + \alpha \frac{T_v R_l}{g}$$

$$\alpha = 1 - \frac{p_{\mathrm{low}}}{p_{\mathrm{up}} - p_{\mathrm{low}}} \cdot \ln\left(\frac{p_{\mathrm{low}}}{p_{\mathrm{up}}}\right)$$

Roeckner et al. (2003)



# Finally negligible biases

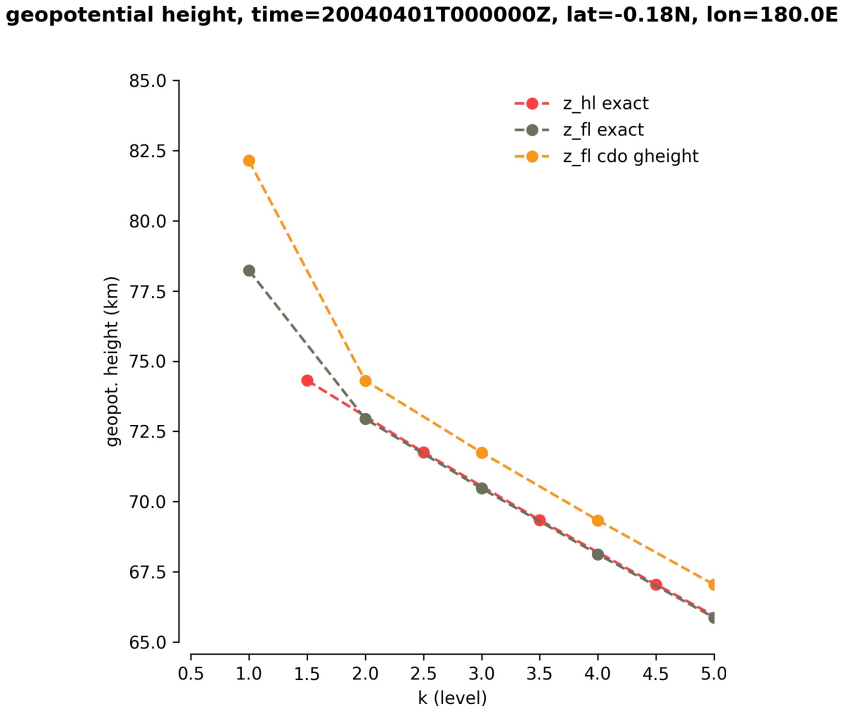



Plots: Marco Giorgetta

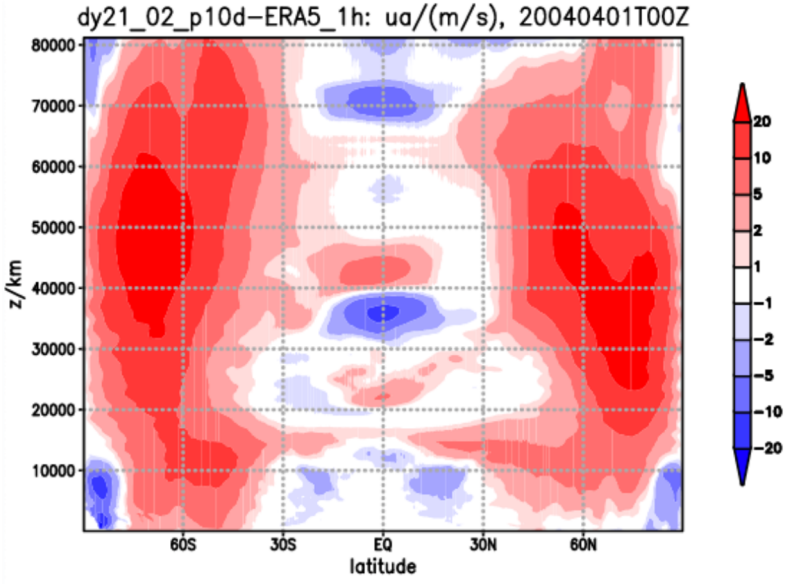
→ Geopotential height calculated by CDO operator gheight is not exact!

# How does the cdo operator gheight work?

$$z_{\mathrm{up}} = z_{\mathrm{low}} + \frac{T_v R_l}{g} \cdot \ln\left(\frac{p_{\mathrm{low}}}{p_{\mathrm{up}}}\right)$$



# Second pitfall: ERA5 horizontal wind data




Plot: Marco Giorgetta

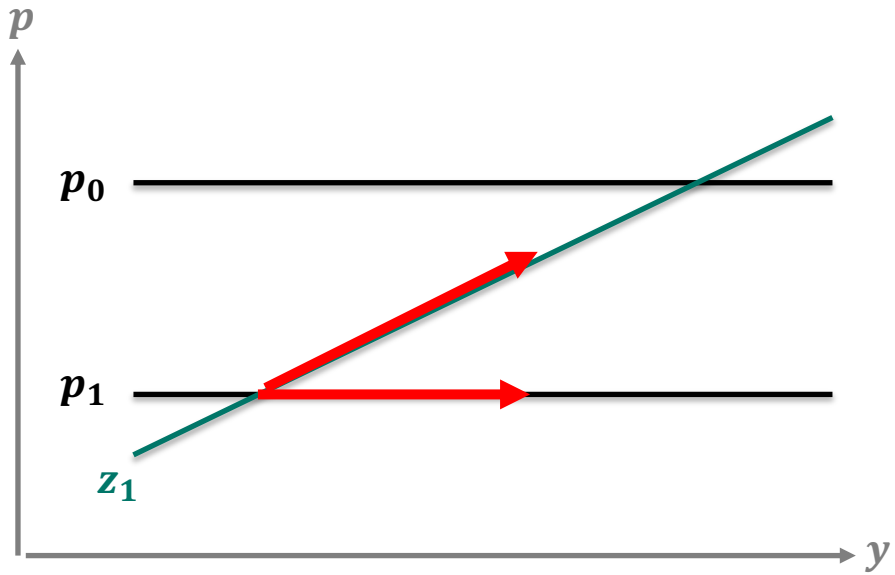
- ERA5 horizontal wind data on model levels is provided in **m·cos($\varphi$)/s** instead of **m/s**
- Scaling by 1/cos($\varphi$) reduces zonal wind bias significantly
- Documented in DKRZ Readme

# Third pitfall: Transformation of vertical velocity

**Vertical velocity in $p$-coordinates $\omega$ (Pa/s) must be transformed to $w$ (m/s)!**

General transformation of a derivative $\partial x/\partial y$ from $p$-coordinates to $z$-coordinates:

$$\left(\frac{\partial x}{\partial y}\right)_z = \left(\frac{\partial x}{\partial y}\right)_p + \frac{\partial x}{\partial p}\left(\frac{\partial p}{\partial y}\right)_z$$



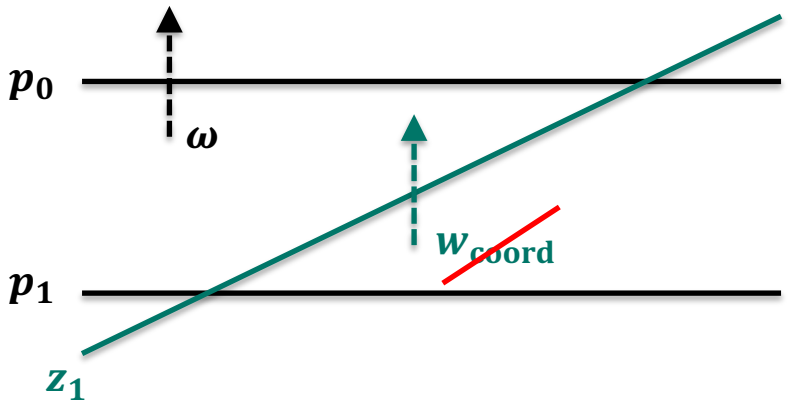

From: Holton and Hakim (2012), An Introduction to Dynamic Meteorology

# Transformation of vertical velocity

$$\left(\frac{\partial z}{\partial t}\right)_z = \left(\frac{\partial z}{\partial t}\right)_p + \frac{\partial z}{\partial p}\left(\frac{\partial p}{\partial t}\right)_z$$

$$w = \left(\frac{\partial z}{\partial t}\right)_p + \frac{\partial z}{\partial p} \cdot \omega$$

$$w = \left(\frac{\partial z}{\partial t}\right)_p - \frac{\omega}{\rho g}$$




$$w = w_{\text{coord}} - \frac{\omega}{\rho g}$$

$$w \approx -\frac{\omega}{\rho g}$$

**Is $w_{\text{coord}}$ really negligible?**

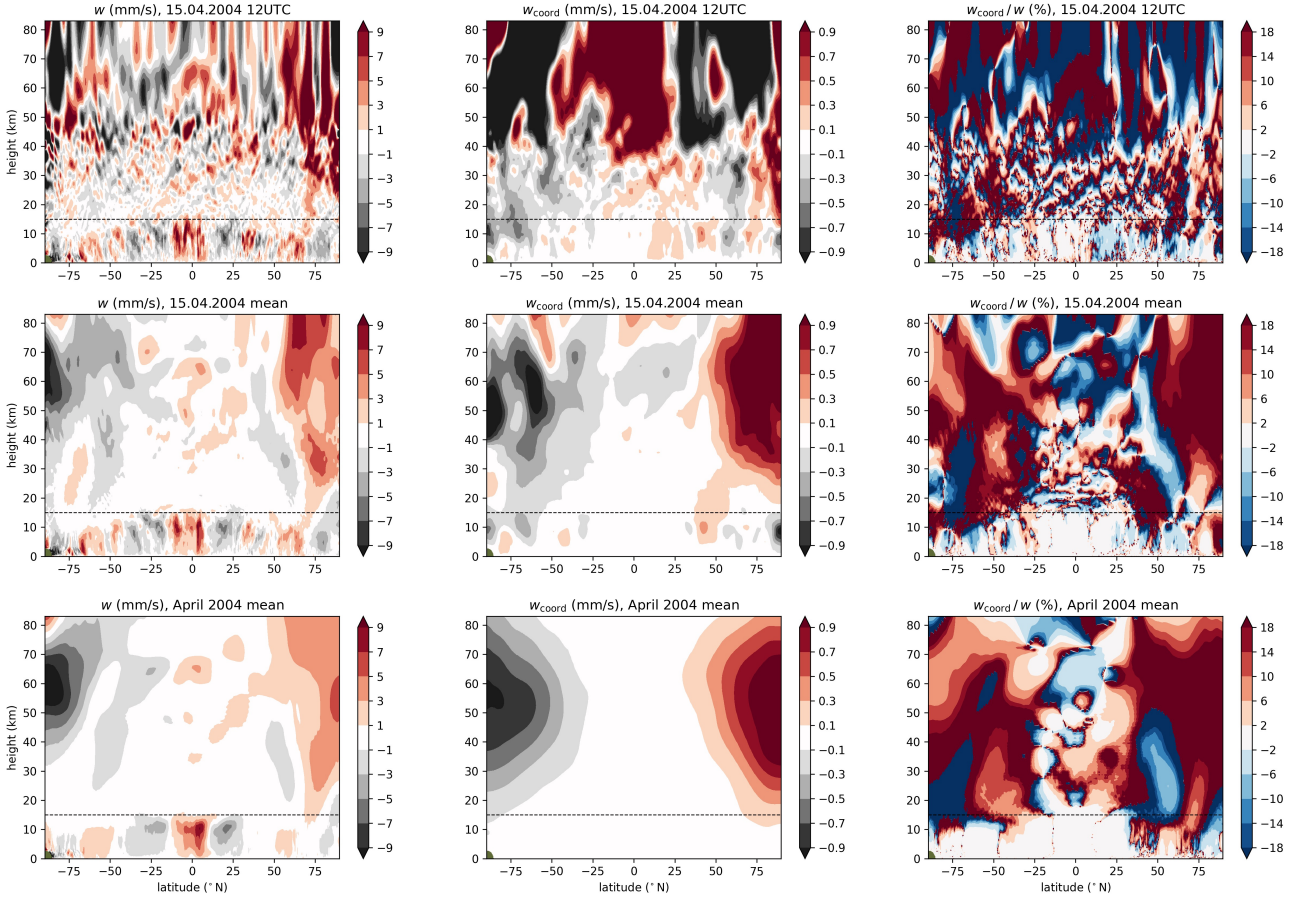
# Importance of $w_{\mathrm{coord}}$ on different timescales

$$w = w_{\mathrm{coord}} - \frac{\omega}{\rho g}$$

Instantaneous:

Daily mean:

Monthly mean:

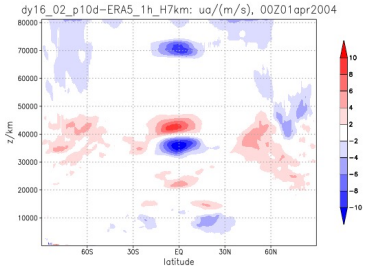


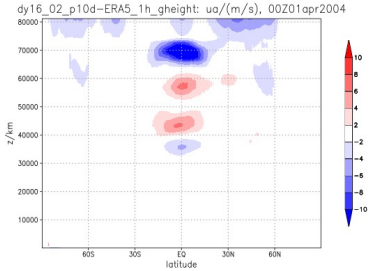
# Conclusion

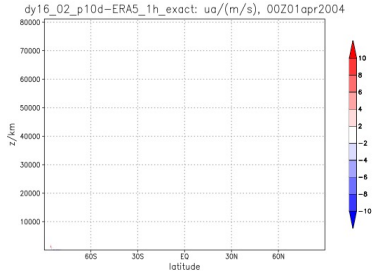
- Doublecheck the unit of provided ERA5 horizontal wind data
- Comparison on z-levels: Coordinate transformation using a fixed scale height is very inaccurate (e.g., warm bias of ~10 K in upper troposphere)
- CDO operator gheight still inaccurate due to missing final interpolation step
- Transformation of vertical velocity: Approximation inaccurate in the stratosphere and mesosphere

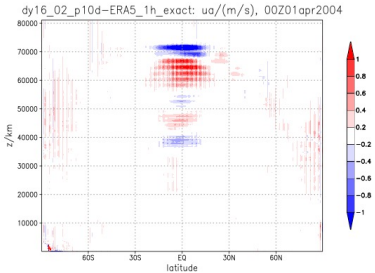
**Careless vertical coordinate transformation can lead to significant artificial biases!**

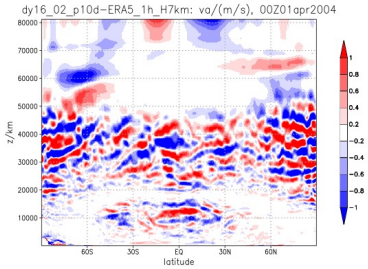
# Finally negligible biases

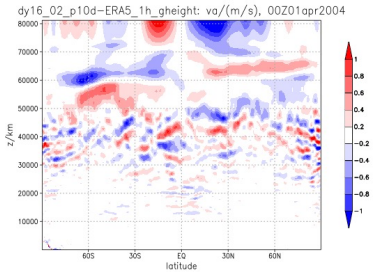


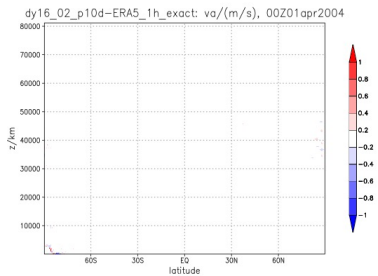



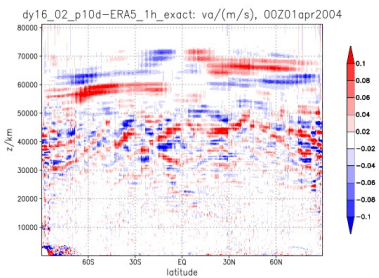



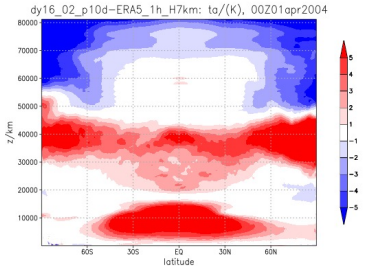



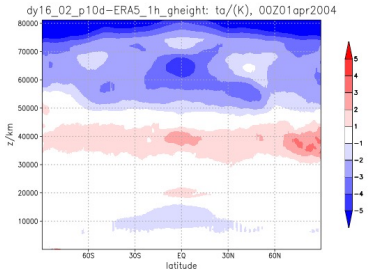



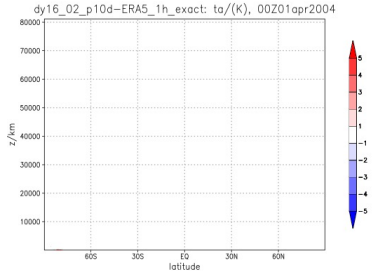



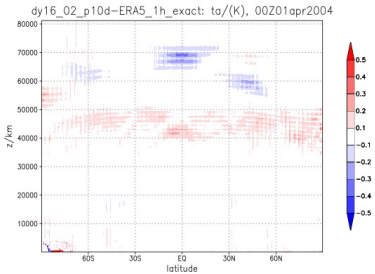




Plots: Marco Giorgetta