

S1 Meteorological data preprocessing

S1.1 Calculation of geopotential heights

Various applications require information about the geopotential heights of the air parcels. MPTRAC calculates geopotential heights from the pressure level data of temperature and specific humidity as well as surface pressure and geopotential height at the surface. Starting from the surface pressure p_0 and geopotential height Z_0 , the geopotential height increments $\Delta Z_{i,i+1}$ between the neighboring pressure levels of a vertical column are calculated and added up onto each other to determine the geopotential height Z_j of the corresponding pressure level p_j .

The geopotential height differences $\Delta Z_{i,i+1}$ between neighboring pressure levels p_i and p_{i+1} are calculated from

$$\Delta Z_{i,i+1} = \frac{R}{g} \frac{T_{v,i} + T_{v,i+1}}{2} \left[\ln \left(\frac{p_i}{p_{i+1}} \right) \right], \quad (\text{S1})$$

where $R = 287.058 \text{ J kg}^{-1} \text{ K}^{-1}$ is the specific gas constant of dry air and $g = 9.80665 \text{ m s}^{-2}$ is the conventional standard gravitational acceleration. The virtual temperature T_v is computed from air temperature T and water vapor volume mixing ratio $x_{\text{H}_2\text{O}}$ according to

$$T_v = T [1 + (1 - \epsilon)x_{\text{H}_2\text{O}}], \quad (\text{S2})$$

where $\epsilon = M_{\text{H}_2\text{O}}/M_a \approx 0.622$ is the ratio of the molar mass of dry air, $M_a = 28.0644 \text{ g mol}^{-1}$, and the molar mass of water vapor, $M_{\text{H}_2\text{O}} = 18.01528 \text{ g mol}^{-1}$. The water vapor volume mixing ratio $x_{\text{H}_2\text{O}}$ relates to specific humidity q via

$$x_{\text{H}_2\text{O}} = \frac{q}{\epsilon}. \quad (\text{S3})$$

As an example, Fig. S1 shows maps of geopotential heights at selected pressure levels in the troposphere and stratosphere as calculated with the meteorological data processing code of MPTRAC using ERA5 data on 1 January 2017, 00:00 UTC for input. We verified the geopotential height calculations of MPTRAC by comparing them to geopotential height data provided directly along with the ERA5 reanalysis. Figure S2 shows that the zonal root mean square (RMS) differences between the geopotential heights from MPTRAC and ERA5 are less than 10–15 m at 5–500 hPa. At pressure levels above 1–2 hPa, interpolation and integration errors start to accumulate and the RMS differences increase to 20–50 m. With relative errors being always less than 0.2%, this comparison indicates reasonable accuracy of the method used to obtain geopotential heights from ERA5 data in the troposphere and stratosphere with MPTRAC.

S1.2 Calculation of potential vorticity

Potential vorticity is defined as

$$P = \alpha (2\boldsymbol{\Omega} + \nabla \times \mathbf{u}) \cdot \nabla \theta, \quad (\text{S4})$$

where α is the specific volume, $\boldsymbol{\Omega}$ the angular velocity vector of the Earth's rotation, \mathbf{u} the three-dimensional vector velocity relative to the rotating Earth, and θ the potential temperature. Potential temperature is defined as

$$\theta = T \left(\frac{p_0}{p} \right)^\kappa \quad (\text{S5})$$

for the reference pressure $p_0 = 1000 \text{ hPa}$ and $\kappa = R/c_p \approx 0.286$ with $R = 287.058 \text{ J kg}^{-1} \text{ K}^{-1}$ being the specific gas constant of dry air and $c_p = 1003.5 \text{ J kg}^{-1} \text{ K}^{-1}$ being the specific heat capacity at a constant pressure. Both, potential vorticity and potential temperature are dynamical tracers. In the absence of friction and heat sources, P is a materially conservative property that remains constant for each particle.

On pressure levels, Ertel's potential vorticity P can be calculated from temperature T and horizontal wind (u, v) as

$$P = g \left[-\frac{\partial \theta}{\partial p} \left(\frac{\partial v}{\partial x} - \frac{1}{\cos \phi} \frac{\partial u}{\partial y} + f \right) + \frac{\partial v}{\partial p} \frac{\partial \theta}{\partial x} - \frac{\partial u}{\partial p} \frac{\partial \theta}{\partial y} \right]. \quad (\text{S6})$$

The Coriolis parameter f at latitude ϕ is calculated from

$$f = 2\Omega \sin \phi, \quad (\text{S7})$$

with Earth's rotation rate $\Omega = 7.2921 \times 10^{-5} \text{ rad s}^{-1}$. The horizontal partial derivatives in Eq. (S6) are evaluated by means of the central differencing scheme. The partial derivatives with respect to pressure are evaluated similarly, but need to consider the irregular vertical grid spacing. The numerical implementation in our code follows a software package to calculate potential vorticity developed by Barlow (2017). Potential vorticity is reported in potential vorticity units, $1 \text{ PVU} = 1.0 \times 10^{-6} \text{ m}^2 \text{ s}^{-1} \text{ K kg}^{-1}$.

As an example, Fig. S3 shows potential vorticity at different pressure levels in the troposphere and stratosphere as calculated with the meteorological data processing code of MPTRAC using ERA5 data on 1 January 2017, 00:00 UTC as input. We verified the potential vorticity calculations of MPTRAC by comparing the results to data provided along with the ERA5 reanalysis. Corresponding zonal mean RMS differences are shown in Fig. S4. The mean relative differences of the potential vorticity data of MPTRAC and ERA5 are mostly below $\pm 5\%$ at mid and high latitudes at the pressure levels from 1 to 500 hPa considered here. At low latitudes, the relative errors may become larger. However, Fig. S4 shows that the absolute RMS differences in the tropics are similar to the RMS differences in the extratropics. Somewhat larger differences of the PV calculations at specific pressure levels and latitudes are related to different factors, including elevated terrain (at 30–40°N and 500 hPa), the subtropical jet (at 30–40°S and 200 hPa), and deep convection (at 0–10°N and 100 hPa).

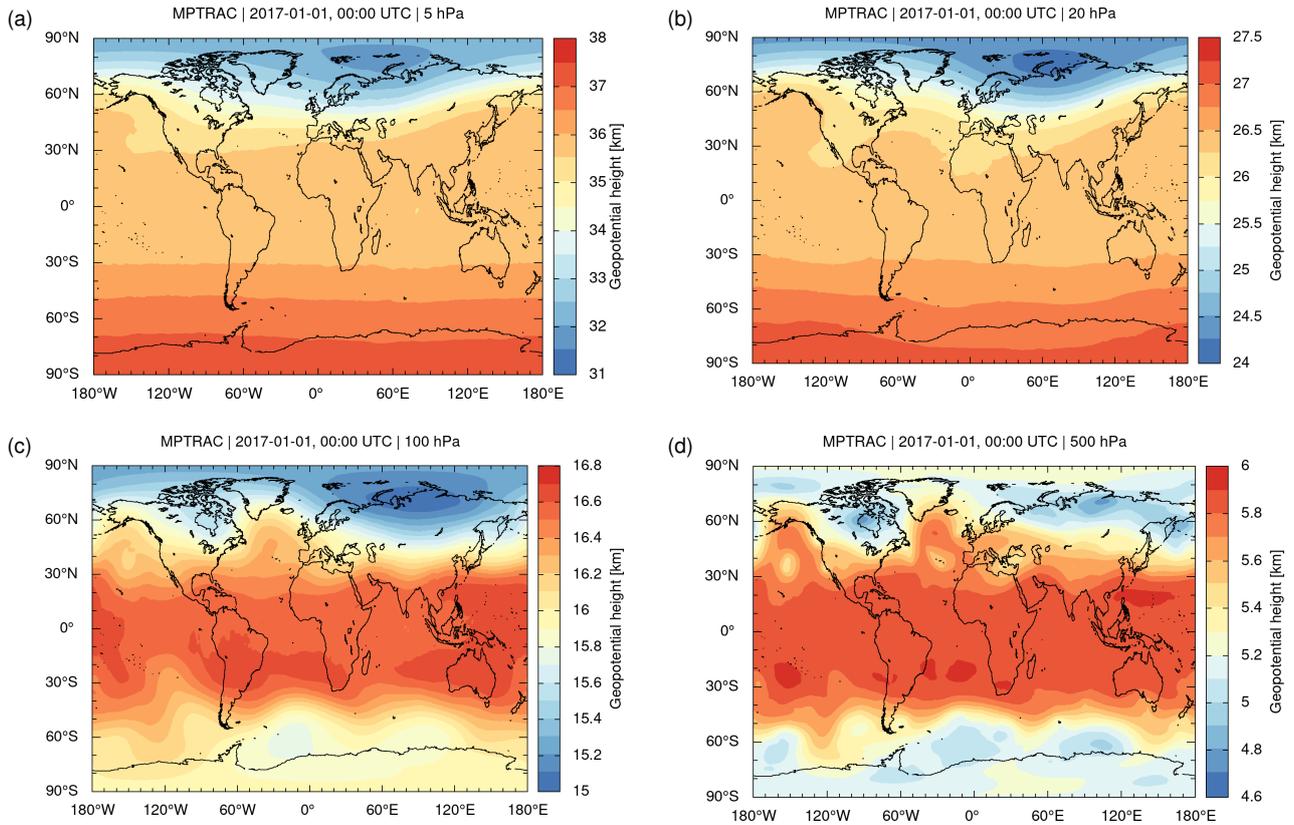


Figure S1. Geopotential heights at 5, 20, 100, and 500 hPa as calculated with the meteorological data processing code of MPTRAC from ERA5 input data for 1 January 2017, 00:00 UTC.

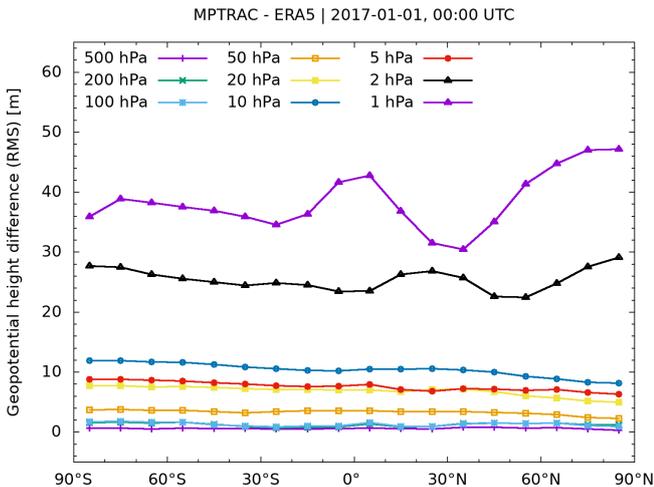


Figure S2. Zonal root mean square (RMS) differences between geopotential heights calculated by the meteorological data processing code of MPTRAC and from ERA5 directly at pressure levels from 500 to 1 hPa on 1 January 2017, 00:00 UTC.

S1.3 Determination of the tropopause

In the MPTRAC model, we implemented a code to estimate the pressure level of the thermal lapse rate tropopause as defined by the World Meteorological Organization (WMO). We also implemented a code to estimate the pressure level of the second thermal tropopause, the cold point, and the dynamical tropopause. In order to improve the vertical resolution of the calculations, we first interpolated the reanalysis temperatures, geopotential heights, and potential vorticity data from the pressure levels of the meteorological input data to a fine vertical grid (100 m grid spacing in log-pressure altitude coordinates). Cubic spline interpolation was applied as this is expected to yield a more realistic representation of real temperature profiles than linear interpolation, in particular near the tropical tropopause region. The vertical range of the tropopause determination was restricted to pressure levels between 47 – 530 hPa or 4.5 – 21.5 km in log-pressure altitude. If the algorithm failed to identify a tropopause in that vertical range, a missing value is reported.

The pressure levels of the first and second thermal tropopause were estimated following the definition of the World Meteorological Organization (WMO, 1957). The first tropopause is defined as the lowest level at which the lapse

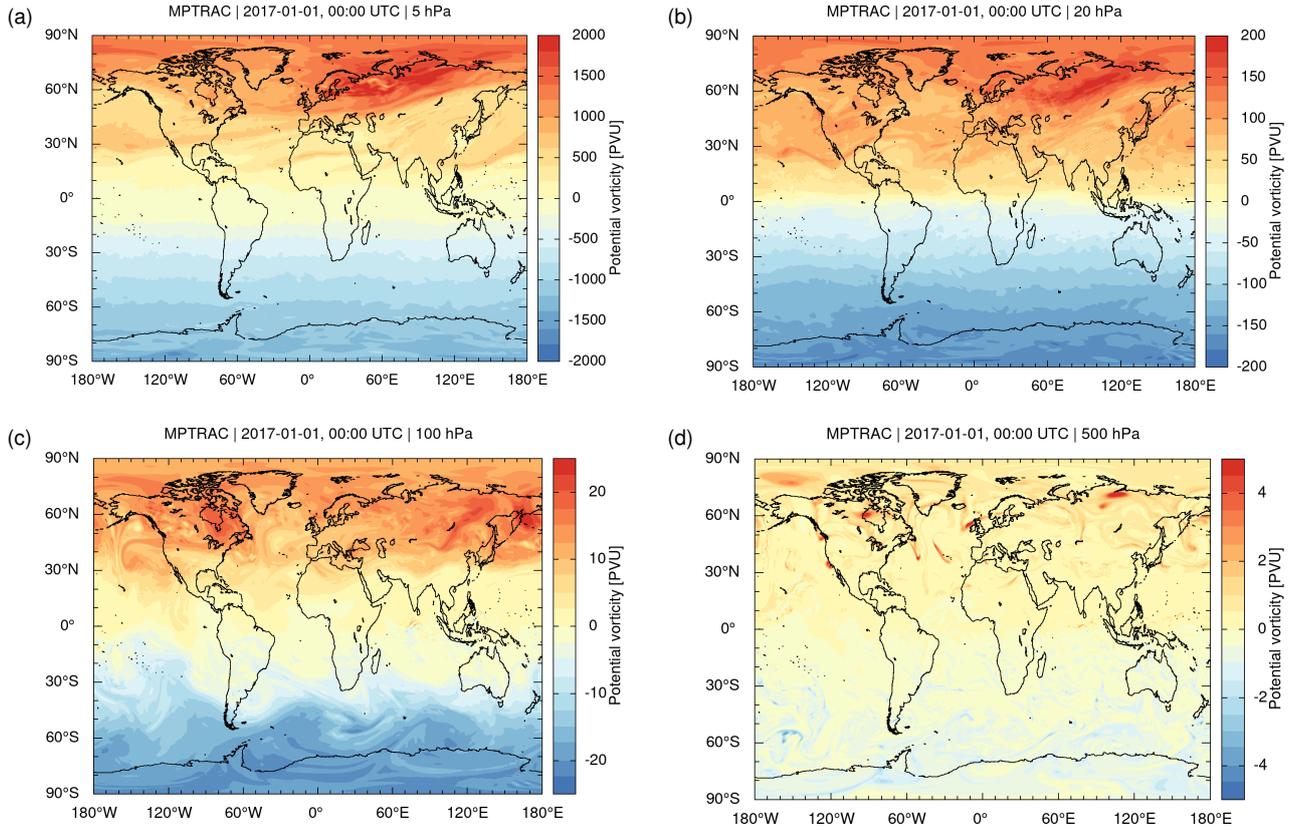


Figure S3. Potential vorticity at 5, 20, 100, and 500 hPa as calculated with the meteorological data processing code of MPTRAC from ERA5 input data for 1 January 2017, 00:00 UTC.

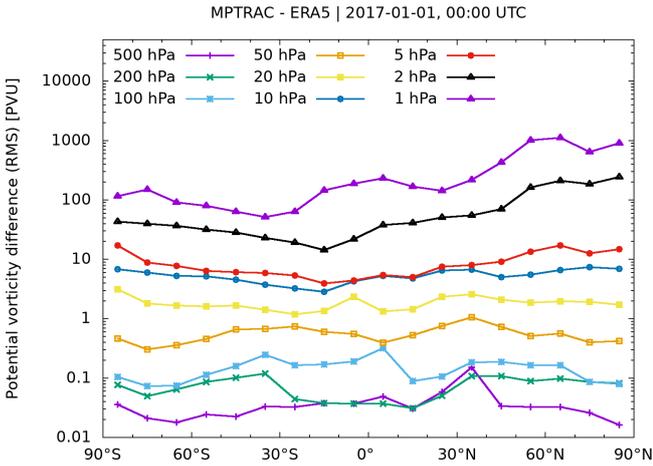


Figure S4. Zonal root mean square (RMS) differences between potential vorticity calculated by the meteorological data processing code of MPTRAC and ERA5 at pressure levels from 500 to 1 hPa on 1 January 2017, 00:00 UTC.

rate falls to $2^{\circ}\text{C km}^{-1}$ or less, provided the average lapse rate between this level and all higher levels within 2 km also does not exceed $2^{\circ}\text{C km}^{-1}$. If above the first tropopause the aver-

age lapse rate between any level and all higher levels within 1 km exceeds $3^{\circ}\text{C km}^{-1}$, then a second tropopause is defined by the same criterion. Considering that the input data are given on pressure levels, the lapse rate Γ is calculated using the hydrostatic equation and the ideal gas law according to

$$\Gamma = -\frac{dT}{dz} = -\frac{dT}{dp} \frac{dp}{dz} = \frac{gp}{RT} \frac{dT}{dp}, \quad (\text{S8})$$

with temperature T , geopotential height z , pressure p , standard gravity g , and specific gas constant of dry air R . Local lapse rates Γ_i based on temperature T_i on the pressure level p_i of the fine vertical grid are therefore calculated according to

$$\Gamma_i = \frac{g}{R} \left(\frac{p_{i+1} + p_i}{T_{i+1} + T_i} \right) \left(\frac{T_{i+1} - T_i}{p_{i+1} - p_i} \right). \quad (\text{S9})$$

The local lapse rates Γ_i on the refined vertical grid are averaged over the given height ranges of the WMO criterion in order to estimate the tropopause pressure.

Another quantity of interest is the cold point. The cold point corresponds to the local minimum of the temperature profile in the upper troposphere and lower stratosphere at a given time and location. We calculated it by applying a

search algorithm to the temperature values T_i at the pressure levels p_i of the refined vertical grid. We calculated the cold point globally, but it is typically meaningful only in the tropics, i. e., within a latitude range of about 30°S to 30°N . At mid and high latitudes the cold point is often not well-defined as the temperature profiles tend to become isothermal in the upper troposphere and lower stratosphere region, in particular in the winter season.

Finally, we also estimated a dynamical tropopause based on potential vorticity at mid and high latitudes and potential temperature in the tropics. The pressure level p_i of the dynamical tropopause is found as the lowest level at which either the absolute value of the potential vorticity first exceeds a threshold of 3.5 PVU or potential temperature first exceeds a threshold of 380 K. Previous studies typically used thresholds between 2 and 4 PVU to define the dynamical tropopause. Here we decided for a value of 3.5 PVU following Hoerling et al. (1991) and Hoinka (1998), showing that statistical deviations between the thermal tropopause and the dynamical tropopause are minimized for that value.

As an example, zonal mean geopotential heights and temperatures of the first and second thermal tropopause, the cold point, and the dynamical tropopause as estimated with MPTRAC from ERA5 input data on 1 January 2017 are presented in Fig. S5. The daily mean zonal mean tropopause data from MPTRAC are in accordance with first thermal tropopause data extracted from Global Positioning System (GPS) radio occultation observations on the same day. A more detailed description of the extraction of tropopause information with MPTRAC as well as the application of the data to identify stratospheric ice clouds is provided by Zou et al. (2020, 2021).

S1.4 Calculation of cloud properties

Some applications of the MPTRAC model require the calculation of specific cloud properties. In particular, the parametrization of wet deposition requires an estimate of the cloud top pressure p_{ct} , the cloud bottom pressure p_{cb} , and the total column cloud water c_l . The cloud pressure levels p_{ct} and p_{cb} are determined from the 3-D fields of CLWC, CRWC, CIWC, and CSWC. We vertically scan the columns of these data to determine p_{ct} and p_{cb} as the uppermost and lowermost mid-level where at least one of the variables changes from zero to non-zero, respectively. At the same time, the total column cloud water c_l is obtained by vertical integration,

$$c_l = \int [\text{CLWC}(z) + \text{CRWC}(z) + \text{CIWC}(z) + \text{CSWC}(z)] \rho(z) dz. \quad (\text{S10})$$

Using the hydrostatic equation,

$$\frac{dz}{dp} = -\frac{1}{g\rho}, \quad (\text{S11})$$

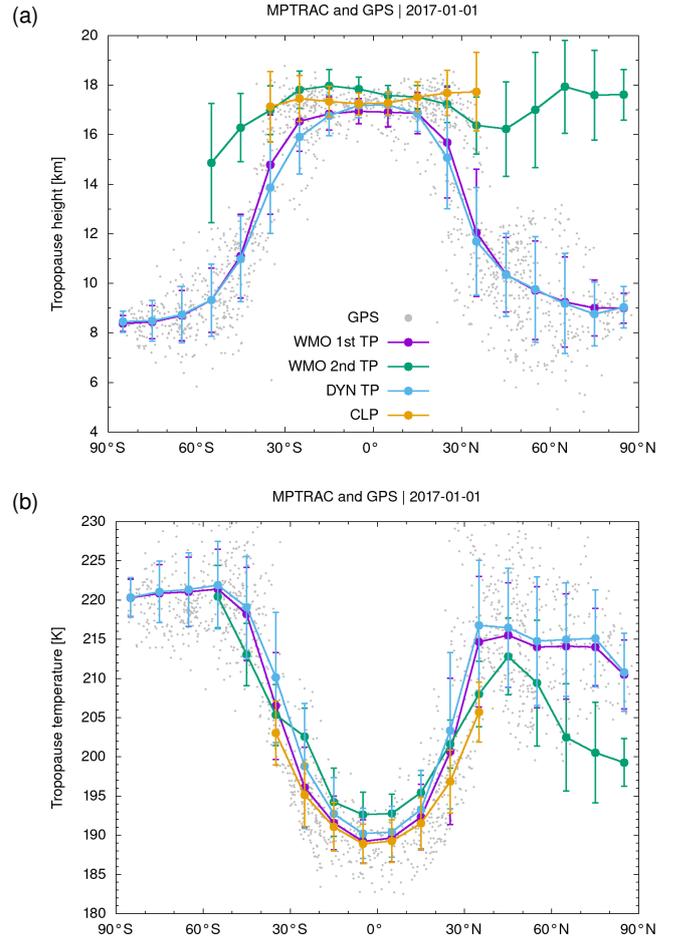


Figure S5. Zonal mean (a) geopotential heights and (b) temperatures of the WMO first lapse rate tropopause (WMO 1st TP), the second lapse rate tropopause (WMO 2nd TP), the dynamical tropopause (DYN TP), and the cold point (CLP) as retrieved with MPTRAC from ERA5 input data for 1 January 2017. Data for the WMO first lapse rate tropopause from individual GPS profiles on the same day are shown for comparison (gray dots).

Eq. (S10) can be numerically evaluated as

$$c_l = \sum_i \frac{p_i - p_{i+1}}{2g} [\text{CLWC}(p_i) + \text{CLWC}(p_{i+1}) + \text{CRWC}(p_i) + \text{CRWC}(p_{i+1}) + \text{CIWC}(p_i) + \text{CIWC}(p_{i+1}) + \text{CSWC}(p_i) + \text{CSWC}(p_{i+1})]. \quad (\text{S12})$$

S1.5 Calculation of convective available potential energy

The convective available potential energy (CAPE) is a measure of the maximum buoyancy of an undiluted air parcel and therefore related to the potential updraft strength of thunderstorms. CAPE is calculated by integrating vertically the local buoyancy of an air parcel from the level of free convection

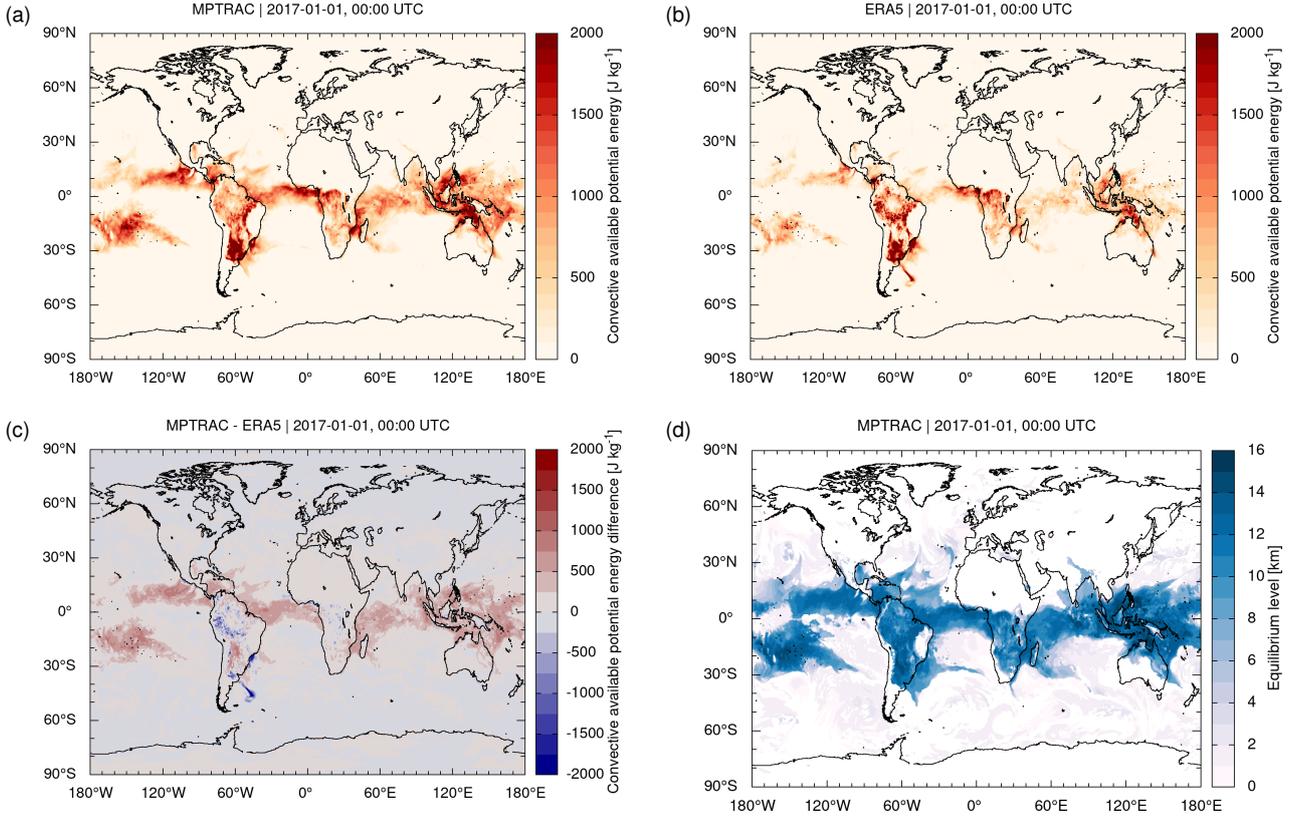


Figure S6. Convective available potential energy (CAPE) on 1 January 2017, 00:00 UTC from (a) the MPTRAC meteorological data processing code, (b) the ERA5 reanalysis, and (c) the differences of MPTRAC minus ERA5. The heights of the equilibrium levels found in the CAPE calculations with MPTRAC are shown in (d).

(LFC) to the equilibrium level (EL),

$$\text{CAPE} = \int_{z_{\text{LFC}}}^{z_{\text{EL}}} g \left(\frac{T_{v,ap} - T_{v,env}}{T_{v,env}} \right) dz, \quad (\text{S13})$$

where $T_{v,ap}$ is the virtual temperature of the air parcel and $T_{v,env}$ is the virtual temperature of the environment. The level of free convection (LFC) is the altitude where the temperature of the environment decreases faster than the moist adiabatic lapse rate of a saturated air parcel at the same level. Above the LFC, the equilibrium level (EL) is the altitude at which a rising parcel of air is at the same temperature as its environment.

In order to establish the vertical range of integration, the lifting condensation level (LCL) needs to be found first. The LCL is formally defined as the height at which the relative humidity (RH) of an air parcel reaches 100 % with respect to liquid water when it is cooled by dry adiabatic lifting. To find the LCL, the potential temperature and water vapor volume mixing ratio of an air parcel near the surface need to be calculated first. A common procedure is to calculate MLCAPE50, which is obtained using the mixed-layer mean $\bar{\theta}$ and $\bar{x}_{\text{H}_2\text{O}}$ for the lowermost 50 hPa above the surface. Starting from surface pressure, the air parcel pressure p is continuously re-

duced and T is recalculated from Eq. (S5) by keeping θ constant until RH approaches 100 %.

The relative humidity over liquid water is calculated from

$$\text{RH}(p, T, x_{\text{H}_2\text{O}}) = \frac{p_w(p, x_{\text{H}_2\text{O}})}{p_{\text{sat}}(T)} \cdot 100\%, \quad (\text{S14})$$

with partial water vapor pressure,

$$p_w(p, x_{\text{H}_2\text{O}}) = \frac{p x_{\text{H}_2\text{O}}}{1 + (1 - \epsilon)x_{\text{H}_2\text{O}}}, \quad (\text{S15})$$

and the saturation pressure over water (WMO, 2018),

$$p_{\text{sat}}(T) = 6.112 \text{ hPa} \exp\left(\frac{17.62(T - T_0)}{243.12 \text{ K} + T - T_0}\right), \quad (\text{S16})$$

with $T_0 = 273.15 \text{ K}$.

Starting from the air parcel pressure at the LCL, we sequentially reduce the pressure in steps Δp of about 100 m in log-pressure height. The corresponding layer width Δz in terms of geopotential heights is calculated from Eq. (S1). The temperature change $\Delta T = -\Gamma_m \Delta z$ is calculated by means of the moist adiabatic lapse,

$$\Gamma_m(T, r) = g \frac{RT^2 + L_V r T}{c_{pd} RT^2 + L_V^2 r \epsilon}, \quad (\text{S17})$$

with specific gas constant of dry air R , specific heat capacity of dry air at constant pressure c_{pd} , latent heat of vaporization of water $L_v = 2.501 \times 10^6 \text{ J kg}^{-1}$, and the mixing ratio of the mass of water vapor to the mass of dry air,

$$r = \frac{\epsilon x_{\text{H}_2\text{O}}}{1 - \epsilon x_{\text{H}_2\text{O}}}. \quad (\text{S18})$$

The water vapor volume mixing ratio $x_{\text{H}_2\text{O}}$ is recalculated to keep RH at 100%. With this procedure, the virtual temperature of the air parcel $T_{v,ap}$ can be calculated along the pressure steps and compared to the virtual temperature of the environment $T_{v,env}$. If $T_{v,ap} > T_{v,env}$, the LFC is found. The EL is found as the first level above the LFC, where $T_{v,ap} < T_{v,env}$. CAPE is calculated by vertically integrating the relative differences of the virtual temperature $T_{v,ap}$ and $T_{v,env}$ from the LFC to the EL according to Eq. (S13).

Figure S6 compares the results of the CAPE calculations of MPTRAC with CAPE data provided along with the ERA5 reanalysis on 1 January 2017, 00:00 UTC. Qualitatively, the data sets compare well and show similar patterns on the global scale. However, it is also found that CAPE maxima from MPTRAC may exceed the ERA5 data by up to 500 to 1000 J kg^{-1} in the tropics (Fig. S6c). Figure S7a compares tail distributions of the CAPE data from MPTRAC and ERA5. This more detailed analysis shows that CAPE values larger than about 20 J kg^{-1} are up to 5 percentage points more frequent in MPTRAC compared to ERA5.

Note that CAPE calculations can be quite sensitive to parameter choices, such as the depth of the surface layer being used to calculate $\bar{\theta}$ and $\bar{x}_{\text{H}_2\text{O}}$. Tests with different depths of the surface layer (50 or 100 hPa) with the MPTRAC code showed already quite significant variations in CAPE. For this reason, we attribute the differences between the MPTRAC and ERA5 CAPE data seen here to different methodologies and parameter choices applied for the CAPE calculations.

Both, CAPE and the height of the equilibrium levels are required as input data for the convection parametrization implemented in MPTRAC. For reference, Fig. S6d shows the log-pressure height of the equilibrium levels found in the CAPE calculations with MPTRAC. Corresponding data from ERA5 were not available for comparison. Figure S7b shows the log-pressure height of the equilibrium levels versus CAPE. A running mean through the data indicates a tight correlation between CAPE and the height of the equilibrium level. CAPE values less than 20 J kg^{-1} typically correspond to equilibrium levels below 2 km whereas CAPE values larger than 1000 J kg^{-1} may correspond to equilibrium levels as high as 16 km in tropical convection.

S1.6 Determination of the planetary boundary layer

Most recently, a code was added to the MPTRAC model to estimate the height of the planetary boundary layer (PBL) using the bulk Richardson number method (Vogelezang and Holtslag, 1996). The bulk Richardson number at height z

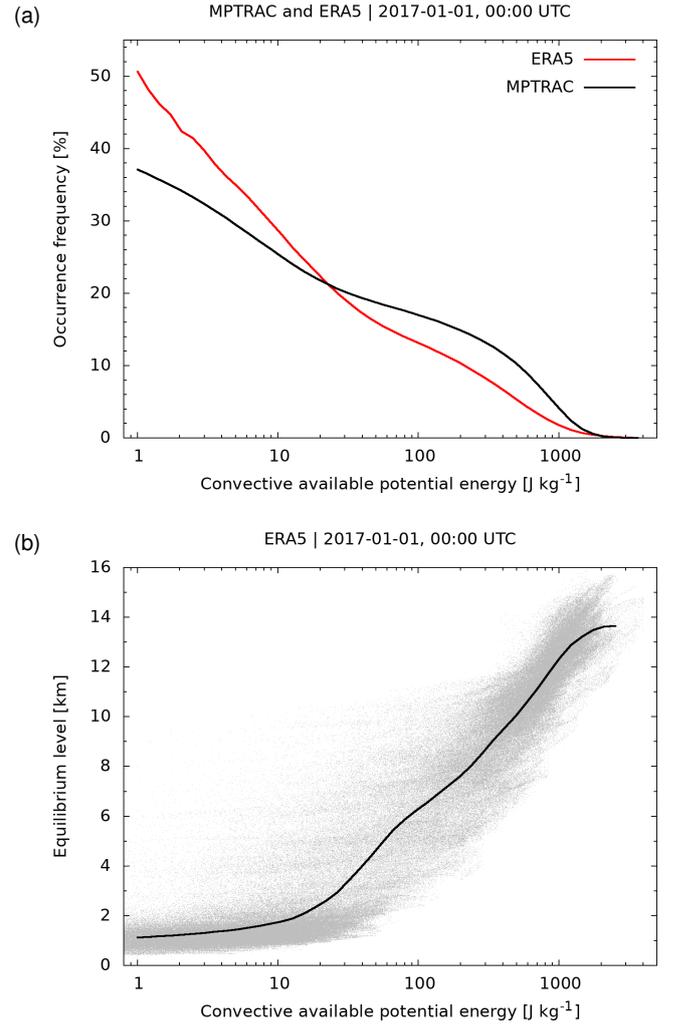


Figure S7. Convective available potential energy (CAPE) estimated with MPTRAC using ERA5 data. (a) The occurrence frequencies of CAPE as estimated from the MPTRAC meteorological data processing code and the ERA5 reanalysis and (b) equilibrium levels versus CAPE for individual grid boxes of the reanalysis data (gray dots) and a running mean through the data (black curve). All data refer to 1 January 2017, 00:00 UTC.

above the surface is calculated from

$$\text{Ri}(z) = \frac{(g/\theta_{vs})(\theta_{vz} - \theta_{vs})(z - z_s)}{(u_z - u_s)^2 + (v_z - v_s)^2}, \quad (\text{S19})$$

where θ_v is the virtual potential temperature and the index s denotes the surface. The height of the PBL is found as the vertical level where Ri exceeds a critical value of 0.25, considering that layers with Richardson numbers less than this critical value are dynamically unstable and likely to become or remain turbulent. The bulk Richardson number method is generally considered suitable for both stable and convective boundary layers, identifies a non-negative height in all cases, and is not strongly dependent on vertical resolution of the input data (Seidel et al., 2012).

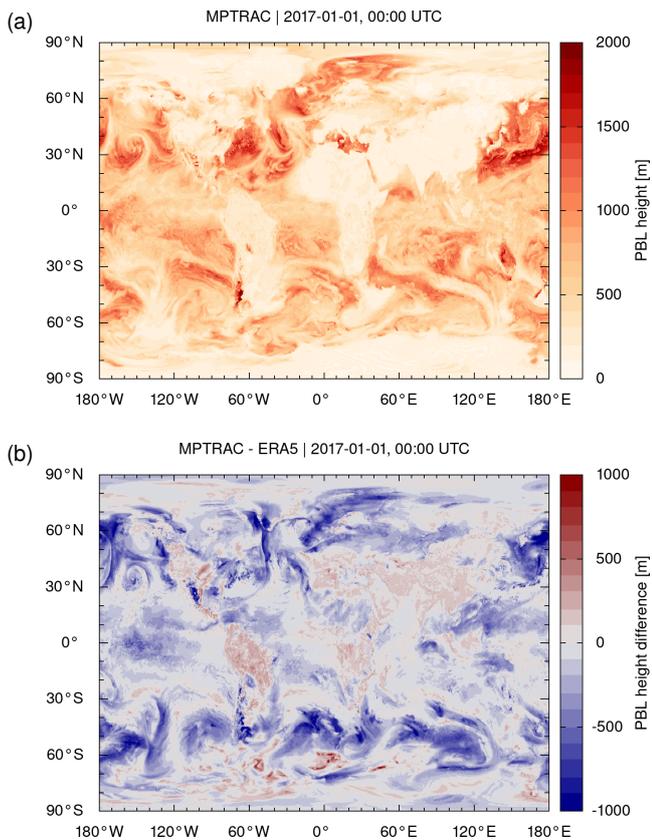


Figure S8. Contour surfaces show (a) the planetary boundary layer (PBL) height from the MPTRAC meteorological data processing code and (b) the differences with respect to the ERA5 reanalysis for 1 January 2017, 00:00 UTC.

As an example, Fig. S8 shows PBL heights on 1 January 2017, 00:00 UTC as estimated with the method implemented into the MPTRAC model as well as the differences with respect to data provided directly along with ERA5. Overall, this comparison shows that the large-scale patterns of the PBL heights are similar, but also that some systematic differences are present. In particular, the PBL heights from MPTRAC over mid-latitude ocean regions can be 500 to 1000 m lower than the ERA5 data. Consistent with other studies (Zhang et al., 2014), tests showed that the calculation of PBL heights is particularly sensitive to the choice of the critical value for Ri and the surface data. We also found it necessary, to impose an empirical lower limit of $(5 \text{ m s}^{-1})^2$ for the denominator of Eq. (S19), to account for surface friction velocity and to achieve numerical stability of the results.

References

- Barlow, M.: mathewbarlow/potential-vorticity: minor fix, <https://doi.org/10.5281/zenodo.1069032>, 2017.
- Hoerling, M. P., Schaack, T. K., and Lenzen, A. J.: Global Objective Tropopause Analysis, *Mon. Wea.*

Rev., 119, 1816–1831, [https://doi.org/10.1175/1520-0493\(1991\)119<1816:GOTA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1991)119<1816:GOTA>2.0.CO;2), 1991.

Hoinka, K. P.: Statistics of the Global Tropopause Pressure, *Mon. Wea. Rev.*, 126, 3303–3325, [https://doi.org/10.1175/1520-0493\(1998\)126<3303:SOTGTP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<3303:SOTGTP>2.0.CO;2), 1998.

Seidel, D. J., Zhang, Y., Beljaars, A., Golaz, J.-C., Jacobson, A. R., and Medeiros, B.: Climatology of the planetary boundary layer over the continental United States and Europe, *J. Geophys. Res.*, 117, <https://doi.org/10.1029/2012JD018143>, 2012.

Vogelezang, D. and Holtslag, A. A.: Evaluation and model impacts of alternative boundary-layer height formulations, *Bound.-Lay. Meteorol.*, 81, 245–269, <https://doi.org/10.1007/BF02430331>, 1996.

WMO: *Meteorology A Three-Dimensional Science: Second Session of the Commission for Aerology*, WMO Bull., iv, 134–138, 1957.

WMO: *Guide to Instruments and Methods of Observation*, WMO-No. 8, 2018.

Zhang, Y., Gao, Z., Li, D., Li, Y., Zhang, N., Zhao, X., and Chen, J.: On the computation of planetary boundary-layer height using the bulk Richardson number method, *Geosci. Model Dev.*, 7, 2599–2611, <https://doi.org/10.5194/gmd-7-2599-2014>, 2014.

Zou, L., Griessbach, S., Hoffmann, L., Gong, B., and Wang, L.: Revisiting global satellite observations of stratospheric cirrus clouds, *Atmos. Chem. Phys.*, 20, 9939–9959, <https://doi.org/10.5194/acp-20-9939-2020>, 2020.

Zou, L., Hoffmann, L., Griessbach, S., Spang, R., and Wang, L.: Empirical evidence for deep convection being a major source of stratospheric ice clouds over North America, *Atmos. Chem. Phys.*, 21, 10 457–10 475, <https://doi.org/10.5194/acp-21-10457-2021>, 2021.