The Lagrangian analysis tool LAGRANTO – version 2.0

M. Sprenger and H. Wernli

Institute for Atmospheric and Climate Science, ETH Zurich, Switzerland

Received: 3 February 2015 – Accepted: 6 February 2015 – Published: 20 February 2015

Correspondence to: M. Sprenger (michael.sprenger@env.ethz.ch)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Lagrangian trajectories are widely used in the atmospheric sciences, for instance to identify flow structures in extratropical cyclones (e.g., warm conveyor belts) and long-range transport pathways of moisture and trace substances. Here a new version of the Lagrangian analysis tool LAGRANTO (Wernli and Davies, 1997) is introduced, which offers considerably enhanced functionalities: (i) trajectory starting positions can be described easily based on different geometrical and/or meteorological conditions; e.g., equidistantly spaced within a prescribed region and on a stack of pressure (or isentropic) levels; (ii) a versatile selection of trajectories is offered based on single or combined criteria; these criteria are passed to LAGRANTO with a simple command language (e.g., “GT:PV:2” readily translates into a selection of all trajectories with potential vorticity (PV) greater than 2 PVU); and (iii) full versions are available for global ECMWF and regional COSMO data; core functionality is also provided for the regional WRF and UM models, and for the global 20th Century Reanalysis data set.

The intuitive application of LAGRANTO is first presented for the identification of a warm conveyor belt in the North Atlantic. A further case study then shows how LAGRANTO is used to quasi-operationally diagnose stratosphere–troposphere exchange events over Europe. Whereas these example rely on the ECMWF version, the COSMO version and input fields with 7 km horizontal resolution are needed to adequately resolve the rather complex flow structure associated with orographic blocking due to the Alps. Finally, an example of backward trajectories presents the tool’s application in source-receptor analysis studies. The new distribution of LAGRANTO is publicly available and includes simple tools, e.g., to visualize and merge trajectories. Furthermore, a detailed user guide exists, which describes all LAGRANTO capabilities.
1 Introduction

The calculation of air parcel trajectories has become a widely used approach in different areas of atmospheric research. Several trajectory tools have been developed and are currently available to the community for computing backward and forward trajectories using gridded wind fields, typically from (re-)analysis fields or numerical model simulations. Examples of such trajectory tools are FLEXTRA (Stohl et al., 1995), the NASA Goddard Trajectory Model (Schoeberl and Newman, 1995), HYSPLIT, the Hybrid Single Particle Lagrangian Integrated Trajectory Model (Draxler and Hess, 1998), the UGAMP offline trajectory model (Methven, 1997; Methven et al., 2001), and LAGRANTO, the Lagrangian Analysis Tool (Wernli and Davies, 1997). They all solve iteratively the same trajectory equation

\[
\frac{Dx}{Dt} = u(x),
\]

where \( x = (\lambda, \phi, p) \) is the position vector in geographical pressure coordinates and \( u = (u, v, \omega) \) the three-dimensional wind vector. The tools differ mainly in terms of spatial interpolation (e.g., Seibert, 1993; Stohl et al., 2001), the number of iterations used for one time step, the handling of the lower boundary, and, an important aspect for the user, in terms of their functionality and the details of the output. For instance, HYSPLIT can be operated directly on the web, which makes it extremely straightforward to obtain trajectory information. LAGRANTO, in contrast, has been designed to flow-dependently select meaningful trajectories based on objective criteria.

In this paper, a new version of this well-established trajectory tool is presented, which significantly generalizes and extends the possibilities in determining starting points, selecting trajectories using combinations of criteria, and tracing a comprehensive set of information along the selected trajectories. This new version 2.0 of LAGRANTO enables novel options for research with air parcel trajectories, and aims at achieving a level of flexibility for the user that goes beyond the options of previously available trajectory tools. Before we describe in detail the functionality of this new version of LA-
GRANTO and present potential application, we first summarize some highlights of historical research using trajectories, and then a selection of applications of LAGRANTO 1.0 during the last almost 20 years. This summary should put the calculation of trajectories in a historical context and illustrate the very broad range of research questions, for which a Lagrangian analysis using simple air parcel trajectories can be useful.

Shaw and Lempfert (1906) were probably the first who used surface trajectories, calculated based upon surface weather charts, to analyze and visualize the airflow in extratropical cyclones. Decades later, Reed (1955) and Danielsen (1961) pioneered the calculation of isentropic trajectories from radiosonde data for studying the descent of stratospheric intrusions associated with upper-level fronts. Kleinschmidt (1950) estimated the pathway of trajectories ascending in the warm sector of extratropical cyclones and the diabatic potential vorticity modification along this flow. These are three very early examples for the usefulness of air parcel trajectories in studies on the structure and dynamics of extratropical weather systems. Since the late 1970s – using now wind fields from reanalyses or model simulations – trajectories have been frequently used for investigating different types of atmospheric flow phenomena including extratropical cyclones (e.g., Whitaker et al., 1988; Kuo et al., 1992), orographic flows (e.g., Buzzi and Tibaldi, 1978; Steinacker, 1984), and stratosphere–troposphere exchange (e.g., Buzzi et al., 1984; Vaughan et al., 1994; Newman and Schoeberl, 1995). When developing LAGRANTO in the early 1990s, it was the major objective to develop a versatile tool for the Lagrangian analysis of weather systems and their dynamics. But clearly, at that time trajectories have been also successfully used in other areas of the atmospheric sciences, for instance for transport studies of pollution (Tyson et al., 1996) and Saharan dust (Swap et al., 1992) – see also the review by Stohl (1998).

LAGRANTO has been first used to objectively identify coherent airstreams in case studies of extratropical cyclones, in particular warm conveyor belts (WCBs, see also Sect. 2) and dry intrusions (Wernli and Davies, 1997; Wernli, 1997), and later also sting jets (Gray et al., 2011) and cold conveyor belts (Schemm and Wernli, 2014). Recently, WCBs have been calculated for the ECMWF ensemble prediction system for
producing WCB probability forecasts (Schäfler et al., 2014; Rautenhaus et al., 2015), and for the entire ERA-Interim data set leading to a comprehensive global climatology of WCBs (Madonna et al., 2014; Pfahl et al., 2014). Also, LAGRANTO trajectories helped analyzing how the moist outflow of tropical cyclones modifies the downstream midlatitude jet (Grams et al., 2013).

Other important research areas for the application of LAGRANTO were (i) the identification of stratosphere–troposphere exchange events (e.g., Wernli and Davies, 1997; Sprenger and Wernli, 2003; Skerlak et al., 2014), (ii) the quantitative analysis of moisture sources and transport (e.g., Sodemann et al., 2008; Knippertz and Wernli, 2010), (iii) the interpretation of trace gas and isotope measurements from in-situ and remote sensing instruments (e.g., Prévôt et al., 1997; Calisesti et al., 2001; Koch et al., 2002; Pfahl et al., 2008), (iv) cloud microphysics (e.g., Fueglistaler et al., 2003; Joos and Wernli, 2012; Brabec et al., 2012), and last but not least orographic flows (e.g., Kljun et al., 2001; Miltenberger et al., 2013; Würsch and Sprenger, 2015). This short list clearly illustrates the very broad range of research themes, for which a versatile Lagrangian analysis tool can be highly valuable. These possibilities are further increased due to the novel features of LAGRANTO version 2.0, as illustrated in the following section and the examples given in Sect. 4.

2 An introductory example: identification of a warm conveyor belt

Warm conveyor belts (WCBs) are coherent flow structures, which transport moist near-surface air from the warm sector of an extratropical cyclone upward to near-tropopause levels (e.g., Browning, 1990). Typically this transport occurs during a 48 h ascent period and is accompanied by a poleward movement of the strongly precipitating air parcels. Here we take this atmospheric flow feature as a first example to illustrate the application of LAGRANTO. Our aim is to detect a WCB in the North Atlantic associated with an extratropical cyclone in January 2009, investigated in detail by Joos and Wernli (2012).
We start with the formal definition of the problem: first, starting positions are initialized within a suitable domain over the North Atlantic and Europe at 12:00 UTC 29 January 2009. For this case we choose the domain from 60°W to 20°E and 30 to 80°N, which encloses the overall cyclone evolution. We choose starting positions in this domain that are horizontally equidistant with 80 km horizontal spacing and extend vertically from 1030 to 790 hPa with 30 hPa vertical spacing. Second, trajectories are calculated from all starting positions 48 h forward in time and then the trajectories are selected that ascend at least 600 hPa, i.e., whose pressure decreases by at least 600 hPa. And finally, for this introductory example, potential temperature is traced along the selected WCB trajectories. Note the somewhat brute-force approach taken: first, many trajectories are calculated (here about 26 700), and then the few ones (here 76) fulfilling the imposed WCB characteristics are selected and further studied.

To accomplish the proposed aim, the wind and potential temperature fields must be available on netCDF files for the time period from 12:00 UTC 29 January to 12:00 UTC 31 January 2009 (see Sect. 3 for further details). Then the starting positions can be created with a command (all details can be found in the supplementary material, see example 1)

```
startf 20090129_12 startf.2 ...
...’region.eqd(1,80) @ ...
... profile(1030,790,8) @ hPa’
```

This command creates a file `startf.2` with starting coordinates (longitude, latitude, pressure) covering the whole North Atlantic. The horizontal domain is specified in a text file `regionf` with a line entry “1 -60 20 30 80”, where label 1 identifies the region and the other values define the rectangular longitude/latitude box as specified above. The suffix `.eqd` tells LAGRANTO that the starting positions shall be equidistantly distributed within the box with a spacing of 80 km. The horizontal distribution of the starting positions is shown in Fig. 1. The vertical level of the starting points is specified by `profile(1030 790,8) @ hPa`, which intuitively translates to 8 equidistant levels between 1030 and 790 hPa. Many other options exist to create starting points, as discussed later.
In the next step, the forward trajectories are calculated with the LAGRANTO command `caltra`:

```
  caltra 20090129_12 20090131_12 ...
  ... startf.2 traj.4 -j
```

The starting positions are taken from `startf.2`, the file created above, and the output of the trajectory calculation, i.e., the 6-hourly trajectory positions, are written to the file `traj.4`. The output interval is set by default to the interval of the input files (6 h for ERA-interim). Note that different file formats can be specified according to the suffix of the filenames: suffix 2 (or equivalently suffix “ti”) refers to a text file, whereas suffix 4 (equivalently “du”) to a Fortran binary dump. The latter is very compact (reduced disk space needed) and can be quickly read by other LAGRANTO commands. The order of the two dates in the call, `20090129_12 followed by 20090131_12` defines the direction of the trajectory calculation (here forward, since the second date is after the first one) and the duration (here 48 h). Note that the format to specify dates is `YYYYMMDD_HH(mm)`, where YYYY, MM, DD, HH, and mm refer to the year, month, day, hour, and minutes, respectively. If no value is given for the minutes mm, then 00 is used as the default. LAGRANTO is able to handle a high frequency of input files; in principle, they could be available every minute. A special option is the so-called “jump flag” – `-j` to handle trajectories crossing the lower boundary due to numerical deficiencies (as discussed further in Sect. 3.3).

Note that so far trajectories are calculated for the whole North Atlantic, and it is only now that an objective selection criterion is used to extract WCB trajectories from this large set. The LAGRANTO command to do this is called `select`, and an appropriate call for the present purpose, corresponding to the requirement mentioned above, is

```
  select traj.4 wcb.1 'GT:p(DIFF):600:0,48'
```

The criterion selects all trajectories for which the pressure (p) difference (DIFF) between time 0 and time 48 h is greater (GT) than 600 hPa. The selected trajectories, i.e., their 6-hourly positions, are written to a new trajectory file `wcb.1`, which is now in
ASCII format sorted by trajectory, i.e. the whole first trajectory is listed in the ASCII file, followed by an empty line, after which the second trajectory is listed and so on. Note, in the previous call to *startf* the format 2 was used. There, this was advantageous because only one single time was saved (the starting position) and the entries are not separated by an empty line.

The final step is accomplished with the command *trace*, which allows to trace meteorological fields along trajectories. Here, we already have selected the WCB trajectories and would like to know the evolution of potential temperature (*TH*) along each trajectory. The appropriate LAGRANTO call is

```
trace wcb.1 wcb.1 -f TH 1
```

which adds a column with potential temperature values to the existing trajectory file *wcb.1*. The last parameter 1 defines a scaling factor, which applies to *TH* before it is written to the output file. Note, if several fields shall be traced along the trajectories, they can be listed in an external file (see Sect. 3.3). As a result, the file *wcb.1* lists, for all the selected WCB trajectories, their position (*λ*, *φ*, *p*) and potential temperature, every 6 h from time 0 to 48. The first few lines from the resulting trajectory file are shown here as an example:

```
Reference date 20090129_1200 / ...
   ... Time range  2880 min

<table>
<thead>
<tr>
<th>time</th>
<th>lon</th>
<th>lat</th>
<th>p</th>
<th>TH</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>-22.72</td>
<td>33.96</td>
<td>859</td>
<td>294.151</td>
</tr>
<tr>
<td>6.00</td>
<td>-18.04</td>
<td>35.58</td>
<td>734</td>
<td>298.055</td>
</tr>
<tr>
<td>12.00</td>
<td>-14.25</td>
<td>39.49</td>
<td>555</td>
<td>306.311</td>
</tr>
<tr>
<td>18.00</td>
<td>-11.23</td>
<td>45.85</td>
<td>406</td>
<td>312.269</td>
</tr>
<tr>
<td>24.00</td>
<td>-9.46</td>
<td>53.83</td>
<td>370</td>
<td>311.663</td>
</tr>
</tbody>
</table>
```
In the header the reference time, which corresponds to time 0.00, and the time range are given. Then, headings are provided for the five columns: time (in hour.min); longitude and latitude (in deg); pressure (in hPa) and potential temperature (in K). For each 48 h trajectory information is provided on 9 lines, followed by an empty line.

The selected WCB trajectories are shown in Fig. 1, colored with potential temperature. A well established WCBs can be discerned in the eastern North Atlantic, starting near surface to the south-west of the Iberian peninsula and ascending rapidly while proceeding nortward before finally being deflected to the east. A more detailed analysis of this WCB example follows in Sect. 4.1.

### 3 The structure of LAGRANTO

In this section additional, mainly technical, details about LAGRANTO are discussed. An overview of a typical LAGRANTO call is shown in Fig. 2 as a flowchart. First, the input files must be prepared in netCDF format, then the starting positions are defined from which the forward or backward trajectories are calculated. Depending on the application, a subsample of trajectories is selected based on their pathway, and/or additional meteorological and geometrical parameters are traced along the trajectories. Further selections of trajectories can then rely on these additional fields. Finally, the trajectories can be visualized as either individual trajectories or in the form of a trajectory density field. Note that modularity is a characteristic feature of LAGRANTO, which allows at any time to select another trajectory subsample, trace additional fields, or even base new trajectory calculations on the output of previous ones. In the following, the key aspects of LAGRANTO are discussed in greater detail.

<table>
<thead>
<tr>
<th>Time</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Pressure</th>
<th>Potential Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>30.00</td>
<td>-7.87</td>
<td>62.48</td>
<td>313</td>
<td>312.610</td>
</tr>
<tr>
<td>36.00</td>
<td>2.76</td>
<td>71.83</td>
<td>299</td>
<td>311.725</td>
</tr>
<tr>
<td>42.00</td>
<td>32.38</td>
<td>71.24</td>
<td>269</td>
<td>311.074</td>
</tr>
<tr>
<td>48.00</td>
<td>32.20</td>
<td>62.84</td>
<td>259</td>
<td>310.461</td>
</tr>
</tbody>
</table>
3.1 Data structure

LAGRANTO calculations are based on netCDF files that contain three-dimensional wind fields on a regular longitude/latitude grid, either globally or for a limited domain. In the vertical, different level types are supported: (i) ECMWF hybrid $\sigma - p$ model levels, (ii) constant pressure levels, and (iii) isentropic levels for isentropic trajectories. All fields must be horizontally and vertically de-staggered, in particular the vertical wind must be provided at the same grid points as the horizontal wind components. Since ECMWF disseminates its products in GRIB format (formerly GRIB1, now GRIB2), a conversion into netCDF is necessary. To this aim, a conversion tool is provided together with LAGRANTO. Alternatively, the conversion can be done using the CDO tools (https://code.zmaw.de/projects/cdo), which requires a subsequent adjustment of the CDO-netCDF files to the needs of LAGRANTO. Specific information for using LAGRANTO with data not from ECMWF will be given in Sects. 4.3 and 5.1.

LAGRANTO expects a specific file and file name convention: the zonal wind $U$ (in m s$^{-1}$), the meridional wind $V$ (in m s$^{-1}$) and the vertical wind $\Omega$ (in hPa s$^{-1}$) must be available on a primary file named $P\{\text{date}\}$, where the date format is YYYYMMDD_HH, as mentioned above. Note that LAGRANTO does not recognize whether the wind fields originate from ECMWF analyses or forecasts. In addition to the three wind components, surface pressure $P_S$ (in hPa) must be provided on the $P$ files for level type (i) for calculating the full 3-D pressure field on the model levels and identifying trajectories intersecting the topography. For the calculation of secondary fields, which accordingly are named $S\{\text{date}\}$ an additional tool $p2s$ is available. It allows, for instance, to calculate potential vorticity or potential temperature on the ECMWF hybrid grid. LAGRANTO will automatically check whether $S$ files are available in addition to the $P$ files, and e.g. if starting positions are specified on isentropic levels assumes that potential temperature is found on the $S$ file.

Trajectory position information can be written in different formats: (i) ASCII, (ii) netCDF, or (iii) Fortran dumps. Whereas (i) is generally favored for the final results,
(ii) and (iii) are preferable for inter-module communication, e.g., when the trajectory information is passed from the calculation to the selection tool. Of course, LAGRANTO comes with a suite of auxiliary tools (see Sect. 5.2), which among other things allow to convert between the different output formats.

3.2 The specification of starting points (startf)

Starting positions for trajectory calculations are provided most easily as a text file with three entries per line: longitude (from \(-180\) to \(180\)), latitude (from \(-90\) to \(90\)), and pressure (in hPa). Every line then specifies a trajectory starting point. All air parcels specified in one so-called starting file will have the same start and end time of the trajectory calculation, as later described in Sect. 3.3.

As an alternative to manually constructing such a list of starting positions, LAGRANTO offers a tool (startf) to create more sophisticated starting files, e.g., when the manual listing would be too cumbersome either because of its length or the involved calculations. Basically, startf allows to describe starting positions based on geometrical and/or meteorological criteria. For instance, the specification \texttt{box.eqd[-20,30,-30,30,10]} defines horizontal starting positions in a box in the tropical Atlantic, extending from \(20^\circ\) W to \(30^\circ\) E and from \(30^\circ\) S to \(30^\circ\) N. The positions would be equidistantly distributed in this box (every 10 km, specified by the fifth parameter). In addition to a simple rectangular box, startf offers several other options to define the starting points in the horizontal (see Table 1), e.g., on the model grid or within an arbitrary spherical polygon, i.e., a polygon with longitude/latitude vertices on the sphere.

The choice of the vertical position of the starting points is passed to startf with two further arguments, separated by \(@\), where the first defines the “heights” of the starting points and the second the vertical unit. As an example, \texttt{profile[300,400,10] @ hPa} defines 10 equidistant starting levels between 300 and 400 hPa. If the second parameter was changed to \(K\), the 10 levels would be between the 300 and 400 K isentropic levels. For instance, \texttt{profile[300,400,3] @ K} would set the three vertical levels to 300, 350 and 400 K. A comprehensive list of possibilities to define the vertical positions is given in the
LAGRANTO reference guide (supplementary material), and a few examples are listed in Table 1.

In addition, starting positions can be defined based on meteorological criteria. For instance, with an additional argument $PV:GT:2$ passed to `startf`, only starting points with potential vorticity larger than 2 PVU will be selected. Of course, such a criterion makes only sense if a specific date is given because the PV field varies strongly in time. In fact, a date has to be passed to `startf` as a first parameter and only becomes relevant if a vertical coordinate other than hPa is applied.

Note that the applicability of meteorological selection criteria originates directly from two characteristics of LAGRANTO: (i) starting files created with `startf` are themselves trajectory files, and (ii) the selection tool `select` (see Sect. 3.4) operates on any trajectory file.

### 3.3 The calculation of trajectories (caltra and trace)

**Trajectory calculation:** Given a file with starting positions, the LAGRANTO tool `caltra` calculates forward or backward trajectories, where an intuitive syntax defines the direction: e.g., `caltra 20130719_06 20130720_06` calculates 24 h forward trajectories starting at 06:00 UTC 19 July 2013, and `caltra 20130720_06 20130719_06` correspondingly calculates 24 h backward trajectories starting at 06:00 UTC 20 July 2013. LAGRANTO correctly handles global and hemispheric domains, i.e., the crossing of the poles and the date line is fully supported. If wind fields are available only within a limited domain, trajectories crossing the lateral boundary are marked to do so with a missing value flag.

A trajectory, $\mathbf{x}(t)$, is obtained from the numerical integration of Eq. (1). The key elements of any numerical trajectory calculation are the temporal discretization and spatial interpolation. For the numerical time step, LAGRANTO uses a variant of the approach outlined by Petterssen (1956), which has also been used by, e.g., Whitaker et al. (1988) and Kuo et al. (1992). Starting at time $t$ at the position $\mathbf{x}$, and taking the wind at this position $\mathbf{u}(\mathbf{x}, t)$, the first iteration of the new position $\mathbf{x}^*$ at time $t + \Delta t$ is obtained from
the following “forward” time step
\[ x^{(1)} = x + u(x, t) \cdot \Delta t \],

where \( \Delta t \) denotes the time step. This first iteration only uses wind information at the starting position. For further iterations, the wind vector is averaged between the starting position and the previously estimated ending position, i.e.,

\[ u^{(1)} = \frac{1}{2}[u(x, t) + u(x^{(1)}, t + \Delta t)] \],

and the second iteration of the new position then reads

\[ x^{(2)} = x + u^{(1)} \cdot \Delta t \].

This simple scheme is also shown schematically in Fig. 3. With LAGRANTO, three iterative steps are applied per default. On request the number of iterations can be changed. The time step, \( \Delta t \), is either determined automatically as \( 1/12 \) of \( \tau \), the data time interval, or it can be manually set. For the spatial interpolation, LAGRANTO uses a bilinear interpolation in the horizontal and a linear interpolation in pressure in the vertical.

Special attention must be given to trajectories approaching the surface. If a trajectory crosses the lower boundary, i.e., if its pressure exceeds the surface pressure \( PS \), LAGRANTO stops the integration of this trajectory and its coordinates remain fix at the last position before hitting the surface. To allow the air parcel to proceed its course, LAGRANTO offers an option (the so-called “jump flag”) to artificially relaunch the air parcel \( 10 \text{ hPa} \) above the surface whenever it crosses the lower boundary (i.e., the integration is continued from the pressure level \( PS - 10 \text{ hPa} \)). The rationale for this very pragmatic procedure is that the accuracy of trajectories in the turbulent planetary boundary layer, given the typically very limited temporal resolution of the wind field is anyway critical, and care must be taken when interpreting individual trajectories that travel through the boundary layer\(^1\).

\(^1\)This is the starting point of so-called Lagrangian particle dispersion models, which statistically consider the effects of subgrid-scale turbulence on the flow of grid-scale air parcels.
**Tracing fields:** Whereas *caltra* delivers the coordinates of the air parcel’s path, *trace* allows to trace additional 3-D and 2-D meteorological fields along the computed trajectory. The fields must be available either on the primary P or secondary S files, or they can be calculated (online) based on the meteorological and positional information already available on a trajectory file. Note, however that the online calculation of meteorological fields is typically much less efficient than if they are pre-computed on a secondary input file and then can readily be traced along the trajectories.

Additionally, fields can be traced at positions offset from the air parcel’s path. For instance, *T:*−100 hPa traces the temperature field 100 hPa above the trajectory. If this position was below the ground, then a missing data value is set. Such a calculation can be useful, for instance, to assess the static stability in the layer beneath the trajectory or to determine the horizontal temperature gradient along the trajectory (e.g., with *T:*+100 km[lon] and *T:*+100 km[lon].

Note that LAGRANTO’s modularity allows to easily extend the list of fields traced along the trajectories, i.e., if an additional field turns out to be of interest, then no recalculation of the whole trajectory is necessary. A simple invocation of *trace*, e.g., in the form

```
trace trajectory.1 trajectory.1 ... 
... -f TH:+50hPa 1
```

adds two columns to the trajectory file, with values of potential temperature 50 hPa below and 50 hPa above the air parcels position, respectively. Furthermore, another option allows to trace the average, maximum or minimum values of a variable in a circle around the air parcel, which is useful to characterize the surrounding of the considered parcel (see reference guide in the Supplement).
3.4 Objective selection criteria (select)

A powerful aspect of LAGRANTO is its capability to select subsets of trajectories from a pre-calculated set, using a large variety of selection criteria. Many (thousands to millions) trajectories can be calculated in a first step and then a selection can be applied according to geometrical and/or meteorological criteria. In essence, *select* takes an input trajectory file, applies a filter specified by simple commands, and then writes out the selected trajectories that fulfill the filter conditions. On request, the output can be in the form of a boolean list or as a list with the trajectory number in the original file. The latter allows, for instance, to work with sets of trajectories (set used in a mathematical sense): Say one set contains all trajectories that pass over China, another all trajectories that pass over the United States. Then, it will be straightforward to find the common set of trajectory numbers being part in both sets. The resulting set of trajectory numbers can, of course, be transferred back into trajectory files, using the LAGRANTO *extract* tool. In the following, the selection criteria are further discussed. A glimpse at the possibilities can be inferred from Table 2, listing some common criteria.

**General structure:** Selection criteria are specified in the following way: `{COMMAND:FIELD:ARGUMENTS:TIME}`. The first parameter describes the operator that will be applied on the field indicated by the second parameter. As a typical example, one might want to test whether the trajectory position falls within a certain band of latitudes, say between 20–30° N. The corresponding criterion for *select* would be `{IN:lat:20,30:ALL(ANY)}`, where the time parameter needs some further explanations. *ALL(ANY)* means that the criterion is checked for all times of the trajectory, but the criterion has to hold only at least once. This selects trajectories that pass through the latitude band at any time. Instead, with *ALL(ALL)* it can be enforced that the trajectories remain in the band at all times, or with *ALL(NONE)* trajectories can be selected that are never located in the latitude band. Furthermore, the list of times when the criterion is applied can be specified in greater detail: *FIRST(ANY)* and *LAST(ANY)* refer to the first and last time steps of the trajectory (in this case the specification *ANY* is obsolete.
and can be omitted), and \textit{6,12,18(ANY)} applies the band criterion only at the times 6, 12, and 18 h (where time 0 h corresponds to the starting time of the trajectory).

**Difference and integral criteria:** A more detailed specification of the parameter \{\textit{FIELD}\} allows more refined trajectory selections. Often, temporal changes of a certain variable along the trajectory are used to select interesting subsets. \textit{select} supports this with the field extension \textit{DIFF}: for instance, \textit{GT:TH(DIFF):10:LAST,FIRST} computes the difference of potential temperature between the last and first trajectory position and checks whether it is larger than 10. In addition, trajectories can be selected based upon the integral of a certain variable (e.g., total precipitation \textit{RTOT}) along the trajectory path. In this case, the criterion \textit{GT:RTOT(SUM):10:ALL} integrates the field \textit{RTOT} over all time steps of the trajectory (indicated by the time parameter \textit{ALL}) and then checks whether the sum is greater (\textit{GT}) than 10. Note that no unit is specified with the command; therefore the valid unit is determined by the unit of the field on the trajectory file\(^2\). Similarly, one can imagine that the minimum \textit{RTOT(MIN)} or maximum \textit{RTOT(MAX)} of a field are of interest for the selection process, or it’s average \textit{RTOT(MEAN)} or temporal variance along the trajectory \textit{RTOT(VAR)}.

**Positional criteria:** There are many criteria that are based only on the air parcels’ positions. For this type of criteria, \textit{select} offers “built-in fields” that can be given as the second parameter even if they are not listed in the trajectory file. Examples for such parameters are \textit{DIST}, \textit{DIST0}, \textit{INBOX}, \textit{INPOLYGON}\(^3\) and \textit{INCIRCLE}. If air parcel should be selected that have traveled a certain distance, a criterion like \textit{GT:DIST:1000:48} checks whether the total distance traveled by the air parcel until time 48 h exceeds 2

\[^2\]Note that for this criterion to work, the field \textit{RTOT} must have been traced along the trajectories before.

\[^3\]The \textit{INPOLYGON} field relies on spherical polygons which are given as a list of longitude/latitude vertices and a single latitude/longitude point inside the polygon. The vertices define the boundary of the polygon, and the single point’s aim is only to specify which region is inside and which outside. Then use is made of the algorithm by Bevis and Chatelain (1989) to decide if a trajectory point falls within or outside of the given spherical polygon.
1000 km; or if DIST is replaced by DIST0 whether it has moved at least 1000 km away from its starting position, measured along a great circle. Of particular interest are criteria that rely on regions, which have to be crossed or avoided by the selected trajectories. Here select supports different types of region specifications. TRUE:INPOLYGON:file:ALL(ANY) tests whether the trajectory passes through (TRUE) a spherical polygon, e.g., a country border, which is specified as a list of coordinates in an external file. A simpler region specification is based on circles: for instance FALSE:INCIRCLE−40,50,500:LAST excludes (FALSE) all trajectories that are located at their final time step (LAST) within the circle with a radius of 500 km centered at 40° W, 50° N. Finally, rectangular regions given by longitude and latitude coordinates can be set, either with the inline command INBOX (see an example in Sect. 4.1) or by specifying the coordinates in an external file.

Logical combinations: LAGRANTO’s modularity guarantees that select can be invoked iteratively, e.g., the user can do a pre-selection of trajectories and later do more refined selections. If two select calls are invoked subsequently, this corresponds to a logical AND combination. Logical OR can be applied within a single call of select. If trajectories should be selected that experienced either a lot of diabatic heating (e.g., an increase of potential temperature of more than 10 K) or a strong ascent (e.g., a pressure decrease of more than 500 hPa), the following command can be used: GT:TH(DIFF):10:LAST,FIRST | LT:p(DIFF):500:FIRST,LAST. The second operator could also be written equivalently as GT:p(DIFF):500:LAST,FIRST. Note the logical OR operator |. The corresponding AND operator is &, and the order of operator execution follows standard Boolean algebra.

Combining the criteria of selection criteria comes with a pitfall: if for instance testing that an air parcel has a PV value greater than 2 PVU and is located above 300 hPa, a naive call would be GT:PV:2:ALL(ANY) & LT:p:300:ALL(ANY). However, this call only guarantees that the trajectory has once reached a PV value above 2 PVU and surpassed the 300 hPa surface; however the two conditions might be met at different times.
For enabling a coincident combination of conditions, a special feature of select can be used, namely the definition of label fields.

**Label fields:** To enforce that the two criteria in the previous example are met at the same time, a two-step procedure can be used. First, \(GT:PV:2:1(LABEL)\) and \(LT:p:300:2(LABEL)\) sets labels at all trajectory times when one or both of the conditions are met: label 1 if only the PV condition applies, 2 if only the pressure condition applies, and both labels if both criteria apply. Then, in a second step, the trajectories can be selected using these labels: \(ALL:LABEL:1,2:ALL(ANY)\) will select all trajectories for which labels 1 and 2 are both set simultaneously at any time along the trajectory, i.e. if \(PV > 2\) PVU and \(p < 300\) hPa. Note the special command \(ALL\), which can only be applied for the \(LABEL\) feature of \(select\). Two other specific commands are \(ANY\) and \(NONE\), which accepts trajectories for which any or none of the labels are set. The labeling of trajectory times can be used also in a different way. Suppose that all trajectory times are marked when the air parcel stays in a certain domain, e.g., over Iceland. Then we would like to select trajectories associated with precipitation \((RTOT)\) at exactly these times, i.e., when the trajectories are above Iceland. This can be achieved by first setting a label 1 when the air parcel is located over Iceland. Then the criterion \(GT:RTOT:0:LABEL(ANY)\) selects all trajectories that are associated with precipitation \((RTOT > 0)\) at any time over Iceland. Note, since in this example only one single label is set, \(LABEL(ANY)\) will select all trajectory times when this label is set (air parcel over Iceland). Together with the criterion \(GT:RTOT:0\) all trajectory times over Iceland and with precipitation are considered.

### 4 Illustrative examples

In this section, a selection of possible applications of LAGRANTO 2.0 are briefly presented to further illustrate the versatility of the tool. They relate to airflows in extratropical cyclones (Sect. 4.1), stratosphere–troposphere exchange (Sect. 4.2), orographic...
flows (Sect. 4.3), and the analysis of the air mass origin at a hypothetical measurement site (Sect. 4.4).

4.1 Refined analysis of a WCB

In Sect. 2, a warm conveyor belt (WCB) has been identified in the North Atlantic. Here, we extend the analysis of this WCB and thereby illustrate the exploratory application of different selection criteria. The starting point is the trajectory file wcb.1 created in Sect. 2, which contains trajectories with an ascent of at least 600 hPa within 48 h and starting from between 1050 and 790 hPa.

A first refined analysis could look at WCB trajectories that experience a different ascent, e.g., at least 550 hPa. The corresponding selection can be achieved for instance with the following criterion:

GT:p(CHANGE):550:FIRST, LAST

The field specification p(CHANGE) considers the pressure at the first and final trajectory time and checks if the modulus of the pressure change exceeds 550 hPa. Note that this criterion would also identify trajectories that descend more than 550 hPa. However, if we apply the selection criterion to wcb.1 it is clear that only ascending air parcels are captured.

It might be of special interest to consider WCB trajectories starting from an industrial region, e.g., the east coast of North America. Such a rectangular domain might extend from 80 to 60° W and 30 to 60° N. To select these trajectories, a region criterion must be applied:

TRUE:INBOX:-80,-60,30,60:FIRST

using the built-in field INBOX, which is logically true if the air parcel position falls within the specified rectangular box, and the first operator TRUE tests if the condition is met. The last argument specifies when the condition has to be fulfilled, here at the first time of the trajectory.
Whereas the selection criteria applied so far only rely on the path of the air parcels, the full power of explorative analysis comes into play if also meteorological fields are considered. For instance, one can consider only trajectories that experience extensive diabatic heating, as manifest in an increase of potential temperature:

\[ \text{GT:TH(DIFF):10:48,0} \]

selects trajectories with more than 10 K diabatic heating between 0 and 48 h, i.e., the difference of the potential temperature value at the end of the ascent minus the one at the beginning is larger than 10 K.

The next two examples are more complicated. The first one looks at trajectories with a non-zero ice water content (IWC) during a period longer than 12 h. We assume that IWC has been traced along the trajectories. First we mark all trajectory times with \( IWC > 0 \) with the label 1, and then we determine for each trajectory the time period during which the label is set and compare it to the required minimal time. This two-step approach translates into two corresponding selection criteria: determine for each trajectory the time period during which the label is set and compare it to the required minimal time.

\[ \text{GT:IWC:0:1(LABEL)} \]
\[ \text{GT:LABEL(NONZERO):2:ALL} \]

The second command counts the number of trajectory timesteps for which a label is set, considering all times (ALL), and compares this to the required minimum number of time steps (GT), which here is 2 (corresponding to 12 h for 6 hourly input data).

Also the second more complicated example relies on label setting. We would like to select all trajectories that loose at least 1.5 g kg\(^{-1}\) water in a certain domain, e.g., to assess whether the WCB trajectories significantly contribute to the precipitation in this domain. Let’s assume that the region is specified by a list of longitude/latitude vertices of a spherical polygon in a file named `greenland`. The first step would be to mark (with label 1) all trajectories crossing this region; then the change of specific humidity (\( Q \))
is considered during the time interval when the air parcel is within the marked region. The translation into selection commands is again straightforward:

\[
\text{TRUE:INPOLYGON:greenland:1(LABEL)} \\
\text{GT:Q(CHANGE):1.5:LABEL}
\]

i.e., the change in specific humidity, i.e. $Q(CHANGE)$, is found between the times for which the label is set. Hence, the last argument \textit{LABEL} in the criterion specifies again a list of trajectory times, more specifically those for which a \textit{LABEL} ist set. Of course, these times depend on the trajectory considered, i.e. it is not a fixed set of times.

### 4.2 Stratosphere–troposphere exchange

Air mass transport across the tropopause has been often quantified using Lagrangian techniques (see introduction). In this context, deep stratospheric intrusions are of particular interest, for instance because they can transport ozone-rich air to low-tropospheric levels. In this example we show how a quasi-operational forecasting system of stratospheric intrusions over Europe has been setup with LAGRANTO, which is used since more than a decade to inform measurement sites about potentially interesting episodes (e.g., Trickl et al., 2010). The setup is as follows: (i) ECMWF operational forecasts provide the meteorological fields for the trajectory calculations; (ii) every 12 h, four-day forward trajectories are calculated, starting in the entire region covering the North Atlantic and western Europe ($80^\circ$ W to $20^\circ$ E, $40^\circ$ N to $80^\circ$ N) and with starting heights between 250 and 600 hPa; (iii) from this large set of trajectories a subset is selected. The selected trajectories must initially reside in the stratosphere (i.e., their initial PV value must be larger than 2 PVU), and they must descend within four days by more than 300 hPa; and (iv) it will be verified if these “stratospheric intrusion trajectories” pass nearby over a measurement site, hence providing “hit tables” specific for each site.

The procedure described above can readily be transformed into a corresponding sequence of LAGRANTO calls, once the required input netCDF files are ready. The
first step with LAGRANTO is to define the starting positions once (we take the same
for all trajectory calculations), with the following criterion applied in `startf` (the complete
call can be found in the Supplement):

```plaintext
box.eqd[-80,20,40,80,50] @ ...
... profile(250,600,20) @ hPa
```

This specifies the equidistant (50 km) starting points in the desired box and 20 verti-
cal levels between 250 and 600 hPa. In principle, it would be possible to first calculate
all forward trajectories and then select the ones that are within the stratosphere at the
starting time. However, this approach is resource demanding. It is a better strategy to
reduce the starting points to those with a PV value larger than 2 PVU before calculating
the trajectories. Hence, we apply `select` to the starting file. The selection criterion is

```
GT:PV:2:FIRST
```

Figure 4a shows the starting positions at 250 hPa after this criterion is applied on
a particular day. The descent criterion can only be applied once the forward trajectories
have been calculated, i.e., after the call `caltra 20090 118_00 20090 122_00 startf.2
traj.1`. Note that for convenience, LAGRANTO comes with a little tool `newtime`, which
can be used to get the date string `20090 122_00` corresponding to 96 h (4 days) after
the starting date. The appropriate call is `newtime 20090 118_00 96`. Now, it is possible
to select the trajectories that descend more than 300 hPa within 4 days because all
positional information is ready on the trajectory file. The selection criterion is

```
GT:P(RANGE):300:ALL & ...
... LT:P(DIFF):0:FIRST,LAST'
```

Here, the first part guarantees that the maximum minus the minimum pressure ex-
ceeds the requested 300 hPa, and the second part makes sure that the air parcel is
at a lower altitude at the end of the 4 day period than at the beginning. The selected
trajectories are shown in Fig. 4b.
Now, we might want to assess whether one of these stratospheric intrusion trajectories passes near a surface measurement site, e.g., the Jungfraujoch (JFJ) in the Swiss Alps at a height of 3471 m a.s.l. and at 7.985° E, 46.547° N. For testing whether a trajectory passes this point within a certain distance in the horizontal (here 100 km), the function \texttt{INCIRCLE} can be used, which has been described in Sect. 3.4:

\begin{verbatim}
TRUE:INCIRCLE:7.985,46.547,100:ALL(ANY)
\end{verbatim}

From the large number of descending trajectories shown in Fig. 4b, six fulfill this geometric criterion (Fig. 4c). They split into two distinct clusters, both originating in the western Atlantic, but exhibiting a very different evolution in the vertical (Fig. 4d).

4.3 Orographic flows

Orographic flows pose a particular challenge for trajectory analyses due to the complex topography, which influences the wind field on small scales. Here we consider an example where the flow is blocked by the Alps, a mountain chain with about 1000 km west–east and 100 km north–south extension. The horizontal resolution of the ECMWF model is not fully sufficient to accurately model the fine-scale flow. Therefore, we here used the LAGRANTO version adapted to the limited-area model COSMO (see Sect. 5.1 for further details). The horizontal resolution of the meteorological fields is 7 km, on 60 vertical levels and the wind fields are available every hour from the operational COSMO analyses provided by the Swiss national weather service.

Here, we consider the case of a north-westerly flow approaching the Alps. The trajectories are started at 750, 1500, 2250 and 3000 m a.s.l. every hour during the period from 09:00 UTC 28 January 2009 to 09:00 UTC 30 January 2009. First, the starting positions are defined along a line well upstream of the Alpine barrier (Fig. 5a). The corresponding specification for \texttt{startf} is (for, e.g., the 1500 m starting height)

\begin{verbatim}
line(0,16,49,52,100) @ level(1500) @ m
\end{verbatim}
where 100 equidistant points are defined along a line between the points $0^\circ$ E, $49^\circ$ N and $16^\circ$ E, $52^\circ$ N, all at a height of 1500 m. The initial distance from the Alps was chosen larger than the Rossby radius of deformation, which roughly determines how far the mountains influence the upstream flow (Pierrehumbert, 1985). Note that in the COSMO version of LAGRANTO, starting heights are given in m, in accordance with the model’s vertical coordinate. If pressure is also available on the P files, of course the starting height can also be given in hPa. Internally, LAGRANTO will then convert the pressure heights into geometrical heights and will do all the further calculations with these latter vertical coordinate. It is one main feature of startf to offer this versatility to the user.

The trajectories are then calculated with the following call, where the starting positions are on the file startf_1500m.2 and the trajectories are written to a file traj_1500m.1:

```
caltra 20100128_09 20100130_09 ...
   startf_1500m.2 traj_1500m.1 -j
```

Note that the jump flag (-j) is set, hence all trajectories crossing the lower boundary are artificially lifted to 10 m above the surface and allowed to move on. Figure 5b shows the resulting trajectories, indicating a nice example of orographically blocked flow: the air parcels are horizontally deflected around the Alpine barrier.

An overview how the path of the trajectories depends on the starting height can be readily obtained with the visualization tool quickview included with LAGRANTO (see Sect. 5.3), e.g.,

```
quickview traj_750m.1
quickview traj_3000m.1
```

Whereas the trajectories starting at low levels are strongly deflected by the Alpine barrier, the flow becomes more and more unperturbed if initialized at higher levels (Fig. 6).

We next consider the question whether air parcels pass over the Gotthard pass ($8.56^\circ$ E, $46.57^\circ$ N) in the Swiss Alps:
Here, all trajectories are selected which pass through a circle with 18 km radius around the Gotthard pass. For the previously calculated trajectories, this is the case for a single trajectory. To check whether the flow of this air parcel is adiabatic or whether it is influenced by diabatic processes, one can consider the potential temperature evolution in the trajectory file. In case that this field is not available on the primary and secondary input files, LAGRANTO still allows to trace this field along the trajectories, provided that temperature and pressure are already available on the trajectory file or are found on the input files. A call `trace gotthard.1 gotthard.1 -f TH 1.`, assuming that the single trajectory is saved on file `gotthard.1`, will automatically decide whether the field is ready for tracing on an input file or whether it can be computed online from existing fields on the trajectory file. In our case potential temperature changes from 289.6 K at the beginning to 292.2 K at time 43 h when the air parcel crosses the lateral boundary of the COSMO domain. Therefore this particular flow is almost isentropic.

Sometimes it is very interesting to investigate the vertical structure of the atmosphere along a specific trajectory, for instance, the mountain-crossing air parcel discussed above. To this end, LAGRANTO’s `profile` tool can be used. At each trajectory time and position, it extracts not only the value of a meteorological field at the air parcel’s position, but a complete vertical profile starting from the surface to the upper troposphere. The resulting field can then be visualized as a time-height plot (see Fig. 5c) and this allows studying processes that occur above or below the considered trajectory (e.g., the evolution of the PBL underneath the trajectory, the vertical extension of a cloud, or the potential occurrence of gravity wave activity). The example shown in Fig. 5c was created by the call

```
timeres gotthard.1 gotthard.1 -min 15
profile gotthard.1 profile.nc
```

which takes the trajectory in the input file `gotthard.1` and creates a netCDF file with the corresponding vertical profiles of all meteorological fields listed in `tracevars` (see
Sect. 3.3). In order to increase the time resolution of the trajectory from 1 h to 15 min, the LAGRANTO tool timeres (see Table 3) was called before profile. Between model output times a linear (or on demand cubic spline) interpolation is applied. Note, the vertical range of the pseudo profile and/or the meteorological fields included can all be passed as options to the tool (for details see the LAGRANTO reference guide and Sect. 5.4).

As a second example, we would like to find all trajectories that reach 90 % relative humidity (RH) in a region “around the Alps”, more precisely when they are less than 300 km away from the 1500 m height contour of the Alps. Such a criterion might be of interest in assessing if and how very moist air parcels are affected by the Alps. The region of interest is saved as a spherical polygon in a file alps_300km.txt, which lists first one point within the polygon and then all points along its boundary. The first few lines of the file are hence,

8.568 46.572  
18.711 47.648  
18.667 47.898  
18.592 48.133  
18.479 48.374  
18.340 48.596  
18.182 48.794  
18.023 48.960  
...

In Fig. 5d the spherical polygon is shown by filled yellow circles. Shown are also the trajectories reaching at least 90 % relative humidity within this polygon and the position where this happens (filled magenta circles). They are selected by a three-step selection criterion:
where the first marks (with label 1) all trajectory positions and times inside the spherical polygon. Analogously, the second criterion marks (with label 2) instances where the relative humidity exceeds the 90% threshold. Finally, it is checked whether the two labels are set simultaneously, and if so the trajectory is selected. In total, 72 of the 100 trajectories fulfill this sophisticated criterion: they are shown in Fig. 5d.

4.4 Origin of air masses at observational site

In the previous examples, only forward trajectories were considered. Here, we are interested in the origin of air masses arriving at a measurement site. This asks for backward trajectories that are released in the surrounding of the measurement site. Technically, the calculation of kinematic backward trajectories is straightforward and is done by simply rotating the wind vectors by 180°.

As an illustrative example, we consider here the origin of air masses arriving at the ABC-Pyramid Atmospheric Research Observatory in the Himalayas (Bonasoni et al., 2008). This station, situated at 5079 m a.s.l., monitors ozone and other atmospheric constituents since March 2006. LAGRANTO is used to provide daily a quasi-operational air mass source diagnostic (http://evk2.isac.cnr.it). The forecast trajectories help predicting whether the air masses arriving at Pyramid station are, for instance, stratospheric or originating from the planetary boundary layer.

Figure 7 shows as an example, re-calculated with the new version of LAGRANTO, 120 h backward trajectories started in a circle around the ABC-Pyramid site from the 500 hPa level. Note that the starting points are equidistantly distributed, with a spacing of 10 km, in a circle of radius 100 km around the measurement site (86.80° E, 27.95° N). This approach allows to quantify the degree of horizontal coherence of the air arriving at the station and to identify different air streams that are potentially involved. LAGRANTO
offers an easy way to create starting positions displaced from the actual point of interest. In the present example, the following option for `startf`:

```plaintext
circle.eqd(86.80,27.95,100,10) @ ...  
... list(500) @ hPa
```
defines equidistant grid points in a circle, given by \((\lambda, \phi)\) of the center and the radius 100 km. The distance between starting points is given by the last parameter 10 km for `circle.eqd`. As a starting height we choose 500 hPa only.

Given the start file `startf.2`, backward trajectories can be calculated with `caltra`. For instance, if the date of interest is 00:00 UTC 26 Oct 2007, 5 day backward trajectories can be calculated with `caltra 20071031_18 20071026_00 startf.2 traj.1 -j`, which writes the trajectories to `traj.1`. Hence, backward trajectories are simply selected by providing the start and end date in the correct order. The resulting trajectories are shown in Fig. 7, colored with pressure to show the height evolution of the different air streams.

Different aspects of the backward trajectories can now be studied in more detail. For instance, it is possible to separate long-range transport from local transport by using the intrinsic LAGRANTO field `DISTO`, which measures the great circle distance between an air parcel’s position at a specific time and its initial position:

```plaintext
GT:DIST0:1000:LAST
```
selects a trajectory if the distance between the initial and final time (indicated by `LAST`) exceeds 1000 km. Or, it might be of interest to select only the trajectories that descend more than 250 hPa. Again, an intrinsic field can be used:

```plaintext
GT:VERT0:250:ALL(ANY)
```
Here, `VERTO` considers the pressure difference between the initial time and any other time step along the trajectory. Note the difference to the call

```plaintext
GT:p(DIFF):250:FIRST,LAST
```
which would only consider the pressure difference between the initial and final time steps.

A further question could address if the air parcels arriving at ABC-Pyramid observatory had recently contact with the planetary boundary layer (PBL), potentially identifying polluted air masses. To check if this applies to some of the selected trajectories, we consider their vertical distance from the top of the PBL. Assuming that the pressure at the PBL top is available on the input files as a variable “PBL” and traced along the trajectories, then the tool tracal\(^4\) can be used to calculate this vertical distance and to add it as a new column to the trajectory file (which is called traj.1 in this case):

\[
\text{tracal traj.1 traj.1 'DPBL=P-PBL'}
\]

Using this new variable, a criterion can be invoked for selecting the trajectories that had contact with the PBL (note that for such trajectories DPBL becomes negative):

\[
\text{GT:DPBL:0:ALL(ALL)}
\]

In the present example, this is never the case, i.e., we can readily assume that no PBL polluted air is directly transported to the ABC-Pyramid measurement site.

Finally, the origin of the air masses at Pyramid station might be assessed on a climatological basis. Here, we take one year (2007) of backward trajectories arriving at Pyramid station at 500 hPa. The basic strategy of the trajectory calculation is the same as the one outlined before. However, only one single starting position at “exactly” the Pyramid station’s coordinates is set and the backward trajectories are calculated based on the ERA-interim re-analysis of ECMWF. All backward trajectories are merged into one large trajectory file using LAGRANTO tool mergetra, comprising \(4 \times 365\) trajectories (starting every 6 h). This large number of trajectories can no longer be visualized in the way that each single trajectory is shown: the figure would be completely filled with trajectories, not allowing to see any structure. Instead, trajectory densities can be used,

\^4The tracal calculations are based on the Fortran String Utilities developed by George Benthien. They can be downloaded from URL www.gbenthien.net.
i.e., all trajectories are gridded onto a regular latitude/longitude grid and then for each grid point the number of trajectories associated with this grid point is shown (Fig. 7b). If all trajectories are merged into one single file `alltraj.1` with `mergetra`, it is very easy to get the trajectory density. To this aim, LAGRANTO's `density` tools must be called:

```
density alltraj.1 alltraj.nc
```

The gridded trajectory density is then available on the netCDF file `alltraj.nc` and can readily be visualized.

## 5 Special features and special versions

### 5.1 LAGRANTO for other models

LAGRANTO was initially developed for ECMWF analysis and forecast data, but meanwhile several other models are supported. In particular, a fully functional version exists for the non-hydrostatic regional model COSMO (Baldauf et al., 2011). The COSMO and the ECMWF versions differ mainly because of differences in the model grid: COSMO operates on a rotated longitude/latitude grid and its vertical levels are based on geometrical height instead of pressure. Another important difference, of course, is the higher spatial and temporal resolution of COSMO, and the limited domain. Currently, COSMO is run operationally by the German and Swiss weather services with horizontal resolutions of about 7 and 2 km and output is archived at 1 h time steps. A COSMO version of LAGRANTO is run operationally at MeteoSwiss to provide short-term forecasts of the dispersion of pollutants and radioactive emissions. Recent research applications of COSMO trajectories include the calculation of moisture origins for measurements of stable water isotopes (Aemisegger et al., 2014) and for heavy precipitation events (Winschall et al., 2014), the Lagrangian analysis of cirrus cloud observations (Cirisan et al., 2014), and the detailed analysis of foehn flows (Miltenberger et al., 2013). It is noteworthy that the last of these studies also describes the implementation of an on-
line trajectory module within COSMO, which is based on the LAGRANTO algorithms but calculates the trajectories during the model integration using wind fields at every model time step.

Preliminary LAGRANTO version, for which not all functionalities have been implemented and tested yet, are also available for the UK Met office Unified Model (UM), the Weather Research and Forecasting (WRF) modeling system and the 20CR re-analysis data set (Compo et al., 2011). Whereas UM is based on a rotated longitude/latitude grid, WRF uses a different horizontal grid structure. For an example of a UM model study based on LAGRANTO see Elvidge et al. (2014).

When working with data from COSMO or the UM, additional tools can be used to facilitate the conversion between geographical and rotated coordinates. For instance, `geo2rot -lonlat -170 45` followed by `geo2rot trainp traout` first defines a coordinate system with the rotated north pole at 170° W, 45° N (used for the grid of the COSMO or UM model simulation), and then the geographic coordinates in the input trajectory file `trainp` are converted to rotated ones and written to `traout`. This is particularly helpful if the COSMO-trajectories are to be shown together with COSMO model output, i.e., if the air parcel positions are overlaid directly on meteorological fields from the COSMO model.

### 5.2 Tools

LAGRANTO comes with a set of auxiliary programs that facilitate working with trajectories. Some of them have already be mentioned in the preceding sections. A full list of these tools can be found in the reference guide (see supplementary material) and a few are briefly introduced here. First, `trainfo file` gives some meta information about the trajectory file. For instance, `trainfo file dim` provides the dimensions, i.e., the number of trajectories, time steps and fields (columns) in the trajectory file; and `trainfo file list` list all trajectories irrespective of the format in which they are stored.

A second tool `reformat` allows to change the format of the trajectory file. The different file formats have a different compression level, which can be essential if large traject-
tory files have to be archived or passed from one LAGRANTO program (e.g., caltra) to another (e.g., trace). A typical conversion is reformat file.nc file.ls, which converts from netCDF to ASCII format. The most compact and fastest format is a Fortran dump reformat file.ls file.du, at the cost of losing human readability. Furthermore, a small number of trajectories can be converted to KML format for visualization with Google Earth (see below).

Often, it is of interest to combine different trajectory files or to extract certain features from a single file. For instance, two trajectory files file1 and file2 cover the same time period and are started from the same positions, but contain complementary meteorological fields along the trajectories. Then mergetra file1 file2 file will merge the two trajectory files and write a new resulting trajectory file. mergetra can also merge forward and backward trajectories starting at the same time and position. In contrast, the tool extract allows to extract pieces of information from a trajectory file. In particular, specified columns or times can be extracted; or selected trajectories given by their index number; or extract file -startf writes all starting positions of a trajectory file to a list.

5.3 Trajectory visualization

Different tools are provided for trajectory visualization, based on (i) Matlab, (ii) NCL, and (iii) Python. Example scripts show how LAGRANTO trajectory files can be imported and the trajectories displayed on geographical maps. The trajectories can be color-coded according to meteorological fields. For instance, pressure can be used to reveal the height evolution of the trajectories. The most convenient visualization relies on Python, for which LAGRANTO comes with a quickview tool to instantly visualize trajectories (as already presented in Sect. 4.3). The most basic call is

quickview file

Most trajectory plots in this paper are prepared with Matlab.
which reads the trajectory file and shows all trajectories on a equidistant cylindrical projection. The geographical domain is automatically chosen to include all trajectories. On request it can be manually set. Furthermore, it is possible to select a field on the trajectory file that is used for coloring them, for instance

```
quickview -v RH -d nh trajectory
```

displays the trajectories in the Northern Hemisphere (nh) and colors them according to their relative humidity (RH). Two different geographical projections are supported: equidistant cylindrical and polar stereographic.

Finally, LAGRANTO can write a KML file for a limited number of trajectories, which then can be loaded into Google Earth for visualization and further study (see example in Fig. 8). A KML file essentially needs the latitude/longitude position of the air parcels and the geometrical height a.s.l. Whereas the horizontal position is part of all valid trajectory files, the height needs special consideration. If the geopotential height is available on the file, as is the case for every LAGRANTO trajectory based on the COSMO model, this will be inserted into the KML file when invoking the file conversion to the KML format. However, since ECMWF uses pressure as the vertical coordinate, the height is not available on a typical ECMWF trajectory file. In this case, LAGRANTO assumes a US standard atmosphere to scale from pressure to height a.s.l. when converting to the KML format.

### 5.4 Availability, installation and documentation

LAGRANTO is disseminated over an SVN repository (https://svn.iac.ethz.ch/websvn/pub), where the different versions are available as distinct releases. LAGRANTO has been successfully installed on Linux platforms, using different Fortran compilers (PGI, Intel, gfortran). The installation is controlled with a shell script and details of the installation are listed in a document. In addition to Fortran, several other software tools are needed, in particular Unix csh, Perl and partly Python.
Finally, a first-step user guide and a detailed reference guide are provided. The latter lists the details of all options for all core programs (e.g., for the trajectory selection).

6 Conclusions

This article described in some detail a new version of the Lagrangian analysis tool LAGRANTO and presented some illustrative examples of its application. Compared to the original version of LAGRANTO, which has been in use since almost 20 years, many important parts of the code have been completely rewritten. LAGRANTO allows to calculate forward and backward trajectories based on fields from different NWP models (ECMWF, COSMO, UM, WRF), to trace meteorological fields along these trajectories, and to select subsamples based on kinematic, geographical and meteorological criteria. The very special features of LAGRANTO, compared to other trajectory tools, are

1. the highly flexible definition of trajectory starting positions,
2. the versatile and iterative selection of subsets of trajectories from a previously calculated larger ensemble,
3. availability of the tool for global datasets (e.g., ECMWF analyses) and very high-resolution numerical models (e.g., convection-permitting COSMO simulations).

Also thanks to these characteristics, which facilitate exploratory research with trajectories, LAGRANTO is already well established as a useful research tool for atmospheric dynamics, air chemistry, and cloud microphysics. It is run either online during the integration of a NWP model (Miltenberger et al., 2013) or, in most cases, as a stand-alone offline Lagrangian analysis tool (as described in this article). With the improved new release of offline LAGRANTO, new capabilities are introduced that further increase the options for trajectory-based research. Since the access to high-quality data from NWP model might be limited for potential users, we also plan to develop a web-based version of LAGRANTO, which will allow to perform trajectory calculations over a web interface.
similar to the already existing HYSPLIT model. This step will complete the trias of LAGRANTO trajectory models available to the research community: (i) COSMO-based online trajectories; (ii) offline trajectories using input fields from different models calculated by the user on-site; and (iii) offline trajectories provided by web-LAGRANTO, which can be invoked from a user-friendly web-GUI.

The Supplement related to this article is available online at doi:10.5194/gmdd-8-1893-2015-supplement.

Acknowledgements. We thank MeteoSwiss for providing access to ECMWF analysis and forecast data. The LAGRANTO user community is acknowledged for feedback on shortcomings of the tool and for suggesting novel features. We are particularly grateful to Stephan Pfahl (ETH Zurich) for developing the first COSMO version of LAGRANTO, to Suzanne L. Gray (University of Reading) for helping to develop the UM version, and to Andrew Penny (Naval Postgraduate School, Monterey) for setting up and testing the WRF version. Thanks also to Nicolas Piaget and Sebastian Schemm (ETH Zurich), who provided the Python and NCL visualisation tools, respectively.

References


Table 1. Some examples for the creation of starting files (see Sect. 3.2). The specification is passed to LAGRANTO by program \textit{startf}. A typical call would be of the form: \{\textit{startf YYYY-MM-DD HH startf criterion}\}, where the first argument defines the starting date of the calculation, \textit{startf} is the name of the files with the starting positions and the criterion, as given in the left column, specifies the positions. In the right hand column, only the part in bold font of the specification.

| point(−10,50) @ list(450 500,550) @ hPa | single position (10° W, 50° N) at 450, 500, 550 hPa |
| point(−10,50) @ list(450 500,550) @ hPa,agl | 450, 500, 550 hPa a.g.l. |
| \textbf{box.grid(−10,-5,40,50)} @ level(300) @ hPa | all grid points in box 10–5° W / 40–50° N |
| box.grid(−10,-5,40,50) @ level(320) @ K | on 320 K isentrope |
| \textbf{polygon.eqd(greenland,20)} @ level(100) @ hPa,agl | equidistant (20 km) points within polygon "greenland" |
| point(−10,50) @ \textbf{profile(100 900,10)} @ hPa | vertical profile (10 values) from 100–900 hPa |

1933
Table 2. Overview examples of selection criteria. A typical call of LAGRANTO program *select* takes the following form: `{select inp.trajectory out.trajectory criterion}`, where the first two arguments provide the input and output trajectory files and the third one defines the selection criterion, as exemplarily given in the left column. Only the part in bold font of the specification is explained in the right column. For a complete list of all options with examples, consider the reference guide in the Supplement.

<table>
<thead>
<tr>
<th>Selection Criterion</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>GT:PV:2:ANY</td>
<td>potential vorticity (PV) larger than 2 PVU at any time</td>
</tr>
<tr>
<td>GT:PV:2:ALL</td>
<td>PV larger than 2 PVU at all time</td>
</tr>
<tr>
<td>GT:PV(MAX):2:ALL</td>
<td>PV maximum larger than 2 PVU</td>
</tr>
<tr>
<td>GT:PV(DIFF):2:FIRST,LAST</td>
<td>PV difference larger than 2 PVU</td>
</tr>
<tr>
<td>LT:DIST:1000:LAST</td>
<td>length of trajectory less than 1000 km</td>
</tr>
<tr>
<td>FALSE:INCIRCLE:8,45,200:ANY</td>
<td>not within circle (radius 200 km) centered at 8° E, 45° N</td>
</tr>
</tbody>
</table>
Table 3. Several tools included in LAGRANTO. The last two tools are only available for the COSMO and UM version of LAGRANTO, i.e. the models which work on a rotated geographical coordinate system. For a detailed description of usage, including all options, please consult LAGRANTO’s reference guide.

- **extract**: extract columns, specified times or single trajectories
- **mergetra**: merge different trajectory files
- **reformat**: change the format of a trajectory file (ASCII, netCDF)
- **timeres**: change time resolution of trajectories by linear interpolation
- **tracal**: do simple calculation with trajectory file
- **trainfo**: list a trajectory file or extract meta information
- **geo2rot**: change from true to rotated longitudes/latitudes (for COSMO version)
- **rot2geo**: change from rotated to true longitudes/latitudes (for COSMO version)
Figure 1. Lagrangian identification of a warm conveyor belt (WCB). Two-day forward trajectories are started at 12:00 UTC 29 January 2009 over the North Atlantic between 1030 and 790 hPa (30 hPa steps). The starting points are equidistantly distributed (80 km horizontal spacing; blue dots). Trajectories ascending more than 600 hPa within 48 h are selected as WCB trajectories and colored according to potential temperature (in K).
**Figure 2.** Flowchart showing the typical steps in a LAGRANTO trajectory calculation (LAGRANTO tools are written in bold). The input files include the 3-D wind fields at several time steps. Starting positions (longitude/latitude/pressure) are defined with `startf`. Then trajectories are calculated with `caltra`, and possibly subsamples of trajectories are selected with `select`. Along these trajectories additional fields are traced with `trace`, which might then ask for further `select/trace` iterations. Finally, the resulting trajectories can be further analysed or visualized: `quickview` shows individual trajectories; `density` calculates trajectory densities on a geographical grid, and `profile` allows studying the air columns the trajectory is passing through. For further details, see text (Sect. 3).
Figure 3. Schematic representation of the iterative Eulerian time step used in LAGRANTO. The air parcel's position at time \( t \) is given as \( x \). The velocity field at this position is \( u(x) \). Based on the position, the velocity and the time step, a new air parcel position can be calculated with a simple Euler forward step: \( x^* = x + \Delta t \cdot u(x) \). However, a simple Euler forward step is not numerically accurate. Therefore, in a next iteration a new velocity \( u^* \) is computed as the average of the velocities at the initial position and at the forward projected position. A new air parcel position \( x^{**} \) is then calculated based on this refined guess of the velocity. In LAGRANTO, three such iterative steps are repeated to obtain the final new position of the air parcel.
Figure 4. Example for a stratosphere–troposphere exchange (STE) forecast based on LAGRANTO trajectories. (a) Equidistant (50 km horizontal spacing) starting position at 250 hPa in the stratosphere (PV > 2 PVU, in color shading) at 00:00 UTC 18 January 2009; (b) Selection of 96 h forward trajectories that descend more than 300 and start above 400 hPa (coloring according to pressure); (c) Refined selection of descending trajectories that pass the Jungfraujoch (JFJ) measurement site within 100 km horizontal distance. The color shading gives the averaged wind speed at 400 hPa during the four days and the dots mark the position of the air parcels at the given times; (d) vertical evolution of JFJ trajectories, colored according to their PV value. The dots mark the time and pressure when geographically the air parcel is closest to JFJ.
Figure 5. (a) Starting positions (black dots) and 48 h forward trajectories (grey lines) started at 09:00 UTC 28 January 2009; additionally the topography is shown in color, and the wind arrows at 500 hPa at the trajectory starting time; (b) trajectory passing over the Gotthard pass. Time intervals of 1 h are marked along the trajectory by filled circles; (c) vertical profiles along the Gotthard trajectory with time as the horizontal axis and height a.s.l as the vertical axis. Potential temperature (in K) is shown in color, vertical velocity (in m s\(^{-1}\)) as black (upward) and blue (downward) contours. The vertical evolution of the trajectory is shown as a bold black line with markers (filled circles) every hour; (d) trajectories that reach at least 90 % relative humidity within the spherical polygon outlined by the small yellow dots, corresponding to a distance of 300 km of the 1500 m height contour of the Alps. The yellow dot over Switzerland defines the interior of the spherical polygon. Magenta dots mark the positions where RH > 90 %. 

1940
Figure 6. 48 h forward trajectories started at 09:00 UTC 28 January 2009 along a line north of the Alps. The two different panels correspond to different starting heights of the trajectories: (a) 750 m a.s.l. and (b) 1500 m a.s.l. The individual trajectories are colored according to their height. All panels were created with the visualization tool quickview included in LAGRANTO (see Sect. 5.3).
Figure 7. (a) 120 h backward trajectories started within 100 km from the ABC-Pyramid Atmospheric Research Observatory in the Himalayas (see text) at 18:00 UTC 31 October 2007. The starting height is 500 hPa and the trajectories are colored according to their pressure; the position of the air parcels 48 and 96 h before arrival are marked with black dots; (b) density of 120 h backward trajectories from Pyramid station started at 500 hPa for the whole year 2007 (for details, see text).
Figure 8. Trajectories as in Fig. 5d, but converted into KML format, which can then be added as a separate layer to a Google Earth map.