**Farewell package, Pernilla Kühn, 1.08.2025**

**Folder structure**

* Figures – all final figures included in my master’s thesis
* Software\_levante
  + convertERA5 packaage to convert ERA5 data to lat/lon coordinates, can be used for individual parameters or to get all data for a FLEXPART run (still needs to be preprocessed using starting prepare\_flexpart.py from flex\_extract, and resort produced EA files)
  + flexpart\_v11\_start\_multiple

skripts to prepare and start FLEXPART runs for GOSAT retrievals and In-situ measurements

* + NA\_inversion

Skripts to prepare, run and first analysis of CO2 flux inversion

* Software\_atmo

Prepare\_TM5-4DVar data diurnal cycle scaling factors, mole fraction fields and preprocess ObsPack data (see README for details)

* Data, see */work/bb1170/RUN/b382762/data/README.md*

**ERA5 data conversion (Levante):**

For details see: /work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/convert\_era5/README.md

* Install flex\_extract (see IUPEDIA Flexpart)
* Install ERA5 conversion package (https://github.com/ATMO-IUP-UHEI/PK\_FarewellPackage/tree/main/software\_levante/convert\_era5)
* Use */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/convert\_era5/convert\_era5/prepare\_flexpart.sh* to run all necessary steps for FLEXPART preparation, (alternatively steps can also be run individually):

Before running: adapt file paths listed in README.md, adapt config file (e.g. /work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/convert\_era5/convert\_era5/config/config\_test.yaml)

The tree steps are:

* 1. Convert ERA5 data on Levante and combine into required files (convert to lat/lon coordinates, produces ANOG\_\_ML, ANOG\_\_SL …. files) using */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/convert\_era5/convert\_era5/Start\_ConvertEra5.sh*
  2. Flex\_extract preprocessing, creates *EAYYMMDDHH* files using */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/convert\_era5/convert\_era5/prepare\_flexpart.sh*
  3. Resorting parameters in created EA files (need specific order for FLEXPART) using */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/convert\_era5/convert\_era5/Bulkstart\_prepare\_EA\_files.sh*

(Sometimes the preprocessing failed, but worked when restarting it a second time)

* Create AVAILIBLE file (file containing list of EA files, necessary for FLEXPART) using */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/mkAVAIL\_v7.py* (see IUPEDIA Flexpart)

Run with mk\_AVAILABLE.py -m EA -p *outputpath* -a *outputpath* to save AVAILIBLE file in same directory as EA files

* Delete old EA files to save storage

**GOSAT measurement data (Levante):**

* Transfered relevant yearly GOSAT RemoTeC v2.4.0 data files from Atmo server '/mnt/data/users/eschoema/GOSAT/NetCDF\_files\_withCH4/fp\_short\_fil\_corr\_201202\_*YYYY.*nc '
* Create GOSAT positions csv files (RemoTeCv2.4.0\_YYYYMMDD.csv, used to setup FLEXPART runs) containing sounding position, time and measurement value, using */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/flexpart\_v11\_start\_multiple/create\_GOSATpositions.py*

Using config file*: config\_path configs/* *gosat\_config\_RemoTeC240.yaml*

For details about skript see: /work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/flexpart\_v11\_start\_multiple/README.md

**In-situ measurement data (Atmo & Levante):**

On Atmo server:

(For detail about skript see */work/bb1170/RUN/b382762/master\_farewell\_package/software\_atmo/README.md*)

* Used obspack\_co2\_1\_GLOBALVIEWplus\_v5.0\_2019-08-12 dataset, '/net/dsvr-02/mnt/data2/users/eschoema/ObsPack/obspack\_co2\_1\_GLOBALVIEWplus\_v5.0\_2019-08-12/data/nc/'
* Select data, calculate 4h mean with */work/bb1170/RUN/b382762/master\_farewell\_package/software\_atmo/selectObsPackData.py*
* Selects surface & tower measurements within defined region & time period, CarbonTracker assimilation flag CT\_assim=1, only uses highest tower inlet
* transfer created file to Levante (e.g. *obspack\_co2\_1\_GLOBALVIEWplus\_v5.0\_2019-08-12\_sel\_mean\_combined.nc*)

On Levante:

For details about skript see (/work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/flexpart\_v11\_start\_multiple/README.md)

* Create IS positions csv files (*ISpositions\_YYYYMMDD.csv*, used to setup FLEXPART runs) containing: start of 4h period, location, inlet height and CO2 concentration, with */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/flexpart\_v11\_start\_multiple/create\_ISpositions.py*

Using config file*: config\_path configs/IS\_config.yaml*

**Runing Flexpart (Levante):**

For details about skript see (/work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/flexpart\_v11\_start\_multiple/README.md)

* Instal FLEXPART v11 (see IUPEDIA Flexpart)
* Create necessary files and directories for FLEXPART runs (output directories, initial release condition files, options directories, pathname files)
  + For GOSAT total column releases: */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/flexpart\_v11\_start\_multiple/prepare\_GOSATruns.py*

Using config file: configs/options\_config\_RemoTeC240.yaml

* + For in-situ 4h releases: */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/flexpart\_v11\_start\_multiple/prepare\_ISruns.py*

Using config file: configs/options\_config\_IS.yaml

* Start Flexpart runs using specifications in pathnames files

Three possibilities to start files:

* 1. Single flexpart run: */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/flexpart\_v11\_start\_multiple/slurm\_flexpart\_v11.sh*
  2. For all pathnames files in one directory (all within one slurm job): */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/flexpart\_v11\_start\_multiple/slurm\_start\_multiple.sh*
  3. Multiple pathnames directories (in separate slurm jobs): */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/flexpart\_v11\_start\_multiple/Bulkstart\_multiple.sh*

**TM5-4DVar (Atmo & Levante):**

On Atmo Server:

For more details about skript see */work/bb1170/RUN/b382762/master\_farewell\_package/software\_atmo/README.md*

Fluxes:

* Get diurnal cycle scaling using */work/bb1170/RUN/b382762/master\_farewell\_package/software\_atmo/prep\_TM5-4DVar.py* on Atmo server

Calculates total scaling factors based on TM5-4DVar 3hourly prior data, interpolate to hourly resolution, need to adapt region and time period in main function

Transfer resulting file (*high\_res\_total\_scaling\_RemoTeC+IS.nc*) to Levante

Mole fraction fields:

* Cut desired region, calculate pressure at TM5-4DVar layer boundaries, calculate xco2

*/work/bb1170/RUN/b382762/master\_farewell\_package/software\_atmo/prep\_TM5-4DVar.py* on Atmo server

Transfer created files (xco2\_mean\_YYYYMMDD.nc) to Levante

On Levante:

For more details about skript see: */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/NA\_inversion/README.md*

* Transfer files from Atmo server to Levante: ‘/mnt/data/users/eschoema/TM5Inversion/glb3x2\_20220413/new\_gridded\_flux/’
  + *flux\_1x1\_RemoTeC\_2.4.0+IS.nc* for using RemoTeC\_2.4.0+IS version
  + *flux\_1x1\_prior.nc*
* cut to region, get total flux in necessary units, weekly time resolution using */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/NA\_inversion/prepare\_inversion.py* with PREP\_TM5\_4DVAR\_REF\_FLUXES=True flag

**Preprocessing for Inversion (Levante):**

Using skript: */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/NA\_inversion/ prepare\_inversion.py*

For more details about skript see */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/NA\_inversion/README.md*

Measurements:

* Calculate background values for in-situ and GOSAT releases

with PROCESS\_FLEX\_RUNS=True flag

(or can be done separately for in-situ and GOSAT measurements, see README)

* Saves values into csv files (in *FLEXPART\_directory/RemoTeCv240/TM5-4DVar\_estimate/YYYY\_MM/* and *FLEXPART\_directory/insitu/TM5-4DVar\_estimate/YYYY\_MM/...*)

Footprints:

* Get 1x1 footprints for all releases in a month in one xarray dataset

with GET\_HIGH\_RES\_FOOTPRINTS=True flag

* Apply TM5-4DVar diurnal cycle scaling, sum to weekly footprint values

with GET\_TM5\_4DVAR\_SCALED\_FOOTPRINTS=True flag

* Combine footprints for entire time period into one dataset, coarsen to 2x2° resolution

with COARSEN\_HIGH\_RES\_FOOTPRINT=True flag

Saves all footprint with weekly 2x2 resolution in e.g.: */work/bb1170/RUN/b382762/data/FarewellPackage\_test/Flexpart/insitu/prep\_footprints/scaled\_weekly/high\_res\_scaled\_footprints\_20100601-20100606\_2x2\_weekly.nc*

(for measurement time period of 20100601-20100606)

**Running the Inversion (Levante):**

* Install pyinverse ([GitHub - ATMO-IUP-UHEI/pyinverse: Solve various inverse problems](https://github.com/ATMO-IUP-UHEI/pyinverse))
* Using skript: */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/NA\_inversion/ run\_inversion.py*

But run with slurm: sbatch run\_inv.sh

(For more details about skript see */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/NA\_inversion/README.md*)

* Adapt configurations given in inversion\_config/config.yaml

Set time period of measurements to be assimilated, output directory, resolution, with/no correlation, footprint scaling, TM5-4DVar background dataset

Optionally: run additional inversion with flat prior, save entire averaging kernel, filter Gosat measurements …

**First Analysis (Levante):**

Using skript: */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/NA\_inversion/ analyze\_inversion.py*

For more details about skript see */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/NA\_inversion/README.md*

* Calculate modelled measurement values
  + For our inversion: using GET\_CO2\_VALS flag

Multiplies footprints (select scaled footprints when including diurnal cycle) with posterior fluxes from performed inversion

* + For TM5-4DVar: using GET\_HIGH\_RES\_PRIOR\_CO2 and GET\_HIGH\_RES\_TM5\_POSTERIOR\_CO2 flags

Multiplies fluxes with high resolution footprints (1x1 and hourly)

* Plot time series of prior, posterior ad TM5-4DVar reference 2x2 inner domain using PLOT\_INNER\_2X2\_FLUXES flag (or 4x4 inversion entire domain using PLOT\_4x4\_FLUX\_MAP)
* Plot histogram of differences between modelled values and measurements using PLOT\_DIFF\_HIST flag
* Create mask and geodataframe for easter, western US using GET\_EAST\_WEST\_MASKS flag

**Plots in master’s thesis (Levante):**

See */work/bb1170/RUN/b382762/master\_farewell\_package/software\_levante/MA\_plots.ipynb*, descriptions in respective sections