

Improvements in Extpar

Internship January to May 2020 at C2SM

- Improvements to the source code
 - Logger
 - Removed ~ 200 compiler warnings
 - Python rewrite
- Organization of the input datasets
 - Collect all input data into Git-LFS repository
- Improvements to the code infrastructure
 - Enhanced testing with Jenkins

Source Code

```
Flake data buffer filename: flake_buffer.nc
ENTER write_netcdf_buffer_flake
write out ext_par_flake_cosmo.nc
Enter write_netcdf_cosmo_grid_flake
===== flake_to_buffer done =====
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:43]
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:44] ==== Code information used to build this binary ====
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:45]
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:46] Compile-Date .....: 2020-09-07 05:58:10
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:47] Code is modified ..: modified
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:48] Library name .....: extpar-5.4
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:49] Compiled on .....: mlogin
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:50] Revision number ...: d3ac853241d6f4d31ff9147465544e7bfaa98877
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:51] Checkout-Date .....: 2019-11-28 13:02:25
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:52] Binary name .....: /mnt/lustre01/scratch/b/b381001/extpar/test/testsuite/work/mch/c7_globe/extpar_consistency_check.exe
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:53] Compiled by .....: b381001
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:54] Compiler version ..:
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:55] Current start time : 2020-09-07 08:04
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:56] ==== End of code information ====
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/info_extpar.f90:57]
  lradtopo: T
  nhoriz:      24
           0
  io extpar:      0
  raw data:      0
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/extpar_consistency_check.f90:595] isoil_data: 1
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/extpar_consistency_check.f90:597] ldeep_soil: F
  Warning *** ifill valley has to be 1 or 2 ***
           *** set ifill valley = 1 (default value)! ***
  namelist_grid_def: INPUT_grid_org
  igrd_type:      2
  domain_def_namelist: INPUT_COSMO_GRID
  after reading namelist INPUT_COSMO_GRID
  ie_tot, je_tot:      601      421
  ke_tot:      0
  Allocated lon_rot and lat_rot
  Allocated lon_geo, lat_geo, no_raw_data_pixel
  Cosmo domain coordinates determined with calculate_cosmo_target_coordinates
  target grid tg:      601      421      1 -22.281927349247017      42.281927349247006      31.718572811288968      59.300000000000004
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/extpar_consistency_check.f90:714] Land use dataset      : GLC2000
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/extpar_consistency_check.f90:715] Land use lookup table: Heise_2005
[2020-09-07 08:04:39][mnt/lustre01/scratch/b/b381001/extpar/src/extpar_consistency_check.f90:793] Read INPUT_CHECK for COSMO
  soil_buffer_file = soil_buffer.nc
```

Source Code

```
==== Code information used to build this binary ====

Compile-Date .....: 2020-09-06 13:56:54
Code is modified ..: clean
Library name .....: extpar-v5.4
Compiled on .....: mlogin
Revision number ...: 906e396a3270816a69f8344f75a84bc92b009b33
Checkout-Date .....: 2020-07-30 16:37:29
Binary name .....: ./extpar_consistency_check.exe
Compiled by .....: b380729
Compiler version ..: GCC 6.4.0
Current start time : 2020-09-06 16:04
==== End of code information ====

===== start consistency_check =====

===== read namelist =====

Scale separation active
Enter routine: read_namelist_extpar_orosmooth
***WARNING: ifill_valley has to be 1 or 2 => set ifill_valley = 1 (default value)!
Exit routine: read_namelist_extpar_orosmooth
Enter routine: init_target_grid
Enter routine: read_namelist_extpar_grid_def
Enter routine: allocate_com_target_fields
Exit routine: allocate_com_target_fields
Enter routine: calculate_cosmo_target_grid_coordinates
Exit routine: init_target_grid
Land use dataset   : GLC2000
Land use lookup table: Heise_2005
Urban data ISA active
Urban data AHF active
Enter routine: read_namelist_extpar_check_cosmo
Number of special points: 0
Tile mode: 0
Flake correction: T
Exit routine: read_namelist_extpar_check_cosmo
Tile mode for EXTPAR is set to tile_mode=      0 tile_mask= F

===== allocate fields =====

Enter routine: allocate_lu_target_fields
Enter routine: allocate_add_lu_fields
Enter routine: allocate_soil_target_fields
Enter routine: allocate_isa_target_fields
Enter routine: allocate_add_isa_fields
Enter routine: allocate_ndvi_target_fields
Enter routine: allocate_era_target_fields
Exit routine: allocate_era_target_fields
Enter routine: allocate_topo_target_fields
Exit routine: allocate_topo_target_fields
Enter routine: allocate_aot_target_fields
Exit routine: allocate_aot_target_fields
Enter routine: allocate_cru_target_fields
```

- Each Extpar program writes to own logfile
- 3 types of messaging levels
- Helpful for debugging
- In parallel consistent formatting of entire Fortran code base
~ 1 month of work
- Remove around 200 compiler warnings

Source Code

- Merge `sgsl_to_buffer` into `extpar_topo_to_buffer`
 - Remove redundant/duplicate code
 - Less maintenance
 - Include preprocessing for GLOBE into Extpar
 - `lprepoc_oro = .true.`

For Aster no matching preprocessing could be found!

Source Code – Python rewrite

- Python/CDO rewrite of 4 Fortran subprograms
 - Albedo, T_CRU, NDVI, Emissivity
 - Generalization of approach introduced for Icon with release 5.0
 - Full support for Icon and Cosmo-grids
 - Seamless integration into existing Extpar workflow
- Advantages
 - Much better (and more) interpolation methods by CDO
 - Much faster for large Icon model grids
 - Convenience of Python for maintenance

Source Code – Python rewrite

```
#-----  
#-----  
logging.info('')  
logging.info('===== CDO: remap to target grid =====')  
logging.info('')  
  
# calculate weights  
utils.launch_shell('cdo', '-f', 'nc4', '-P', omp, f'gendis,{grid}',  
                  raw_data_alb_1, weights)  
  
# regrid 1  
utils.launch_shell('cdo', '-f', 'nc4', '-P', omp,  
                  f'setrtoc,-1000000,0.02,0.02',  
                  f'-remap,{grid},{weights}', raw_data_alb_1, alb_cdo_1)
```

```
#-----  
#-----  
logging.info('')  
logging.info('===== reshape CDO output =====')  
logging.info('')  
  
alb_nc_1 = nc.Dataset(alb_cdo_1, "r")  
  
if (ialb_type == 1):  
    alb_nc_2 = nc.Dataset(alb_cdo_2, "r")  
    alb_nc_3 = nc.Dataset(alb_cdo_3, "r")  
  
if (igrid_type == 1):  
  
    # infer coordinates/dimensions from CDO file  
    ie_tot = len(alb_nc_1.dimensions['cell'])  
    lon = np.rad2deg(np.reshape(alb_nc_1.variables['clon'][:],  
                                (1, 1, ie_tot)))  
    lat = np.rad2deg(np.reshape(alb_nc_1.variables['clat'][:],  
                                (1, 1, ie_tot)))  
  
    je_tot = 1  
    ke_tot = 1
```

Organization of the input datasets

Problem

- Different data-files used at DKRZ, CSCS or other facilities
 - Same name, different data
 - Different name, same data
 - Not all users could access all data
 - No documentation at all

Organization of the input datasets

Problem

- Different data-files used at DKRZ, CSCS or other facilities
 - Same name, different data
 - Different name, same data
 - Not all users could access all data
 - No documentation at all

Solution

- Collected all existing data sets for Extpar into Git-LFS repo
 - Shared place where all users/developers can work on
 - At least some sort of documentation
 - Same workflow as normal git (commits, pull request, ...)
 - Around 380 GB (compressed) of data

Infrastructure

- Enhanced testing for Mistral, Tsa and Daint with Jenkins
- GCC, NAG and Intel compiler
- Two test to ensure coding standards for the future
- Allow roundoffs for pre-defined fields
- And many other small improvements

Read more about it:

- [Release Notes](#)
- [Official Documentation](#)
- [Extpar Input Data \(Git-LFS\)](#)
- Or ask me:
- `jonas.jucker@env.ethz.ch`