FUME – a new open source emission processor for air quality models

N. Benešová[3], M. Belda[2,4], K. Eben[1,5], J. Geletič[1,5], P. Huszár[2,4], P. Juruš[1,2], P. Krč[1,5], J. Resler[5], and O. Vlček[3]

[1] Institute of Computer Science, Czech Academy of Sciences
[2] Faculty of Transportation Sciences, Czech Technical University in Prague
[3] Czech Hydrometeorological Institute
[4] Faculty of Mathematics and Physics, Charles University, Czech Republic
[5] Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague

Work was supported by the grant of the Technology Agency of the Czech Republic, project No. TA04020797.

Presentation outline

- Motivation
- Philosophy behind FUME
- Description of
  - configuration
  - inputs
  - special modules
  - outputs
- Computational efficiency
- Current status & outlook

Motivation

- **What do we need?**
  - emission for AQ model
  - Chemically speciated emission fluxes in model grid (or points)

- **What do we have?**
  - Annual totals of basic pollutants in different spatial resolution and geometries
  - Activity data for emission calculation
  - All (im)possible file formats, units, naming conventions, etc.

Motivation

- Need for file preprocessing, chemical speciation, time & space disaggregation
- Only very few emission processors publically available and widely used
  - SMOKE – need of emission preprocessing, US-specific some things were hardcoded and not possible to change flexibly
  - after long-time experience with SMOKE we decided to write our own emission processor (EP)

Demands on new EP

- Flexible and configurable tool for emiss. proc.
- Independent on specific input data
- Applicable in different scales (local / regional / continental)
- Not limited to any geographical coordinate systems
- Configurable chains of emissions processing

Demands on new EP

- Easy implementation of new modules
- Easy involvement of external models (e.g. MEGAN) through common interface
- Outputs usable for different types of models / application
- Alloving for output also on irregular grids and general polygons
- Reporting and QA/QC (need to to keep as much info as possible)

Demands on new EP – fun with data ☺

- **Different geometries:**
  - Point
  - Area – regular grid, irregular polygons
  - Line
  - Only implicit geometry instructions
  - …
- **Different naming conventions:**
  - NOX = NOx = Nox_as_NO2
  - PM2_5 = PM25
  - …
- **Different units**
  - sometimes differ even in the same inventory file
- **Different geographical projections**

FUME

- Technologies used:
  - PostgreSQL v9.5 or newer
  - PostGIS v2.2 or newer
  - Python v3.5 or newer

    + some Python libraries (see fume/doc/QUICK-START.txt)
- https://fume-dev.github.io/fume GNU GPL v3.0

Configuration files

Source/emission categories

- list of emission sources categories

```
"cat_id"   ,"name"                                                    ,"parent"
0          ,"default category",
2000000    ,"SNAP 2 - NON-INDUSTRIAL COMBUSTION PLANTS"               ,0
3002000000 ,"POL: SNAP 2 - NON-INDUSTRIAL COMBUSTION PLANTS"          ,2000000
```

- Every category has it's parent – used for speciation and time disaggregation
- Category definition is completely upon user (standard categories e.g. SNAP can be also used)

Configuration files

Chemical speciation splitting inventory pollutants into model species. **FUME accepts two ways of chemical speciation:**

SMOKE format (GSPRO file)

- direct splits for from inventory pollutants to model species
- mole or mass based split factors

```
# GSPRO for gaseous pollutants
#
CB05
# main pollutants - mapping to themselves
0,"SO2","SO2"    ,1, 64.064, 1
0,"NO" ,"NO"     ,1, 30.006, 1
0,"NO2","NO2"    ,1, 46.006, 1
0,"CO" ,"CO"     ,1, 28.010, 1
0,"NH3","NH3"    ,1, 17.031, 1
0,"CH4","CH4"    ,1, 16.04 , 1
0,"BZN","BENZENE",1, 78.11 , 1
#
#   Passant (2002) - profile 125: Domestic wood combustion"
2020204, "NMBVOC", "PAR" , 0.120, 36.457, 0.046
2020204, "NMBVOC", "OLE" , 0.055, 36.457, 0.042
2020204, "NMBVOC", "XYL" , 0.013, 36.457, 0.039
2020204, "NMBVOC", "ETH" , 0.471, 36.457, 0.362
2020204, "NMBVOC", "ETOH", 0.288, 36.457, 0.364
2020204, "NMBVOC", "ETHA", 0.139, 36.457, 0.115
2020204, "NMBVOC", "UNR" , 0.037, 36.457, 0.015
...
...
```

Configuration files

Chemical speciation splitting inventory pollutants into model species. **FUME accepts two ways of chemical speciation:**

SMOKE format (GSPRO file)

- direct splits for from inventory pollutants to model species
- mole or mass based split factors

Profiles of particular chemical compounds

- accepts profiles of chemical compounds
- assignment of chemical compounds to chemical mechanism species must be provided
- format consistent with output files of Chemical Speciation Database by Carter, 2015 (J Air Waste Manag Assoc, 65 (10))

both ways can be combined together

Configuration files

Time profiling

- Default: factors for month / day of week / hour of day
- Optional: explicit time factors for each timestep

tv_series.csv | model_list.csv | inventory_species.csv | USER_GUIDE.rst

```
cat_id     , time_loc             , tv_factor
3102020200, 2015-01-01 00:00:00, 0.75758
3102020200, 2015-01-01 01:00:00, 0.7177
3102020200, 2015-01-01 02:00:00, 0.7177
...
3102020200, 2015-12-31 21:00:00, 1.92906
3102020200, 2015-12-31 22:00:00, 1.42893
3102020200, 2015-12-31 23:00:00, 0.60015
...
...
```

- Possibility to include special module for calculation of time factors e.g. based on meteorology (residential heating).

Emission inputs

inventory_input.txt – input files and their assignment to the certain inventory. Option to filter data on input to dtb.

inventory_name	file_name	file_path	file_info_path
	CZ_ZSJ_2016	CzechRep/geom/zsj_2016.shp	CzechRep/geom/zsj_2016.info
# TNO MACC-III v 1.1 emissions			
TNO_3	TNO_MACC_3_v1_1_2011	TNO_MACC_3_v1_1/TNO_3_v1_1_2011.txt	TNO_MACC_3_v1_1/TNO_3_v1_1_2011.info
TNO_3	TNO_MACC_3_v1_1_2011	TNO_MACC_3_v1_1/TNO_3_v1_1_2011.txt	TNO_MACC_3_v1_1/TNO_3_v1_1_2011.info
# POLAND – Malopolska road transport			
POL_MP_road	POL_MP_road_bochenski	Poland/Malopolska/roads_bochenski.csv	Poland/Malopolska/roads_bochenski.info
POL_MP_road	POL_MP_road_brzeski	Poland/Malopolska/roads_brzeski.csv	Poland/Malopolska/roads_brzeski.info
POL_MP_road	POL_MP_road_dabrowski	Poland/Malopolska/roads_dabrowski.csv	Poland/Malopolska/roads_dabrowski.info

set_name (opt.)	filter (opt.)
TNO_MACC_3_v1_1_2011_REST_EUR	ISO3!=CZE,ISO3!=SVK,ISO3!=POL
TNO_MACC_3_v1_1_2011_SVK_!SNAP2	ISO3=SVK,SNAP!=2
POL_MP_road_1	
POL_MP_road_2	
POL_MP_road_3	

Emission inputs – info files

csv with point sources

```
# type of file
file_type       = text
field_delimiter = ';'
text_delimiter  = '"'
encoding        = 'utf8'

# number of lines before header to be skipped
skip_lines = 0

# source type
src_type = P

category  = SNAP
source_id = ID_EMIT, ID_SRC, ID_OUTLET

# point source parameters
height      = HGT , meter
diameter    = DIAM, m
temperature = TEMP, degC
velocity    = SPED, m/s

# geometry
crd_east  = x_gauss
crd_north = y_gauss
EPSG      = 28403
```

Emission inputs

Categories mapping and different naming conventions / units solved within inventory.

inventory_name	file_name	file_path	file_info_path
	CZ_ZSJ_2016	CzechRep/geom/zsj_2016.shp	CzechRep/geom/zsj_2016.info
# TNO MACC-III v 1.1 emissions			
TNO_3	TNO_MACC_3_v1_1_2011	TNO_MACC_3_v1_1/TNO_3_v1_1_2011.txt	TNO_MACC_3_v1_1/TNO_3_v1_1_2011.info
TNO_3	TNO_MACC_3_v1_1_2011	TNO_MACC_3_v1_1/TNO_3_v1_1_2011.txt	TNO_MACC_3_v1_1/TNO_3_v1_1_2011.info
# POLAND – Malopolska road transport			
POL_MP_road	POL_MP_road_bochenski	Poland/Malopolska/roads_bochenski.csv	Poland/Malopolska/roads_bochenski.info
POL_MP_road	POL_MP_road_brzeski	Poland/Malopolska/roads_brzeski.csv	Poland/Malopolska/roads_brzeski.info
POL_MP_road	POL_MP_road_dabrowski	Poland/Malopolska/roads_dabrowski.csv	Poland/Malopolska/roads_dabrowski.info

Emission inputs

Categories mapping and different naming conventions / units solved within inventory.

category_<inventory_name>.csv

```
orig_id,cat_id
1      , 1000000
2      , 2000000
...
71     , 7000000
72     , 7000000
73     , 7000000
...
```

species_<inventory_name>.csv

```
inv_specie_name, ep_specie_name, inv_unit
NOX_as_NO2     , NOX           , t/year
CO             , CO            , t/year
SO2            , SO2           , t/year
C6H6           , BENZENE       , t/year
benzo_a_pyrene , BAP           , g/year
...
```

Emission inputs

calculate_pollutants.csv

```
"category","expression"

3007000000,"NO=0.652*NOX-0.652*NO2, NMBVOC=VOC-BZN-CH4"

2002000000,"BZN=0.00044*NMVOC, NMBVOC=NMVOC-BZN"
```

Meteorology inputs

Currently option to use WRF or ALADIN meteorology. Can be easily extended to others (MCIP format, RegCM etc.)

Processing of data

User configurable transformation chains with reasonable default

Example transformations

- Filtering
- Masking
- Gridding
- Increase/decrease of emission by factor
- …

- Speciation
- Time disaggregation

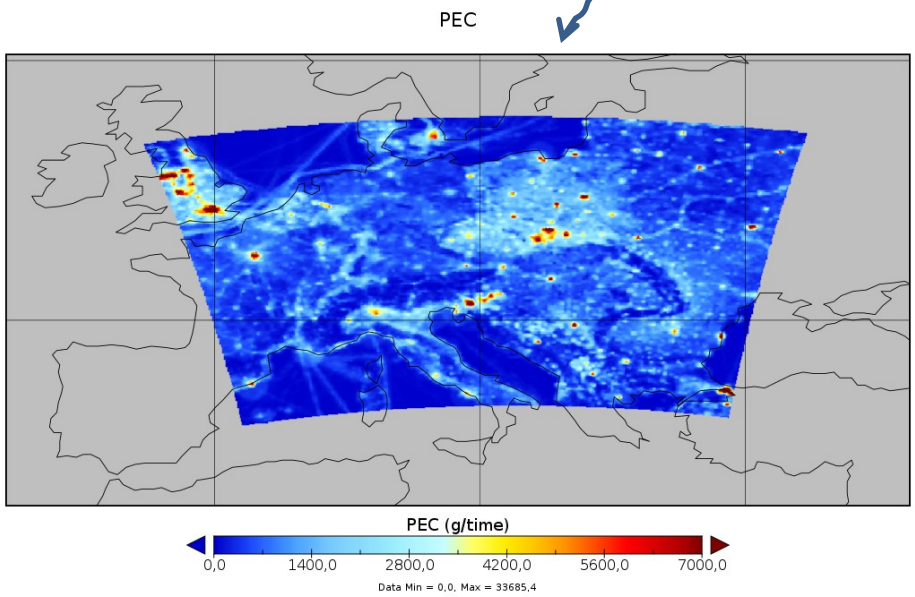
- Calculation of external models (currently MEGAN for biogenic emiss, NH3agri – NH3 emission from agriculture)

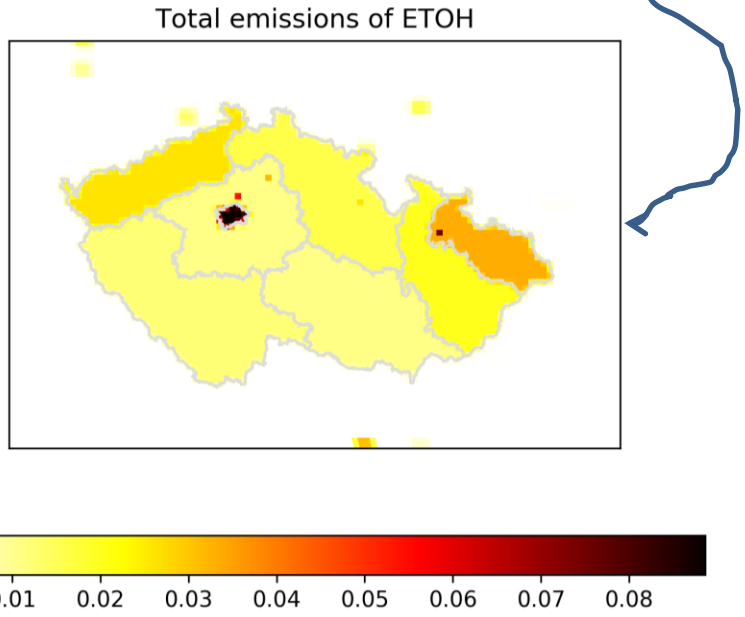
Processing of outputs

Outputs are done through the provider-reciever bus

- Parallel processing of different outputs
- Easily configurable
- Easy addition of new output module

Currently outputs for CAMx, CMAQ and graphical plotter





Computational efficiency

Import of data to the database

- 4 mil sources (450 000 point, 150 000 line, 3 400 000 area; 43 mil combinations of source-category-species)
- 80 min (done once)

Calculation for domain covering Czech Rep. and Slovakia –
– 182 x 101 cells (ca 900 000 sources in domain)

- case preparation, intersecting and other spatial operations
- 33 min (done once if grid or other case parameters do not change)
- Calculation 25 hours of emissions for CAMx (point and area files, 27 species)
- 44 min (time disaggregation most demanding)

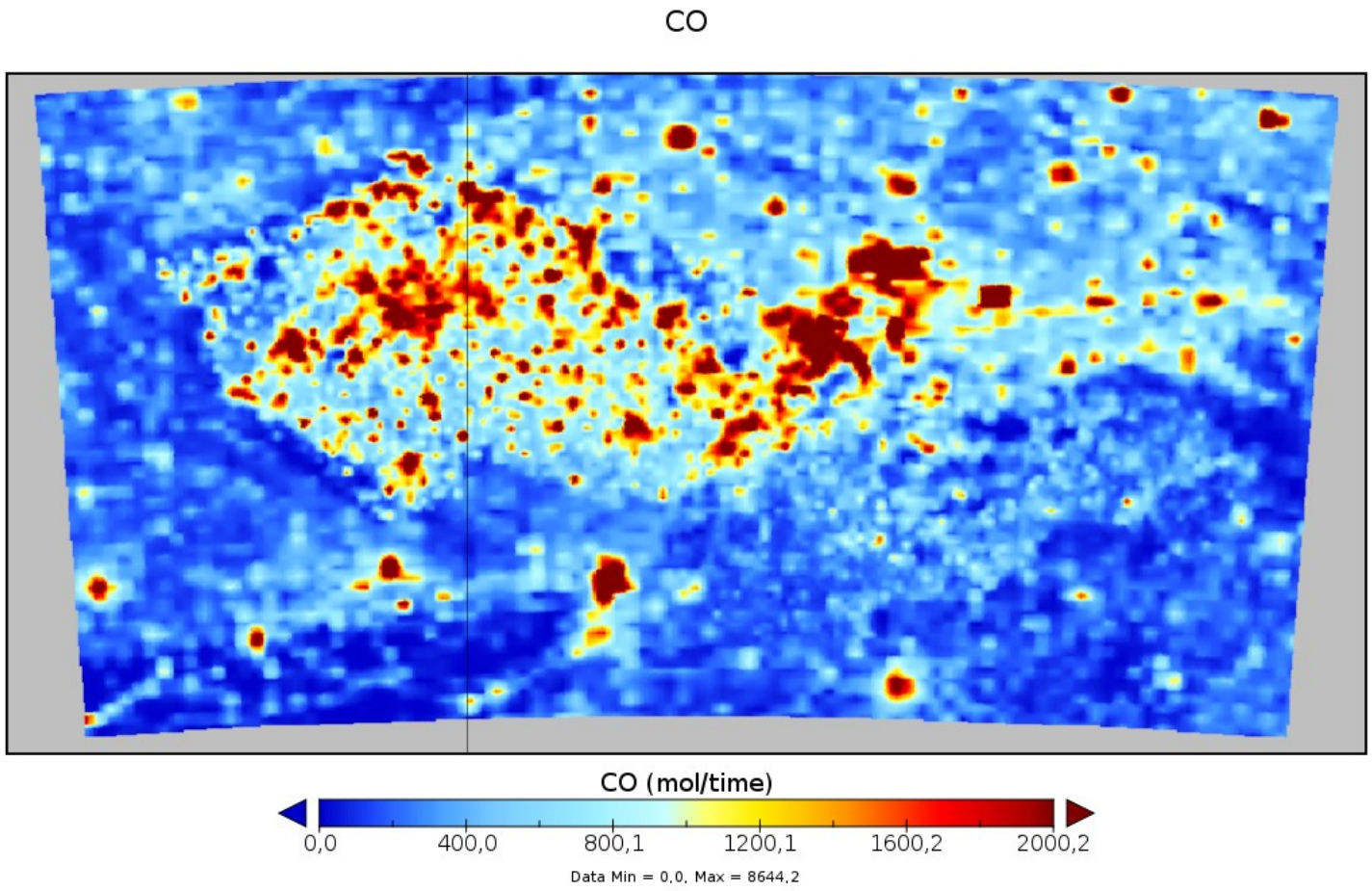
Current status of FUME

- Being used in several projects (LIFE IP Malopolska, Urbi Pragensi…)
- Still under development
- Feedback and contribution from modelling community welcome
- We are ready to provide advice for new FUME users
- Improve documentation
- Improve dedicated web page www.fume-ep.org, which will contain also sample data

Outlook & future plans

- Improve efficiency – data handling / parallelization
- Additional possibilities to process emissions
- User-friendly usage/reporting/logging
- World-wide user community ☺…

Thank you!


CO
CO (mol/time)
0,0
400,0
800,1
1200,1
1600,2
2000,2
Data Min = 0,0, Max = 8644,2