Universitat Rovira i Virgili (URV) y Universitat Oberta de Catalunya (UOC)

Master in Computational and Mathematical Engineering

FINAL MASTER PROJECT

Area: Modeling and Simulation

Implementation and evaluation of a parallelization strategy for the HERMESv3 atmospheric emission modeling system

https://earth.bsc.es/gitlab/es/hermesv3_bu

Author: Carles Tena Medina
Director: Alberto García Villoria
Co-Director: Marc Guevara Vilardell
Co-Director: Mario César Acosta Cobos

Barcelona, June 13, 2020

Dr./Dra. Alberto García Villoria, certifies that the student Carles Tena Medina has elaborated the work under his/her direction and he/she authorizes the presentation of this memory for its evaluation.

Director's signature:

Credits/Copyright

FINAL PROJECT SHEET

Title:	Emission sector manager
Author:	Carles Tena Medina
Tutor:	Alberto García Villoria, Marc Guevara Vilardell, Mario César Acosta Cobos
Date:	06/2020
Program:	Master in Computational and Mathematical Engineering
Area:	Modelling and Simulation
Language:	English
Key words	Emissions, Bottom-Up, MPMD

Dedicatory / Quote

I want to dedicate that work to Yoli who is giving me her support wherever she was.

Acknowledgments

I would fist acknowledge Kim Serradell who have encouraged me to keep a continuous formation and make my working environment as comfortable as possible, he has also provide me all the facilities that are on his hands to do that work as my chief on the Computational Earth Science group of the Barcelona Supercomputing Center.

In the same way and with the same enthusiasm I would to acknowledge Marc Guevara who is not only a work college, he has become also a friend. Without Marc's expertise that work would never have existed.

I will also acknowledge my colleges Mario Acosta and Miguel Castrillo for their daily feedback on the evolution of the HERMESv3 software that have help a lot on the resolution of any problem that I had, and will have.

Miriam Olid is also a special person in my life that makes easier to manage all the working environment problems that can occur in the office and I do not want to forget an acknowledge for her.

Also my parents Ana and Jose Antonio who have helped me to become the person that I am today. They have always been there when I needed them. And also my cousins Neus, Mar, and Yoli that are like sisters.

And at last, but not least, Manel Cantos, my partner, who supports me in my day to day, listens to me even if he do not understand what I was saying and makes my life happier waking up next to him every day.

Abstract

Air pollution remains as one of the major environmental issues. In this context, air quality modelling systems are seen as a key tool that can forecast air pollution episodes and provide guidance for the design of effective emission abatement measures. A combination and integration of multiple numerical models is needed to simulate the release, transport, transformation and deposition of air pollutants, namely meteorological, emissions and chemical transport models. Among these components, emission models are considered to be the highest source of uncertainty to the air quality modelling results. In order to reduce this uncertainty, the development of emission models capable of producing detailed and very high spatial and temporal emissions is mandatory.

To achieve that, the geo-location and temporal distribution of these emissions have to be provided with the highest resolution possible required by the chemical transport model. However, the sources emissions are derived from many different sources grouped by sectors such as road transport, shipping or livestock management, which implies the integration of multiple estimation algorithms and the management of large amounts of statistical and geo-localized datasets that are provided in multiple formats.

The High-Elective Resolution Modelling Emission System version 3 (HERMESv3), a tool developed by the Earth Science department of the Barcelona Supercomputing Center, is a objected oriented python based model developed for high computing environments. HERMESv3 can compute the emissions using different bottom-up approaches applicably for multiple sources of emission sectors.

This work presents and describes the sector managing system developed within HERMESv3, that allows to scatter the computational resources between the different sector calculators needs in the most efficient way possible.

HERMESv3 can be applicable to any European region on several user-defined grid projections to obtain the necessary inputs for different atmospheric chemistry transport models (CMAQ, WRF-Chem and MONARCH) as well as a street-level dispersion model (R-LINE).

Keywords: Emission model, Air quality, Multiple-Program Multiple-Data, HPC, Bottom-up

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Since the time that human beings developed the ability to reason, we have been trying to understand the physics and the behaviour of the world surrounding us.

Along this time, knowledge has been developed and with it we have found equations and parametrizations that allow us simulating most of the physical and chemical processes that describe the world we live in. Nevertheless, these formulas have become too complex to solve them without any additional help besides the human brain.

Computer sciences have been a key factor in the development and improvement of the computational numerical modeling and the computational fluid dynamics. These tools have allowed us to reproduce the physical behaviour in models and to achieve the desired output using a short amount of time thanks to different available high computing programming strategies.

Thanks to computer sciences, these equations and parametrizations have been introduced into computational models with the aim of simulating the physical and chemical processes that occur in different parts of the world. Multiple examples can be provided, but for the specific field of Earth sciences the main examples are: climate models, meteorological models, chemical transport models, ... and emission models like the one presented in this work.

In all the aforementioned cases, the models use as a basis a coordinate system that divides the planet into a 3D grid (Figure 1.1). All the equations, parametrizations and input data defined within the model are executed individually within each grid cell of the model domain. Depending on the application, the modeling domain can have a global, regional or local coverage, as well as higher or lower resolution (i.e. size of the grid cells).

Computational sciences have evolved very fast during the last decades and now, thanks to high performance computing (HPC) strategies, it is possible to produce more accurate results in less time. High-resolution has been identified as one essential element of the development of

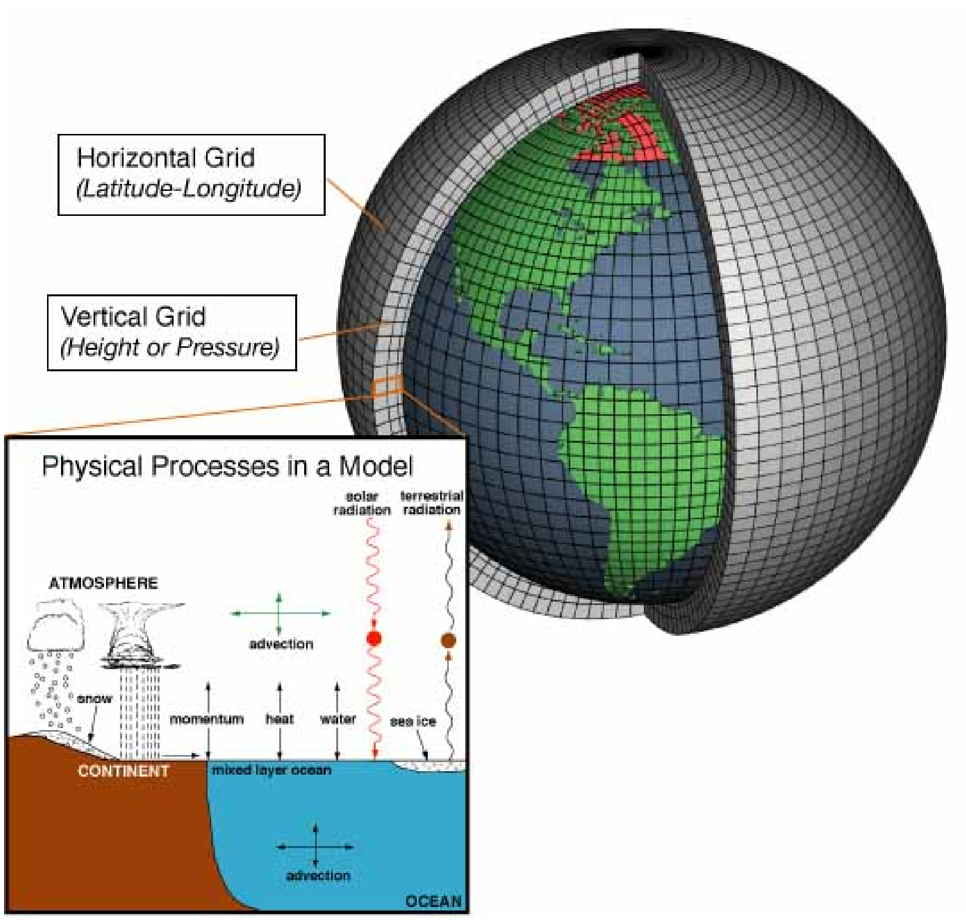

Figure 1.1: 3D grid representation of an earth system model with a detail from the column physics process [28]

General Circulation Models (GCM) to reproduce key climate process with higher fidelity than the conventional GCMs.

One of the most used parallelization strategies in the field of Earth Sciences is the domain decomposition. It follows the rule 'divide and conquer' to minimize the problem into small problems to reduce the amount of data needed to solve the problem.

The goal of obtaining a high resolution product in a short time of calculation is impossible to achieve without using that parallelization strategy in a flexible way.

Traditionally Earth models have been developed using C, C++ or FORTRAN programming languages. Python have appear in 1991 and uses the Object-Oriented Programming language, in an abstract way. Programs developed with Python makes readable and easy to understand the steps that the program is doing. Python can also use different tools, developed as libraries, that allow to do complex tasks in an easier way like apply machine learning algorithms, mathematical operations, geospatial operations, among others.

Taking benefit from all the points presented before, it has been developed the High-Elective

Resolution Modeling Emission System version 3 (HERMESv3) Bottom-Up module that will be explained and evaluated, from a computational point of view, in this document.

HERMESv3 is an open-source Python based software that can estimate the emissions of a region using different python libraries, some of the strong pros of HERMESv3 is their flexibility on customization and the geo-processing of the data into as high resolution as demanded.

All the developments of that work have been possible thanks to the whole Earth Science department of the Barcelona Supercomputing Center taking benefit of all the expertise knowledge acquired during the years.

In addition of the computational description of the HERMESv3 software, the air pollution problem and how is it modelled have to be explained first. To have a detailed description of an air quality modeling system each model involved listed and briefly explained making emphasis on the emission models.

1.2 Earth Science department of Barcelona Supercomputing Center

The Barcelona Supercomputing Center [8] is a public institution with multidisciplinary research teams as well as the host of MareNostrum [10] supercomputer.

BSC is divided into research departments like Computer Sciences, Life Sciences, Computer Applications in Science and Engineering and Earth Science. Each one of these groups explore different research lines but with lot of synergy between them.

Earth Science department, lead by Francisco J Doblas Reyes, has the mission to develop, improve and implement global and regional models for air quality and climate forecasting and their applications. The whole group is divided into 4 teams: Climate prediction, Atmospheric Composition, Earth System Services and Computational Earth Science [9] which I belong to.

"Climate Prediction group aims to develop regional and global climate prediction capability for time scales ranging from a few weeks to a few decades into the future (sub-seasonal to decadal climate prediction)."

"Atmospheric Composition group aims to better understand and predict the spatiotemporal variations of atmospheric pollutants along with their effects upon air quality, weather and climate."

"Earth System Services group facilitates knowledge and technology transfer of state-of-the-art research at local, national and international levels."

"Computational Earth Sciences group is a multidisciplinary team with different IT profiles that interacts closely with all the other groups of the department. The group provides help and guidance to scientists on the technical issues relating to their work and develops a framework for the most efficient use of HPC resources.”

Since 2007, Earth Science Department are providing daily air quality forecast using the CALIOPE air quality modeling system [5]. The next sections explains what an air quality modeling system is and the main components, but first we are going to introduce CALIOPE.

1.2.1 CALIOPE

CALIOPE [5] is an operation air quality forecasting system developed and maintained by the Earth Sciences department of the Barcelona Supercomputing Center. The system operationally provides forecast of air quality for the next 24/48 hours for Europe, Spain and Catalonia at resolutions 27x27 km, 4x4 km and 1x1 km respectively in their own web page [36].

CALIOPE is an air quality modeling system focused to study and forecasting the air pollution on the atmosphere. It is a compound of different models that some of them are widely use by the community like WRF [40] (weather) or CMAQ [29] (chemistry) and the other ones are in house developments like DREAM [32], [6] (dust) and HERMES (emissions) (extended explanation on section 1.4).

Other software that are still under development is coupling the meteorological Nonhydrostatic Multi-scale Model on the B-grid (NMMB [31]) with a chemistry module in order to obtain benefits from the feedback of the both modules. That coupled model called MONARCH [1] will substitute WRF and CMAQ models in the CALIOPE system, but it is still on the testing phase.

This work is focused on the HERMES emission model. HERMES has evolve during the years and taken benefit from all the knowledge acquired with the previous versions HERMESv3 has born. But before to go to the HERMESv3 details the next sections are going to introduce the framework where HERMESv3 is involved.

1.3 Air pollution

Air pollution is currently one of the most serious environmental and health risks (see World Health Organization publication [30]).

The problem of air pollution is driven by a combination of pollutants released from multiple anthropogenic and natural activities, including road transport, manufacturing industries or forest fires, among others.

Most governments have sign agreements, (Paris 2015 [44], Kyoto 2009 [12] and Long-range Transboundary Air Pollution 1979 [43]) to reduce the impact of the emissions that their countries are releasing to the atmosphere. They are applying, or studding, different mitigation strategies to reduce the amount of emitted pollutants for certain areas.

These emission policies should be evaluated in advance, using air quality systems to assess their effectiveness. To make these types of studies, a flexible tool to compute the release of primary pollutants and emission scenarios is needed in combination with the rest of components of an air quality modeling system.

Air quality modeling systems are also used to provide daily forecasts, like CALIOPE [36] for Spain, to show the people how is the status of the air quality.

1.4 Air quality modeling system

An air quality modeling system is a combination and integration of multiple numerical models.

Some institutions are working on providing air quality forecasts simulations, like CALIOPE (see 1.2.1) or COPERNICUS [14], using different methodologies and software.

The Copernicus Atmosphere Monitoring Service (CAMS) is a service implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF), that provides continuous data and information on atmospheric composition. CAMS is part of the Copernicus earth observation program call, which is managed and coordinated by the European Commission

The main numerical models in air quality are the meteorological model, the emission model and the chemical transport model (CTM) to simulate the release, transport, transformation and deposition of air pollutants. Figure 1.2 sows the workflow and needs of a generic air quality modeling system.

All the models used in the air quality modeling system have to work using the same domain definition. That definition comes from a grid with specific parameters like projection, resolution and spatial coverage.

Meteorological models, like WRF [40], are commonly used for many purposes and are widely maintained by many institutions. That has allow the models to grow and improve the forecast accuracy by reducing the uncertainty of their products over the software time evolution.

In an air quality modeling system, the meteorological outputs are used as input for the chemical transport model as well as the emission model.

Chemical transport models (CTM), like CMAQ [29], need to be fed by the meteorological forecast as well as the emissions estimation in order to simulate concentrations of gases and aerosols toxic compounds, as well as their deposition.

Within this modeling chain, emission models are know to be the largest contributor to the overall uncertainty of the air quality system results. This uncertainty is related to multiple factors, including the input data (i.e. activity and emission factors) and estimation approaches

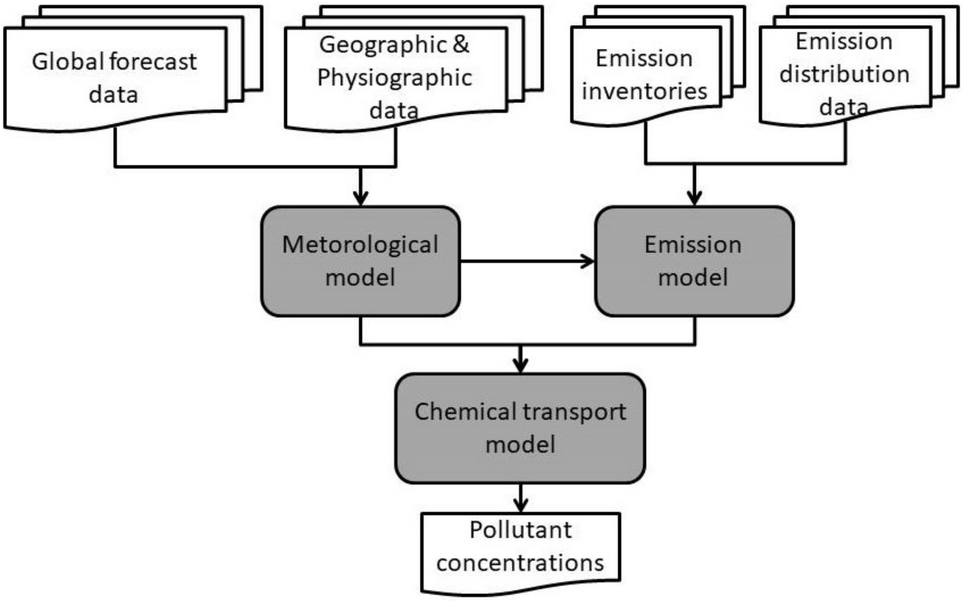

Figure 1.2: Air quality modeling system

used to compute the emissions. To reduce the uncertainty, the best emission estimation has to fed the chemical transport model. That's why the emission model is the key element of an air quality modeling system.

1.5 Emission models

An emission model is a tool to calculate the amount of emissions released by a certain source or activity in a specific location and time.

There are two calculation approaches to obtain the emission product: top-down and bottom up. The used methodology depends mainly on the data availability. A brief definition of that both approaches is presented in the next section ass well as the needs and requirements that the CALIOPE emission model have to accomplish.

1.5.1 Bottom-up and Top-down approaches

As we can see in the Figure 1.3 both methods are used to obtain similar results.

Top-Down approaches use the already emission inventories reported by the administrations and tries to increase the resolution using spatial proxies to distribute that emissions horizontally.

CAMS products [16] are trying to improve the emission inventories. Their resolution products are 1/10°x 1/20°(longitude latitude, ∼ 6x6 km over central Europe) but it isn't enough for a high resolution domain (up to 1x1 km) or even street level.

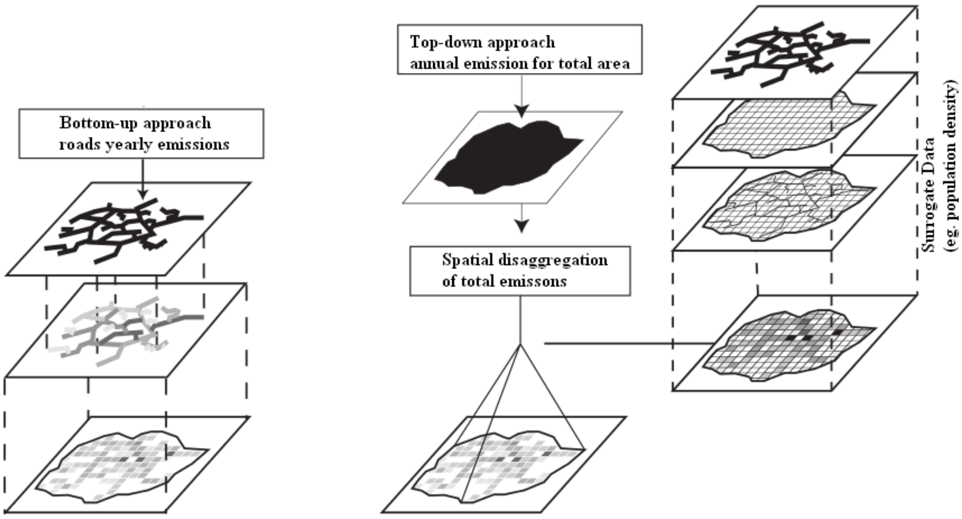

Figure 1.3: Top-Down and Bottom-Up emission method approaches [34]

With that approach is omitted the local activity data of the region and is not differentiating between the different source types inside the same the emission sectors.

Some of the most used emission models that follow the top-down approach are SMOKE, [2] and HEMCO [25].

Bottom-Up approaches need lots of geo-location, statistical and meteorological input data to get the source detailed activity scenario. That scenario is needed to calculate one by one of the source contribution for the region of the study.

To make very detailed emission scenarios is needed to have the local activity data as detailed as possible to make flexibility on the customization of that emissions. A high resolution product or scenario must use a Bottom-up approach with the most accurate geolocalized data.

Some of them are the VOLT'AIR model [26] for the agricultural crop fertilizer, the VEIN model [23], which provides bottom-up exhaust and evaporative road transport emissions, or the MetVed model [17] for estimating residential wood combustion emissions.

1.5.2 Emission source sectors

The emission calculation can be summarized as an emission factor EF multiplied by an activity factor AF and the temporal factor TF of the date time to simulate.

$$E = EF * AF * TF \tag{1.1}$$

The EF depends on the evaluated source (e.g. certain type of gasoline vehicle, fertilized applied on a crop, type of cattle, ...) and the TF is a compound of statistical data to get the hourly factor for the time series evaluated.

But the AF depends on the quantity of that source are appearing on the area of study, grid cell, as well as meteorological data in some cases, and it have to be calculated over the source shape.

Not all of the different types of emissions follows the previous formula and to differentiate the numerical process calculation for each source the emission types have been grouped into different emission sectors suggested by the EMEP/EEA 2016 report.

1.5.2.1 Point sources

Energy and manufacturing industrial point sources are estimated in that sector.

Thanks to the georeferenced multi-point shapefile with the needed information each emissions is horizontal allocated to the nearest grid cell of the working domain.

Using meteorological data it is able to vertically distribute the emissions along the vertical levels applying the plume rise approach [20].

1.5.2.2 Road transport

Most of the road transport emissions are calculated on the road link level for each one of the vehicle categories.

Average traffic flow, mean vehicle speed and vehicle fleet composition are part of the information provided in that multi-line shapefile with the road links.

Exhaust emissions

Exhaust emissions are the ones derived from the use of the vehicle engine during the circulation.

Hot-exhaust emissions are very different between vehicle type and must be calculated by vehicle category but excluding the electrical and hybrid ones.

The cold-start emissions are the ones derived from the first minutes of the engine start and are dependent from the outdoor temperature.

Non exhaust

The non exhaust emissions are the particulate matter derived from the abrasion process of the road-surface, tyre and brake wear and resuspension. That emissions comes from each type of vehicles including the non-exhaust emitting ones.

The resuspension also depends on the state of the road-surface. A wet road-surface do not allow to raise particulate matter as well as a dry one so can be used the precipitation to calculate that wet factor for the resuspension emissions.

Gasoline evaporation

Gasoline evaporation emissions are not on the road link level, that are the emissions derived from the constant evaporation of the gasoline during the parking time. That emissions are distributed over the population area [38] that is normally the place where the vehicles are parked.

That type of emissions are also meteorological dependent by the outdoor temperature.

1.5.2.3 Agriculture

Agricultural emissions are the ones derived by the animal husbandry and the cultivation process.

Livestock

NH_3 emissions derived from the manure management activities are calculated on the livestock sector.

Each one of the animal types (coats, cattle, pigs, chickens, ...),as well as their subcategories, have different statistical involvement on that emissions so must be calculated by any animal subtype category.

Emissions have to be also distributed over the amount of animals [15] by destination grid cell.

Fertilizers application

The coming emissions from the cultivation process of fertilization are calculated on that sector.

That emissions are divided by all the fertilizers products but applied over the crop distribution area.

Crop distribution area are not coming from any source but applying different mappings between land uses [39] and crop area statistics by NUTS level 2 is obtained the crop distribution.

Crop operations

As well as the fertilizers applications, crop operation emissions also needs the crop distribution to horizontal distribute the particulate matter released during soil cultivation and crop harvesting activities.

The temporal distribution, in that case, uses soil cultivation and harvesting calendars [37].

1.5.2.4 Residential and commercial combustion

Emissions from the combustion of natural gas, liquefied petroleum gas (LPG), heating diesel oil, wood and coal derived from residential and commercial small combustion plants are estimated in that emission sector.

Residential and commercial combustion sector is also using the population distribution [38] to allocate the amount of fuel used by destination grid cell in order to do the calculation by fuel type.

The usage of the fuel to heating building temperature is influenced with the outdoor meteorological temperature and the Heating Degree Day indicator [35] is used as a proxy variable to temporal distribute the emissions.

1.5.2.5 Other mobile sources

Exists other mobile sources, not depending on the road link level, such as the shipping and recreational boats emissions, the agricultural machinery as well as the emissions produced on the airport area (i.e. cruiser, ferry, general cargo, ...).

Shipping in port areas

The emission over the port area are derived from the mail engine and auxiliary engine during the manoeuvring and hoteling operations of the different vessels.

That emissions are calculated by engine category and operation phase performed over two different shapefiles one by phase.

A specific weight is also applicable to the different docks as a function of its usage to get a better distribution.

Recreational boats

Pleasure boat activities emissions are also calculated by vessel type along the coast line.

To the author's knowledge, no other national emission inventory is reporting that source type.

The spatial proxy of that sector is based in the location of each recreational dock and a simulated distribution depending on the distance to the stationary dock.

Agricultural machinery

Tractors and harvesters emissions are also occurring during the agricultural operations over the crop distribution area.

That emission are calculated in that sector differentiating between phase and agricultural equipment type.

Landing and take-off cycles in airports

Aircraft landing and take-off (LTO) emission occurring at airports are calculated depending on the plane engine and phase. In that sector are including the approach, landing, taxi-in, post taxi-in, pre taxi-out, taxi-out, take-off and climb-out phases over each airport.

The climb-out and approach phases of the planes are distributed vertically based on a extension over the altitude of the airport roadways.

1.5.3 Emission model requirements

The emission model have to manage the preparation and execution of all the emission sectors dealing with all the different input data and methodologies for the emission sources calculation.

It also must to accomplish the following requirements

1.5.3.1 Flexible and usable

The main requirement is to be flexible in order to obtain customized emission scenarios using different approaches over the emission sectors.

It has to be user-friendly and highly configurable with understandable configuration options.

1.5.3.2 Geospatial input data

The calculation part of each sector must be executed on the source level to do not lose geo-localized precision.

The model must also includes several functionalities for automatically manipulating and performing spatial operations on geo-referenced objects (shapefiles and raster files) performing different post-GIS operations.

Once the emission is calculated on each source shape them have to be dumped on the high custom definition grid to write an unique emission file.

1.5.3.3 User defined domain

The user can define multitudes of regions of study with different resolution and map projections.

The emission model has to be able to work using the most common used map projections like regular lat-lon, Lambert Conformal Conic, mercator and rotated pole projections.

1.5.3.4 Multiple output formats

The emission output files provided by the emission model must follow the formatting requirements of multiple chemical transport model.

CMAQ [29], WRF-Chem [33], MONARCH [1] and R-LINE [41] (to go throw street level dispersion modeling [7]) are the models that will need as input the emission estimation as a file and it have to accomplish the formatting requirements of the different CTMs.

1.5.3.5 Working environment

The emission model has to run on a HPC environment such as MareNostrum 4 [10].

The general-purpose block of MareNostrum4 consists of 48 racks housing 3456 nodes with a grand total of 165,888 processor cores and 390 Terabytes of main memory. Compute nodes are equipped with (see Table 1.1).

Those computational resources have to be used in the most possible efficient way. There are multiple metrics to evaluate the efficiency of a parallel program. The study of that metrics are in the results chapter 3 of that document.

So the scattering of the resources between the different sectors have to be optimal.

2 sockets Intel Xeon Platinum 8160 CPU with 24 cores each @ 2.10GHz for a total of **48 cores per node** L1d 32K; L1i cache 32K; L2 cache 1024K; L3 cache 33792K
96 GB of main memory 1.880 GB/core, 12x 8GB 2667Mhz DIMM (216 nodes high memory, 10368 coreswith 7.928 GB/core) 100 Gbit/s Intel Omni-Path HFI Silicon 100 Series PCI-E adapter
10 Gbit Ethernet
200 GB local SSD available as temporary storage during jobs

Table 1.1: Compute nodes characteristics of MareNostrum 4

1.5.3.6 Execution time limit

Daily air quality 48 hours forecast is also needed and the whole process have to end before the publishing date time because users need the data in a certain time limit. So the portion of time devoted to the emission model mustn't overpass the 15 minutes limitation.

Due to that restriction the computations and the input/output process have to be performed in the most optimized way.

1.6 HERMESv3

The High-Elective Resolution Modeling Emission System version 3 (HERMESv3) estimates atmospheric emissions for use in multiple air quality models. HERMESv3 is composed of two independent modules named global-regional (HERMESv3_GR [22]) and bottom-up (HERMESv3_BU [21]).

During 2005 and 2006, the first version of HERMES [4] was developed for the purpose to be use in CALIOPE (1.2.1) by the Atmospheric Composition group of the BSC. HERMESv2 [19] introduces the improved emission estimation methodologies and data until 2013. HERMES-Mex [18] is the first emission modeling system fully written in Python developed by BSC for the Mexico City's Secretariat of the Environment (SEDEMA).

Thanks to the previous versions, HERMESv3 has been developed as the official emission modeling system of the Earth science department.It is an open-source and objected oriented python system prepared to run into a high computing environment divided into two modules:

The Global/Regional module combines existing gridded inventories with user-defined vertical, temporal and speciation profiles. That module uses mostly Top-Down approaches to obtain the emission file for the CTM.

That model have the purpose to be used for global domains or low resolution domains. It's also useful to take benefit from that module for the locations, inside the studied region, that we do not have as detailed information as the Bottom-Up module needs. An example of that usage is to use the Global/Regional module for the frontier countries of the country to be analyzed and the Bottom-Up module for the inside country emissions.

The Bottom-Up module combines scientific estimation strategies based on (but not limited to) the calculation methodologies reported by the European EMEP/EEA air pollutant emission inventory guidebook.

HERMESv3 Bottom-Up combines meteorological data and emission factors along with local data. That local data is not restricted by the country, it can be used for any European region. That local data is formed by statistical usage data and geo-localized data with many inputs types of GIS (Geographic Information System) formats such as rasters or shapefiles.

As have been explained before, is needed to obtain an emission estimation into the highest possible resolution. So this document is focused on the deep analysis, from a computational point of view, of the Bottom-Up implementation.

In a previous work of HERMESv3 have been developed individual programs that can simulate, using Bottom-Up approaches, the different emission sector work in order to obtain individual emission files by source.

Table 1.2 is summarizing all the developed sectors in HERMESv3_BU and all the input data needed to do the calculations.

This work is presenting the managing of that emission sector estimators with an engine able to scatter the computational resources using Multi-Program Multiple-Data (MPMD) methodologies.

To make that analysis is needed first to evaluate their needs, special operations and components.

1.6.1 Sector classification

To get the most accuracy result, the emissions have to be calculated directly by source and all the sectors explained before have been analysed depending on their shapes to be grouped into four emission source types. Table 1.3 is showing that classification.

1.6.1.1 Point source emissions

The point sources emissions are the ones that can be allocated into a single coordinates, like a chimney.

That emissions can be easily distributed into the grid cell summing all the point sources that are allocated into the same cell polygon.

HERMESv3_BU can also distribute them vertically using two approaches. Multiplying the

Sector	User-dependent files	Built-in files	External files
Point sources	• Point sources shapefile • Temporal profiles CSV files • Speciation profiles CSV files • Meteorological files (only if plume rise is activated) (hourly 4D temperature, 4D U/V wind components, PBL height, Obukhov length, friction velocity) (1)		
Road transport	• Road network shapefile • Temporal profiles CSV files • Fleet composition profiles CSV file	• Emission factors CSV files • Speciation profiles CSV files	• ERA5 meteorological files (hourly 2-m temperature and precipitation) (2)
Residential & commercial combustion	• Energy consumption at NUTS3 CSV file	• Emission factors CSV files • Temporal profiles CSV files • Speciation profiles CSV files	• JRC global human settlement population grid • JRC global human settlement city model grid • ERA5 meteorological files (daily 2-m temperature) (2)
Shipping in ports	• Hoteling & manoeuvring shapefiles • Vessel's operations CSV file	• Emission factors CSV files • Vessel's technology CSV file • Load factor CSV file	
Aviation (LTO)	• Airports, runways and air trajectories shapefiles • Plane operations CSV file • Temporal profiles CSV files	• Emission factors CSV files • Speciation profiles CSV files	
Recreational boats	• Recreational boat units, load factor, working hours, nominal engine power CSV files • Spatial distribution raster file	• Emission factors CSV file	
Livestock	• Livestock split and adjusting factors CSV file	• Emission factors CSV files • Speciation profiles CSV files	• FAO gridded livestock of the world version 3 • ERA5 meteorological files (daily 2-m temperature and 10-m wind speed) (2)
Agricultural crop operations		• Emission factors CSV file • Temporal profiles CSV files • Speciation profiles CSV files	• CORINE Land Cover land uses
Agricultural machinery	• Equipment units and nominal engine power and working hours CSV file	• Emission factors CSV files • Deterioration factors CSV file • Temporal profiles CSV files • Speciation profiles CSV files	• CORINE Land Cover land uses
Agricultural fertilizers	• Fertilizer rate CSV file • Crop calendar CSV file • Ration of cultivated to fertilised area CSV file • Share of fertilizer type per crop CSV file	• Fertilizer related emission factor parameter CSV file • Temporal profiles CSV file • Speciation profiles CSV file	• ISRIC soil pH and CEC data • CORINE Land Cover land uses • ERA5 meteorological files (daily 2-m temperature and 10-m wind speed) (2)
(1) These meteorological parameters are not provided by ERA5 and therefore need to be derived from other models (2) ERA5 is proposed since it is open data. Users can alternatively use the outputs from other meteorological models			

Table 1.2: Classification of HERMESv3_BU input data files per sectors [21]

height of the chimney by a factor or using a Plume rise approach [20] that allocates the emissions vertically depending on meteorological data.

Sector	Point	Area	Gridded	Line
Point sources	X			
Road Transport			X	X
Residential and commercial combustion			X	
Shipping ports		X		
Aviation (LTO)		X		
Recreational boats			X	
Livestock			X	
Agricultural crop operations			X	
Agricultural machinery			X	
Agricultural fertilizers			X	

Table 1.3: Sector source types

1.6.1.2 Area source emissions

The area source emissions are the ones that have to be distributed over the region that occupies that sector.

For example, airports or shipping ports have a defined polygon with a different shape. That emissions are calculated on the full shape of the source and them are weighted distributed by the occupied area on each grid cell.

1.6.1.3 Gridded source emissions

The Gridded source emission types are the ones that depends on other input data type like rasters.

That rasters may contain land uses, population, crop types, or other types of relevant data. The same emission sector can use different raster information but some simplified examples are the agriculture sector using crop types or residential combustion using population.

The way to calculate that emissions are directly in the grid cell. To get the necessary information, all the needed data has to be mapped from the raster resolution to the grid one.

1.6.2 Geographic information system (GIS) operations

HERMESv3_BU have post-GIS operations for automatically manipulating and performance spatial operations.

The most important types of them are explained bellow:

1.6.2.1 Read/Write/Create

The principal spatial operation is to read, write or create vector-based spatial data using the ESRI shapefile format.

That operation includes the projection changes between the different input projections to work always using a latitude-longitude projection (WGS84) for the rest of spatial operations.

1.6.2.2 Clip

The clip is used to work only in a portion of your domain.

There are several ways to create the clip. The default one is the boundary polygon that include the input grid; But, if the region of interest is only a portion of the full grid, a shapefile with the polygon can be delivered to HERMESv3_BU to use it or even to create a user-defined polygon one delivering to HERMESv3 several points in the latitude-longitude projection.

That operation overlays the clip with the sector data to extract from it only the data that lies within the area defined by the clip polygon.

1.6.2.3 Allocation

The allocation is the method that map each point sources into the destination grid cell.

That method have to probe that each element is complete inside the polygon that define the grid cell.

1.6.2.4 Joint

The Joint method is the one responsible to mach information between two types of shapes.

For example it can be used matching a shapefile with countries with another shapefile and the resultant data is the second shapefile but with the information of the country that each element belong.

1.6.2.5 Raster

HERMESv3_BU can also transform a Raster into polygons with data in order to be able to use them using the other GIS operations.

1.6.2.6 Weighted distribution

That operation is the responsible to get some source emissions (like area ones) into the grid cell.

That operation calculate the percentage of the source that goes into the cell polygon and distribute them in a weighted way.

The most common strategy is to use the portion of area from the source in the destiny cell. Otherwise the line sources have to use the portion of length to get that weight.

1.6.2.7 Nearest point

To extract the meteorology of a certain area is performance using that operation.

The meteorological data are saved in NetCDF [42] format that distributes the information in a matrix with the location of the centroid of each grid cell with their latitude and longitude.

The nearest point operation are used to calculate the nearest NetCDF cell to each emission source to get from there the meteorological data needed for the calculation.

1.6.3 Workflow

HERMESv3_BU can be divided into tree main blocks, Initialization, Calculation and Writing.

Each one of the tree blocks depends on the previous one as well as some input data to be configured as is shown in the Figure 1.4.

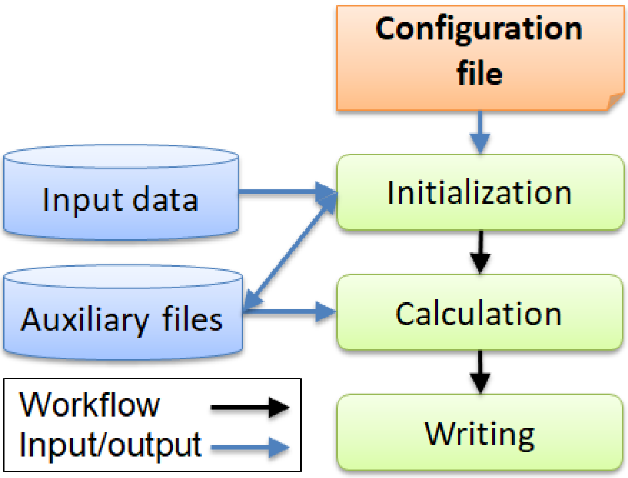

Figure 1.4: HERMESv3 Bottom-Up workflow

1.6.3.1 Initialization

The first thing that need the Initialization is a configuration file where is detailed all the specifications and needs that the simulation have to take into account. Some of the arguments that this configuration file has are the date to simulate, the number of time steps, grid definition type and parameters, as well as all the paths to the different input files needed, rough geo-spatial data, administrative boundaries, meteorological data, emission factors, profiles for the vertical and temporal distribution, speciation profiles, among others.

You can notice in the Figure 1.4 that are a bi-directional arrow between the Initialization block and the Auxiliary files. This mean that the needed auxiliary files will be created or loaded if them already exists.

The first thing that this block does is the creation of the Grid. Depending on the input parameters HERMESv3 is able to generate a grid into as high resolution as demanded following different of projection types as Regular Lat-Lon, Lambert Conformal Conic, Mercator and Rotated pole projection. This grid will be saved as a shapefile into the auxiliary file directory.

Another general thing that the Initialization block does is the generation of the clip. The clip is a polygon where is detailed the region of study to avoid calculation above this area. That clip will be also stored as a shapefile with an unique polygon (or an multi-polygon in the case of contain islands or complex geometries)

Once the Grid and the Clip are created, the Initialization does the most part of the geo-spatial operations. That ones are the most time and resources consuming so them have to be calculated once. With that geo-spatial operations is obtained the spatial proxies with the activity factors (e.g. the amount of each animal type by grid cell)

1.6.3.2 Calculation

The calculation block is the responsible to calculate the different emissions for the selected period as well as to apply the spatial proxies to get the emissions on the corresponding grid cell.

For some of the sectors that need meteorological data that varies from one time step to an other, that block is also getting that needed variables in the grid cell using the nearest meteorological value to each emission source.

That meteorological data is used to calculate the emissions or to differentiate the equations that can variate the formula depending on that variables. E.g. The emission factor for the cold start emissions of the road traffic sector is different between some range of temperature.

The last thing that the calculation core does is the speciation. The speciation is the transformation from the input pollutants to the output ones. E.g. NO and NO2 emission are a fraction from the NOx ones. That profiles may be different between the different sources. E.g. For the biomass burning in the residential combustion the Formaldehyde pollutant is approximately the 0.02% of the total NMVOC (Non-methane volatile organic compounds) while for the rest of materials (e.g. natural gas) is the 0.61%.

1.6.3.3 Writing

The usual way to store the data of an earth model is using the NetCDF[42] files.

A NetCDF file is a type of file that stores gridded information such a matrix with several dimensions. The most common dimensions used to store the date are time, vertical level, Y and X.

A common convention have been developed by the community and improved during the years. That convention is the CF-1.8 convention [13] and HERMESv3 has the capability to write the NetCDF in that way bu several chemical transport model (CTM), like CMAQ (IO API [11]) or WRF-Chem [27], uses their own conventions and HERMESv3 can follow them. That block will also do the unit change if the CTM convention requires it.

Chapter 2

Methodology

This chapters presents the main developments related to the improvement of the computational performance of HERMESv3. Note that the section does not present a detailed description of how the emissions are estimated within HERMES, but it is rather focused on the computational behaviour of the general parallelization of the different sector types 2.1 and how the central process units (CPU) resources are distributed along the portions of the calculation 2.2. This portions divide the total number of data among the number of elements. Additionally, it is explained how HERMESv3 provides and creates the output files 2.3 and all the needed steps to achieve that.

For a complete and detailed description of the input datasets used by HERMESv3 and the equations implemented in the model, the author refers to Guevara et al. (2020) [21].

2.1 Calculation parallelization

As have been explained, as first step, individual programs have been developed to compute each one of the sectors separately. The objective of hat section is is to explain how them have been parallelized.

Each one of the sectors have different needs and different methods to be calculated, so, individual strategies have been implemented for each sector but all of them follow the same rule: Divide and conquer.

HERMESv3 is taking benefit from two parallelization technologies vectored operations and Message Passing Interface (MPI).

Vectored operations allow to do the same instruction for a group of elements by using a single call saving a lot of time. It is easier to explain using an example:

You can perform tree sums one by one (12+4=16, 2+8=10 and 15+5=20), using 3 calls to the instruction of sum, or to make only one call summing two bigger numbers that are containing all of them (012002015 + 004008005 = 016010020).

That technology allows HERMESv3 to do the same operation for a group of data using less instructions.

Taken benefit from the requirement that the emissions have to be calculated on each source type (each road link, each point source, each airport, ...) to do not loose precision, we can avoid communication between the different components during the emission calculation process. MPI allows to split the amount of data without sharing memory between them and it is why that technology have been chosen.

That simplifies the parallelization because it is only needed to spread the sources into the amount of processes dedicated to each sector by distributing the input data as soon as possible.

But there is a big limitation on the area sources. It is not possible to use more processes than the number of sector sources. E.g. For the aviation sector, HERMESv3 is only considering 31 airports so, it is not possible to use more than 31 processes for the aviation sector. That limitation is a design choice due to the sectors that are having that limitation are the fastest ones and only a few process are needed to finish on time.

When the calculation is finished, it is time to start communications in order to write them, but this is explained in the section 2.3.2.

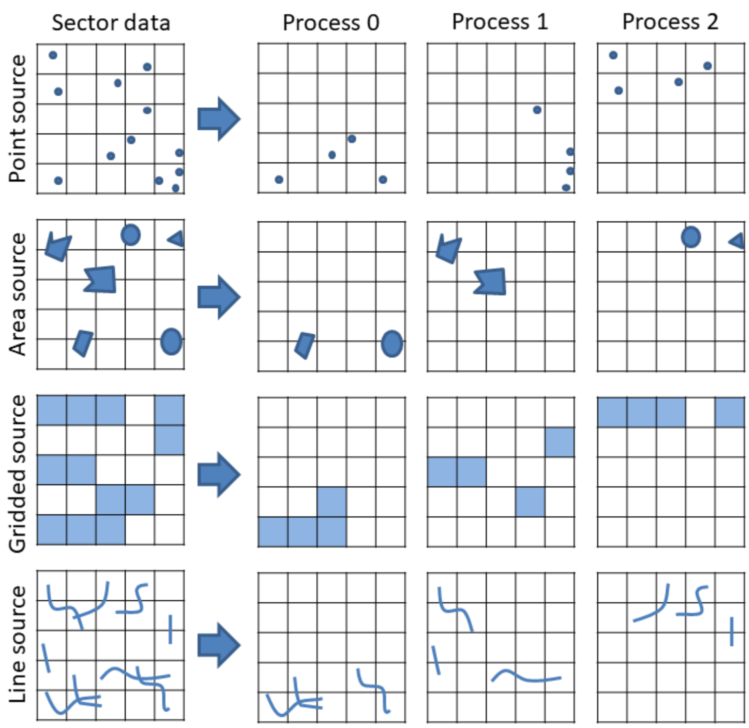

Figure 2.1: Sector parallelization

Figure 2.1 is showing the strategy, in a simplified example, that is simulating 4 different sectors using 12 processes. Each sector belong to each one of the emission source types and each one of them is using 3 different processes.

You can notice how the shapes of each emission source can be distributed to get a sub-portion, in order to divide the number of calculations in a balanced way. How the groups of process are distributed over the sectors are explained in the next section 2.2.

The parallelization of the sectors is done using domain decomposition in the horizontal plane (2 Dimensions), though part of the calculation process could be distributed by the the emissions vertically. When HERMESv3 is selecting the destination of each source type, it is taken the entire column of the grid to distribute them. However, the vertical levels cannot be distributed over the processes because the strategy is focused by source distribution. This is why the vertical distribution of the emissions is not going to be part of this work because we are only focusing on how to speedup the computational time and not how the emissions of each different sector are calculated.

That strategy is similar among the different emission source types but with their own peculiarities and limitations. The differences on the strategies are explained in the subsequent sections.

2.1.1 Point source parallelization

The simplest emission source type is the point source 1.6.1.1. We have the location of all the point sources that can belong to one, and only one, destination grid cell.

When the input data is read, the number of point sources are divided into groups with similar length. Each group will be calculated in parallel by sending each group to the involved process for the calculation of that sector. That distribution have to be in the most balanced way possible in order to ensure load balance.

The most complex part of the calculation of the point source emissions is the vertical allocation. This is not related to the parallelization because when we are selecting where the point source belong, we are selecting the entire vertical column of the destination grid cell.

Just to name, HERMESv3 can use two methods to vertically allocate the emission: multiplying the height of the source (e.g. chimney) by an increment, or using plume rise approaches [20] that need detailed information of the source (disseminated gas velocity, temperature, diameter, ...), as well as the meteorological conditions (air temperature, pressure, ...).

2.1.2 Area source parallelization

As the point source emission types, the first step of that strategy is to scatter all the elements of the sector between the amount of processes available. You can see an example in the second row of the Figure 2.1.

But achieve load balance among processes is not fully possible because each one of the sources have not the same amount of data to calculate. For example, the process that has a big airport like 'Barajas' or 'El Prat' have more plane types and plane operations compared with small ones like 'Burgos' or 'Albacete (that example is shown on the Table 2.1.2) and due to the amount of calculations the time spent on each airport may vary.

That unbalancing on the area source emission sectors is not quite relevant do to that ones are the fastest sectors. A better balancing analysis will waste more time that the spent time due to the unbalancing.

Airport	**2015 operations**
Airport of Barcelona (El Prat)	288,576
Airport of Madrid (Barajas)	366,316
Airport of Albacete	328
Airport of Burgos	929

Table 2.1: Somme airport 2015 operations

2.1.3 Line source parallelization

The road transport sector is distributed using the same approach of the area and point ones (an example is shown on the last row of the Figure 2.1). Due to the problem explained with the airport example, a similar thing makes that the road traffic sector is the most unbalanced one.

The road links are spread between the dedicated processes in order to obtain a similar number of road links in each process. But that spread is done by the internal order of the input data.

Usually, this data is written by zones because it make no sense to save consecutively spatial data that appear far from other one. That means that the congested regions (urban areas) are all close allocated in memory among them and the spread of the sources group them in the same process, opposed, the rural roads are all together.

The road traffic emissions are calculated by each vehicle type and therefore, the congested areas have large vehicle composition. That unbalancing problem happen because the line sources are not calculated by road link, instead of that, each line are duplicated to obtain as many duplicated road links as vehicle types are participating on that road link.

In a first step we have balanced distributing the amount of road links, but when we transform the data from road links to one road link for each vehicle type, the memory and the amount of calculations increased exponentially. Thus, the processes with more urban area road links will spend more time than the ones that have more rural road links.

Lot of work have to be done in order to improve that sector. A deeper analysis of the different strategies that can be applied are described in the 4.2.4 section.

2.1.4 Gridded source parallelization

The strategy developed for the gridded source emissions is divided into two sections.

Part of the input data of that sector are georeferenced data using the raster format. Each raster has his own projection and resolution containing global data. The first part is to transform the raster data into shapes in order to be used. To achieve that is important to screen the data to get an small raster file containing only the data that is inside the working domain. That operation is done only by the master process by using the clip GIS operation 1.6.2.2. When the clipped raster is done, each process involved on that sector read a sub-domain of that data and transform the raster data into polygons.

The second part of that goal is to obtain the information on the grid cell. So, each process with the sub-domain of the raster data, map the data into the grid cells that are intersecting with by using the weighted distribution GIS Operation 1.6.2.6.

HERMESv3_BU calculates that emissions by grid cell, but, a cell can have information from different raster polygons and them can be in different processes. An unique process, the master one, have to gather all the data and eliminate the duplicated cells by summing them.

Once the master process has the unified data without duplicated cells, the data is scattered among the different processes in a balance way using a similar length for subgroups of row-major consecutive cells. The reason to make the subgroups in row-major order is to avoid extra communication during the data gathering (2.3.2 section).

That second part is shown, in an example, on the third row of the Figure 2.1.

2.2 Sector manager

Once we were able to run each individual sector separately in a parallel way is needed to run all of them simultaneously.

The objective is to be able to run each one of the emission calculation sector program into an unique computational job spreading the resources among them. To achieve that a Multiple Programs, Multiple Data (MPMD) strategy was implemented.

The common way to apply a MPMD strategy is while launching the job and it is possible using C or FORTRAN programs but Python doesn't allow to do that in that way. Like HERMESv3_BU is fully written in Python, it has been created a sector manager for the purpose to apply the MPMD strategy by distributing the amount of processes available during the running time. Due to that, the amount of processes that have to be devoted to each sector must be provided in advance to that sector manager using the configuration file.

As it has been explained, HERMESv3 uses the MPI technology to distribute the work. The way that MPI controls each one of the processes and all the communication between them is by using the MPI communicator object. That MPI communicator describes each one of the processes and is the tool to manage them or to perform the communications.

The most important operations that a communicator can do is to control a piece of code that only one process have to do, to put a barrier that all of the processes have to wait until all of them have arrived to the barrier as well as the communications properly like send or receive messages.

With the information of the amount of processes devoted to each sector, the sector manager takes the MPI Global communicator and split them into as many pieces as sectors with the provided dimensions 2.2. Each one of the sectors takes their own splitted communicator in order to use that small one instead of the MPI global one.

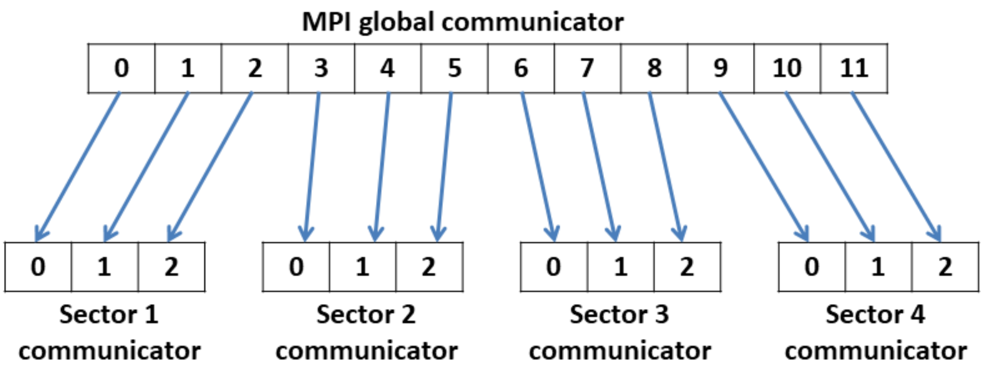

Figure 2.2: MPI split communicator

Each small program that calculate each emission sector has the ability to check if a communicator is given to be used or, if it is not given, to use the global MPI communicator to perform the calculation block.

That strategy allows to run different programs, without communication between them, in the same computational job by spreading the resources along the different programs using the split communicator. Otherwise, when it is needed the communication between the different sectors (e.g. writing part 2.3.3), can be used the global MPI communicator.

A particular case occurs for the sectors that are sharing operations like the agricultural ones.

All the sectors that depends on the crop distribution (crop operations, fertilizers and agricultural machinery) are getting two types of communicators, the sector one properly and a bigger common one containing the total amount of processes devoted to each individual sector.

Firstly the common big part to calculate the crop distribution is calculated using the bigger communicator. When it is done, each small communicator takes the whole crop distribution, divide it in sub-domains and spread them along the processes involved in that sector.

2.3 Writing

The last, but not least, block of HERMESv3_BU is the writing part.

The writing block is composed by two blocks: The data gathering section 2.3.2 and the properly writing section 2.3.3; but before starting to gather and write the emissions have to be placed on the grid cells.

2.3.1 Emissions on cell

Prior to start the writing block, each source shape have to be stored into the corresponding grid polygons.

When the emission calculation is finished, each process have the emissions calculated for a sample of source shapes. As it is has been explained in the 2.1 section, each process has the same type of emission sources, but the only ones that are already mapped into the grid are the gridded source ones (section 1.6.1.3).

All the rest of sources have to be mapped to the grid too. Each process applies the correspondent weighted distribution 1.6.2.6 to get them on the grid cell.

Figure 2.3 is showing the four types of emission sources in the columns differentiating with colors the sources that are allocated in the same process. Since each one of the colors means the process data, the example of the figure is showing that each sector has 3 processes having a total of 12 processes for the complete simulation.

May happen, that some near sources are sharing the same destination cell therefore each processor have to sum the emissions that occur in the same grid cell in order to have unique cell emissions on each processor. In the example provided by the Figure 2.3 it happen on the bottom right cell of the point source that have two orange point source inside the same grid cell.

Otherwise is it not important if the same cell appear in different processes (slashed textured

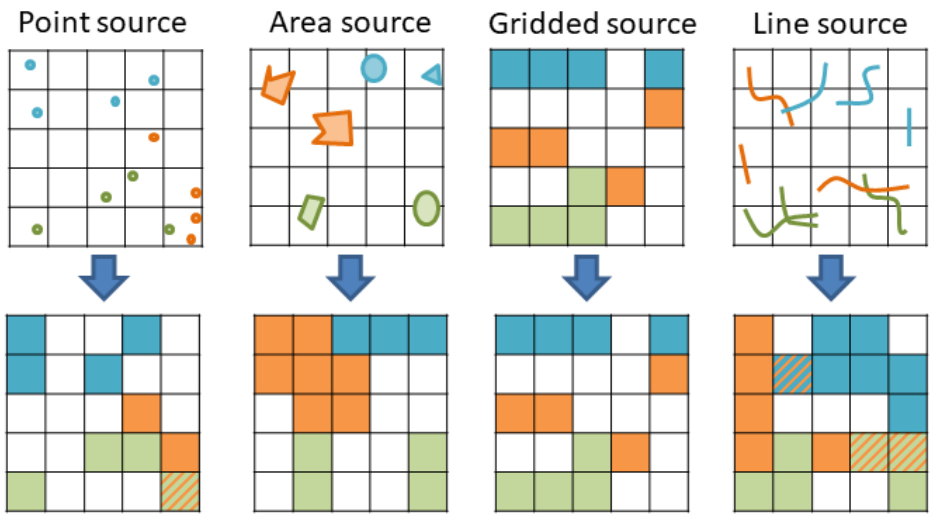

Figure 2.3: Emission source to grid cell

colors on the Figure 2.3), including the processes dedicated to the same sector, and since it is not significant, it will be explained in the next section.

Once the emissions are all of them already in gridded format, HERMESv3 is ready to start the Gathering part.

2.3.2 Gathering data

As it has been explained in the writing step of the HERMESv3_BU workflow 1.6.3.3 the emissions are stored using NetCDF output format. An early adopted functionality of the NetCDF4 writing (01/11/2017) was the capability to write in parallel and that feature is also implemented on HERMESv3.

Due to that feature, many processes can write in the same NetCDF file at the same time. That functionality can be managed by HERMESv3 throw choosing in advance the number of processes devoted to write the NetCDF in the configuration file.

But, before writing, it is needed to gather all the data spread along all the calculation processes. In that point it doesn't matter what kind of emission source it comes from because all of them are already gridded. All the gridded emissions have to be sent to the processes devoted to write and each one of them is the responsible to write an specific region of the NetCDF.

2.3.2.1 NetCDF regions

Python writes the information in row-major order, explained in detail in the next section 2.3.3.1, because of that, each writing process region contains only complete rows of the grid, that regions have to be distributed over the correspondent writing process. That regions are also balanced in the way that all the writing processes have similar number of rows to write.

Two different strategies to balance the writing part can be chosen for the writing block, Both are going to be explained using the example of the Table 2.2.

	Balanced	Unbalanced
Write process 0	3	3
Write process 1	3	3
Write process 2	2	3
Write process 3	2	1

Table 2.2: Writing row distribution over the writing processes

That is an example of a simulation that have 10 rows and we want to spread them into 4 writing processes. The content of the table is the amount of rows that each region is containing. On one hand, we can see how in the balanced way only differs by one the number of total rows devoted on each region. On the other hand, we have the same number of rows in each writing process, except the last one that differs containing the rest of rows.

That unbalanced strategy is used when you want to chunk the NetCDF (explained in the section 2.3.3.2) while the balanced strategy always write a similar number of elements.

2.3.2.2 Writing communication pattern

Once explained the writing regions, the first step is to localize the data of the different regions in order to group a piece of data to be sent to the writing process devoted to that zone. Figure 2.4 shows an example of how the data can be distributed.

When the writing process receives all the emissions is time to sum the ones that occur in the same cell, in order to obtain the total emission by cell. As the emissions are written once the calculation is finished, the writing processes will be some of the same ones of the calculation processes instead of having processes only devoted to the writing part.

The Table 2.3 is showing the communication pattern developed for HERMESv3 in order to reproduce the example of the Figure 2.4. To get a further explanation of the communication pattern we are going to summarize the explanation of the steps that appear on the table:

1. Split the data into as many regions as writing processes have.

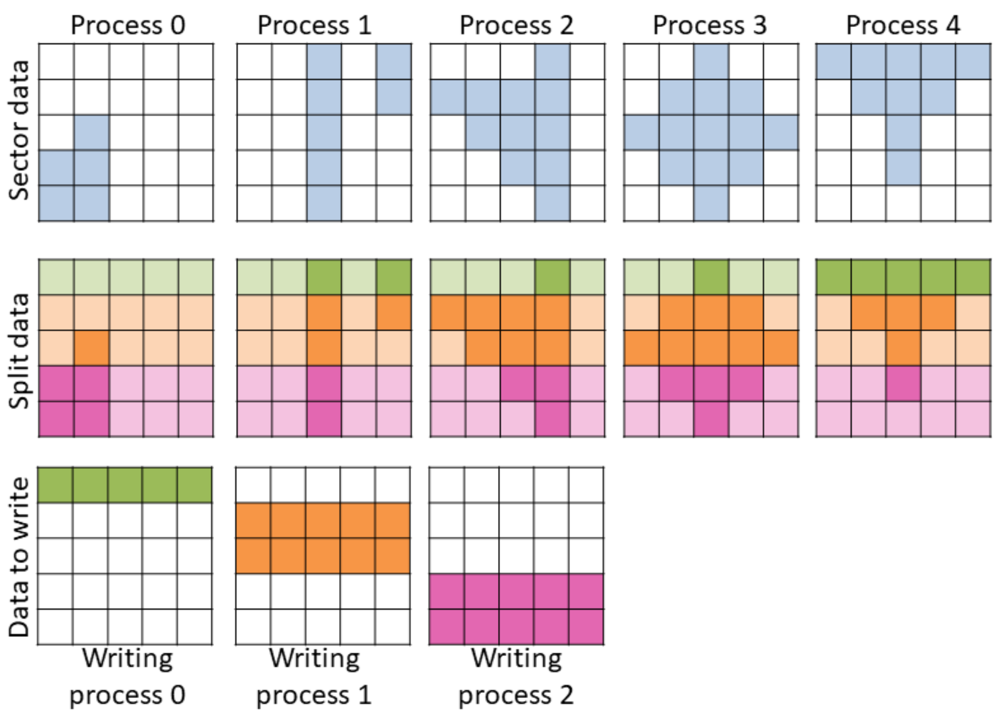

Figure 2.4: Gathering data to write

2. Send the region data to the first writing processor. Notice that the process 0 doesn't have to send the data to itself.

3. The first process receive the data from the other processes.

4. Send the region data to the second writing processor. Notice that the process 1 doesn't have to send the data to itself.

5. The second process receive the data from the other processes.

6. Send the region data to the third writing processor. Notice that the process 2 doesn't have to send the data to itself.

7. The third process receive the data from the other processes.

8. Each writing processor concatenate the data in order to have all the data together.

9. At last, all that concatenated data may contain emissions on duplicated grid cells, for that reason them have to be summed.

All the communications in that section are blocking. The reason to make them blocking is because the messages contains Python objects and they have to be packed to be sent and unpacked when received.

	Process 0	Process 1	Process 2	Process 3	Process 4
step 1	Split by region	Split by region	Split by region	Split by region	Split by region
step 2	Store my region	Send to 0	Send to 0	Send to 0	Send to 0
step 3	Receive 1,2,3,4				
step 4	Send to 1	Store my region	Send to 1	Send to 1	Send to 1
step 5		Receive 0,2,3,4			
step 6	Send to 2	Send to 2	Store my region	Send to 2	Sent to 2
step 7			Receive 0,1,3,4		
step 8	Concatenate	Concatenate	Concatenate		
step 9	Sum duplicated	Sum duplicated	Sum duplicated		

Table 2.3: Writing communication pattern

If non-blocking communications where used, instead of the blocking ones, firstly it has had to pack and send all the messages without proving that they have been received. That packed message is duplicating the memory and if lot of packets are prepared to be sent without receiving them, may we loose someone due to automatic memory release in Python.

Using blocking communications that problem does not occur because the messages are sent and received one by one.

Finally, when all the data is gather, HERMESv3 is ready to start the NetCDF writing part.

2.3.3 Writing

As it has been explained in the 1.6.3.3 section, HERMESv3 can write the NetCDF output using different conventions.

But, the way to write a NetCDF is not convention dependent, so we are going to explain furthermore the principal characteristics that the NetCDF writing have.

A NetCDF, usually, stores matrix information so the data have to contain the full matrix including the non emissions zone having zeros on that corresponding cells. When the emissions are stored in the 4D matrix (time, level, y, x) is time to start the properly writing.

2.3.3.1 Row-Major order writing

A NetCDF can be writing using two approaches depending on the programming language that is used. FORTRAN uses the Col-Major order while C, C++ or Python, for example, are using the Row-Major order.

Both strategies are shown in the Figure 2.5.

The figure showed in the example is following the balanced distribution of the Table 2.2. Where each point means a call to the internal NetCDF instruction to write. When that instruction is call, all the consecutive data in the rows (or columns) are written without any other

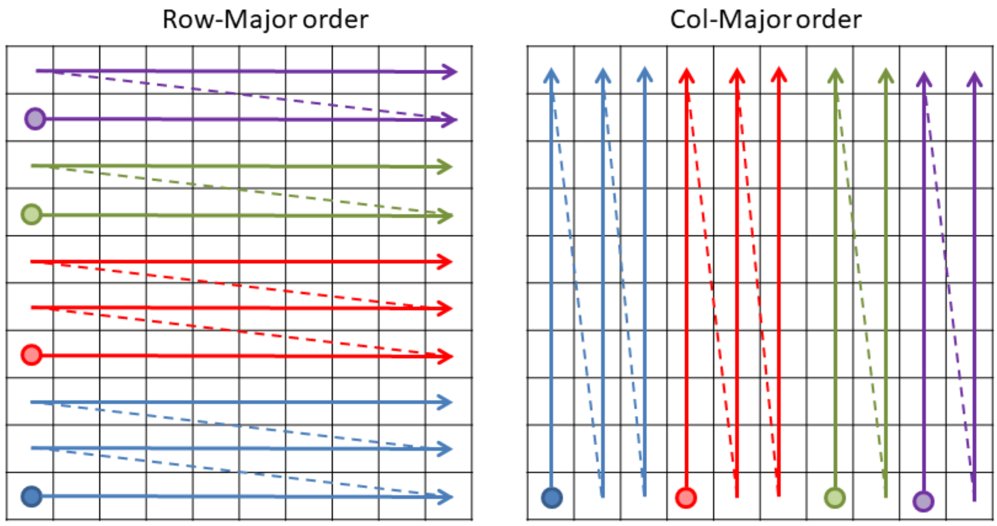

Figure 2.5: Row-Major order vs Col-Major order

call to that write instruction. It is also considered as consecutive data the next row above the last written row so it will be also written using the same call.

That why HERMESv3 are storing the data using complete rows. If HERMESv3 where not following that rule, each separate element will need a NetCDF write call and them are high time consumers.

2.3.3.2 Chunking

When we are talking about NetCDF chunking we have to think that the NetCDF is splitted into sub-domains treated as individual NetCDF files.

The result will be the same with chunking or without it but the internal storage of the data is completely different. To explain further the chunking we are going to use the Row-Major order example, the one that HERMESv3 uses, with help of the Figure 2.6

As it was explained in the previous section, data is internally stored consecutively for each element of the row and the next row is stored next to the previous one. The chunking allows to store consecutively the vertical levels of different sub-domains.

But the chunking has a limitation, these chunks, or pieces of the NetCDF files, must have the same dimensions (only excluding the last chunk). For that reason the unbalanced strategy to scatter the rows shown in the Table 2.2 have to be used in order to have the sub-domains in the chunking format.

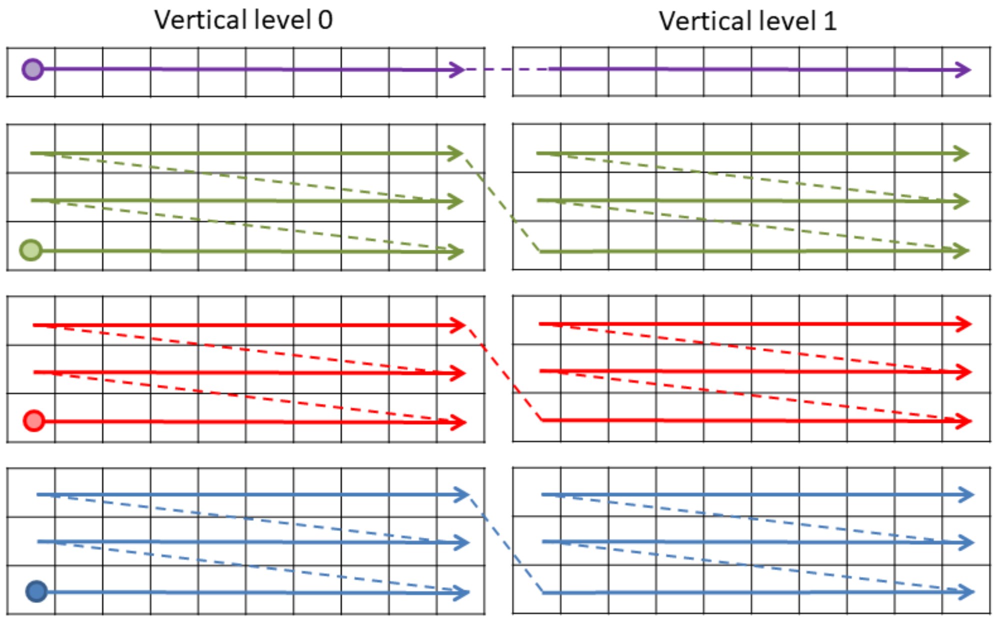

Figure 2.6: NetCDF chunking

2.3.3.3 NetCDF compression

A NetCDF may contain millions of values and they are very storage expensive. Due to that the compression of the NetCDF is allowed in HERMESv3.

That compression is an internal property that you can enable while starting the NetCDF creation, but the writing is very time consuming compared to the non compressed writing. You have to consider if you want the compression because that increment on time does not only occur while writing, it will also appear when you want to read it so, because of that, is an option that you can select on the HERMESv3 configuration file. Using that option you can choose the compression level setting it as 0 for no compression and 9 as the highest value. A normal compression uses the compression level 4.

The main reason to do not choose to activate the compression is when the emissions are only devoted to feed a chemical transport model (CTM) immediately and, once finished, the emission file will be deleted. That always happen on air quality models that you only want to evaluate the air pollution and not the emissions.

But HERMESv3 have a limitation on the NetCDF compression, it is only possible to perform the compression using an unique processor.

Further exploration on that topic have to be done (see 4.2.5).

Chapter 3

Results

To analyse the new parallel implementation, HERMESv3_BU has been executed in a domain that covers the entire Iberian Peninsula with a spatial resolution of 4km by 4km. This domain is currently used by the CALIOPE system [36] (Barcelona Supercomputing Center air quality forecast) to provide operational daily air quality forecast simulations for the next 24 and 48 hours.

Once the domain is defined 3.1 the execution times of each sector using different number of processes have been analyzed 3.2 as well as their speedup 3.3 and efficiency 3.4 to get the best process configuration and evaluate the new implementation 3.5. With that configuration have been analysed the writing options and the best amount of processes devoted to that part 3.6.

3.1 Defining experiment

To test HERMESv3_BU, several simulations have been done. All of them are using the same domain characteristics in order to evaluate the same type of run and the same region (Spain). The chosen domain configuration for testing purpose is the most complex one that has to be used in the CALIOPE daily air quality forecasts.

The study domain covers the entire Iberian peninsula using a Lambert Conformal Conic projection (LCC) at an horizontal spatial resolution of 4km x 4km. The main characteristics of the domain are summarised in Table 3.1. This domain fulfills all the requirements of the section 1.5.3.

The list of pollutants included in the simulations are as follows: CH_4, CO, CO_2, NH_3, NMVOC, NOx, PM_{10}, PM_{25}, PMC and SOx. This list includes the main criteria pollutants as well as green house gases.

To illustrate the emission results obtained with HERMESv3_BU, Figure 3.1 shows the total amount of hourly NOx emitted by the sum of all sectors at 15:00 UTC on November, 29th

Grid resolution	4x4 km
Grid dimensions	397x397 cells
Vertical levels	48
Time resolution	1h
Time steps	48 (two days)

Table 3.1: Iberian peninsula domain characteristics

2016.

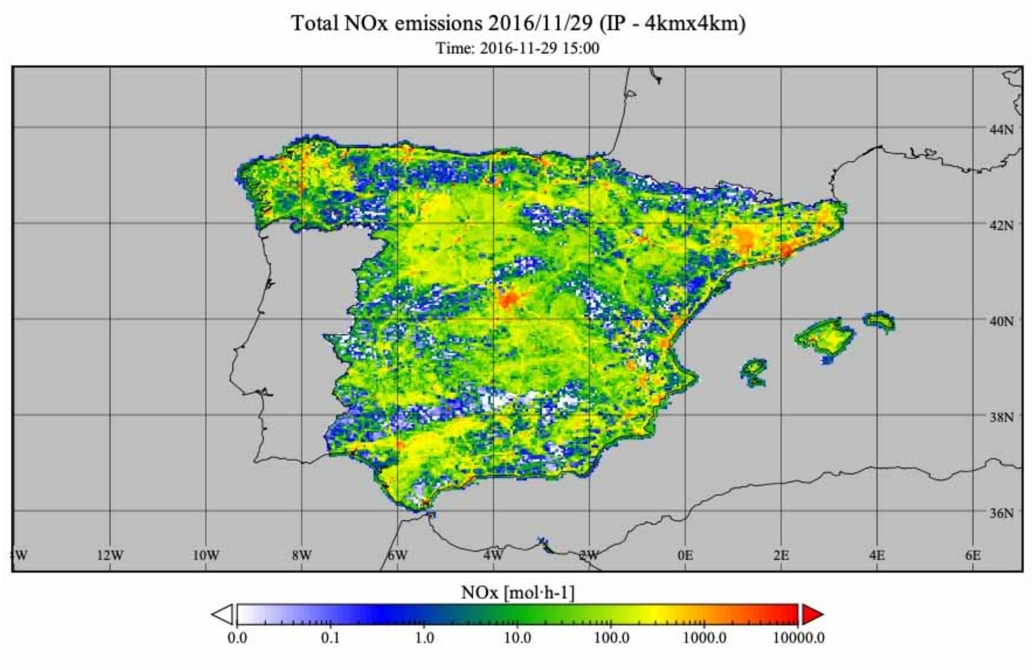

Figure 3.1: HERMESv3_BU simulation on 2016/11/29 for NOx pollutant at 15:00 UTC

It is very clear in the figure that the main contributor to NOx emissions is the urban traffic in the main cities like Madrid or Barcelona and it can be seen also, how well defined are the principal highways that are connecting the different urban areas. Background emissions in agricultural areas are also observed, which correspond to the use of agricultural machinery during this period of the year.

In order to evaluate the efficiency of the new implementation and analyze the overhead introduced by the new parallel approach, individual simulations by sector have been done with different configuration of the number of processes used by each sector, in order to evaluate each sector separately. The main objective is to choose the minimum number of processes that each sector must use in order to finish the execution in less than 15 minutes (following the limit requirement presented in 1.5.3.6.

3.2 Time results

A total of 500 simulations have been performed to test HERMESv3_BU. This number is the results of the 10 HERMESv3_BU available sectors using 10 different number of processes for each sector. Each simulation has been repeated 5 times in order to have a more accurate timings by averaging them and evaluate the variability of the machine, avoiding the outliers times that may occur due to hardware problems.

Figure 3.2 summarises the time results. The Y axes is containing the time in seconds and the X axes the number of processes used in a logarithmic base 2 scale.

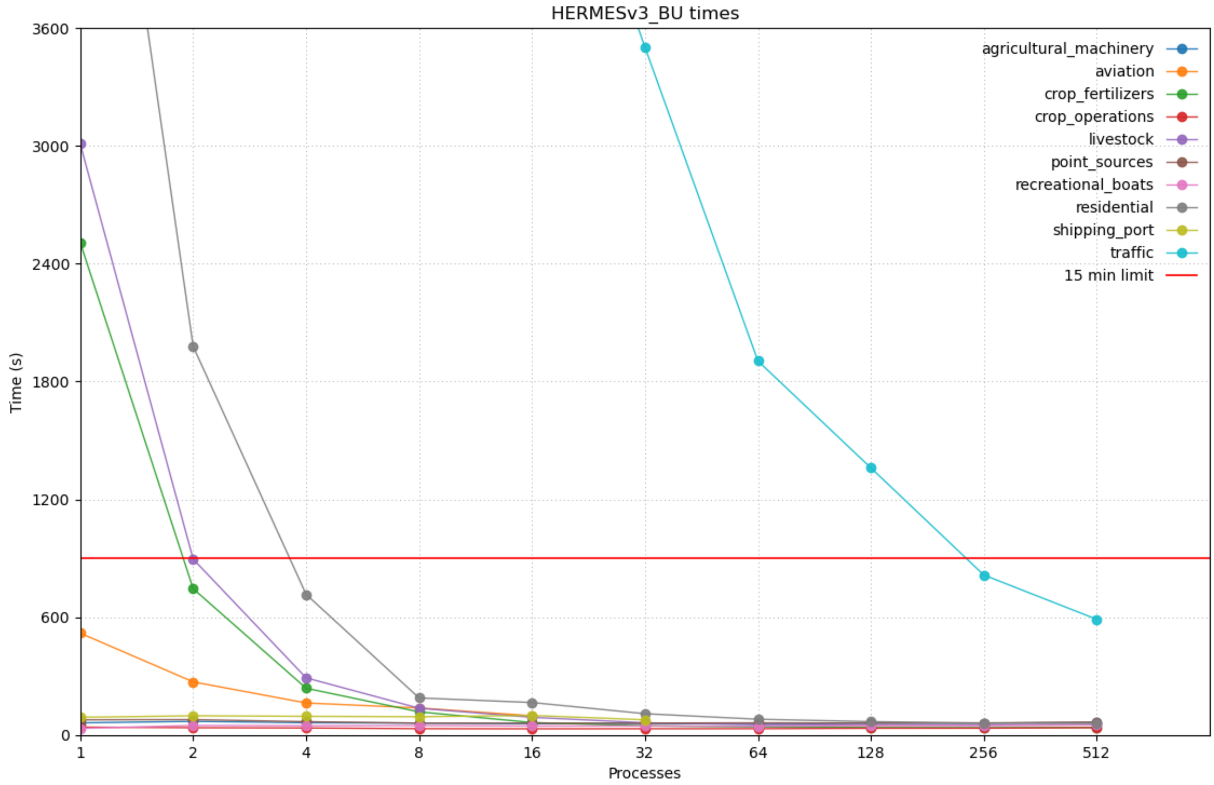

Figure 3.2: HERMESv3_BU times by sector

The red line in the plot highlights the constraint of 15 minutes defined in the requirements section (see 1.5.3.6), and allows identifying for each sector which of the tested configurations are accomplishing it.

As it is observed, most of the sectors are always below the 15 min threshold, except for crop fertilizers, residential combustion, livestock and traffic, which have some configurations exceeding this time limit (e.g. traffic sector when using 128 processes). This analysis emphasises the study of that sectors in order to better understanding what is happening.

Table 3.2 presents an overview of the sectors that are fulfilling that restriction.

processes	Agricultural machinery	Aviation	Crop fertilizers	Crop operations	Livestock	Point sources	Recreational boats	Residential	Shipping port	Traffic
1	True	True	False	True	False	True	True	False	True	False
2	True	True	True	True	True	True	True	False	True	False
4	True	True	True	True	True	True	True	True	True	False
8	True	True	True	True	True	True	True	True	True	False
16	True	True	True	True	True	True	True	True	True	False
32	True	Fail	True	True	True	True	True	True	True	False
64	True	Fail	True	True	True	True	True	True	Fail	False
128	True	Fail	True	True	True	True	True	True	Fail	False
256	True	Fail	True	True	True	True	True	True	Fail	True
512	True	Fail	True	True	True	True	True	True	Fail	True

Table 3.2: HERMESv3_BU simulated sectors by processes that are accomplishing the 15 minutes restriction.

In that table is marked as 'False' the configuration options that overpass the limit and with 'Fail' the ones that have failed. The sectors that have failed are the area source sector ones (1.6.1.2) that overpass the limit of processes permitted by the number of sources that the sector has (as it have been said in the section 2.1.2). The Iberian peninsula domain used in HERMESv3 includes 31 airports and 48 harbours, so, the simulations that are using more processes of that limit have been failed and are marked as 'Fail' in the Table 3.2.

That analysis have screened some configurations but, to chose the best configuration one, the speedup (3.3) and the efficiency (3.4) have to be evaluated as well.

3.3 Speedup

The first metric that is going to be evaluated is the speedup

$$S = \frac{T_s}{T_p} \tag{3.1}$$

where S is the speedup, T_s the time spent in serial mode (using only 1 process) and T_p the time spent by using the p number of processes.

The speedup metric was established by Amdahl's law [3] which was developed to analyse the improvement in time of the same simulation using different resources. The speedup starts at 1, for the simulations done in serial mode, and the ideal speedup is to have the same speedup as processes have been used.

It is normal to expect that using the double number of processes the time will be reduced to the half but is very hard to get that perfect speedup. Otherwise, it is not usual to have speedup over the ideal one but surprising results have been obtained by analysing the HERMESv3_BU speedup.

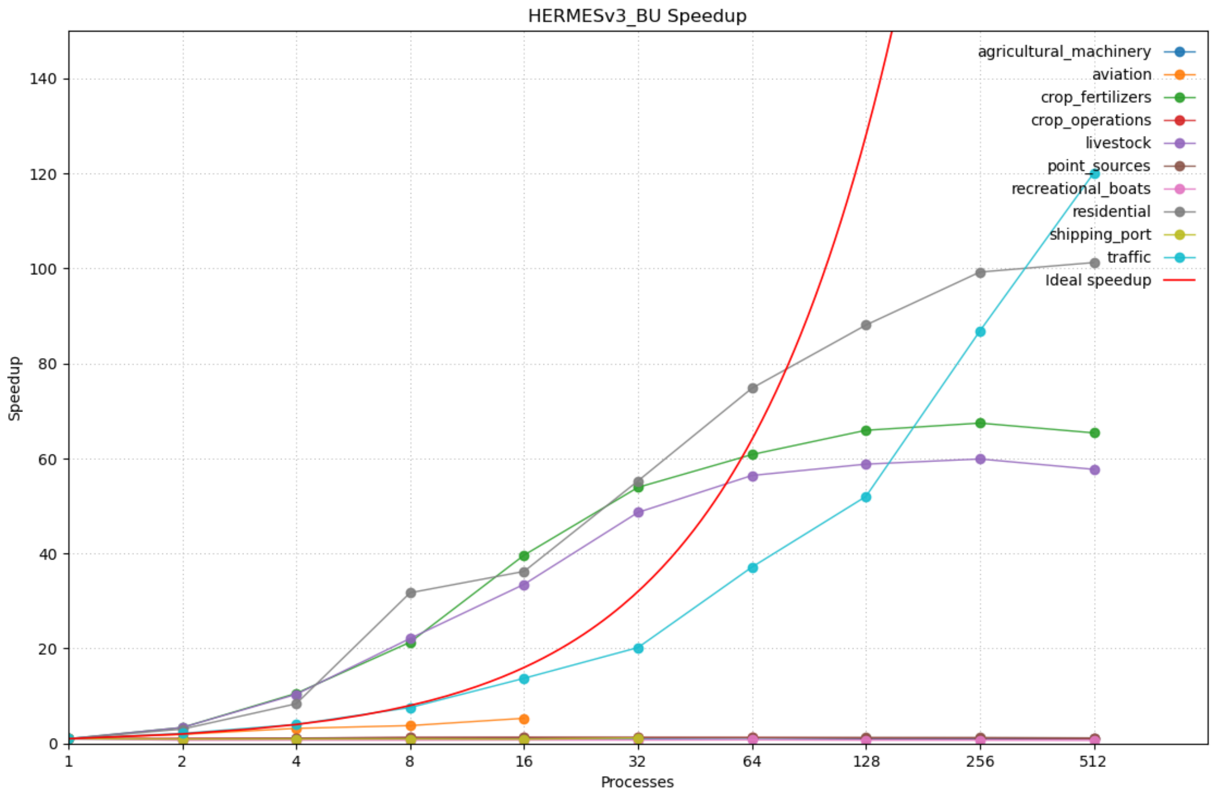

Figure 3.3: HERMESv3_BU speedup by sector

Figure 3.3 is showing the obtained speedup for each test done and HERMESv3_BU is having very good results for the residential combustion, crop fertilizers and livestock sector. That is due to all of them are gridded source sectors and all the calculation part are very well distributed over the processes. On the other hand, all of that sectors saturates the speedup at some point, due to the the writing part overhead, that is common on all the test done.

This writing cost can be noticed when the total speedup is compared to the speedup evaluation for only the calculation block (evaluating only the work done in the section 2.1) of each sector avoiding the writing part (Figure 3.4). These results prove that the main overhead comes from the writing part.

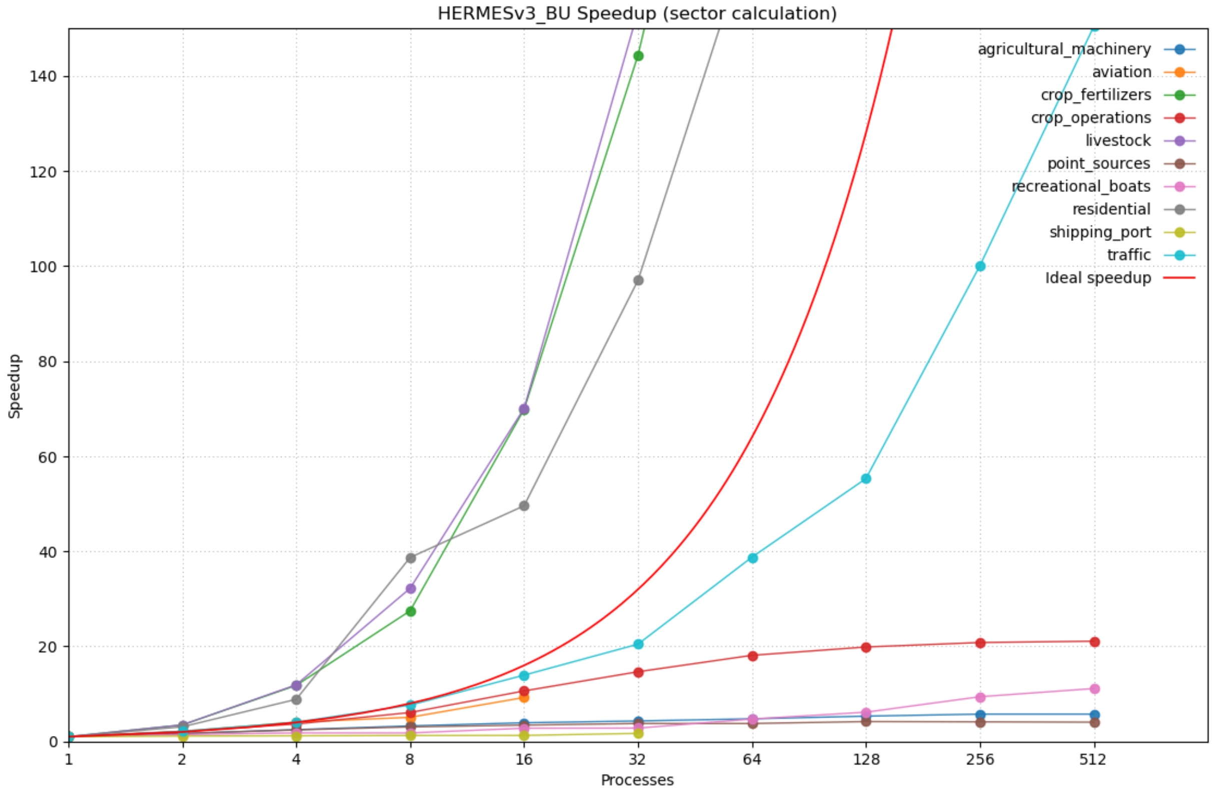

Figure 3.4: HERMESv3_BU speedup by sector (only sector calculation)

As it has been said, it is not usual to have an over-speedup but all that sectors calculations mentioned before are having it. That occurs due to the amount of data that the sector needs to do the calculation.

The over-speedup can be explained by the efficiency usage of the cache memory. A more completed description of the cache memory and its function can be read on [45] but we are going to do a brief summary.

MareNostrum4 has three levels of cache memory with different sizes (L1d 32K; L1i cache 32K; L2 cache 1024K; L3 cache 33792K (1.5.3.5)) and the data is moved from RAM memory to the different caches in order to be used. To do the calculation, the data needed has to be allocated on the L1 cache and, if it is not there, have to go to the next level of cache to found it. This effect is called cache locality.

This process is repeated over all the memory levels until the data is found. Then, it copies a full block of data containing the needed data as well as the subsequent values until fill the lowest cache level with the needed data. That procedure will copy more data than needed because is expected to be used in the next instructions.

If small pieces of data are needed, it is normal to expect that can be found in the lower levels of cache instead to go to the upper memory levels. Therefore less movements of memory have to be done. That is what is happening with the crop fertilizes, livestock and residential combustion sector and that explain the over-speedup of that sectors.

Road traffic sector is also having good results on speedup although them are under the ideal speedup line. As it has been explained in the section 2.1.3 the traffic sector is the most complex one and is very difficult to improve the results but it is a future work to explore (see 4.2.4 section). Part of that low speedup comes from the work to duplicate each road link by the involved vehicle composition that is very expensive in time. It has been also explored the possibility to save that expanded road links as an auxiliary file, but due to the amount of road links and vehicle types it is also more time consuming, the reading of that auxiliary file, than the creation. It has to be also taken into account that the formula of the road link emissions is quite complex and it varies depending on the calculated velocity for each vehicle type and road link. Lot of work can be done in order to improve that sector that was not expected for the complete parallelization of the application.

The rest of the sectors analyzed has not a good speedup because is very difficult to improve sectors with not too much calculation time. Except aviation sector one, all the rest of sectors are very low times even with only one process. Aviation sector with 2 or 4 process is also getting very good speedup values.

3.4 Efficiency

A second metric used to evaluated the parallelization of HERMESv3_BU is the efficiency, which is defined as follows:

$$E = \frac{S}{p} = \frac{T_s}{p * T_p} \tag{3.2}$$

where $S = \frac{T_s}{T_p}$ is the speedup, p the number of processes used, T_s the time spent in serial mode (using only 1 process) and T_p the time spent using the p number of processes.

The efficiency can be used to evaluate the energy consumption. The simulations with a high efficiency will consume less energy obtaining the same result as the high consumer ones with low efficiency. Is it expected that the efficiency metric have to be provided, in a near future, in order to grant an allocation resource in the supercomputers to run simulations. Therefore the resources requested that can prove to have a good efficiency, will have a better priority.

Ideal efficiency has values over 1 to do a better usage of the energy waste by the running time. So, an efficiency of one is the same that have an ideal speedup.

The Figure 3.4 is showing the efficiency results of all the tests done for HERMESv3_BU.

As the efficiency is very related with the speedup, the different sectors show a similar

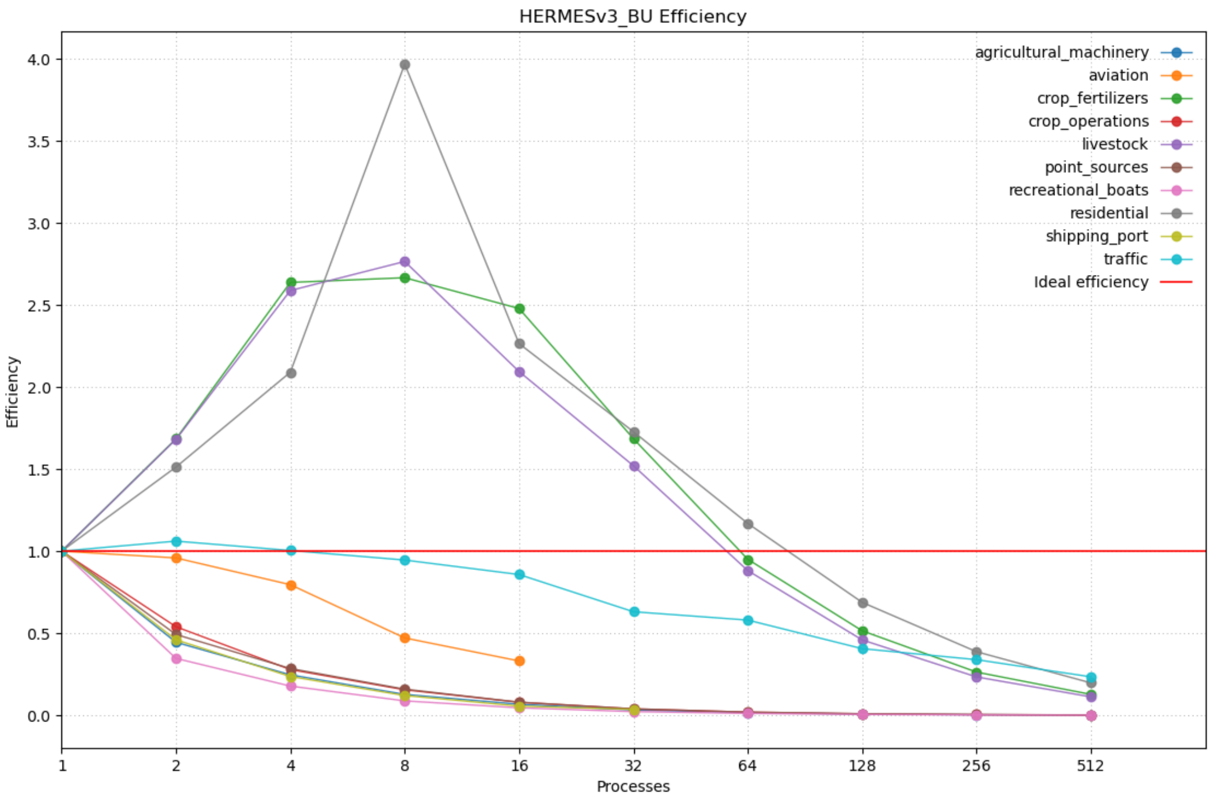

Figure 3.5: HERMESv3_BU efficiency by sector

behavior compared to speedup results.

At that point, we can assume that the best configuration to simulate each sector separately is the one that fits in minimum time required and has the highest efficiency. But that conclusion cannot be applied to choose the best configuration for a full simulation with all the sectors since is not easy when we want to calculate all the emission sectors in the same simulation. As it has been said in the writing section 2.3.3, when the emission calculation is finished, the processes that have spent less time have to wait to the most consuming ones to start the writing part. Because of that, we cannot take into account only the efficiency. It has to be also taken into account a combination of all the metrics and the spending time of calculation. Next section will explain the approach followed to find the best configuration.

3.5 Choosing the best configuration

To choose the best configuration, each one of the sectors and their metrics have to be evaluated separately. According to the results showed in figure X and figure Y, we will epxlain the best number of processes per sector.

To this point, it is easy to select the sectors that consume less time by selecting the configuration of only one process: agricultural machinery, crop operations, point sources, recreational boats and shipping port.

For the aviation sector we are going to choose 2 processes because that combination have also performed a good efficiency and the time consuming is more much below of the time limit. If one process was selected, we were only 5 minutes below the limit.

The efficiency of the residential combustion sector using 8 processes is impressive therefore we are going to take that value as the chosen one.

Even the best efficiency for crop fertilizes and Livestock is not 4 we are choosing that configuration because the efficiency is better and the time consumed is closed to the best one possible.

At that point we have selected a total of 23 processes. For the traffic sector we would take 256 as the minim value but another thing have to be considered before taking that decision.

HERMESv3_BU has to be run in the MareNostrum4 platform (see requirement 1.5.3.5) and we know that each node has 48 cores. Therefore, a multiple of 48 should be chosen as the total number of processes selected.

To explain better that restriction imagine that you want to execute a program with 50 processes, therefore 2 nodes (96 cores) would be provided in exclusive to your simulation allocating 25 processes in each node. 46 processes would be wasted in that simulation. That is why is better to choose multiple of 48 processes in total if the overhead introduced when we use more processes does not affect to the efficiency or speedup significantly.

Returning to the previous point we have already booked 23 processes and we have to take at least 256 for the traffic one. That amount of processes sum a total of 279 but the immediately multiple of 48 higher value to that is 288 (6 MareNostrum nodes). If we finally use 288 processes for the total full simulation 265 can be used for the traffic sector.

Table 3.5 is summarizing the results of that study.

Sector	**processes**
Agricultural machinery	1
Aviation	2
Crop fertilizers	4
Crop operations	1
Livestock	4
Point sources	1
Recreational boats	1
Residential combustion	8
Shipping port	1
Road traffic	265

Table 3.3: Chosen processes distribution by sector to get a full simulation of HERMESv3_BU

Using that configuration the total amount of time spent to run a full simulation of HERMESv3_BU is 12 minutes and 36.1681 seconds. It has been proved that this simulation accomplish all the requirements.

3.6 Writing configuration

Once the selection of the processes has been done, we are going to evaluate what type of writing and how many processes we are going to select.

The simulations for the writing test purpose have been run using the configuration provided by the previous study. Figure 3.6 is showing the time spent to write with the chunked write and the standard one.

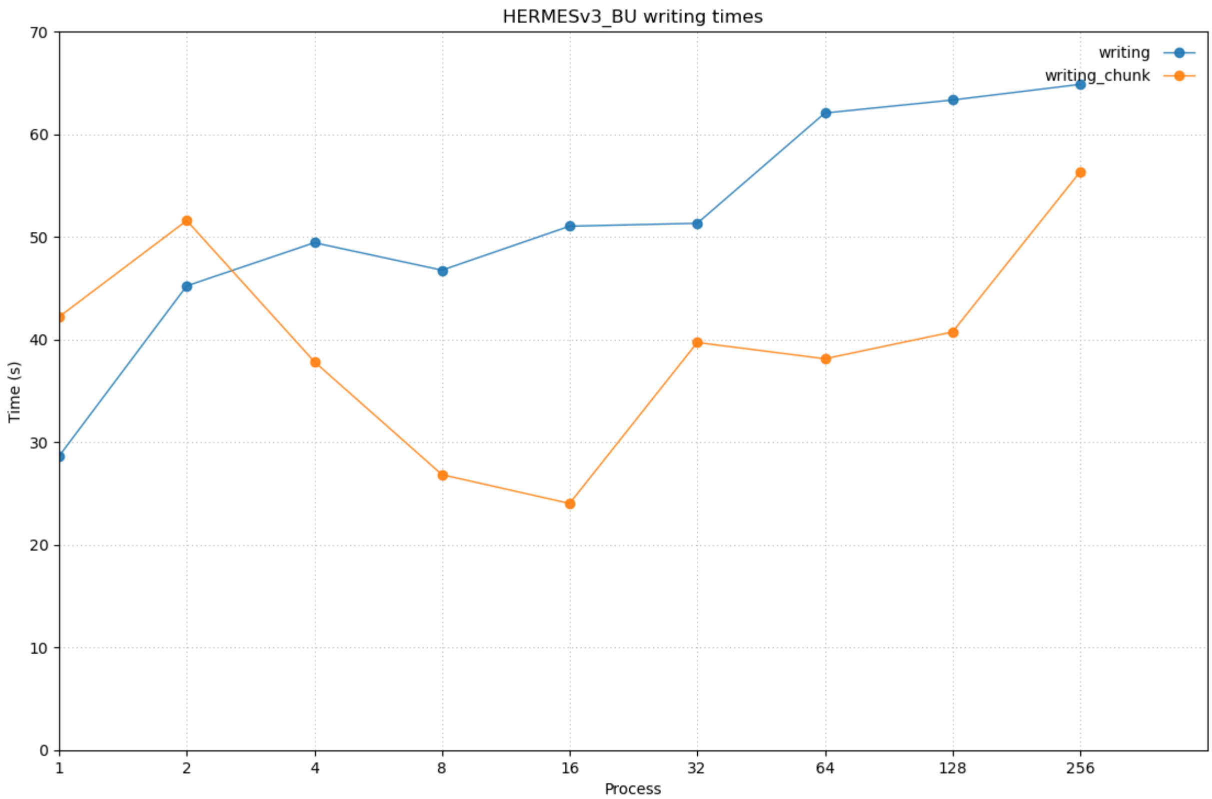

Figure 3.6: HERMESv3_BU writing times

On the one hand, the normal writing increase the time consumed proportionally to the processes used, due to the effect of have more NetCDF write calls. On the other hand, we have the writing by chunks that is having a different behaviour as it is explained below.

Firstly, it has an increase of time in the two processes simulation. That is due to the time consumed by the chunk creation. For the subsequent simulations, the chunk creation is compensated by the time spent to write so, we are having a reduction on the time consumed.

The rest of simulations, with more than 16 processes devoted to write, are increasing the time spent and the creation of the chunks are not as well compensated as the previous ones.

The best option to chose the writing processes can be the one that the writing part is chunking the NetCDF using 16 processes. But the difference between this way and the normal serial option is too small.

Additionally, it has to be taken into account that, similar that the compressed NetCDF, if the emission file have to be read by a chemical transport model, the reading processes will have to read from different chunks and it is more time consuming than read from an unchunked file. So, finally, the chosen option to write a NetCDF is going to be the normal serial one (using only one process to write.

As it has been said in the section 2.3.3.3 HERMESv3 can compress the NetCDF only while using one process but it is very time consuming. Even that option has not to be used for air quality system simulations, we are going to evaluate the compression for that purposes that the only area of study is the emissions and the emission file has to be saved.

Write type	**Size (Bytes)**	**Size (GB)**	**Time**
Uncompressed	29063130108	27.067	12 minutes 36.1681 seconds
Compressed	262144	0.168	16 minutes 39.7230 seconds

Table 3.4: Comparing size and time to compress a NetCDF file

Table 3.6 shows the differences in size and time by using the compression (using a level of compression of 4) with the no compression one.

As we can see, the compressed file is occupying only 0.62% of the total amount of memory occupied by the uncompressed one but increasing by 4 minutes the total amount of time.

It is a useful option if you want to save the emissions but not a good one for operational runs due to the decompression that the Chemical Transport Model while reading them.

Chapter 4

Conclusions

4.1 Overview

This work presents the description and evaluation of the first and novel parallel strategy developed within the HERMESv3 atmospheric emission modelling system. The HERMESv3 system is used to compute high resolution atmospheric emission from multiple sources, including point sources, road transport, residential and commercial combustion, shipping ports, aviation, recreational boats, livestock, agricultural crop operations, agricultural machinery and agricultural fertilizers.

For each sector, a specific parallelization strategy was defined, implemented and evaluated as well as the strategies to write the output file, in order to achieve the best efficiency possible for each different sector using different parallelization approaches.

The HERMESv3 code is distributed free of charge under the licence GNU GPL v3.0. through the Earth Science Gitlab repository (https://earth.bsc.es/gitlab/es/hermesv3_bu). The repository also includes a complete wiki page with the description and usage of each one of the HERMESv3_BU modules.

Thanks to the work described on that document, HERMESv3_BU is operationally implemented in the CALIOPE system and is being used to produce the forecast over Spain and Catalonia. And also it has been reached these outcomes:

- Input data can be provided with shape formats without any pre-processing.
- Output files can be provided in multiple formats.
- It can be run in the MareNostrum environment using different number of processes.
- It uses the computational resources in the most efficient way to save energy and time.
- HERMESv3 can simulate several sectors in the same run.

Despite presenting highly accurate and detailed modelling methodologies, and the improvements achieved in terms of consumption of computational resources, there are still some shortcomings associated with the model:

- The processes distribution have to be known, or analyzed, in advance manually.
- It is not possible to run different sectors using one process.
- Road traffic sector uses lot of memory and time.
- It is not possible to produce a compressed NetCDF in parallel.
- It is not overlapping calculation with the writing part.

All that shortcomings are deeper analysed on the 4.2 section in order to be explored in the future.

4.1.1 Usages

HERMESv3_BU is currently being used as the official emission model at the Earth Sciences Department of the Barcelona Supercomputing Center. The tool has been implemented as the emission core of the CALIOPE air quality forecasting system, which operationally provides air quality forecast for Spain and the region of Catalonia for the next 24 and 48 hours at a spatial resolution of 4km x 4km and 1km x 1km, respectively. The results produced for the region of Catalonia are being used by the Generalitat de Catalunya as a complementary information for their air quality management activities.

HERMESv3 is also being used to analyze the impact of the COVID-19 effect on the air quality levels due to the emission reduction.

Thanks to HERMESv3 is an open source model, everyone can use it by downloading them from their GitLab repository and we are expecting to have lots of users that can take benefit from our developments.

4.2 Future work, ideas for improvement and limitations

HERMESv3_BU is an continuous development software that, even with all it advantages, can be improved during the time.

It is listed some of the futures works, ideas and limitations related to the HERMESv3 evolution.

4.2.1 Choosing online the sector processes distribution

As it have been said in the 3.5 section, the best configuration has to be chosen with an analysis of all their components.

A possibility is to estimate the time consumed by each sector by using a formula that may depends on the quantity of input data for each sector in order to obtain an estimation of the time that it will spend.

That will allow to do a simpler study in order to get an automatic processes distribution over the different sector calculators.

4.2.2 Several sectors calculation on the same process

Due to a design decision over the sector manager 2.2, it is not possible to run more than one sector using the same process.

A possibility of improvement is to gather the calculation sectors that spend less time into an unique process.

4.2.3 Emission calculation as pool of emission sources

An idea for the improvement of the emission calculation is to create a pool of emission sources instead of having them divided with the sector manager.

That idea is to have a stack of emission sources and each process has to take a group of sources to calculate them. That will allow to do not have processes waiting without operations to do, which are waiting for the rest of sectors that have not finished yet.

It has to be taken also into account the dimension of those groups of emissions in order to take benefit from the vectored operations (see section 2.1). Also it has to be taken into account all the amount of communications to send a lots of groups of emissions. That complex communication patterns have to be developed in order to know when a process is ready to calculate other group.

4.2.4 Better implementation of road traffic emissions

As it is explained in previous sections, the road traffic sector is the largest in time and memory consumption among all the HERMESv3_BU modules, as well as the worst balanced one.

It has to be deeper analysed to try to improve the internal algorithm, in order to calculate the emissions using a better usage of the memory, exploring different strategies:

- Different data structures have to be analysed.

- Try to save as an auxiliary file the common operations that may repeat over the daily simulation.
- Identify if the same operation is repeated over several road links in order to do them once.
- To distribute vehicle types in road links instead to do it by road links.
- Divide the road transport sector into small sectors to overlap the different emission type calculations: exhaust, non exhaust and gasoline evaporation.

4.2.5 Parallel compression

As previously highlighted, the parallel compression is not still allowed in HERMESv3.

We have to remark that a NetCDF4 file is build over HDF5 and in the release 1.10.2 of the HDF5 library (April 2018) they have added the capability to write a HDF5 files with parallel compression.

This is a very new implementation and hard to configure. It has to be understood that the python NetCDF4 library is installed over the C NetCDF4 library one, and that one is installed over HDF5 and other libraries with more dependencies.

The way that all the module dependencies are installed may add the capability of the parallel compression of a NetCDF4 file. A further exploration on that topic has to be explored.

However, it is known that the parallel compression must follow the rule of the chunking and that work have been done in advance.

4.2.6 Asynchronous write by time step

With the purpose to overlap calculation and writing process, both parts have to be re-formulated. It is needed to change each sector code in order to calculate the emissions by time step and send them to write while calculating the next one.

That means that have to be increased the number of processes in order to have some specific ones devoted only to write. This is a typical approach used by several models in Earth Science where there is a IO server for the writing process, though not trivial for the implementation. Those servers can be devoted to write or to be used as interface to communicate the emission model with the chemical transport model in order to have a complete online model.

4.2.7 HERMESv3 as online emission model inside MONARCH

As it has been said previously (see section 1.2), MONARCH [1] is the online Meteorological and Chemical transport model that is going to substitute the current ones (WRF for the weather and CMAQ for the chemistry) in a future.

MONARCH is still under development but it is planned that in the near future HERMESv3 will have to be used as an online component of MONARCH.

To achieve that, modifications in the workflow of HERMES has to be done in order to achieve the calculation by time step as well as some modifications to couple HERMESv3 and MONARCH, in order to obtain the online model. That coupling has to communicate between a FORTRAN model (MONARCH) and a Python one (HERMESv3).

As an example, the COSMO chemical transport model [24] has an online emission module developed with Python and we have to analyse the implementation in order to learn the state of the art of that topic.

Bibliography

[1] A. Badia, O. Jorba, A. Voulgarakis, D. Dabdub, C. Pérez García-Pando, A. Hilboll, M. Gonçalves, and Z. Janjic. Description and evaluation of the multiscale online nonhydrostatic atmosphere chemistry model (nmmb-monarch) version 1.0: gas-phase chemistry at global scale. *Geoscientific Model Development*, 10(2):609–638, 2017.

[2] Baek B.H. SMOKE (Sparse Matrix Operator Kerner Emissions) Modeling System, 2019. https://www.cmascenter.org/smoke Accessed: 2020-06-09.

[3] Mike Bailey. Parallel Programming: Speedups and Amdahl's law. *Oregon State University*, pages 1–7, 2019.

[4] José Baldasano, Leonor Güereca, Eugeni Lopez, Santiago Gassó, and Pedro Jimenez-Guerrero. Development of a high-resolution (1km×1km, 1h) emission model for Spain: The High-Elective Resolution Modelling Emission System (HERMES). *Atmospheric Environment*, 42:7215–7233, 2008.

[5] José María Baldasano Recio, Oriol Jorba Casellas, Santiago Gassó Domingo, M. Teresa Pay Pérez, and Gustavo Arevalo Roa. Caliope: sistema de pronóstico operacional de calidad del aire para Europa y España. *XXXII Jornadas Científicas de la Asociación Meteorológica Española.Meteorología y Calidad del Aire*, pages 1–6, 2012.

[6] Sara Basart, Carlos PéRez, Slodoban Nickovic, Emilio Cuevas, and José MaríA Baldasano. Development and evaluation of the BSC-DREAM8b dust regional model over northern Africa, the mediterranean and the middle east. *Tellus, Series B: Chemical and Physical Meteorology*, 64(1), 2012.

[7] Jaime Benavides, Michelle Snyder, Marc Guevara, Albert Soret, Carlos Pérez García-Pando, Fulvio Amato, Xavier Querol, and Oriol Jorba. Caliope-urban v1.0: coupling r-line with a mesoscale air quality modelling system for urban air quality forecasts over barcelona city (spain). *Geoscientific Model Development*, 12:2811–2835, 07 2019.

[8] BSC. Barcelona supercomputing center. https://www.bsc.es/ Accessed: 2020-06-05.

[9] BSC. Earth science deparment of the barcelona supercomputing center. https://www.bsc.es/discover-bsc/organisation/scientific-structure/earth-sciences Accessed: 2020-06-05.

[10] BSC. Marenostrum4. https://www.bsc.es/es/marenostrum/marenostrum/informacion-tecnica Accessed: 2020-04-13.

[11] CMAS. Ioapi. https://www.cmascenter.org/ioapi Accessed: 2020-04-13.

[12] European Community and International Environmental. Kyoto Protocol Status of Ratification. *Review of European Community and International Environmental Law*, 7(2):214–217, 2009.

[13] Brian Eaton, Jonathan Gregory, Bob Drach, Karl Taylor, Steve Hankin, John Caron, Rich Signell, Phil Bentley, Greg Rappa, Heinke Höck, Alison Pamment, and Martin Juckes. NetCDF Climate and Forecast Metadata Conventions, v 1.6. . *Http://Cf-Pcmdi.Llnl.Gov/Documents/Cf-Conventions/Latest-Cf-Conventions-Document-1*, 1.6(December), 2011.

[14] European Comission. en @ www.copernicus.eu. https://www.copernicus.eu Accessed: 2020-05-13.

[15] Marius Gilbert, Gaëlle Nicolas, Giuseppina Cinardi, Thomas Boeckel, Sophie Vanwambeke, G. Wint, and Timothy Robinson. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific Data*, 10 2018.

[16] C. Granier, S. Darras, H. Denier van der Gon, J. Doubalova, N. Elguindi, B. Galle, M. Gauss, M. Guevara, J.-P. Jalkanen, J. Kuenen, C. Liousse, B. Quack, D. Simpson, and K. Sindelarova. The copernicus atmosphere monitoring service global and regional emissions (april 2019 version), 2019.

[17] H. Grythe, S. Lopez-Aparicio, M. Vogt, D. Vo Thanh, C. Hak, A. K. Halse, P. Hamer, and G. Sousa Santos. The metved model: development and evaluation of emissions from residential wood combustion at high spatio-temporal resolution in norway. *Atmospheric Chemistry and Physics*, 19(15):10217–10237, 2019.

[18] M. Guevara, C. Tena, A. Soret, K. Serradell, D. Guzmán, A. Retama, P. Camacho, M. Jaimes-Palomera, and A. Mediavilla. An emission processing system for air quality modelling in the Mexico City metropolitan area: Evaluation and comparison of the MOBILE6.2-Mexico and MOVES-Mexico traffic emissions. *Science of the Total Environment*, 584-585, 2017.

[19] Marc Guevara, Francesc Martínez, Gustavo Arévalo, Santiago Gassó, and José M. Baldasano. An improved system for modelling spanish emissions: Hermesv2.0. *Atmospheric Environment*, 81:209 – 221, 2013.

[20] Marc Guevara, Albert Soret, Gustavo Arévalo, Francesc Martínez, and José Baldasano. Implementation of plume rise and its impacts on emissions and air quality modelling. *Atmospheric Environment*, 99:618–629, 12 2014.

[21] Marc Guevara, Carles Tena, Manuel Porquet, Oriol Jorba, and Carlos Pérez Garc\'\ia-Pando. HERMESv3, a stand-alone multi-scale atmospheric emission modelling framework - Part 2: The bottom-up module. *Geoscientific Model Development*, 13(3):873–903, mar 2020.

[22] Marc Guevara, Carles Tena, Manuel Porquet, Oriol Jorba, and Carlos Pérez García-Pando. HERMESv3, a stand-alone multi-scale atmospheric emission modelling framework – Part 1: global and regional module. *Geoscientific Model Development*, 12(5):1885–1907, may 2019.

[23] S. Ibarra-Espinosa, R. Ynoue, S. O'Sullivan, E. Pebesma, M. D. F. Andrade, and M. Osses. Vein v0.2.2: an r package for bottom–up vehicular emissions inventories. *Geoscientific Model Development*, 11(6):2209–2229, 2018.

[24] Michael Jähn, Gerrit Kuhlmann, Qing Mu, Jean-Matthieu Haussaire, David Ochsner, Katherine Osterried, Valentin Clément, and Dominik Brunner. An online emission module for atmospheric chemistry transport models: Implementation in COSMO-GHG v5.6a and COSMO-ART v5.1-3.1. *Geoscientific Model Development Discussions*, 0:1–20, 2020.

[25] C Keller, M Long, Robert Yantosca, Arlindo Da Silva, Steven Pawson, and D Jacob. HEMCO v1.0: a versatile, ESMF-compliant component for calculating emissions in atmospheric models. *Geoscientific Model Development*, 7, 2014.

[26] Hamaoui-Laguel Lynda, Frederik Meleux, Matthias Beekmann, Bertrand Bessagnet, Sophie Génermont, Pierre Cellier, and Laurent Létinois. Improving ammonia emissions in air quality modelling for france. *Atmospheric Environment*, 92:584–595, 08 2014.

[27] NCCAR. Wrf conventions. https://www.ncl.ucar.edu/Applications/wrfnetcdf.shtml Accessed: 2020-05-26.

[28] NOAA. A model based on ocean and atmosphere interactions, 2020. https://celebrating200years.noaa.gov/breakthroughs/climate_model Accessed: 2020-06-05.

[29] US EPA Office of Research and Development. Cmaq, August 2019. For up-to-date documentation, source code, and sample run scripts, please clone or download the CMAQ git repository available through GitHub: https://github.com/USEPA/CMAQ.

[30] World Health Organization. Ambient (outdoor) air pollution, May 2018. https://www.who.int/en/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health: 2020-06-05.

[31] C. Pérez, K. Haustein, Z. Janjic, O. Jorba, N. Huneeus, J. M. Baldasano, T. Black, S. Basart, S. Nickovic, R. L. Miller, J. P. Perlwitz, M. Schulz, and M. Thomson. Atmospheric dust modeling from meso to global scales with the online NMMB/BSC-Dust model – Part 1: Model description, annual simulations and evaluation. *Atmospheric Chemistry and Physics*, 11(24):13001–13027, 2011.

[32] Carlos Pérez, Slobodan Nickovic, Goran Pejanovic, José María Baldasano, and Emin Özsoy. Interactive dust-radiation modeling: A step to improve weather forecasts. *Journal of Geophysical Research Atmospheres*, 111(16), 2006.

[33] Jordan G Powers, Joseph B Klemp, William C Skamarock, Christopher A Davis, Jimy Dudhia, David O Gill, Janice L Coen, David J Gochis, Ravan Ahmadov, Steven E Peckham, Georg A Grell, John Michalakes, Samuel Trahan, Stanley G Benjamin, Curtis R Alexander, Geoffrey J Dimego, Wei Wang, Craig S Schwartz, Glen S Romine, Zhiquan Liu, Chris Snyder, Fei Chen, Michael J Barlage, Wei Yu, and Michael G Duda. The Weather Research and Forecasting Model: Overview, System Efforts, and Future Directions. *Bulletin of the American Meteorological Society*, 98(8):1717–1737, 2017.

[34] Salvador Enrique Puliafito, David Gabriel Allende, Fernando Castro, Rafael Pedro Fernandez, and Pablo Cremades. New Approaches for Urban and Regional Air Pollution Modelling and Management. In Farhad Nejadkoorki, editor, *Advanced Air Pollution*, Chapters. IntechOpen, May 2011.

[35] R. G. Quayle and H. F. Diaz. Heating degree day data applied to residential heating energy consumption., 1980.

[36] Baldasano Recio; José María; Jorba Casellas; Oriol; Gassó Domingo; Santiago; Pay Pérez; M. Teresa; Arevalo Roa;. Caliope. http://www.bsc.es/caliope Accessed: 2020-04-13.

[37] William J Sacks, Delphine Deryng, Jonathan A Foley, and Navin Ramankutty. Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography*, 19(5):607–620, 2010.

[38] S. Schiavina, M.; Freire and K. MacManus. Ghs population grid multitemporal (1975-1990-2000-2015), 2019a. european commission, 2019.

[39] Copernicus Land Monitoring Service. Corine land cover, 2019. https://land.copernicus.eu/pan-european/corine-land-cover Accessed: 2020-04-15.

[40] J. B. Klemp J. Dudhia D. O. Gill Z. Liu J. Berner W. Wang J. G. Powers M. G. Duda D. M. Barker Skamarock, W. C. and X.-Y. Huang. A description of the advanced research wrf version 4. *NCAR Technical Notes Note NCAR/TN-556+STR*, 2019.

[41] Michelle Snyder, Akula Venkatram, David Heist, Steven Perry, William Petersen, and Vlad Isakov. Rline: A line source dispersion model for near-surface releases. *Atmospheric Environment*, 77:748–756, 10 2013.

[42] UCAR. Netcdf. https://www.unidata.ucar.edu/software/netcdf Accessed: 2020-04-13.

[43] United Nations. CHAPTER : Chapter XXVII. Environment TITLE : 1. Convention on Long-range Transboundary Air Pollution. Geneva, 13 November 1979. *United Nations, Treaty Series*, 1302(November), 1979.

[44] United Nations. Summary of the Paris Agreement. *United Nations Framework Convention on Climate Change*, pages 27–52, 2015.

[45] Palash Volvoikar. How Does CPU Cache Work and What Are L1, L2, and L3?, 2019. https://www.makeuseof.com/tag/what-is-cpu-cache/ Accessed: 2020-05-31.