VI-SEEM DREAMCLIMATE SERVICE

DUŠAN VUDRAGOVIĆ,* LUKA ILIĆ,† PETAR JOVANOVIĆ,‡ SLOBODAN NIČKOVIĆ,§ ALEKSANDAR BOGOJEVIĆ,¶ AND ANTUN BALAŽ‖

Abstract. Premature human mortality due to cardiopulmonary disease and lung cancer is found in epidemiological studies to be correlated to increased levels of atmospheric particulate matter. Such negative dust effects on the human mortality in the North Africa – Europe – Middle East region can be successfully studied by the DREAM dust model. However, to assess health effects of dust and its other impacts on the environment, a detailed modelling of the climate for a period of one year in a high-resolution mode is required. We describe here a parallel implementation of the DREAM dust model, the DREAMCLIMATE service, which is optimised for use on the high-performance regional infrastructure provided by the VI-SEEM project. In addition to development and integration of this service, we also present a use-case study of premature mortality due to desert dust in the North Africa – Europe – Middle East region for the year 2005, to demonstrate how the newly deployed service can be used.

Key words: DREAM model, dust effects, human mortality, VI-SEEM project, application service

AMS subject classifications. 68W10, 68M14, 68N30

1. Introduction. Exposure to airborne mineral dust particles can significantly influence human health. Atmospheric dust particles are primarily driven by mesoscale and synoptic processes, and may be present in high concentrations near the sources and carried over long distances while having adverse health effects. Drought and desertification, as climate-related changes and human activities such as changes in land use, affect potential dust sources of fine particulate matter in arid areas. Therefore, numerical modelling with sufficiently high resolution of the processes of the atmospheric dust cycle that drive dust emissions and transport is a useful approach to assessment of the potential health effects of exposure to dust.

The previously developed Dust REgional Atmospheric Modeling (DREAM) system [1] is a component of a comprehensive atmospheric model designed to simulate and predict the atmospheric cycle of mineral dust aerosols. The DREAM provides a climatology of dust based on long-term re-analysis of the model. It is widely used by the research and operational dust forecasting communities in more than 20 countries, including its recent use in a series of NASA-funded projects [2, 3, 4, 5] dealing with health aspects of dust suspended in the air. The Institute of Physics Belgrade group, which is a partner in the Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS) project of the World Meteorological Organization, uses DREAM to provide daily dust forecasts to the SDS-WAS model inter-comparisons and validation activities. Also, it is used for investigation on how fine particulate matter contributes to air pollution in North Africa – Europe – Middle East region.

To assess health effects of dust in the region and other dust impacts on the environment, it is usual to consider at least a one-year modelling climatology for the given region. In this case this was achieved by solving the DREAM model in a high-resolution mode with the horizontal grid resolution of 15 km. Such a high resolution model is capable to accurately describe the behaviour of small-scale dust sources in the desert areas (Sahara, Middle East), as well as the mesoscale atmospheric conditions. However, due to numerical complexity it requires a parallelised version of the DREAM code, which we created and optimised for usage on high-performance computing infrastructures available today.

---

*Scientific Computing Laboratory, Center for the Study of Complex Systems, Institute of Physics Belgrade, University of Belgrade, Serbia (dusan.vudragovic@ipb.ac.rs).

†Environmental Physics Laboratory, Institute of Physics Belgrade, University of Belgrade, Serbia (luka.ilic@ipb.ac.rs).

‡Scientific Computing Laboratory, Center for the Study of Complex Systems, Institute of Physics Belgrade, University of Belgrade, Serbia (petar.jovanovic@ipb.ac.rs).

§Environmental Physics Laboratory, Institute of Physics Belgrade, University of Belgrade, Serbia (slobodan.nickovic@ipb.ac.rs).

¶Scientific Computing Laboratory, Center for the Study of Complex Systems, Institute of Physics Belgrade, University of Belgrade, Serbia (aleksandar.bogojevic@ipb.ac.rs).

‖Scientific Computing Laboratory, Center for the Study of Complex Systems, Institute of Physics Belgrade, University of Belgrade, Serbia (antun.balaz@ipb.ac.rs).

In parallel to development of the DREAM model, a number of initiatives were crucial for enabling high-quality climate research in the region. This was achieved by providing e-Infrastructure resources, application support and training through the VI-SEEM project [6], funded by the EU H2020 programme. The project brings together regional e-Infrastructures in order to build capacity and better utilise synergies, as well as to provide improved service within a unified virtual research environment for several inter-disciplinary scientific user communities. The overall aim is to offer a user-friendly integrated e-Infrastructure platform for regional cross-border scientific communities in climatology, life sciences, and cultural heritage. This includes integration of computing, data, and visualisation resources, as well as services, models, software solutions and tools. The VI-SEEM virtual research environment provides the support to scientists in a full lifecycle of collaborative research.

By efforts of the DREAM code developers and the VI-SEEM support team, the DREAM model was successfully refactored and tuned for usage on high-performance computing infrastructures in a form of the DREAMCLIMATE service, presented here. Section 2 briefly describes the DREAM model, which is capable of producing results in the required high-resolution mode for a one year period. The DREAMCLIMATE service is presented in detail in Section 3, while Section 4 describes produced datasets and main results. By using an order of magnitude finer DREAM model grid than available before, we perform a detailed analysis of dust impacts to public health.

2. DREAM model. Premature human mortality due to cardiopulmonary disease and lung cancer is found in epidemiological studies to be correlated to increased levels of atmospheric particulate matter, in particular to long-term exposure to particulate matter with an aerodynamic diameter smaller than $2.5\,\mu$m. In order to estimate the premature mortality caused by the long-term exposure to airborne desert dust, we use results of the DREAM gridded model dust climatology of fine particulate matter and dust concentrations. This analysis follows the previous study [7] that indicates that there is a large number of premature deaths by cardiopulmonary disease and a significant number of deaths by lung cancer, mostly in the dust belt region neighbouring Sahara and Middle East deserts.

The DREAM model is developed as an add-on component of a comprehensive atmospheric model and is designed to simulate and/or predict the atmospheric cycle of mineral dust aerosols. It solves a coupled system of the Euler-type partial differential nonlinear equations for dust mass continuity, one equation for each particle size class, which is one of the governing prognostic equations in an atmospheric numerical prediction model [8, 9, 10]. The DREAM model takes into account all major processes of the atmospheric dust cycle. During the model simulation, calculation of the surface dust emission fluxes is made over the model cells declared as deserts. A viscous sub-layer parameterisation regulates the amount of dust mass emission for a range of near-surface turbulent regimes. Once injected into the air, dust aerosols are driven by the atmospheric dynamics and corresponding physical quantities: by turbulence in the early stage of the process, when dust is lifted from the ground to the upper levels; by winds in later phases of the process, when dust travels away from the sources; and finally, by thermodynamic processes, rainfall and land cover features that provide wet and dry deposition of dust over the Earth surface.

The model is implemented as a bundle of Fortran programs and libraries. These components are divided into three groups: the preprocessing system, the model operational system, and post-processing and visualisation tools. The preprocessing consists of two phases. The first is the setup in which the simulation domain, model configuration and interpolation of terrestrial data are defined. These parameters are mostly hard-coded and any change to parameters in this phase requires recompilation. The second stage of preprocessing is interpolation of the meteorological input data from the global meteorological model to the current simulation domain, as well as a setup of initial boundary conditions for the dust model. The model operational system is the main component, and it runs the numerical integration program. Post-processing and visualisation tools include GrADS [11] with conversion from Arakawa E-grid to geo-referenced grid and plots.

The code is predominantly written in the style of the Fortran 77 standard. Some of the more pressing constraints of the standard were the lack of support for dynamic memory allocation and command line arguments. These two constraints required for a number of parameters to be hard-coded. As a consequence, this limited the number of users who could use the application independently, and the number of parallel tests that could be ran at once. Recompilation also requires a deep technical knowledge of the implementation itself, which reduces

usability and dissemination of the model.

3. DREAMCLIMATE service. Within the framework of the VI-SEEM project, the DREAM model was successfully re-factored and tuned for usage on high-performance computing infrastructures. The DREAM-CLIMATE service was developed and deployed using the VI-SEEM infrastructure modules. Configuration of the considered physical system is separated from the source code of the application, and all relevant parameters are grouped into a single configuration file. Such an improved configuration approach enabled more user-friendly way to configure various model setups, without the need for each user to dive into the code and technical details of the implementation. This also enables multiple users to run their model instances independently. Important additional improvements include significant reduction of the disk-space consumption, as well as standardisation of its usage through an environment-module approach.

Configuration files follow the format of the Python configuration parser, which is a convenient, flexible, and powerful way for parsing configuration files. It uses simple INI style configuration syntax, i.e., a text file with a basic structure composed of sections, properties, and values. Parameters are divided into sections which are designated by square brackets. Within one section, each parameter is specified in a separate line and its name and value are delimited by the equals sign. In-line comments are also permissible and corresponding lines begin with a semicolon. In addition to this, a support for variable interpolation is included as well.

The DREAM processing stages remain similar to the original version of the code, and consist of the pre-processing, the model operational processing, and the post-processing phase. Majority of changes are related to reducing the complexity of configuration in the setup stage of preprocessing. In a typical use-case, a user begins the simulation project by loading the environment module for the DREAMCLIMATE service, which sets the environment paths for the commands used to initialise and prepare the DREAM model simulation. Afterwards, by invoking the `dreamclimate_init` script the default configuration file is created in the working directory and files needed for a configuration of the local simulation instance are created in the `.dreamclimate` subdirectory. After the parameters are set in the configuration file, the `dreamclimate_reconfig` script is called to execute the setup stage, which encapsulates recompilation of the components, depending on the parameters changed. The resulting binaries, which are used to run simulation, are placed in the `.dreamclimate/bin` directory. This step isolates each user's simulation instance from others and enables multiple instances to work without interference. The next step in this stage generates and interpolates vegetation and soil texture for the forecast domain, by calling the `gt30mounth`, `gt30source`, `gt30vegetadirect`, `text4eta`, and `texteta` components.

After the setup, preprocessing continues by invoking the `dreamclimate_preproc` script whose role is to prepare input data for the Eta model grid. This script invokes the following components:

- `climsst` – horizontal grid (IMT, JMT) Eta model indexing from the SST as a function of the month,
- `anecw` – horizontal grid (IMT, JMT) Eta model indexing from global initial data,
- `pusiWRF` – set of the vertical variables and vertical interpolation of the pressure to sigma surfaces,
- `const` – conversion of the initial fields in Eta model coordinates from 2D horizontal (IMT, JMT) indexing to 1D (IMJM), definition of dummy initial boundary soil moisture and temperature values, and calculation of the constants needed for the 1D version of the soil model,
- `dboco` – creation of the boundary condition files,
- `gfdlco2` – interpolation of the transmission functions grid, for which the transmission functions have been pre-calculated, to the grid structure.

This preprocessing step produces binary files interpolated to the model grid (i.e., Arakawa E-grid) in the output directory specified in the configuration file. All the routines of the model itself, which describe atmospheric processes including the dust cycle, are built into the main executable file. This is a parallel MPI program that runs the simulation and is submitted to the job scheduling system using the job description script, which is automatically generated earlier in the setup stage. The post-processing includes the conversion of the main GrADS output file from the Arakawa E-grid to the GrADS grid. These steps are handled by the `dreamclimate_post-process` script.

Many of the configuration parameters in the generated configuration file have sensible default values, to minimise the need for users to search through lengthy lists of output file locations. The domain parameters of interest for configuring the model itself, inside the ALLINC section, are:

- TLM0D – longitude of the centre point of the domain,

- TPH0D – latitude of the centre point of the domain,
- WBD – western boundary of the domain with respect to the centre point (always less than 0),
- SBD – southern boundary of the domain with respect to the centre point (always less than 0),
- DLMD – longitudinal model grid resolution,
- DPHD – latitudinal model grid resolution,
- DTB – time step of the model, which depends on DLMD and DPHD values by means of the Couranf-Friedrichs-Lewy (CFL) criteria,
- LM – the number of vertical levels.

Another set of commonly changed model parameters are dimensions of the model grid. These are grouped in the PARMETA section of the configuration file:

- IM – the number of mass grid points along the first row, essentially half of the total number of grid points in the west-east direction, due to the horizontal staggering of mass and wind points,
- JM – the number of rows in the north-south direction.

These parameters also influence the number of processes and the topology of the MPI parallel execution.

The rest of the parameters in the configuration file specify paths for input, output and intermediate files. With these paths defined during configuration, a significant reduction in disk space usage was achieved, as the data files no longer need to be copied together with the code, and no longer have to be in fixed relative locations.

The DREAMCLIMATE service is deployed during the first VI-SEEM development access call at the PARADOX high-performance computing cluster [12], hosted by the Scientific Computing Laboratory, Center for the Study of Complex Systems of the Institute of Physics Belgrade. This cluster is part of the VI-SEEM infrastructure, and consists of 106 working nodes. Working nodes (HP ProLiant SL250s Gen8) are configured with two Intel Xeon E5-2670 8-core Sandy Bridge processors, at a frequency of 2.6 GHz and 32 GB of RAM. The total number of CPU-cores available in the cluster is 1696, and each working node contains an additional GP-GPU card (NVIDIA Tesla M2090) with 6 GB of RAM. The peak computing power is 105 TFlops. The PARADOX provides a data storage system, which consists of two service nodes (HP DL380p Gen8) and 5 additional disk enclosures. One disk enclosure is configured with 12 SAS drives of 300 GB each (3.6 TB in total), while the other four disk enclosures are configured each with 12 SATA drives of 2 TB (96 TB in total), so that the cluster provides around 100 TB of storage space. Storage space is distributed via a Lustre high-performance parallel file system that uses Infiniband QDR interconnect technology, and is available on both working and service nodes.

Although the DREAMCLIMATE code is a copyright-protected software, it can be obtained for research purposes with the permission of the principal investigator (S. Ničković). Therefore, the DREAMECLIMATE service source code is only internally available at the VI-SEEM code repository [13], as well as a module at the PARADOX cluster software repository. Transfer of the software to third parties or its use for commercial purposes is not permitted, unless a written permission from the author is received.

4. Produced datasets and results. Using the DREAMCLIMATE service at PARADOX during the first VI-SEEM call for production use of resources and services, we produced a dataset with the aerosol optical thickness and surface dust concentration for the one-year period. We selected the year 2005 for this analysis, which serves as an example and demonstrates usability of DREAMCLIMATE service. The dataset covers wide region of North Africa, Southern Europe and Middle East in 30 km horizontal resolution with 28 vertical levels, and is made publicly available via the VI-SEEM data repository [14].

In addition to this initial dataset, we also produced a dataset with a higher resolution of 15 km for the same region and period of time. The global mean DREAMCLIMATE-modelled dust concentration for year 2005 is presented in Fig. 4.1.

Using the human health impact function introduced in Refs. [15, 16], we can relate the changes in pollutant concentrations to the changes in human mortality, and estimate the global annual premature mortality due to airborne desert dust. For this, we use as a baseline the mortality rate estimated by the World Health Organization (WHO) Statistical Information System on the country-level based on the International Classification of Diseases 10th Revision (ICD-10) classification, and regional data from the WHO Global burden of disease for countries with no data. Population statistics we used for the year 2005 is based on the United Nations Department of Economic and Social Affairs (UNDES 2011) database, while gridded global population numbers
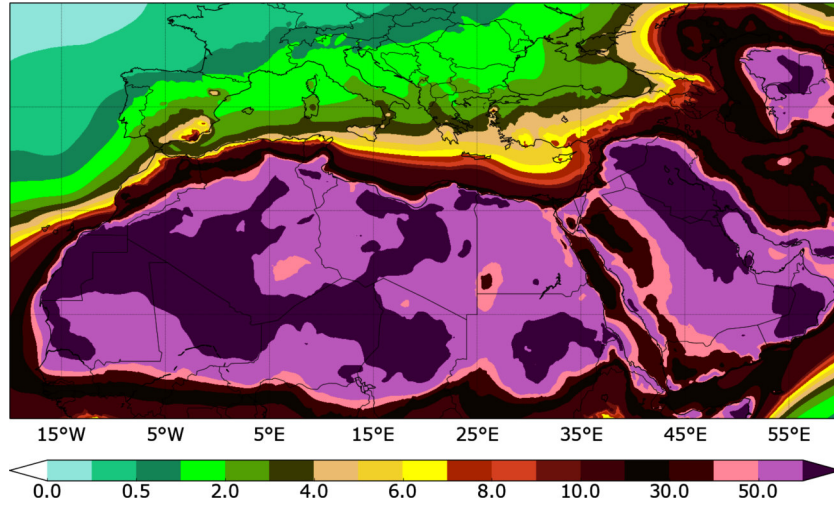
FIG. 4.1. *Calculated mean dust concentrations in* $\mu g/m^3$*, obtained from the DREAMCLIMATE model. The model integration area covers region of North Africa, Southern Europe and Middle East, with 15 km horizontal resolution in 28 vertical levels for the year 2005.*

TABLE 4.1
Total CPD and LC premature mortalities for the threshold concentrations between 0 and 10 $\mu g/m^3$*.*

Baseline concentration (in $\mu g/m^3$)	0	5.0	7.5	10
CPD premature mortality (in thousands)	765	615	567	524
LC premature mortality (in thousands)	14.8	10.2	9.1	8.4

are taken from the Columbia University Center for International Earth Science Information Network (CIESIN) database. We used the population cohort of 30 years and older in the health impact function.

Applying the health impact function to the considered population, the DREAM model output suggests a significant contribution of desert dust to premature human mortality. For the global background of dust concentration of 7.5 $\mu g/m^3$ i.e., threshold below which no premature mortality occurs, the estimated premature mortality (per grid cell) by cardiopulmonary disease (CPD) and lung cancer (LC) is illustrated in Fig. 4.2. In total, around 570,000 premature deaths in the model domain are predicted to occur during a one-year period, as a negative consequence of dust. According to our results, top five countries with the highest induced CPD-mortality in the year 2005 are: Egypt with 74,000; Iraq with 67,000; Iran with 50,000; Nigeria with 46,000; Sudan with 45,000. On the other hand, top five countries with the highest induced LC-mortality in the same year are: Iraq with 1,200; Iran with 900; Sudan with 800; Egypt with 800; Uzbekistan and Turkey with 500 premature deaths each.

We also investigated the sensitivity of our results on the value of the threshold concentrations, which is above assumed to be 7.5 $\mu g/m^3$. Table 4.1 gives the obtained total CPD and LC premature mortalities for the threshold concentrations between 0 and 10 $\mu g/m^3$. This analysis is presented to showcase capabilities of the model and the developed DREAMCLIMATE service, and can be efficiently used to study desired regions and time periods if the required input data are provided.

5. Conclusions. Using the VI-SEEM project infrastructure and services, we have successfully re-factored the DREAM atmospheric model. We have developed and implemented the DREAMCLIMATE service, which is tuned for usage on high-performance computing infrastructures available today. In order to demonstrate a typical use-case, we have produced a dataset with the aerosol optical thickness and surface dust concentration for the one-year period for the wide region of North Africa, Southern Europe and Middle East. We have used
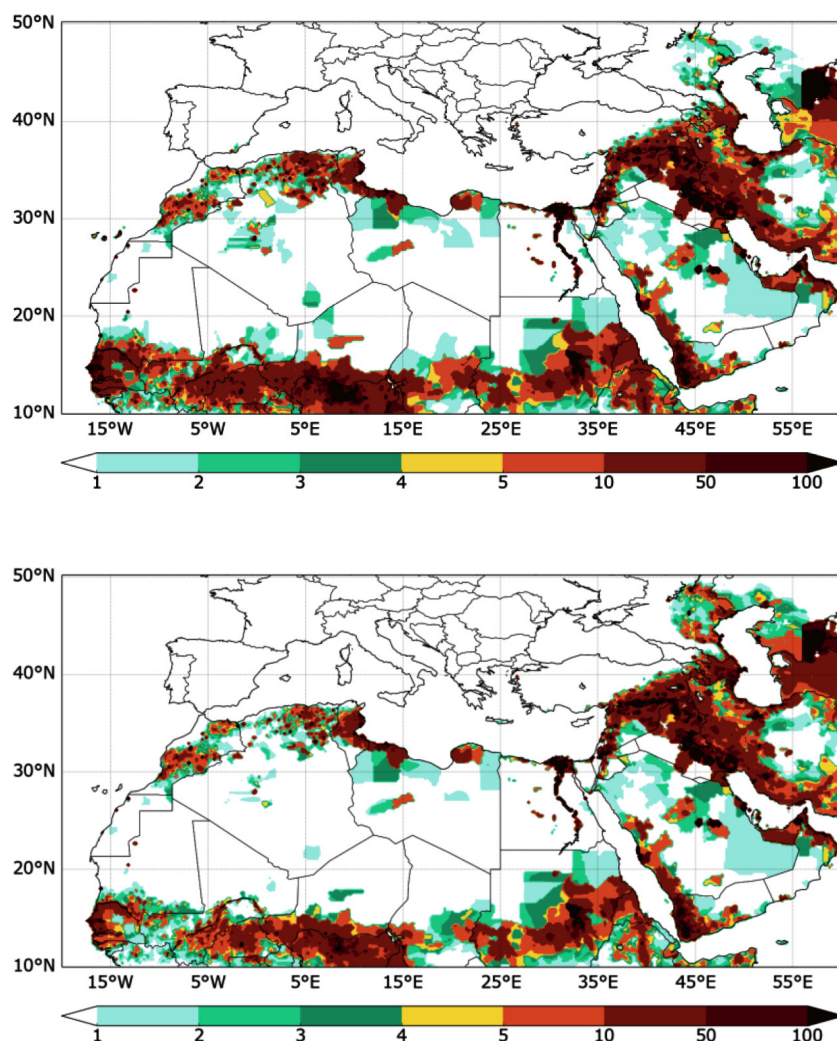
FIG. 4.2. *Estimated global premature mortality per grid cell by cardiopulmonary disease (top) and lung cancer (down) due to the long-term exposure to desert dust with an aerodynamic diameter smaller than 2.5 μm, calculated by the VI-SEEM DREAM-CLIMATE service.*

both the 30 km and the 15 km horizontal resolution, with 28 vertical levels. To showcase how results of the DREAMCLIMATE service can be applied, using the human health impact function and calculated global fine particulate matter concentrations, we have estimated the premature mortality caused by the long-term exposure to airborne desert dust with an aerodynamic diameter smaller than 2.5 μm for the year 2005 in the considered region. The results show that the large total number of premature deaths (around 570,000) in the model domain is mainly due to cardiopulmonary disease, but a significant number of deaths is also caused by lung cancer. The model also shows high sensitivity of the results on the threshold concentration, which is a significant parameter of relevance to public health.

REFERENCES

[1] S. Ničković, G. Kallos, A. Papadopoulos and O. Kakaliagou, *A model for prediction of desert dust cycle in the atmosphere*, J. Geophys. Res., 106 (2001), pp. 18113-18129.

[2] D. Yin and W. A. Sprigg, *Modeling Airbourne Mineral Dust: A Mexico - United States Trans-boundary Perspective*, in Southwestern Desert Resources, W.Halvorson, C. Schwalbe and C. van Riper, eds., University of Arizona Press, Tucson, AZ, (2010), pp. 303–317.

[3] D. Yin, S. Nickovic and W. A. Sprigg, *The impact of using different land cover data on wind-blown desert dust modeling results in the southwestern*, Atmos. Environ., 41 (2007), pp. 2214–2224.

[4] W. A. Sprigg, S. Ničković, J. N. Galgiani, G. Pejanović, S. Petković, M. Vujadinović, A. Vuković, M. Dacić, S. DiBiase, A. Prasad and H. El-Askary, *Regional dust storm modeling for health services: The case of valley fever*, Aeolian Res., 14 (2014), pp. 53–73.

[5] A. Vuković, M. Vujadinović, G. Pejanović, J. Andrić, M. R. Kumjian, V. Djurdjević, M. Dacić, A. K. Prasad, H. M. El-Askary, B. C. Paris, S. Petković, S. Ničković and W. A. Sprigg, *Numerical simulation of "an American haboob"*, Atmos. Chem. Phys., 14 (2014), pp. 3211–3230.

[6] D. Vudragović, P. Jovanović and A. Balaž, *VI-SEEM Virtual Research Environment*, 10th RO-LCG Conference, Sinaia, Romania, 26-28 October 2017.

[7] D. Giannadaki, A. Pozzer and J. Lelieveld, *Modeled global effects of airborne desert dust on air quality and premature mortality*, Atmos. Chem. Phys., 14 (2014), pp. 957–968.

[8] Z. I. Janjić, *The Step-Mountain Coordinate: Physical Package*, Mon. Wea. Rev., 118 (1990), pp. 1429-1443.

[9] Z. I. Janjić, *The Step-mountain Eta Coordinate Model: Further developments of the convection, viscous sublayer and turbulence closure schemes*, Mon. Wea. Rev., 122 (1994), pp. 927-945.

[10] S. Ničković and S. Dobricić, *A model for long-range transport of desert dust*, Mon. Wea. Rev., 124 (1996), pp. 2537–2544.

[11] Grid Analysis and Display System (GrADS),
http://cola.gmu.edu/grads/

[12] V. Slavnić, *Overview of Grid and High Performance Computing activities in Serbia*, 7th RO-LCG Conference, Bucharest, Romania, 3–5 November 2014.

[13] DREAMCLIMATE code at the VI-SEEM Code Repository,
https://code.vi-seem.eu/petarj/dreamclimate

[14] DREAMCLIMATE datasets at the VI-SEEM Data Repository,
https://repo.vi-seem.eu/handle/21.15102/VISEEM-86

[15] S. C. Anenberg, L. W. Horowitz, D. Q. Tongand and J. J. West, *An estimate of the global burden of anthropogenic ozone and fine particulate matter on premature human mortality using atmospheric modeling*, Environ. Health Perspect., 118 (2010), pp. 1189-1195.

[16] J. Lelieveld, C. Barlas, D. Giannadaki and A. Pozzer, *Model calculated global, regional and megacity premature mortality due to air pollution*, Atmos. Chem. Phys., 13 (2013), pp. 7023-7037.