REQUEST FOR A SPECIAL PROJECT 2026–2028

MEMBER STATE: The Netherlands, KNMI………………………………………………………………..…………

Principal Investigator[1]: Dr. Ir. Natalie E Theeuwes…………………………………………………………..……

Affiliation: Royal Netherlands Meteorological Institute (KNMI).……………………..……

Address: Utrechtseweg 297
3731 GA De Bilt
……………………………………………………………………..

Other researchers: Ulf Andrae (SMHI, Sweden), Eric Bazile (Meteo France, France), Maarten van Reeuwijk (ICL, UK), Jean Pierre Chaboureau (University of Toulouse, France), Björn Maronga, Nooshin Nowzamani (University of Hannover, Germany), Melanie Rochoux (CERFECS, France), Fredrik Jansson, Stephan de Roode (TU Delft, the Netherlands), Dirk Lauwaet (VITO, Belgium), Jan Mateu (BSC), Clara García-Sánchez (TUDelft, The Netherlands)

……………………………………………………………………………………………………..……

Project Title: …UrbanAIR…………………………………..………………………………………………..……

……………………………………………………………………………………………………..……

To make changes to an existing project please submit an amended version of the original form.)

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SP** ………….	
Starting year: (A project can have a duration of up to 3 years, agreed at the beginning of the project.)	2025 if possible; otherwise 2026	
Would you accept support for 1 year only, if necessary?	YES ☒	NO ☐

Computer resources required for project year:		**2026**	**2027**	**2028**
High Performance Computing Facility	[SBU]	300 M	300 M	300 M
Accumulated data storage (total archive volume)[2]	[TB]	200	400	700

EWC resources required for project year:		**2026**	**2027**	**2028**
Number of vCPUs	[#]			

[1] The Principal Investigator will act as contact person for this Special Project and, in particular, will be asked to register the project, provide annual progress reports of the project's activities, etc.
[2] These figures refer to data archived in ECFS and MARS. If e.g. you archive x GB in year one and y GB in year two and don't delete anything you need to request x + y GB for the second project year etc.
[3]The number of vGPU is referred to the equivalent number of virtualized vGPUs with 8GB memory.

Total memory	[GB]			
Storage	[GB]			
Number of vGPUs[3]	[#]			

Continue overleaf.

Principal Investigator: Natalie E Theeuwes……………………………………

Project Title: UrbanAIR.………………………………………………

Extended abstract

UrbanAir general project description

The goal of UrbanAIR is to develop a new digital twin that supports decision-makers in urban areas to deal with urban design dilemmas in atmospheric heat and air quality to maximise the health and socio-economic well-being of its citizens affected by climate change. It will provide critical tools for climate adaptation and hazard control through urban design and planning, including very high-resolution model components of the urban atmosphere. UrbanAIR is designed by a consortium that covers the full value chain to revolutionize digital twin platforms by starting from the perspective of the end user. Through co-creation with the end users and a balanced evaluation of the decision criteria, the overall objective of UrbanAIR is to yield a dynamic, user-friendly infrastructure integrated into the Destination Earth infrastructure that empowers municipalities and industries to face urgent urban climate risks. The scales in the atmospheric models in UrbanAIR cover the full range from the regional to the neighbourhood level. This innovative multiscale approach is achieved through the development of software interfaces for the modular coupling of atmospheric models. AIbased emulators allow for the acceleration of these computationally expensive models, which, together with the application of advanced data assimilation techniques, allows the quantification of risks and uncertainties for the UrbanAIR scenarios. Corresponding behavioural models simulate the human response to changes in climate and associated hazards. The resulting scenarios form the input to the objective evaluation of the criteria for decision-making. With these science-based tools for scenario simulation of natural and human behaviour, reliable risk assessment, and balanced decision analysis, UrbanAIR will develop tools and the infrastructure to support decision-makers in cities. This will pave the way for effective climate adaptation by developing tools for a safer, healthier, and more resilient future.

Plan for numerical modelling

One of the large parts of UrbanAIR is running through a hierarchy of models, going from the regional scale to the street scale (figure 1). Using the output of the extremes and climate DT of Destination Earth (DestinE; Wedi et al., 2025), UrbanAIR will create the possibility to downscale weather and climate to street-scales. This is needed to answer user questions related to urban planning, hazard response and climate adaptation.

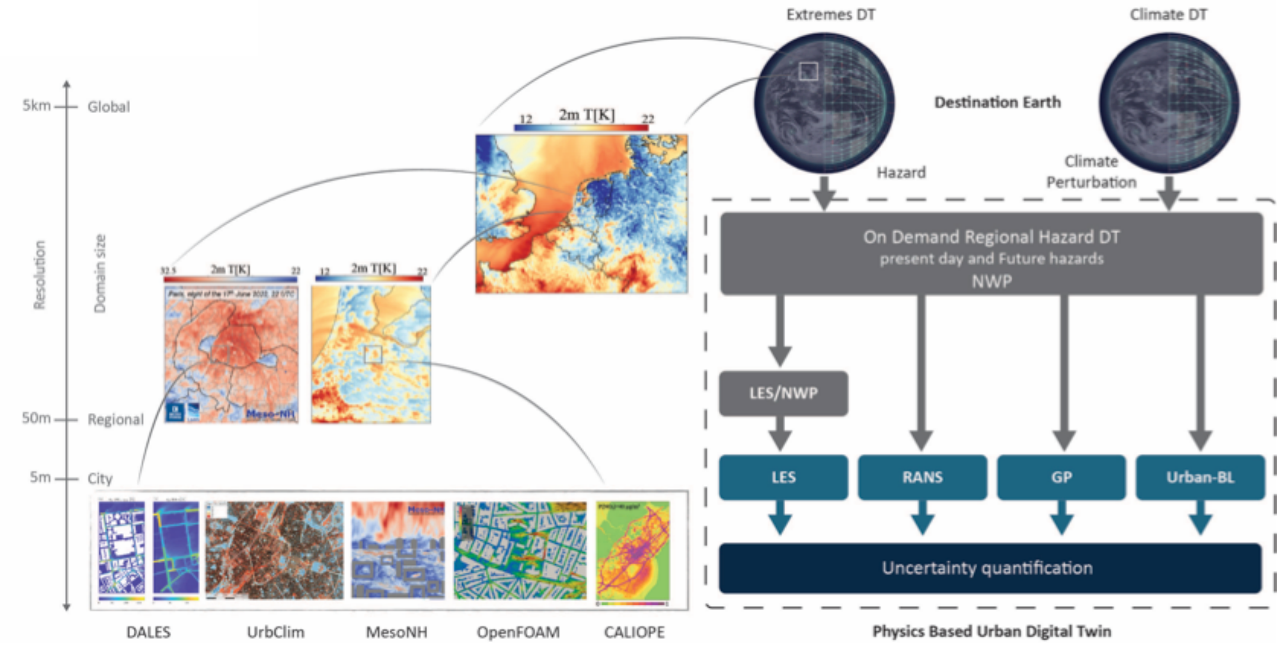

Figure 1 Schematic of the proposed hierarchy of models used in UrbanAIR

For each user question a user journey will be defined with a corresponding modelling chain. For example, a user might want to respond to an air quality hazard event, which would need a downscaling towards the city scale. However, if an urban planner wants to know how to reorganise a part of a city for climate adaptation. Several street-scale scenario simulations are necessary. To be able to facilitate these different modelling chains, several models need to be developed to (1) work at the scales they are intended for in this project, (2) be able to be coupled to one another in a flexible workflow and (3) be able to produce simulations of scenario's useful to answer user requirements.

The two global digital twins (DTs) developed in Destination Earth: one focused on climate adaptation and another on weather-induced extremes, are both targeting horizontal grid-spacing of 3–5 km. While this is suitable for synoptic-scale processes, it remains insufficient to resolve mesoscale-to-microscale phenomena, particularly in complex urban environments, where street-scale processes are more relevant. Here, building-resolving resolutions (~5 m) might be required.

However, direct nesting from global DTs to microscale LES models is infeasible due to turbulence spin-up issues and the so-called "terra incognita" of atmospheric modelling at sub-kilometre scales. To address this gap, DestinE's DE_330 initiative introduces regional downscaling at ~500 m resolution using models such as HARMONIE-AROME and AROME. An intermediate resolution (100/50 m) is needed to bridge this nesting hierarchy, though this significantly increases computational demands and limits temporal/spatial coverage. These city-scale simulations can be done using a numerical weather prodection (NWP) as used in DestinE's DE_330, but at a grid spacing of ~200/100 m. Or by large scale Large Eddy Simulation models.
UrbanAIR proposes a multiscale, multi-model simulation framework to resolve urban atmospheric processes across scales. This includes:

1. **Numerical Weather Prediction (NWP):** Regional mesoscale models (e.g., AROME, HARMONIE-AROME) operating at ~100 m to a few kilometers. These models provide regional-scale, meteorologically consistent simulations suitable for downscaling, but cannot resolve urban geometries explicitly.
2. **Large Eddy Simulations (LES):** High-fidelity turbulence-resolving models (PALM, DALES, Meso-NH) capable of simulating unsteady flow around complex urban geometries via immersed boundary methods. LES remains computationally prohibitive for long-term or large-domain simulations.
3. **Reynolds-Averaged Navier-Stokes (RANS):** Computationally efficient, steady-state models suited for scenario-based assessments and urban design applications. RANS lacks the ability to capture transient and turbulent structures.

The model developments needed for each of the model types are described below.

1. Numerical Weather Prediction (NWP)

- **Turbulence parameterization:** Implementation and refinement of advanced three-dimensional turbulence closure schemes (e.g., AROME) are necessary to improve horizontal and vertical mixing and subgrid-scale flux representation at these finer scales. LES results will serve as a benchmark to evaluate when high-fidelity schemes are warranted.
- **Boundary turbulence spin-up:** Development and evaluation of perturbation-based boundary forcing techniques to generate realistic turbulent and convective structures at domain boundaries, with reduced dependence on large-domain simulations or extensive nesting hierarchies.
- **Sub-hourly boundary condition exchange:** Enhancements to the model infrastructure to allow for high-frequency (e.g., timestep-resolved) boundary updates during nesting from coarser (~750 m) to finer (~100 m) scales.
- **Climate signal integration (PGW):** Construction of modified boundary conditions incorporating climatological anomalies from the Climate DT or CMIP6 projections, enabling pseudo-global warming simulations of future extremes.
- **LES boundary support:** Generation of dynamically consistent, high-resolution output fields to serve as initial and lateral boundary conditions for LES and RANS models operating at finer resolutions.

2. Large Eddy Simulation (LES) Models

- **Boundary condition coupling:** Development of robust interfaces for ingesting high-resolution, time-resolved boundary conditions derived from 100–200 m NWP simulations, including turbulence initialization and lateral inflow spin-up strategies.
- **Model optimization and scalability:** Enhancements in computational performance and parallel efficiency to support operational use, including potential model acceleration via reduced physics or adaptive meshing strategies.
- **Pseudo-global warming capability:** Implementation of PGW approaches within LES frameworks to simulate future extreme events under altered climate forcings by perturbing initial and boundary conditions.
- **Urban geometry ingestion:** Automated pipelines for integrating high-resolution urban morphology and surface data (e.g., building footprints, land use) using standardized formats across use cases.

- **Validation with observational data:** Utilization of dense urban observational datasets (e.g., from Bristol campaigns) to calibrate LES outputs and improve representation of urban canopy drag, radiative transfer, and subgrid-scale turbulence.

3. Reynolds-Averaged Navier-Stokes (RANS) Models

- **Initialization from upstream models:** Use outputs from LES or NWP models to generate boundary conditions for steady-state urban flow simulations, with emphasis on replicating realistic wind and stability conditions.
- **Automated urban domain generation:** Develop workflows to create geometrically accurate RANS model domains using high-resolution building and land-use data derived from open-source geospatial databases.
- **Multi-fidelity model integration:** Participate in the construction of a comprehensive urban flow database combining LES and RANS simulations for multiple inflow directions and atmospheric stability classes, enabling reconstruction of high-resolution fields under varied scenarios.
- **Scenario-based applications:** Employ RANS models for computationally efficient analysis of steady urban flow fields, pollutant dispersion, and urban ventilation performance under diverse meteorological and planning configurations.

In addition, development of automated toolchains and scripting interfaces to facilitate data flow, model coupling, and task execution across scales and modelling paradigms need to be worked out. By systematically advancing each modelling component and ensuring cross-scale coupling, UrbanAIR will enable accurate, high-resolution atmospheric simulations tailored to urban climate extremes under both present and future conditions.

Note that this request for resources is solely to work together on the scientific developments necessary to run the model chains in UrbanAIR. For the piloting of the UrbanAIR digital twin dedicated resources will be requested on EURO-HPC's.

Description of the codes used

The ACCORD NWP code

The ACCORD NWP model [Seity et al., 2011, Gleeson et al., 2024] is used for operational weather prediction in more than 20 countries in Europe. It's also used in the Destination Earth project DE330 targeting extreme weather by sub kilometric simulations. For the current study we aim to apply the configurations AROME and HARMONIE-AROME on 100-200m resolution.

PALM

The PALM model is based on Fortran organised in a series of Fortran files. Within these files the differential equations of the model core are solved and discretised in space and time. Spatial discretisation is realised through finite differences on a cartesian grid and temporal discretisation is achieved with a third-order Runge Kutta scheme. The advection of dynamic and scalar quantities is realised using a fifth-order scheme after Wicker and Skamarock (2002). The Poisson equation can either besolved directly using Fast-Fourier Transform (FFT), a tri-diagonal matrix, and backwards FFT, or using an iterative multigrid method. The code is parallelised with a 2D domain decomposition and the communication between the subdomains is achieved by Message Passing Interface (MPI). In hybrid mode additional thread parallelization is realized in a hybrid approach using Open Multiprocessing (OpenMP). The PALM code is optimised to run on CPU. Data I/O is steered with NetCDF files. More information on the model, its implementation and contact information of the model developers can be found on the model website: https://palm.muk.uni-hannover.de/trac . Due to the efficient parallelization of PALM the code has been run on many supercomputer architecturessince 1998 (e.g. IBM Regatta, NEC-SX6, SGI Altix ICE, Cray XT4, CrayXC40). PALM is generally characterised by outstanding scalability characteristics. These can be further improved when the problem size increases with increasing number of PEs. Also, the total performance increases for larger model domains. Figure 1 shows a benchmark test on the Cray XC40 for a model domain of 21603 grid points and using up to 11,520 PEs. The figure x clearly shows the good performance even for high PE numbers.

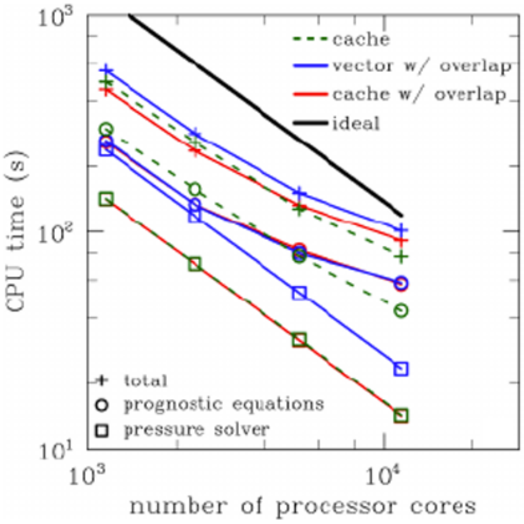

Fig.1. Scalability of PALM on the Cray XC40 supercomputer of HLRN. Simulations were performed with a computational grid of 2160^3 grid points (Intel-Ivy Bridge CPUs). Shown are data for up to 11520 PEs with cache (red lines) and vector (blue lines) optimisation and overlapping during the computation (FFT and tri-diagonal equation solver), enabled (dashed green lines). Measurement data are shown for the total CPU time (crosses), the prognostic equations (circles), and for the pressure solver (boxes).

DALES/uDALES

DALES, the Dutch Atmospheric Large Eddy Simulation, has been used to model clouds, convection, and turbulence since the 1990s, often in idealized settings with periodic boundary conditions [Heus et al, 2010]. Recently open lateral boundary conditions have been implemented [Liqui Lung et al., 2024], making it possible to nest DALES in a regional weather model. In this case DALES will be nested in HARMONIE. The finer scale version of DALES (uDALES) has been developed with immersed boundary method and a resolved energy balance to resolve buildings [Owens et al, 2024].

MesoNH

The Meso-NH code [Lac et al., 2018] is the mesoscale non-hydrostatic atmospheric model of the French research community, jointly developed by CNRS, Météo-France and the University of Toulouse. For the current study, we aim to run simulations at 100m resolution starting from HARMONIE and AROME forecasts.

OpenFOAM

OpenFOAM is a collocated unstructured finite-volume code written in C++. Within this project we plan to use different solvers, such as buoyantBoussinesqPimpleFoam and buoyantBoussinesqSimpleFoam to incorporate buoyancy effects into Reynolds-Averaged Navier-Stokes simulations, where velocity and pressure fields are coupled. OpenFOAM uses MPI to solve the governing equations in parallel and it can be compiled directly on either intel or gnu C++ compilers. In our computational fluid dynamics workflow, the following standard packages are used: FFTW, Lapack, Intel MKL, NetCDF, MPI, OpenFOAM.

In-house deep leaning model:

Our initial deep learning model, consisting of approximately 50 million parameters and trained on 504 simulations with 256×256×9 input resolution, required 162 hours on a single NVIDIA RTX 3090 GPU to train. The storage required for the baseline model and its corresponding dataset was approximately 7 GB. To scale this work to our full dataset of 8,640 simulations with a higher input resolution of 512×512×9, we estimate a storage requirement of roughly 300 GB. This includes both the extended dataset and larger model checkpoints. In addition, to support additional experimentation including hyperparameter tuning, exploration of different model architectures (e.g., spatial attention modules, GNNs, Transformers), and validation across multiple cities and atmospheric stability conditions, we request an initial allocation of 10,000 RTX 3090-equivalent GPU hours. If additional resources are required for extended model exploration, with more urban configurations, or increased input resolutions, we will submit a supplemental request based on the initial results.

Justification of SBU's and storage

To be able to answer the scientific questions and develop the model infrastructure, we request 300 Milion SBU's and 200 TB of storage per year. An approximation of how this will be subdivided between the different models and institutes is given below.

HARMONIE-AROME
SMHI
Investigation of 3D turbulence, perturbations at the boundaries, and running test cases
40 simulations @ 500 m 48 h forecasts * 90 000 SBU's = 3 600 000 SBU's + 4 TB of storage
80 simulations @ 200 m 24 h forecasts * 110 000 SBU's = 8 800 000 SBU's. + 8 TB of storage
80 simulations @ 100 m 24 h forecasts * 225 000 SBU's = 18 000 000 SBU's + 8 TB of storage
Total 30.4 Milion SBUs + 20 TB storage per year

KNMI
Future weather simulations, urban representation, and running test cases
40 simulations @ 500 m 48 h forecasts * 90 000 SBU's = 3 600 000 SBU's + 4 TB of storage
80 simulations @ 200 m 24 h forecasts * 110 000 SBU's = 8 800 000 SBU's. + 8 TB of storage
80 simulations @ 100 m 24 h forecasts * 225 000 SBU's = 18 000 000 SBU's + 8 TB of storage
Total 30.4 Milion SBUs + 20 TB storage per year

AROME
Investigation of 3D turbulence, urban representation, and running test cases
40 simulations @ 500 m 48 h forecasts * 90 000 SBU's = 3 600 000 SBU's + 4 TB of storage
80 simulations @ 200 m 24 h forecasts * 110 000 SBU's = 8 800 000 SBU's. + 8 TB of storage
80 simulations @ 100 m 24 h forecasts * 225 000 SBU's = 18 000 000 SBU's + 8 TB of storage
Total 30.4 Milion SBUs + 20 TB storage per year

PALM
Preprocessing 240 simulations *17 429 SBU's = 3 834 336 SBU's + 2 TB of storage
Production 6 simulations * 5 017 144 SBU's = 30 102864 SBU's + 64 TB of storage
Adjusted production 6 * 2 558 569 SBU's = 15 351 414 SBU's + 4 TB of storage
Total of 44.1 Milion SBU's + 70 TB storage per year

DALES
54 simulations @ 100m 24h forecasts * 840 000 SBU's = 45 400 000 SBU's + 20 TB of storage
For a domain with 2048x2048 points = 205x205 km

UDALES
14 simulations with 8192 cores for 24 hours = 49.5M SBUs + 5 TB of storage

Meso-NH
University of Toulouse
54 simulations @ 100m 24h forecasts * 840 000 SBU's = 45 400 000 SBU's + 20 TB of storage
For a domain with 2048x2048 points = 205x205 km

OpenFOAM
TUDelft
Investigating the effect of stability on urban RANS simulations uncertainty and developing a fast prediction tool for deployment
10 800 RANS simulations = 7 500 000 SBU''s + 10 TB of storage
Deep learning model = 10.000 *RTX 3090-equivalent GPU hours*

OpenFOAM
VITO
Investigating the effect trees on air quality, ventilation and heat stress:
10 800 RANS simulations = 7 500 000 SBU''s + 10 TB of storage

Workflows
UKRI-STFC
Workflows to define and test interfaces:
7 simulations of 24h on 4 nodes = 2 200 000 SBU''s + 5 TB of storage

References

Gleeson, E., Kurzeneva, E., De Rooy, W., Rontu, L., Martín Pérez, D., Clancy, C., et al. (2024). The Cycle 46 Configuration of the HARMONIE-AROME Forecast Model. Meteorology, 3(4), 354-390.

Heus, T., van Heerwaarden, C. C., Jonker, H. J., Pier Siebesma, A., Axelsen, S., Van Den Dries, K., et al (2010). Formulation of the Dutch Atmospheric Large-Eddy Simulation (DALES) and overview of its applications. *Geoscientific Model Development*, *3*(2), 415-444.

Lac, C., Chaboureau, J. P., Masson, V., Pinty, J. P., Tulet, P., Escobar, J., et al. (2018). *Overview of the Meso-NH model version 5.4 and its applications, Geosci. Model Dev., 11, 1929–1969*.

Liqui Lung, F. P., Jakob, C., Siebesma, A. P., & Jansson, F. R. (2023). Open Boundary Conditions for Atmospheric Large Eddy Simulations and the Implementation in DALES4. 4. *Geoscientific Model Development Discussions*, *2023*, 1-25.

Owens, S. O., Majumdar, D., Wilson, C. E., Bartholomew, P., & van Reeuwijk, M. (2024). A conservative immersed boundary method for the multi-physics urban large-eddy simulation model uDALES v2. 0. *Geoscientific Model Development*, *17*(16), 6277-6300.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., et al (2011). The AROME-France convective-scale operational model. *Monthly Weather Review*, *139*(3), 976-991.

Wedi, N., Sandu, I., Bauer, P., Acosta, M., Andersen, R. C., Andrae, U., et al. (2025). Implementing digital twin technology of the earth system in Destination Earth. *Journal of the European Meteorological Society*, *3*, 100015.

Wicker, L. J., and Skamarock, W. C. (2002). Time-splitting methods for elastic models using forward time schemes. *Monthly weather review*, *130*(8), 2088-2097.