On Application of Artificial Neural Network Methods in Large-eddy Simulations with Unresolved Urban Surfaces

Igor Esau (Corresponding author)

G.C. Rieber Climate Institute of the Nansen Environmental and Remote Sensing Center

Thormohlensgt. 47, 5006, Bergen, Norway

Tel: 47-55-205-876 E-mail: igor.ezau@nersc.no

The research leading to these results has received funding from the European Union's Seventh Framework Programme FP/2007-2011 under grant agreement n°212520

Abstract

Micro-meteorological aspects of city comfort, land use management and air quality monitoring are rapidly growing areas of applications where environmental turbulence-resolving or large-eddy simulation (LES) models play the central role. Complex details of the urban surface morphology however remain unresolved or poorly resolved in the state-of-the-art LES due to severe limitations from computer facilities. The LES code LESNIC is applied in this study to simulate turbulent flow interaction with elements of the urban surface morphology. The study investigates a possibility to utilize a trained three layers' artificial neural network (ANN) to parameterize the flow-to-surface interactions in the coarse LES where surface features are unresolved. It is concluded that the ANN can be a robust predictor for scalar concentrations and components of the surface stress tensor in the urban sub-layer with unresolved scalar sources and surface morphology. It has been noted however that dynamic training of the ANN may require more computational resources than adequate refinement of the LES resolution.

Keywords: Urban meteorology, Large-eddy simulations, Artificial neural networks, Turbulence

1. Introduction

Meteorological applications, e.g. city comfort, land use management and air quality monitoring, require development of turbulence-resolving large-eddy simulation (LES) models as well as approaches to utilize them in simulations of turbulent flows over poorly resolved complex surface features. Figure 1 gives an example of such an urban surface obtained from the MEGAPOLI database of the Paris surface morphology (Sievinen et al., 2009). Although the digital elevation model of Paris is sufficiently smooth and large-scale as it can be easily noticed in Figure 1, the building elevations on the top of it cannot be resolved with a moderately fine mesh in LES as the buildings have a lot of the elevation variability on scales 100 m and less. Elsewhere in the world, more complicated surface features may be required to incorporate in the model (e.g. Hj Kamaruzaman & Mohd Hasmadi, 2009). The under-resolved surface structures can be incorporated in the model boundary conditions in three ways: traditional parameterization methodologies (e.g. Britter and Hanna, 2003; Bou-Zeid et al., 2009); adjoint optimisation model (e.g. Losch and Wunsch, 2003); and artificial neural network technique (ANN). The latter approach is addressed in this study. We will investigate applicability and utility of the three layers' ANN in a priori numerical experiments. ANN is a statistical technique, which solves a set of parametric function equations, called static nonlinear maps, with empirically tuned coefficients or weights to relate the model variables to unresolved quantities of interest. The weights are obtained in so-called training process searching for the global minimum of a cost function in multi-dimensional parameter space. ANN were found useful in many closely related applications, e.g. turbulence control (Lee et al., 1997; Ayhan et al., 2004), turbulent flow simulations (Panigrahi, et al., 2003; Hocevar et al., 2004), air quality monitoring (Morabito and Versaci, 2002; Mammarella et al., 2006). In particular, Kukkonen et al. (2003) evaluated performance of five ANN methods, a linear regression statistical model and a hydrodynamic model using measurements of urban chemical concentrations at two stations in central Helsinki from 1996 to 1999. On average, the ANN methods demonstrated better agreement with data than other types of models. Mammarella et al. (2006) reported the use of the ANN in the fully operational air quality monitoring system "ATMOSFERA".

In the past, limited computational resources restricted the LES applications to the case of turbulent flows over homogeneous flat surfaces. More interesting applications like urban "hot-spots" of anthropogenic impact on the Earth's system were not fully addressed. In this study, the LES code LESNIC developed at the Nansen Environmental and Remote Sensing Center (Esau, 2004) will be utilized with the ANN application to elements

of an urban surface.

2. Introduction

2.1 Artificial neural network in modeling of surface stress

Limitations on computational resources and a lack of knowledge to parameterize sub-grid scale processes in coarsely resolved simulations have motivated a number of attempts to make shortcuts between practically important but unresolved quantities, like concentrations in the urban sub-layer, and more readily simulated ones, like the mean wind. The ANN application approach exploits persistent dynamical patterns on resolved scales created by the unresolved interactions between fixed surface features and the turbulent flow. Suppose a certain feature of boundary conditions causes a perturbation in modelled variables. For instance, a building, unresolved with the given model mesh, perturbs the flow. In this case, one can introduce a statistical correction included as an additional artificial force into the model. Since the exact structure of this force is unknown, it is approximated with a matrix of stochastic empirical coefficients and a set of linear equations that is in fact the ANN. The coefficients are chosen in such a way to minimize the difference between the model and the observations.

Mathematically, the ANN can be considered as an optimal estimator of unresolved quantities using a finite time series of resolved-scale solutions (Moreau et al., 2006). In this study, a three-level ANN (one input layer of 16 neurons and two hidden layers of 4 an 1 neurons) with a back-propagation learning function and the tan-sigmoid transfer function is considered. The neuron distribution was selected optimal by trail-and-fail approach for the LES data. The ANN could be written as

$$q(x_i) = \sum_{m=1}^{M} C_{mi} \sum_{j=1}^{N} A_j \tanh\left(\sum_{k=1}^{N_k} B_{jk} X_m(x_k) + b_{jm} \right) + a_m \,. \tag{1}$$

Here $q(x_i)$ is a predicted quantity, e.g. a chemical concentration or velocity stress, given at positions x_i; $X(x_k)$ is the input vector of predictors given at positions x_k, which could be different from x_i. Other symbols in Eq. (1) denote empirical weights A_j, B_{jk}, C_{mi} and thresholds b_{jm} and a_m. The weights and thresholds are to be "trained", i.e. empirically fitted using one or another statistical approximation, on a sub-set of available data $q(x_i)$ and $X(x_k)$. Thus, the correspondence between x_i and x_k sets of positions will be implicitly built-in into the weights during this training process. This ANN feature will be used in its application to urban LES studies where the training process will implicitly account for the correspondence between the fixed positions of the unresolved surface features and the fixed positions of the mesh nodes in the model. In this particular application, the weights and thresholds are obtained by a conjugate-gradient method. We seek for minimization of the residual error on a sub-set of 5 min data sampled from the fine resolution LES run as the input vector $X(x_k)$ and the coarse resolution LES runs as the output vector $q(x_i)$. Finally, M, N are numbers of neurons on the hidden layers of the ANN and N_k is the number of neurons linked to each input vector element. Actual calculations have been run in MATLAB using neural network toolbox (*nnet*) by Demuth and Beale. In this sense, the application of the ANN in this study is similar to the approach chosen by Manjunatha et al. (2010). The ANN application requires preconditioning through the model feedback, as large predictor vector dimension leads to a classical ill-conditioned problem of the linear algebra (Hsieh and Tang, 1998).

A correct way to introduce the unresolved boundary conditions into the LES model is through the sub-grid scale model. The sub-grid scale model describes the integral effect of unresolved flow features on the resolved scales. Sarghini et al. (2003) pioneered application of ANN to dynamic sub-grid scale modelling. LES equations are obtained through filtering of Navier-Stokes equations. The filtering result in appearance of a sub-grid stress term

$$\tau_{ij} = \overline{u_i u_j}^{h} - \overline{u}_i^{h} \overline{u}_j^{h} - 2 l_{cut}^2 \left| \overline{S}_{ij}^{h} \right| \overline{S}_{ij}^{h} \,, \tag{2}$$

which must be modelled. Here the overbar with the superscript *h* denotes filtering operator at scale twice the grid scale while the grid filtering is done implicitly by the numerical scheme; $l_{cut} = C_s \Delta$ is the mixing length scale; C_s is the Smagorinsky constant; Δ is the LES mesh resolution. The strain rate tensor is

$$S_{ij} = \frac{1}{2} \left(\nabla_i u_j + \nabla_j u_i \right), \tag{3}$$

where ∇_i denotes a gradient in *i*-th direction. As Eq. (2) reveals, C_s is the parameter to be tuned to the specific unresolved features of the flow. Here we consider the following numerical experiments to test the ANN utility in the described problem of the flow around large unresolved obstacles below the first computational layer.

The LESNIC model run with developed atmospheric turbulent boundary layer with the geostrophic wind speed 4 m s^{-1} at latitude 45 degrees over homogeneous flat surface with the roughness length scale 0.01 m. The model domain was 1000 m by 750 m by 500 m in streamwise, spanwise and vertical directions with the mesh 64 by 48 by 64 grid nodes. The boundary layer evolved over one hour of model time. At the end of its evolution, the boundary layer occupied ½ of the domain depth. The last 3-dimensional state of the model was stored and later used as initial conditions to obtain additional 30 min runs both with homogeneous and heterogeneous surfaces. The simulations were run for truly neutral boundary layers with passive scalars without feedback effects on the flow. The lateral boundary conditions were periodic in all runs. At the surface, the turbulent flux of momentum was computed instantly and pointwise using the log-law and therefore did not account directly to the surface features. The ANN was utilized in the following way. The 3 layer back-propagation feeding ANN was trained on 5 min data from the 30 min additional LES runs described above. The input vector for the ANN was composed of components of the tensor in Eq. (3) and its module as

$$X = [S_{11}, S_{12}, S_{13}, S_{21}, S_{22}, S_{23}, S_{31}, S_{32}, S_{32}, |S_{ij}|]^T . \quad (4)$$

Since the tensor is symmetric, its off-diagonal components receive larger weights in the input vector.

Let us first consider the effect of the ANN in the Ekman layer simulations over flat, homogeneous surface. Figure 2 compares histograms of C_s computed with Eq. (2) only and with Eq. (2) following the ANN correction. Equation (2) represents a dynamic mixed model as described in Esau (2004). This model uses significant resolved scale information, obtained through an explicit discrete Gaussian filtering, to estimate the magnitude of the sub-filter stress. This so-called scale-similarity information becomes insufficient when a large fraction of the turbulent kinetic energy is unresolved. The mismatch between the required within the inertial interval of scales and the resolved spectral energy fluxes is compensated by lowering C_s in the dynamic-mixed model in Eq. (2) (for details see Pope, 2004). However, too low C_s results in numerical instability. Numerous ad hoc approaches have been proposed to address the low C_s values. Desirable stabilizing modifications of the C_s distribution introduced by the chosen ANN are clearly seen in Figure 2. The small values of C_s, which destabilize the simulations, were almost eliminated and too large values ($C_s > 0.4$) occurred considerably less frequent in the flow during the analysed 30 min interval of time. Thus, we can conclude that sub-grid modelling with ANN correction may improve performance and stability of LES even in simulations over homogeneous surfaces.

The next step is to introduce a surface heterogeneity in form of rough spots that represent completely unresolved obstacles or just surface heterogeneity of any other kind. This experiment is LES H1 in Table 1. The surface properties remind unchanged during the simulations. Hence ANN can be trained to correct the simulations specific features of the perturbed flow. This training is achieved in the study trough the run with twice as fine resolution, which served as observations for the training process. The run LES-H1 had only surface roughness modified. The roughness was set to 5 m in a square of ¼ of the domain surface area; whereas the rest of the domain had roughness of 0.01 m. Figure 3 shows the mean values of C_s in the coarse resolution run without ANN (Figure 3a) and with ANN (Figure 3b) correction. The added value of the ANN correction is clearly seen in localized modification of C_s and therefore the surface sub-grid fluxes as related in Eq. (2). Figure 4 shows that ANN was able to improve the vertical profile of averaged C_s as well. ANN has automatically introduced modifications, which were previously left for *ad hoc* fitting as described in Balaras et al. (1995). The observed changes in the mean Smagorinsky coefficient profiles as well as in their spatial distribution and clipping of the extreme values are welcome modifications made through ANN implementation to LES closure problem over roughness heterogeneous surface.

At the next step, a cubic obstacle was defined in the boundary conditions of the coarse resolution run (LES-H2). The obstacle was under-resolved in the sense that only the 1st model level was modified to have immersed boundary conditions. The log-law was applied only at the upward facing facet of the obstacle, whereas the simple non-slip conditions were applied on the lateral facets. The obstacle in LES-H2 is of 12 m tall (2 computational levels) located in ¼ of the domain area. Thus, the LES-H1 run represented a resolved flow response on unresolved elements of urban morphology, and LES-H2 represents a corresponding response on semi-resolved elements. Figure 5 shows values of C_s obtained as in Figure 3 but for LES-H2 experiment. An impenetrable obstacle indeed has strong impact on C_s values and their spatial distribution. As in Glazunov (2009), C_s is distributed asymmetrically around the obstacle with larger values in the front (upstream) facet and smaller values at the rear (downstream) facet. Remarkably, ANN correction was able to account for the localization of the smallest C_s on the downstream sides of the obstacle and very close to the centre of the obstacle front facet. Those features were obtained in a fine resolution high quality LES by Glazunov (2009) but hardly reproduced by the dynamic closure in this LES. It is also worth noticing that the time needed to estimate

C_s with the ANN correction is about 2 orders of magnitude less than time used in the Glazunov's LES (personal communication) and an order of magnitude less than time used to compute it in the present high-resolution LES.

2.2 Artificial neural network in modeling of concentrations

One of the most important and interesting applications of ANN in conjunction with LES is the modelling of the scalar concentrations (pollution). The problem is to predict the concentration field $q(\vec{x},t)$ on the basis of the velocity field $\vec{u}(\vec{x},t)$, which is known with a limited accuracy, and $q(x_k,t-k\delta t)$, which is known only in a few positions x_k. Using the control LES run, the application was computed for 1 min forecast of concentrations in the near field of a pollution source in the turbulent boundary layer. The forecast was computed 25 times over 25 min of model time. Initially, the ANN has been trained on 5 min data in order to produce $q(\vec{x},t)$. To do the training, we use the same LES run assuming that the LES data supplant the real atmospheric observations. To avoid self-correlation, the data computed for 30 min before the forecast period have been used. In practice, the observations and the LES together should be used for the ANN training. The technology is known (e.g. Mamarella et al., 2006) but since it is not used here, we will not describe it. The input vector was composed of components of velocity and the historical value with δt = 1 min. This choice is to very much degree arbitrary. It could be motivated only by an expert opinion that it is difficult to get the observational data of higher frequency. Hence, the input vector was defined as $X=[u_1(t),u_2(t),u_3(t),q(t-\delta t)]^T$ for each mesh node x_k. Figure 6 shows concentration field averaged over 30 min and within the layer $\pm$15 m above and below the injection source. The source was located at the height of 18 m above the surface so that the polluted wake meanders in three dimensions. The spots of high concentration are created by the wake meandering in the vertical direction. A nontrivial feature is a rarefied scalar concentration directly downstream from the source on the wake axes recognized also in Babatunde and Enger (2002). The ANN predicts robustly short-term fluctuations of the concentration but produce some noise in the absence of a real signal. The overall weighted correlation between the simulated concentration and the ANN averaged forecast was 0.92 in this experiment. The forecast interval of 1 min would probably seem too short for a practical use but the reader should keep in mind that we are analysing the ANN as an addition to the LES hydro-dynamical model. Thus, the ANN role is utilized only for predictions on unresolved intervals of time. The advection across larger distances and longer time scales are explicitly resolved in the model. Indeed over 1 min event with the wind speed of 4 m s^{-1} a polluted air volume would be transported over 240 m or more than 10 grid intervals in the present experiment. With typical wind of 10 m s^{-1}, the same volume will be transported over 600 m or 50 grid intervals.

3. Discussion: Compatibility of LES and ANN approaches

Artificial Neural Networks (ANN) were seen in several publications (e.g. Azamathulla et al., 2008) as a promising method to reduce dimensionality of models required to describe complex properties of the flow dynamics, dispersion in the flow and the flow interaction with the surface. These promises had not been met. One, probably insurmountable problem has been pointed out by Milano and Koumoutsakos (2002). To present it, these authors noticed analogy between the ANN and the Proper Orthogonal Decomposition (POD) techniques. The linear ANN could be used for the POD analysis. Alternatively, the POD analysis could be considered as another training method to reduce dimensionality of the problem. In the case of POD, the training procedure is simply singular value decomposition retaining N first eigenvectors. As with POD, the dimension of the problem for any realistic flows is too large and untreatable. The flow domain is to be decomposed on small independent sections to circumvent the problem. If those sections are essentially interdependent, the decomposition becomes loose in the sense that essential part of the information will be lost.

In the ANN, the computational expenses are concentrated on calculation of the weight matrix W. Let us show that the POD and the linear ANN are equivalent. The POD could be written as

$$\begin{aligned} a(t) &= \Phi^T \cdot u(t) \\ [u(t)]^{FFT} &= \Phi \cdot a(t) \end{aligned} \tag{5}$$

where $a(t)$ is the eigenvalue vector of the size $n_x n_y n_z n_{variables}$, Φ is the eigenvector matrix, and $u(t)$ is the transposed variables' vector. The ANN could be formulated as

$$\begin{aligned} x &= W_1 \cdot u(t) \\ [u(t)]^{FFT} &= W_2 \cdot x \end{aligned} \tag{6}$$

where W_1 and W_2 are the ANN weight matrices of the size $n_x n_y n_z n_{variables}$ to be found from a mini-max problem.

Now obvious that while the POD problem can be hugely accelerated by choice of the orthogonal harmonic

functions, the Fast Fourier Transformation (FFT) and the fast partial decomposition algorithms, the ANN problem should involve essentially slow iterative optimization problem solvers. The problem complexity increase with non-linear networks that can be expressed as

$$\begin{aligned} x &= W_1 \cdot \tanh(W_2 \cdot u(t)) \\ [u(t)]^{FFT} &= W_4 \cdot \tanh(W_3 \cdot x) \end{aligned} \tag{7}$$

where already 4 weight matrix are to be obtained in the training process. Milano and Koumoutsakos (2002) reported the largest trained network of just $n_x n_y n_z$ = 42 x 39 x 16 = 26 208 elements, which result in the size of the weight matrix W_1 of 6 710 272 elements. Large dimensionality of the optimization problem makes it not only computationally difficult but also ill-conditioned. Thus, any flow non-locality virtually prohibits the application of the ANN. Even neglecting non-locality, the problem remains untreatable.

There is however a serious deficiency in the Milano and Koumoutsakos analysis. They considered a global optimization problem in the sense that they reduce dimensionality within each of sub-domains solving the global mini-max problem. In the application approach proposed and tried in this study, the low-dimensional problem has been formulated for each mesh cell. In fact, the resolved-scale dimension of the problem here is $n_x n_y n_z = 1$. Even admitting non-locality of the surface element impact, the dimension $n_x n_y n_z$ remains O(10) or typically about O(100 m) in the length scale terms and therefore treatable. We simply use the scaling of the problem imposed by the specifics of the urban-type environment where the buildings are different but not particularly variable in size. Thus, the Milano and Koumoutsakos approach is insightful. But the ANN application for the considered urban problems is not obviously disadvantages.

Another approach to the ANN applications has been proposed by Jimenez et al. (1997). It refers to rigorous LES formalism, exemplified in e.g. Cook and Riley (1994) model. The formalism demands to express the sub-grid quantity q' through resolved quantity $\overline{q}$ in a universal scale-similar manner. Cook and Riley exploited the fact that any quantity defined at [0;1], i.e. normalized q, and subjected to stochastic turbulent fluctuations u with the Gaussian distribution function should have the β-distribution

$$\beta(x;\overline{q},\sigma_q^2) = x^{a-1}(1-x)^{b-1}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}, \quad a = \overline{q}\left(\frac{\overline{q}(1-\overline{q})}{\sigma_q^2}-1\right), b = \frac{a}{\overline{q}} - a, \tag{8}$$

where $\Gamma(a)$ is the gamma function of Euler With knowledge of the β-distribution, the true values of concentration and its ensemble-averaged value over a grid cell are related as

$$< q >= \int_{\Delta} q(x)\beta(x;\overline{q},\sigma_q^2)dx. \tag{9}$$

In principle, Equations (8)-(9) allow estimations of the concentration within the model grid cell with a minimal root mean square error. As Jimenez et al. (1997) argued, the envisaged model provides excellent results for a homogeneous flow even at high Re. The situation is however more complicated in heterogeneous urban sub-layer. Here, space-time filtering applied to compute σ_q^2 and ultimately $< q >$ is not valid any more due to presence of the vertical walls. The resolved variations of the concentration are far too small to represent scale similarity.

4. Conclusions

The micro-meteorological modelling tool, large-eddy simulation model, is well established for the research purposes. In order to utilize this tool in practical applications, unresolved surface and emission features with effect on the resolved turbulence have to be represented. The study addressed the problem investigating utilization of the artificial neural networks with the turbulence resolving large-eddy simulation (LES) model LESNIC. *A priori* analysis has been applied for idealized, relatively coarse resolution numerical experiments, namely, the experiments with turbulent neutrally stratified boundary layers over homogeneous surface, the surface with very large roughness patch and the surface with an obstacle. It is shown that the filtering property of the three layers' ANN is useful in improvement of numerical stability and physical consistency of the dynamic sub-grid closure in LES. The 3-layer tan-sigmoid ANN with back-propagation training over rather short interval of 5 min outperformed the usual engineering clipping approach to obtain the dynamic values of the Smagorinsky constant. Contrary to the scepticism induced by the earlier applications of the ANN to the sub-domain reduction in dimensionality, the ANN application for each of dynamical sub-regions (usually just of a few computational cells) in the LES domain was found promising and computationally efficient. The key to the problem is the fact that the small scale turbulent perturbations quickly decay. Thus the radius of locality of the problem is small and

limits the size of the weight matrices in the ANN. The ANN is shown to be a robust predictor for the pollutant scalar concentration in the urban sub-layer with unresolved scalar sources.

References

Ayhan, T., Karlik, B., & Tandiroglu, A. (2004). Flow geometry optimization of channels with baffles using neural networks and second law of thermodynamics. *Comput. Mech.*, 33, 139–143.

Azamathulla, H. Md., Ab Ghani, A., Zakaria, N. A., Kiat, C. C., & Siang, L. C. (2008). Knowledge Extraction from Trained Neural Network Scour Models. *Modern Applied Science*, 2(4), 52.

Babatunde, A. J., & Enger, L. (2002). The role of advection of fluxes in modelling dispersion in convective boundary layers. *Quart. J. Roy. Meteorol. Soc.*, 128, 1589-1607.

Balaras, E., Benocci, C., & Piomelli, U. (1995). Finite-difference computations of high Reynolds number flows using the dynamic subgrid-scale model. *Journal Theor. Comput. Fluid Dynamics*, 7(3), 207-216.

Britter, R. E., & Hanna, S.R. (2003). Flow and dispersion in urban areas. *Annu. Rev. Fluid Mech.,* 35, 469–496

Bou-Zeid E., Overney J., Rogers B.D., & Parlange, M. B. (2009). The Effects of Building Representation and Clustering in Large Eddy Simulations of Flows in Urban Canopies. *Boundary-Layer Meteorol.*, 132(3), 415-436

Cook, A. W., & Riley, J. J. (1994). A sub-grid model for equilibrium chemistry turbulent flows. *Phys. Fluids*, 6, 2868.

Esau, I. (2004). Simulation of Ekman boundary layers by large eddy model with dynamic mixed subfilter closure, *Environmental Fluid Mechanics*, 4, 273–303.

Jimenez, J., Linan, A., Rogers, M. M., & Higuera, F. J. (1997). A priori testing of sub-grid models for chemically reacting non-premixed turbulent shear flows. *J. Fluid Mech*., 349, 149-171.

Hj Kamaruzaman, J., & Mohd Hasmadi, I. (2009). Mapping and Quantification of Land Area and Cover Types with Landsat in Carey Island, Selangor, Malaysia. *Modern Applied Science*, 3(1), 42.

Hocevar, M., Sirok, B., & Grabec, I. (2004). A Turbulent-Wake Estimation Using Radial Basis Function Neural Networks. *Flow Turbul. Combustion,* 74, 291–308.

Hsieh, W. W., & Tang, B. (1998). Applying neural network models to prediction and data analysis in meteorology and oceanography. *Bull. Am. Meteorol. Soc.*, 79, 1855-1870.

Glazunov, A. V. (2009). Large-eddy simulation of turbulence with the use of a mixed dynamic localized closure: Part 2. Numerical experiments: Simulating turbulence in a channel with rough boundaries. *Izvestiya Atmospheric And Oceanic Physics*. 45(1), 25–36.

Kukkonen, J., & 10 co-authors. (2003). Extensive evaluation of neural network models for the prediction of NO2 and PM 10 concentrations, compared with a deterministic modelling system and measurements in central Helsinki. *Atmos. Environ.*, 37, 4539–4550.

Mammarella, M. C., Grandoni, G., Fedele, P., Sanarico, M., & Di Marco, R. A. (2006). Neural networks for predicting and monitoring urban air pollution: the ATMOSFERA intelligent automatic system. *Accademia dei Lincei* (preprint) [Online] Available: http://www.atmosfera.enea.it/en/evoluzione.htm (June 14, 2010).

Manjunatha, R., Badari Narayana, P., & Hema Chandra Reddy, K. (2010). Application of Artificial Neural Networks for Emission Modelling of Biodiesels for a C.I Engine under Varying Operating Conditions, *Modern Applied Science*, 4, 77–89.

Morabito, F. C., & Versaci, M. (2002). Wavelet neural network processing of urban air pollution. *Proceedings of Int. Joint Conference on Neural Networks*, 1, 432 – 437.

Moreau, A., Teytaud, O., & Bertoglio, J. P. (2006). Optimal estimation for large-eddy simulation of turbulence and application to the analysis of sub-grid models. *Phys. Fluids*, 18(10), 10510-10511.

Milano, M., & Koumoutsakos, P. (2002). Neural network modeling for near wall turbulent flow. *Journal Comput. Phys.*, 182, 1-26.

Lee, C., Kim, J., Babcock, D., & Goodman, R. (1997). Application of neural networks to turbulence control for drag reduction. *Phys. Fluids,* 9(6), 1740 -1751.

Losch, M., & Wunsch, C. (2003). Bottom Topography as a Control Variable in an Ocean Model, *Journal Atmos. Oceanic Technol.*, 20, 1685–1696.

Panigrahi, P. K., Dwivedi, M., Khandelwal, V., & Sen, M. (2003). Prediction of Turbulence Statistics Behind a

Square Cylinder Using Neural Networks and Fuzzy Logic. *Journal Fluids Engineering.* 125(2), 385-387.

Pope, S. (2004). Ten questions concerning the large-eddy simulation of turbulent flows. *New Journal Phys.*, 6, 35

Sarghini, F., de Felice, G., & Santini, S. (2003). Neural networks based sub-grid scale modeling in large eddy simulations. *Computers & Fluids*, 32, 97–108.

Sievinen P., Praks J., Hallikainen M., Koskinen J., Hellsten, A., & Kukkonen, J. (2009). Urban morphology retrieval by means of remote sensing for the modelling of atmospheric dispersion and micro-meteorology. *Digest IEEE International Symposium on Geoscience and Remote Sensing (IGARSS-09)*, Cape Town, South Africa, 4 pp.

Table 1. Domain-averaged correlation between actual, i.e. simulated in dynamic LES, sub-grid turbulent stress in the high resolution run and the stress in the control run with the ANN correction using the input vector X over post training period of 25 min.

Experiment	Correlation	Comments
Homogeneous surface	0.585	Low correlation due to filtering of too low C_s values in ANN; fraction of $C_s < 0.15$ is 48%.
Heterogeneous surface roughness, LES-H1	0.595	Domain-averaged correlation improves as there is generally higher C_s . Fraction of $C_s < 0.15$ is 40%.
Heterogeneous surface with an obstacle, LES-H2	0.581	Domain-averaged correlation improves as there is generally higher C_s . Fraction of $C_s < 0.15$ is 39%.

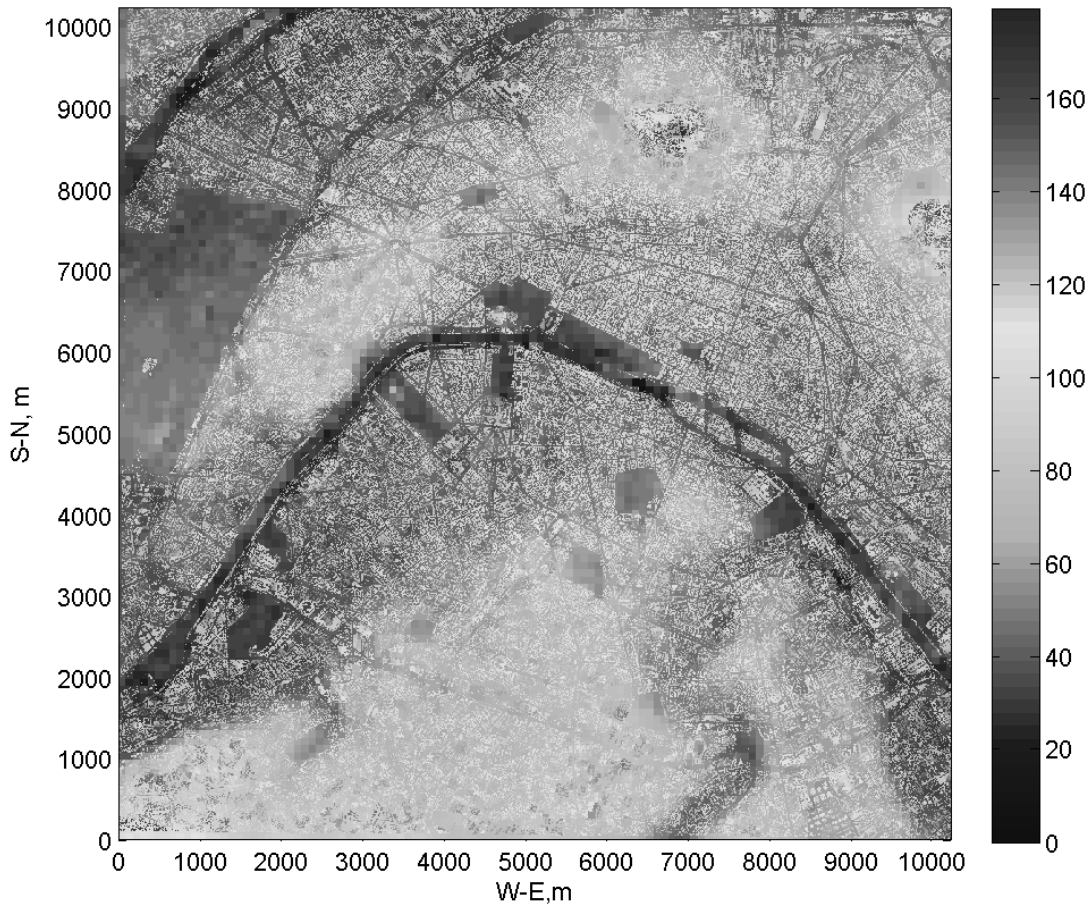


Figure 1. The surface elevation obtained for Paris area with the resolution of 10 m (1 m around Place Concordia). The gray scale indicates the elevations in meters. Courtesy MEGAPOLI team.

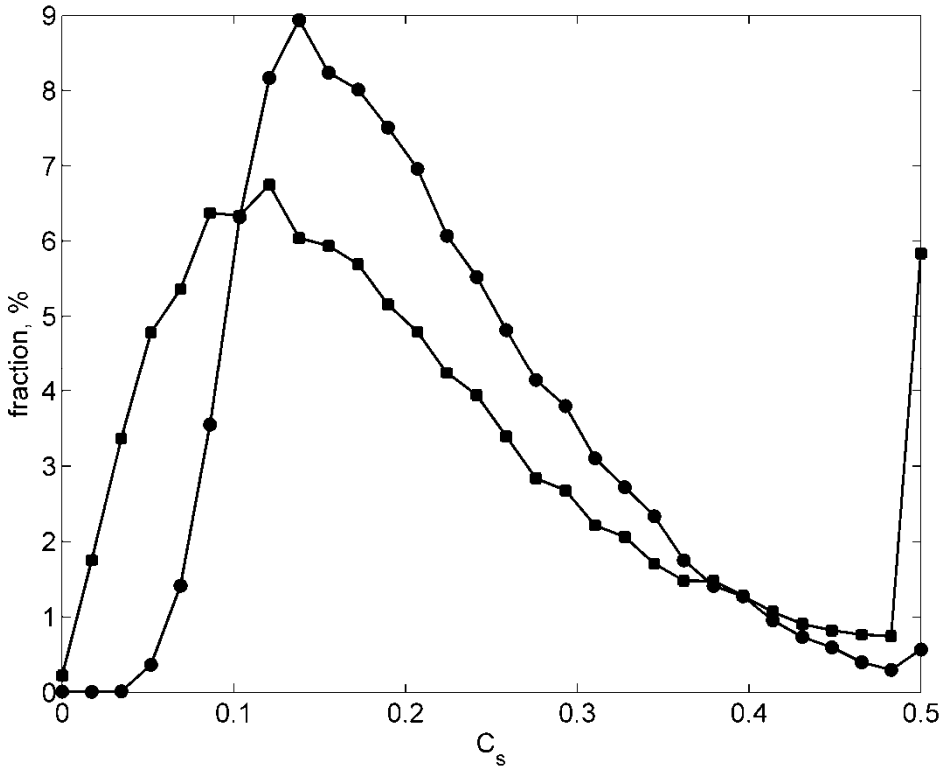
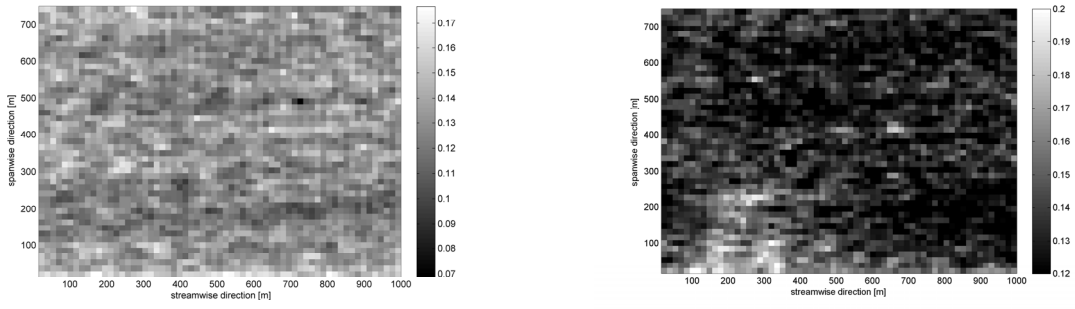
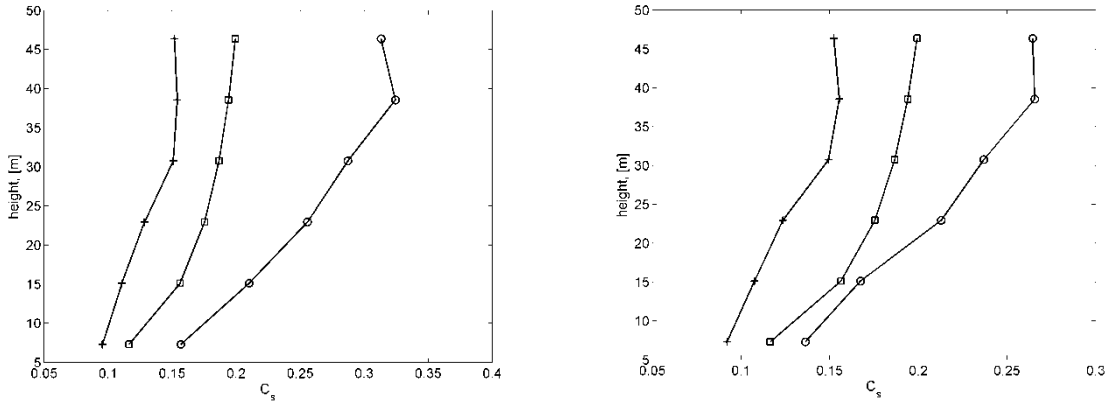


Figure 2. Distribution of C_s at the levels 1 through 6 in the LES-H1 without (squares) and with ANN (dots) correction. In ideal case, the distributions should be similar; the additional information resulting in the shown modification appears from the scale-similarity assumption, i.e. the averaged behavior of the fine-resolution run used for training remains valid for the coarse-resolution run.



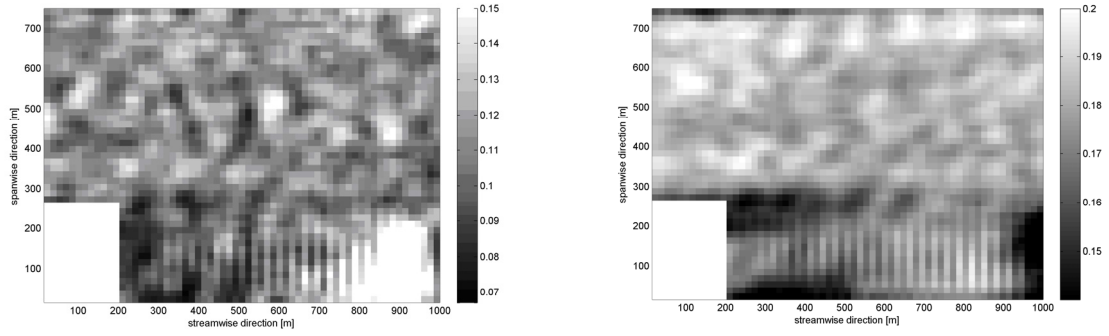
(a) (b)

Figure 3. Horizontal map of C_s at the 1st level in the coarse resolution LES-H1 without (a) and with ANN (b) correction. Larger C_s (light areas) corresponds to larger sub-grid fluxes in response to the resolved velocity shear.



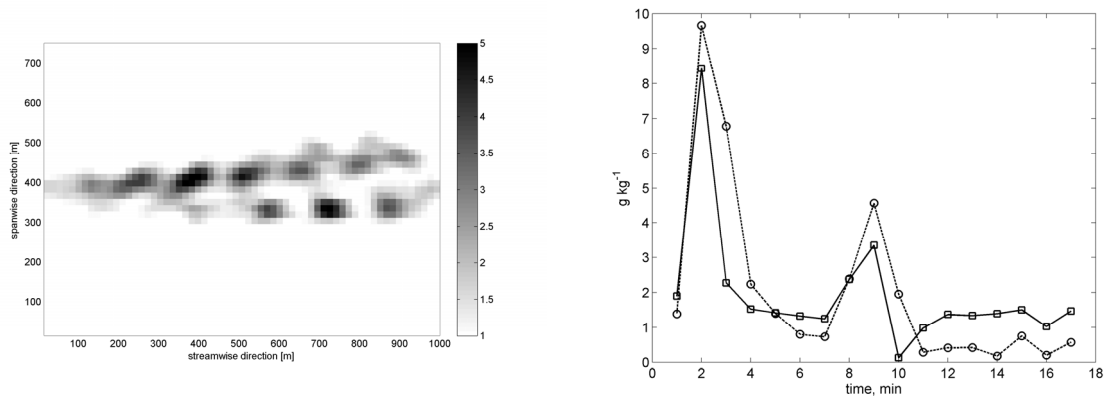
(a) (b)

Figure 4. Averaged vertical profiles of Smagorinsky coefficient, C_s , in the downstream (a) and aside (b) sub-domains in LES-H1 experiment: crosses – dynamically calculated coefficient; circles – ANN modified coefficient; squares – ad hoc fitting from Balaras et al. (1995).



(a) (b)

Figure 5. The same as in Figure 3 but for LES-H2 experiment. Pay attention that in Figure 3 the turbulent fluxes are computed over the roughness spot itself, whereas in this experiment the flow does not penetrate into the obstacle.



(a) (b)

Figure 6. (a) The averaged over 30 min field of concentration (g kg^{-1}) of a passive scalar released from unresolved source at 18 m above the surface. (b) The sequence of the 1 min concentration forecasts with ANN (squares) using pollution data from homogeneous LES experiment (circles).