Empirical orthogonal function analysis of coherent structures in a neutral atmospheric surface layer

T. Dubos, P. Drobinski and P. Carlotti

Institut Pierre Simon Laplace / Laboratoire. de Météorologie Dynamique, Palaiseau, France
Institut Pierre Simon Laplace / Service d'Aéronomie, Paris, France
Centre d'étude des tunnels, Bron, France

Abstract

Three-dimensional empirical orthogonal functions (EOFs) representing the statistically most energetic structures are extracted from a high resolution large-eddy simulation of a neutral atmospheric boundary layer. Simulated flow patterns near the ground show so-called streaks, regions near the surface of alternating high and low speed fluid organized into nearly linear bands, with horizontal spacing of several hundred meters, oriented up to $30°$ relative to the geostrophic wind, that evolve through a continuous cycle of generation, growth, decay and re-formation.

The leading EOF at a given wavenumber is found to explain a significant amount of energy for wave-vectors approximately at right angle with the streaks. The structure of such EOFs is interpreted qualitatively and quantitatively in terms of vertical profiles and momentum flux. The interaction between the mean flow and the EOFs is analyzed from an energy budget.

Corresponding author :
dubos@lmd.polytechnique.fr

1 Introduction

Coherent structures in the form of near-surface streaks are ubiquitous features of large-eddy simulations (LES) of the planetary boundary layer (PBL) in which shear plays an important role in the dynamics (Deardorff (1972); Moeng and Sullivan (1994); Drobinski and Foster (2002); Carlotti (2002)). Their existence also received recent support from field observations (Drobinski et al. (2003)). Beyond this qualitative observations, there is a need for a quantitative analysis of their individual structure as well as their role in the dynamics and energetics of the flow. There have been theoretical attempts to interpret streaks in terms of optimal linear perturbations of an Ekman layer (Foster (1997); Drobinski and Foster (2002)). Another way is to extract from a numerical simulation the most recurrent structures and to analyze their contribution to the simulated dynamics.

One extraction method is Empirical orthogonal function (EOF) analysis, also known as proper orthogonal decomposition (POD). Such a correlation technique was initially applied to turbulence by Lumley and others in classical flows (pipe, channel, boundary layer) (Holmes et al. (1996, 1997)). Such techniques have long been used in climatic studies to detect large-scale correlation patterns and derive low-dimensional models. However the lack of datasets having sufficient coverage and resolution prevented until recently this type of analysis in the planetary boundary layer. Wilson and Wyngaard performed an EOF analysis of a weakly convective atmospheric boundary layer, extracting EOFs some of which could be identified as gravity waves or boundary-layer rolls (Wilson and Wyngaard (1995)). Nevertheless a reliable representation of near-ground structures such as streaks requires a higher resolution which could be achieved only very recently (Carlotti (2002)).

In this work, we extract the EOFs of the flow simulated in (Carlotti (2002)), then analyze their contribution to the transport of momentum and the energy balance in this flow.
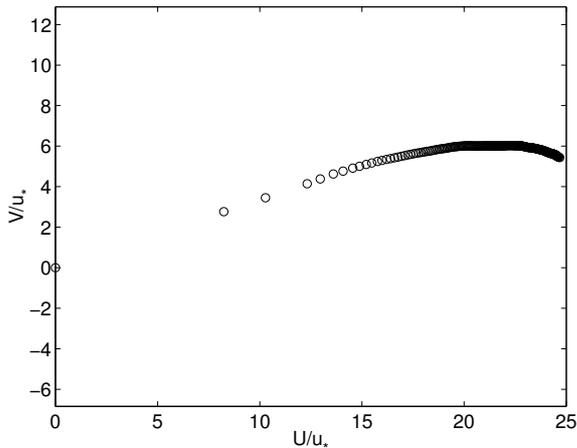
Figure 1: Mean wind hodograph normalized by the friction velocity $u_* = 0.42 m \cdot s^{-1}$.

Figure 2: Snapshot of the three velocity components $U + u$, $V + v$, w in $m \cdot s^{-1}$ at altitude $z = 60m$.

2 Mean flow properties

2.1 LES simulation

The simulation runs the non-hydrostatic LES model Méso-NH in a box with size $(L, l, H) = 3km \times 1km \times 750m$ along the x, y, z axes respectively and periodic conditions on the vertical boundaries. The mesh cell is a cube of side $6.25m$, corresponding to a resolution $N_x \times N_y \times N_z = 480 \times 160 \times 120$. The Navier-Stokes equations include a turbulent viscosity depending on the local subgrid kinetic energy which obeys a complementary prognostic equation. Boundary conditions are rigid-lid and rough ground. The velocity field is decomposed into a mean flow $(U(z), V(z), 0)$ forced by large-scale pressure gradient and turbulent fluctuations $(u(x, y, z, t), v(x, y, z, t), w(x, y, z, t))$. The time-averaged wind hodograph is presented in figure 1. The velocity is rescaled by the friction velocity $u_* = 0.42 m \cdot s^{-1}$. The wind is oriented about $12°$ left to the x direction in high layers and about $20°$ left to the long direction close to the ground, following an Ekman spiral continued by a log-layer near the surface. In the first 100m, the wind fluctuations form streaky structures roughly aligned with the ground wind (Carlotti (2002)). These structures appear clearly on horizontal cuts of the velocity field (figure 2).

2.2 Reynolds' stresses

We consider the averaged energy balance for the resolved turbulent kinetic energy $e(x, y, z, t) = (u^2 + v^2 + w^2)/2$:

$$\left\langle \frac{\partial e}{\partial t} \right\rangle + \langle uw \rangle \frac{\mathrm{d}U}{\mathrm{d}z} + \langle vw \rangle \frac{\mathrm{d}V}{\mathrm{d}z} = -\varepsilon(z) \quad (1)$$
$$\frac{\partial}{\partial z} \langle pw \rangle + \frac{\partial}{\partial z} \langle we \rangle + \frac{\partial}{\partial z} \left(\nu_t \frac{\partial e}{\partial z} \right)$$

Assuming statistical stationarity, the evolution term $\langle \partial e / \partial t \rangle$ vanishes. We present in figure 3 the shear production $\langle uw \rangle \frac{\mathrm{d}U}{\mathrm{d}z} + \langle vw \rangle \frac{\mathrm{d}V}{\mathrm{d}z}$ and dissipation $\varepsilon(z)$, which are the only terms contributing to the vertically-averaged budget.

3 EOF / POD analysis

3.1 Principle

Consider a random signal represented as an $N-$dimensional vector ϕ. We wish to find an orthonormal set of basic vectors ϕ_i such that the projection coefficients a_i are *mutually independent* random variables :

$$\phi = \sum_i a_i \phi_i \quad (2)$$

2
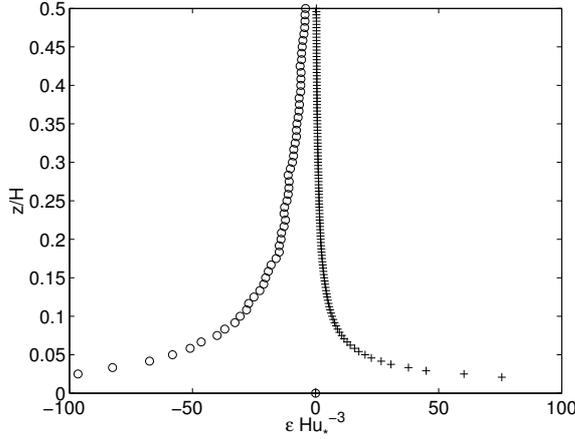
Figure 3: Resolved TKE budget : shear production (circles) and viscous dissipation (crosses) normalized by u_*^3/H.

A necessary condition for this is obtained by considering the self-correlation matrix, a Hermitian matrix :

$$\langle \phi\phi^* \rangle = \sum_{i,j} \left\langle a_i a_j^* \right\rangle \phi_i \phi_j^* = \sum_i \left\langle |a_i|^2 \right\rangle \phi_i \phi_i^*$$

where ϕ^* stands for the conjugate-transpose of ϕ and $\langle \rangle$ stands for the statistical (ensemble) average. It appears from this expression that the ϕ_i are the eigenvectors of the self-correlation correlation matrix with eigenvalues $\lambda_i = \left\langle |a_i|^2 \right\rangle$. One can not in general guarantee that the the coefficients $a_i = \phi_i^*\phi$ will be *mutually independent*. However they satisfy the weaker orthogonality property that

$$Re \left\langle a_i^* a_j \right\rangle = 0 \text{ when } i \neq j.$$

In practice the statistical average $\langle \cdot \rangle$ is computed on a set of realizations of the random signal ϕ. Our set of realizations consists of 14 snapshots of the whole velocity field taken at different instants. Thus we extract the recurrent spatial flow patterns and drop any temporal or dynamical information from the signal.

Due to the statistical orthogonality of the projection coefficients a_i any quantity that has a quadratic dependence of the signal can be on average be split into individual contributions from the EOFs. Indeed, such a quantity can be written $q(\phi) = \phi^* A \phi$ with A some $N \times N$ Hermitian matrix. Upon averaging one gets :

$$\langle q \rangle = A : \langle \phi\phi^* \rangle = \sum_i \lambda_i q(\phi_i) \qquad (3)$$

Quadratic quantities of particular interest here are the energy and the Reynolds' stresses.

3.2 Fourier and EOF analysis of the velocity fluctuations

Due to periodic boundary conditions in the x and y directions, it is possible to write the velocity field in Fourier representation :

$$u(x,y,z,t) = Re \sum_{m,n} \hat{u}_{mn}(z,t) \exp 2i\pi \left(\frac{mx}{L} + \frac{ny}{L} \right)$$

where the horizontal wave-vector $\mathbf{k} = (k_x, k_y) = 2\pi(m/L, n/l)$ is quantized according to the box dimensions L and l. The advantage is that, assuming statistical invariance on horizontal translations, the cross-correlation between Fourier coefficients $\hat{u}_{mn}(z,t)$ with different wave vectors is zero. This allows to solve the eigenvalue problem separately for each Fourier mode.

As the scalar product underlying the EOF analysis, we choose the box-averaged kinetic energy $\int (\hat{u}\hat{u}^* + \hat{v}\hat{v}^* + \hat{w}\hat{w}^*) \, \mathrm{d}z/H$. Hence for each m, n we compute the self-correlation matrix σ^{mn} made of the $N_z \times N_z$ blocks $\sigma_{zz'}^{mn}$ of size 3×3 defined as :

$$\left\langle \begin{matrix} \hat{u}_{mn}^*(z)\hat{u}_{mn}(z') & \hat{v}_{mn}^*(z)\hat{u}_{mn}(z') & \hat{w}_{mn}^*(z)\hat{w}_{mn}(z') \\ \hat{u}_{mn}^*(z)\hat{v}_{mn}(z') & \hat{v}_{mn}^*(z)\hat{v}_{mn}(z') & \hat{w}_{mn}^*(z)\hat{w}_{mn}(z') \\ \hat{u}_{mn}^*(z)\hat{w}_{mn}(z') & \hat{u}_{mn}^*(z)\hat{w}_{mn}(z') & \hat{w}_{mn}^*(z)\hat{w}_{mn}(z') \end{matrix} \right\rangle .$$

The values of z and z' are the $N_z = 120$ altitudes resolved by the model. The eigenvalues E_i^{mn} and eigenvectors $(\hat{u}_i^{mn}(z), \hat{v}_i^{mn}(z), \hat{w}_i^{mn}(z))$ of the $3N_z \times 3N_z$ Hermitian matrix σ^{mn} finally provide the desired EOFs and their energetic weight. The corresponding flow patterns have a vertical structure described by $(\hat{u}_i^{mn}(z), \hat{v}_i^{mn}(z), \hat{w}_i^{mn}(z))$ and a sinusoidal horizontal dependence. We sort the energies E_i^{mn} in descending order $E_1^{mn} > E_2^{mn} > \ldots$ Notice that by construction, $E^{mn} = \sum_i E_i^{mn}$ is exactly the spectral density of kinetic energy at wave vector $\mathbf{k}$.

3
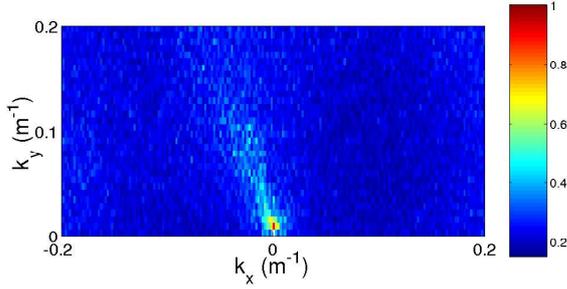
Figure 4: Energy fraction E_i^{mn}/E^{mn} explained by the first EOF of each Fourier mode as a function of the wave vector $\mathbf{k} = 2\pi\,(m/L, n/l)$.

4 Individual EOFs

4.1 Significant EOFs

We display in figure 4 the energy fraction E_1^{mn}/E^{mn} explained by the first ($i = 1$) EOF at each wave vector. This quantity is an indicator of the statistical significance of the first EOF. Suppose for instance that for a given wave vector $E_1^{mn}/E^{mn} = 1$. Then all projection coefficients a_i with $i > 1$ must be zero in decomposition (2). Thus the flow (within this Fourier mode) has the same vertical structure at all instants, but appears with a random amplitude given by $|a_1|$ and at a random position given by the phase of a_1. Conversely, if the energy is equipartitioned among all EOFs, the signal is white-in-space noise. So wave vectors with a significant ($\geq 50\%$) explained energy fraction can be said to be "strongly structured".

In wave space, such wave vectors are found to lie close to the line $m + n = 0$, corresponding to phase lines nearly parallel to the surface wind and to the streaks observed qualitatively. Thus the EOF analysis gives an objective, quantitative basis to this qualitative observation, with well-defined statistical properties.

4.2 Flow structure

Due to incompressibility, the flow corresponding to a single Fourier mode or to several Fourier modes with parallel wave vectors can be conveniently described in terms of an across-wave-vector horizontal velocity u' and an along-wave-vector stream function, from which the vertical velocity w and the along-wave-vector horizontal velocity v' are derived. This description is equivalent to Squire's transformation in the context of normal-mode stability analysis (Foster (1997)).

The first EOFs are found to be concentrated close to the ground with a vertical extension comparable to their horizontal wavelength. Thus low-wavenumber EOFs ($-m = n = 1, 2$) have a vertical scale of about $200 \sim 400$ m and are rather reminiscent of boundary-layer rolls (figure 5). Streaks are more likely to be represented by higher-wavenumber EOFs ($-m = n = 4, 5$) (figure 6).

4.3 Turbulent fluxes

From Parseval's theorem, the horizontally-averaged turbulent kinetic energy $e(z,t)$ and shear production $S(z,t)$ are the sum of contributions from each Fourier mode :

$$
\begin{aligned}
e(z,t) &= \iint \left(u^2 + v^2 + w^2\right) \frac{\mathrm{d}x\mathrm{d}y}{2Ll} \\
&= \sum_{mn} e^{mn}(z,t) \\
e^{mn}(z,t) &= \left(\hat{u}\hat{u}^* + \hat{v}\hat{v}^* + \hat{w}\hat{w}^*\right)/2
\end{aligned}
$$

$$
\begin{aligned}
S(z,t) &= \iint w \left(u\frac{\mathrm{d}U}{\mathrm{d}z} + v\frac{\mathrm{d}V}{\mathrm{d}z}\right) \frac{\mathrm{d}x\mathrm{d}y}{Ll} \\
&= \sum_{mn} S^{mn}(z,t) \\
S^{mn}(z,t) &= Re\left(\left(\hat{u}\frac{\mathrm{d}U}{\mathrm{d}z} + \hat{v}\frac{\mathrm{d}V}{\mathrm{d}z}\right)\hat{w}^*\right)
\end{aligned}
$$

Then the quadratic quantities e^{mn} and S^{mn} may be decomposed on average into contributions from the individual EOFs:

$$
\begin{aligned}
\langle e^{mn}(z,t)\rangle &= \sum_i E_i^{mn} e_i^{mn}(z) \\
e_i^{mn}(z) &= \left(|\hat{u}_i^{mn}(z)|^2 + |\hat{v}_i^{mn}(z)|^2 + |\hat{w}_i^{mn}(z)|^2\right)/2
\end{aligned}
$$

$$
\begin{aligned}
\langle S^{mn}(z,t)\rangle &= \sum_i E_i^{mn} S_i^{mn}(z) \\
S_i^{mn}(z) &= Re\left(\left(\hat{u}_i^{mn}(z)\frac{\mathrm{d}U}{\mathrm{d}z} + \hat{v}_i^{mn}(z)\frac{\mathrm{d}V}{\mathrm{d}z}\right)\hat{w}_i^{mn*}(z)\right)
\end{aligned}
$$

4
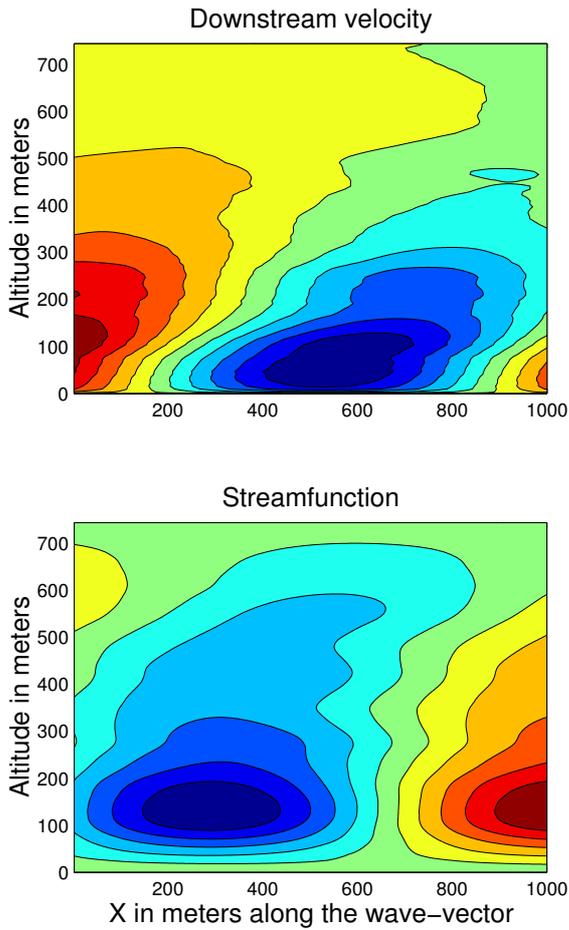
Figure 5: Flow structure of the first EOF at wave-vector $\mathbf{k} = 2\pi(-2/L,\, 2/l)$. The horizontal axis is parallel to the wave-vector. Top : contours of cross-wave-vector (along-roll) velocity. Bottom : contours of along-wave-vector stream function (regularly spaced).
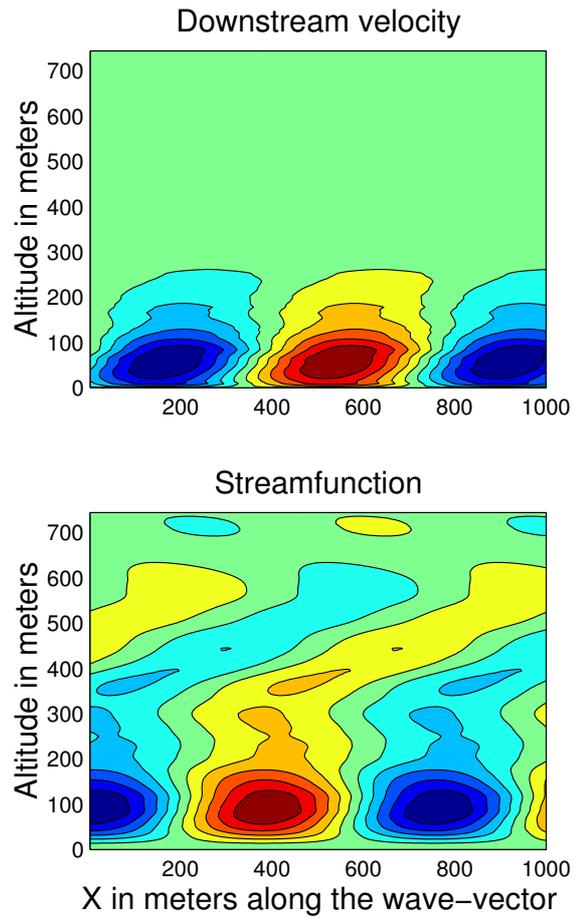
Figure 6: Flow structure of the first EOF at wave-vector $\mathbf{k} = 2\pi(-4/L,\, 4/l)$. The horizontal axis is parallel to the wave-vector. Top : contours of cross-wave-vector (along-roll) velocity. Bottom : contours of along-wave-vector stream function (regularly spaced).
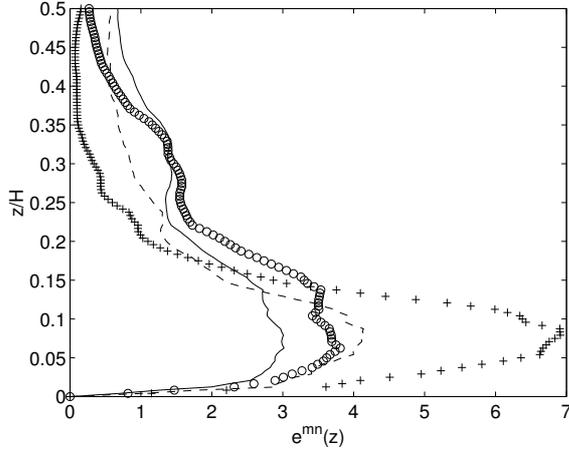
Figure 7: Vertical repartition e_1^{mn} of the kinetic energy of the first EOFs for Fourier modes $-m = n = 2$ (circles) and $-m = n = 4$ (crosses). Vertical repartition e^{mn} of the total kinetic enery contained in the Fourier modes $-m = n = 2$ (solid) and $-m = n = 4$ (dashed)

Figure 8: Vertical profile of $-S_1^{mn}$ (contribution of the first EOFs to the shear production) for Fourier modes $-m = n = 2$ (circles) and $m = -n = 4$ (crosses). Vertical profile of the total shear production S^{mn} for the Fourier modes $-m = n = 2$ (solid) and $-m = n = 4$ (dashed)

We display in figure (7) the vertical profiles of $e_1^{mn}(z)$ for $-m = n = 2$ and $-m = n = 4$. Since EOFs have by construction unit norm, $e_1^{mn}(z)$ is an adimensional quantity whose $z-$average is 1. In both cases, it can be seen that the energy of the first EOF is concentrated closer to the ground than the total energy of the corresponding Fourier mode.

We displaiy in figure 8 the shear production of energy. Most of it is solely due to the first EOF, except very close to the ground where the flux intensity contributed by the first EOF decreases much faster than the total contribution by the corresponding Fourier mode. The maximum shear production for the first EOF of Fourier mode $(-m = n = 4)$ is attained at an altitude $z \simeq 60m$ which corresponds to the typical altitude up to which streaks are observed.

5 Combinations of EOFs

By (3), second-order statistics can be expressed as the sum of independent contributions from each EOF. Indications of the relevance of the EOF basis to represent the flow are given by the con-

tribution of different sets of EOFs to dynamically important quadratic quantities such as the energy and the shear production. We shall consider and compare the contributions due to four sets of EOFs :

- the set $A = (\hat{u}_i^{mn}(z), \hat{v}_i^{mn}(z), \hat{w}_i^{mn}(z))_{imn}$ all EOFs of all Fourier modes, whose contribution is by definition the average over the full flow field

- the set $B = (\hat{u}_1^{mn}(z), \hat{v}_1^{mn}(z), \hat{w}_1^{mn}(z))_{mn}$ of the first EOF of each Fourier mode

- the set $C = (\hat{u}_i^{-m,m}(z), \hat{v}_i^{-m,m}(z), \hat{w}_i^{-m,m}(z))_{i,m}$ of all EOFs of the line of Fourier modes previously identified as "strongly structured". The contribution of this set to quadratic quantities is equal to the spectral contribution of these modes.

- the set $D = (\hat{u}_1^{-m,m}(z), \hat{v}_1^{-m,m}(z), \hat{w}_1^{-m,m}(z))_m$ of the first EOF of each "strongly structured" Fourier mode.

6
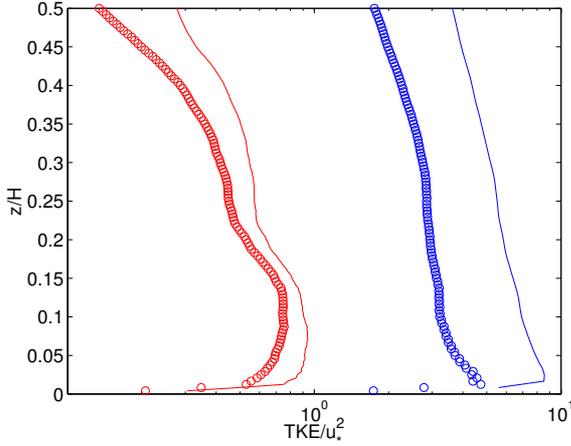
Figure 9: Profiles of turbulent kinetic energy contained in sets of EOFs A (solid blue), B (blue circles), C (solid red) and D (red circles). NB : semi-logarithmic scale.

Figure 10: Profiles of shear production of energy by sets of EOFs A (solid blue), B (blue circles), C (solid red) and D (red circles). NB : semi-logarithmic scale.

So the sets B and C are distinct subsets of A, and D is the intersection of sets B and C.

5.1 Energy profiles

The contributions to the turbulent kinetic energy $e(z)$ due to the four different sets considered are displayed in figure 9. The first EOFs of all Fourier modes capture about 25% of the total kinetic energy, at any altitude. This is not small since retaining only the first EOFs is equivalent to projecting onto a basis of $N_x \times N_y$ vectors instead of $N_x \times N_y \times N_z$.

The line $m + n = 0$ of "strongly structured" modes represents by itself about 10% of the TKE. Within these modes, the first EOFs now capture about 50% of the TKE. This was to be expected since these EOFs were identified as capturing a particularly high proportion of the TKE of their Fourier mode

5.2 Turbulent fluxes

The contributions to the shear production of energy $-S(z)$ due to the four different sets considered are displayed in figure 9. Now the first EOFs of all Fourier modes capture about 50% of the total kinetic energy. This indicates that first EOFs

are more efficient than the average at extracting energy from the mean flow.

The line $m + n = 0$ of "strongly structured" modes contributes for about 10% of $S(z)$. Within these modes, the first EOFs now represent a fraction of $S(z)$ close to 100% except very close to the ground. This fall-off can in fact also be remarked on the contribution of set B. This would indicate that spatially organized motions contribute by about 50% to the extraction of the energy from the mean flow, except for a small layer ($\sim 20m$) very close to the ground where disorganized motion is preponderant. However we might be reaching the resolution limits of the model at such altitudes of 2 or 3 grid cells above the ground.

6 Discussion

The EOF analysis aims at representing a signal as a superposition of independant, elementary patterns. In our case, the boundary conditions and the assumed statistical homogeneity imply that these flow patterns are monochromatic Fourier modes on the horizontal. The leading EOF at a given wavenumber was found to explain a significant amount of energy for wave-vectors approx-

imately at right angle with the streaks. We described the vertical structure of these EOFs. It appeared that the EOFs for wavenumbers $-m = n \geq 4$ are good candidates to represent the qualitatively observed streaks. We analyzed their interaction with the mean flow and found that they are more efficient at extracting energy from the mean flow that the average fluctuations.

The "strongly structured" EOFs are efficient at extracting energy from the mean flow and suffer little viscous dissipation due to their relatively smooth vertical structure and moderate wavenumber. Thus the dominant mechanism for their decay is the nonlinear interaction with other EOFs. The study of these interactions require the derivation of an energy budget detailed at the EOF level. Wilson and Wyngaard (1995) decompose Reynolds' energy budget into contributions from the EOFs. This budget involves analogues of the terms of shear production, pressure transport and dissipation, together with a term describing the EOF-EOF interactions. In a different way, a vertically-averaged view of the energy transfers between EOFs can be obtained by Galerkin-projecting the flow dynamics onto the orthonormal basis provided by the EOFs, resulting in a prognostic equation for the projection coefficients a_i^{mn}. The different terms contributing to $\partial(a_i^{mn*} a_i^{mn})/\partial t$ will provide the desired information.

In both cases the average value of nonlinear terms imply cubic statistics which can not be obtained from the EOFs and their energetic weight. Wilson and Wyngaard obtain the nonlinear terms as the remainder of the energy budget when all other terms have been computed. At this degree of detail, this can give a precise result only if the energy budget is derived from the discrete equations simulated by the model and not from their continuous formulation. This question is under current investigation.

References

Carlotti, P. (2002). Two-point properties in atmospheric turbulence very close to the ground : comparison of a high resolution LES with theoretical models. *Bound.-Layer Meteor.*, 104:381–410.

Deardorff, J. (1972). Numerical Investigation of Neutral and Unstable Planetary Boundary Layer. *J. Atoms. Sci.*, 29:91–115.

Drobinski, P., Carlotti, P., Newson, R., Banta, R., Foster, R., and Redelsperger, J.-L. (2003). The structure of the near-neutral atmospheric surface layer. *J. Atmos. Sci., in press*.

Drobinski, P. and Foster, R. (2002). On the origin of near-surface steaks in the neutrally-stratified planetary boundary layer. *Boundary-Layer Meteorol.*, 108:247–256.

Foster, R. (1997). Structure and energetics of optimal Ekman layer perturbations. *J. Fluid. Mech.*, 333:97–123.

Holmes, P., Lumley, J., and Berkooz, G. (1996). *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press.

Holmes, P., Lumley, J., Berkooz, G., Mattingly J.C., and Wittenberg R.W. (1997). Low-dimensional models of coherent structures in turbulence. *Phys. Reports*, 287(4):338–384.

Moeng, C. and Sullivan, P. (1994). A Comparison of Shear- and Buoyancy-Driven Planetary Boundary Layer Flows. *J. Atmos. Sci.*, 51:999–1022.

Wilson, D. and Wyngaard, J. (1995). Empirical Orthogonal Function analysis of the weakly convective atmospheric boundary layer. Part II : Eddy Energetics. *J. Atmos. Sci.*, 53(6):824–841.