2D Modeling of Turbulent Flow around a Cylindrical Storage Tank by Artificial Neural Networks

Pierre Lauret, Frederic Heymes, Laurent Aprin, Anne Johannet, Pierre Slangen

2D Modeling of Turbulent Flow around a Cylindrical Storage Tank by Artificial Neural Networks

Pierre Lauret*, Frédéric Heymes, Laurent Aprin, Anne Johannet, Pierre Slangen

Laboratoire de Génie de l'Environnement Industriel (LGEI), Ecole des Mines d'Alès, Alès, France
pierre.lauret@mines-ales.fr

Loss of containment of storage tanks can involve atmospheric dispersion of toxic or flammable gases. The accident of Buncefield oil storage depot in 2005 and the Viareggio LPG explosion in 2009 tragically illustrate the potential consequences of dispersion phenomena in complex environments. Gas dispersion modeling in a complex geometry is a tricky task because of the effects of obstacles, which involves for example a high level of turbulence. Indeed, flows around obstacles have specific behavior like boundary layers effects and recirculation zones. Different approaches exist to estimate atmospheric dispersion, depending on the modeling strategy. Gaussian models solve advection-diffusion equation (ADE) assuming several hypotheses, in particular uniform flow and turbulence, but are not very efficient in congested areas. On the other hand, Computational Fluid Dynamics (CFD) computes very likely turbulent flows and dispersion by solving non-linear Navier-Stokes equations for congested areas. Turbulence can be modeled by considering turbulent kinetic energy k and its dissipation rate ó (k-ε model standard). A full 2D CFD modeling requires solving 5 equations (continuity, momentum on x and y axes, transport of k and ó). These numerous calculations induce long computing time on large areas with fine meshing. Once the wind field is calculated, dispersion of a pollutant can be rapidly computed by ADE.As the major part of computing time of CFD models is dedicated to determine the turbulent flow, Artificial Neural Networks (ANN) were thus investigated to calculate x and y velocities and turbulent diffusion coefficient D_t. This machine learning method is a powerful statistical tool as it is able to reproduce accurately any nonlinear and dynamic behavior from a database without any physical assumption. This study focused on turbulent flows around cylindrical storage tanks, with a diameter in the range [10 m – 52 m]. Database is designed by RANS G – ó CFD model. Several neural networks solutions are proposed and their efficiency is compared and discussed in terms of quality, real-time applicability and real-life plausibility. Four criteria (coefficient of determination, factor of two, fractional bias and normalized mean squared error) are used to evaluate the model. While the accuracy is kept within satisfying criteria values, computational time is reduced by a factor of 600.

1. Introduction

Danger is inherent to industrial facilities. Indeed, many hazardous materials are usually stored in numerous tanks connected together by combination of pipes, pumps and others process equipment. In the next section, analysis of previous accidents is focused on the near field of the leakage.

1.1 Recent accidents analysis

Several atmospheric dispersion accidents occurred in the past decade, involving leakage from storage tank. These usually lead to explosion and/or fire. In this context, unleaded petrol evaporation, dispersion and explosion that occurred at Buncefield Terminal, United Kingdom in December 2005 is a case of study. The overfilling of one of the larger tank of Buncefield Oil Storage Depot releases about 260 m^3 of petrol from the roof. The liquid fell and had been fragmented into spray of droplets, implying evaporation which produced dense cloud of flammable vapors. Evolution of such a cloud is induced by weather conditions, stable and humid with low wind velocities (< 3m.s^{-1}) in this case. Water vapor condensed in the cloud and produced a

visible mist that has been captured by security cameras. Evolution of the cloud is well known and highly influenced by the presence of obstacles: bund, tree lane, tanks (Gant and Atkinson, 2011). Finally, ignition of the cloud led to a massive vapor cloud explosion, and, by domino effect (Heymes et al., 2014), to oil fire that last over 2 days. This accident reveals the importance of the perturbation of the flow by storages with different shapes. This work is focused on forecasting flows around cylinder.

1.2 Flow around cylinder

Flows around common shapes have been evaluated through field and small scale experiments. Taneda (1977) compared several laminar and turbulent regimes around sphere and vertical cylinder to identify different behaviors including a recirculation zone in the wake of the obstacle for sufficient Reynolds. For higher Reynolds number, Hosker Jr (1985) identifies several zones for uniform inlet flow around cylinder (Figure 1). A displacement zone is located upwind of the obstacle (a), where a high pressure is applied on the cylinder. Boundary layers are created on the edges of the obstacle (b). Separating point location depends on the flow turbulence. The turbulent recirculation zone or cavity is observed directly behind the obstacle (c). The minimum pressure is observed on body at wake centerline. Due to the increase of turbulence and velocity orientation, presence of dispersed gas last. At the cavity closure point begin a far field zone characterized by vortex generation and turbulent wake (d).

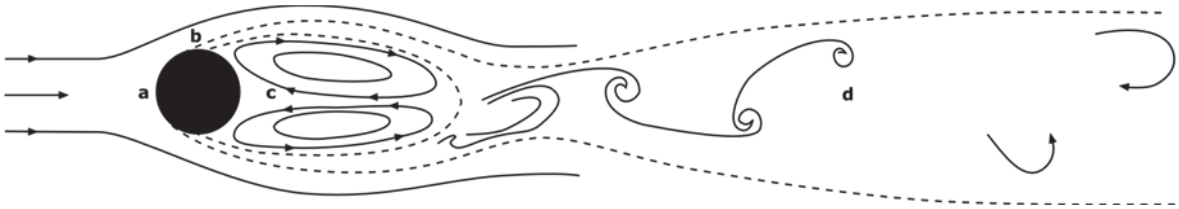

Figure 1: Typical 2D mean flow field around a cylinder, based on Hosker (2005) with - a: flow separation – b: boundary layers creation – c: recirculation zone – d: turbulent wake

Unsteady aspect of the turbulence is noticed, especially in the far field of the obstacle. Flow in the wake is perturbed up to 50 to 100 body diameters. Moreover, the flow structures around a cylinder are considered 2-D (Ozgoren et al., 2011). In the following study, mean velocity field are considered at 2 meters of altitude (corresponding to values of Reynolds number superior to 10^6). A fully turbulent separation for a flow around two-dimensional cylinder is achieved for Reynolds number around 3.10^5 (Schlichting, 1979).
Several experimental investigations of atmospheric dispersion around obstacles have been made in full-scale. Those are necessary to feed models (both reduced scale and numerical) and evaluate them. In this goal, Mavroidis et al. (2003) investigated the behavior of plumes originating from sources locations differing from the axis between center of obstacle and mean wind direction. Recirculation zone is of primary importance, as the measured concentrations increase when the source is displaced slightly from the axis. Moreover, comparisons between full-scale and wind tunnel experiments show the same trend: highest mean concentrations were overestimated by the wind tunnel, credited to the more complex scales of turbulence observed in the field.

1.3 Flow modeling

Computational Fluid Dynamics (CFD) models are eulerian models solving continuity and momentum equations on a mesh to obtain wind fields. When dealing with atmospheric flows, it is necessary to take turbulence into account. Several models can be detailed. In the scope of this study, RANS (Reynolds Averaged Navier Stokes equations) standard $k-\epsilon$ model is considered. k represents the turbulent kinetic energy and ϵ the linked dissipation rate. Transport equations of these variables are used to close the system of equations through the modeling of turbulent viscosity μ_t :

$$\mu_t = \rho C_\mu \frac{k^2}{\epsilon} \quad (1)$$

With ρ the density of the fluid and C_μ, a model constant equal to 0.09.
Air compressibility is generally neglected when Mach number is less than 0.3. In case of atmospheric flow at ground level, Mach number is less than 0.06 for a maximum wind speed of 20 $m.s^{-1}$ so that density is considered as constant. These CFD models give accurate results for turbulence. Accuracy is usually better as

the mesh is finer, but implies an increase of computation time. In this study, a database was generated using a classical 2D RANS standard $k-\epsilon$ model with turbulence set chosen as recommended by Richards and Hoxey, (1993). Some limitations exist when dealing with wind engineering simulations using these models, especially concerning the degradation of wind and turbulence profile. However, the aim of this work was to check if neural networks are able to simulate a modeling turbulent flow around an obstacle, and not to question the ability of CFD modeling to predict atmospheric flow. To investigate this question, own limitations of CFD models were not considered.

1.4 Artificial Neural Networks

Artificial Neural Networks (ANNs) are powerful non-linear fitting tools based on statistical modeling. They are generally used when the process to model is not fully known. Two properties are essential: the universal approximation (Hornik et al., 1989), and the parsimony (Barron, 1993). Thanks to these properties ANNs are able to predict efficiently future behaviors on never encountered situations. Information about the non-linear phenomenon to simulate or forecast must be provided using a database. As a non-linear fitting tool, ANN generally acts as a black-box: the physics cannot be extracted from the results. Nevertheless, ANN can be used to forecast physical phenomenon, presenting powerful models (Kong A Siou et al., 2011). A neuron is a nonlinear, parameterized, bounded function. Variables are assigned to the inputs of the neuron. Output of a neuron is the result of nonlinear combination of the inputs, weighted by the parameters and using an activation function. Sigmoid s-shaped functions are generally used. A neural network is the composition of several neurons. Parameters calibration is done through application of an algorithm using the training database and designed to decrease the model error, in this work the Levenberg-Marquardt method is adopted (Hagan and Menhaj, 1994). The function realized by the ANN is continuously tested on a disjoined set of examples, namely the stop set. This last set is employed to avoid overtraining using early stopping (Sjöberg et al., 1995). Lastly, performances of the model must be measured on another set, never used during training or stopping: test or validation set.

2. 2D horizontal Flow around cylinder forecasting using Artificial Neural Networks

In the present work, the aim was to check relevance of ANN to predict main characteristics of a turbulent flow around a cylinder. It noticed that, this study follows a previous work which consisted to compare machine learning tools to model the atmospheric dispersion (Lauret et al., 2014). This model was designed especially for emergency management or anticipating situation. It has to be effective and time computation efficient.

2.1 Database creation

In case of a flow around a cylinder and so as to apply atmospheric dispersion equations, the required parameters are velocities in the direction of the incident flow, on orthogonal direction, and information of turbulence. These data are given by the turbulent viscosity, which is not constant on the entire plane. It is directly linked to the turbulent diffusion coefficient by the Schmidt number S_{ct}:

$$D_t = \rho \frac{\mu_t}{S_{ct}} \tag{2}$$

This number is taken equal to 0.7 in the following. Thus, 72 different CFD simulations of 2D flow around cylinder in neutral stability conditions at altitude $z = 2\,m$ are generated combining different inlet velocities and diameter. Values used are respectively included in the interval [2 ; 10] m.s^{-1} with a step of 1 m.s^{-1} and in the interval [10 ; 52] m with a step of 6 m. This database represents more than 8 million values of velocities and turbulent diffusion coefficient. In order to build a database for neural networks, it is necessary to sample it. Stratified random sampling is proposed in this work using concentrations as class parameter. The database is thus created by selecting E_C examples in I_c classes in order to limit the total number of examples at 30 000. This database is divided in three sets: training set, validation set (or stop set) and test set. Validation set is used to avoid overfitting and test set is used to evaluate the performance of the model (Kong A Siou et al., 2012).

2.2 Variables selection and neural network architecture

Inputs of the neural networks have to be representative of the flow and easily determinable in real life (Lauret et al., 2013). Three types of input can be determined: the inlet velocity, cylinder diameter and coordinates where output is evaluated. These inputs are reported in the following table:

Table 1: Neural networks inputs

Inputs	Targets		
	U_x/U_{ini}	U_y/U_{ini}	D_t
Angle formed by the axis ray and cylinder center to location ray (rad)	x	x	x
Distance from center of the cylinder (m)	x	x	x
Inlet Velocity (m.s-1)	x	x	
Obstacle radius (m)	x	x	
Characteristic number of the flow ($R.U_{ini}$)			x

ANNs used in this work are two-layer perceptron with hyperbolic tangent as activation function for the first layer and linear function for output layer. Several trainings are done to ensure that the best neural network is used in simulations. Optimization is done through three different ways: initialization of neurons parameters, number of neurons in hidden layer and database sampling. Training results are discussed in these terms.

2.3 Results on training database

ANN evaluation is usually done through the coefficient of determination or Nash criterion calculation (R^2) on the test sets. This criterion evaluates the data fitting of the ANN model for a $[-\infty; 1]$ range, with 1 the perfect match and 0 the forecasting result equals to the mean value of test set. Results of best training for each flow parameter are reported in the following table:

Table 2: ANN training results for U_x/U_{ini}, U_x/U_{ini}, and D_t

Flow parameter	U_x/U_{ini}	U_y/U_{ini}	D_t
I_c best number	20	40	20
E_c best number	30	20	20
Number of neurons in hidden layer	20	20	20
R^2 value	0.985	0.997	0.999
Training duration (s)	906	3,731	2,775

Values of R^2 are below 0.98 and thus correspond to correct accuracy. It is noticed that training duration can last more than one hour on a classical workstation. Nevertheless, training duration is a long process while simulations are very fast as observed in the following paragraph.

3. Results on situations independent from training and model selection

To assess the performance of such a model, it is necessary to evaluate the neural networks against unlearned data. Nine different test cases are then evaluated containing three different inlet velocity values (2.5, 5.5 and 9.5 m.s^{-1}) and three different cylinder diameters (12, 26 and 50 m). These values are taken to be representative of low, medium and high values of each parameter. Four performance criteria are used. Coefficient of determination shows a general accuracy of the model. Factor of two (FAC2) corresponds to the fraction of the values between an half and twice the observed values. Best value is 1. Fractional bias (FB) represents the systematic error. Best value is 0. FB positive values represent underestimation and negative values represent overestimation. Normalized mean square error (NMSE) represents the global error. Best value is 0.

Results of the nine test cases are reported in the Table 3. Considering velocity on the x-direction, the coefficient of determination and the factor of two show a good agreement with CFD results with values superior to 0.94. Systematic error is at a low level with only one case overestimating CFD values. Global error is low. Considering velocity on the y-direction, coefficient of determination gives values equal to 0.98 but the factor of two is low, less than 0.48. This can be explained by the high number of values near zero that can be lightly over or under estimated and this produced high level of error when ratio is used. The same problem occurs when the fractional bias is evaluated (and consequently normalized mean square error), giving unrepresentative values. Nevertheless, estimation of absolute error is thereafter detailed.

Considering turbulent diffusion coefficient, values of coefficient of determination and factor of two are both higher than 0.9. Nevertheless, test cases with a large diameter are less well predicted. The same observation can be done on the systematic and total error with a magnitude of ten between 50 diameters test cases and remaining test cases. Moreover, fractional bias indicates overestimation of turbulent diffusion coefficient values for the major part of test cases except the ninth.

Table 3: Performance criteria for the forecasting of x and y velocities and the coefficient of determination by the neural network on nine test cases.

Test cases	Diameter (m)	Velocity (m.s-1)	U_x				U_y		D_t			
			R^2	FAC2	FB	NMSE	R^2	FAC2	R^2	FAC2	FB	NMSE
1	12	2.5	0.98	0.99	10^{-3}		0.98	0.46	0.98	0.99	-1.2×10^{-2}	1.3×10^{-3}
2	12	5.5	0.98	0.99	10^{-4}		0.98	0.47	0.98	0.99	-1.1×10^{-2}	1.2×10^{-3}
3	12	9.5	0.97	0.99	-10^{-4}		0.98	0.48	0.98	0.99	-7.5×10^{-3}	1.1×10^{-3}
4	26	2.5	0.96	0.99	10^{-3}		0.98	0.45	0.98	0.99	-2.3×10^{-3}	4.0×10^{-3}
5	26	5.5	0.96	0.99	10^{-3}	10^{-3}	0.98	0.45	0.98	0.99	-7.1×10^{-3}	3.7×10^{-3}
6	26	9.5	0.96	0.99	10^{-3}		0.98	0.46	0.98	0.99	-3.4×10^{-3}	3.4×10^{-3}
7	50	2.5	0.94	0.99	10^{-3}		0.98	0.44	0.93	0.95	-3.2×10^{-2}	3.7×10^{-2}
8	50	5.5	0.95	0.99	10^{-3}		0.99	0.45	0.92	0.94	-1.8×10^{-2}	3.8×10^{-2}
9	50	9.5	0.96	0.99	10^{-3}		0.98	0.44	0.91	0.93	2.1×10^{-2}	4.1×10^{-2}

Based on the mean values of performance criteria listed above, ANN modeling tends to correctly forecast characteristics of the flow. The following pictures represent the relative error between CFD reality and ANN modeling of U_x, U_y, D_t for test case 9, which seems to be the more difficult to forecast by the ANN.

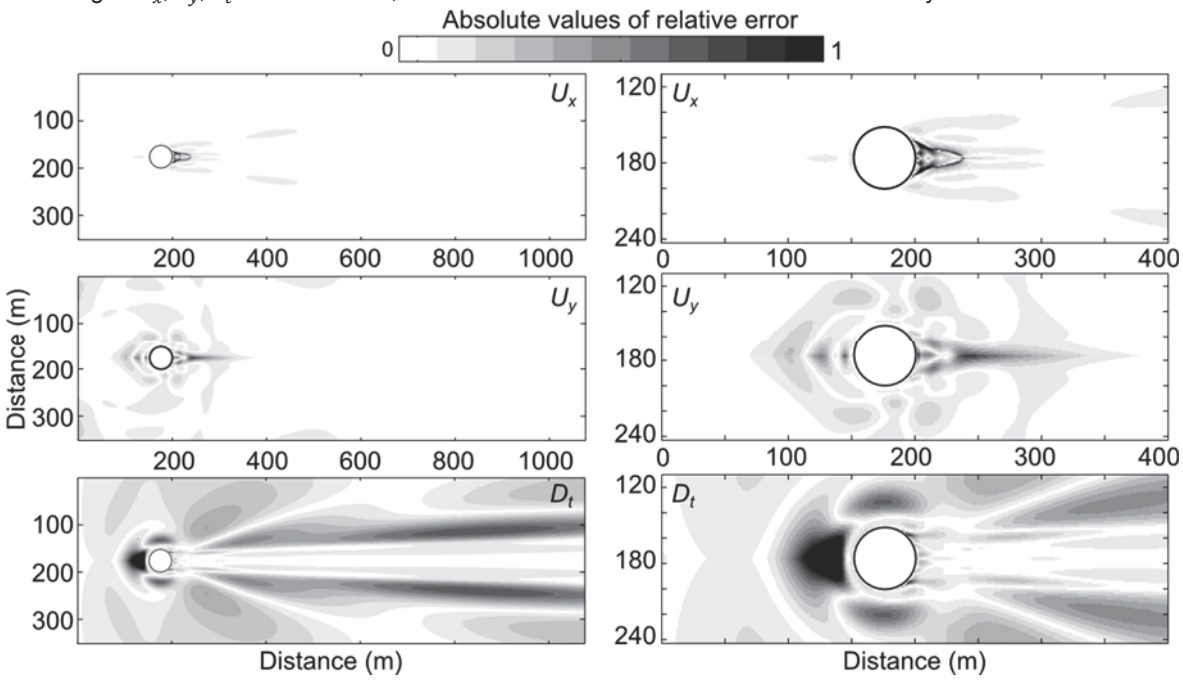

Figure 2: Absolute values of relative error between ANN and CFD of x and y velocities and turbulent diffusion coefficient for test case 9 – global view (left) and zoom near the 50 m cylinder (right)

For determination of U_x, relative error is not significant except in the near wake of the cylinder, in the recirculation zone. Errors are more important when considering forecasting of U_y. Deficient values find here are mainly due to the mean value of zero that can induce high level of relative error even with low level of absolute error. D_t is the most difficult parameter to forecast. Indeed, errors can be found through the whole domain. Nevertheless, wake of the cylinder is correctly modeled. The worst values modeled are in the front of the cylinder and on the side of the far wake. ANN underestimates these values: values obtained are less than values deduced from atmospheric conditions and thus represent non-logical values. Possible explanations are that sampling method enhanced high values compared to low ones and that the focused zone corresponds to high gradient changes. These results have to be compared in terms of computation time for each model. Duration depends on the domain to model and the resolution required. In this work, space step is 1 m representing more than 350 000 node values on a structured mesh. This is comparable to CFD case with unstructured mesh. Time computation is extremely different with about 20 minutes for RANS $k-\epsilon$ CFD model and less than two seconds for the ANN model.

4. Conclusions

The work presented here shows the forecasting feasibility of flow characteristics by artificial neural networks. Several steps are needed to design the network e.g. forming a database, sampling it and selecting the optimized ANN. To be real time effective, the ANN model has to use data directly available. Here, data used are only wind velocity at the inlet and obstacle diameter. Hence, ANN models the entire 2D velocity and turbulent diffusion coefficient field. Performance criteria show satisfying values with coefficient of determination superior to 0.9 for all training-independent test cases. Despite these results, several limited area are less well forecasted, specifically nearby the obstacle. Investigations have to be done to improve forecasting in these areas. Otherwise, using the ANN model instead of CFD model improves significantly time computation by a factor of 600. Knowledge of the wind field and the turbulent diffusion coefficient allows calculation of the dispersion of a passive pollutant around a cylinder. Several methods can thus be used, using particle tracking or solving the advection diffusion equation. Further work will be focused on the performance improvement, implementation of gas dispersion and addition of multiple obstacles with different shapes. In this perspective, using concatenation of separate cases is considered in first approach.

References

Barron, A. R., 1993, Universal approximation bounds for superpositions of a sigmoidal function, IEEE Transactions on Information Theory, 39(3), 930–945, DOI:10.1109/18.256500

Gant, S. E., & Atkinson, G. T., 2011, Dispersion of the vapour cloud in the Buncefield Incident, Process Safety and Environmental Protection, 89(6), 391–403, DOI:10.1016/j.psep.2011.06.018

Hagan, M. T., & Menhaj, M. B., 1994, Training Feedforward Networks with the Marquardt Algorithm, IEEE Transations on Neural Networks, 5(6), 989–993.

Heymes F., Aprin L., Slangen P., Lapébie E., Osmont A., Dusserre G., 2014, On the Effects of a Triple Aggression (Fragment, Blast, Fireball) on an LPG Storage, Chemical Engineering Transactions, 36, 355–360.

Hornik, K., Stinchcombe, M., & White, H., 1989, Multilayer Feedforward Networks are Universal Approximators, Neural Networks, 2, 359–366.

Hosker Jr, R. P., 1985, Flow and Diffusion Near Obstacles, D. Randerson (Ed.), Atmospheric Science and Power Production, Technical Information Centre, US Department of Energy, Washington, District of Columbia, 241–326.

Kong A Siou, L., Johannet, A., Borrell, V., & Pistre, S., 2011, Complexity selection of a neural network model for karst flood forecasting: The case of the Lez Basin (southern France), Journal of Hydrology, 403(3-4), 367–380, DOI:10.1016/j.jhydrol.2011.04.015

Kong A Siou, L., Johannet, A., Valérie, B. E., & Pistre, S., 2012, Optimization of the generalization capability for rainfall–runoff modeling by neural networks: the case of the Lez aquifer (southern France), Environmental Earth Sciences, 65(8), 2365–2375, DOI:10.1007/s12665-011-1450-9

Lauret, P., Heymes, F., Aprin, L., Johannet, A., Lapébie, E., & Osmont, A., 2014, Atmospheric Turbulent Dispersion Modeling Methods using Machine learning Tools, Chemical Engineering Transactions, 36, 517–522.

Lauret, P., Heymes, F., Aprin, L., Johannet, A., Munier, L., & Lapébie, E., 2013, Near Field Atmospheric Dispersion Modelling on an Industrial Site Using Neural Networks, Chemical Engineering Transactions, 31, 151–156, DOI:10.3303/CET1331026

Mavroidis, I., Griffiths, R. F., & Hall, D. J., 2003, Field and wind tunnel investigations of plume dispersion around single surface obstacles, Atmospheric Environment, 37(21), 2903–2918, DOI:10.1016/S1352-2310(03)00300-5

Ozgoren, M., Pinar, E., Sahin, B., & Akilli, H., 2011, Comparison of flow structures in the downstream region of a cylinder and sphere, International Journal of Heat and Fluid Flow, 32, 1138–1146, DOI:10.1016/j.ijheatfluidflow.2011.08.003

Richards, P. J., & Hoxey, R., 1993, Appropriate boundary conditions for computational wind engineering models using the k-epsilon turbulence model, Journal of Wind Engineering and Industrial Aerodynamics, 46-47, 145–153.

Schlichting, H., 1979, Boundary-layer theory, McGraw-Hill Book Company, Inc., New York,

Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Deylon, B., & Glorennec, P. Y., 1995, Nonlinear Black-Box Modeling in System Identification: a Unified Overview. Automatica, 31(12), 1691–1724.

Taneda, S, 1977, Visual study on unsteady separated flows around bodies, Progress in Aerospace Sciences, 17, 287–348.