Basic course of tropospheric composition modelling

Course outline

Auditorium: Master and PhD student specializing in atmospheric physics and meteorology. Required courses include higher mathematics (PDE, ODE) and statistics, basic chemistry and physics.

Schedule: one lecture per week, one seminar per 2 weeks, home assignments

Credits: 5 points

The aim of the course is to provide the basic knowledge about construction and application of atmospheric chemistry-transport models. The size of the course will not allow detailed consideration of specific parameterizations but practical examples will be given at the seminars, home computer practice and specific tasks.

- overview of the atmospheric composition problems, their main types, areas of research and applications
- basics of forward and adjoint atmospheric advection-diffusion equation (AD): derivation, main features, terms
- continuity equation and mass budget
- advection term of the AD, its representation in models. Types of dispersion models
- diffusion term of the AD, its representation in models. Turbulent closures and parameterizations
- chemical transformations of the atmospheric tracers, representation in the AD
- removal terms in the AD. Dry and wet deposition, degradation.
- input and output data. Links with meteorological and impact-assessment models. On-line and off-line coupling with meteorological models.
- model-measurement comparison. Representativeness of point observations. Sub-grid variability of modeled data.

Main literature:

Seinfeld, J., Pandis, S. (2006) Atmospheric chemistry and physics. From air pollution to climate change. *J.Wiley & sons, Inc.*, ISBN 978-0-471-72018-8.

Jacobson, M. (1999) Fundamentals of atmospheric modelling. *Cambridge Univ. press*. ISBN 0-521-63717-1.

Nieuwstadt, F.T.M. & van Dop, H. (1982) Atmopspheric turbulence and air pollution modelling. D.Reidel publishing company. ISBN 90-277-1365-6.

Jacob, D. (1999) Introduction to atmospheric chemistry. *Princeton Univ. press*. ISBN 0-691-00185-5.

Holton, J.R. (2004) An introduction to dynamic meteorology. *Elsevier Academic Press*, ISBN-13: 978-0-12354015-7. International Geophysics series, **88**.

Vertical structure of the atmosphere

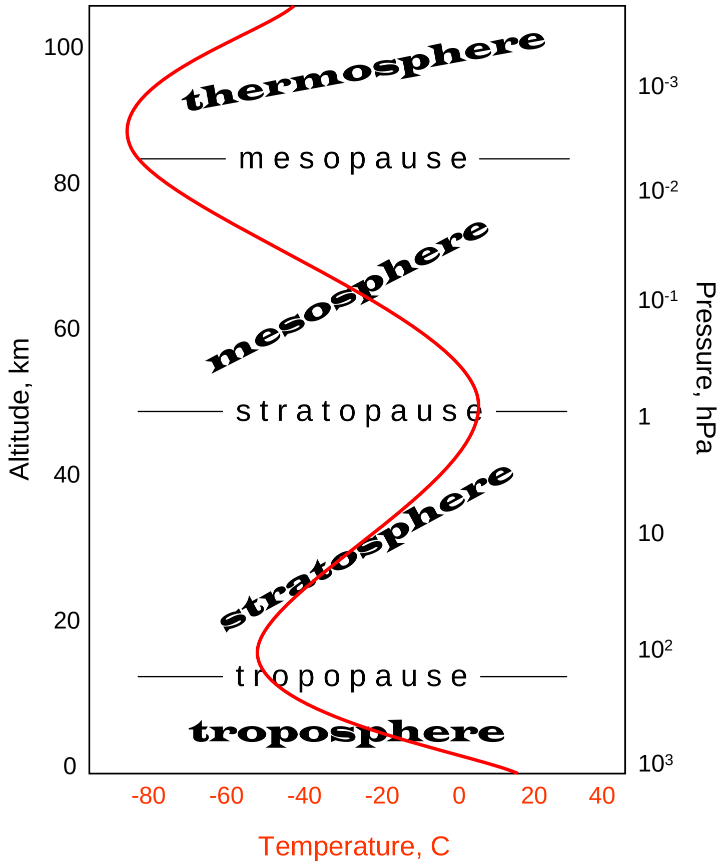


Figure 1. Vertical layers of the atmosphere

Troposphere. The lowest layer, extends up to tropopause (8-15 km, i.e. 300-100hPa), absolute temperature decreases with altitude with a mean rate of ~5^0 km^{-1} (dry-air lapse rate 9.7 K km^{-1}). Strong mixing.

Stratosphere. Extends from tropopause to stratopause at ~50km (~1 hPa). Temperature increases with altitude (absorption of solar UV radiation by ozone), vertical mixing is slow.

Mesosphere. Extends from stratopause to mesopause at ~85km (~0.01 hPa). Absolute temperature decreases with altitude, comparatively rapid vertical mixing. Mesopause is the coldest point in the atmosphere.

Thermosphere. Region above the mesopause, extends to hundreds km (pressure down to 10^{-2} Pa). Temperature grows with altitude, the hottest region due to solar radiation absorption by N_2 and O_2.

Ionosphere covers upper part of mesosphere and lower part of thermosphere. Characterised yb high concentrations of ions produced by photoionization.

Exosphere. The upper-most region of the atmosphere. High-energy molecules escape to the open space.

Troposphere. Tropos is a Greek word meaning mixing. The name was given by Sir Napier Shaw (British meteorologist). Most of processes are controlled by the distance from the earth surface, including the temperature decline, boundary layer in the lowest part of the troposphere, etc. Tropopause by WMO definition is the layer where the lapse rate falls down to 2 K km^{-1} and the lapse rate averaged between this level and any level within next 2 km does not exceed 2 K km^{-1}. The troposphere contains 80% of the total atmospheric mass.

Reasons for vertical motions come from: (i) sun-heated surface initiating convection, (ii) convergence / divergence of horizontal flows, (iii) horizontal flows over non-flat topographic elements, (iv) convective forces caused by the latent heat release due to water condensation.

Examples of the tropospheric composition problems

There can be numerous criteria for classification of the atmospheric composition problems. For instance, they cover a wide range of the spatial and, correspondingly, temporal scales:

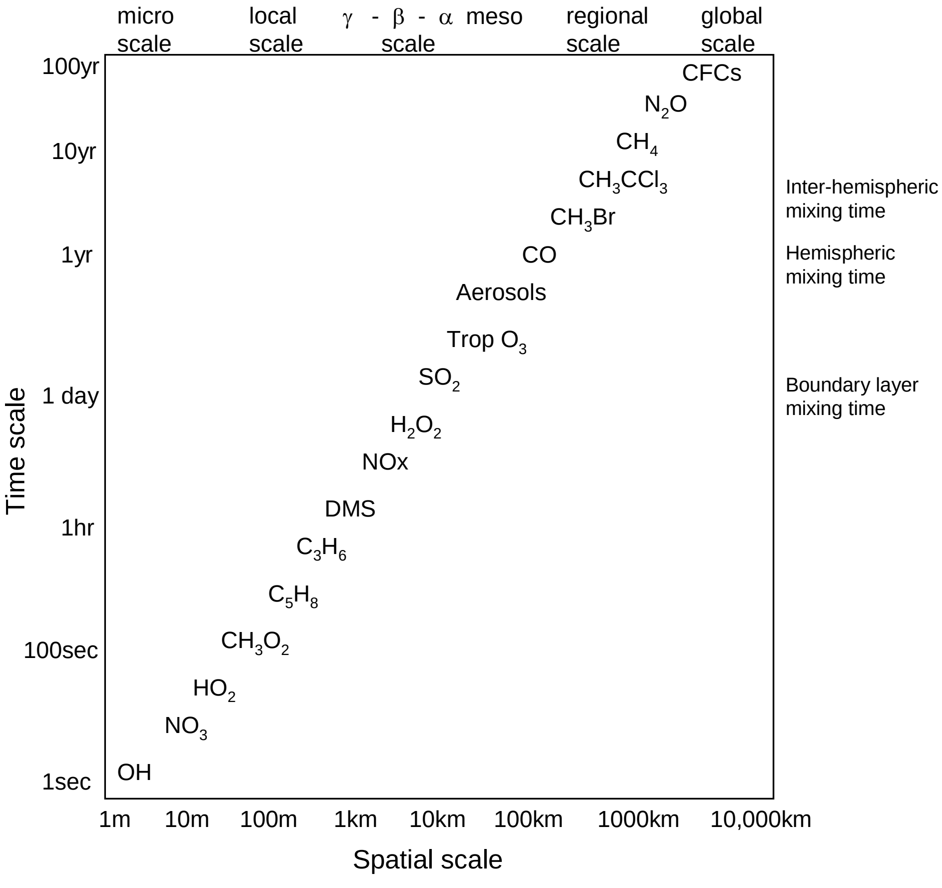


Figure 2. Spatial and temporal scales of variability of some of main atmospheric tracers.

Typically recognized scales: from local to γ-β-α mesoscale to regional to continental to global. Respectively, the species can be segregated in accordance with their characteristic lifetime and spatial scales of variability. Apart from horizontal scales, vertical scales can also be considered for some phenomena that involve more than one of the major atmospheric layers.

Table 1. Examples of scales of atmospheric composition problems and phenomena

Phenomenon	Horizontal scale	Vertical scale	Temporal scale
Urban air pollution	1-10 km	< 1 km	minutes - hours
Mesoscale air pollution (impact of cities, industrial areas)	10-100	ABL and FT up to a few km	hours - days
Regional air pollution (e.g., acidification, toxic aerosols)	100-1000	troposphere	days - weeks
Distribution of multi-media and long-living pollutants	1000-10,000	troposphere	months - years
Stratospheric ozone depletion	10,000-global	trpoposphere + stratosphere	years – decades
Greenhouse gas increase and climate forcing	1000-global	troposphere	years - decades
Aerosol – climate interaction	100-10,000	troposphere	years
Troposphere – stratosphere exchange	1 - 1000	troposphere and lower stratosphere	hours - days

Process- and scale-interactions.

All the atmospheric composition phenomena are inter-connected, so as the parts of the atmosphere. The cycling of many atmospheric constituents starts at local sources, often of anthropogenic but also natural origin, include up-scaling of the transport in both vertical and horizontal dimensions, chemical and physical inter-actions with other species and atmospheric phenomena, removal from the atmosphere, etc. Depending on the tracer, the range of scales where the specific contribution is important can vary from local to global. Example: CO, which is emitted by cars and other anthropogenic and some natural sources, takes part in urban smog formation, regional-scale ozone production, and, finally being converted to CO_2, contributes to climate forcing and interactions.

Basic terms, definitions and relations

The atmospheric **tracer** is a substance (in gaseous or aerosol form) that is transported being embedded in the atmospheric flows but does not provide any significant feedback to these flows.

Atmospheric composition is an objective characterization of the chemical constituents of air, as well as their interaction and transformation. **Air quality** is the narrower term characterizing the impact of the chemical composition to human organism. **Chemical weather**, analogously to meteorological weather, characterizes multi-scale distributions and variabilities of gases and aerosols at different time scales in light of their impacts on human health, ecosystems, the meteorological weather, and climate.

Atmospheric lifetime.

A mass budget conservation law in any imaginary volume of air at all scales is the fundamental physical principle:

$Inflow - Outflow + Emission - Removal = Accum$

(1)
$$\frac{dQ}{dt} = (F_{in} - F_{out}) + E - R$$

Stationary condition:

(2)
$$F_{in} + E = F_{out} + R$$

If some amount of a substance is present, one can introduce a formal characteristic time needed for removal of this amount (assuming that the removal rates are stable):

(3)
$$\tau = \frac{Q}{F_{in} + E} = \frac{Q}{F_{out} + R}$$

Note that, contrary to relaxation time introduced in theory of differential equations, the atmospheric life time depends on both production/loss mechanisms and the actual amount of species in the atmosphere.

Atmospheric pressure and density variations along altitude

Consider a small volume of air $dx \times dy \times dz$ that is in balance with gravity forces. The pressure at the top and the bottom is $p(z+dz)$ and $p(z)$, respectively. The difference is to be in balance with the gravitational force due to mass of this volume:

$$(p(z+dz) - p(z))\, dx\, dy = \rho\, g\, dx\, dy\, dz$$

From where:

(4)
$$\frac{dp}{dz} = -\rho(z)\, g$$

From equation of state $p = \rho R T$ ($R=8.31\ J\ mole^{-1}K^{-1}$), we finally get:

(5)
$$\frac{dp}{dz} = -\frac{p(z)}{RT} g$$

or

(6)
$$\frac{d \ln p}{dz} = -\frac{g}{RT} = -\frac{1}{H(z)}, \text{ where } H(z)=RT/g$$

Since temperature in the atmosphere varies within a factor of two while pressure changes over many order of magnitude (10-fold in the troposphere alone), one can substitute *T(z)* with some mean value to get a characteristic height scale of the atmosphere. For such a constant *H*, the pressure drops exponentially:

(7)
$$p(z) = p_0 e^{-z/H}$$

If mean tropospheric temperature of 253 K is taken, *H=7.4 km*.

Lagrangian and Eulerian coordinate systems, total and partial derivatives

The formal derivations of the physical fundamental laws (conservation of mass as the most-important for our purposes) requires consideration of small (infinitesimal) volumes of air with imaginary boundaries delineating the volume. In **Lagrangian** coordinate system this volume (approximately) follows the path of the air parcel where the volume is located. The volume contains 'marked air particles', which are always the same but exhibit the influence of varying in space external conditions. In **Eulerian** coordinate system, the control volume is fixed at the surface while the 'air particles' flow through it.

A rate of change of some field variable (e.g. concentration) following the parcel motion is called total derivative. A rate of change of the variable at the fixed point is called local derivative, which is actually just a partial derivative over time taken at a specific place.

To derive the relation between these two derivatives, let's consider a concentration *c*, which is taken in Cartensian *x-y-z* coordinates. For a given air parcel, its position will be a function of time $x=x(t), y=y(t), z=z(t)$. Following the parcel, *c* is seen a function of only time, so its variation is a "true" or total derivative of time: *dc/dt*. To relate the total derivative of concentration and its fixed-point local derivative, let's assume that at the point $c(x_0,y_0,z_0,t_0)=c_0$. Then after the parcel has moved to $x_0+\delta x$, $y_0+\delta y$, $z_0+\delta z$, $t_0+\delta t$, the concentration will change to $c_0+\delta c$ with δc that can be expressed via Taylor series as:

$$\delta c = \left(\frac{\partial c}{\partial t}\right)\delta t + \left(\frac{\partial c}{\partial x}\right)\delta x + \left(\frac{\partial c}{\partial y}\right)\delta y + \left(\frac{\partial c}{\partial z}\right)\delta z + ...$$

Dividing with δt and noting that the δc is the total derivative following the parcel motion, we get:

$$\frac{dc}{dt} = \lim_{\delta t\to 0}\frac{\delta c}{\delta t} = \left(\frac{\partial c}{\partial t}\right) + \left(\frac{\partial c}{\partial x}\right)\lim_{\delta t\to 0}\frac{\delta x}{\delta t} + \left(\frac{\partial c}{\partial y}\right)\lim_{\delta t\to 0}\frac{\delta y}{\delta t} + \left(\frac{\partial c}{\partial z}\right)\lim_{\delta t\to 0}\frac{\delta z}{\delta t} = \frac{\partial c}{\partial t} + \frac{\partial c}{\partial x}\frac{dx}{dt} + \frac{\partial c}{\partial y}\frac{dy}{dt} + \frac{\partial c}{\partial z}\frac{dz}{dt}$$

Setting $\frac{dx}{dt} \equiv u$, $\frac{dy}{dt} \equiv v$, $\frac{dz}{dt} \equiv w$ and regrouping, finally obtain:

(8)

$$\frac{dc}{dt} = \frac{\partial c}{\partial t} + \left(\frac{\partial c}{\partial x} u + \frac{\partial c}{\partial y} v + \frac{\partial c}{\partial z} w \right) = \frac{\partial c}{\partial t} + \mathbf{U} \bullet \nabla c$$

Hence, the total derivative of a species concentration along the trajectory of the air parcel motion consists of partial derivative of the concentration due to internal processes and the contribution of the transport term.

The same equation for the total derivative can be obtained from a chain differentiation rule:

$$\frac{dc(x,y,z,t)}{dt} = \frac{\partial c}{\partial t} + \frac{\partial c}{\partial x}\frac{dx}{dt} + \frac{\partial c}{\partial y}\frac{dy}{dt} + \frac{\partial c}{\partial z}\frac{dz}{dt} = \frac{\partial c}{\partial t} + \left(\frac{\partial c}{\partial x} u + \frac{\partial c}{\partial y} v + \frac{\partial c}{\partial z} w \right) = \frac{\partial c}{\partial t} + \mathbf{U} \bullet \nabla c$$

Mass conservation law. Continuity equation

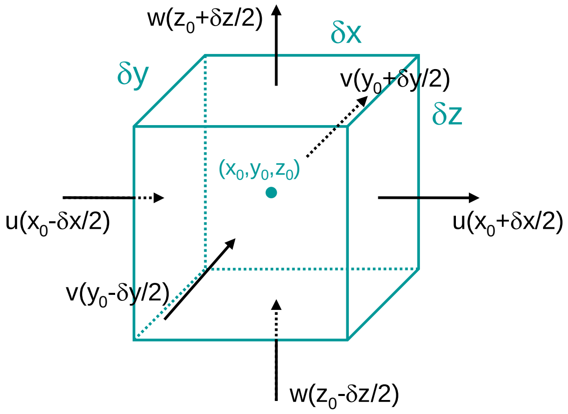


Figure 3. An illustration of the mass conservation law

Let's consider a small volume of air $\delta x \times \delta y \times \delta z$. Let this volume contains some tracer substance with the mass confined within the volume M and concentration varying in time and space $c(x,y,z,t)$. Then the mass conservation law at a point $\mathbf{r}_0=(x_0, y_0, z_0)$ would require that a sum of emission $E(\mathbf{r}_0)$ per unit volume minus sinks $R(\mathbf{r}_0)$ per unit volume plus the difference between the in- and out-flow of the substance is zero. For one-dimensional case and x-axis directed along the wind speed, this yields:

(9)

$$\frac{\partial M}{\partial t} = \left(u(x_0 - \frac{\delta x}{2}, y_0, z_0)\ c(x_0 - \frac{\delta x}{2}, y_0, z_0) - u(x_0 + \frac{\delta x}{2}, y_0, z_0) cC(x_0 + \frac{\delta x}{2}, y_0, z_0) \right) \delta y\, \delta z + \\ + (E - R) \delta x\, \delta y\, \delta z$$

Dividing (9) with $\delta x\ \delta y\ \delta z$ and setting $\delta x \to 0$, $\delta y \to 0$, $\delta z \to 0$, $\delta t \to 0$, one can obtain the mass conservation law for one-dimensional case in differential form:

(10)

$$\frac{dc}{dt} = \frac{\partial c}{\partial t} - \frac{\partial (uc)}{\partial x} + E - R$$

An extension to 3-D case reads:

(11)

$$
\begin{aligned}
\frac{\partial c}{\partial t} &= -\frac{\partial(uc)}{\partial x} - \frac{\partial(vc)}{\partial y} - \frac{\partial(wc)}{\partial z} = \\
&= -\nabla(\mathbf{u}c) + E - R \\
&= -(\mathbf{u}\nabla)cC - c(\nabla \bullet \mathbf{u}) + E - R
\end{aligned}
$$

In tensor notations, this cornerstone equation can be written as:

(12)

$$
\begin{aligned}
\frac{\partial c}{\partial t} &= -\frac{\partial}{\partial x_i}(\mathbf{u}c) + E - R \\
&= -u_i \frac{\partial c}{\partial x_i} - c\frac{\partial u_i}{\partial x_i} + E - R
\end{aligned}
$$

Comparison of (11) with the total derivative formulation (8) leads to equation for the total derivative over time (along the air parcel trajectory):

(13)

$$\frac{dc}{dt} = \frac{\partial c}{\partial t} + (\mathbf{u}\nabla)c = (\mathbf{u}\nabla)c - \nabla(\mathbf{u}c) + E - R = -c(\nabla \bullet \mathbf{u}) + E - R$$

Equations (11) and (13) are the forms of the **continuity equation** written for the particular tracer substance. It also describes the transport and often called as transport or **dispersion equation** for the tracer.

The same continuity equation can be evidently written for air density itself, with emission sources and removal processes equal to zero:

(14)

$$\frac{\partial \rho_a}{\partial t} = -\frac{\partial(u\rho_a)}{\partial x} - \frac{\partial(v\rho_a)}{\partial y} - \frac{\partial(w\rho_a)}{\partial z} = -\nabla(\mathbf{u}\rho_a) = -(\mathbf{u}\nabla)\rho_a - \rho_a(\nabla \bullet \mathbf{u})$$

(15)

$$\frac{d\rho_a}{dt} = \frac{\partial \rho_a}{\partial t} + (\mathbf{u}\nabla)\rho_a = -\rho_a(\nabla \bullet \mathbf{u})$$

The form of continuity equation for air density (14), (15) represents the approximation of a compressible air with non-uniform density distribution.

For many dispersion problems a simplified consideration under the **incompressible air** assumption is entirely sufficient. Under such consideration, volume of a selected air parcel does not change in time, thus having $\partial \rho_a / \partial t = 0$ and the continuity equation for air turning into:

(16)

$$\nabla(\mathbf{u}\rho_a) = (\mathbf{u}\nabla)\rho_a + \rho_a(\nabla \bullet \mathbf{u}) = 0$$

Even further simplification can be sometimes possible in the problems confined into the planetary boundary layer where constant uniform air density distribution can be a reasonable assumption: $\rho_a = const$. This turns the continuity equation for air into the simplest form. The air mass conservation then requires the solenoidal wind field, i.e. zero wind divergence:

(17)
$$(\nabla \bullet \mathbf{u}) = 0$$

However, most of problems require accurate representation of processes in thick vertical layers. In this case the air density cannot be considered uniform and at least its vertical variability has to be considered (incompressibility assumption still holds reasonably well in most cases). To represent the corresponding variation of the tracer concentration, it may be then convenient to introduce the mixing ratio of the tracer:

(18)
$$q = \frac{amount\_tracer}{amount\_total} \approx \frac{c}{\rho_a}$$

Then, the dispersion equation for the tracer can be written for mixing ratio:

(19)
$$\frac{\partial(q\rho_a)}{\partial t} = \frac{q\partial\rho_a}{\partial t} + \frac{\rho_a\partial q}{\partial t} = -\nabla(\mathbf{u}q\rho_a) = -\rho_a(\mathbf{u}\nabla)q - q(\mathbf{u}\nabla)\rho_a - q\rho_a(\nabla \bullet \mathbf{u})$$

Substituting the continuity equation for air (14) into (19) and dividing the result with the air density, we obtain the transport equation expressed via the tracer mixing ratio:

(20)
$$\frac{\partial q}{\partial t} = -(\mathbf{u}\nabla)q + E_q - R_q$$

Here E_q and R_q are emission sources and removal terms expressed in the mixing-ratio units: $E_R=E/\rho_a$, $R_R=R/\rho_a$.

The same equation in tensor notations will be:

(21)
$$\frac{\partial q}{\partial t} = -u_i\frac{\partial q}{\partial x_i} + E_q - R_q$$

Importantly, there is a difference between the transport equations for concentration (11), (12) and the same equation written for the mixing ratio (20), (21). The additional term $c(\nabla \bullet \mathbf{u})$ in the concentration equation represents the impact of the wind divergence onto the tracer concentration. Since this term has the same impact on air density, it cancels out when substituting the air continuity equation. Formulations become identical only under homogenous air density assumption.

Mean motion and turbulence

The transport equations for concentration (11), (12) and for mixing ratio (20), (21) have been derived from the mass budget considerations, which are fulfilled at every time moment. Respectively, wind used in the derivation is the "instant" wind. However, in real conditions such consideration has the well-known problem since the flows are always turbulent. To take this into account, the standard procedure of Reinolds decomposition is used to split the instant wind into "mean" wind **U** and "fluctuation" **u'**. Similarly, the concentration is split to mean C and fluctuation c' components:

(22)
$$\mathbf{u} = \mathbf{U} + \mathbf{u}', \quad c = C + c'$$

For **u’** it is assumed that it is a stochastic process with zero mean. Transport equation (11) then can be split into mean and fluctuation components.

(23)

$$\frac{\partial (C+c')}{\partial t} = -\nabla((\mathbf{U}+\mathbf{u'})(C+c')) + E - R$$

Opening-up the brackets, taking the average and remembering the zero mean of the fluctuations, we obtain:

(24)

$$\overline{\frac{\partial (C+c')}{\partial t}} = \frac{\partial C}{\partial t} = \overline{-\nabla((\mathbf{U}+\mathbf{u'})(C+c')) + E - R} = \\ = -\nabla(\mathbf{U}C) - \nabla\overline{(\mathbf{u'}c')} + \overline{E} - \overline{R}$$

In tensor notations:

(25)

$$\frac{\partial C}{\partial t} = -\frac{\partial}{\partial x_i}(U_i C) - \frac{\partial}{\partial x_i}\overline{(u_i' c')} + \overline{E} - \overline{R}$$

Strictly speaking, in real conditions we do not have any ensemble of the realizations of the wind and concentration fluctuations. Instead, one has to assume the ergodicity of the processes and replace ensemble averaging with time averaging. Then the bars in (24) mean averaging over specific time period, sufficiently long to reach zero of the turbulence-related fluctuations but sufficiently short so that the variations of the mean wind and concentration can be neglected. Also, one has to assume that fluctuations of the concentration are much less than its mean value and its time derivative. This allows neglecting it in the left-hand part of the equation.

One of the important consequences of the switch form ensemble averaging to time averaging is that the mean equation (24) is valid only “in-average” and, strictly speaking, is not suitable for treatment of instant concentrations or fluxes.

Closure problem.

Equation (24) has two variables to be determined: mean *C* and its fluctuation *c’* (actually, $\overline{\mathbf{u'}c'}$), which means that the system of equations is not closed. The value of *C* depends on the first-order term describing the covariance between the fluctuations of wind and concentrations. An attempt to derive the equation for $\overline{\mathbf{u'}c'}$ can be made by subtracting (24) from (23), multiplying with **u’** and averaging. For simplicity, let’s do the exercise in 1D case:

(26)

$$\overline{\left(\frac{\partial (C+c')}{\partial t} - \frac{\partial C}{\partial t}\right)u'} = -\overline{u'\frac{\partial u'c'}{\partial x}} - \overline{u'\frac{\partial Uc'}{\partial x}} - \overline{u'\frac{\partial u'C}{\partial x}} + \overline{u'\frac{\partial \overline{u'c'}}{\partial x}} = \\ = -\left[\frac{\partial \overline{u'u'c'}}{\partial x} - \overline{u'c'\frac{\partial u'}{\partial x}}\right] - \left[\frac{\partial U\overline{u'c'}}{\partial x} - \overline{Uc'\frac{\partial u'}{\partial x}}\right] - \overline{u'\frac{\partial u'C}{\partial x}}$$

Already now the first term in the right-hand-side of (26) represents the second moment of the concentration fluctuations. Thus, an attempt the equation for the concentration contains the first moment of its fluctuations, the equation for the first moment contains the second one, etc. Therefore, under no circumstances we can obtain the closed set of equations.

The chain of equations (24)/(25) and (26) that expresses the dependence of lower moments on higher ones has to be broken at some point with an assumption that the higher-order moment depends somehow on the lower-order one(s). That would result in equal number of equations and unknowns.

K-theory: first-order closure

The most-popular closure approach is based on so-called mixing length theory, which involves physical (or mechanistic) interpretation of turbulence.

Let's consider a liquid moving along the x-axis with some feature (e.g. concentration) varying along y-axis. Let's consider the turbulence phenomenon as a random relocation of parcels of the fluid across the main wind direction, i.e. along the y-axis (Figure 4). Such parcels, being relocated with some finite speed, keep their properties over some time and, consequently, over some distance (respectively denoted as t_a and λ_a in Figure 4) and then mix back to the main liquid accepting its features at the new location.

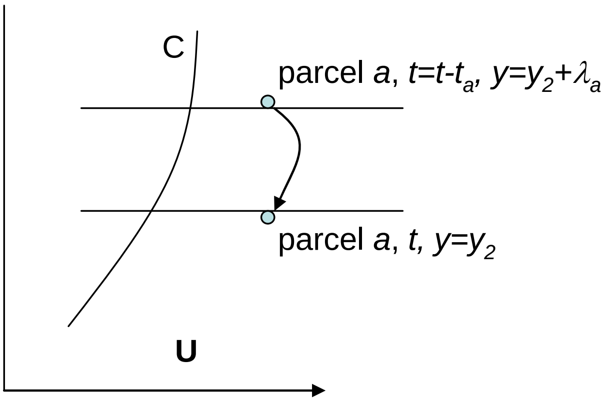


Figure 4. An illustration of the mixing length principle

If a parcel a of a liquid is at $y=y_2+\lambda_a$ at time moment $t=t-t_a$ and then relocated to y_2 during time interval t_a, we obtain that the fluctuation of the concentration at the point y_2 and time t will be the difference of the mean concentration at these two points - y_2 and $y_2+\lambda_a$:

(27)
$$c'(y_2,t) = C(y_2+\lambda_a, t-t_a) - C(y_2,t)$$

This can be expanded in Taylor series:

(28)
$$c'(y_2,t) = \lambda_a \frac{\partial C}{\partial y} - t_a \frac{\partial C}{\partial t} + \frac{\lambda_a^2}{2}\frac{\partial^2 C}{\partial y^2} + \frac{t_a^2}{2}\frac{\partial^2 C}{\partial t^2} + \dots$$

Let's also assume the stationarity of the case to eliminate the time derivatives. Let also the value of λ_a be small in comparison with other characteristic scales of the flow, in particular, the scales of substantial wind or concentration changes. Then we can neglect all higher terms in (28) and the concentration fluctuation becomes a function of mean field features and the mixing length λ_a:

(29) $$c'(y_2,t)=\lambda_a\frac{\partial C}{\partial y}$$

Multiplying it by *v'*$_a$ the turbulent fluctuation in the y-direction associated with the *a*-eddy becomes:

$$v'c'=v'\lambda_a\frac{\partial C}{\partial y}$$

After averaging over all the fluctuations, we obtain:

$$\overline{v'c'}=\overline{v'\lambda_a}\frac{\partial C}{\partial y}$$

To roughly estimate the term $\overline{v'\lambda_a}$, let's introduce the maximum distance over which the parcels maintain their features as *L* and the turbulence intensity as $\hat{\upsilon}=\sqrt{\overline{v'^2}}$. Then

(30) $$\overline{v'c'}=-const\;L\;\hat{v}\;\frac{\partial C}{\partial y}=-K\frac{\partial C}{\partial y}$$

Where *const* is some positive value reflecting the extent of correlation of the relocation distances and the wind fluctuations. Here the *K* proportionality is the eddy diffusivity for scalar, such as tracer concentration.

The main physical value of this idea is that turbulent fluxes are now connected with the gradients of the mean fields and the transport equation will contains only one unknown variable – mean concentration *C*:

(31) $$\mathbf{F}=\overline{\mathbf{u'}c'}=-\mathbf{K}\nabla C\;\text{ or: }\;F_i=\overline{u_i'c'}=-K_{ij}\frac{\partial C}{\partial x_j}$$

Substituting this term into (24) and (25) we obtain the well-known dispersion equation:

(32) $$\frac{\partial C}{\partial t}=-\nabla(\mathbf{U}C)+\nabla\mathbf{K}\,\nabla C+\overline{E}-\overline{R}$$

or:

(33) $$\frac{\partial C}{\partial t}=-\frac{\partial}{\partial x_i}(U_iC)+\frac{\partial}{\partial x_i}K_{ij}\frac{\partial C}{\partial x_j}+\overline{E}-\overline{R}$$

This form of the equation does not contain any more the small-scale characteristics, instead including the unknown "turbulent diffusion coefficient" or "eddy diffusivity" K_{ij}, which, in general case, is a tensor with components that must be measured or somehow parameterized from empirical data.

In order to conclude the derivations, let's convert the transport equation (33) into the mixing ratio terms. In order to do that one has to notice that speculations (27) - (30) and resulting assumption (31) are valid only when the density of the fluid is constant. Indeed, if the parcel after its relocation appears to be of different density than the surrounding liquid it will adjust its density by changing its volume much faster than the concentration approaches the equilibrium. It means that the tracer concentration in the parcel will be adjusted as well and with the same proportion as the density. Hence, for compressible liquid equation (27) should be written to the concentration weighted with the liquid density, i.e. for the mixing ratio:

(34)
$$q'(y_2,t) = Q(y_2+\lambda_a, t-t_a) - Q(y_2,t)$$

Here *q'* and *Q* are mixing ratio fluctuation and the mean value, respectively.

Then the K-theory main formula (31) becomes:

$$\mathbf{F} = \rho\,\overline{\mathbf{u}'q'} = -\rho\,\mathbf{K}\nabla(C/\rho) \quad \textbf{or:} \quad F_i = \rho\,\overline{u_i'c'} = -\rho\,K_{ij}\frac{\partial(C/\rho)}{\partial x_j}$$

The final form of the transport equation for the compressible air will be:

(35)
$$\frac{\partial C}{\partial t} = -\nabla(\mathbf{U}C) + \nabla\rho\,\mathbf{K}\,\nabla(C/\rho) + \overline{E} - \overline{R}$$

or:

(36)
$$\frac{\partial C}{\partial t} = -\frac{\partial}{\partial x_i}(U_iC) + \frac{\partial}{\partial x_i}\rho\,K_{ij}\frac{\partial(C/\rho)}{\partial x_j} + \overline{E} - \overline{R}$$

The dispersion equation (35), (36) has to be accompanied with initial and boundary conditions, which can depend on specific problem and in comparatively general case can look like the following:

(37)
$$\begin{aligned}
&C(t=0) = C_0(\mathbf{x})\\
&C(t,\mathbf{x}) = C_{\partial\Omega}(t,\mathbf{x}),\ \mathbf{x}\in:\partial\Omega:\ U_n<0\\
&\frac{\partial C}{\partial n}(t,\partial\Omega) = 0:\ U_n \ge 0\\
&\frac{\partial C}{\partial z}(t,\mathbf{x}) = 0:,\ \mathbf{x}\in:\Gamma_H\\
&K_{33}\frac{\partial C}{\partial z}(t,\mathbf{x}) = v_d C(z_0):,\ \mathbf{x}\in:\Gamma_0
\end{aligned}$$

Here $\partial\Omega$ is lateral boundary (western, southern , eastern and northern vertical planes limiting the domain), Γ_H and Γ_0 are the upper and lower domain limits, *n* is the normal vector starting at a domain border and directed outside, C_0 is the initial distribution of the concentrations, $C_{\partial\Omega}$ is the concentration distribution at the lateral domain boundaries; finally, v_d is the proportionality coefficient with a unit of speed and usually called dry deposition velocity.

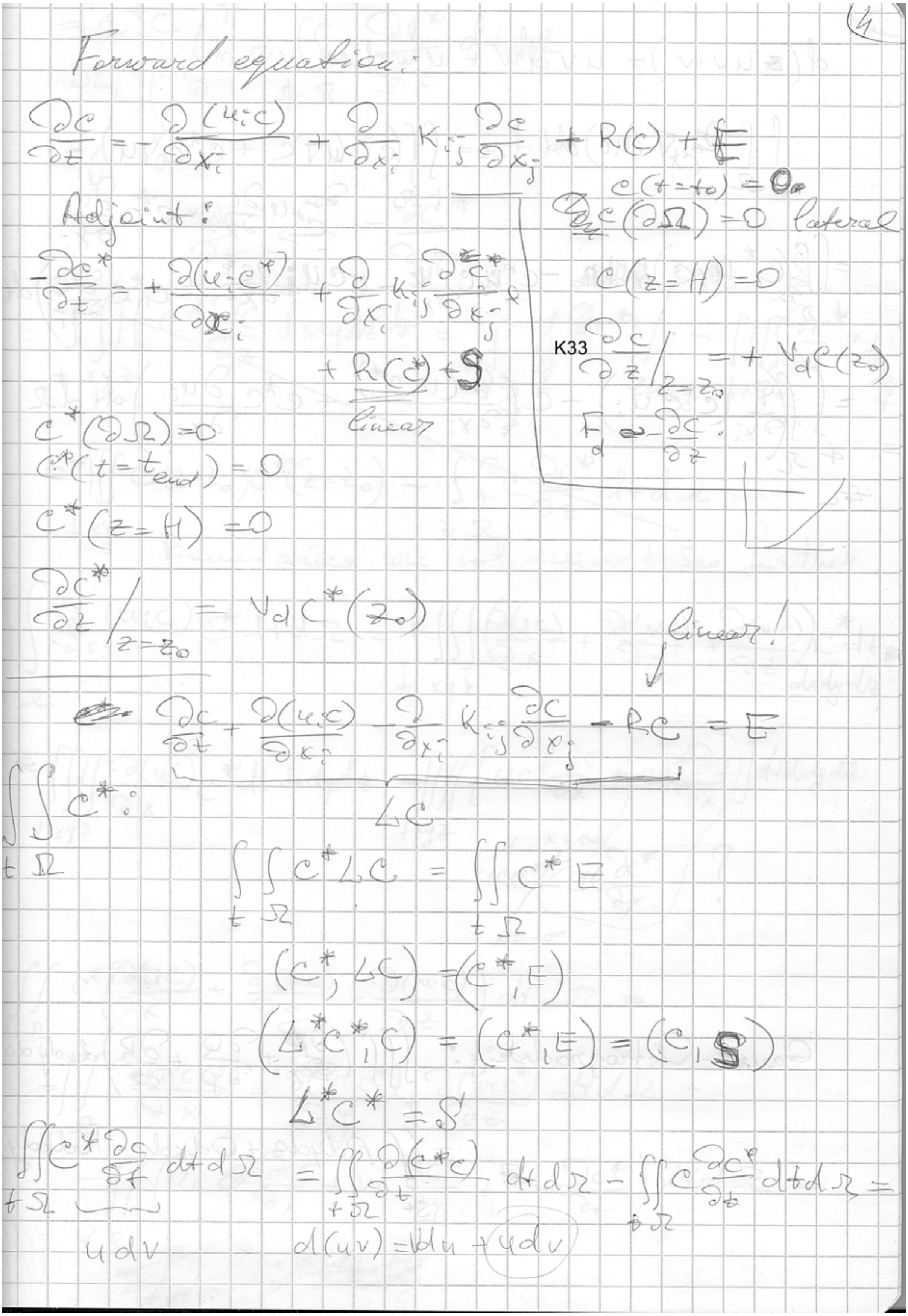
A physical sense of the boundary conditions in (37) is that (i) inflow into the domain is dictated by the outside concentration distribution $C_{\partial\Omega}$, (ii) outflow is free, (iii) there is no exchange through the upper domain boundary, and (iv) turbulent diffusion flow at the lower domain boundary is determined by the vertical diffusion coefficient, dry deposition velocity coefficient and concentration at some near-surface level z_0.

The main weakness of the K-approach, as well as the mixing-length theory, is that the unknown terms describing the turbulent fluxes are replaces with unknown eddy diffusivity coefficient, which is a function of time, space, scale and even some model technicalities.

Let briefly list the main problems associated with K_{ij}.

1. Since turbulent eddies are 3-D structures, there is no justification for assuming this tensor to be diagonal. It is rather an experimental result that in most practically important cases it can be taken diagonal without drastic consequences.
2. Since the spectrum of turbulence is inhomogeneous, the actual values will depend on the scale of the problem and the model resolution. Indeed, the larger the model grid cell is the larger eddies are to be considered as unresolved turbulence and, consequently, the higher will be K_{ij}.
3. In presence of a horizontal or vertical wall the maximum size of eddies is limited with the distance to this wall, which leads to dependence of K_{ij} on this distance.
4. The eddy diffusivity even depends on the plume characteristics: the wider and longer plume includes larger eddies as turbulent motions inside the plume, thus its further expansion happens faster.
5. Finally, the specific resolution of the dispersion, as well as meteorological, problems becomes “no-go area” (this, however, is a problem of all current parameterizations). These are the resolutions, which are too coarse to resolve explicitly e.g. the convective cells (a scale of ~1km) but too fine to put them entirely into subgrid processes.

Adjoint advection-diffusion equation



(4)

Forward equation:

$$\frac{\partial c}{\partial t} = -\frac{\partial (u_i c)}{\partial x_i} + \frac{\partial}{\partial x_i} K_{ij} \frac{\partial c}{\partial x_j} + R(c) + E$$

$c(t=t_0) = 0$

$c(\partial\Omega) = 0$ lateral

$c(z=H) = 0$

K33 $\left.\frac{\partial c}{\partial z}\right|_{z=z_0} = + V_d c(z_0)$

$F_d \sim -\frac{\partial c}{\partial z}$

Adjoint:

$$-\frac{\partial c^*}{\partial t} = +\frac{\partial (u_i c^*)}{\partial x_i} + \frac{\partial}{\partial x_i} K_{ij} \frac{\partial c^*}{\partial x_j} + \underbrace{R(c^*)}_{\text{linear}} + S$$

$c^*(\partial\Omega) = 0$

$c^*(t=t_{end}) = 0$

$c^*(z=H) = 0$

$\left.\frac{\partial c^*}{\partial z}\right|_{z=z_0} = V_d c^*(z_0)$

linear!

$$\underbrace{\frac{\partial c}{\partial t} + \frac{\partial (u_i c)}{\partial x_i} - \frac{\partial}{\partial x_i} K_{ij} \frac{\partial c}{\partial x_j} - Rc}_{Lc} = E$$

$\int_t \int_\Omega c^* :$

$$\int_t \int_\Omega c^* Lc = \iint_{t\,\Omega} c^* E$$

$$(c^*, Lc) = (c^*, E)$$

$$(L^* c^*, c) = (c^*, E) = (c, S)$$

$$L^* c^* = S$$

$$\iint_{t\,\Omega} \underbrace{c^* \frac{\partial c}{\partial t}}_{u\,dv} dt\,d\Omega = \iint_{t\,\Omega} \frac{\partial (c^* c)}{\partial t} dt\,d\Omega - \iint_{t\,\Omega} c \frac{\partial c^*}{\partial t} dt\,d\Omega =$$

$$d(uv) = v\,du + \boxed{u\,dv}$$

6

Formally, the dispersion problem:

$$\frac{\partial c}{\partial t} + \frac{\partial (u_i c)}{\partial x_i} - \frac{\partial}{\partial x_i} \rho_a K_{ij} \frac{\partial (c/\rho_a)}{\partial x_j} + R(c) = E$$

Boundary conditions: $c = c_{\partial\Omega}$ $\partial\Omega$, $u_n \leq 0$ no inflow

n is ~~assumed~~ a normal vector directed outside Ω

$$\frac{\partial c}{\partial n} = 0 : \partial\Omega, \; u_n \geq 0$$

$$\rho K_{33} \frac{\partial c/\rho}{\partial z} = V_d c : x_3 = x_3^0$$

$$\frac{\partial c/\rho}{\partial z} = 0 : x_3 = H$$

This will be called the forward dispersion problem.

$$\int_t\int_\Omega c^* : \int_t\int_\Omega c^* \frac{\partial c}{\partial t} \, d\Omega \, dt + \int_t\int_\Omega c^* \frac{\partial (u_i c)}{\partial x_i} \, d\Omega \, dt - \int_t\int_\Omega c^* \frac{\partial}{\partial x_i} \rho_a K_{ij} \frac{\partial \left(\frac{c}{\rho}\right)}{\partial x_j} \, d\Omega \, dt$$

$$+ \int_t\int_\Omega c^* R(c) \, d\Omega \, dt = \int_t\int_\Omega c^* E \, dt \, d\Omega$$

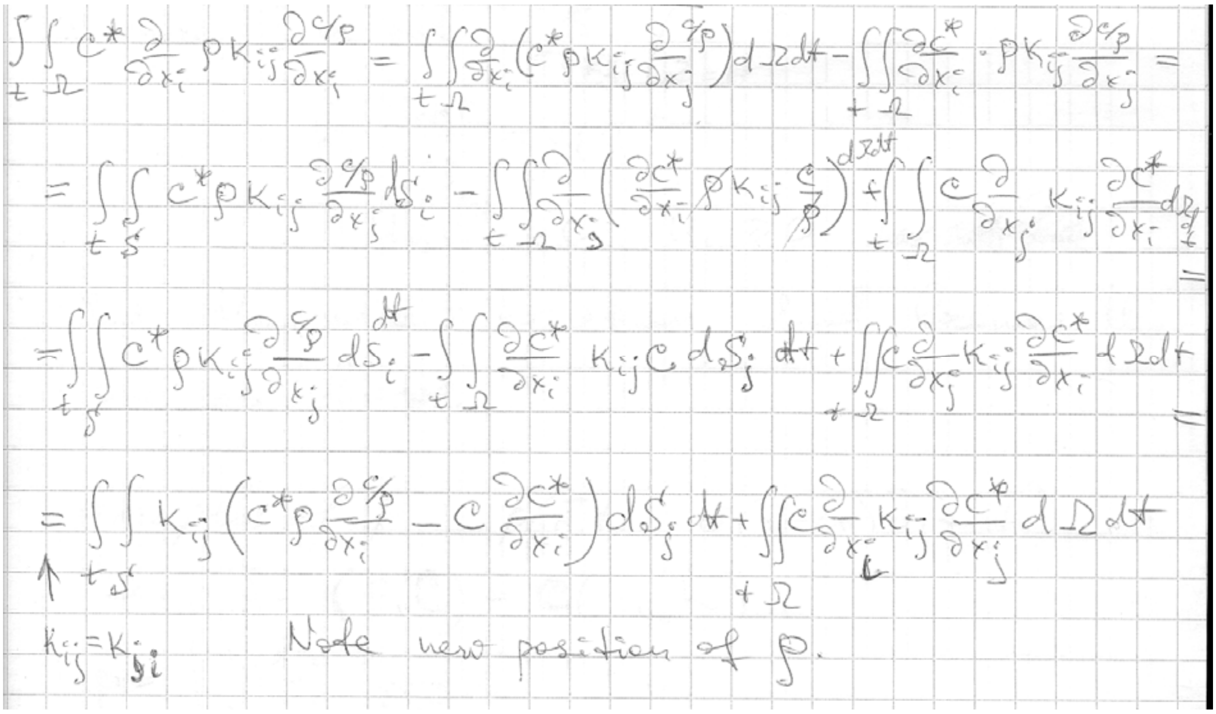
One-by-one:

$$\int_t\int_\Omega c^* \frac{\partial c}{\partial t} \, d\Omega \, dt = \int_\Omega c^* c \Big|_0^T d\Omega - \int_t\int_\Omega c \frac{\partial c^*}{\partial t} \, d\Omega \, dt$$

$$\int_t\int_\Omega c^* \frac{\partial (u_i c)}{\partial x_i} \, d\Omega \, dt = \int_t\int_\Omega \frac{\partial (u_i c c^*)}{\partial x_i} \, d\Omega \, dt -$$

$$- \int_t\int_\Omega u_i c \frac{\partial c^*}{\partial x_i} \, d\Omega \, dt \underset{\substack{\uparrow \\ \text{G-O fla}}}{=} \int_t\int_{\partial\Omega} c c^* \underbrace{\left(u_1 dx_2 dx_3 + u_2 dx_1 dx_3 + u_3 dx_1 dx_2\right)}_{= u_n dS_n}$$

$$- \int_t\int_\Omega c \, u_i \frac{\partial c^*}{\partial x_i} \, d\Omega \, dt.$$



$$\int_t\int_\Omega c^* \frac{\partial}{\partial x_i} \rho K_{ij} \frac{\partial c/\rho}{\partial x_i} = \int_t\int_\Omega \frac{\partial}{\partial x_i}\left(c^*\rho K_{ij}\frac{\partial c/\rho}{\partial x_j}\right) d\Omega dt - \int_t\int_\Omega \frac{\partial c^*}{\partial x_i}\cdot \rho K_{ij}\frac{\partial c/\rho}{\partial x_j} =$$

$$= \int_t\int_S c^*\rho K_{ij}\frac{\partial c/\rho}{\partial x_j} dS_i - \int_t\int_\Omega \frac{\partial}{\partial x_j}\left(\frac{\partial c^*}{\partial x_i}\cancel{\rho} K_{ij}\frac{c}{\cancel{\rho}}\right) d\Omega dt + \int_t\int_\Omega c\frac{\partial}{\partial x_j} K_{ij}\frac{\partial c^*}{\partial x_i} d\Omega dt =$$

$$= \int_t\int_S c^*\rho K_{ij}\frac{\partial c/\rho}{\partial x_j} dS_i\, dt - \int_t\int_\Omega \frac{\partial c^*}{\partial x_i} K_{ij} c\, dS_j\, dt + \int_t\int_\Omega c\frac{\partial}{\partial x_j} K_{ij}\frac{\partial c^*}{\partial x_i} d\Omega dt =$$

$$= \int_t\int_S K_{ij}\left(c^*\rho\frac{\partial c/\rho}{\partial x_i} - c\frac{\partial c^*}{\partial x_i}\right) dS_j\, dt + \int_t\int_\Omega c\frac{\partial}{\partial x_i} K_{ij}\frac{\partial c^*}{\partial x_j} d\Omega dt$$

↑

$K_{ij}=K_{ji}$ Note new position of ρ.

Therefore, our adjoint equation becomes:

$$-\frac{\partial c^*}{\partial t} - u_i \frac{\partial c^*}{\partial x_i} - \frac{\partial}{\partial x_i} K_{ij} \frac{\partial c^*}{\partial x_j} + R c^* = p$$

We also have to require: ~~e.g.,~~ $C(t) = C(t+T)$

$$c^*(t) = c^*(t+T)$$

or: $C(0) = 0$ or....

$$c^*(T) = 0$$

$$\int_t \int_S c c^* u_n \, dS_n \, dt = \int_t \int_{\partial\Omega} c c^* u_n^+ \, dS_n \, dt$$

$c = 0 /_{u_n < 0}$ — if $C_\Omega = 0$; $w = 0 /_{z_0, H}$

$$\int_t \int_S K_{ij} \left(c^* p \frac{\partial c/p}{\partial x_i} - c \frac{\partial c^*}{\partial x_i} \right) dS_j \underset{\substack{\uparrow \\ \text{skip horiz. diff.} \\ \text{at the border.}}}{=} - \int_t \int_{S_{z_0}} \left(c^* V_d C - K_{33} c \frac{\partial c^*}{\partial x_3} \right) dS_{z_0} +$$

$$+ \int_t \int_{S_H} \left(-c \frac{\partial c^*}{\partial x_3} \right) dS_H$$

Therefore, if we put: $\left. \frac{\partial c^*}{\partial x_3} \right|_{z=H} = 0$

$$K_{33} \frac{\partial c^*}{\partial x_3} = V_d c^*$$

$$u_n c^* = 0 \quad u_n \geq 0$$

Then the adjoint equation indeed becomes the one as above

Examples of P: $p(\vec{x}, t) = \delta(\vec{x} - \vec{x}_0)\, \delta(t - t_0)$

Then $J = (p, c) = c(x_0, t_0)$

$$p = \begin{cases} 1 & \vec{x} \in D \\ 0 & \vec{x} \in \Omega \setminus D \end{cases}$$

Outline of analytical approaches to solution of advection-diffusion equation

Current section will consider a couple of examples illustrating the ways to solve the advection-diffusion equation analytically and highlighting the role of its components.

Uniqueness of the solution of ADE

Before moving into this direction, one has to ensure that the solution is unique at least in some boundary conditions – and find out the limitations, if any.

Let assume that we have two solutions C_1 and C_2 for the equation (36). Then let $\varphi=C_2-C_1$ be the difference of these two solutions, which start from the same point, i.e. $\varphi(0)=0$. Assuming linearity of the chemical transformation term and noting that the emission term is independent on C and the same for both solutions, we end up with the following equation for φ:

(38)
$$\frac{\partial\varphi}{\partial t} = -\frac{\partial}{\partial x_i}(U_i\varphi) + \frac{\partial}{\partial x_i}\rho K_{ij}\frac{\partial(\varphi/\rho)}{\partial x_j} - R\varphi$$

Let's now multiply it with φ and integrate over the considered domain and time period. For advection term, let's also assume the density-independent continuity equation for wind: *div* $\boldsymbol{U}=0$ and recall that for incoming flows through the boundary surface for both solutions are identical, i.e. this term will be zero for their difference. Then:

$$\int_0^T\int_\Omega \varphi\frac{\partial}{\partial x_i}(U_i\varphi)d\Omega\,dt = \int_0^T\int_\Omega \frac{\partial}{\partial x_i}(U_i\varphi^2/2)d\Omega\,dt = \int_0^T\int_S \frac{U_n\varphi^2}{2}dS_n\,dt =$$
$$= \int_0^T\int_S \frac{U_n^+\varphi^2}{2}dS_n^+\,dt - \int_0^T\int_S \frac{U_n^-\varphi^2}{2}dS_n^-\,dt$$

For turbulent diffusion term, let's simplify the equation by assuming constant density:

$$\int_0^T\int_\Omega \varphi\frac{\partial}{\partial x_i}\rho K_{ij}\frac{\partial(\varphi/\rho)}{\partial x_j}d\Omega\,dt = \int_0^T\int_\Omega \left(\frac{\partial}{\partial x_i}\varphi K_{ij}\frac{\partial\varphi}{\partial x_j} - K_{ij}\frac{\partial\varphi}{\partial x_i}\frac{\partial\varphi}{\partial x_j}\right)d\Omega\,dt =$$
$$= \int_0^T\int_S \varphi K_{nn}\frac{\partial\varphi}{\partial x_n}dS_n\,dt - \int_0^T\int_\Omega K_{ii}\left(\frac{\partial\varphi}{\partial x_i}\right)^2 d\Omega\,dt$$

Bringing together all terms, we obtain:

(39)
$$\int_\Omega \frac{\varphi^2}{2}d\Omega\bigg|_0^T + \int_0^T\int_S \frac{U_n^+\varphi^2}{2}dS_n^+\,dt + \int_0^T\int_\Omega K_{ii}\left(\frac{\partial\varphi}{\partial x_i}\right)^2 d\Omega\,dt + R\int_0^T\int_\Omega \varphi^2 d\Omega\,dt = \int_0^T\int_S \varphi K_{nn}\frac{\partial\varphi}{\partial x_n}dS_n\,dt$$

On the left-hand side, all terms are essentially non-negative and can be zero only if φ itself is zero, i.e. the two solutions are identical. Therefore, it is enough for the uniqueness of the solution to have the boundary conditions, which make the right-hand-side term zero for $\varphi=0$ and negative otherwise.

For the boundary conditions (37), this term becomes:

$$\int_0^T \int_S \varphi K_{nn} \frac{\partial \varphi}{\partial x_n} dS_n\, dt = -\int_0^T \int_S \varphi K_{nn} \frac{\partial \varphi}{\partial z} dS_n\, dt = -\int_0^T \int_{\Gamma_0} v_d K_{nn} \varphi^2 dS_n\, dt$$

Substituting it into (39) we obtain that the only φ satisfying the equation is zero, i.e. the initial and boundary conditions (37) lead to unique solution of the dispersion problem.

One has to keep in mind numerous assumptions and simplifications made during the derivation. Some of them, such as constant-density requirement can be lifted up at a price of somewhat bulkier derivation.

Examples of analytical solution of ADE

Analytical solution of ADE in general case is impossible, therefore a set of strong assumptions has to be made in order to obtain the solution for some specific cases. Below we consider a few such cases in order to illustrate the meaning of each of them.

Example 1.

The simplest case to consider is pure advection of a passive tracer with zero diffusion term and constant wind speed, along which we then direct the x-axis. Let also the boundaries be far away, so that the boundary conditions do not affect the solution at the interval considered. The problem becomes one-dimensional with the initial condition defined for $t=0$:

(40)
$$\frac{\partial C(t,x)}{\partial t} + u \frac{\partial C(t,x)}{\partial x} = 0, \quad C(0,x_0) = C_0(x_0)$$

Since wind component u is constant in time and space, one can consider the characteristic line $\partial x / \partial t = u$, i.e. $x=ut+x_0$, and, following the method of characteristics, look for the solution of (40), which must constant along this line:

(41)
$$C(t, x_0 + ut) = C(0, x_0) = C_0(x_0).$$

From (41), it is seen that the solution of the pure-advection equation with constant wind and given initial conditions represents, in fact, a move of this initial shape along the x-axis with the speed u – as one would intuitively expect.

Example 2.

Let now consider both advection and diffusion terms, still with constant parameters in both time and space and with a single-point non-zero initial condition. Again directing x-axis along the wind, we obtain the following form of ADE:

$$\frac{\partial C(t,x)}{\partial t}+u\frac{\partial C(t,x)}{\partial x}=K_{xx}\frac{\partial^2 C}{dx^2}+K_{yy}\frac{\partial^2 C}{dy^2}+K_{zz}\frac{\partial^2 C}{dz^2},\quad C(0,\mathbf{x})=C_0\delta(\mathbf{x})$$

Let’s look for the solution as: $C(t,\mathbf{x})=c_x(t,x)c_y(t,y)c_z(t,z)$. Then:

$$\frac{\partial c_x(t,x)}{\partial t}+u\frac{\partial c_x(t,x)}{\partial x}=K_{xx}\frac{\partial^2 c_x}{dx^2},\quad c_x(0,x)=c_{0x}\delta(x)=C_0^{1/3}\delta(x)$$

$$\frac{\partial c_y(t,x)}{\partial t}=K_{yy}\frac{\partial^2 c_y}{dy^2},\quad c_y(0,y)=c_{0y}\delta(y)=C_0^{1/3}\delta(y)$$

$$\frac{\partial c_z(t,x)}{\partial t}=K_{zz}\frac{\partial^2 c_z}{dz^2},\quad c_z(0,z)=c_{0z}\delta(z)=C_0^{1/3}\delta(z)$$

Let’s look for a solution via Fourier transform:

$$\Sigma(t,\alpha)=F[c_x(t,x)]=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}c_x(t,x)e^{-i\alpha x}dx$$

(42)

$$\int_{-\infty}^{\infty}\left[\frac{\partial c_x(t,x)}{\partial t}+u\frac{\partial c_x(t,x)}{\partial x}-K_{xx}\frac{\partial^2 c_x}{dx^2}\right]e^{-i\alpha x}dx=$$
$$=\frac{\partial\Sigma}{\partial t}+iu\alpha\Sigma+\alpha^2K_{xx}\Sigma=0,\qquad \Sigma(0,\alpha)=\frac{C_0^{1/3}}{\sqrt{2\pi}}$$

Solution for (42) is straightforward:

(43)

$$\Sigma(\alpha,t)=\frac{C_0^{1/3}}{\sqrt{2\pi}}\exp\left(-(i\alpha u+K_{xx}a^2)t\right)$$

Inverting the transform, we obtain:

$$c_x(t,x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\Sigma(t,\alpha)e^{i\alpha x}d\alpha$$

Deal with the exponent first by completing the square:

$$(K_{xx}\alpha^2+iu\alpha)t-ix\alpha=K_{xx}t\alpha^2-i\alpha(x-ut)=K_{xx}t\alpha^2-i\alpha(x-ut)+\frac{(x-ut)^2}{4K_{xx}t}-\frac{(x-ut)^2}{4K_{xx}t}=$$

$$=\left(\alpha\sqrt{K_{xx}t}-i\frac{x-ut}{2\sqrt{K_{xx}t}}\right)^2-\frac{(x-ut)^2}{4K_{xx}t}$$

Denoting $\eta=\alpha\sqrt{K_{xx}t}-i\frac{x-ut}{2\sqrt{K_{xx}t}},\quad d\eta=\sqrt{K_{xx}t}\;d\alpha$, we can write the solution as:

$$c_x(t,x)=\frac{C_0^{1/3}}{2\pi\sqrt{K_{xx}t}}\exp\left(-\frac{(x-ut)^2}{4K_{xx}t}\right)\int_{-\infty}^{\infty}e^{-\eta^2}d\eta=\frac{C_0^{1/3}}{2\sqrt{\pi K_{xx}t}}\exp\left(-\frac{(x-ut)^2}{4K_{xx}t}\right)$$

Evidently,

$$c_y(t,y) = \frac{C_0^{1/3}}{2\sqrt{\pi K_{xx} t}} \exp\left(-\frac{y^2}{4K_{xx}t}\right), \quad c_z(t,z) = \frac{C_0^{1/3}}{2\sqrt{\pi K_{xx} t}} \exp\left(-\frac{z^2}{4K_{xx}t}\right)$$

Then the final solution becomes:

(44)

$$C(t,\mathbf{x}) = \frac{C_0}{8(\pi t)^{3/2}\sqrt{K_{xx}K_{yy}K_{zz}}} \exp\left(-\frac{(x-ut)^2}{4K_{xx}t} - \frac{y^2}{4K_{yy}t} - \frac{z^2}{4K_{zz}t}\right)$$

The formulations (44) constitute the simplest form of so-called Gaussian model equations. Such models were the first ones used in numerous dispersion applications where the assumptions behind the derivation are more or less fulfilled. The most-crucial ones are the requirements of constant wind and eddy diffusivity terms, which can be considered reasonable only at local scales. Numerous modifications of the assumptions expanded the range of applicability of the Gaussian formulations to more realistic profiles of eddy diffusivity and varying wind. However, the applicability of such systems for multi-source situation with complicated chemistry is very problematic due to still-extreme level of simplifications and inevitable linearity requirement. Another assumption, which cannot be lifted is the stationarity of the pattern.

Approaches to numerical solutions of ADE

For atmospheric composition tasks, by far the most widely used approach is direct numerical solution of the dispersion equation. The rest of the course will be dedicated to this topic.

Before starting the in-depth analysis, let’s attempt to solve seemingly the simplest part of the problem – pure advection equation for a passive tracer (40). For the explicit algorithm:

(45)
$$\frac{C_i^{k+1} - C_i^k}{\Delta t} = -u\frac{C_i^k - C_{i-1}^k}{\Delta x}$$

and for the implicit algorithm:

(46)
$$\frac{C_i^{k+1} - C_i^k}{\Delta t} = -u\frac{C_i^{k+1} - C_{i-1}^{k+1}}{\Delta x}$$

In the above basic discretization schemes, super-script indices correspond to time axis and the sub-scripts ones – to x-axis. To analyze the quality of the obtained solution, let’s expand *C* in Taylor series:

(47)
$$C(t,x) = C_i^k + \frac{\partial C}{\partial t}\bigg|_i^k (t-t_k) + \frac{\partial C}{\partial x}\bigg|_i^k (x-x_i) + \frac{\partial^2 C}{\partial t^2}\bigg|_i^k \frac{(t-t_k)^2}{2} + \frac{\partial^2 C}{\partial x^2}\bigg|_i^k \frac{(x-x_i)^2}{2} + \ldots$$

Substituting (47) into (45), we obtain the equation for concentration for the explicit scheme (noting that sub-script *i-1* corresponds to $-\Delta x$ and super-script *k+1* – to Δt):

$$\frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} = u\frac{\partial^2 C}{\partial x^2}\frac{\Delta x}{2} - \frac{\partial^2 C}{\partial t^2}\frac{\Delta t}{2}$$

Differentiating the original equation (40) with *t*, one can obtain that:

$$\frac{\partial^2 C}{\partial t^2} = u^2\frac{\partial^2 C}{\partial x^2}$$

Then

(48)
$$\frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} = \left(\frac{u\Delta x - u^2\Delta t}{2}\right)\frac{\partial^2 C}{\partial x^2}$$

Hence, the advection equation after discretization turned into advection-diffusion equation with the diffusivity term dependent on both physical parameter *u* and numerical setup. Its comparison with typical eddy diffusivity values can highlight the level of the problem. For normal atmospheric conditions in middle latitudes, $K_{xx}<100\ m^2sec^{-1}$, $u\sim10\ m\ sec^{-1}$. Typical setup of a regional-scale dispersion problem operates with $\Delta x\sim10km=10{,}000m$, $\Delta t\sim10min=600sec$. The artificial numerical viscosity coefficient will then be ~$5*10^4$ m^2 sec^{-1}, i.e. 100-1000 times stronger than any realistic turbulent diffusion. It is equal to zero only in case $\Delta x=u\Delta t$, which in realistic cases can be fulfilled only occasionally.

More, this diffusion coefficient $D_{num} = (u\Delta x - u^2\Delta t)/2$ can be both positive and negative. In the latter case, the problem is ill-posed according to Hadamard stability criterion, which means that the solution is unstable with regard to any changes in the initial conditions.

The threshold at which D_{num} changes its sign is when the Courrant number $Courant = u/(\Delta x/\Delta t) = 1$.

The scheme is unstable for any *u<0*, which requires re-orientation of the *x*-axis so that it is always co-directed with the wind.

The situation somewhat improves for implicit scheme (46) where stability is kept for any *u>0* but it for *u<0* again the re-orientation of the axis is inevitable to keep the scheme stability:

$$\frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} = -\left(\frac{u\Delta x + u^2\Delta t}{2}\right)\frac{\partial^2 C}{\partial x^2} \tag{ 49}$$

Note that in both (48) and (49) adding the "true" diffusion term does not change anything because this term is much smaller than the numerical terms.

Other discretization schemes are possible.

Symmetrical scheme:

$$\frac{\partial C}{\partial x} = \frac{C_{i+1}^k - C_{i-1}^k}{2\Delta x} \tag{ 50}$$

seemingly gives the second-order approximation for the derivatives. However, its application in the time-explicit algorithm leads to a disaster (S&P, p.1130) because the obtained algorithm is unstable for any model steps.

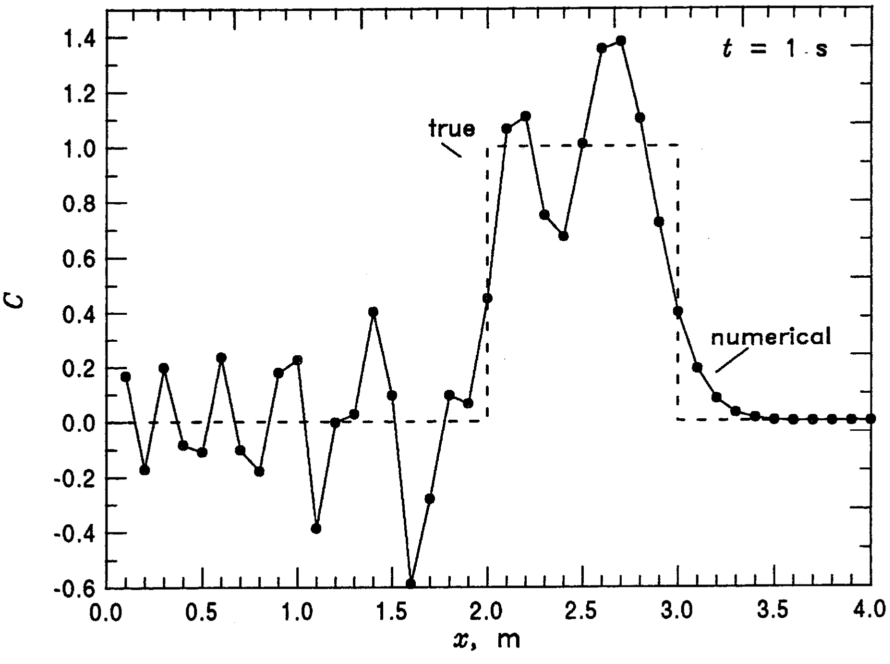


Figure 5. Example of numerical solution with symmetrical discretization **(50**) after 50 time steps. Wind blows from left to right, initial squared-shape plume position is from *x=1m* and *x=2m*.

Indeed, substitution of Taylor expansion into (46), (50) leads to:

(51)
$$\frac{\partial C}{\partial t}+u\frac{\partial C}{\partial x}=\left(\frac{-u^2\Delta t}{2}\right)\frac{\partial^2 C}{\partial x^2}+\left(\frac{u\Delta x^2}{6}\right)\frac{\partial^3 C}{\partial x^3}$$

which is unstable (seen from negative numerical diffusion term).

Another semi-implicit discrete algorithm, second-order in space and time, is the Crank-Nicolson scheme, which mixes-up the *t* and $t+\Delta t$ time moments:

(52)
$$\frac{\partial C}{\partial x}=\mu_c\frac{C_{i+1}^{k+1}-C_{i-1}^{k+1}}{2\Delta x}+(1-\mu_c)\frac{C_{i+1}^{k}-C_{i-1}^{k}}{2\Delta x}$$

The discrete equation then reads as:

(53)
$$\frac{C_i^{k+1}-C_i^k}{\Delta t}=-u\left[\mu_c\frac{C_{i+1}^{k+1}-C_{i-1}^{k+1}}{2\Delta x}+(1-\mu_c)\frac{C_{i+1}^{k}-C_{i-1}^{k}}{2\Delta x}\right]$$

The Crank-Nicolson parameter μ_c is usually taken as 0.5, thus mixing the present- and past-time concentration values into single time-step computations.

More schemes involve various interpretations of the first and second derivatives. An extensive review can be found in the course literature.

However, all direct-discretization schemes suffer severely from numerical viscosity or similar-type problems, eventually failing to reproduce even the simplest shapes (Figure 6).

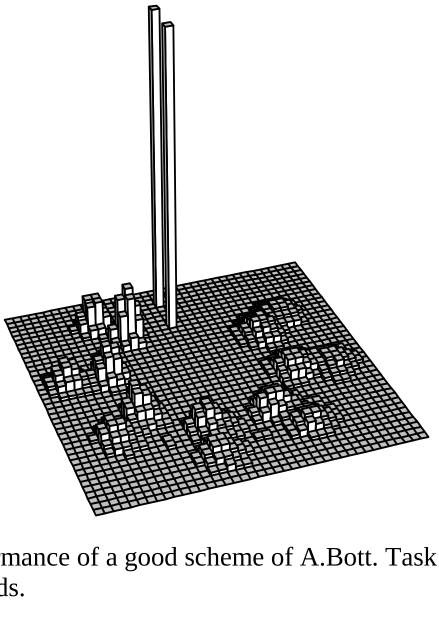


Figure 6. An example of performance of a good scheme of A.Bott. Task: rotation of two separated single-grid-cell non-zero concentration fields.

Physical meaning of the numerical viscosity is demonstrated by diagram in Figure 7.

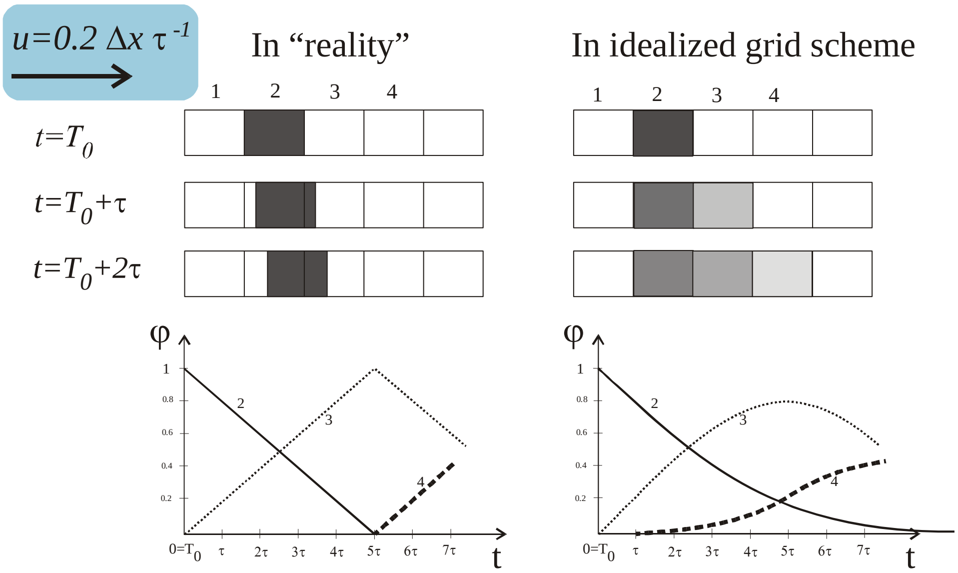


Figure 7. An illustration of the principle of numerical viscosity.

Options for further development of the discrete algorithms

- more and more sophisticated discretization schemes
- semi-Lagrnagian schemes: determine the transport distance and direction at each time step via single-time-step backward trajectory from the destination point that points at the location of the air mass coming to this point at the time *t+1*. Then the concentration is interpolated from the source place to the destination one – taking surrounding into account.
- finite-elements methods where a few grid cells are approximated with prescribed analytical function – usually some polynomials. A set of discrete equations is then derived to minimize the error of representation of actual pattern with these piece-wise approximations. The most-popular approach here was suggested by Galerkin, who required that the error of such approximation be orthogonal to the piecewise polynomial space itself.
- flux schemes are the variations of the finite-difference methods when the polynoms are used for evaluation of the fluxes at the borders of the grid cells.

The main features of the advection schemes

- global mass conservation
- local mass conservation
- positive definition of the scheme (appearance of negative concentration)
- numerical stability and robustness to errors in initial and boundary conditions
- phase error (error in the propagation speed of some characteristic elements of the pattern)
- non-monotonicity of the scheme

Specific additions suggested by M.Galperin in his works

- information-based approach and cause-result principle. In the advection term the information is spread with wind, so that a cell, which does not “know” about the coming or passed plume must not be disturbed by the scheme. In ideal case, such approach eliminates numerical viscosity.
- global consideration. The scheme should “see” the whole computation domain rather than a few adjacent cells. The Courant criterion for the scheme stability is then modified into $uD/\Delta t<1$, where D is the size of the domain.
- advancing the subgrid information utilization (essentially, already considered by the semi-Lagrangian and flux schemes) that keeps and preserves the pattern finer than the formal grid spacing.

Discrete adjoint equation vs discretized adjoint equation

Considering the ways of solving the adjoint dispersion equation with non-ideal advection-diffusion discrete algorithms, one has to keep in mind that numerical solution of the adjoint equation (???eq ref???) will not satisfy the adjointness requirement (????eq ref???). Hense, the duality of forward and adjoint problems will be lost. Since this duality is one of the strongest motivations to consider the adjoint formalism, one might wish to keep the exact adjointness of the numerical solution. Such choice forces selection of a specific numerical algorithm for the adjoint equation discretization, which depends on the forward algorithm.

Consider a general form of the forward equation:

(54)
$$\frac{\partial\varphi}{\partial t}+A\varphi=E,\quad \varphi(t=0)=\varphi_0$$

Here **A** is a linear operator in the appropriate Hilbert space with the corresponding scalar product.

Its discrete form it will read as:

$$\frac{\varphi^{k+1}-\varphi^{k}}{\Delta t}+\mathbf{A}^{h}\varphi^{k}=E,\quad \varphi^{k}(k=0)=\varphi_0\,,$$

Where φ^k is a column $(\varphi_1^k,...\varphi_n^k)^T$ defined in points $i=1\ldots n$, of the computation domain and in points $k=1\ldots K$ of the time axis. Matrix operator $\mathbf{A^h}$ is defined in the grid space $\Phi_1=\{\varphi^k\}$

Let also introduce the grid space $\Phi_2=\{\varphi=\{\varphi^k\},k=\overline{0,K}\}$ and define the scalar products in the following way:

$$(\varphi^k,\tilde{\varphi}^k)_1=(D\varphi^k,\tilde{\varphi}^k)=\sum_{i=1}^{n}d_i\varphi_i^k\tilde{\varphi}_i^k$$

$$(\varphi,\tilde{\varphi})_2=\sum_{k=0}^{K}(\varphi^k,\tilde{\varphi}^k)_1\Delta t$$

Here **D** is a diagonal positively defined matrix $D=diag\{d_i\}$. In the simplest case $d_i=\Delta x$, $i=1..n$.

With regard to Φ_2 space, the equation (54) can be written in the condensed form:

(55)

$$L^h \varphi = E$$

Multiplying (55) with $\tilde{\varphi}$, we get:

$$(L^h \varphi - E, \tilde{\varphi})_2 = \sum_{k=0}^{K} [(\varphi^k - \varphi^{k-1}, \tilde{\varphi}^k)_1 + \Delta t (A^h \varphi^{k-1}, \tilde{\varphi}^k)_1 - (E, \tilde{\varphi}^k)_1] = 0$$

Expanding the product term-by-term:

$$\begin{aligned}
&\sum_{k=0}^{K} [(\varphi^{k+1} - \varphi^k, \tilde{\varphi}^k)_1] = \sum_{k=0}^{K} [(\varphi^{k+1} - \varphi^k, \tilde{\varphi}^k + \tilde{\varphi}^{k+1} - \tilde{\varphi}^{k+1})_1] = \\
&= \sum_{k=0}^{K} [(\varphi^{k+1}, \tilde{\varphi}^{k+1})_1 - (\varphi^k, \tilde{\varphi}^k)_1 + (\varphi^{k+1}, \tilde{\varphi}^k - \tilde{\varphi}^{k+1})_1 + (\varphi^k, \tilde{\varphi}^{k+1} - \tilde{\varphi}^{k+1})_1] = \\
&= (\varphi^K, \tilde{\varphi}^K)_1 - (\varphi^0, \tilde{\varphi}^0)_1 + \sum_{k=0}^{K} [(\varphi^{k+1}, \tilde{\varphi}^k - \tilde{\varphi}^{k+1})_1]
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{k=0}^{K} [\Delta t (A^h \varphi^k, \tilde{\varphi}^k)_1] = \sum_{k=0}^{K} [\Delta t (D A^h \varphi^k, D^{-1} D \tilde{\varphi}^k)] = \sum_{k=0}^{K} [\Delta t (D^{-1} D A^h \varphi^k, D \tilde{\varphi}^k)] = \\
&= \sum_{k=0}^{K} [\Delta t (\varphi^k, (A^h)^T D \tilde{\varphi}^k)] = \sum_{k=0}^{K} [D^{-1} D \Delta t (\varphi^k, (A^h)^T D \tilde{\varphi}^k)] = \sum_{k=0}^{K} [D \Delta t (\varphi^k, D^{-1} (A^h)^T D \tilde{\varphi}^k)] = \\
&= \sum_{k=0}^{K} [\Delta t (\varphi^k, D^{-1} (A^h)^T D \tilde{\varphi}^k)_1]
\end{aligned}$$

We end up with the equation:

(56)

$$\sum_{k=0}^{K} [(\tilde{\varphi}^k - \tilde{\varphi}^{k+1}, \varphi^k)_1 + \Delta t (D^{-1} A^h D \tilde{\varphi}^k, \varphi^k)_1 - \Delta t (E, \tilde{\varphi}^k)_1] = 0$$

Let's require that $\tilde{\varphi}$ satisfies the equation:

(57)

$$\begin{aligned}
&\frac{\tilde{\varphi}^k - \tilde{\varphi}^{k+1}}{\Delta t} + D^{-1} A^h D \tilde{\varphi}^k = p, \\
&\tilde{\varphi}^K = \tilde{\varphi}^{K+1} = 0
\end{aligned}$$

We end up with the known duality statement, this time for discrete problem:

$$J = (f, \tilde{\varphi})_2 + (\varphi^0, \tilde{\varphi}^0)_2 = (\varphi, p)_2 = J$$

Remark on dynamic-adjoint formalism

In many cases, exact formulation of the discrete matrix A^h is cumbersome or simply non-existent. Instead, the advection equation solution can be represented via an algorithm with possible logic forks and joints, resolving some implicit problems, etc. In such cases the application of (57) formulations is impossible. Instead, one can note that the solution of adjoint equation for the given point in space and time is the sensitivity of the forward solution to variations of the emission intensity. Indeed, from the main duality equation, obtain:

$$J = (C, p) = (C^*, E)$$
$$\nabla_E J = (\nabla_E C, p) = (C^*, \nabla_E E) = (C^*, e)$$

Where e is a unit matrix. Here we used an independence of both sensitivity p and sensitivity distribution C^* from emission. If then one puts $p = \delta(x - x_1)\delta(t - t_1)$, the sensitivity of the concentration at this point and at the particular time will be the solution of the appropriate adjoint equation:

(58)
$$\frac{\partial C(t_1, x_1)}{\partial E} = (C^*, e), \qquad L^* C^* = \delta(x - x_1)\delta(t - t_1)$$

This is another fundamental feature of the adjoint equation: it describes the sensitivity of the forward equation results to uncertainties in the input data. With little efforts, the equations similar to (58) can be written for model internal parameters and other external data. We will see its practical application further in the data assimilation problem. Due to duality, the sensitivity of the adjoint solution is evidently the solution of the corresponding forward equation.

Construction of adjoints for the specific realizations of the numerical algorithms can also benefit from (58). Indeed, if somehow the sensitivity of the forward solution is extracted from the code, it will be the solution of the adjoint equation for the given time and space.

Elements of splitting algorithms

So far we considered the pure-advection equation. Extension to the other processes requires solving the following problem. If a finite-difference scheme is applied to the basic equation, the results will be very poor and extremely costly. Creation of a separate advection scheme, a diffusion scheme, a chemical scheme, etc requires a methodology of their combination and a proof that the resulting field will have some connection to the solution of the original equation. The numerical methods dealing with the problem comprise an own direction in numerical mathematics called splitting algorithms.

The idea of the time splitting is based on the separate consideration of each cause of the time derivative of the concentration at each time step:

(59)
$$\frac{\partial C}{\partial t} = \sum_{pr} A_{pr}(C)$$

Here a set of A_{pr} represents a set of processes contributing to the overall change of the concentration. However, solving the equation in such a form using global time-space finite-difference discretization would lead to prohibiting costs of the simulations. Therefore, at the next step of splitting suggests entirely separate consideration of each of the processes with subsequent summing-up their impacts:

(60)

$$\frac{\partial C}{\partial t} = \sum_{pr} \frac{\partial C_{i\_pr}}{\partial t}$$

$$\frac{\partial C_{i\_pr}}{\partial t} = A_{i\_pr}(C), \quad i\_pr = \overline{1, N_{pr}}$$

Note that this formulation allows flexible global time step for mixing-up the specific contributions, while individual time steps for solving the process-related equations can vary depending on the process nature and time scale.

In discrete case, the representation of time derivative via $\Delta C = C(t + \Delta t) - C(t)$ creates two more problems. Firstly, the individual operators can be applied in parallel:

$$\Delta C_{i\_pr} = A_{pr}\, C(t)\, \Delta t, \quad i\_pr = \overline{1, N_{pr}}$$

or sequentially:

$$\begin{aligned} &\Delta C_1 = A_1\, C(t)\, \Delta t', \quad C_1 = C(t) + \Delta C_1 \\ &\Delta C_{i\_pr} = A_{i\_pr} C_{i\_pr-1} \Delta t', \quad C_{i\_pr} = C_{i\_pr-1} + \Delta C_{i\_pr}, \quad i\_pr = \overline{2, N_{pr}} \end{aligned}$$

or in a mix. The results will evidently be different and both schemes have their positive and negative sides.

One of popular approaches is so-called symmetric operator splitting when the sequence of operators is applied twice with half-time-step and the second time it is applied in the opposite order:

$$\begin{aligned} &\Delta C_1 = A_1\, C(t)\ \Delta t'/2, \quad C_1 = C(t) + \Delta C_1 \\ &\Delta C_{i\_pr} = A_{i\_pr} C_{i\_pr-1} \Delta t', \quad C_{i\_pr} = C_{i\_pr-1} + \Delta C_{i\_pr}, \quad i\_pr = \overline{2, N_{pr}} \end{aligned}$$

after which

$$\begin{aligned} &\Delta C_{N\_pr} = A_{N\_pr}\, C_{N\_pr}\ \Delta t'/2, \quad C_{N\_pr-1} = C_{N\_pr} + \Delta C_{N\_pr} \\ &\Delta C_{i\_pr} = A_{i\_pr} C_{i\_pr+1}\, \Delta t'/2, \quad C_{i\_pr} = C_{i\_pr+1} + \Delta C_{i\_pr}, \quad i\_pr = \overline{N_{pr} - 1, 1, -1} \end{aligned}$$

This, however, also has own weaknesses.

Let's consider a general case of the operator splitting (59) - (60) for two operators $A_1 \geq 0$ and $A_2 \geq 0$. Let the general evolution-type equation is presented in a finite-difference form:

(61)

$$\frac{\partial \varphi}{\partial t} + (A_1 + A_2)\varphi = f$$

The A_1 and A_2 are matrices, both non-negatively defined, with coefficients dependent on time. Further we assume sufficient smoothness of all the coefficients and solutions. Let's consider their representation at a time interval $t_j \leq t \leq t_{j+1}$:

(62)

$$\Lambda_i^k = A_i(t_{k+1/2}), \quad \Lambda^k = \Lambda_1^k + \Lambda_2^k$$

Writing down the sequential Crank-Nicolson's algorithms for these two operators, we obtain:

(63)

$$\frac{\varphi^{k+1/2}-\varphi^{k}}{\Delta t}+\Lambda_1^k\frac{\varphi^{k+1/2}+\varphi^{k}}{2}=0,$$
$$\frac{\varphi^{k+1}-\varphi^{k+1/2}}{\Delta t}+\Lambda_2^k\frac{\varphi^{k+1}+\varphi^{k+1/2}}{2}=0$$

The unknown intermediate variable $\varphi^{k+1/2}$ can be excluded, which leads to the direct connection of discrete concentrations at *k-th* and *k+1-st* time steps. From the first equation:

$$\varphi^{k+1/2}=\left(\varepsilon+\frac{\Delta t}{2}\Lambda_1^k\right)^{-1}\left(\varepsilon-\frac{\Delta t}{2}\Lambda_1^k\right)\varphi^k$$

Substituting it to the second one, obtain:

(64)

$$\varphi^{k+1}=\left(\varepsilon+\frac{\Delta t}{2}\Lambda_2^k\right)^{-1}\left(\varepsilon-\frac{\Delta t}{2}\Lambda_2^k\right)\varphi^{k+1/2}=$$
$$=\left(\varepsilon+\frac{\Delta t}{2}\Lambda_2^k\right)^{-1}\left(\varepsilon-\frac{\Delta t}{2}\Lambda_2^k\right)\left(\varepsilon+\frac{\Delta t}{2}\Lambda_1^k\right)^{-1}\left(\varepsilon-\frac{\Delta t}{2}\Lambda_1^k\right)\varphi^k$$

In order to get the order of approximation, one has to open-up the brackets:

$$\varphi^{k+1}=\left(\varepsilon-\frac{\Delta t}{2}\Lambda_2^k+\left(\frac{\Delta t}{2}\Lambda_2^k\right)^2\right)\left(\varepsilon-\frac{\Delta t}{2}\Lambda_2^k\right)\left(\varepsilon-\frac{\Delta t}{2}\Lambda_1^k+\left(\frac{\Delta t}{2}\Lambda_1^k\right)^2\right)\left(\varepsilon-\frac{\Delta t}{2}\Lambda_1^k\right)\varphi^k=$$
$$=\left(\varepsilon+\frac{\Delta t}{2}\left(-2\Lambda_2^k-2\Lambda_1^k\right)+\left(\frac{\Delta t}{2}\right)^2\left(2\left(\Lambda_2^k\right)^2+2\left(\Lambda_1^k\right)^2+4\Lambda_2^k\Lambda_1^k\right)\right)\varphi^k=$$
$$=\left(\varepsilon-\Delta t\Lambda^k+\frac{\Delta t^2}{2}\left(\left(\Lambda_2^k\right)^2+\left(\Lambda_1^k\right)^2+2\Lambda_2^k\Lambda_1^k\right)\right)\varphi^k$$

If Λ_1 and Λ_2 commute, the approximation will be of the second order with regard to time. If not, the first order.

However, the requirement of the matrix commutation can be eliminated if the two-step scheme is considered for symmetrical $2\Delta t$ period $t_{j-1}\le t\le t_{j+1}$. The idea of the symmetrization is based on sequential application of the finite-difference algorithms (63) - (64).

$$\frac{\varphi^{k-1/2}-\varphi^{k-1}}{\Delta t}+\Lambda_1^k\frac{\varphi^{k-1/2}+\varphi^{k-1}}{2}=0,$$
$$\frac{\varphi^{k-1}-\varphi^{k-1/2}}{\Delta t}+\Lambda_2^k\frac{\varphi^{k}+\varphi^{k-1/2}}{2}=0,$$
$$\frac{\varphi^{k+1/2}-\varphi^{k}}{\Delta t}+\Lambda_2^k\frac{\varphi^{k+1/2}+\varphi^{k}}{2}=0,$$
$$\frac{\varphi^{k}-\varphi^{k+1/2}}{\Delta t}+\Lambda_1^k\frac{\varphi^{k+1}+\varphi^{k+1/2}}{2}=0$$

Again eliminating the intermediate variables, it is straightforward to obtain the operator connecting *k-1-st* and *k+1-st* time steps:

$$\varphi^{k+1} = \left(\varepsilon + \frac{\Delta t}{2}\Lambda_1^k\right)^{-1}\left(\varepsilon - \frac{\Delta t}{2}\Lambda_1^k\right)\left(\varepsilon + \frac{\Delta t}{2}\Lambda_2^k\right)^{-1}\left(\varepsilon - \frac{\Delta t}{2}\Lambda_2^k\right)\times$$
$$\times\left(\varepsilon + \frac{\Delta t}{2}\Lambda_2^k\right)^{-1}\left(\varepsilon - \frac{\Delta t}{2}\Lambda_2^k\right)\left(\varepsilon + \frac{\Delta t}{2}\Lambda_1^k\right)^{-1}\left(\varepsilon - \frac{\Delta t}{2}\Lambda_1^k\right)\varphi^{k-1}$$

Again opening-up the brackets, it is straightforward to see that:

$$\varphi^{k+1} = \left(\varepsilon - 2\Delta t\Lambda^k + 2\Delta t^2\left(\Lambda^k\right)^2 + \ldots\right)\varphi^{k-1}$$

regardless the commutation of the operators. Thus, if the algorithms (63) - (64) are applied sequentially in reverse order, over a double time step we obtain the second order of approximation with regard to time.

Without going into further details, one can extend the considered case to arbitrary number of operators. If they are applied sequentially and then the second time in reverse order, the resulting solution will have the second order of accuracy with regard to time.

It is also important to say that the above formulations pose only quite general requirements to the individual discrete operators. Therefore, it does not matter how exactly each of the processes is represented in the discrete model. As long as they are positively defined and sufficiently smooth, their sequential application will result in reasonable formulations for the time derivative for the overall concentration field.

Example of implementation of numerical algorithms

In this section we outline the solution of the diffusion equation along the vertical:

(65)
$$\frac{dC}{dt} = \frac{\partial}{\partial z}K_z\frac{\partial C}{\partial z}$$

With boundary conditions (37).

One of the most-significant problems appearing while considering the vertical dimension is a large variety of spatial scales involved. Thus, near the surface the typical length scale is height, while aloft this parameter does not play a significant role. Therefore, all atmospheric models have strongly inhomogeneous vertical coordinate systems, which have to be taken into account in the discrete algorithms. Let's consider the classical scheme with *n* layers with thicknesses h_i, central points z_i, and distances between the central points Δz_i. (Figure 8). Note that for such scheme, the thickness of layers and the distance between their centre points are connected via: $\Delta z_i = 0.5\,(h_{i+1} + h_i)$

Let the diffusion coefficient K_z is represented at the borders of the layers: $K_z(i), i=0..n$. Let also represent the vertical derivative in the center of the level *i* as follows:

$$\left.\frac{\partial\xi}{\partial z}\right|_{z_i} = \frac{\xi_{i+1/2} - \xi_{i-1/2}}{h_i}, \qquad \left.\frac{\partial\xi}{\partial z}\right|_{z_{i+1/2}} = \frac{\xi_{i+1} - \xi_i}{\Delta z_i}, \qquad \left.\frac{\partial\xi}{\partial z}\right|_{z_{i-1/2}} = \frac{\xi_i - \xi_{i-1}}{\Delta z_{i-1}}$$

Then discrete form of (65) will be written as follows:

(66)
$$\left.\frac{dC}{dt}\right|_i = \frac{1}{h_i}\left[K_{i+1/2}\frac{\varphi_{i+1}-\varphi_i}{\Delta z_i} - K_{i-1/2}\frac{\varphi_i-\varphi_{i-1}}{\Delta z_{i-1}}\right] = \\ = \varphi_{i-1}\frac{K_{i-1/2}}{h_i\Delta z_{i-1}} - \varphi_i\left(\frac{K_{i-1/2}}{h_i\Delta z_{i-1}} + \frac{K_{i+1/2}}{h_i\Delta z_i}\right) + \varphi_{i+1}\frac{K_{i+1/2}}{h_i\Delta z_i}$$

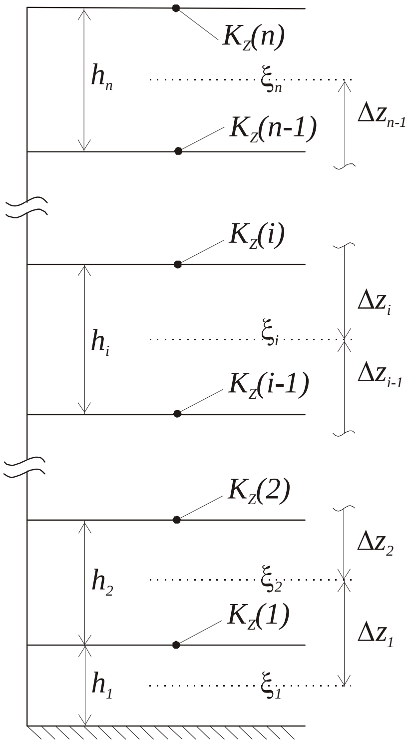


Figure 8. A scheme for vertical discretization of the advection-diffusion equation

In the widely used Crank-Nicolson algorithm, we assume that with regard to time $\varphi_i = (\varphi_i^k + \varphi_i^{k+1})/2$, where the *k* index is running along the time dimension.

In some models, the diffusion coefficient K_z can be directly available at half-levels *i+1/2, i-1/2*. However, in many cases it has to be obtained from the values at the centre of the layers. It should be stressed that simple interpolation of this parameter leads to unphysical results, such as violation of the local mass conservation. Correct averaging rule has to be derived keeping the conservation laws in mind.

Let's derive the correct representation of the vertical diffusion coefficient in a general discrete scheme. Let K_z be known as a continuous function of height *z*: $K_z(z)$. Let some finite-thickness layer $\Delta z = z_2 - z_1$ be free from local sources and sinks of a tracer. Let also the mean vertical wind is zero, so that its concentration is determined exclusively by diffusion. Then, following the K-closure formulations, the flux from z_1 to z_2 must be constant:

(67)
$$F(z) = const = K_z(z)\frac{\partial\xi}{\partial z}, \quad z \in [z_1, z_2]$$

By definition of the mean diffusion coefficient:

$$F_{mean}(z) = \overline{K_z}\,\frac{\xi(z_2) - \xi(z_1)}{\Delta z}$$

Noting that $F_{mean}=F(z)=const$ throughout the layer, one can integrate (67) from z_1 till z_2:

$$\frac{\partial \xi}{\partial z} = F_{mean}\,\frac{1}{K_z(z)},$$

$$\xi(z_2) - \xi(z_1) = F_{mean}\int_{z_1}^{z_2}\frac{1}{K_z(z)}\,dz$$

Finally obtaining:

(68)
$$\overline{K_z} = \frac{F_{mean}(z_2 - z_1)}{\xi(z_2) - \xi(z_1)} = \frac{z_2 - z_1}{\int_{z_1}^{z_2}\frac{1}{K_z(z)}\,dz}$$

From equation (68) one can notice that even a thin layer with small eddy diffusivity has a strong blocking effect – due to inverse K_z under the integral. Therefore, if the input data and the model formulations allow, it is always advisable to use the continuous definition of K_z since the discrete approximations always have a chance to miss such a layer.

If the K_z values are available only at z_i points, the averaging (68) turns into summation of reciprocals:

$$\overline{K_z}(z) = \frac{z_2 - z_1}{\dfrac{z - z_1}{K_z(z_1)} + \dfrac{z_2 - z}{K_z(z_2)}}$$

Here z is the point to which the mean coefficient is to be applied.

With $K_{i+/-1/2}$ defined, one can approach solving the diffusion equation (66). To illustrate the application, let’s consider the implicit first-order scheme with regard to time:

$$\frac{\varphi_i^{k+1} - \varphi_i^k}{\tau} = \varphi_{i-1}^{k+1}\,\frac{K_{i-1/2}^{k+1}}{h_i\Delta z_{i-1}} - \varphi_i^{k+1}\left(\frac{K_{i-1/2}^{k+1}}{h_i\Delta z_{i-1}} + \frac{K_{i+1/2}^{k+1}}{h_i\Delta z_i}\right) + \varphi_{i+1}^{k+1}\,\frac{K_{i+1/2}^{k+1}}{h_i\Delta z_i}$$

Here we assume that the values of K are known also for the future time, thus leaving only φ as the unknown. Time index at K is then omitted for the sake of compactness of the derivation.

Let’s define the matrix **A** as follows:

(69)

$$\mathbf{A}=\begin{bmatrix} -\frac{K_{1-1/2}}{h_1\Delta z_0}-\frac{K_{1+1/2}}{h_1\Delta z_1} & \frac{K_{1+1/2}}{h_1\Delta z_1} & 0 & \dots & 0 \\ \frac{K_{2-1/2}}{h_2\Delta z_1} & -\frac{K_{2-1/2}}{h_2\Delta z_1}-\frac{K_{2+1/2}}{h_2\Delta z_2} & \frac{K_{2+1/2}}{h_2\Delta z_2} & 0\dots & \dots 0 \\ 0\dots. & \frac{K_{i-1/2}}{h_i\Delta z_{i-1}} & -\frac{K_{i-1/2}}{h_i\Delta z_{i-1}}-\frac{K_{i+1/2}}{h_i\Delta z_i} & \frac{K_{i+1/2}}{h_i\Delta z_i} & \\ 0\dots. & & & & \dots.0 \\ 0\dots. & \dots & 0 & \frac{K_{n-1/2}}{h_n\Delta z_{n-1}} & -\frac{K_{n-1/2}}{h_n\Delta z_{n-1}}-\frac{K_{n+1/2}}{h_n\Delta z_n} \end{bmatrix}$$

It is easy to see that the matrix **A** is the three-diagonal matrix, which first and last rows determine the boundary conditions. Then (65) turns into its discrete form:

$$\frac{d\varphi}{dt}=\mathbf{A}\varphi$$

For implicit scheme it will read:

$$\frac{\varphi^{k+1}-\varphi^{k}}{\Delta t}=\mathbf{A}\varphi^{k+1}$$

And its solution will be:

$$\varphi^{k+1}=(\varepsilon-\Delta t\mathbf{A})^{-1}\varphi^{k}$$

Here ε is a unit matrix. Note that $(\varepsilon - \Delta t\mathbf{A})$ is also a three-diagonal matrix. For its inversion, an elegant **swift** method has been developed. In order to demonstrate it, let's rename the elements in (69), so that the sub-diagonal elements will be a_i, diagonal be b_i and above-diagonal be c_i:

$$a_i=-\frac{K_{i-1/2}}{h_i\Delta z_{i-1}},\quad b_i=1+\Delta t\left(\frac{K_{i-1/2}}{h_i\Delta z_{i-1}}+\frac{K_{i+1/2}}{h_i\Delta z_i}\right),\quad c_i=-\frac{K_{i+1/2}}{h_i\Delta z_i}$$

(70)

$$\varepsilon-\Delta t\,\mathbf{A}=\begin{bmatrix} b_1 & c_1 & 0 & \dots & 0 \\ a_2 & b_2 & c_2 & 0.. & ..0 \\ 0.. & a_i & b_i & c_i & ..0 \\ 0\dots & & \dots & & \dots 0 \\ 0\dots & & ..0 & a_n & b_n \end{bmatrix}$$

Then, let us introduce two new variables via recursive formula:

$$p_1=-\frac{c_1}{b_1},\quad q_1=\varphi_1^k$$

$$p_{i+1}=-\frac{c_{i+1}}{b_{i+1}+p_i a_{i+1}},\quad q_{i+1}=\frac{\varphi_{i+1}^k-q_i a_{i+1}}{b_{i+1}+p_i a_{i+1}},$$

$$q_n=\frac{\varphi_n^k-q_{n-1}a_n}{b_n+p_{n-1}a_n}$$

Both q and p can be computed directly following the recursion, which constitutes the forward part of the swift cycle. The return part gives the new concentrations:

$$\varphi_n^{k+1} = q_n$$

$$\varphi_i^{k+1} = q_i + p_i \varphi_{i+1}^{k+1}$$