

2 Particle tracking

David Brickman, Bjørn Ådlandsvik, Uffe H. Thygesen, Carolina Parada, Kenneth Rose, Albert J. Hermann, and Karen Edwards

Particle-tracking models form the backbone of three-dimensional models of fish early life. These models use predictions of current velocities and diffusivities from hydrodynamic models to calculate the movement of individual particles in space and time. The goal of this section is to provide a set of recommendations for particle tracking in estuary and ocean modelling. Because the motivation comes principally from its application to biophysical modelling, the case of biologically active particles is specifically considered. The first part of this section presents, in a concise form, the essential aspects of best practices for particle tracking. Extra material is contained in Annexes 1–5. The second part presents a number of cases designed to test the performance of a particle-tracking routine.

2.1 Best practices for particle tracking

What makes particle tracking in an oceanographic (biophysical modelling) context different from tracking in an atmospheric context? The simple answer is that, historically, development of particle-tracking theory and techniques in the atmosphere was concerned principally with the atmospheric boundary layer, with an emphasis on correctly describing the statistics of dispersion for time-scales shorter than the Lagrangian time-scale (T_L), the time-scale at which velocity fluctuations are correlated. Generally, the computations were done for short periods (minutes to hours) and in one or two dimensions (for which analytic models exist; see Wilson *et al.*, 1981; Legg and Raupach, 1982; Thomson, 1987). These Lagrangian stochastic models (LSMs), or “random flight models”, are mathematically complicated, but are valid at all time scales (except below the Kolmogorov microscale, where viscosity becomes relevant; Thomson, 1987; Rodean, 1996). In addition, a critical problem of buoyant particles, “the trajectory crossing problem”, has only approximate solutions for LSMs (Sawford and Guest, 1991; Olia, 2002).

For biophysical modelling in the aquatic realm, we tend to be interested in time-scales longer than T_L and in three-dimensional drift for periods as long as several months. Another crucial difference is that many biophysical particles (representing planktonic larvae) have directed swimming motions that must be incorporated into the particle-tracking algorithm. This necessitates the use of random displacement models (RDMs, also known as random walk models). These models are valid for time-scales $\gg T_L$ (T_L vertical = 3–10 min; T_L horizontal = 1–8 d (near surface; greater at depth)). That the time-scales of interest in the ocean are not always $\gg T_L$ (especially on the horizontal plane) means that the use of RDMs in oceanographic particle tracking can be considered a “best-we-can-do” approach.

2.1.1 Choice of model

For the reasons outlined above, an RDM is recommended for oceanographic application. If we assume that the turbulence at each point is isotropic in the horizontal (i.e. its local statistics are invariant to rotations around a vertical axis), then turbulence is characterized by the horizontal diffusivity $K_{11} = K_{22}$ and the vertical diffusivity K_{33} . The three-dimensional RDM then becomes (Rodean, 1996):

$$dx_i = \left[U_i(\bar{x}, t) + \frac{\partial K_{ii}(\bar{x}, t)}{\partial x_i} \right] dt + (2K_{ii}(\bar{x}, t) dt)^{1/2} Q_i, \quad (1)$$

where dx_i is the displacement in the i th direction ($i = 1, 2, 3 = x, y, z$), U_i is the velocity, \bar{x} denotes three-dimensional position, t is time, K_{ii} is the eddy diffusivity, dt is the time-step, and Q is a Gaussian random variable with zero mean and unit variance. The term for the spatial derivative of the diffusivity $\partial K_{ii}(\bar{x}, t)/\partial x_i$ is a drift correction term required to remove erroneous aggregations, or evacuations, of particles (see Hunter *et al.* (1993); Visser (1997) for other formulations of the RDM). This term is required in order to maintain a well-mixed condition (WMC), that is, the requirement that an initial uniform concentration of particles remains uniform for all time (Brickman and Smith, 2002). For most applications, the algorithm based on Equation (1) will use circulation model output to provide the velocity and diffusivity fields. These fields exist on discrete grids, which may be problematic (see below).

2.1.2 Time discretization

The RDM is a stochastic differential equation, which in practice is solved using a discretization technique. The two commonly used are the Euler and Runge–Kutta routines. The former is a simple, first-order forward discretization routine, which generally executes quickly but is subject to truncation errors and (possible) instabilities. The latter is a higher order routine that is numerically more accurate. In the absence of turbulence, a higher order differencing scheme is recommended.

In the presence of turbulence, the choice of discretization technique is less obvious, because the precision gained by a high-order routine could be lost as a result of the “noise” of the turbulence. To examine this possibility, experiments were performed comparing the Euler and the Runge–Kutta routines for two different analytic flow-fields plus a turbulent component (see Annex 1). Histograms were created of the difference between endpoint positions for the two routines for 5000 different particle releases. These histograms resembled zero-mean Gaussian distributions, indicating that the difference between the two routines was random, not systematic. This suggests that the error introduced by use of an Euler stepping routine, in the presence of turbulence, itself looks “turbulent” and may reduce concerns about the relative accuracy of this scheme. Although the Euler scheme may be adequate for some situations, the effect of different discretization techniques on biological predictions has not been investigated and should be assessed in the context of specific modelling objectives.

2.1.3 Choice of time-step

In an RDM, as in any numeric algorithm for discretizing a continuous-time phenomenon, the time-step should be smaller than time constants of the system. This leads to upper bounds on the time-step (Thomson, 1987; Wilson and Flesch, 1993). The exception to this general rule is the Lagrangian time-scale characterizing the decorrelation of turbulent velocity fluctuations. RDMs are accurate descriptions of turbulent dispersal only on time-scales larger than the Lagrangian time-scale, so there is no reason to force the time-step below the Lagrangian time-scale.

For pure stationary diffusion in one dimension with diffusivity $D(z)$ ($\text{m}^2 \text{s}^{-1}$), the time constants $D/(\partial D/\partial z)^2$ and $1/|\partial^2 D/\partial z^2|$ describe when the expected change in diffusivity is larger than the diffusivity itself and, therefore, provide upper bounds on the time-steps. The time-scale of vertical mixing will, in most applications, be significantly larger; for Couette flow (the flow between two planes moving relative to each other), the half-time of the slowest mode of vertical mixing is $H^2(\log 2)/(8 \max_z$

($D(z)$), where H is the water depth. This time-scale can be used as a rough measure of vertical mixing in other flows as well, or more accurate time-scales can be obtained analytically or numerically for the specific flow.

Additional time-scales may characterize horizontal motion or other (e.g. biophysical) processes. The chosen time-step must ensure that all processes are accurately resolved. For an example of the effects of different choices of time-step see Annex 2.

2.1.4 Number of particles

A single-particle trajectory in a turbulent flowfield can be considered one trial of a statistical ensemble of which we are interested in the ensemble-averaged behaviour. If too few particles are released in a particle-tracking experiment, it is possible that the trajectories are polluted by statistical outliers and do not satisfactorily represent the desired ensemble average. There is a risk that this can lead to erroneous conclusions. Although there is no generic answer to this problem, we recommend that at least some tests be done to check whether or not sufficient particles are being used; for example, an experiment to measure the concentration of particles in some downstream grid cell at a given time after release (where concentration = # particles in grid cell/total number released) and repeating this experiment for an increasing total number of particles. This concentration, as a function of the total number of particles, will stop fluctuating when a sufficient number of particles are being used. For more details on such techniques, see Brickman and Smith (2002). In general, the oceanographic literature contains numerous instances of poorly performed particle-tracking experiments. The basic premise of performing a particle-tracking experiment should be the ability to do it correctly. There is no excuse for using too few particles.

2.1.5 Choice of random number generator

The random number generator should perform well enough to ensure that the results are not artefacts of the particular algorithm. Some fairly common random number generators have been demonstrated to be flawed; these generators have typically been included in general-purpose development environments, as opposed to environments designed specifically for scientific computing. The typical problems with poor generators are short periods and correlation in the random numbers. Short periods mean that the sequence of random numbers repeats itself too soon. Correlation in the random numbers may result in incorrect dispersion: either too weak or too strong, depending on the correlation pattern. Both flaws seriously undermine the credibility of the study.

There is no reason to use a random number generator with insufficient performance. It may be easier to obtain and install a state-of-the-art generator than to determine the properties of the built-in generator. Currently, the "Mersenne Twister" seems to be the strongest algorithm; this is, for example, the default generator in R and is also available in Matlab. C source code, made by the original designers of the algorithm, is available at <http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/emt.html>. Source code in other languages and a list of libraries that include the algorithm can be found at <http://www.Wikipedia.org> under Mersenne Twister.

A general introduction to random number generators can be found in Ross (2001) and similar textbooks on stochastic simulation. The standard tool for verifying built-in random number generators is Marsaglia's Diehard battery of tests (see <http://www.stat.fsu.edu/pub/diehard/>).

2.1.6 Boundary conditions

The boundary conditions for an RDM are similar to those for an ocean circulation model, that is, a condition of no flux through the boundaries. For an RDM, this means that no particle should cross a boundary or, equivalently, that particle numbers should be conserved. This is important because the calculation of particle concentrations, or probability density functions (PDFs), can be incorrect if particles are lost from the domain. This boundary condition is enacted as a reflection scheme. However, the requirements of this scheme can be non-trivial, as certain properties have to be maintained upon reflection, especially the WMC. An incorrect reflection scheme can lead to spurious particle concentrations near boundaries. For an LSM, these requirements have been determined (Wilson and Flesch, 1993), but for the RDM, they are less clear. In practice, many of the theoretical requirements for a boundary reflection scheme are not met, but this does not seem to have any great effect on the result (Legg and Raupach, 1982).

Owing to the various uncertainties in the theory and practice of boundary reflection schemes, no simple best practice can be recommended, except to state that such a routine is required for a valid particle-tracking model. Experience indicates that these schemes can be complicated to code and should be tested carefully before proceeding.

2.1.7 Additional considerations

2.1.7.1 The use of discrete circulation model fields

Most particle-tracking models rely on space- and time-discretized fields from an ocean circulation model. A number of problems can arise because of this, including interpolation within grid cells near model boundaries and the use of discretized turbulence quantities.

- **Interpolation within grid cells near model boundaries.** Circulation models typically have no slip and no flux boundary conditions on velocity, so that flow runs parallel with closed boundaries. The determination of the velocity within such boundary cells can be complicated, especially where flows are “turning corners” following a coastline. This can result in particles erroneously crossing a boundary as a result of the combination of velocity and time-step, or drifting in an incorrect direction (see test case, Section 2.2.2 Flow around an obstacle). The addition of turbulence to this process is a further complication resulting in the expenditure of significant coding and execution time handling particle tracking near boundaries. The best practice recommendation in this case is to be aware of this problem and to check carefully that the algorithm is performing correctly.
- **The use of discretized turbulence quantities.** Circulation models can produce discontinuous turbulence fields, particularly in the vertical dimension. The particle-tracking model (Equation 1) requires values and derivatives of these quantities, which can lead to problems in the correct prediction of particle positions if these fields are sufficiently non-smooth (Brickman and Smith, 2002; Thygesen and Ådlandsvik, 2007). A solution can be to smooth these fields before use (Brickman and Smith, 2002; North *et al.*, 2006), but it is difficult to determine the degree to which this is necessary or successful in a complicated model setting. The best advice in this case is to be aware of this problem, proceed carefully, and check that the algorithm is performing correctly whenever possible.

2.1.7.2 Backwards particle tracking

In problems of egg/larval drift, we often have an estimate of the distribution of eggs or larvae, provided by survey data, but incomplete knowledge of the release area of the propagules. In other words, we often have more data at the endpoint than at the starting point. One benefit of the particle-tracking technique is the ability to reverse time and perform backward particle tracking in order to find the most likely origin for observed propagules. For example, we consider the case of truly planktonic particles in a flowfield u that is divergence-free and does not cross boundaries. In this case, it is reasonable to use the simple one-dimensional, time-reversed, Euler scheme:

$$x_{t-dt} = x_t - u(x_t)dt + \nabla K(x_t)dt + \sqrt{2K(x_t)dt} Q, \quad (2)$$

where Q has the same meaning as in Equation (1). Starting from the final position and time (x_t, t_t) when the simulation reaches the starting time t_0 , the density of larvae at any position x_0 will be proportional to the likelihood function of the initial condition x_0 , viewed as an unknown parameter. (For more details on this example, see Thygesen, in prep.). Other papers on biophysical backward particle tracking include Batchelder (2006) and Christensen *et al.* (2007). A paper to be recommended from the atmospheric literature is Flesch *et al.* (1995).

2.1.7.3 Coupling particle tracking with continuous fields from NPZ models

There are several issues to consider when coupling particle-tracking models to the continuous fields generated by nutrient-phytoplankton-zooplankton (NPZ) models. The continuous fields are the spatially explicit, physics-related outputs (e.g. velocities used for advection-diffusion movement of the particles) and biologically related outputs (e.g. zooplankton densities as prey for the particles) generated by the NPZ model. Some of these issues relate to the quality of these continuous fields, whereas other issues relate to the mechanics of how the particles are coupled to the fields.

The first issue is the quality of the outputted fields from the NPZ, including the overall stability of the NPZ model, the realism of the NPZ-related parameter values, the formulation of the predation-closure terms used to impose mortality on the zooplankton, and the information on model performance provided by data assimilation and validation efforts (see Annex 3).

The second issue also influences the quality of the fields and involves the way in which the NPZ submodel is coupled to the physics model. Issues such as whether the NPZ is run online or offline with the physics, and the compatibility of the spatial and temporal resolutions between the NPZ and physics models, affect the realism and quality of the outputted NPZ fields (see Annex 4).

The third issue relates to how the particles are coupled to the NPZ fields (see Annex 5), for example, whether or not a sufficient number of particles (e.g. larval fish) are followed in order to properly represent their interactions with prey patchiness, the fact that one-way coupling prevents trophic feedback from the particles to their prey and from prey exhibiting avoidance behaviours or other responses, and the degree to which movement of particles (e.g. larval fish) is purely physics-driven or involves active behaviour (e.g. vertical migration, swimming). Addressing the patchiness, trophic feedback, and prey-response issues requires the NPZ and particle-tracking models to be solved simultaneously using a large number of particles. How to meld advective and behaviour modes of movement remains an open question. Both the active behaviour of the particles and the reactions of the prey can change the trajectories of the particles (individuals in the model) and the predicted densities of the prey.

2.2 Test cases

In this section, we present a number of test cases designed to test the performance of a particle-tracking routine and illustrate problems that can arise when interpolating near boundaries.

2.2.1 Vertical distribution of buoyant particles

2.2.1.1 Purpose

The purpose is to test how well the particle-tracking code handles buoyant particles, especially in relationship to the surface and bottom boundary conditions.

2.2.1.2 Background

The need to handle non-neutral particles arises in many applications, including phytoplankton, sediments, or, in this test case, fish eggs. The stationary case was treated by Sundby (1983). The general problem is easily handled in the Eulerian (concentration-based) setting. A Matlab toolbox was developed by Ådlandsvik (2000). This point of view has been adopted for the sampling of anchovy and sardine eggs using the Continuous Underwater Fish Egg Sampler (CUFES; Boyra *et al.*, 2003). For particle tracking, the binned random walk part of this test case was given by Thygesen and Ådlandsvik (2007).

2.2.1.3 Analytical solution

This test case considers a one-dimensional water column with non-neutral particles with a buoyant velocity w and eddy diffusivity K . The vertical coordinate z points upwards, with $z=0$ at bottom and $z=H$ at the surface. The concentration ϕ of particles is governed by the Eulerian conservation law,

$$\frac{\partial \phi}{\partial t} + \frac{\partial}{\partial z}(w\phi) = \frac{\partial}{\partial z} \left(K \frac{\partial \phi}{\partial z} \right). \quad (3)$$

The boundary conditions are zero flux through the surface:

$$w\phi = K \frac{\partial \phi}{\partial z}, \quad z = 0, H. \quad (4)$$

The solution evolves towards a stationary solution where the flux is zero in the whole water column. With constant coefficients, this ordinary differential equation gives a truncated exponential distribution. With $m = w/K$ and a vertical integrated concentration Φ , this can be written

$$\phi = \Phi \frac{m}{e^{mH} - 1} e^{mz}. \quad (5)$$

This has mean height above bottom

$$\mu = H - \frac{1}{m} + \frac{H}{e^{mH} - 1}, \quad (6)$$

and variance

$$\sigma^2 = \frac{2e^{mH} - m^2 H^2 - 2mH - 2}{m^2 (e^{mH} - 1)} - \mu^2. \quad (7)$$

Further details are given in Sundby (1983) and Ådlandsvik (2000).

2.2.1.4 Specification

The specific values used for this test case are given in Table 2.2.1. These values give a stationary mean depth (from surface) of 9.25 m and a standard deviation of 8.34 m. The particles are released 12.5 m above bottom, and the simulation time is 48 h.

Table 2.2.1. Variable settings for the buoyant test case.

VARIABLE	VALUE	UNIT
H	40	m
w	0.001	m s ⁻¹
K	0.01	m s ⁻²

2.2.1.5 Continuous random walk model

The continuous random walk model (i.e. RDM) for this problem with constant coefficients is implemented in a Euler–Forward fashion by,

$$Z^{n+1} = Z^n + w\Delta t + \sqrt{2K\Delta t} Q, \tag{8}$$

where Z is displacement and Q is a random variable with zero mean and unit variance. The boundary conditions are more difficult; the usual reflective boundary scheme at the surface,

$$Z^{n+1} \leftarrow 2H - Z^{n+1} \text{ 1, if } Z^{n+1} > H, \tag{9}$$

corresponds to

$$\frac{\partial \phi}{\partial z} = 0, \tag{10}$$

which differs from the correct boundary condition in Equation (4). In fact, the analytical stationary solution has the maximum of the derivative at the surface.

The number of particles in this test case is 40 000. Two different time-steps, 5 and 30 min, are considered, and a Gaussian distribution is used for the random walk. The 5 min case has also been run with a uniform (top-hat) distribution for the random component. The reflective boundary condition is applied. For the plot, the particles have been counted in 1 m bins.

The result demonstrates that the RDM solutions are good (Figure 2.2.1) except when they are close to the surface, where they underestimate the concentration. The height of the boundary zone depends on when the particle movement is influenced by the boundary, that is, the length scales $w\Delta t$ and $\sqrt{2K\Delta t}$. In this case, the shape of the random walk distribution influences the result, where the Gaussian shape is superior to the top-hat. This is probably caused by the top-hat distribution giving higher probabilities further from the mean, making the random walk “feel” the boundary at longer distance.

2.2.1.6 Binned random walk

The binned random walk does not have boundary problems because it is constructed by finite volume methods for the advection-diffusion equation (see Thygesen and Ådlandsvik, 2007). The water column was discretized into eight uneven bins, with

depths of 10, 5, 5, 5, 5, 3, and 2 m, counted from the bottom. The time-step used was 5 min, and both the first-order upstream and a second-order scheme were considered. The results are given in Figure 2.2.2. This figure also shows the analytical solution, averaged into the same bins. The upstream solution shows too much mixing: underestimating the concentration near the surface and overestimating it near the bottom. The second-order method follows the analytical solution well but overshoots near the surface.

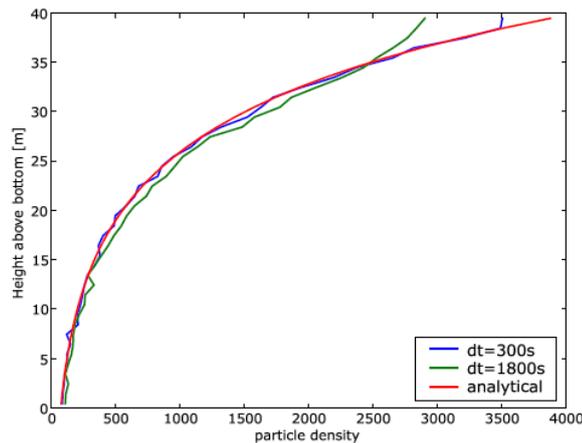


Figure 2.2.1. Results for the continuous random walk model.

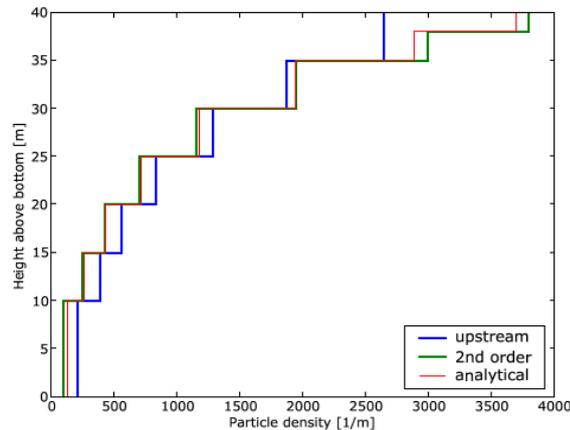


Figure 2.2.2. Results for the binned random walk model.

2.2.2 Flow around an obstacle

2.2.2.1 Purpose

The purpose is to test how different horizontal advection implementations handle a curved flowfield and a land obstacle.

2.2.2.2 Background

Non-rotational flow around a cylinder is one of the classical examples considered in almost all hydrodynamics textbooks. Of particular interest is the book by Bennett (2006), which takes a Lagrangian point of view.

2.2.2.3 Analytical considerations

The example is considered in a coastal oceanographic setting; the cylinder becomes a circular island. As the example is symmetric, only the upper half is considered. That

is, we consider a straight coast at $y=0$ with ocean in the upper half plane ($y>0$) and with a half-circular peninsula with centre $(x_0, 0)$ and radius R .

The steady non-rotational flow is given by a stream function

$$\psi = \frac{u_0 R^2 y}{(x-x_0)^2 + y^2} - u_0 y, \tag{11}$$

where u_0 is the along-coast velocity far from the obstacle. The stream function is normalized so that the land boundary is given by the contour $\psi=0$. The flow follows the streamlines, that is, isolines of ψ with higher values to the right, more precisely

$$u = -\frac{\partial \psi}{\partial y} = u_0 - u_0 R^2 \frac{(x-x_0)^2 - y^2}{((x-x_0)^2 + y^2)^2} \tag{12}$$

and

$$v = \frac{\partial \psi}{\partial x} = -2u_0 R^2 \frac{(x-x_0)y}{((x-x_0)^2 + y^2)^2}. \tag{13}$$

According to Bennett (2006), it is unlikely that analytical expressions will be found for the time-dependent particle movement in this example. Bennet does, however, provide an analytical description of stream lines. The “exact” solution shown below is obtained by using converged Runge–Kutta with a small time-step (36 s), using the analytical expression above for the velocities without interpolation. The dashed stream lines are simply obtained by contouring the discretized version of the stream function.

2.2.2.4 Specification

A domain of length L along the coast and width W is considered. The peninsula centre is at $x=0.5 L$ and the radius $R=0.32 W$. The numerical values are specified in Table 2.2.2.

Table 2.2.2. Variable settings for the peninsula test case.

Variable	Value	Unit
L	100	km
W	50	km
u_0	1	m s ⁻¹

The domain is discretized by $\Delta x = \Delta y = 1$ km. The grid coordinates are chosen so that grid cell (i, j) has its centre at $(x, y) = (i\Delta x, j\Delta x)$ for $i=0, \dots, 99$ and $j=0, \dots, 49$. The velocities are sampled in an A-grid, that is, in the grid centres. Denoting the velocity arrays U and V , we have

$$U(i, j) = u(i\Delta x, j\Delta x), \quad V(i, j) = v(i\Delta x, j\Delta x), \tag{14}$$

where u and v are given by the analytical formulas above. The velocities are set to zero at land, that is, where $\psi \leq 0$, in particular $U(i, 0) = V(i, 0) = 0$. The initial particle distribution is 1000 particles on a line perpendicular to the coast:

$$X_k = 3, \quad Y_k = 0.45 + 0.045k \quad \text{for} \quad k = 1, \dots, 1000. \tag{15}$$

The simulation time is 24 h, for which the particles would be transported 86.4 km with the reference velocity u_0 .

2.2.2.5 Simulations

The first-order Euler forward and the Runge–Kutta fourth-order method are considered. Both methods are used here with bilinear interpolation to interpolate from the grid-cell centres to the particle positions. The treatment of boundaries is simple, with the zero land velocity interpolated to the particle position and no reflection scheme implemented. This procedure may leave particles on land, but in the absence of turbulence, this was not considered to be important. A time-step of 1 h was used for both methods. The results from this test are presented in Figure 2.2.3. Far from the peninsula, both methods recapture the exact solution (green, red, and black symbols overlap). Close to the peninsula, the Euler method fails, leaving a trail of particles clearly separated from the peninsula. The Runge–Kutta method performs better, leaving a tiny tail of particles very close to the peninsula that do not overlap those produced by the exact solution.

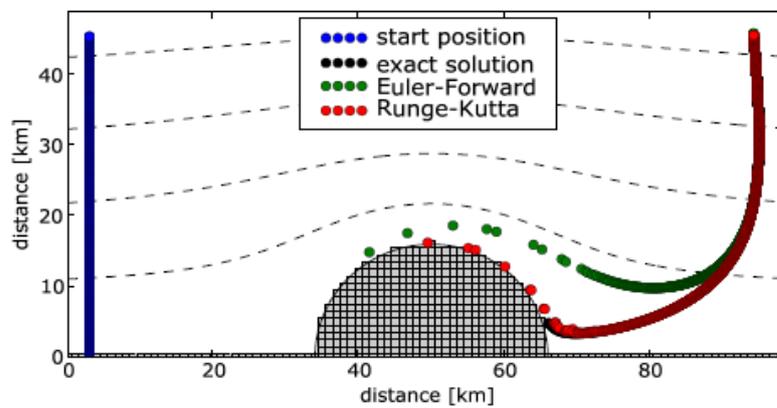


Figure 2.2.3. Peninsula test case.

The velocities from the formulas above are also defined for $\psi > 0$, giving a circulation within the “peninsula”. Using these velocities in the interpolation and intermediate Runge–Kutta steps gives a reference solution with ideal land treatment. This land treatment makes the Runge–Kutta indistinguishable from the exact solution and also improves the results from the Euler method. These results are shown in Figure 2.2.4 in which symbols for the Runge–Kutta method overlap those of the exact solution.

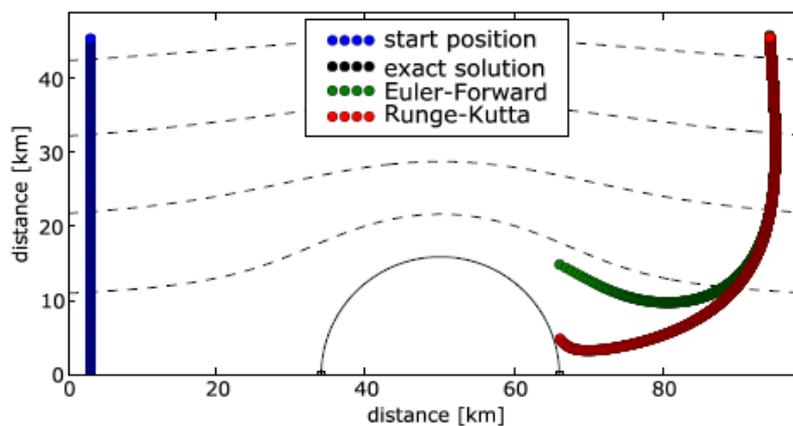


Figure 2.2.4. Peninsula test case with circulation within the “peninsula”.

2.2.2.6 Comment

This test was designed to demonstrate the difference between the Euler method and higher order methods, such as Runge–Kutta, and to point out problems associated with interpolation near boundaries. No random walk diffusion has been applied, which could reduce the advantage of higher order methods (see Annex 1). Also, shorter time-steps improve the performance of both models and may decrease the difference between them.