Supporting Information for "A dimensionless parameter for predicting convective self-aggregation onset in a stochastic reaction-diffusion model of tropical radiative-convective equilibrium"

Giovanni Biagioli[1,2] and Adrian Mark Tompkins[2]

[1]University of Trieste, Trieste, Italy

[2]Abdus Salam International Center for Theoretical Physics (ICTP), Trieste, Italy

Contents of this file

1. Figures S1 to S7

Introduction This file contains an overview of the numerical procedure designed to solve the diffusion-reaction prognostic equation for R, eqn. (3) in the paper. A full description of idealized tests is also included and the corresponding results are shown in Figs. S1-S3.

An example of time evolution of diagnostics for the two metrics $d_{\mathrm{max,clr}}$ and $d_{\mathrm{max,nn}}$ introduced in the main manuscript, calculated directly from the numerical routine, is then reported in Fig. S4 for the same set of simulations presented in Fig. 3. In particular, the theoretical model (12), which is used in the definition of the dimensionless parameter (15), is seen to exactly match the maximum inter-convective nearest neighbour distance to be expected when randomly throwing points onto a bounded region of the plane (Fig. S5).

Additional figures (Figs. S6-S7) are finally to be compared with Fig. 10 and motivate the reliability of the quantity γ, defined in the main manuscript (eqn. 14), in capturing the transition between aggregated and non-aggregated states in the (τ_{sub}, K) space independent of the domain size (Fig. S6) or the horizontal grid spacing (Fig. S7). In these cases the a_d-dependency is not considered, as a_d is kept at its default value, $a_d = 14.72$, but apparently the critical isopleth of $K\tau_{\mathrm{sub}}(L\bar{d})^{-1}$ that fits the transition regime is the same across these ensembles of simulations.

1. Numerics

The numerical solution of the governing equation (3) uses second-order finite differences in space and a Strang-type operator splitting scheme in time (Strang, 1968).

Splitting techniques are commonly advocated in meteorological applications with multiple time scales (e.g., Beljaars et al., 2018) and, in general, when a differential problem involves many physical processes. Commonly, it is not possibly or numerically efficient to attempt the integration of the equations by means of a single solver method. A computationally sustainable alternative is thus offered by decomposing the system into sub-groups of processes and using different suitable and advantageous methods for each group (e.g., Hundsdorfer & Verwer, 2007), with tendencies from each treated sequentially in time. In detail, given a generic, scalar partial differential equation

$$\frac{\partial u}{\partial t} = f(u),$$

where f can be regarded as a spatial partial differential operator, a two-term decomposition of the RHS is considered:

$$f(u) = f_1(u) + f_2(u).$$

The Strang splitting approach is to perform half a time step with the operator f_2, followed by a full time step with f_1 and another half step with f_2 (or vice versa), with the tendency from the previous process added to provide the initial value of the subsequent process(es). In formulae,

$$\begin{aligned}
&\frac{\partial}{\partial t}u^* = f_2(u^*), \quad t_n < t \le t_{n+\frac{1}{2}}, \quad u^*(t_n) = u_n, \\
&\frac{\partial}{\partial t}u^{**} = f_1(u^{**}), \quad t_n < t \le t_{n+1}, \quad u^{**}(t_n) = u^*\left(t_{n+\frac{1}{2}}\right), \\
&\frac{\partial}{\partial t}u^{***} = f_2(u^{***}), \; t_{n+\frac{1}{2}} < t \le t_{n+1}, \; u^{***}\left(t_{n+\frac{1}{2}}\right) = u^{**}(t_{n+1}),
\end{aligned}$$

the subscript n refers to the temporal discretization, and the overall solution is given by $u_{n+1} = u^{***}(t_{n+1})$. This sequential procedure normally introduces an error at each integration step. The Strang splitting is second-order accurate for sufficiently smooth solutions (e.g., LeVeque, 2007), provided that each subproblem is treated with a method of such accuracy at least.

In the stochastic model presented here, we separate the diffusion and subsidence components of eqn. (3) from the convective source term, i.e.,

$$f_1(R) = K\left(\delta_x^2 + \delta_y^2\right)R - \frac{R}{\tau_{\text{sub}}}, \;\; f_2(R) = \frac{(R_c - R)}{\tau_c}\mathcal{H}(p_c(R) - X),$$

as this will not produce any splitting error in the non-convective grid points, where the second operator vanishes. The difference operator $\delta_\cdot^2$ is the second-order centered difference approximation of the second derivative. For the problem involving f_2, the analytical solution is derived, in order to reduce as much as possible any integration error, related to the application of numerical methods, which can be incurred in addition to the splitting error.

The diffusion-reaction semidiscrete problem

$$\frac{\partial R}{\partial t} = K\left(\delta_x^2 + \delta_y^2\right) R - \frac{R}{\tau_{\text{sub}}} \tag{1}$$

is solved by means of a properly modified version of the classical Peaceman-Rachford Alternating Direction Implicit (ADI) method (Peaceman & Rachford, 1955). The scheme consists of splitting 2D problems into two separate steps, treating implicitly only one spatial operator at a time and therefore performing line-by-line solution of smaller, generally structured, independent sets of equations. It can be regarded as a perturbed formulation of the Crank-Nicholson scheme, whose application to eqn. (1) would yield

$$\left(1 - \beta\delta_x^2 - \beta\delta_y^2 + \omega\right) R_{j,k}^{n+1} = \left(1 + \beta\delta_x^2 + \beta\delta_y^2 - \omega\right) R_{j,k}^{n}, \tag{2}$$

where the superscript n indicates discrete time steps, the subscripts j, k refer to the horizontal square grid, $\beta = K\dfrac{\Delta t}{2\Delta x^2}$, $\omega = \dfrac{\Delta t}{2\tau_{\text{sub}}}$, Δt denoting the time step and Δx the horizontal spacing. Eqn. (2) can be factorized and rearranged as

$$\left(1 - \beta\delta_x^2 + \omega\right)\left(1 - \beta\delta_y^2\right) R_{j,k}^{n+1} = \left(1 + \beta\delta_x^2 - \omega\right)\left(1 + \beta\delta_y^2\right) R_{j,k}^{n} + \left(\beta^2\delta_x^2\delta_y^2 - \omega\beta\delta_y^2\right)\left(R_{j,k}^{n+1} - R_{j,k}^{n}\right),$$

where the last term on the RHS can be proved to be $\mathcal{O}(\Delta t^3)$ and is therefore negligible for small Δt. Thus, the application of the ADI method implies the sequential solution of the systems

$$\left(1 - \beta\delta_x^2 + \omega\right) R_{j,k}^{n+\frac{1}{2}} = \left(1 + \beta\delta_y^2\right) R_{j,k}^{n}, \tag{3}$$

$$\left(1 - \beta\delta_y^2\right) R_{j,k}^{n+1} = \left(1 + \beta\delta_x^2 - \omega\right) R_{j,k}^{n+\frac{1}{2}}, \tag{4}$$

and the two-step scheme (3)-(4) can be shown to be unconditionally stable and second order in both space and time, hence convergent.

If periodic boundary conditions are assigned, the linear systems resulting from the discretizations turn out to be circulant tridiagonal and can be easily solved by using Fast

Fourier Transform algorithms (e.g., Cooley & Tukey, 1965). Note that, if the splitting procedure had not been adopted, the matrices would not have been circulant or constant with time owing to the triggering of convective events, whose corresponding terms contribute to the entries of the main diagonal.

The convergence properties of the new ADI solver are assessed against the assumption of initial top hat or Gaussian profiles. The results of convergence tests under time step and grid refinements are shown in Figs. S1 and S2, respectively. In the first case (Fig. S1), the problem in eqn. (1) is considered, with $K = 10^4$ m^{2}s^{-1}, $\tau_{\rm sub} = 10$ days, over a domain with size $L = 300$ km and spacing $\Delta x = 2$ km, and initial condition given by

$$R_0(\mathbf{x}) = R(\mathbf{x}, 0) = \begin{cases} 1 & \text{for } \mathbf{x} \in [x_1, x_2] \times [y_1, y_2] \\ 0.8 & \text{elsewhere} \end{cases}, \tag{5}$$

where $x_1, y_1 = 140$ km and $x_2, y_2 = 160$ km. The top-hat configuration is a good test bench due to presence of very sharp discontinuities. It is well known that, if implicit discretization is adopted, there is no stability constraint on the time step, as it is for explicit numerical solver, e.g., explicit Euler, for which the following relationship (diffusive stability criterion) is to be satisfied to ensure stability:

$$\Delta t \leq \frac{\Delta x^2}{4K}. \tag{6}$$

Nevertheless, for implicit schemes, a limit of the type (6) still serves as a measure of accuracy (e.g., Ferziger et al., 2002), hence, labelling $\xi = 4K\dfrac{\Delta t}{\Delta x^2}$, we impose several time steps corresponding to a range of values of ξ. Convergence is apparent in the time slices of Fig. S1 (dashed and dash-dotted lines), even though the approximation for $\xi = 6$ exhibits a spurious oscillation at beginning, which is then damped and rapidly disappears.

Spatial convergence properties are examined, as also demanded by a resolution sensitivity study mentioned in the paper. The results of a grid refinement analysis performed on the problem eqn. (1) with $K = 5 \times 10^3$ m^{2}s^{-1}, $\tau_{\text{sub}} = 10$ days and Gaussian initial distribution with $\mu = 150$ km and $\sigma = 5$ km are shown in Fig. S2. The horizontal spacing is successively halved ranging from $\Delta x = 2$ km to 250 m, and the time step Δt is such that $\xi = 0.25$ in all cases. At $t = 600$ s (dashed lines), the profiles are almost insensitive to the resolution, then any error is eventually smoothed down and, at $t = 3600$ s (dash-dotted), the curves are nearly indistinguishable.

In case profiles with sharper discontinuities are prescribed, the method is still able to provide reasonably good approximations, despite exhibiting a more pronounced sensitivity to both the time step size and the resolution.

In the full system, the action of convection is to continuously introduce local delta function perturbations into the R distribution, with sharp gradient at between the convective point and the surrounding grid cells. This is particularly challenging for the numerics and could possibly amplify numerical errors. Further sets of tests are thus conducted, aimed at quantifying the impact of the errors associated with the operator splitting. Fig. S3 charts the results obtained for different values of K, $K = 10^4$ m^{2}s^{-1} (solid lines) and $K = 2.5 \times 10^4$ m^{2}s^{-1} (dashed lines), and ξ (colors), all else being kept fixed ($\tau_{\text{sub}} = 12$ days, $a_d = 14.72$, $L = 300$ km, $\Delta x = 2$ km, $R_0 = 0.8$). Convergence to the same statistically steady solutions is apparent, even though, in the low-diffusion case, the approximation for $\xi = 3$ ($\Delta t = 300$ s) yields some differences in both the R spatial mean and standard deviation final equilibrium state. Interestingly, both the transition to the self-aggregated state and the following evolution do not vary monotonically with Δt, and we attribute this

effect to the large stochastic component present in the modelled system. The time-step dependency almost entirely disappears for higher values of K, as they require the use of a smaller time step (via eqn. 6), which also leads to less severe splitting errors. For the experiments presented in the paper, in general, the time step is chosen so that $\xi < 1$.

References

Beljaars, A., Balsamo, G., Bechtold, P., Bozzo, A., Forbes, R., Hogan, R. J., ... Wedi, N. (2018). The numerics of physical parametrization in the ECMWF model. *frontiers*, *6*, doi:10.3389/feart.2018.00137.

Cooley, J. W., & Tukey, J. W. (1965). An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation*, *19*(90), 297–301. doi: 10.2307/2003354

Ferziger, J. H., Peric, M., & Street, R. L. (2002). *Computational methods for fluid dynamics*. Berlin Springer.

Hundsdorfer, W., & Verwer, J. G. (2007). *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*. Springer Science & Business Media.

LeVeque, R. J. (2007). *Finite Difference Methods for Ordinary and Partial Differential Equations.* https://epubs.siam.org/doi/book/10.1137/1.9780898717839.

Peaceman, D. W., & Rachford, J., H. H. (1955). The Numerical Solution of Parabolic and Elliptic Differential Equations. *Journal of the Society for Industrial and Applied Mathematics*, *3*(1), 28–41. doi: 10.1137/0103003

Strang, G. (1968, September). On the Construction and Comparison of Difference Schemes. *SIAM Journal on Numerical Analysis*, *5*(3), 506–517. doi: 10.1137/0705041

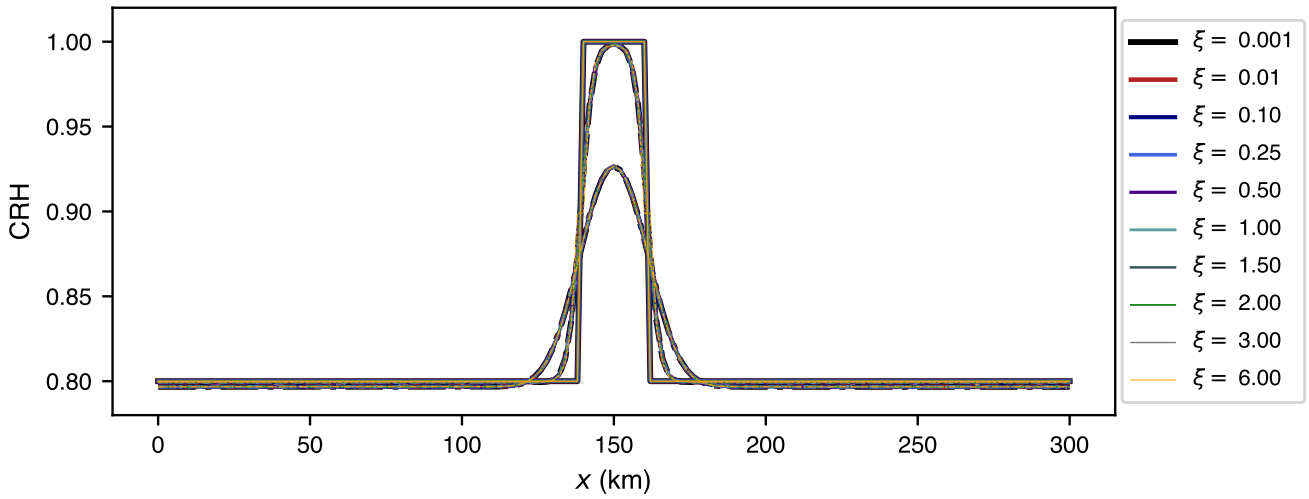

Figure S1. Computed solutions along the section $y = 150$ km for the problem (1) with $K = 10^4$ m^{2}s^{-1}, $\tau_{\text{sub}} = 10$ days, $R_0(\mathbf{x})$ as specified in (5). Shown are the initial profile (solid lines) and the numerical approximations for different time step choices at $t = 600$ s (dashed) and 3600 s (dash-dotted). $\xi = 1$ corresponds to $\Delta t = 100$ s.

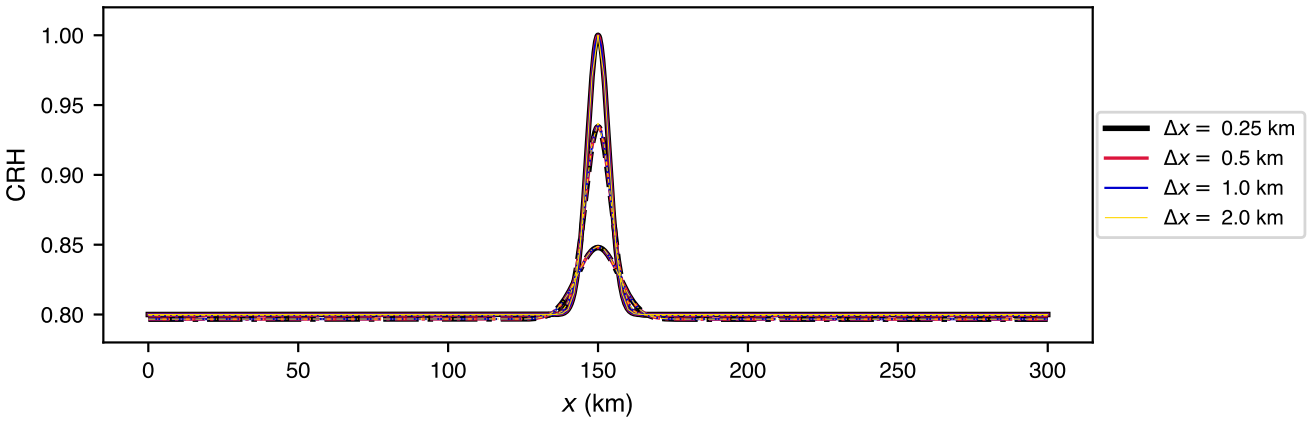

Figure S2. Results of a grid refinement study conducted on the problem (1) with $K = 5 \times 10^3$ m^{2}s^{-1}, $\tau_{\text{sub}} = 10$ days and a Gaussian initialization. Shown are the solutions along the section $y = 150$ km, at times $t = 0$ (solid lines) $t = 600$ s (dashed) and 3600 s (dash-dotted).

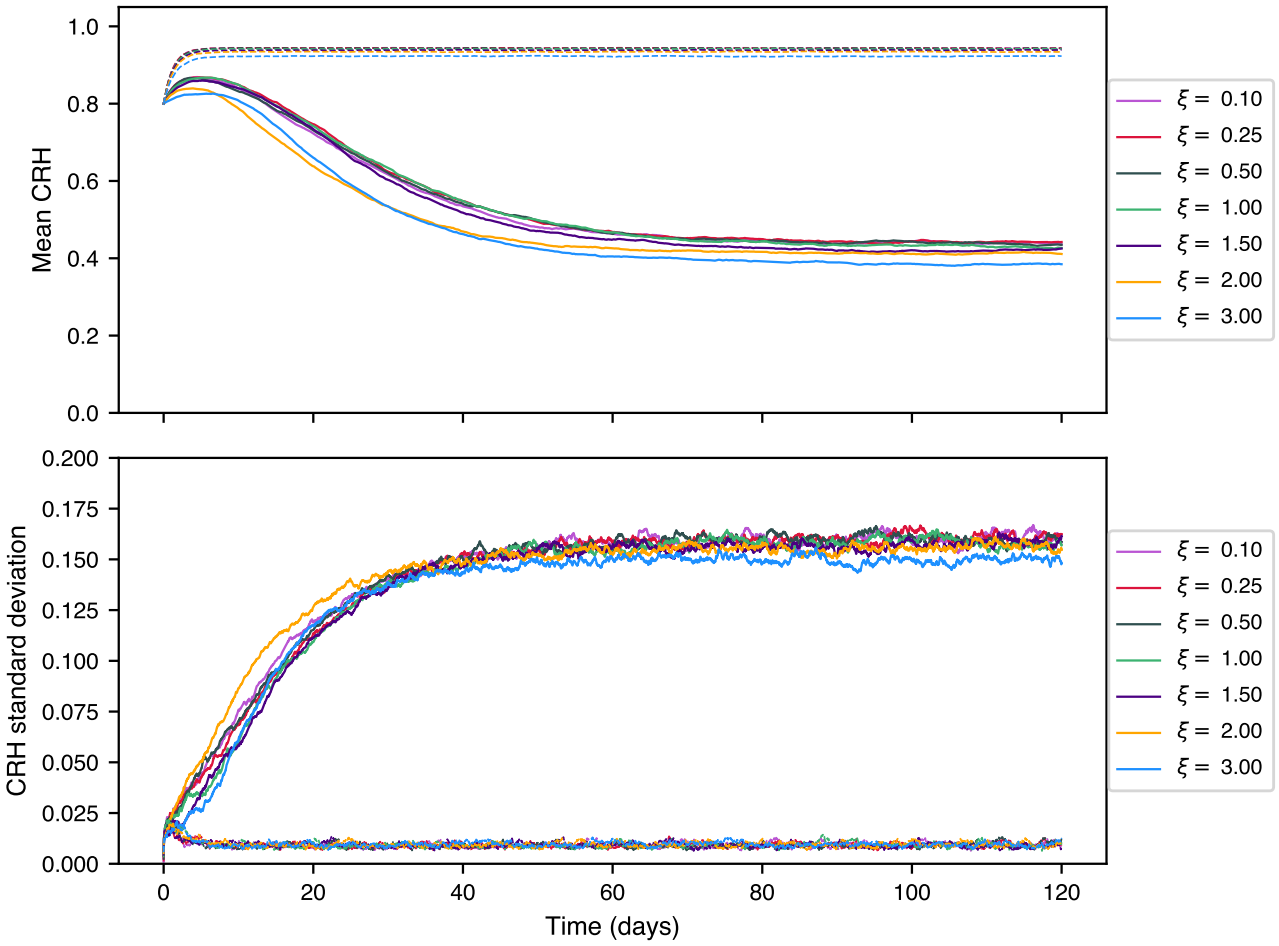

Figure S3. Statistics of the CRH distribution for different runs of the model in terms of Δt in the cases $K = 10^4$ m^{2}s^{-1} (solid lines), 2.5×10^4 m^{2}s^{-1} (dashed lines), with $\tau_{\text{sub}} = 12$ days and $a_d = 14.72$. $\xi = 1$ corresponds to $\Delta t = 100$ s and $\Delta t = 40$ s, respectively.

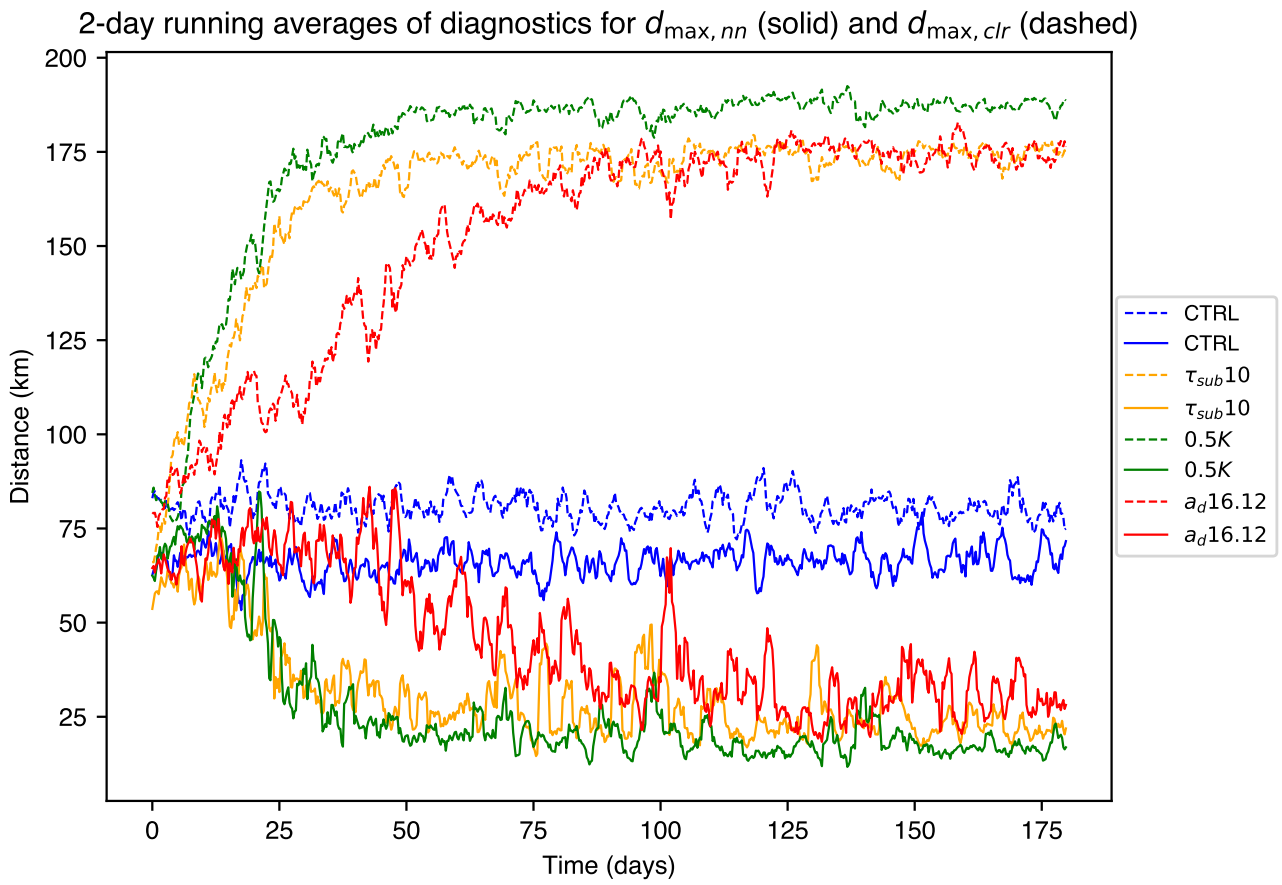

Figure S4. Evolution of the maximum inter-convection nearest neighbour distance (solid lines) and the largest distance from a non-convective to the nearest convective grid cell (dashed), as diagnosed from the numerical model, for the runs of Figure 3. A 2-day running mean has been applied.

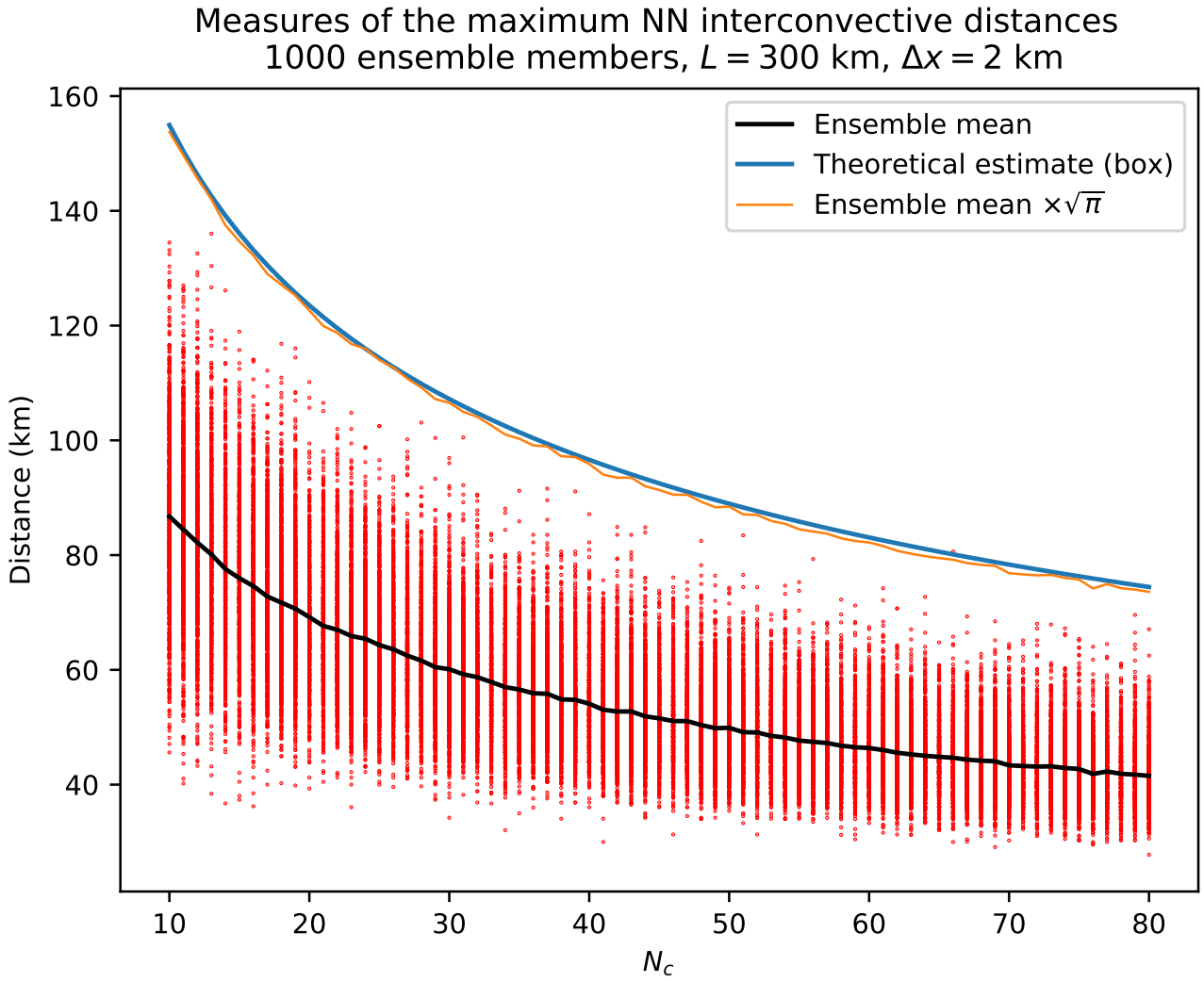

Figure S5. Comparison between the results of a multi-run ensemble of experiments and the theoretical estimate (12). Red dots illustrate the maximum nearest neighbour distances between N_c objects thrown onto a 300 × 300 km domain with 2 km resolution, whereas the black line represents the ensemble mean distance. Examining nearest neighbour distances implies that there are no events within a radius r of the base point, whereas the theory in the paper involves the void probabilities for a square box. The ratio of the area of a circle of radius r to a box of size d is $\sqrt{\pi}$, and we see that multiplying the ensemble mean by this factor (orange line) reproduces exactly the theoretical curve for $\bar{d}_{\mathrm{max,nn}}$ (blue line).

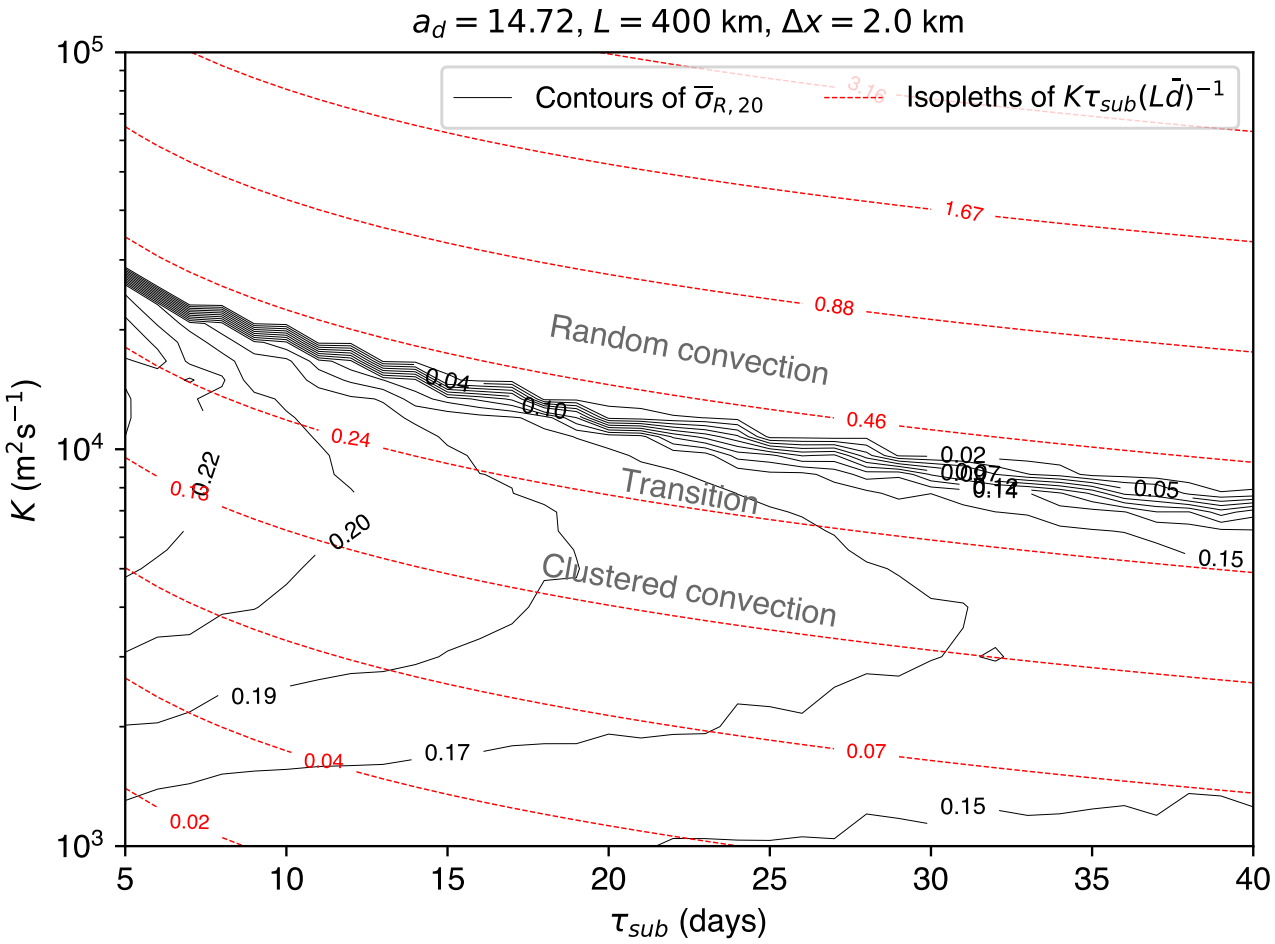

Figure S6. As in Figure 10 but with larger domain, $L = 400$ km.

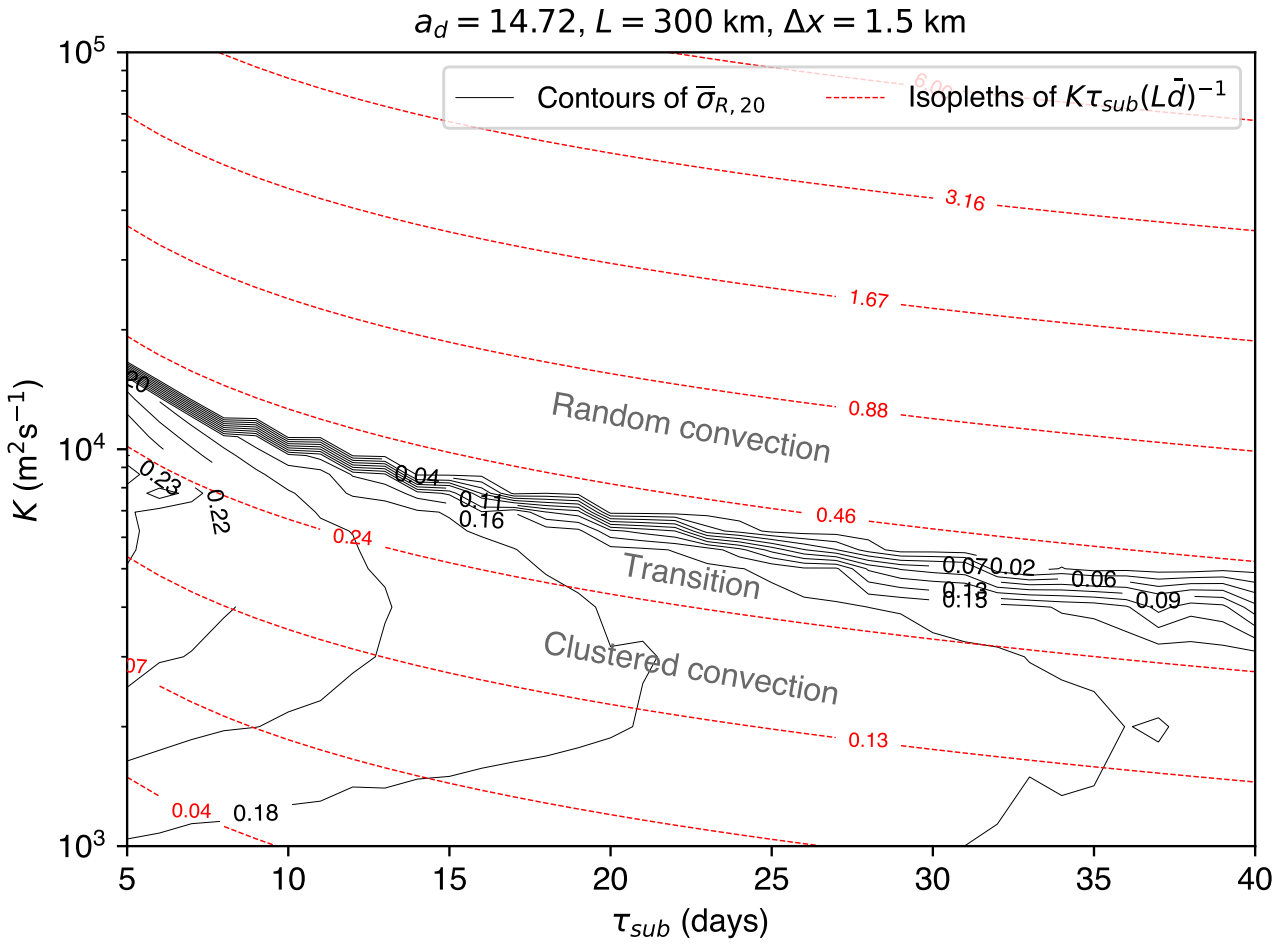

Figure S7. As in Figure 10 but with higher resolution, $\Delta x = 1.5$ km.