Documentation RT code

Jordi Petchamé-Guerrero[1,2,3] and Jesus Carrera[2,3]

[1]Neurite Lab, Sant Joan de la Salle 42, 08022 Barcelona, Spain
[2]Geosciences Department, Institute of Environmental Assessment and Water Research (IDAEA), Severo Ochoa Excellence Center, Consejo Superior de Investigaciones Científicas (CSIC), Jordi Girona 18-26, 08034 Barcelona, Spain
[3]Associated Unit - Hydrogeology Group (UPC-CSIC), 08034 Barcelona, Spain

email address: jordi.petchame.guerrero@gmail.com
ORCID numbers:

- Jordi Petchamé-Guerrero: 0000-0001-8476-0020

- Jesus Carrera: 0000-0002-8054-4352

1 Introduction

The basic WMA equation for target j and time step $k+1$ can be expressed as

$$\mathbf{c}_j^{k+1} = \sum_i \lambda_{ij}^k \hat{\mathbf{c}}_i^k + \mathbf{R}_j^k \tag{1.1}$$

where $\hat{\mathbf{c}}$ accounts for all the different waters affecting the solute mass balance, λ are the mixing ratios and $\mathbf{R}$ represents contributions of reactions to concentrations, which can be computed explicitly or implicitly. Transport is restricted to the mixing between waters, which makes it decoupled from chemical reactions.

2 Initialisation

Let us denote by $\mathbf{d}$ the data in the input file. The nonlinear system to be solved is

$$\mathbf{f}(\mathbf{c}_1) = 0 \tag{2.1}$$

where the residues are

$$f_i(\mathbf{c}_1) = \begin{cases} c_i - d_i, & i \in IC_1 \\ \mathbf{U}_{i,aq} \cdot \mathbf{c}_{aq} - d_i, & i \in IC_2 \\ \gamma_i(\mathbf{c}_1)c_i - d_i, & i \in IC_3 \\ \mathbf{S}_{e_{cstr},1} \cdot \log_{10} \mathbf{a}_1 + S_{e_{cstr},cstr} \log_{10} d_i - \log_{10} K_{e_{cstr}}, & i \in IC_4 \end{cases} \tag{2.2}$$

The sets IC_j are the indices of primary species such that $icon = j$. The subscript $cstr$ refers to the constrain used for the mass action law, whereas the subscript aq denotes aqueous species. In the case of equilibrium with a phase, there are two options: logarithmic (last equation in (2.2)) or arithmetic

$$\forall i \in IC_4, \qquad f_i(\mathbf{c}_1) = d_i^{\nu_{cstr}} \prod_{j \in I_{cstr_1}} a_j^{\nu_j} - K_{e_{cstr}} \tag{2.3}$$

The Jacobian for the logarithmic case is

$$\forall j \in I_1, \qquad \frac{df_i}{dc_j} = \begin{cases} \delta_{ij}, & i \in IC_1 \\ U_{i,j} + \sum_{k=1}^{n_{2,aq}} U_{i,n_1+k} \frac{dc_{2_k}}{dc_j}, & i \in IC_2 \\ \frac{d\gamma_i}{dc_j}c_i + \gamma_i \delta_{ij}, & i \in IC_3 \\ \frac{1}{c_j \log 10} \sum_{k=1}^{n_p} S_{e_{cstr},k} \left(\frac{d \log_{10} \gamma_k}{d \log_{10} c_j} + \delta_{kj} \right), & i \in IC_4 \end{cases} \tag{2.4}$$

where I_1 is the set of primary species indices. Notice that in the last expression the Jacobian of secondary species with respect to primary species is not present. This is because the equilibrium reaction is written in terms of primary aqueous species.

The activity coefficients of aqueous species for ionic strengths less than 1 can be written as

$$\log_{10} \gamma_i = -\alpha_i \frac{A z_i^2 \sqrt{I}}{1 + \beta_i \sqrt{I}} + \gamma_i I, \qquad i \in I_{aq} \tag{2.5}$$

where $\alpha_i, \beta_i, \gamma_i$ are constants. The general expression for the derivative with respect to the ionic strength is

$$\frac{d \log_{10} \gamma_i}{dI} = -\alpha_i \frac{A z_i^2}{2\sqrt{I} \left(1 + \beta_i \sqrt{I}\right)^2} + \gamma_i, \qquad i \in I_{aq} \tag{2.6}$$

For Debye-Hückel, $\alpha_i = 1, \gamma_i = 0$, therefore

$$\frac{d \log_{10} \gamma_i}{dI} = \frac{\log_{10} \gamma_i}{2I \left(1 + \beta_i \sqrt{I}\right)}, \qquad i \in I_{aq} \tag{2.7}$$

The log-Jacobian of (2.5) is

$$\frac{d \log_{10} \gamma_i}{d \log_{10} c_j} = \frac{d \log_{10} \gamma_i}{dI} \frac{dI}{dc_j} \frac{dc_j}{d \log_{10} c_j}, \qquad i \in I_{aq} \tag{2.8}$$

with

$$\frac{dI}{dc_i} = \begin{cases} \frac{1}{2} \left(z_i^2 + \sum_{j=n_p+1}^{n_{aq}} z_j^2 \frac{\partial c_j}{\partial c_i} \right) & i \in I_1 \\ \frac{1}{2} z_i^2 & i \in I_{2,aq} \end{cases} \tag{2.9}$$

and

$$\frac{dc_i}{d \log_{10} c_i} = c_i \log 10 \quad \forall i \in I_{aq}. \tag{2.10}$$

To compute $\frac{\partial \mathbf{c}_{2,aq}}{\partial \mathbf{c}_1}$ we must solve the linear system

$$\left(\mathbf{I}_{n_2} - \mathbf{S}_{e_1}^* \frac{\partial \log_{10} \gamma_1}{\partial \log_{10} \mathbf{c}_2} + \frac{\partial \log_{10} \gamma_2}{\partial \log_{10} \mathbf{c}_2} \right) \frac{d \log_{10} \mathbf{c}_2}{d \log_{10} \mathbf{c}_1} = \mathbf{S}_{e_1}^* \left(\frac{\partial \log_{10} \gamma_1}{\partial \log_{10} \mathbf{c}_1} + \mathbf{I}_{n_1} \right) - \frac{\partial \log_{10} \gamma_2}{\partial \log_{10} \mathbf{c}_1} \tag{2.11}$$

Since

$$\frac{dI}{d\mathbf{c}_{aq}} = \frac{1}{2} \mathbf{z}^2 \begin{pmatrix} \mathbf{I}_1 & 0 \\ \frac{d\mathbf{c}_{2,aq}}{d\mathbf{c}_1} & \mathbf{I}_{2,aq} \end{pmatrix}, \qquad \mathbf{z} = (z_1, z_2, \ldots, z_{n_{aq}}) \tag{2.12}$$

we do not need to compute the derivatives of activity coefficients with respect to concentrations explicitly. Instead, we write the equations in terms of (2.6):

$$\frac{d \log_{10} \gamma_{aq}}{d \log_{10} \mathbf{c}_{aq}} = \frac{\log 10}{2} \left(\frac{d \log_{10} \gamma_{aq}}{dI} \otimes \mathbf{z}^2 \right) \begin{pmatrix} \mathbf{I}_{n_1} & 0 \\ \frac{d\mathbf{c}_{2,aq}}{d\mathbf{c}_1} & \mathbf{I}_{n_{2,aq}} \end{pmatrix} \operatorname{diag}(\mathbf{c}_{aq}) \tag{2.13}$$

Using the identity

$$\frac{d\mathbf{c}_2}{d\mathbf{c}_1} = \operatorname{diag}(\mathbf{c}_2) \frac{d \log_{10} \mathbf{c}_2}{d \log_{10} \mathbf{c}_1} \operatorname{diag}\left(\frac{1}{\mathbf{c}_1}\right) \tag{2.14}$$

The linear system is

$$\mathbf{A} \frac{d \log_{10} \mathbf{c}_2}{d \log_{10} \mathbf{c}_1} = \mathbf{B} \tag{2.15}$$

where

$$\mathbf{A} = \mathbf{I}_{n_2} - \frac{\log 10}{2} \left[\mathbf{S}_{e_1}^* \left(\frac{d \log_{10} \gamma_1}{dI} \otimes \mathbf{z}_2^2 \right) - \frac{d \log_{10} \gamma_2}{dI} \otimes \mathbf{z}_2^2 \right] \operatorname{diag}(\mathbf{c}_2) \tag{2.16}$$

and

$$\mathbf{B} = \mathbf{S}_{e_1}^* \left[\frac{\log 10}{2} \left(\frac{d \log_{10} \gamma_1}{dI} \otimes \mathbf{z}_1^2 \right) \operatorname{diag}(\mathbf{c}_1) + \mathbf{I}_{n_1} \right] - \frac{\log 10}{2} \left(\frac{d \log_{10} \gamma_2}{dI} \otimes \mathbf{z}_1^2 \right) \operatorname{diag}(\mathbf{c}_1) \tag{2.17}$$

Matrices $\mathbf{A}$ and $\mathbf{B}$ are computed using (2.6). Once we have computed the Jacobian of secondary species with respect to primary species, we can compute (2.13) and therefore the Jacobian of the residue (2.4).

2.1 Pseudoalgorithm

1. Set $i = 0$

2. Set initial guess $\mathbf{c}_1^i$

3. Set $j = 0$

4. Set initial guess $(\mathbf{c}_2^i)^j$:

$$\left(\mathbf{c}_2^i\right)^j = (1 + \mu)\,\mathbf{c}_2^{i-1} - \mu\mathbf{c}_2^{i-2}, \quad 0 \leq \mu \leq 1 \tag{2.18}$$

   where $\mathbf{c}_2^{-1} = \mathbf{c}_2^{-2} = \mathbf{c}_2^0$.

5. Compute $\mathbf{c}_2^i(\mathbf{c}_1^i)$ with Picard method:

$$\log_{10}(\mathbf{c}_2^i)^{j+1} = \mathbf{S}_{e_1}^* \log_{10}\left((\gamma_1^i)^j\mathbf{c}_1^i\right) + \log_{10}\mathbf{K}_e^* - \log_{10}(\gamma_2^i)^j$$

   If

$$\left\|\frac{(\mathbf{c}_2^i)^{j+1} - (\mathbf{c}_2^i)^j}{(\mathbf{c}_2^i)^j}\right\| < \varepsilon_r$$

   then set $\mathbf{c}_2^i = (\mathbf{c}_2^i)^{j+1}$. Otherwise, set $j = j + 1$. If $j > n_{it,max}$, go to step 3 and change initial guess.

6. Compute $\mathbf{a}(\mathbf{c}_1^i)$

7. Compute Newton residue $\mathbf{f}(\mathbf{c}_1^i)$ using (2.2). If $\left\|\mathbf{f}(\mathbf{c}_1^i)\right\| < \varepsilon_a$ then set $\mathbf{c} = \mathbf{c}^i$ and exit the algorithm (convergence is attained).

8. Compute $\frac{\partial \log_{10}\gamma}{\partial \log_{10}\mathbf{c}_2}$ and $\frac{\partial \log_{10}\gamma}{\partial I}$

9. Solve linear system (2.15) for each primary species

10. Compute $\frac{\partial \log_{10}\gamma}{\partial \log_{10}\mathbf{c}_1}$

11. Compute Jacobian of Newton residue $\frac{\partial \mathbf{f}\left((\mathbf{c}_1)^i\right)}{\partial(\mathbf{c}_1)^i}$ using (2.4)

12. Solve linear system

$$\frac{\partial \mathbf{f}\left((\mathbf{c}_1)^i\right)}{\partial(\mathbf{c}_1)^i}\Delta(\mathbf{c}_1)^i = -\mathbf{f}\left((\mathbf{c}_1)^i\right)$$

   where $\Delta(\mathbf{c}_1)^i = (\mathbf{c}_1)^{i+1} - (\mathbf{c}_1)^i$. If

$$\left\|\frac{\Delta(\mathbf{c}_1)^i}{(\mathbf{c}_1)^i}\right\| < \varepsilon_r$$

   algorithm does not converge. You must either increase tolerance or change initial guess primary concentrations. Go to step 1.

13. Update $\mathbf{c}_1$ with control factor $\alpha \in (0,1)$:

$$(\mathbf{c}_1)^{i+1} = (\mathbf{c}_1)^i + \Delta(\mathbf{c}_1)^i$$

$$(\mathbf{c}_1)^{i+1} = \begin{cases} \alpha(\mathbf{c}_1)^i & \text{if } (\mathbf{c}_1)^{i+1} \leq \alpha(\mathbf{c}_1)^i \\ \frac{(\mathbf{c}_1)^i}{\alpha} & \text{if } (\mathbf{c}_1)^{i+1} \geq \frac{(\mathbf{c}_1)^i}{\alpha} \end{cases}$$

14. Set $i = i + 1$. If $i > n_{it,max}$ then the algorithm does not converge fast enough. Follow the same procedure as in step 12.

15. Go to step 3.

3   Mathematical formalism

Reactive transport is the mass balance of chemical species subject to transport and chemical reactions (Carrera et al., 2022). For each i-th species, we must solve the following system of equations:

$$\phi \frac{\partial c_i}{\partial t} = M_{ii} \mathcal{L}(c_i) + \phi S_e^T r_e + \phi S_K^T r_K, \qquad i = 1, \ldots, n_{sp}, \tag{3.1}$$

$$S_e \log a = \log K_e, \tag{3.2}$$

$$r_K = r_K(c) \tag{3.3}$$

where $\phi \in (0,1]$ is the porosity, $c_i \geq 0$ is the concentration of species i, $\mathbf{M} \in \mathbb{R}^{n_{sp} \times n_{sp}}$ is a diagonal matrix with 1 if species i is mobile and 0 otherwise, $\mathcal{L}(c) = -\mathbf{q} \cdot \nabla c + \nabla \cdot (\mathbf{D} \nabla c) + f_w (c_r - c)$ is the transport operator in its advective form, S_e is the equilibrium stoichiometric matrix, S_K is the kinetic stoichiometric matrix, r_e contains the equilibrium reaction rates (expressed per unit volume of water), r_K contains the kinetic reaction rates (same units as r_e), a represents species activities, and K_e are the equilibrium constants.

The number of unknowns is $n_{sp} + n_e + n_K$ where n_{sp} is the total number of species, n_e is the number of equilibrium reactions and n_K is the number of kinetic reactions. We must solve these equations for each target (cell, node, or particle depending on the solution method); therefore, the total number of unknowns is $n \times (n_{sp} + n_e + n_K)$.

Equation (3.1) needs to be solved subject to appropriate initial ($c_i(x,t) = c_{i,0}(x)$ throughout the model domain) and boundary conditions, which we write as

$$M_{ii} \left(-\mathbf{q} \mathbf{c}_i + \mathbf{D} \nabla \mathbf{c}_i \right) \cdot \mathbf{n} = M_{ii} \mathbf{q} \mathbf{c}_{i,BC} \cdot \mathbf{n}.$$

For convenience, we write concentrations in matrix form, where rows represent species and columns are targets. This is not the conventional matrix notation; therefore, we must transpose the matrices obtained from the standard Eulerian numerical schemes used to solve (3.1). Let us denote by $\mathbf{c}_{ext} = \left(\mathbf{c} \mid \mathbf{c}_e \right) \in \mathbb{R}^{n_{sp} \times (n+n_{inf})}$ the extended concentrations matrix comprising both $\mathbf{c} \in \mathbb{R}^{n_{sp} \times n}$, matrix of concentrations at all targets, and $\mathbf{c}_e \in \mathbb{R}^{n_{sp} \times n_{inf}}$, matrix of concentrations of external waters (those entering the model domain associated to water sources and those at the boundaries). Let us also denote by $\mathbf{A}_{ext} = \left(\mathbf{A} \mid \mathbf{E} \right) \in \mathbb{R}^{n \times (n+n_{ext})}$ the extended transport matrix comprising both the transport matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ and the external matrix $\mathbf{E} \in \mathbb{R}^{n \times n_{ext}}$ of external water fluxes contributions to the solute mass balance. The discretised version of (3.1) becomes

$$\frac{\partial \mathbf{c}}{\partial t} \mathbf{F} = \mathbf{M} \mathbf{c}_{ext} \mathbf{A}_{ext}^T + \mathbf{S}_e^T r_e \mathbf{F} + \mathbf{S}_K^T r_K \mathbf{F} \tag{3.4}$$

where $\mathbf{F} \in \mathbb{R}^{n \times n}$ is the water storage matrix, which is diagonal if and only if the mesh is homogeneous ($F_{ii} = \phi_i$ in differential formulations such as FD, $F_{ii} = V_i \phi_i$ in integral formulations such as lumped FEM or FV).

Let $\mathbf{S}_{e,nc} \in \mathbb{R}^{n_e \times n_{nc}}$ be the sub-matrix of $\mathbf{S}_e$ that corresponds to the n_{nc} species with non-constant activity. To simplify the system of equations, we pre-multiply equation (3.4) by a component matrix $\mathbf{U} \in \mathbb{R}^{n_p \times n_{sp}}$, which is the orthogonal subspace of $\mathbf{S}_{e,nc}^T$, that is, $\mathbf{U} \mathbf{S}_{e,nc}^T = 0$, which leads to

$$\frac{\partial \mathbf{u}}{\partial t} \mathbf{F} = \tilde{\mathbf{M}} \mathbf{u}_{ext} \mathbf{A}_{ext}^T + \mathbf{U} \mathbf{S}_K^T r_K \mathbf{F} \tag{3.5}$$

where $n_p = n_{sp} - n_e - n_c$ is the number of primary species (that is, the basis of the component matrix), n_c is the number of constant activity species, $\mathbf{u}_{ext} = \mathbf{U} \mathbf{c}_{ext,nc}$, and $\tilde{\mathbf{M}}$ is the same matrix as $\mathbf{M}$ but with dimension $n_p \times n_p$ instead.

4   Water Mixing Approach

Soler-Sagarra et al. (2022) proposed solving the reactive transport problem as a sequence of reactive mixing operations in the shape of (1.1). They argued that any transport formulation (i.e. *master equation*) can be written in the form of equation (1.1), but did it only for explicit solution methods, which restricts the time step value. To extend this solution to implicit or time-weighted schemes, we need to rewrite equation (3.4) as equation (1.1). That is, we need to find the mixing ratios, λ_{jm}^k. This requires specifying both the form of the transport equation and the solution method. For the former, we use the ADE since it is the standard formulation of transport (?), while for the latter we start with the Eulerian advective form proposed by Petchamé-Guerrero and Carrera (2024) which we will generalise to others. The formula for computing the mobile species concentrations in the whole mesh in the time step $k+1$ results from approximating the time derivative in equation (3.4) by its incremental ratio, expressing $\mathbf{c}^{k+\theta_t} = \theta_t \mathbf{c}^{k+1} + (1 - \theta_t) \mathbf{c}^{k+\theta_t}$ and solving for $\mathbf{c}^{k+1}$, which produces

$$\mathbf{c}^{k+1} = \mathbf{c}_{ext}^k \mathbf{\Lambda}^k + \mathbf{R}^{k+\theta_r} \mathbf{B}^k \tag{4.1}$$

where $\mathbf{R}^{k+\theta_r} = \Delta t^k \mathbf{P}$, $\mathbf{B}^k = \left(\mathbf{I} - \theta_t \mathbf{A}_{k,ext}^T \mathbf{F}^{-1}\right)^{-T}$, $\boldsymbol{\Lambda}^k = \left(\mathbf{I} + (1-\theta_t)\mathbf{A}_{k,ext}^T \mathbf{F}^{-1}\right)^T \mathbf{B}^k$ and $\mathbf{A}_{k,ext} = \Delta t^k \mathbf{A}_{ext}$. The resulting matrix that multiplies $\mathbf{c}_{ext}^k$ can be interpreted as containing the mixing ratios of all the waters that mix with each other. Therefore, the first term in the r.h.s. of equation (4.1) represents conservative mixing, which henceforth we will denote by $\tilde{\mathbf{c}}^k$. In the case of explicit transport, since $\theta_t = 0$, equation (4.1) becomes much simpler, because $\mathbf{B}^k = \mathbf{I}$ and $\boldsymbol{\Lambda}^k = \left(\mathbf{I} + \mathbf{A}_{k,ext}^T \mathbf{F}^{-1}\right)^T$. Notice that equation (4.1) still reflects the mass balance of mobile species, therefore, the term $\mathbf{R}^{k+\theta_r}$ represents the total reaction amounts during the $k+1$-th time step. To compute this term, we must choose an optimal time weighting factor θ_r which will depend on the reaction and time step.

4.1 Consistent approach

The consistent WMA scheme gives the following system of equations for each target water j to be solved at every time step $k+1$:

$$\mathbf{u}_j^{k+1} = \sum_{i \in I_j} \lambda_{ji}^k \hat{\mathbf{u}}_i^k + \mathbf{U}_j \sum_{i \in I_j} \beta_{ji}^k \mathbf{S}_{K,nc,i}^T \mathbf{R}_{K,i}^{k+\theta_r} \tag{4.2}$$

where I_j is the set of target waters that mix with target water j (including itself), $\mathbf{S}_{K,nc,i}$ is the kinetic stoichiometric matrix associated to target water i without constant activity species, $\theta_r \in [0,1]$ is the reaction time weighting factor $\left(\mathbf{c}_j^{k+\theta_r} = \theta_r \mathbf{c}_j^{k+1} + (1-\theta_r)\mathbf{c}_j^k\right)$, β^k are the mixing ratios associated to kinetic reaction amounts $\mathbf{R}_K^{k+\theta_r} = \Delta t^k \mathbf{r}_K^{k+\theta_r}$ (i.e. kinetic reaction rates multiplied by time step), and λ^k are the mixing ratios of solute concentrations. The mixing ratios will depend on the transport time weighting factor θ_t. The mixing ratios pseudomatrices might have columns with different dimensions, depending on the amount of waters that mix with each other at each target water, hence the use of the word "pseudo". It is important to point out that the mixing ratios apply to mixing waters as well as boundary or recharge waters, therefore we use the hat notation for the concentrations of mobile species in all of these different waters. More precisely,

$$\hat{\mathbf{c}} = \left(\mathbf{c} \mid \mathbf{c}_e\right)$$

and

$$\hat{\mathbf{u}} = \mathbf{U}\hat{\mathbf{c}}$$

where $\mathbf{c}_e$ represents the external waters. This method was already considered by Soler-Sagarra et al. (2022) who solved reactive transport in an isochrone mesh.

The main disadvantage of this equation is that, in the implicit case, transport and reactions are not decoupled, and therefore it is not possible to parallelise the equations for each target water.

Another issue is that the concentrations now depend on the reaction terms of different waters. This makes sense since the waters that mix after a certain time step will have suffered changes due to the chemical reactions involved.

4.2 Lumped approach

4.2.1 Lumping reaction rates

In order to decouple the components concentrations, since the mixing ratios pseudomatrices will generally be diagonally dominant, we use the *lumped* pseudomatrix $\mathbf{B}^L$ which is diagonal and its entries are given by

$$\beta_{jj}^L = \sum_{i=1}^{n_j} \beta_{ji}$$

where n_j is the number of target waters that mix with target water j (including itself). Therefore we can separately compute the component concentrations at each target. Moreover, the sum of mixing ratios for each target will be 1 if and only if they are stable and the diagonal terms of the storage matrix are identical (another advantage of Lagrangian methods). Henceforth we will assume that this is the case.

4.2.2 Euler explicit in chemical reactions

The explicit WMA scheme gives the following system of equations for each target water j to be solved at every time step $k+1$:

$$\mathbf{u}_j^{k+1} = \mathbf{U}_j \tilde{\mathbf{c}}_j^k + \Delta t^k \mathbf{U}_j \mathbf{S}_{K,nc,j}^T r_K\left(\mathbf{c}_j^k\right) \tag{4.3}$$

where

$$\tilde{\mathbf{c}}_j^k := \sum_{i \in I_j} \lambda_{ji}^k \hat{\mathbf{c}}_{nc,i}^k. \tag{4.4}$$

Once we have computed $\mathbf{u}_j^{k+1}$, we speciate to compute the concentration of variable activity species $\mathbf{c}_{nc_j}^{k+1}$ (see Appendix C). Then, we use the mass balance of the secondary variable activity species to compute the equilibrium reaction rates $\mathbf{r}_e$. The method we have chosen is *ordinary least squares*, which consists in premultiplying the mass balance equation by $\mathbf{S}_{e_2,nc}$ and solving the resulting linear system for the mean equilibrium reaction amounts $\overline{\mathbf{R}}_e = \Delta t \mathbf{r}_e$ (see Appendix D).

4.2.3 Euler implicit in chemical reactions

Let us define the vector $\tilde{\mathbf{u}}_j^k$ as

$$\tilde{\mathbf{u}}_j^k := \mathbf{U}_j \tilde{\mathbf{c}}_j^k. \tag{4.5}$$

If we consider an implicit scheme for the chemical reactions term in the WMA method, we obtain

$$\mathbf{u}_j^{k+1} = \tilde{\mathbf{u}}_j^k + \Delta t^k \mathbf{U}_j \mathbf{S}_{K,nc,j}^T r_K \left(\mathbf{c}_j^{k+\theta_r} \right) \tag{4.6}$$

where $\theta_r \in (0,1]$ is the reaction time weighting factor, $\mathbf{c}_j^{k+\theta_r} = \theta_r \mathbf{c}_j^{k+1} + (1-\theta_r)\mathbf{c}_j^k$. To solve this implicit equation, we use Newton-Raphson with the primary concentrations as argument. The algorithm details can be found in Appendix E.

An alternative of this scheme is to do the time weighting directly in the kinetic reaction rates instead of the concentrations, that is,

$$\mathbf{u}_j^{k+1} = \tilde{\mathbf{u}}_j^k + \Delta t^k \mathbf{U}_j \mathbf{S}_{K,nc,j}^T \mathbf{r}_{K,j}^{k+\theta_r} \tag{4.7}$$

where $\mathbf{r}_{K,j}^{k+\theta_r} = \theta_r r_K \left(\mathbf{c}_j^{k+1} \right) + (1-\theta_r) r_K \left(\mathbf{c}_j^k \right)$.

4.3 Hybrid approaches

There are some possibilities to ensure decoupling in the implicit case without applying the lumping technique. In order to keep the WMA consistent and not lose the decoupling benefit, we can rewrite (4.2) as

$$\mathbf{u}_j^{k+1} = \tilde{\mathbf{u}}_j^k + \mathbf{U}_j \tilde{\mathbf{R}}_{K,j}^{k+\theta_r} + \beta_{jj}^k \mathbf{U}_j \mathbf{S}_{K,nc,j}^T \mathbf{R}_{K,j}^{k+\theta_r} \tag{4.8}$$

where the term $\tilde{\mathbf{R}}_{K,j}^{k+\theta_r}$ represents the concentration changes (after mixing) in variable activity species during a time step due to kinetic reactions. The idea is that this term is already computed for each target water j together with $\tilde{\mathbf{u}}_j^k$, and therefore we only have to iterate at target water j.

There are different options to solve this equation. An option is to apply the lumping technique only in the current target water j and use the mixed reaction rates of the previous time step on the remaining target waters. Therefore, the mixing kinetic term would only depend on the concentrations at the previous time step:

$$\tilde{\mathbf{R}}_{K,j}^{k+\theta_r} = \Delta t^k \sum_{i \in I_j \setminus \{j\}} \beta_{ji}^k \mathbf{S}_{K,nc,i}^T r_K \left(\mathbf{c}_i^k \right). \tag{4.9}$$

Another reasonable method would be to solve from upstream to downstream waters, and update the kinetic reaction rates at each target water iteration. In this case, we have two options:

- Option 1:

$$\tilde{\mathbf{R}}_{K,j}^{k+\theta_r} = \Delta t^k \sum_{i \in U_j} \beta_{ji}^k \mathbf{S}_{K,nc,i}^T r_K \left(\mathbf{c}_i^{k+\theta_r} \right) + \Delta t^k \sum_{i \in D_j} \beta_{ji}^k \mathbf{S}_{K,nc,i}^T r_K \left(\mathbf{c}_i^{k+\theta_r,est} \right) \tag{4.10}$$

- Option 2:

$$\tilde{\mathbf{R}}_{K,j}^{k+\theta_r} = \Delta t^k \sum_{i \in U_j} \beta_{ji}^k \mathbf{S}_{K,nc,i}^T \mathbf{r}_{K,i}^{k+\theta_r} + \Delta t^k \sum_{i \in D_j} \beta_{ji}^k \mathbf{S}_{K,nc,i}^T \mathbf{r}_{K,i}^{k+\theta_r,est} \tag{4.11}$$

where *est* stands for estimated, U_j includes the indices of upstream waters with respect to target water j and D_j the downstream waters.

4.3.1 Estimating reaction amounts

Estimating the downstream kinetic reaction rates is not straightforward and will generally depend on the problem being solved.

- Option 1: If we expect the reaction rates to be similar throughout the whole domain (for instance, if we have mineral zones without reaction fronts) then it might be useful to take the reaction rate of the closest upstream water as an estimate for the downstream waters that belong to the same "kinetic" zone as the closest upstream water, as an anticipation of what will happen. If a downstream water has a different set of kinetic reactions, then we take the value of the previous time step.

- Option 2: Since the kinetic reaction rates will generally be smooth and monotonic with respect to time, it is possible to extrapolate the estimated kinetic rates at the next time step by a linear combination of the two previous values, that is,

$$\mathbf{r}_K^{k+\theta_r,est} = (1 + \mu\theta_r)\,\mathbf{r}_K^k - \mu\theta_r\mathbf{r}_K^{k-1}, \quad 0 \le \mu \le 1. \tag{4.12}$$

- Option 3: While in Option 2 we use the first derivative for the extrapolation, if the reaction rates are not smooth enough, we can use the second derivative to obtain more accuracy. In this case, we will need the values for three time steps instead of just 2. The idea is to approximate the second derivative of the next time step as the derivative of the actual time step. The formula is

$$\frac{\mathbf{r}_K^{k+1} - 2\mathbf{r}_K^k + \mathbf{r}_K^{k-1}}{\Delta t^2} = \frac{\mathbf{r}_K^k - 2\mathbf{r}_K^{k-1} + \mathbf{r}_K^{k-2}}{\Delta t^2}$$

and therefore

$$\mathbf{r}_K^{k+\theta_r,est} = (2\theta_r + 1)\mathbf{r}_K^k - 3\theta_r\mathbf{r}_K^{k-1} + \theta_r\mathbf{r}_K^{k-2}.$$

- Option 4: Another possibility is to estimate them assuming that there is no mixing and using the estimated concentrations in the next time step. We would have to solve the following least squares linear system to obtain the estimated mean kinetic reaction amounts:

$$\left(\mathbf{S}_{K,nc}\mathbf{S}_{K,nc}^T\right)\overline{\mathbf{R}}_K^{k,est} = \mathbf{S}_{K,nc}\left(\mathbf{c}_{nc}^{k+1,est} - \mathbf{c}_{nc}^k\right). \tag{4.13}$$

Once we have solved (4.13) we compute the estimated kinetic reaction rate using the definition:

$$\overline{\mathbf{R}}_K^{k,est} = \Delta t^k\mathbf{r}_K^{k+\theta_r,est}.$$

5 Biofilm growth

We follow the model from Wang et al. (2024):

$$r_{k,im,j,b}^g = \mu_b \frac{c_{D,im,j}}{K_{D,b} + c_{D,im,j}} \frac{c_{A,im,j}}{K_{A,b} + c_{A,im,j}} \frac{I_{A,b}}{I_{A,b} + c_{I,im,j}} \left(1 - \frac{\theta_{bf}}{\phi}\right) c_{im,j,b}, \qquad j = 1,\ldots,N_{im}, \quad b = 1,\ldots,N_b \tag{5.1}$$

where

$$c_{im,j,b} = \rho_b \frac{\theta_{b,j}}{\theta_{bf,j}}$$

assuming constant biomass density ρ_b.

$$r_{k,im,j,b} = \frac{1}{Y}r_{k,im,j,b}^g \tag{5.2}$$

$$r_{k,im,j,b}^d = d_b c_{im,j,b}, \qquad j = 1,\ldots,N_{im}, \quad b = 1,\ldots,N_b \tag{5.3}$$

$$\rho_b \frac{\partial\theta_{b,j}}{\partial t} = \theta_{bf,j}r_{g,im,j,b}, \qquad j = 1,\ldots,N_{im}, \quad b = 1,\ldots,N_b \tag{5.4}$$

where

$$r_{g,im,j,b} = S_{k,im}^g r_{k,im,j,b}^g - S_{k,im}^d r_{k,im,j,b}^d$$

7

A   Convergence Analysis

Let us assume we have a matrix $\mathbf{c} \in \mathbb{R}^{n_{sp} \times n}$ of concentrations, where the rows represent species and columns represent targets. RT of these species can be written as

$$\mathbf{U}\frac{\partial \mathbf{c}}{\partial t}\mathbf{F} = \mathbf{UM}\mathcal{L}(\mathbf{c}) + \mathbf{US}_K^T \mathbf{r}_K \tag{A.1}$$

Let us consider an explicit time integration scheme such that we can write

$$\mathbf{c}^{k+1} = \Phi(\mathbf{c}^k, \Lambda)$$

where k is a time step and Λ is a set of parameters of the equation that are stationary (i.e. they are independent of time). We will make use of the *Contraction Mapping theorem* Name (2022):

Theorem 1 *Let (M, d) be a complete metric space and $F : M \to M$ a contraction, that is, $\exists \mu \in \mathbb{R}, 0 < \mu < 1$, such that*

$$d\left(F(x), F(y)\right) \leq \mu d(x, y) \quad \forall x, y \in M.$$

Then there exists a unique fixed point $x^* \in M$ of F. Moreover, for any choice of $x_0 \in M$, the recursive sequence

$$x_{m+1} = F(x_m), \quad m \geq 0$$

converges to x^*.

Therefore, we can find a critical time step Δt to ensure convergence of our time integration method. WLOG, let us consider the Euler explicit scheme in (A.1):

$$\mathbf{U}\frac{\mathbf{c}^{k+1} - \mathbf{c}^k}{\Delta t}\mathbf{F} = \mathbf{UM}\mathcal{L}(\mathbf{c}^k) + \mathbf{US}_K^T r_K(\mathbf{c}^k) \tag{A.2}$$

We can rewrite this equation so as to isolate $\mathbf{c}^{k+1}$:

$$\mathbf{Uc}^{k+1} = \mathbf{U}\left(\mathbf{c}_a^k \mathbf{B}^T + \Delta t \mathbf{F}^{-1}\mathbf{s}\right) + \Delta t \mathbf{US}_K^T r_K(\mathbf{c}^k)\mathbf{F}^{-1}$$

where $\mathbf{B}$ is the mixing ratios matrix and $\mathbf{s} = (r_1 c_{r_1}, \ldots, r_n c_{r_n})^T$. If we consider the iteration map

$$\Phi(\mathbf{c}^k, \lambda) = \mathbf{c}_a^k \mathbf{B}^T + \Delta t \mathbf{F}^{-1}\mathbf{s} + \Delta t \mathbf{S}_K^T r_K(\mathbf{c}^k)\mathbf{F}^{-1}$$

we can apply Theorem 1. Let $m > l \geq 0$. Let $n_1, n_2 \in \mathbb{N}$ and let $\|\cdot\|$ be a sub-multiplicative matrix norm in $\mathbb{R}^{n_1 \times n_2}$ (that is, $\|AB\| \leq \|A\|\|B\| \forall A, B \in \mathbb{R}^{n_1 \times n_2}$) induced by the same vector norm in $\mathbb{R}^{n_1}, \mathbb{R}^{n_2}$. We have

$$\left\|\mathbf{U}\left(\Phi\left(\mathbf{c}^m\right) - \Phi\left(\mathbf{c}^l\right)\right)\right\| = \left\|\mathbf{U}\left(\mathbf{c}_a^m - \mathbf{c}_a^l\right)\mathbf{B}^T + \Delta t \mathbf{US}_K^T\left(r_k(\mathbf{c}^m) - r_k(\mathbf{c}^l)\right)\mathbf{F}^{-1}\right\|$$

Applying the triangle inequality and the sub-multiplicativity of the matrix norm, we get

$$\left\|\mathbf{U}\left(\Phi\left(\mathbf{c}^m\right) - \Phi\left(\mathbf{c}^l\right)\right)\right\| \leq \|\mathbf{U}\|\left\|\mathbf{B}^T\right\|\left\|\mathbf{c}_a^m - \mathbf{c}_a^l\right\| + \Delta t\|\mathbf{U}\|\left\|\mathbf{S}_K^T\right\|\left\|\mathbf{F}^{-1}\right\|\left\|r_k(\mathbf{c}^m) - r_k(\mathbf{c}^l)\right\|$$

Moreover,

$$\left\|\mathbf{U}\left(\Phi\left(\mathbf{c}^m\right) - \Phi\left(\mathbf{c}^l\right)\right)\right\| \leq \|\mathbf{U}\|\left\|\Phi\left(\mathbf{c}^m\right) - \Phi\left(\mathbf{c}^l\right)\right\| \tag{A.3}$$

Let us assume that r_k is Lipschitz, that is, $\exists K > 0$ such that

$$\left\|r_k(\mathbf{c}_1) - r_k(\mathbf{c}_2)\right\| \leq K\|\mathbf{c}_1 - \mathbf{c}_2\| \quad \forall \mathbf{c}_1, \mathbf{c}_2 \in \mathbb{R}^{n_{sp} \times n}$$

Therefore,

$$\left\|\mathbf{U}\left(\Phi\left(\mathbf{c}^m\right) - \Phi\left(\mathbf{c}^l\right)\right)\right\| \leq \|\mathbf{U}\|\left\|\mathbf{B}^T\right\|\left\|\mathbf{c}^m - \mathbf{c}^l\right\| + \Delta t\|\mathbf{U}\|\left\|\mathbf{S}_K^T\right\|\left\|\mathbf{F}^{-1}\right\|K\left\|\mathbf{c}^m - \mathbf{c}^l\right\|$$

$$= \|\mathbf{U}\|\left(\left\|\mathbf{B}^T\right\| + \Delta t\left\|\mathbf{S}_K^T\right\|\left\|\mathbf{F}^{-1}\right\|K\right)\left\|\mathbf{c}^m - \mathbf{c}^l\right\|$$

Therefore, using (A.3) we deduce that Φ is Lipschitz if and only if

$$\left\|\Phi\left(\mathbf{c}^m\right) - \Phi\left(\mathbf{c}^l\right)\right\| \leq K_\Phi \left\|\mathbf{c}^m - \mathbf{c}^l\right\|$$

where

$$K_\Phi = \left\|\mathbf{B}^T\right\| + \Delta t \left\|\mathbf{S}_K^T\right\| \left\|\mathbf{F}^{-1}\right\| K$$

is the Lipschitz constant of Φ. It remains to find the value of Δt such that Φ is contractive, that is, $K_\Phi < 1$. Clearly, this can only occur if $\left\|\mathbf{B}^T\right\| < 1$. An analysis of stability for the transport equation can be found in Appendix B. Let us denote by $\Delta t_{c,t}$ the critical time step for transport stability. The second necessary condition is that

$$\Delta t \left\|\mathbf{S}_K^T\right\| \left\|\mathbf{F}^{-1}\right\| K < 1 - \left\|\mathbf{B}^T\right\|.$$

which is equivalent to the following condition:

$$\Delta t < \frac{1 - \left\|\mathbf{B}^T\right\|}{\left\|\mathbf{S}_K^T\right\| \left\|\mathbf{F}^{-1}\right\| K} \leq \frac{1 - \rho(\mathbf{B})}{\left\|\mathbf{S}_K^T\right\| \left\|\mathbf{F}^{-1}\right\| K}$$

where $\rho(\cdot)$ denotes the spectral radius. As a result, the critical time step for (A.2) is

$$\Delta t_c = \min \left\{\Delta t_{c,t}, \frac{1 - \left\|\mathbf{B}^T\right\|}{\left\|\mathbf{S}_K^T\right\| \left\|\mathbf{F}^{-1}\right\| K}\right\}$$

The question now is how to obtain the Lipschitz constant of the reaction rates. A sufficient condition for a function to be Lipschitz is to be C^1 and in that case the Lipschitz constant would be $K = \max_{\mathbf{c}} \left\|Dr_k(\mathbf{c})\right\|$. Therefore we must estimate an upper bound to the values of the Jacobian of the reaction rates throughout the time interval we are considering. For instance. let us consider the Organic Carbon aerobic decay reaction:

$$0.25CH_2O + 0.25O_2 \rightarrow 0.25HCO_3^- + 0.25H^+.$$
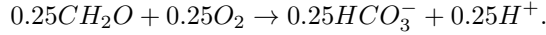
The reaction rate follows Monod kinetics, given by

$$r_k = \mu c_{OC} \frac{c_{O_2}}{k_{O_2} + c_{O_2}}$$

describir formula The expressions for the partial derivatives are

$$\frac{\partial r_k}{\partial c_i} = \begin{cases} \mu \frac{c_{O_2}}{k_{O_2} + c_{O_2}} & \text{if species } i \text{ is } CH_2O \\ r_k \frac{k_i}{c_i(k_i + c_i)} & \text{if species } i \text{ is } O_2 \end{cases}$$

Therefore, an upper bound would be

$$\left\|\frac{\partial r_k}{\partial c_i}\right\| \leq \frac{\mu c_{j,0}}{k_{O_2}}$$

where $c_{j,0}$ is the initial concentration of species $j \neq i$.

Let us now consider a kinetic mineral dissolution/precipitation reaction rate:

$$r_K = \sigma_K \zeta_K e^{-E_a/RT} \sum_{j=1}^{N_p} k_j \left(\Omega^{\theta_j} - 1\right)^{\eta_j} \prod_{i=1}^{N_x} a_{cat_i}^{p_{ji}}$$

The partial derivative with respect to species l is

$$\frac{\partial r_K}{\partial c_l} = \sigma_K \zeta_K e^{-E_a/RT} \frac{\nu_l}{c_l} \sum_{j=1}^{N_p} k_j \eta_j \theta_j \Omega^{\theta_j} \left(\Omega^{\theta_j} - 1\right)^{\eta_j - 1} \prod_{i=1}^{N_x} a_{cat_i}^{p_{ji}}$$

Usually, the parameters θ and η will be 1. In this case, we get

$$\frac{\partial r_K}{\partial c_l} = \sigma_K \zeta_K e^{-E_a/RT} \Omega \frac{\nu_l}{c_l} \sum_{j=1}^{N_p} k_j \prod_{i=1}^{N_x} a_{cat_i}^{p_{ji}}$$

Since for a mineral reaction

$$\Omega = \frac{a_A a_B}{K_e}$$

where A, B are the reactants, we have

$$\frac{\partial r_K}{\partial c_A} = \sigma_K \zeta_K e^{-E_a/RT} \frac{\gamma_A a_B}{K_e} \sum_{j=1}^{N_p} k_j \prod_{i=1}^{N_x} a_{cat_i}^{p_{ji}}$$

and idem for B.

B   Transport Stability Analysis

The discretisation of the stationary flow equation with uniform density using a finite difference scheme yields

$$-\nabla \cdot \mathbf{q} + r = 0 \tag{B.1}$$

$$\frac{q_{i+1/2} - q_{i-1/2}}{\Delta x} = r_i \quad \forall i \in \{1, \dots, n\} \tag{B.2}$$

where i represents the centre of the i-th cell.

Henceforth we will consider a monotonically increasing flow (that is, $r_i > 0 \, \forall i > 0$). We will also assume for simplicity uniform time and space discretisations, stationary flow and sink/source terms, as well as constant dispersion. An Euler discretisation of the ADE gives the following time integration scheme

$$\mathbf{A}\mathbf{c}^{k+1} = \mathbf{B}\mathbf{c}^k + \mathbf{f} \tag{B.3}$$

In (B.3),

$$\begin{cases} \mathbf{A} = \mathbf{I} - \vartheta \mathbf{E} \\ \mathbf{B} = \mathbf{I} + (1 - \vartheta)\mathbf{E} \end{cases} \tag{B.4}$$

where ϑ is the time weighting factor and $\mathbf{E}$ is the matrix that generates the mixing ratios matrix using the scheme proposed in (Petchamé-Guerrero and Carrera, 2024).

$$\begin{cases} E_{i,i-1} = \beta_i + \alpha_{i-1/2} \\ E_{i,i+1} = \beta_i - \alpha_{i+1/2} \\ E_{i,i} = -E_{i,i-1} - E_{i,i+1} - \gamma_i = -2\beta_i - \left(\alpha_{i-1/2} - \alpha_{i+1/2}\right) - \gamma_i \end{cases} \tag{B.5}$$

where

$$\begin{cases} \beta_i = \frac{D\Delta t}{\phi_i \Delta x^2} \\ \alpha_{i\pm 1/2} = \frac{q_{i\pm 1/2}\Delta t}{2\phi_i \Delta x} \\ \gamma_i = \frac{r_i \Delta t}{\phi_i} \end{cases} \tag{B.6}$$

Moreover,

$$f_i = \gamma_i c_{e_i}$$

From the definition of $\alpha_{i\pm 1/2}$ it is easy to notice that

$$\alpha_{i-1/2} - \alpha_{i+1/2} = -\frac{\gamma_i}{2} < 0$$

Therefore

$$E_{i,i} = -2\beta_i - \frac{\gamma_i}{2}$$

To study stability, we choose the L^∞ norm in $\mathbb{R}^n$ and we assume $\vartheta = 0$. We must ensure that

$$\sum_{j=1}^{n} \|B_{i,j}\| \leq 1 \quad \forall i$$

Let us consider row $i \in \{2, \dots, n-1\}$ of $\mathbf{B}$. The stability conditions are ponerlo en funcion de $\Delta t, \Delta x$

- $\beta_i \geq \left\|\alpha_{i\pm 1/2}\right\|$:

  If $B_{i,i} < 0$, then
  $$\beta_i \leq \frac{1}{2}$$

  If $B_{i,i} \geq 0$, then
  $$\begin{cases} \gamma_i \geq 0 \\ \beta_i \leq \frac{1}{2} + \frac{\gamma_i}{4} \end{cases}$$

- $\beta_i < \left\|\alpha_{i\pm 1/2}\right\|$:

  If $B_{i,i} < 0$, then
  $$\begin{cases} 2\beta_i + \frac{\gamma_i}{2} + \left(\alpha_{i-1/2} + \alpha_{i+1/2}\right) \leq 2 \\ \alpha_{i-1/2} + \alpha_{i+1/2} \leq 1 \end{cases}$$

  If $B_{i,i} \geq 0$, then
  $$\begin{cases} \alpha_{i-1/2} + \alpha_{i+1/2} \leq 2\beta_i + \frac{\gamma_i}{2} \quad \text{condicion } \Delta x \\ \alpha_{i-1/2} + \alpha_{i+1/2} \leq 1 \end{cases}$$

C Speciation from components

Once we have computed $\mathbf{u}_j^{k+1}$ explicitly, that is, component concentrations at target j and time step $k+1$, we must speciate to compute the variable activity species concentrations. Let $\mathbf{U}_j := (\mathbf{U}_1 \mid \mathbf{U}_{2,nc})$. This involves solving for each target water j the nonlinear system of equations

$$\begin{cases} f_k\left(\mathbf{c}_{1_j}^{k+1}\right) = \mathbf{U}_1 \mathbf{c}_{1_j}^{k+1} + \mathbf{U}_{2,nc} \mathbf{c}_{2,nc_j}^{k+1} - \mathbf{u}_j^{k+1} & \text{(C.1)} \\ \log_{10} \mathbf{c}_{2,nc_j}^{k+1} = \mathbf{S}_{e,nc_1}^* \log_{10}\left(\gamma_{1_j}^{k+1} \mathbf{c}_{1_j}^{k+1}\right) + \log_{10} \mathbf{K}_e^* - \log_{10} \gamma_{2,nc_j}^{k+1} & \text{(C.2)} \end{cases}$$

where the subscript 1 accounts for primary species and $2, nc$ for secondary variable activity species. We use Newton-Raphson, where f_k is the residual function. The Jacobian is therefore

$$\frac{\partial f_k\left(\mathbf{c}_{1_j}^{k+1}\right)}{\partial \mathbf{c}_{1_j}^{k+1}} = \mathbf{U}_1 + \mathbf{U}_{2,nc} \frac{\partial \mathbf{c}_{2,nc_j}^{k+1}}{\partial \mathbf{c}_{1_j}^{k+1}}. \tag{C.3}$$

To compute $\frac{\partial \mathbf{c}_{2,nc_j}^{k+1}}{\partial \mathbf{c}_{1_j}^{k+1}}$ we must solve the linear system for the logarithms:

$$\left(\mathbf{I} - \mathbf{S}_{e,nc_1}^* \frac{\partial \log_{10} \gamma_{1_j}^{k+1}}{\partial \log_{10} \mathbf{c}_{2,nc_j}^{k+1}} + \frac{\partial \log_{10} \gamma_{2,nc_j}^{k+1}}{\partial \log_{10} \mathbf{c}_{2,nc_j}^{k+1}}\right) \frac{\partial \log_{10} \mathbf{c}_{2,nc_j}^{k+1}}{\partial \log_{10} \mathbf{c}_{1_j}^{k+1}} = \mathbf{S}_{e,nc_1}^* \left(\frac{\partial \log_{10} \gamma_{1_j}^{k+1}}{\partial \log_{10} \mathbf{c}_{1_j}^{k+1}} + \mathbf{I}\right) - \frac{\partial \log_{10} \gamma_{2,nc_j}^{k+1}}{\partial \log_{10} \mathbf{c}_{1_j}^{k+1}} \tag{C.4}$$

Then, we get

$$\frac{\partial \mathbf{c}_{2,nc_j}^{k+1}}{\partial \mathbf{c}_{1_j}^{k+1}} = \text{diag}\left(\mathbf{c}_{2,nc_j}^{k+1}\right) \frac{\partial \log_{10} \mathbf{c}_{2,nc_j}^{k+1}}{\partial \log_{10} \mathbf{c}_{1_j}^{k+1}} \text{diag}\left(\frac{1}{\mathbf{c}_{1_j}^{k+1}}\right) \tag{C.5}$$

Let ε_a be the absolute tolerance and ε_r be the relative tolerance. Let us impose a maximum number of iterations $n_{it,max}$. The algorithm for every time step $k+1$ and target water j is

1. Compute $\mathbf{u}_j^{k+1}$ using (4.3).

2. Set $\mu = 0$.

3. Set $i = 0$ and initialize concentration primary species:
$$\left(\mathbf{c}_{1_j}^{k+1}\right)^i = (1+\mu)\,\mathbf{c}_{1_j}^k - \mu \mathbf{c}_{1_j}^{k-1}, \quad 0 \leq \mu \leq 1 \tag{C.6}$$

4. Given $\left(\mathbf{c}_{1_j}^{k+1}\right)^i$, compute $\left(\mathbf{c}_{2,nc_j}^{k+1}\right)^i$ using (C.2) with the fixed-point method

5. Compute $f_k\left(\left(\mathbf{c}_{1_j}^{k+1}\right)^i\right)$ using (C.1). If $\left\|f_k\left(\left(\mathbf{c}_{1_j}^{k+1}\right)^i\right)\right\| < \varepsilon_a$ then convergence is attained:

   (a) Set $\mathbf{c}_{nc_j}^{k+1} = \left(\mathbf{c}_{nc_j}^{k+1}\right)^i$

   (b) Exit algorithm.

6. Compute $\dfrac{\partial\left(\mathbf{c}_{2,nc_j}^{k+1}\right)^i}{\partial\left(\mathbf{c}_{1_j}^{k+1}\right)^i}$ using (C.5)

7. Compute $\dfrac{\partial f_k\left(\left(\mathbf{c}_{1_j}^{k+1}\right)^i\right)}{\partial\left(\mathbf{c}_{1_j}^{k+1}\right)^i}$ using (C.3).

8. Solve linear system

$$\frac{\partial f_k\left(\left(\mathbf{c}_{1_j}^{k+1}\right)^i\right)}{\partial\left(\mathbf{c}_{1_j}^{k+1}\right)^i}\Delta\left(\mathbf{c}_{1_j}^{k+1}\right)^i = -f_k\left(\left(\mathbf{c}_{1_j}^{k+1}\right)^i\right)$$

   where $\Delta\left(\mathbf{c}_{1_j}^{k+1}\right)^i = \left(\mathbf{c}_{1_j}^{k+1}\right)^{i+1} - \left(\mathbf{c}_{1_j}^{k+1}\right)^i$.

9. If

$$\left\|\frac{\Delta\left(\mathbf{c}_{1_j}^{k+1}\right)^i}{\left(\mathbf{c}_{1_j}^{k+1}\right)^i}\right\| < \varepsilon_r$$

   then go to step 2 and increase μ. If $\mu = 1$, algorithm does not converge. You must either increase tolerance and therefore restart the time loop or change initialisation formula (C.6).

10. Update $\mathbf{c}_{1_j}^{k+1}$ with control factor $\alpha \in (0,1)$:

$$\left(\mathbf{c}_{1_j}^{k+1}\right)^{i+1} = \left(\mathbf{c}_{1_j}^{k+1}\right)^i + \Delta\left(\mathbf{c}_{1_j}^{k+1}\right)^i$$

$$\left(\mathbf{c}_{1_j}^{k+1}\right)^{i+1} = \begin{cases} \alpha\left(\mathbf{c}_{1_j}^{k+1}\right)^i & \text{if } \left(\mathbf{c}_{1_j}^{k+1}\right)^{i+1} \leq \alpha\left(\mathbf{c}_{1_j}^{k+1}\right)^i \\ \dfrac{\left(\mathbf{c}_{1_j}^{k+1}\right)^i}{\alpha} & \text{if } \left(\mathbf{c}_{1_j}^{k+1}\right)^{i+1} \geq \dfrac{\left(\mathbf{c}_{1_j}^{k+1}\right)^i}{\alpha} \end{cases}$$

11. Set $i = i + 1$. If $i > n_{it,max}$ then the algorithm does not converge fast enough. Follow the same procedure as in step 9.

12. Go to step 4.

D   Computation of mean equilibrium reaction amounts

The mass balance equation for the secondary variable activity species in target j at time step $k + 1$ is

$$\mathbf{c}_{2,nc_j}^{k+1} = \tilde{\mathbf{c}}_{2,nc_j}^k + \mathbf{S}_{e_{2,nc}}^T\overline{\mathbf{R}}_e + \mathbf{S}_{K_{2,nc}}^T\overline{\mathbf{R}}_K \tag{D.1}$$

where $\overline{\mathbf{R}}_e = \Delta t\mathbf{r}_e, \overline{\mathbf{R}}_K = \Delta t r_K\left(\mathbf{c}_j^{k+\theta_r}\right)$ or $\Delta t\mathbf{r}_{K,j}^{k+\theta_r}$. Equation (D.1) can be rewritten as

$$\mathbf{S}_{e_{2,nc}}^T\overline{\mathbf{R}}_e = \mathbf{c}_{2,nc_j}^{k+1} - \tilde{\mathbf{c}}_{2,nc_j}^k - \mathbf{S}_{K_{2,nc}}^T\overline{\mathbf{R}}_K \tag{D.2}$$

Premultiplying (D.2) by $\mathbf{S}_{e2,nc}$ we obtain the linear system with unknown $\overline{\mathbf{R}}_e$

$$\mathbf{A}\overline{\mathbf{R}}_e = \mathbf{b} \tag{D.3}$$

where

$$\mathbf{A} = \mathbf{S}_{e2,nc}\mathbf{S}_{e2,nc}^T$$

and

$$\mathbf{b} = \mathbf{S}_{e2,nc}\left(\mathbf{c}_{2,nc_j}^{k+1} - \tilde{\mathbf{c}}_{2,nc_j}^k - \mathbf{S}_{K2,nc}^T\overline{\mathbf{R}}_K\right).$$

E Algorithm lumped implicit WMA

At every time step $k+1$ we solve for each target water j the nonlinear system of equations

$$\begin{cases} f_k\left(\mathbf{c}_{1_j}^{k+1}\right) = \mathbf{U}_j\mathbf{c}_{nc_j}^{k+1} - \tilde{\mathbf{u}}_j^k - \Delta t^k\mathbf{U}_j\mathbf{S}_{K,nc,j}^T r_K\left(\mathbf{c}_j^{k+\theta_r}\right) \\ \log_{10}\mathbf{c}_{2,nc_j}^{k+1} = \mathbf{S}_{e,nc_1}^*\log_{10}\left(\gamma_{1_j}^{k+1}\mathbf{c}_{1_j}^{k+1}\right) + \log_{10}\mathbf{K}_e^* - \log_{10}\gamma_{2,nc_j}^{k+1} \end{cases} \tag{E.1}$$

with Newton-Raphson, where f_k is the residual function. The Jacobian is therefore

$$\frac{\partial f_k\left(\mathbf{c}_{1_j}^{k+1}\right)}{\partial\mathbf{c}_{1_j}^{k+1}} = \mathbf{U}_1 + \mathbf{U}_{2,nc}\frac{\partial\mathbf{c}_{2,nc_j}^{k+1}}{\partial\mathbf{c}_{1_j}^{k+1}} - \theta_r\Delta t^k\mathbf{U}_j\mathbf{S}_{K,nc,j}^T\left(\frac{\partial r_K\left(\mathbf{c}_j^{k+\theta_r}\right)}{\partial\mathbf{c}_{1_j}^{k+\theta_r}} + \frac{\partial r_K\left(\mathbf{c}_j^{k+\theta_r}\right)}{\partial\mathbf{c}_{2,nc_j}^{k+\theta_r}}\frac{\partial\mathbf{c}_{2,nc_j}^{k+1}}{\partial\mathbf{c}_{1_j}^{k+1}}\right) \tag{E.2}$$

To compute $\frac{\partial\mathbf{c}_{2,nc_j}^{k+1}}{\partial\mathbf{c}_{1_j}^{k+1}}$ we must solve (C.4) and (C.5). Let ε_a be the absolute tolerance and ε_r be the relative tolerance. Let us impose a maximum number of iterations $n_{it,max}$. The algorithm for every time step $k+1$ and target water j is

1. Compute mixing ratios $\mathbf{\Lambda}_j^k$ (only if they are time-dependent)

2. Compute $\tilde{\mathbf{c}}_j^k$ using (4.4).

3. Set $\mu = 0$.

4. Set $i = 0$ and initialise concentration primary species using (C.6)

5. Given $\left(\mathbf{c}_{1_j}^{k+1}\right)^i$, compute $\left(\mathbf{c}_{2,nc_j}^{k+1}\right)^i$ using (C.2) with the fixed-point method

6. Given $\left(\mathbf{c}_{1_j}^{k+1}\right)^i$ and $\left(\mathbf{c}_{2,nc_j}^{k+1}\right)^i$, compute $\left(\mathbf{c}_{1_j}^{k+\theta_r}\right)^i$ and $\left(\mathbf{c}_{2,nc_j}^{k+\theta_r}\right)^i$

7. Compute $r_K\left(\left(\mathbf{c}_j^{k+\theta_r}\right)^i\right)$

8. Compute $f_k\left(\left(\mathbf{c}_{1_j}^{k+1}\right)^i\right)$ using (E.1). If $\left\|f_k\left(\left(\mathbf{c}_{1_j}^{k+1}\right)^i\right)\right\| < \varepsilon_a$ then convergence is attained:

    (a) Set $\mathbf{c}_{nc_j}^{k+1} = \left(\mathbf{c}_{nc_j}^{k+1}\right)^i$

    (b) Exit algorithm.

9. Compute $\frac{\partial\left(\mathbf{c}_{2,nc_j}^{k+1}\right)^i}{\partial\left(\mathbf{c}_{1_j}^{k+1}\right)^i}$ using (C.4) and (C.5)

10. Compute $\frac{\partial r_K\left(\left(\mathbf{c}_j^{k+\theta_r}\right)^i\right)}{\partial\left(\mathbf{c}_{1_j}^{k+\theta_r}\right)^i}$ and $\frac{\partial r_K\left(\left(\mathbf{c}_j^{k+\theta_r}\right)^i\right)}{\partial\left(\mathbf{c}_{2,nc_j}^{k+\theta_r}\right)^i}$

11. Compute $\frac{\partial f_k\left(\left(\mathbf{c}_{1_j}^{k+1}\right)^i\right)}{\partial\left(\mathbf{c}_{1_j}^{k+1}\right)^i}$ using (E.2)

12. Solve linear system

$$\frac{\partial f_k \left(\left(\mathbf{c}_{1_j}^{k+1}\right)^i\right)}{\partial \left(\mathbf{c}_{1_j}^{k+1}\right)^i} \Delta \left(\mathbf{c}_{1_j}^{k+1}\right)^i = -f_k \left(\left(\mathbf{c}_{1_j}^{k+1}\right)^i\right)$$

where $\Delta \left(\mathbf{c}_{1_j}^{k+1}\right)^i = \left(\mathbf{c}_{1_j}^{k+1}\right)^{i+1} - \left(\mathbf{c}_{1_j}^{k+1}\right)^i$.

13. If

$$\left\| \frac{\Delta \left(\mathbf{c}_{1_j}^{k+1}\right)^i}{\left(\mathbf{c}_{1_j}^{k+1}\right)^i} \right\| < \varepsilon_r$$

then go to step 3 and increase μ. If $\mu = 1$, algorithm does not converge. You must either increase tolerance and therefore restart the time loop or change initialisation formula (C.6).

14. Update $\mathbf{c}_{1_j}^{k+1}$ with control factor $\alpha \in (0, 1)$:

$$\left(\mathbf{c}_{1_j}^{k+1}\right)^{i+1} = \left(\mathbf{c}_{1_j}^{k+1}\right)^i + \Delta \left(\mathbf{c}_{1_j}^{k+1}\right)^i$$

$$\left(\mathbf{c}_{1_j}^{k+1}\right)^{i+1} = \begin{cases} \alpha \left(\mathbf{c}_{1_j}^{k+1}\right)^i & \text{if } \left(\mathbf{c}_{1_j}^{k+1}\right)^{i+1} \leq \alpha \left(\mathbf{c}_{1_j}^{k+1}\right)^i \\ \frac{\left(\mathbf{c}_{1_j}^{k+1}\right)^i}{\alpha} & \text{if } \left(\mathbf{c}_{1_j}^{k+1}\right)^{i+1} \geq \frac{\left(\mathbf{c}_{1_j}^{k+1}\right)^i}{\alpha} \end{cases}$$

15. Set $i = i + 1$. If $i > n_{it,max}$ then the algorithm does not converge fast enough. Follow the same procedure as in step 13.

16. Go to step 5.

F   Algorithm consistent implicit WMA

Similarly to the lumped case, at every time step $k + 1$ we solve for each target water j the nonlinear system of equations

$$\begin{cases} f_k \left(\mathbf{c}_{1_j}^{k+1}\right) = \mathbf{U}_j \mathbf{c}_{nc_j}^{k+1} - \tilde{\mathbf{u}}_j^k - \mathbf{U}_j \tilde{\mathbf{R}}_{K,j}^{k+\theta_r} + \Delta t^k \beta_{jj}^k \mathbf{U}_j \mathbf{S}_{K,nc,j}^T r_K \left(\mathbf{c}_j^{k+\theta_r}\right) \\ \log_{10} \mathbf{c}_{2,nc_j}^{k+1} = \mathbf{S}_{e,nc_1}^* \log_{10} \left(\gamma_{1_j}^{k+1} \mathbf{c}_{1_j}^{k+1}\right) + \log_{10} \mathbf{K}_e^* - \log_{10} \gamma_{2,nc_j}^{k+1} \end{cases} \tag{F.1}$$

with Newton-Raphson. The Jacobian is now

$$\frac{\partial f_k \left(\mathbf{c}_{1_j}^{k+1}\right)}{\partial \mathbf{c}_{1_j}^{k+1}} = \mathbf{U}_1 + \mathbf{U}_{2,nc} \frac{\partial \mathbf{c}_{2,nc_j}^{k+1}}{\partial \mathbf{c}_{1_j}^{k+1}} - \theta_r \Delta t^k \beta_{jj}^k \mathbf{U}_j \mathbf{S}_{K,nc,j}^T \left(\frac{\partial r_K \left(\mathbf{c}_j^{k+\theta_r}\right)}{\partial \mathbf{c}_{1_j}^{k+\theta_r}} + \frac{\partial r_K \left(\mathbf{c}_j^{k+\theta_r}\right)}{\partial \mathbf{c}_{2,nc_j}^{k+\theta_r}} \frac{\partial \mathbf{c}_{2,nc_j}^{k+1}}{\partial \mathbf{c}_{1_j}^{k+1}} \right). \tag{F.2}$$

The algorithm is exactly the same as in Appendix E except that we now use (F.1) and (F.2) instead of (E.1) and (E.2), and we must compute the term $\tilde{\mathbf{R}}_{K,j}^{k+\theta_r}$ in Step 2 using either (4.9), (4.10) or (4.11).

References

Carrera, J., Saaltink, M. W., Soler-Sagarra, J., Wang, J., and Valhondo, C. (2022). Reactive transport: a review of basic concepts with emphasis on biochemical processes. *Energies*, 15(3):925.

Name, A. (2022). Title of the paper. *Journal Name*, X(Y):Z.

Petchamé-Guerrero, J. and Carrera, J. (2024). An improved scheme for the finite difference approximation of the advective term in the heat or solute transport equations. *Transport in Porous Media*, 151(15):2795–2817.

Soler-Sagarra, J., Saaltink, M. W., Nardi, A., De Gaspari, F., and Carrera, J. (2022). Water mixing approach (wma) for reactive transport modeling. *Advances in Water Resources*, 161:104131.

Wang, J., Carrera, J., Saaltink, M. W., Petchamé-Guerrero, J., Herrera, G. S., and Valhondo, C. (2024). Biofilm growth in porous media well approximated by fractal multirate mass transfer with advective-diffusive solute exchange. *Water Resources Research*, 60(11):e2023WR036872.