

## Analytical solution of the flask-sampling model

Editor comment from Jan Kaiser ([J.Kaiser@uea.ac.uk](mailto:J.Kaiser@uea.ac.uk))

During the pre-discussion phase of the paper "Isotopic signatures of production and uptake of H<sub>2</sub> by soil", the first author Qianjie Chen and I (as Editor) exchanged a series of emails about the mathematical analysis of the flask-sampling model (Appendix A of the paper). Here, I would like to summarise the outcome of our email correspondence because it may be helpful for future applications involving large volume sample collection from chamber incubations. I should emphasise that this comment implies no request on the authors to follow this alternative data analysis approach.

In my correspondence with the first author, I pointed out that the equations of the flask-sampling model could be solved analytically. The analytical solution is entirely equivalent to the numerical solution presented in the paper. Its main advantage is that it allows directly solving the flask-sampling model in terms of  $k$  and  $c_e$  for individual experiments.

To derive the analytical solution, there are three different cases to consider depending on how many flasks are connected to the chamber and are flushed through by the sample stream (Case A: 3 flasks; case B: 2 flasks; case C: 1 flask).

In the following, I demonstrate step-by-step the solution for case A and give the final results in the equivalent cases B and C.

### Case A: 3 flasks

Rewriting Eqns. (A1) to (A4) in Appendix A of the paper, using  $g = f / V(\text{flask})$ ,  $g' = f / V(\text{chamber})$  and  $c_e = P / k$  and renaming the chamber mole fraction  $c = c_0$  for simplicity, gives

$$\frac{dc_2}{dt} = g(c - c_2) \quad (1)$$

$$\frac{dc_3}{dt} = g(c_2 - c_3) \quad (2)$$

$$\frac{dc_4}{dt} = g(c_3 - c_4) \quad (3)$$

$$\frac{dc}{dt} = g'(c_4 - c) + k(c_e - c) \quad (4)$$

To solve this system of coupled differential equations with four variables ( $c$ ,  $c_2$ ,  $c_3$  and  $c_4$ ), it is transformed into an equivalent linear system using repeated differentiation and substitution.

To eliminate  $c_4$ , Eq. (4) is differentiated and Eqs. (3) and (4) are substituted into the result:

$$\begin{aligned} \frac{d^2c}{dt^2} &= g' \frac{dc_4}{dt} - (g' + k) \frac{dc}{dt} \\ &= g'g(c_3 - c_4) - (g' + k) \frac{dc}{dt} \\ &= g'gc_3 - g \left[ \frac{dc}{dt} - kc_e + (g' + k)c \right] - (g' + k) \frac{dc}{dt} \\ &= g'gc_3 + g \left[ kc_e - (g' + k)c \right] - (g + g' + k) \frac{dc}{dt} \end{aligned} \quad (5)$$

To eliminate  $c_3$ , Eq. (5) is differentiated and Eqs. (2) and (5) are substituted into the result:

$$\begin{aligned}
\frac{d^3c}{dt^3} &= g'g \frac{dc_3}{dt} - g(g'+k) \frac{dc}{dt} - (g+g'+k) \frac{d^2c}{dt^2} \\
&= g'g^2(c_2 - c_3) - g(g'+k) \frac{dc}{dt} - (g+g'+k) \frac{d^2c}{dt^2} \\
&= g'g^2c_2 + g \left\{ g[kc_e - (g'+k)c] - (g+g'+k) \frac{dc}{dt} - \frac{d^2c}{dt^2} \right\} - g(g'+k) \frac{dc}{dt} - (g+g'+k) \frac{d^2c}{dt^2} \\
&= g'g^2c_2 + g^2[kc_e - (g'+k)c] - g(g+2g'+2k) \frac{dc}{dt} - (2g+g'+k) \frac{d^2c}{dt^2}
\end{aligned} \tag{6}$$

To eliminate  $c_2$ , Eq. (6) is differentiated and Eqs. (1) and (6) are substituted into the result:

$$\begin{aligned}
\frac{d^4c}{dt^4} &= g'g^2 \frac{dc_2}{dt} - g^2(g'+k) \frac{dc}{dt} - g(g+2g'+2k) \frac{d^2c}{dt^2} - (2g+g'+k) \frac{d^3c}{dt^3} \\
&= g'g^3(c - c_2) - g^2(g'+k) \frac{dc}{dt} - g(g+2g'+2k) \frac{d^2c}{dt^2} - (2g+g'+k) \frac{d^3c}{dt^3} \\
&= g'g^3c - g^2(g'+k) \frac{dc}{dt} - g(g+2g'+2k) \frac{d^2c}{dt^2} - (2g+g'+k) \frac{d^3c}{dt^3} \\
&\quad + g^3[kc_e - (g'+k)c] - g^2(g+2g'+2k) \frac{dc}{dt} - g(2g+g'+k) \frac{d^2c}{dt^2} - g \frac{d^3c}{dt^3} \\
&= g^3(kc_e - kc) - g^2(g+3g'+3k) \frac{dc}{dt} - 3g(g+g'+k) \frac{d^2c}{dt^2} - (3g+g'+k) \frac{d^3c}{dt^3}
\end{aligned} \tag{7}$$

This gives the following set of equations:

$$\frac{d^4c}{dt^4} + (3g+g'+k) \frac{d^3c}{dt^3} + 3g(g+g'+k) \frac{d^2c}{dt^2} + g^2(g+3g'+3k) \frac{dc}{dt} + g^3k(c - c_e) = 0 \tag{8}$$

$$c_2 = c + \frac{1}{g'g^2} \left[ \frac{d^3c}{dt^3} + (2g+g'+k) \frac{d^2c}{dt^2} + g(g+2g'+2k) \frac{dc}{dt} + g^2k(c - c_e) \right] \tag{9}$$

$$c_3 = c + \frac{1}{g'g} \left[ \frac{d^2c}{dt^2} + (g+g'+k) \frac{dc}{dt} + gk(c - c_e) \right] \tag{10}$$

$$c_4 = c + \frac{1}{g'} \left[ \frac{dc}{dt} + k(c - c_e) \right] \tag{11}$$

The following ansatz is used for the time-dependent mole fraction in the chamber:

$$c = a_{31}e^{m_{31}t} + a_{32}e^{m_{32}t} + a_{33}e^{m_{33}t} + a_{34}e^{m_{34}t} + c_e \tag{12}$$

To find the coefficients  $m_{31}$  to  $m_{34}$ , the following equation is solved in terms of  $m$ :

$$m^4 + (g'+3g+k)m^3 + 3g(g'+g+k)m^2 + g^2(3g'+g+3k)m + g^3k = 0 \tag{13}$$

It is possible to find analytical solutions to this quartic equation, e.g. using Cardano's method and express them entirely in terms of the coefficients,  $g'$ ,  $g$  and  $k$ . I have not done this, but instead used a numerical method. In the case of the example given in Fig. 5 of the paper ( $g' = 1/11.4 \text{ min}^{-1}$ ,  $g = 1 \text{ min}^{-1}$ ,  $k = 0.1 \text{ min}^{-1}$ ), this results in the following coefficients (shown with five significant digits):

$$m_{31} = -1.4150 \text{ min}^{-1}$$

$$m_{32} = (-0.84817 + 0.45365i) \text{ min}^{-1}$$

$$m_{33} = (-0.84817 - 0.45365i) \text{ min}^{-1}$$

$$m_{34} = -0.076386 \text{ min}^{-1}$$

The coefficients  $a_{31}$  to  $a_{34}$  can be determined from the initial condition  $c(0) = c_2(0) = c_3(0) = c_4(0) = c_i$ , where  $c_i$  corresponds to the initial  $H_2$  mole fraction in the system.

Using Eqns. (9) to (12), this gives the following set of coefficients:

$$a_{31} = \frac{k[(g' + k)(g' + k + m_{34} + m_{33} + m_{32}) + m_{34}m_{33} + m_{34}m_{32} + m_{33}m_{32}] + m_{32}m_{33}m_{34}}{(m_{31} - m_{32})(m_{31} - m_{33})(m_{31} - m_{34})}(c_e - c_i) \quad (14)$$

$$a_{32} = \frac{k[(g' + k)(g' + k + m_{34} + m_{33} + m_{31}) + m_{34}m_{33} + m_{34}m_{31} + m_{33}m_{31}] + m_{31}m_{33}m_{34}}{(m_{32} - m_{31})(m_{32} - m_{33})(m_{32} - m_{34})}(c_e - c_i) \quad (15)$$

$$a_{33} = \frac{k[(g' + k)(g' + k + m_{34} + m_{32} + m_{31}) + m_{34}m_{32} + m_{34}m_{31} + m_{32}m_{31}] + m_{31}m_{32}m_{34}}{(m_{33} - m_{31})(m_{33} - m_{32})(m_{33} - m_{34})}(c_e - c_i) \quad (16)$$

$$a_{34} = \frac{k[(g' + k)(g' + k + m_{33} + m_{32} + m_{31}) + m_{33}m_{32} + m_{33}m_{31} + m_{32}m_{31}] + m_{31}m_{32}m_{33}}{(m_{34} - m_{31})(m_{34} - m_{32})(m_{34} - m_{33})}(c_e - c_i) \quad (17)$$

Again, using the example given in Fig. 5 of the paper, these coefficients have the following values:

$$a_{31} = 3.0783 \text{ nmol mol}^{-1}$$

$$a_{32} = (6.7484 - 4.8312i) \text{ nmol mol}^{-1}$$

$$a_{33} = (6.7484 + 4.8312i) \text{ nmol mol}^{-1}$$

$$a_{34} = 413.42 \text{ nmol mol}^{-1}$$

The first and fourth terms in Eq. (12) describe exponential decay functions. The second and third terms are conjugate complex and can be re-written as a combination of exponential decay and harmonic functions ( $\cos \omega t$ ,  $\sin \omega t$ ), where  $\omega$  is given by the imaginary parts of  $m_{32}$  and  $m_{33}$ . The time-dependent solutions for  $c_2$ ,  $c_3$  and  $c_4$  then follow directly from equations (9) to (11).

### Case B: 2 flasks

At time  $t_1$ , the second flask is filled. Two further flasks remain connected. This changes the set of differential equations describing the time dependence of the mole fractions to the following:

$$\frac{dc_3}{dt} = g(c - c_3) \quad (18)$$

$$\frac{dc_4}{dt} = g(c_3 - c_4) \quad (19)$$

$$\frac{dc}{dt} = g'(c_4 - c) + k(c_e - c) \quad (20)$$

Equations (10) and (11) also apply in the two-flask case.

The following ansatz is used to find the time-dependent mole fraction in the chamber:

$$c = a_{21}e^{m_{21}t} + a_{22}e^{m_{22}t} + a_{23}e^{m_{23}t} + c_e \quad (21)$$

To find the coefficients  $m_{21}$  to  $m_{23}$ , the following equation is solved in terms of  $m$ :

$$m^3 + (g' + 2g + k)m^2 + g(2g' + g + 2k)m + g^2k = 0 \quad (22)$$

Again, it would be possible to solve this cubic equation analytically, using Cardano's method and express the solution in terms of the coefficients,  $g'$ ,  $g$  and  $k$ . The alternative numerical method results in the following coefficients in the case of the example given in Fig. 5 of the paper:

$$m_{21} = (-1.0522 + 0.30491i) \text{ min}^{-1}$$

$$m_{22} = (-1.0522 - 0.30491i) \text{ min}^{-1}$$

$$m_{23} = -0.083327 \text{ min}^{-1}$$

The initial conditions correspond to the mole fractions at the time  $t_1$  as described by case A. They differ for the chamber and the two flasks, and are equal to  $c_1 = c(t_1)$ ,  $c_{31} = c_3(t_1)$  and  $c_{41} = c_4(t_1)$ .

$$a_{21} = \frac{gg'(c_{31} - c_{41}) - (g' + k + m_{23} + m_{22})[g'(c_{41} - c_1) - k(c_1 - c_e)] + (c_1 - c_e)m_{22}m_{23}}{(m_{21} - m_{22})(m_{21} - m_{23})} \quad (23)$$

$$a_{22} = \frac{gg'(c_{31} - c_{41}) - (g' + k + m_{23} + m_{21})[g'(c_{41} - c_1) - k(c_1 - c_e)] + (c_1 - c_e)m_{21}m_{23}}{(m_{22} - m_{21})(m_{22} - m_{23})} \quad (24)$$

$$a_{23} = \frac{gg'(c_{31} - c_{41}) - (g' + k + m_{21} + m_{22})[g'(c_{41} - c_1) - k(c_1 - c_e)] + (c_1 - c_e)m_{22}m_{23}}{(m_{23} - m_{21})(m_{23} - m_{22})} \quad (25)$$

Again, using the example given in Fig. 5 of the paper, these coefficients are:

$$a_{21} = (-1.2615 + 1.8114i) \text{ nmol mol}^{-1}$$

$$a_{22} = (-1.2615 - 1.8114i) \text{ nmol mol}^{-1}$$

$$a_{23} = 195.12 \text{ nmol mol}^{-1}$$

Just as in case A, the complex exponential functions could alternatively be written as a combinations of exponential decay and harmonic functions. The time-dependent solutions for  $c_3$  and  $c_4$  then follow directly from equations (21), (10) and (11), using the coefficients for case B.

### Case C: 1 flask

At time  $t_2$ , the third flask is filled. One further flask remains connected. This changes the set of differential equations describing the time dependence of the mole fractions again:

$$\frac{dc_4}{dt} = g(c - c_4) \quad (26)$$

$$\frac{dc}{dt} = g'(c_4 - c) + k(c_e - c) \quad (27)$$

Equation (11) also applies in the two-flask case.

The following ansatz is used to find the time-dependent mole fraction in the chamber:

$$c = a_{11}e^{m_{11}t} + a_{12}e^{m_{12}t} + c_e \quad (28)$$

To find the coefficients  $m_{31}$  to  $m_{32}$ , the following equation is solved in terms of  $m$ :

$$m^2 + (g' + g + k)m + gk = 0 \quad (29)$$

In the case of the example given in Fig. 5 of the paper, this results in the following coefficients:

$$m_{11,12} = -\frac{1}{2}(g' + g + k) \mp \sqrt{(g' + g + k)^2 - 4gk}$$

$$m_{11} = -1.0965 \text{ min}^{-1}$$

$$m_{12} = -0.091197 \text{ min}^{-1}$$

The initial conditions correspond to the mole fractions at the time  $t_2$  as described by case B. They differ for the chamber and the two flasks, and are equal to  $c_2 = c(t_2)$ ,  $c_{32} = c_3(t_2)$  and  $c_{42} = c_4(t_2)$ .

$$a_{11} = \frac{g'(c_{42} - c_2) - (k + m_{12})(c_2 - c_e)}{m_{11} - m_{12}} \quad (30)$$

$$a_{12} = \frac{g'(c_{42} - c_2) - (k + m_{11})(c_2 - c_e)}{m_{12} - m_{11}} \quad (31)$$

Again, using the example given in Fig. 5 of the paper, these coefficients are:

$$a_{11} = -0.66385 \text{ nmol mol}^{-1}$$

$$a_{12} = 85.468 \text{ nmol mol}^{-1}$$

A time  $t_3$ , the fourth flask is filled. Its  $\text{H}_2$  mole fraction is calculated from Eqs. (28) and (11).

### Conclusion

The calculations above show that the flask sampling model can be described in analytically closed form as a function of the initial mole fraction, the known time constants of the system ( $g$  and  $g'$ ) and the sampling time points ( $t_1$ ,  $t_2$  and  $t_3$ ). The mole fractions measured in the flasks give a unique solution for the  $\text{H}_2$  uptake time constant  $k$  and the equilibrium mole fraction  $c_e$ , which is equal to  $P / k$ , and thus gives the  $\text{H}_2$  production rate. It would be sufficient to fill only three flasks (including the initial flask), rather than four, to solve for the two unknowns  $k$  and  $c_e$ . In the case presented in the paper (four flasks filled in total), the analytical system is over-determined, which requires an optimisation approach to be taken to find the best fit. This may give different results, e.g. depending on whether root mean squared differences between modelled and measured mole fractions are used, or whether the mole fractions are first logarithmically normalised (as done in the paper by Chen and co-workers).

### Acknowledgements:

Many thanks to Qianjie Chen for valuable comments on the analytical solution to the flask sampling model.