(Luo et al., 2017)

Comment to "Transient dynamics of terrestrial carbon storage: mathematical foundation and its applications "

Markus Müller[1]

[1]School of Integrative Plant Science Cornell College of Agriculture and Life Sciences

Correspondence: Markus Müler(mm2796@cornell.edu)

Abstract. The objections fall into different categories. Firstly, the claim that the "Carbon Storage Capacity" $\mathbf{X}_c$ as defined in this paper is an attractor can be disproved by a counterexample. The inappropriate use of the term "attractor" is particularly unfortunate since it has a definite meaning in the mathematical theory of nonautonomous dynamical systems, which IS the appropriate tool to study "transient dynamics". Together with the many equations the mathematical terminology and it's subtitle the paper creates the impression that it's claims are founded on a proved mathematical argument, which they are not, as the counter example alone shows. This is however not the only problem.

Secondly I will show by counter examples that the solution $\mathbf{X}(t)$ does NOT "chase" $\mathbf{X}_c(t)$ which is the main part of the argument for its supposed attractivity.

Thirdly I object to the choice of nomenclature in the context suggested by the word "transient" in the title and confirmed by the equations which describe nonautonomous systems. In this context the nomenclature is outright misleading. It turns out that the "Carbon Storage Capacity" as mathematically defined in the paper is NOT the capacity of a nonautonomous system to store carbon, that consequently the "Carbon Storage Potential" is NOT the amount of carbon potentially to be stored or lost in an nonautonomous system that neither is "Residence Time" the time of residence of carbon in the system. If $\mathbf{X}_c$ is not chased, what is "chasing time"? These terms clearly promise an applicability to the transient case that simply does not exist. They are defined for autonomous systems that can reach equilibrium, a concept unsuitable for the presented nonautonomous case, the study of which requires generalizations of the equilibrium concept, namely the above mentioned attractors.

All these concerns do not only relate to the paper under discussion but also to several other publications which express or use the same ideas, including a book, talks and lectures (e.g the annual international training courses, or a very recent seminar at Cornell). A list of references will be given in the paper.still to be done If I had not checked them, the claims of the paper would have compromised my own work too.

To repair and control the confusion two things have to be accomplished: Firstly we have to see where the paper contradicts proven mathematical theory, and secondly we have to access how applications of the "traceability framework". are affected by the lack of "mathematical foundation" as presented in the paper. Are there parts still standing on solid ground? Which parts can be partially salvaged by restricting their application to special cases?

As a first step I will show $\mathbf{X}_c(t)$ to be the equilibria of a family of autonomous systems, created by freezing the original system at time t. These frozen systems provide the context in which the variables "chasing time", "residence time" $\mathbf{RT}$ and

"carbon storage potential" $\mathbf{X}_p$ can be interpreted. The real geometric connection between these frozen systems and the original is shown not to be one of the solution "chasing" but the other way around of limited maneuverability of the frozen systems in phase space preventing their equilibria to be too far away from the solution.

Although the paper emphasizes that it supersedes previous quasi steady state approximations based on temporal averages, these averages have been used in some related publications, citing this paper. I will illustrate how those approximations are related to the solution and compare them with both $\mathbf{X}_c$ and real attractors. Until here would actually be enough for a rebuttal. The next part is actually more closely related to my self defense against being bullied into the application of questionable methods but less easy to refute because albeit the paper is very grandiose in its claim that "traceability analysis " is applicable p 156 left low, 4.4.3 Traceability analysis "The two terms on the right side of Eq. (2) can be decomposed into traceable components (Xia et al.,2013) so as to identify sources of uncertainty in C cycle projections " it is extremely vague about how this method (for autonomous systems) is to be applied to non-autonomous systems. Rather than attempting to show that it is not possible I should perhaps better challenge the authors to elaborate (Yiqi accused me repeatedly of not having read the manuscript that supposedly applied the method, albeit until not having shown where it is described. The author of this manuscript has subsequent to our inquiry by email changed his code and the plots...

Apart from the problematic statements concerned with to transient aspects of diagnostic variables which can be easily contradicted without any generalization of the class of models considered in the paper, I want to challenge, firstly, the assumption that this class covers most of the relevant carbon cycle models Quote:page 157 left in the middle "The theoretical framework developed in this study has the potential to revolutionize model evaluation. Our analysis indicates that the matrix equation as in Eqs. (1) and (2). , and secondly, that their differences can be further resolved by the "traceability analysis".

Complementary to the original work, which starts with a review of 250 published papers we will start from the mathematical perspective of general compartmental models and derive the definition used in the paper by gradually *decreasing* generality, pointing out which kinds of carbon cycle models are *excluded* by the assumptions of the paper. In particular I will discuss how the assumption of linearity excludes not only microbial models but also those where carbon rates depend on contents of pools of other elements, e.g. Nitrogen and Phosphorus, which process based flux formulation are excluded by the factorizability assumptions of the compartmental matrix, into "environmental scalars" $\xi(t)$ and "base line rates" k, why this factorization is ambiguous and therefore unsuitable for inter-comparisons of models with respect to these factors, which removes two levels of from the "traceability" hierarchie, claimed to be applicable to this supposedly "transient" analysis.

1 Outline

X_c discussion starts in subsubsection 2.1.1

In section 2 we derive eq. (1) of the paper from the viewpoint of the theory of the most general compartmental systems. This approach is complementary to the original presentation as it is not restricted to actually *published* but to mathematically *possible* models, emphasizing the assumptions to be able to identify more and more specific terms in the mass balance equation. We discuss how these assumption limit the generality of eq (1) including examples of scenarios excluded by them. We then introduce the diagnostic variables of the "Traceability Framework" , also in the order of decreasing generality. Not all diagnostic variables are possible for all compartmental models. We will start in subsubsection 2.1.1 with the expressions to compute"carbon storage capacity" $\mathbf{X}_c$, "carbon storage potential" $\mathbf{X}_p$ and "residence time" $\mathbf{RT}$ since the formulas can be evaluated for all compartmental models. We address provable mathematical properties and the limited circumstances in which these definitions coincide with the meaning suggested by their names. This will already show that the claims about their predictive superpowers have not only not been justified yet but also that they are unsupportable in principle. We then recall existing theory about nonautonomous systems from the literature including definitions for different kinds of attractors, conditions for their existence as well as definitions of times of residence suitable for the transient case.

In section 3 will present some counter examples and illustrate the general observations with figures.

2 Definitions

2.0.1 Derivation of the matrix representation

We adapt the derivation originally found in (Jacquez et al., 1972) in order to refer to it's implication for the generality of the formulation used in (Luo et al., 2017). Mathematically Compartmental Models are most economically described as graphs, where the set of compartments $\mathcal{P}$ and the set of non-negative fluxes $\mathcal{F}$ form the the nodes and edges respectively. Choosing one of $n!$ possible ways to enumerate the set of pools $\mathcal{P} = \{p_0, \ldots, p_n\}$ where p_0 is the outside world, we write the contents of the pools as x_i for $i \in \{1, \ldots, n\}$ and the fluxes as

$$\begin{aligned}\mathcal{F} = &\{I_{0\to j} > 0 \text{ for } j \in \{1,\ldots,n\}\}\\ &\cup \{F_{i\to j} > 0 \text{ for } i \in \{1,\ldots,n\} \text{ and } j \in \{1,\ldots,n\} j \neq i\} \text{ with } F_{i\to j} = 0 \text{ for } x_i = 0\}\\ &\cup \{F_{i\to 0} \text{ for } i \in \{1,\ldots,n\} \text{ with } F_{i\to 0} = 0 \text{ for } x_i = 0\}\end{aligned}$$

where $I_{0\to j}$ are influxes from the outside into the system $F_{i\to j}$ are fluxes between pools and $F_{i\to 0}$ are fluxes out of the system.

In general all fluxes can depend on all the x_i and time t (trough environmental factors like Temperature $T(t)$ and moisture $W(t)$.

The influxes don't have to depend on the x_i but internal and out fluxes must at least depend on their source pool content to guarantee the condition that there is no outflux from an empty pool. For every pool we have a mass balance equation.

$$\frac{d}{dt}x_i = \sum_{j\neq i}(-F_{i\to j}(x_1,\ldots,x_n,t) + F_{j\to j}(x_1,\ldots,x_n,t)) + I_{0\to i}(x_1,\ldots,x_n,t) - F_{i\to 0} \quad \forall i \in \{1,\ldots n\} \tag{1}$$

Assuming continuity of the fluxes with respect to their source pool $F \in \mathcal{C}^1$ we can write them in product form. $F_{i\to j} = r_{j,i}x_i$ for $i \in \{1,\dots n\}, j \in \{0,1,\dots,n\}$ and $j \neq i$

$$\frac{d}{dt}x_i = -\underbrace{\left(r_{i\to} + \sum_{j\neq i} r_{j,i}(x_1,\dots,x_n,t)\right)}_{=m_{i,i}(x_1,\dots,x_n,t)} x_i + \sum_{j\neq i} \underbrace{r_{i,j}(x_1,\dots,x_n,t)}_{-m_{i,j}(x_1,\dots,x_n,t)} x_j + \underbrace{F_{\to i}(x_1,\dots,x_n,t)}_{I_{\to i}} \tag{2}$$

$$= -\sum_j m_{i,j}(x_1,\dots,x_n,t)x_j + I_{\to i}(x_1,\dots,x_n,t) \tag{3}$$

Writing $\mathbf{X} = (x_1,\dots x_n)^T$ for the ordered tuple of all pool contents, and $\mathbf{I} = (I_{\to 1},\dots I_{\to n})^T$ for the ordered tuple of all influxes, we get

$$\frac{d}{dt}\mathbf{X} = \mathbf{I}(\mathbf{X},t) - M(\mathbf{X},t)\mathbf{X} \tag{4}$$

$-M$ is called the Compartmental Matrix. [1]

2.0.2 Matrix decomposition

Together with a start-value $\mathbf{X}_0$ (4) constitutes an "initial value problem" (ivp) which can be solved numerically by moving step by step forward in time.

Without further assumptions the system is "nonautonomous" (since either of $\mathbf{I}$ or M can depend on time t) and "nonlinear" since either M can depend on X or $\mathbf{I}$ can depends on X in a way that cannot be expressed in the form $\mathbf{I}(\mathbf{X},t) = \tilde{\mathbf{I}(\mathbf{t})} + I_{mat}(t)\mathbf{X}$ with the matrix $I_{mat}(t)$ independent of X.

If $m_{i,i}(\mathbf{X},t) \neq 0$ [2] it is possible to factorize $M(X,t)$ into a product $M = A(\mathbf{X},t)K(\mathbf{X},t)$ where K is a diagonal matrix and the matrix A has only ones on it's main diagonal.

If $u = \sum_{k=1\dots n} \mathbf{I}_k \neq 0$ it is possible to determine the dimensionless vector $\boldsymbol{\beta} = \mathbf{I}/u$ where $\sum_{k=1\dots n} \beta_k = 1$ and write $\mathbf{I}(\mathbf{X},t) = \boldsymbol{\beta}(\mathbf{X},t)u(\mathbf{X},t)$ Using these terms we arrive at

$$\frac{d\mathbf{X}}{dt} = B(\mathbf{X},t)u(\mathbf{X},t) - A(\mathbf{X},t)K(\mathbf{X},t)\mathbf{X}$$

with:

$$k_{i,i}(x_1,\dots,x_n,t) = \left(r_{i\to}(x_1,\dots,x_n,t) + \sum_{l\neq i} r_{l,i}(x_1,\dots,x_n,t)\right)$$

$$a_{j,i}(x_1,\dots,x_n,t) = \frac{r_{j,i}(x_1,\dots,x_n,t)}{k_{i,i}(x_1,\dots,x_n,t)} = \begin{cases} = \frac{r_{i,j}(x_1,\dots,x_n,t)}{\left(r_{i\to}(x_1,\dots,x_n,t) + \sum_{l\neq i} r_{l,i}(x_1,\dots,x_n,t)\right)} \text{ for } j \neq i \\ 1 \text{ for } j = i \end{cases} \tag{5}$$

[1] Because the enumeration of the set of pools is arbitrary there are, for a model with n pools actually $n!$ such matrix equations, that all describe the same model.

[2] If $m_{i,i}(\mathbf{X},t) = 0$ for some i then some elements of A become undefined. However, this does not mean that we could not write M as a product, just that A cannot be inferred and we have to pretend to know the $a_{j,i}(x_1,\dots,x_n,t)$ for $j \neq i$ and $\forall(x_1,\dots,x_n,t)$ with $k_{ii}(x_1,\dots,x_n,t) = 0$ although we could never learn them from any observed fluxes. The same arguments holds for $\boldsymbol{\beta}$.

The $k_{i,i}$ can be interpreted as the rate of the total flux out of pool i. The elements of column i of A describe then which fractions of this total outflux is transferred to pool j.

2.0.3 Assumption of Linearity

If we assume the model to be linear and nonautonomous the dependency on X vanishes and we have either

$$\begin{aligned}\frac{d\mathbf{X}}{dt} &= \tilde{\mathbf{I}}(t) + I_{mat}(t)\mathbf{X} - M(t)\mathbf{X} \\ &= \underbrace{\tilde{\mathbf{I}}(t) + (I_{mat}(t) - M(t))}_{L(t)}\mathbf{X} \\ &= \tilde{\mathbf{I}}(t) + L(t)\mathbf{X}\end{aligned} \tag{6}$$

or if we insist on a non-state-dependent inputs

$$\begin{aligned}\frac{d\mathbf{X}}{dt} &= \mathbf{I}(t) - M(t)\mathbf{X} \\ &= \boldsymbol{\beta}(t)u(t) - A(t)K(t)\mathbf{X}.\end{aligned} \tag{7}$$

Eq. (**??**) allows for influxes to be dependent on the receiving pool, e.g. for the influx of carbon through photosynthesis to depend on the the size of the leaf pool. Note that L does not have to be compartmental and therefore not factorizable into A and K. Eq. (7) is that Both exclude certain compartmental models e.g. some with interactions between chemical species. Imagine that some of the pools contain Nitrogen and others Carbon. It is likely that some fluxes out of carbon pools are controlled by the available Nitrogen. Imagine a compartmental system where the startvector consist of Carbon and Nitrogen pool contents: $\mathbf{X} = (c_1, c_2, \ldots, n_1, n_2, \ldots)^T$, then a flux between carbon pools a and b that depends of the content of nitrogen pool c depends on (a part of) the statevector, which makes it nonlinear.

$$\begin{aligned}F_{a\to b}(\mathbf{X}, t) &= r_{c_i\to *}(n_c, t)\mathbf{X}_a \\ &= r_{c_i\to *}(\mathbf{X}, t)\mathbf{X}_a\end{aligned}$$

2.0.4 Assumption of Factorizability, substrate centered versus flux centered description

For many published models the nonautonomous part can be further localized into a diagonal matrix $\xi(t)$ so that we can achieve constant A and K. It is important to realize two points here:

1. This is not possible for all compartmental matrices.
2. In the cases where it is possible it does not uniquely define ξ.

We can discuss (1) from a mathematical and a modeling viewpoint: The linear version of (8) is:

$$k_{i,i}(t) = \left(r_{i\to}(t) + \sum_{l\neq i} r_{l,i}(t) \right)$$

$$a_{j,i}(t) = \begin{cases} \frac{r_{j,i}(t)}{\left(r_{i\to}(t)+\sum_{l\neq i} r_{l,i}(t)\right)} \text{ for } j \neq i \\ 1 \text{ for } j = i \end{cases} \tag{8}$$

From this representation it is clear that the $a_{i,j}$ are only constant if all rates $r_{j,i}$ for $j \in \{0, \ldots, n\}$ contain the *same* time dependent factor $\xi(t)$, which makes the existence of constant A and K an *assumption*. From a modeling point of view the $\xi_{i,i}$ can be seen as a "substrate" dependent rate modifier since it affects everything that leaves the same pool in the same way, whereas $r_{i,*}(t)$ is specific to a single flux an so could be different for different "processes" even if they use the same substrate.

In order to discuss (2) we note that the assumption that we can write $M = A\xi K$ implies that we can also write it as $M = A\tilde{\xi}\tilde{K}$ where $\tilde{\xi} = d\xi$, $\tilde{K} = d^{-1}K$ for any diagonal matrix d. This implies that without further assumptions it is not possible to compute ξ for a given model without a gauge condition like $\xi(T_0, W_0) = 1$ for a some specific temperature T_0 and moisture W_0, which in turn implies that the *baseline residence time* $(AK)^{-1}$ is only defined up to the above mentioned diagonal matrix d. This fact becomes very important when certain properties of models are attributed to either ξ or the *baseline residence time*. Any sensible attribution of this kind has to be shown to be robust to changes of d. A1

2.0.5 Assumption of factorizability of ξ

In some cases we can resolve ξ further.

$$\frac{dX}{dt} = B(t)u(t) - A\xi_{temp}(t)\xi_{mois}(t)KX$$

2.1 Definition of diagnostic variables

2.1.1 Storage capacity X_c and storage potential X_p

We define it using linearity but not factorizability since it is not used. If we solve (4) the solution $\mathbf{X}(t)$ still fulfills the equation

$$X(t) = M^{-1}(t)\left(\mathbf{I}(t) - \frac{d\mathbf{X}}{dt}(t)\right) \tag{9}$$

$$= \underbrace{M^{-1}(t)\mathbf{I}(t)}_{\mathbf{X}_c(t)} - \underbrace{M^{-1}(t)\frac{d\mathbf{X}}{dt}(t)}_{\mathbf{X}_p i(t)} \tag{10}$$

$$= \mathbf{X}_c(t) - \mathbf{X}_p(t) \tag{11}$$

Note: The original publication (Luo et al., 2017, p. 150 after eq.8b) claims that $\mathbf{X}_c$ is an "attractor" and attempts to justify this claim by some geometrical arguments. disproved in the appendix We will show :

1. That $\mathbf{X}_c(t)$ is an attractor (fixed point) not for the solution $\mathbf{X}(t)$ of the original system but for the solution of a related *autonomous* system and how this affects the interpretation of the results.

2. How $\mathbf{X}_c$ and $\mathbf{X}$ are *really* related geometrically.

Consider the autonomous linear system

$$\frac{d}{d\,t}\tilde{\mathbf{X}}_{t_f} = I_{t_f} - M_{t_f}\tilde{\mathbf{X}}_{t_f} \tag{12}$$

with constant M_{t_f} and $\mathbf{I}_{t_f}$ obtained by *freezing* $M(t_f)$ and $\mathbf{I}(t_f)$ at the same time when we compute $\mathbf{X}_{ct_f} = M_{t_f}^{-1}\mathbf{I}_{t_f}$ (assuming generously that M_{t_f} is invertable) If we plug $\mathbf{X}_{ct_f}$ into (12) we get

$$\begin{aligned}\frac{d}{d\,t}\tilde{\mathbf{X}}_{t_f} &= I_{t_f} - M_{t_f}\mathbf{X}_{ct_f}\\ &= 0\end{aligned}$$

So $\mathbf{X}_{ct_f}$ is just the fixpoint of this system. Since it is unique and it's basin of attraction is the whole positive orthant the solution $\tilde{\mathbf{X}}_{t_f}(t)$ of the ivp $\tilde{\mathbf{X}}_{t_f}(t=0) = X(t_f)$ will start moving towards it, but also immediately separate from $\mathbf{X}(t)$. t_f is the last guaranteed time where $\mathbf{X} = \tilde{\mathbf{X}}$. While $\tilde{\mathbf{X}}(t)$ starts "chasing" X_{ct_f} the real solution is affected by new $M(t > t_f)$ and $I(t > t_f)$ and has no obligation to do so. The best prediction of the ability to store carbon at time $t > t_f$ is the solution $\mathbf{X}(t)$. In this sense $\mathbf{X}_{ct_f}$ is a *model within the model* that describes the (less accurate) forecasting under inputs and rates frozen at time t_f.0

Note:

The attractivity of $\mathbf{X}_c$ for $\tilde{\mathbf{X}}$ might partially explain why $\mathbf{X}(t)$ sometimes *looks* as if it was "chasing" $\mathbf{X}_c(t)$. It is certainly *possible* that the changes in $M(t)$ and $\mathbf{I}(t)$ are either too small to hamper the 'pursuit' of $\mathbf{X}$ significantly or occasionally temporarily even help to close the gap. This is just not consistently the case, but possibly even frequently in observed systems.

This also changes the interpretation of $\mathbf{X}_p$ which is, if we start at at time t from the point $\mathbf{X}(t)$, the storage potential of the *autonomous* system in other words the difference to the steady state pool values if the matrix $M(t)$ and inputs $\mathbf{I}(t)$ *would* stop to depend on time.

2.1.2 Residence Time RT

The influx $\mathbf{I}$ can always be written as $\mathbf{I} = \boldsymbol{\beta} u$. Assuming that the pool contents (the components of $\mathbf{X}$) have dimension $mass$ we can infer from (4)) that the components of M have dimension $\frac{1}{time}$. The components of the (inverse) matrix M^{-1} have therefore dimension $time$. Accordingly the product $\mathbf{RT} = M^{-1}\boldsymbol{\beta}$ is a vector of the same shape as $\mathbf{X}$ whose components have dimension $time$. In the context of the Traceability Framework $\mathbf{RT}$ is called *residence time*. but a previous paper (Rasmussen et al., 2016, Proposition 1) contains the proof that this formula describes the mean transit time if the system is linear and *autonomous* and in *equilibrium*. Consequently $\mathbf{RT}$is the residence time of system **??**

Note that:

1. Residence time is neither the time of residence of a *single* particle (Carbon atom) in the system nor for the *mean* over all particles for the following reasons:

(a) In well mixed systems particles can reside in a pool for different times from zero to infinity. What is common to all particles in a certain pool are the probabilities to leave this pool in the next unit of time towards different destinations. In a nonautonomous system this probability can change with time but still does not distinguish between particles.

(b) One could compute the mean of these times over all particles exiting a pool, but even then the result is in general not equal to the above mentioned **RT** Rasmussen et al. (2016); Metzler et al. (2018). The mean residence time would only coincide with the definition above if the system was autonomous. This is related to the interpretation of $\mathbf{X}_c$ as attractor for a related *autonomous* system. It turns out that **RT** is the mean time of residence for this related system.

2. The origin of the term is probably most easily understood as the generalization of a one dimensional rate equation

$$\frac{d}{dt}x = u - mx \tag{13}$$

where r and u are constant. In equilibrium the righthandside of (13) is 0 which allows us to compute the equilibrium content as $x^* = \frac{u}{m}$. The time to replace the same amount of mass (which does not mean that all the particles are replaced) as is stored in the system is therefore $\frac{x^*}{u} = \frac{1}{m}$ This time is called the "turnover time" but *in equilibrium* is equal to the mean residence time $rt = m^{-1}$. If we start with the rate as property of the model the *residence time* can be defined as the inverse of this rate. The above definition is the generalization of this simple relationship to matrices and vectors. The matrix M^{-1} takes the role of the number $\frac{1}{m}$. In the context of the *Transient Traceability Analysis* the matrix M^{-1} is called *Chasing Time.*

Definition 2.1 (Attractor)**.** might go to the appendix Let t represent time and let $f(t, *)$ be a function which specifies the dynamics of the system. That is, if a is a point in an n-dimensional phase space, representing the initial state of the system, then $f(0, a) = a$ and, for a positive value of t, $f(t, a)$ is the result of the evolution of this state after time t. An attractor is a subset A of the phase space characterized by the following three conditions:

1. A is forward invariant under f: if a is an element of A then so is $f(t, a)$, for all $t > 0$.
2. There exists a neighborhood of A, called the basin of attraction for A and denoted $B(A)$, which consists of all points b that "enter A in the limit $t \to \infty$. More formally, $B(A)$ is the set of all points b in the phase space with the following property: For any open neighborhood N of A, there is a positive constant T such that $f(t, b) \in N$ for all real t > T.
3. There is no proper (non-empty) subset of A having the first two properties.

Examples are fixed points, limit cycles or more extravagantly shaped objects like the famous "strange" Lorenz attractor.

2.2 Attractors for the solution X

Since the original work explicitly discusses a *non*-autonomous system the definition of attractor has to be extended. Due to the time dependence of M and $\mathbf{I}$ attractors are now functions of time.

Definition 2.2 (Pullback Attractor)**.** might go to the appendix, look up in Martins book

In fact it has been shown in (Rasmussen et al., 2016) that (under some additional conditions on M) solutions of the linear version of (4) are exponentially stable.

2.2.1 Forward attractors

A consequence of this fact is that every solution regardless of the starting point $\mathbf{X}_0$ is a forward attractor to any other solution starting from a different position. In essence the effect of the starting point on the solution at a later time decreases very quickly (exponentially) with the time difference. This is a remarkable result but does not apply to $\mathbf{X}_c$ which is NOT a solution of the original system.

To show that it is possible (in fact very likely) that $\mathbf{X}_c(t)$ and $\mathbf{X}(t)$ do not necessarily get closer to each other (contrary to the solutions) we construct a counter example where both $\mathbf{X}(t)$ and $\mathbf{X}_c(t)$ are periodic functions in time and so is there difference. Consider the following periodic one pool system.

$$\frac{dx}{dt} = u(t) - kx$$

with $u(t) = 1 + sin(t)$ and $k = 1$. The solution for $x_0 = x(t_0) = \frac{1}{2}$ is given by

$$\begin{aligned}
x(t) &= \frac{1}{2}e^{-t} + e^{-t}\int_0^t (sin(\tau)+1)\frac{1}{e^{-\tau}}\,d\tau \\
&= \frac{1}{2}e^{-t} + e^{-t}\int (sin(\tau)+1)e^{\tau}\,d\tau \\
&= \frac{1}{2}e^{-t} + e^{-t}\frac{1}{2}(e^t\sin(t) - e^t\cos(t) + 2e^t - 1) \\
&= \frac{1}{2}e^{-t} + \frac{1}{2}(\sin(t) - \cos(t) + 2 - e^{-t}) \\
&= \frac{1}{2}(\sin(t) - \cos(t) + 2)
\end{aligned}$$

The carbon storage capacity x_c for this system is given by $x_c(t) = \sin(t) + 1$ It is obvious that $x(t) - x_c(t) = -\frac{1}{2}(\sin t + cost)$ is a periodic function that does not disappear for large t.. A graph would show that the "chasing never ends even after catching up"

2.2.2 Pullback attractor

In the same paper the unique pullback attractor for the system is given. If we adapt to our nomenclature it reads:

$$\mathbf{X}_{pullback}(t) = \int_{-\infty}^{t} \Phi(t,\tau)\mathbf{I}(\tau)d\,\tau \tag{14}$$

where the *State Transition Operator* $\Phi(t,\tau)$ is the solution of the matrix ODE $\frac{d}{dt}\Phi = M(t)$ and describes how much of the material that entered at exactly time τ is still present. We note that although the start value $\mathbf{X}_0$ does not appear in (14) the integral shows that the *whole history* of $I(\tau)$ and $\Phi(\tau,t)$ (and therefore $M(t)$) from $-\infty$ to the present time t contributes.

Contrarily the definition of $\mathbf{X}_c = M^{-1}(t)I(t)$ does only refer to *one* point in time and does not capture the history of either input or rates. This does not only hold for the pullback attractor but is perhaps also the most poignant way to describe the difference between $\mathbf{X}$ and $\mathbf{X}_c$. With the help of the state transition operator we can also express the solution $\mathbf{X}$ as:

$$\mathbf{X}(t) = \Phi(t, t_0)\mathbf{X}_0 + \int_{t_0}^{t} \Phi(t, \tau)\mathbf{I}(\tau) d\,\tau \tag{15}$$

Again we see that any $\mathbf{X}(t)$ is influenced by the whole history of $M(t)$ and $I(t)$ from t_0 till t whereas X_c is a momentary value.

2.3 Temporal means of $\mathbf{X_c}$ and their impact on the size of the error $\mathbf{X_p}$

One consequence of this is that $\mathbf{X}_c$ can change with time much more abruptly than $\mathbf{X}$. The original work (Luo et al., 2017, p. 152 Fig. 5) describes this effect. In light of the above integral representation (15) it is clear that $\mathbf{X}$ can be affected by only infinitesimal small increments in the interval between t and $t + d\,t$ whereas $\mathbf{X}_c$ can change dramatically since it not even necessarily continuous let alone differentiable. (if $\mathbf{I}(t)$ or $M(t)$ are discontinuous).

This shows that the X_p can be arbitrarily large and we can easily construct such cases. In applications of the transient traceability framework $\mathbf{X}_c$ is often replaced by a temporal mean $\bar{\mathbf{X}_\mathbf{c}} = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \mathbf{X}_c(t) d\,t$ which has a smoothing effect similar but not equivalent to (15). The size of the error $\bar{X_p}$ depends on the new hyper-parameter $\Delta t = t_2 - t_1$. This has to be taken into account.

3 Counter Examples

3.1 Geometric observations about the relationship between $\mathbf{X}$ and $\mathbf{X_c}$

3.1.1 surrogate system for the overall mass

We are aiming at something like this:

$$\dot{x} = u(t) - m(t)x$$

where x is the aggregated mass over all pools. The question is how to specify $m(t)$ to insure this.

We start with the special case of a linear but nonautonomous system (which is the subject of (Luo et al., 2017)):

$$\frac{d\mathbf{X}}{dt} = \mathbf{I}(t) - M(t)\mathbf{X}$$

Taking the sum over all pools yields.

$$\sum_{p \in pools} \left(\frac{d\mathbf{X}}{dt}\right)_p = \left(\mathbf{I}(t) - M(t)\mathbf{X}\right)_p$$

With:

$$u = \sum_{p \in pools} (\mathbf{I})_p,$$

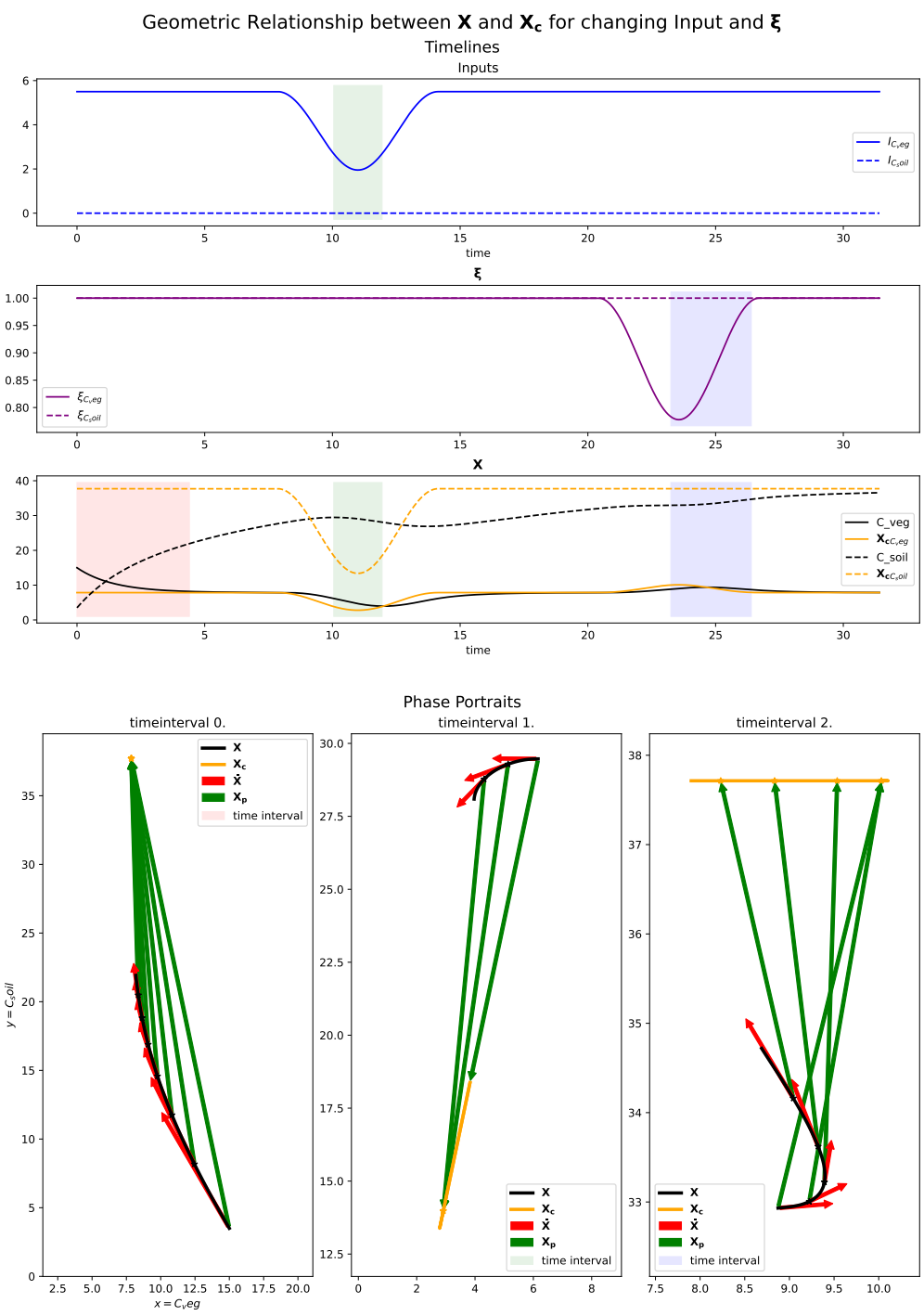
Geometric Relationship between X and X_c for changing Input and ξ
Timelines
Inputs
time
ξ
X
C_veg
C_soil
time
Phase Portraits
timeinterval 0.
timeinterval 1.
timeinterval 2.
X
X_c
Ẋ
X_p
time interval
x = C_veg
y = C_soil

Figure 1.

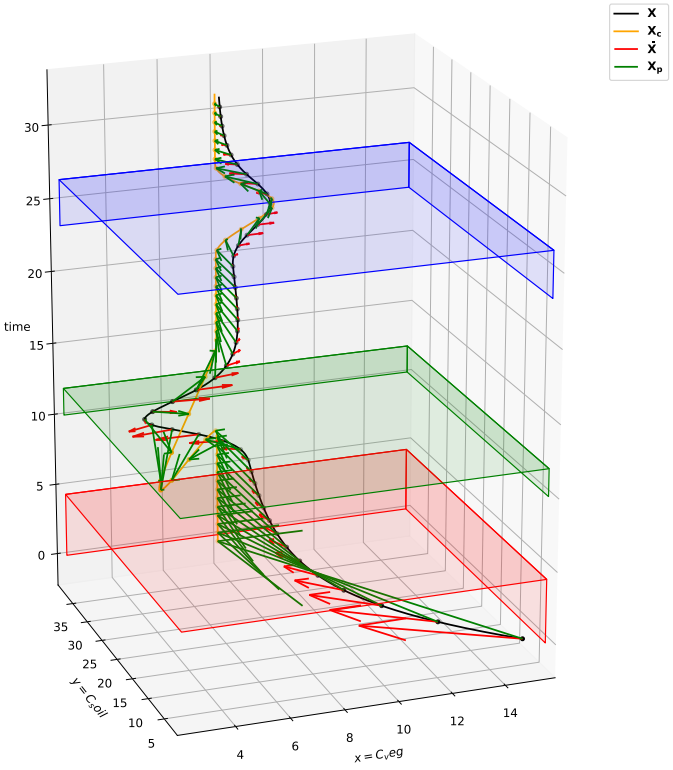
x
x_c
$\dot{x}$
x_p
time
30
25
20
15
10
5
0
35
30
25
20
15
10
5
$y = C_{cell}$
4
6
8
10
12
14
$x = C_{veg}$

Figure 2.

$$x = \sum_{p \in pools} (\mathbf{X})_p \text{ and}$$

$$\sum_{p \in pools} \left(\frac{d\mathbf{X}}{dt} \right)_p = \frac{d}{dt} \sum_{p \in pools} (\mathbf{X})_p = \frac{d}{dt} x$$

We can now try to construct our new system for the combined mass x, in particular we want to find a function for the time dependent rate $m(t)$ such that.

$$\dot{x} = u(t) - m(t)x = \sum_{p \in pools} (\mathbf{I}(t) - M(t)\mathbf{X})_p = u(t) - \sum_{p \in pools} (M(t)\mathbf{X})_p$$

This yields:

$$m(t) = \frac{\sum_{p \in pools} (M(t)\mathbf{X})_p}{\sum_{p \in pools} (\mathbf{X})_p}$$

Notes:

1. m will in general be a function of time $m(t)$ even if the original matrix M is not. This becomes apparent when we consider that we need the vector $\mathbf{X}$ for every t to define $m(t)$, which means that we also must solve the original system at least simultaneously. Assuming that we have done so *before* we we can write a little more obviously.

$$m(t) = \frac{\sum_{p \in pools} (M\mathbf{X}(t))_p}{\sum_{p \in pools} (\mathbf{X}(t))_p}$$

Intuitively this reflects that the original system can have different rates for different pools and so the overall rate depends on how mass is distributed between the pools initially and by the input streams. The fact that the surrogate system is specific to a particular solution of the original system becomes even more apparent if we build surrogate systems for a nonlinear system.

2. Consider a nonlinear systems

$$\frac{d\mathbf{X}}{dt} = \mathbf{I}(\mathbf{X}, t) - M(\mathbf{X}, t)\mathbf{X}$$

Assume that we first solve the system numerically and therefore have $\mathbf{X}(t)$ available. Substituting the solution we get a linear system:

$$\frac{d\mathbf{X}}{dt} = \tilde{\mathbf{I}}(t) - \tilde{M}(t)\mathbf{X}$$

with

$$\tilde{\mathbf{I}}(t) = \mathbf{I}(\mathbf{X}(t), t)$$

and

$$\tilde{M}(t) = M(\mathbf{X}(t), t)$$

Which allows us to construct a one pool surrogate system with the same solution.

The connection between the directions of X_c and X can be formalized in the following observations.

Observation 3.1 (x_p and $\dot{x}$ for the surrogate system)**.** With

$$\begin{aligned}\dot{x} &= u(t) - m(t)x \\ x_c &= \frac{1}{m(t)}u(t) \\ x_p &= x_c - x \\ &= \frac{1}{m(t)}\dot{x}\end{aligned}$$

we have:

$$\operatorname{sign}\dot{x} = \operatorname{sign} x_p \tag{16}$$

Which means that the solution x always increases if $x < x_c$ and always decreases if $x > x_c$. This is clear since $m(t) \geq 0 \quad \forall t$ (in compartmental systems only positive fluxes are allowed) and $m \neq 0$ for x_p and $x_c = \frac{1}{m(t)}u$ to be defined.

Observation 3.2 ($\mathbf{X}_p$ and $\dot{\mathbf{X}}$ are not parallel)**.** An attempt to generalize (16) to vectors in the sense that $\mathbf{X}_p$ and $\dot{\mathbf{X}}$ point in the same direction fails. In Fig. 2 the temporal evolution of $\mathbf{X}_c$ and $\mathbf{X}$ for an example two pool system is plotted, which shows that in general $\mathbf{X}_p$ and $\dot{\mathbf{X}}$ are not pointing in the same direction.

$$\frac{d}{dt}\begin{bmatrix} C_{veg} \\ C_{soil} \end{bmatrix} = \begin{bmatrix} \mathrm{I}_{\mathrm{veg}}(t) \\ 0 \end{bmatrix} + \begin{bmatrix} -(r_{veg2out} + r_{veg2soil0})\xi_{veg}(t) & 0 \\ r_{veg2soil0}\xi_{veg}(t) & -r_{soil2out} \end{bmatrix} \begin{bmatrix} C_{veg} \\ C_{soil} \end{bmatrix} \tag{17}$$

$$\begin{bmatrix} \frac{1}{(r_{veg2out}+r_{veg2soil0})\xi_{veg}(t)} & 0 \\ \frac{r_{veg2soil0}}{r_{soil2out}(r_{veg2out}+r_{veg2soil0})} & \frac{1}{r_{soil2out}} \end{bmatrix} \tag{18}$$

$$\begin{bmatrix} \frac{\mathrm{I}_{\mathrm{veg}}(t)}{(r_{veg2out}+r_{veg2soil0})\xi_{veg}(t)} \\ \frac{r_{veg2soil0}\,\mathrm{I}_{\mathrm{veg}}(t)}{r_{soil2out}(r_{veg2out}+r_{veg2soil0})} \end{bmatrix} \tag{19}$$

This is even the case if we only consider autonomous systems. Although $\mathbf{X}_c$ IS the unique stable fixpoint (and therefore an attractor) for $\tilde{\mathbf{X}}$ the solution of the autonomous system, even $\dot{\tilde{\mathbf{X}}}$ does in general NOT point in the direction of $\mathbf{X}_c - \tilde{\mathbf{X}}$. This is a reflection of the internal pool structure as represented in M. Compartmental systems are in general not obliged to move straight through phase space.

Observation 3.3 (Direction of $\mathbf{X}_p$ and $\dot{\mathbf{X}}$)**.** A minimal attempt to generalize (16) is the following: If $M(t)$ is invertable and if all components of the derivative $\dot{\mathbf{X}}$ have the same sign then all components of the $\mathbf{X}_p = \mathbf{X}_c - \mathbf{X}$ have the same sign too. In other words if $\dot{\mathbf{X}}$ occupies the non negative orthant so does $\mathbf{X}_p$ and if $\dot{\mathbf{X}}$ occupies the non positive orphant so does $\mathbf{X}_p$. In both cases the scalar product $\langle \dot{\mathbf{X}}, \mathbf{X}_p \rangle \geq 0$

Proof: Using that $M(t)$ is invertable we can write the steady state solution $\mathbf{X}^*$ (for the system frozen at time t) for a given constant input $\mathbf{I}$ and Matrix M as

$$\mathbf{X}^* = M^{-1}\mathbf{I} \tag{20}$$

We first convince ourselves that the matrix M^{-1} has indeed only positive elements. For a compartmental system the pool contents and inputs are always positive. This is also true for the steady state values of the frozen system. If we choose $\mathbf{I} = (1,0,\ldots,0)$ we get $\mathbf{X}^* = M_{:,1}$ so we know that the first column of M must be positive. We can do this with any column to obtain $M_{i,j} > 0 \quad \forall i,j$. If we assume further that the derivative $\dot{\mathbf{X}}$ has only positive components then the components of $\mathbf{X}_p$ are sums of products of positive values and have to be positive too. In case all the components of $\dot{\mathbf{X}}$ are negative then the components of $\mathbf{X}_p$ are negative by the same argument. A consequence of this statement is that in case $\dot{\mathbf{X}}$ is either completely in the positive or negative orthant the inner product $\langle \dot{\mathbf{X}}, \mathbf{X}_p \rangle \geq 0$ which means that the angle between them is smaller than a right angle.

Unfortunately this argument only applies to the very special cases that all pools increase or decrease simultaneously. In general $\dot{\mathbf{X}}$ as well as $\mathbf{X}_p$ can have positive and negative components but in this case the knowledge of the positive of M^{-1} is not sufficient to confine the difference in their directions.

Observation 3.4 (Direction of $\mathbf{X}_p$ and $\dot{X}$ in special cases)**.** The previous observation leads to the conjecture that perhaps $\langle \dot{\mathbf{X}}, \mathbf{X}_p \rangle \geq 0$ is generally true, even if $M_{i,j} > 0$ is not sufficient to prove it. There are however special cases where this can be ascertained. If we talk about $\mathbf{X}_c$ at all we assume the existence of a stable fixed point for the frozen system. A necessary (and sufficient) condition for this is that the real parts of all the eigenvalues of M are greater than zero. In case all the eigenvalues are real we can express $\mathbf{X}_p$ and $\dot{\mathbf{X}}$ and the linear mapping encoded by M in terms of tbe eigenvectors $e_1, \ldots e_n$ and eigenvalues $\lambda_1, \ldots \lambda_n$.

$$\begin{aligned}
\langle \dot{\mathbf{X}}, \mathbf{X}_p \rangle &= \langle M\mathbf{X}_p, \mathbf{X}_p \rangle && (21)\\
&= \langle \lambda_1 e_1 x_{p_1} + \ldots \lambda_n e_n x_{p_n}, e_1 x_{p_1} + \ldots e_n x_{p_n} \rangle && (22)\\
&= \lambda_1 x_{p_1}^2 \langle e_1, e_1 \rangle + \cdots + \lambda_n x_{p_n}^2 \langle e_n, e_n \rangle && (23)\\
&> 0. && (24)
\end{aligned}$$

A noteworthy application are two dimensional compartmental linear systems with constant M which can not have complex eigenvalues (Since the complex eigenvalues of real matrices always come in conjugate pairs the matrix of a two dimensional system would describe a rotation with shrinking radius (negative real part). For $\mathbf{X}$ on one of the boundaries of the positive orthant this would lead to a derivative pointing outwards and thus to negative pool contents, violating the assumption of a compartmental system). If the system has more than two dimensions complex eigenvalues become possible and the discussion more complicated.

4 Conclusions

TEXT summary still missing

1. X_c is not an attractor for $\mathbf{X}$ neither is x_c for x (overall mass) All derived variables like $\mathbf{RT}$ and τ have to be interpreted in the context of the 'frozen' systems *in equilibrium*.
2. $\bar{X_c}$ as proxy for $\mathbf{X}$ depends on the time interval (examples show minimal error of about 7/100)
3. ξ is not a model property but a property of the literature In attributions it is saver to look at $k\xi$

Code availability. TEXT

Data availability. TEXT

Code and data availability. TEXT

Sample availability. TEXT

Video supplement. TEXT

Appendix A: ξ

A1 Examples where the actual value of ξ matters or does not matter

E.g. consider variance attribution via covariance as described in the supplementary material S(9) of (Zhou et al., 2018). The relative contribution $RVar(ab)_a$ of variable a to the variance of the product ab is defined as

$$RVar(ab)_a = \frac{cov(ln(a), ln(ab)}{Var(ln(ab)} \tag{A1}$$

As an example for an attribution that is independent of the arbitrary d consider the question of temporal variance of rt to it's factors ξ and br via (temporal) covariance of rt with ξ and k since the temporal covariance $cov(ln(k), ln(rt)) = 0$ but it is also obvious that an attribution to the temporal change of ξ leads to the same result as an attribution to the product ξk, so the distinction between contributions of ξ and k is not necessary. An example where the result does depend on the arbitrary

d is given by the attribution of the variance of $\mathbf{RT}$ of model ensemble to the variance of ξ and k via the same method as the following minimal ensemble of two models shows. If we look at the contribution of ξ to the variance

$$var_{\xi_{rel}} = \frac{cov([ln(\xi_1), ln(\xi_2)]^T, [ln(\frac{k_1}{\xi_1}), ln(\frac{k_2}{\xi_2})]^T)}{var([ln(\frac{k_1}{\xi_1}), ln(\frac{k_2}{\xi_2})]^T)} \tag{A2}$$

If we rewrite the second model in the vector with an arbitrary d we get a different result

Author contributions. TEXT

Competing interests. TEXT

Disclaimer. TEXT

Acknowledgements. TEXT

References

Jacquez, J. A. et al.: Compartmental analysis in biology and medicine, 1972.

Luo, Y., Shi, Z., Lu, X., Xia, J., Liang, J., Jiang, J., Wang, Y., Smith, M. J., Jiang, L., Ahlström, A., Chen, B., Hararuk, O., Hastings, A., Hoffman, F., Medlyn, B., Niu, S., Rasmussen, M., Todd-Brown, K., and Wang, Y.-P.: Transient dynamics of terrestrial carbon storage: Mathematical foundation and its applications, Biogeosciences, 14, 145–161, https://doi.org/10.5194/bg-14-145-2017, 2017.

Metzler, H., Müller, M., and Sierra, C.: Transit-time and age distributions for nonlinear time-dependent compartmental systems, Proceedings of the National Academy of Sciences, 115, 201705 296, https://doi.org/10.1073/pnas.1705296115, 2018.

Rasmussen, M., Hastings, A., Smith, M. J., Agusto, F. B., Chen-Charpentier, B. M., Hoffman, F. M., Jiang, J., Todd-Brown, K. E. O., Wang, Y., Wang, Y.-P., and Luo, Y.: Transit times and mean ages for nonautonomous and autonomous compartmental systems, J. Math. Biol., 73, 1379–1398, https://doi.org/10.1007/s00285-016-0990-8, 2016.

Zhou, S., Liang, J., Lu, X., Li, Q., Jiang, L., Zhang, Y., Schwalm, C. R., Fisher, J. B., Tjiputra, J., Sitch, S., et al.: Sources of uncertainty in modeled land carbon storage within and across three MIPs: Diagnosis with three new techniques, Journal of Climate, 31, 2833–2851, 2018.

to do

1. finish the figure
2. ~~introduce the 1D surrogate systems~~
3. ~~Prove the geometric Observations~~
4. Put the definition for attractors in the text
5. To explain the 1 pools system tendency to "chase" X_c we can rephrase it the other way around to: At a given point $t_f$the derivative is determined by the present $m_f = m(t_f)$ and the present input $i_f = i(t_f)$:$\dot{x} = i_f - m_f x_f$ The derivative at this point determines the possible equilibrium of the frozen linear system with the same derivative at this point. If the derivative is positive we have

$$\begin{aligned} 0 &< if - m_f x \\ &\rightarrow m_f x < i_f \\ \rightarrow x &< \frac{1}{m_f} i_f = x_f* = x_c \end{aligned}$$

So the prospective equilibrium of the frozen system is bigger than the current value.