Information content in time series of litter decomposition studies and the transit time of litter in aridlands

Agustín Sarquis[1, 2] and Carlos A. Sierra[3]

[1]Facultad de Agronomía, Universidad de Buenos Aires, Buenos Aires, 1417, Argentina
5 [2]Instituto de Investigaciones Fisiológicas y Ecológicas Vinculadas a la Agricultura (IFEVA; CONICET-FAUBA), Buenos Aires, 1417, Argentina
[3]Max-Planck-Institut für Biogeochemie, Jena, 07745, Germany

Correspondencia to: Carlos A. Sierra (csierra@bgc-jena.mpg.de)

Abstract. Plant litter decomposition stands at the intersection between carbon (C) loss and sequestration in
10 terrestrial ecosystems. Organic matter during this process experiences chemical and physical transformations that affect decomposition rates of distinct components with different transformation fates. However, most decomposition studies only fit one-pool models that consider organic matter in litter as a single homogenous pool and do not incorporate the dynamics of litter transformations and transfers in their framework. To this extent, compartmental dynamical systems are sets of differential equations that can be used to represent both
15 the heterogeneity in decomposition rates of organic matter, and the transformations it can undergo. Further, a metric that can be used to compare models with different structures is the transit time, that is, the mean age of particles when they are released from a compartment system. In this study, we first asked what model structures are more appropriate to represent decomposition from a publicly available database of decomposition studies in aridlands: *aridec*. For this purpose, we fit one- and two-pool decomposition models
20 with parallel and series structures and used model averaging as a multi-model inference approach. We then asked what the potential ranges of the median transit times of C in aridlands are and what are their relationships with environmental and chemical variables. Hence, we calculated transit time for those models and explored patterns in the data with respect to mean annual temperature and precipitation, solar radiation, the Global Aridity Index, and one litter chemistry trait, the initial lignin content. Median transit time was 1.9 years for the
25 one- and two-pool model with parallel structure, and five years for the two-pool series model. The information in our datasets supported all three models in a relatively similar way, thus our decision to use a multi-model inference approach. After model-averaging, median transit time had values of around three years for all datasets. Exploring patterns of transit time in relation to environmental variables yielded weak correlation coefficients, except for mean annual temperature, which was moderate and negative. Overall, our analysis
30 suggests that the information content in litter decomposition studies often holds little information on the heterogeneity of litter and its transformation rate. Nevertheless, the multi-model inference framework proposed here can help to reconcile theoretical expectations with the information content from field studies and can further help to design field experiments that better represent the complexity of the litter decomposition process.

35 **1 Introduction**

Plant litter decomposition is the process through which plant-derived organic matter is broken down into smaller components. The main biotic driver of decomposition is the metabolic activity of fungi and bacteria (Bradford et al., 2017), but soil fauna can be important too (García-Palacios et al., 2013; Zanne et al., 2022). The magnitude of biotic decomposition is further determined by climate (Gholz et al., 2000) and litter quality (Cornwell et al.,
40 2008). Additionally, abiotic drivers of decomposition like solar radiation can have a large contribution to this process (Méndez et al., 2022). Altogether, plant litter decomposition releases carbon that was fixed by plants back to the atmosphere and mediates soil organic carbon (SOC) formation (Cotrufo et al., 2015). This puts decomposition at a crucial intersection between C loss and sequestration in terrestrial ecosystems. It is thus of great interest to gain a better understanding on how decomposition influences the terrestrial carbon balance
45 and how this process will be affected by global change.
Plant litter is composed of material of different physical and chemical properties that decays at different rates (Adair et al., 2008; Tuomi et al., 2009). However, litter decomposition models commonly assume a single pool that considers the decomposition of organic matter as if it was a homogenous mass pool with a single decomposition constant (Adair et al., 2010). Alternatively, organic matter dynamics can be modelled using

50 compartmental dynamical systems, which are sets of differential equations that can be used to represent both the heterogeneity of organic matter chemical quality, and the transformations plant residues can undergo (Sierra and Müller, 2015). This is achieved with the inclusion of different pools that decompose at different rates. This allows to model the dynamics of labile C compounds that are more readily available for microbial consumption like sugars, and other compounds that have a longer persistence in the litter pool like tannins or

55 lignin. Additionally, it is possible to include interactions between these pools like C transfers from one pool to another. This mass transfer between pools represents the transformation of molecules in litter without the actual loss of C from the litter system (Prescott and Vesterdal, 2021). Compartmental models of decomposition have been successfully applied for decades (Parton et al., 1987; Tuomi et al., 2009), and it has been proven many times that they can be an improvement from the traditional one-pool model (Adair et al., 2008; Cornwell and

60 Weedon, 2014; Manzoni et al., 2012).
Despite the richness of information that can be learned from compartmental models, there still exist limitations for their widespread application. One main limitation is parameter identifiability. This happens because more complex models usually have more parameters and, in some cases, the information contained in time series of litter mass loss may not be enough to estimate those parameters unambiguously (Brun et al., 2001). Depending

65 on the resolution and extension of the time series, it might be possible to obtain different number of parameters from the available data (Sarquis et al., 2022a; Sierra et al., 2015). Consequently, different studies developed under different methodologies and sampling schemes may provide information on different model structures. Further, this limits the application of compartmental models to data from extensive heterogenous databases, since not all parameters might be identifiable for all datasets (Sarquis et al., 2022).

70 It is common to compare model parameters like the decomposition constant when the same model has been applied to many datasets. But, comparing the behavior of models with different structures in the same way is not possible, because decomposition constants of single homogenous pools are not comparable to decomposition constants of specific pools, such as those in compartmental models. Thus, a metric that can be used to compare models with different structures is the transit time of mass in a complex heterogeneous system.

75 Transit time represents the mean age of particles when they are released from a system (Sierra et al., 2017). In the context of litter decomposition studies, transit time can tell us about how long it takes C to exit litter since the start of an experiment. Transit time is a random variable with its own probability distribution, and thus mean and median transit times can be calculated (Sierra et al., 2018). Unlike a single decomposition rate, transit time can be calculated for the bulk of litter when using compartmental models. Transit time contains information

80 from all different C compartments (Lu et al., 2018), and so, it becomes a more useful parameter when making comparisons from models that have different structures.
In this study we used the *aridec* database, which is an open access database of published decomposition studies in aridlands from around the world (Sarquis et al., 2022a). The focus of this database on aridlands stems from how widespread aridlands are, since around 41% of the land surface is classified as arid to some extent (Safriel

85 and Adeel, 2005). This large area represents a wide range of diverse ecosystems, with many shared functional characteristics. For instance, aridlands are usually more sparsely vegetated (Guttal and Jayaprakash, 2007) and this produces a shift in the importance of decomposition drivers in comparison to humid ecosystems. Plant litter under these conditions is more susceptible to solar radiation (Austin and Vivanco, 2006) and desiccation by wind (D'Odorico et al., 2019). Further, water sources other than rain can become more relevant when mean annual

90 precipitation is low (Evans et al., 2020). These unique traits of arid ecosystems probably explain why decomposition rates are not correlated to mean annual precipitation in these systems (Austin, 2011), contrary to what was proposed in the traditional literature (e.g., Meentemeyer, 1978). Furthermore, aridland processes are thought to become more widespread in the future because of aridland expansion (Feng and Fu, 2013) and drought-intensification of humid ecosystems (Grünzweig et al., 2022).

95 Hence, we used the *aridec* database to address the following questions: given the information content in time series of litter decomposition studies, what model structures are more appropriate to represent decomposition from arid ecosystems? From the set of obtained models, what are the potential ranges of the median transit times of C? Moreover, what are the potential relationships between median transit time and environmental variables? We fit one- and two-pool decomposition models with parallel and series structures and used model

100 averaging as a multi-model inference approach. We further calculated transit times for those models and explored patterns in the data in relation to environmental and litter chemical variables.

2 Methods

2.1 Model fitting

First, we used the *aridec* database to fit a group of potential decomposition models. The *aridec* database is a publicly available database of decomposition studies from aridlands across the world (Sarquis et al., 2022a). This database contains bulk litter mass loss data, but it lacks mass loss dynamics of different litter organic matter pools that decompose at different rates (e.g.: soluble carbohydrates, cellulose, lignin). Because of this, we took an inverse-modelling approach that allowed us to estimate the parameters of these unknown pools by fitting the models to mass loss data. This model calibration procedure constitutes a non-linear optimization problem, where the objective is to find parameter values that minimize a measure of badness of fit, like a weighted sum of squared residuals (Soetaert and Petzoldt, 2010). Following this procedure, we obtained a set of parameters for each dataset and fit the dynamics of mass loss for different pools. We did this with the SoilR (Sierra et al., 2012) and the FME (Soetaert and Petzoldt, 2010) packages in R (R Core Team, 2020).

SoilR is a modelling framework that contains a wide set of functions and tools to model soil organic matter decomposition within the R computing platform. Organic matter decomposition in SoilR is represented by systems of linear differential equations that generalize most compartment-based models. A simple general structure to represent litter decay with no inputs follows Equation 1:

$$\frac{dC(t)}{dt} = A\,C(t) \tag{1}$$

$$C(t) = [C_{pool\ 1},\ ...,\ C_{pool\ m}]^{\mathrm{T}}$$

$$A = \begin{bmatrix} -k_1 & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{j1} & \cdots & -k_m \end{bmatrix}$$

Where *C(t)* is a *m x 1* vector with *m* pools of litter mass observed at time *t*, and *A* is a square *m x m* matrix that contains decomposition rates (k_m) for each pool and transfer rates (a_{ij}) between them. These different pools may correspond to different ways in which the quality of the litter is expressed in different studies. For example, they may correspond to different compounds obtained from a specific extraction method (e.g.: water soluble sugars, or acid detergent lignin), or they can be defined by general decay classes such as fast and slow decay compounds. The linear dynamical system represented by Eq. (1), has many different solutions, but we are only interested in the solution that satisfies

$$C\,(t = 0) = C_0 = [total\ C_0 \cdot p_1,\ ...,\ total\ C_0 \cdot p_m]^T \tag{2}$$

where C_0 is a *m×1* vector with the value of initial litter mass content in the different compartments *m*. We set total initial C_0 to be 100% for this analysis and the resulting p_m parameters are the initial proportions of litter in *m* pools.

Before fitting the models, we run a collinearity test following the procedure by Soetaert and Petdzolt (2010) and the results are presented in Sarquis et al. (2022). Briefly, when parameters are functionally related, changes in one of them can be compensated by changes in others. This produces different parameter sets that have similar probability distributions, thus it is impossible to determine a single parameter set for a model (Brun et al., 2001; Sierra et al., 2015). From this analysis, we were able to choose three models: a one-pool model, a two-pool parallel model, and a two-pool series model (Fig. 1). The one-pool model represents mass loss data as a single homogeneous mass compartment and has a single parameter, the decomposition rate *k*. The two-pool model with parallel structure considers litter mass as two distinct compartments that decompose at different rates. Hence, its parameters are the two decompositions rates (k_1 and k_2) and the initial proportion of litter mass in pool one (p_1, from which the proportion of mass in pool 2 can be calculated as $p_2 = 1 - p_1$). Finally, the two-pool series model is similar to the parallel model, but it incorporates the transfer of matter from pool one to pool two after its transformation. This is indicated in the model by the parameter a_{12} (i.e., the transfer rate from pool 1 to pool 2).
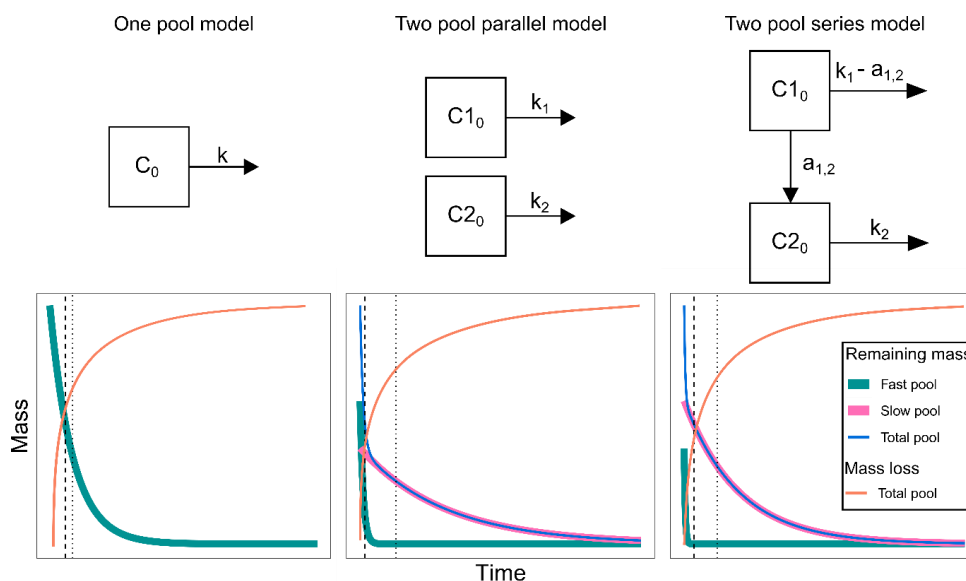
Figure 1: Decomposition models fitted in this study. C_0: total initial carbon content in litter samples; $C1_0$: initial carbon content of the fast-decomposing pool; $C2_0$: initial carbon content of the slow-decomposing pool; k, k_1, k_2: decomposition rates of the total carbon pool, the fast- and the slow-decomposing carbon pools, respectively; $a_{1,2}$: carbon transfer coefficient from the fast-decomposing pool to the slow-decomposing pool.

Specifically for the two-pool series model our collinearity analysis showed that only 20.1% of the datasets produced identifiable results for this model, and only so when we restricted parameter p_1. Restricting or fixing parameters to known values is a way of avoiding collinearity issues. For this purpose, we decided to use initial litter lignin content as a proxy for the p_2 parameter (the initial proportion of mass in pool 2) which is complementary to p_1 ($p_1 + p_2 = 1$). We were limited by the number of datasets that provided initial lignin values in *aridec*. We completed only three of these datasets with information from the TRY database (Kattge et al., 2020). We then completed some of the missing values by averaging lignin data of the same litter types that were already present in *aridec*. Having all the data ready, we proceeded to fit the models mentioned above. All time variables were transformed to monthly timescales to achieve more consistent comparisons.

2.2 Transit time

For each model, we calculated litter mass transit time (Sierra et al., 2017). This concept represents the mean age of the particles when they are released from the bulk litter. Another way to interpret this is the time it takes particles to transit the litter system since the beginning of the experiment. We used a modified version of the Mean Transit Time (MTT) from Sierra et al. (2017) without new litter inputs:

$$MTT = -(1, \ldots, 1) \; A^{-1} \tag{3}$$

For both two-pool models, we used the function transitTime in the SoilR package. This function calculates the mean and median of the distribution of the transit time as well as other quantiles of the distribution. The transit time median is interpreted as the time it takes half the litter mass in a sample to decompose. As a special case, for the one-pool model the MTT can be simply calculated as:

$$MTT = \frac{1}{k} \tag{4}$$

While the Median TT (mTT) can be calculated as:

$$mTT = \frac{\ln 2}{k} \tag{5}$$

Some of the mTT values from the two-pool series models were extremely high and did not make any biological sense. Because of this, we decided to make a cutoff at a mTT of 14.5 years. This value came from fitting this

175 model to the longest data set in aridec which is 10 years long and corresponds to average data of different species at Central Plains Experimental Range in Adair et al. (2017) (S1). We excluded from this study the data sets that exceeded this median transit time cutoff. Finally, after accounting for collinearity, the availability of initial litter lignin data and the mTT cutoff, we were left with 128 data sets from 12 aridec entries.

2.3 Model selection and multi-model inference

180 As a first attempt at model selection, we calculated the Bias Corrected Akaike Information Criteria, which is used for small sample sizes (AICc; Burnham and Anderson, 2002). We used the formula from Shumway and Stoffer (2017):

$$AICc = \log \sigma 2k + \frac{n+k}{n-k-2} \qquad (6)$$

where $\sigma 2k$ is the variance of the model (in this case the mean squared residuals, i.e. sum of squared residuals divided by sample size, MSR hereafter), k is the number of parameters in the model, and n is the sample size or
185 the number of points in each time series. We accounted for the variance as one of the parameters in the formula as Burnham and Anderson (2002) recommend.

A common way of choosing the model with the best fit is by looking at the model with the lowest AIC value. We did this using the *akaike.weights* function from the *qpcR* package. Additionally, we calculated the difference in AICc between the model with the lowest AICc and the other two candidate models (ΔAICc). Since we did not
190 have enough information to choose a single model structure based on AICc (see Results section), we decided to follow a multi-model inference approach (Burnham and Anderson, 2002). We first calculated Akaike Weights using the function *Weights* from the MuMin R package for each model. Akaike Weights can be interpreted as the probability that a model *j* is the best of all *i* candidate models given the data (Lukacs et al., 2010), and are calculated as:

195 $$w_j = \frac{\exp\left(-\frac{1}{2}\Delta AICc_j\right)}{\sum_{i=1} \exp\left(-\frac{1}{2}\Delta AICc_i\right)} \qquad (7)$$

We then calculated new average estimators for the mean and the median transit times as:

$$\hat{\bar{\beta}}_i = \sum_{j=1} w_j \hat{\beta}_{ij} \qquad (8)$$

where $\hat{\beta}_{ij}$ is the *i* parameter estimator $\hat{\beta}$ for each *j* model. This results in estimators of mean (*avgMTT*) and median (*avgmTT*) transit times averaged across models for each database entry.
200 We calculated as well the unconditional variance for each averaged estimator (Burnham and Anderson, 2002; Lukacs et al., 2010) as:

$$v\hat{a}r\left[\hat{\bar{\beta}}_i\right] = \sum_{j=1} w_j \left[MSR_j + \left(\hat{\beta}_{ij} - \hat{\bar{\beta}}_i\right)^2\right] \qquad (9)$$

Finally, we estimated 95% confidence intervals as:

$$\hat{\bar{\beta}}_i \pm cv \sqrt{v\hat{a}r\left[\hat{\bar{\beta}}_i\right]} \qquad (10)$$

205 where *cv* stands for the critical value of a t-distribution for a particular number of degrees of freedom.

We plotted data against environmental variables to explore potential relationships with avgmTT and calculated *r* correlation coefficients. We used data already available in *aridec* like mean annual temperature, mean annual precipitation, and one litter quality variable like initial lignin content. We additionally used Global Aridity Index as calculated in Sarquis et al. (2022) for aridec entries and annual solar radiation from the TerraClimate database
210 (Abatzoglou et al., 2018).

Further, to test whether the data fit an exponential distribution, we calculated the ratio between *avgmTT* and *ln2 * avgMTT*. In an exponential distribution, the median equals *ln2* times the mean. So, if the ratio between the median from our models (avgmTT) and the median calculated as *ln2 * avgMTT* equals 1, that would imply that both medians are equal, and the model follows an exponential distribution. All calculations, modelling and
215 figures were made using R (R Core Team, 2020).

3 Results

We fit three different candidate models for 128 time series of decomposition, which totaled 384 models (see table S2 for model fit). The information in our datasets supported all three models in a similar way. Most times the one-pool model had the lowest AICc values, but close to one-third of the times the two-pool series model had the best fit according to AICc (Fig. 2). Our ΔAICc values were very low (ΔAICc of the 3rd quartile: 1.515), which reinforced the idea that the information available was not enough to choose a single model with the best fit. Following this, we decided for a multi-model inference approach using model averaging, which left us with 128 individual models (see table S3 for model variance and confidence intervals).
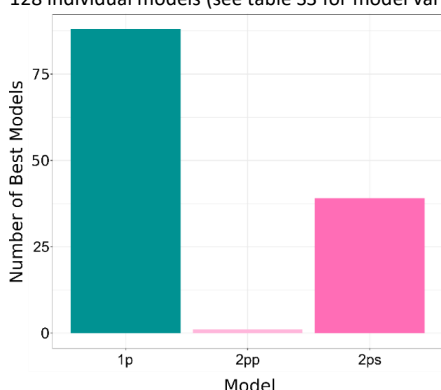

Figure 2: Number of models with the lowest AICc values for each model structure. 1p: one-pool model; 2pp: two-pool parallel model; 2ps: two-pool series model.

Median transit time of plant litter in arid lands after model averaging was within the range of original models (Fig. 3). In this analysis, we only used data from litter decomposed in ambient conditions (without manipulative treatments). One and two-pool parallel models had similar mTT (23.27 ± 9.28 and 23.04 ± 9.65 months, mean ± standard deviation respectively). The two-pool series model had near three-fold mTT values of 60.21 ± 45.80 months. After model-averaging mTT (i.e.: avgmTT) dropped to 36.15 ± 22.20 months.
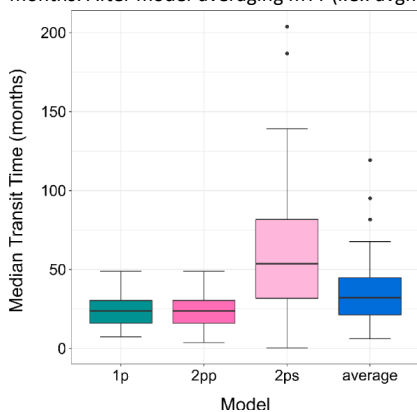

Figure 3: Median transit time (months) for three different models and for the averaged model. Only data for control or ambient treatments were used for this figure. 1p: one-pool model; 2pp: two-pool parallel model; 2ps: two-pool series model.

Looking at the avgmTT alone showed the wide range of time that litter takes to decompose in arid ecosystems (Fig. 4). Exploration of patterns of transit time in relation to environmental variables yielded weak correlation coefficients, except for mean annual temperature which was moderate (r = -0.56). Values of avgmTT at the coldest end ranged between 37 and 65 months, while the warmest site showed values of 8 months (Fig. 4a). This suggests that plant litter in warmer aridlands decomposes faster than in colder sites.

6
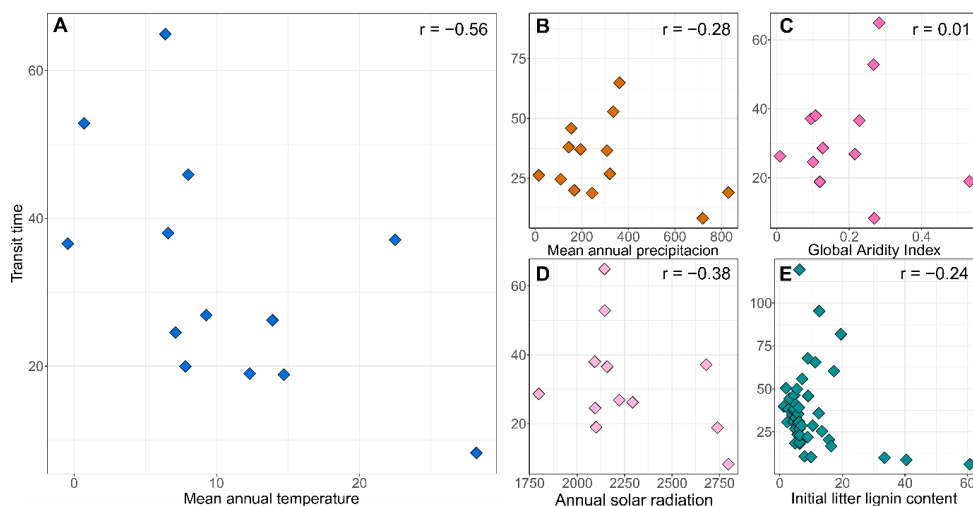
Figure 4: Transit time (months) versus (A) mean annual temperatures (°C), (B) mean annual precipitation (mm), (C) global aridity index, (D) annual solar radiation, and (E) initial litter lignin content (%). Only data for control or ambient

245 **treatments were used for this figure. Each diamond represents the mean avgmTT for different values of a variable at each site. Correlation coefficients (*r*) are displayed.**

Calculating the quotient between the avgmTT and avgMTT times the natural logarithm of two showed contrasting results (Fig. 5). Fourty-two percent of the models in this analysis had values near to zero, which suggests that those models did not follow an exponential distribution. Only 15 % of the models had values

250 between 0.9 and 1.0. Complementarily, this suggests that those models had a near exponential distribution.
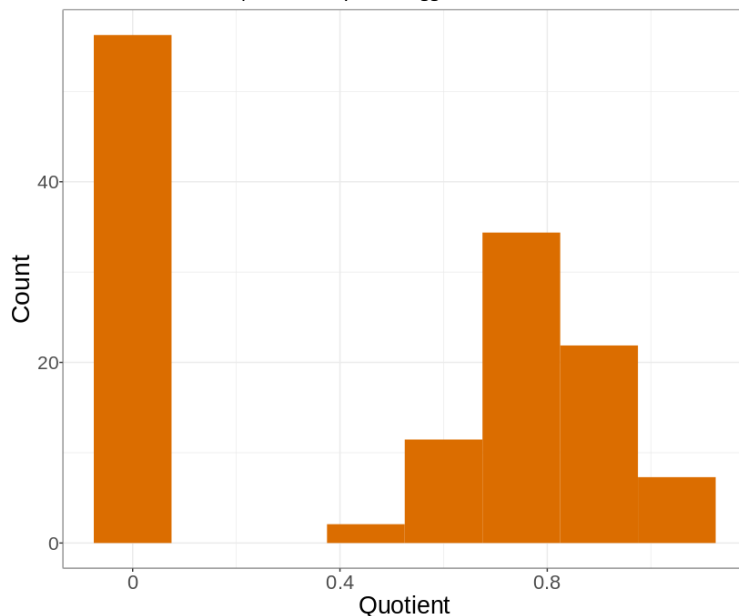

Figure 5: Histogram of frequency for the quotient of the median transit time and the natural logarithm of two times the mean transit time from average models. Bars represent the number of models for a range of values of the quotient.

4 Discussion

255 We asked as our first question: what model structures are more appropriate to represent decomposition in aridlands? After fitting three different models to the data in aridec we found that there was not enough information to choose a unique model judging by their AICc values (Fig 2), so instead we took a multi-model approach (Burnham and Anderson, 2002). This allowed us to incorporate the dynamics of all three models in our results by using AICc weights (Lukacs et al., 2010). In this way, our predictions of transit time in arid lands include

260 the differences in litter chemistry and their effects on decomposition, instead of just considering the bulk of litter as a homogenous C pool. This type of information theoretical approach like model averaging is not novel, but is still underused in ecological studies (Grueber et al., 2011).
       However, before fitting complex compartmental models, researchers should take into consideration the issue of collinearity. In a previous study, we found that most of the time series in the *aridec* database could only be

265 fitted to simpler models with less than three parameters (Sarquis et al., 2022a). This was because the information contained in those time series of litter decomposition was not sufficient to inform more complex models, for example: models with three distinct C pools with transfer coefficients between them. This lack of information in the data caused collinearity between parameters, which in turn made it impossible to identify a single set of parameters for each model (Brun et al., 2001; Sierra et al., 2015). Some of these limitations probably

270 come from the short number of sampling points in most decomposition studies (Sarquis et al., 2022a), which lowers the degrees of freedom available and limits our capacity to model complex organic matter dynamics. The fact that complex models cannot be obtained from the data suggests that we should put more attention into designing field experiments that can better inform about model structures that are more consistent with our current understanding of litter heterogeneity and transformations (Prescott and Vesterdal, 2021).

275 Our second question was: what are the potential ranges of the median transit times of C in litter for aridlands? This part of our study yielded some new insights into the biogeochemistry of arid environments. Median transit time from one- and two-pool decomposition models without interactions were similar and showed that half of the C in litter is lost after almost 2 years in the field (Fig. 3). However, results from the two-pool model with series structure were almost three times higher. This is explained by the C transfer from the fast-decomposing

280 pool to the slow-decomposing pool, which slows down C release from litter. After model-averaging, we obtained intermediate values of median transit times of around three years (Fig. 3).
       Interestingly, this connects back to the problem of model parameter identifiability. Most decomposition studies carried out in aridlands last for only a year (Sarquis et al., 2022a). But our results show that decomposition of litter in arid environments can take on average six times longer until all C exits the system. This means that most

285 field decomposition studies are not capturing the entire dynamics of C release through time. In turn, this has consequences for potential future research because the information that is not contained in data cannot be retrieved by modeling techniques (Brun et al., 2001). If we aim to incorporate field data into complex Earth system models, we need to take into consideration the study length to capture the entire decomposition process. We acknowledge this might seem excessive given academic times go usually faster than litter

290 decomposition in aridlands. However, successful long-term litter decomposition projects exist and can be a potential solution to this issue (e.g.: LIDET; Gholz et al., 2000).
       We asked as our third question: what are the relationships between median transit time and environmental and litter chemical variables? From the set of five variables that we used to explore these relationships, only mean annual temperature showed a moderate correlation with median transit time from average models (Fig. 4a).

295 The importance of temperature as a climatic driver of decomposition is well documented (Zhang and Wang, 2015), both through its positive effects on microbial activity (Sinsabaugh et al., 1991) and its increase of photochemical emissions (Day and Bliss, 2020). Moreover, the correlation with mean annual precipitation was weak (Fig. 4b). This was rather expected since it has been long known that precipitation fails to explain patterns of decomposition rates in aridlands (Austin, 2011).

300 As a final remark, we explored what can transit time teach us about the distribution of decomposition models. After calculating the quotient of the median transit time and the natural logarithm of two times the mean transit time from average models, only 15 % of the models had values close to one (Fig. 5). This is indicative that for most cases models did not follow an exponential distribution since the median of an exponential distribution equals ln2 times the mean. The negative exponential model of decomposition has been the standard for litter

305 and soil organic matter decomposition studies since at least five decades ago (Olson, 1963). This connects back to our first results where the one-pool exponential model was not chosen by our information theoretical approach (Fig. 2). Previous studies found similar results where the negative exponential one-pool model did not rank first for the entirety of the datasets considered (Adair et al., 2008; Cornwell and Weedon, 2014; Manzoni et al., 2012). Testing which distributions fit best the data beforehand is a must and future decomposition studies

310 should test whether the single-pool negative exponential model is actually the best model to fit.

5 Conclusions

315 Although our theoretical understanding of the litter decomposition process is based on the assumption that plant litter is chemically and physically heterogeneous, and undergoes multiple transformations, time series of litter decomposition studies contain only relatively little information on litter heterogeneity and its transformation rates. However, we have shown that a multi-model inference approach helps to reconcile theoretical understanding with information content in observed datasets of litter decomposition. In particular, the combination of AIC model averaging applied to a metric that is independent of model structure, the transit time, provides an inference framework that is useful to understand decomposition dynamics. This framework
320 could help us get a better insight into the chemical transformations of organic matter in litter and soil, and how soil organic matter respond to changes in the environment.

We recognize some limitations for modelling these complex structures arise from field study designs that do not capture the entire decomposition process. This limits the quantity and the quality of the information contained. We recommend that future field decomposition studies incorporate in their designs some strategy to better
325 capture the dynamics of different organic matter pools in litter. This could be done by either measuring the proportion of each compound through time, or by increasing sampling times and study length. The two latter can help gain a better fit and avoid collinearity when using an inverse-modelling approach as in this study. We further encourage researchers to fit models other than the one-pool model, when possible.

6 Code and data availability

330 The aridec database version 1.0.2 is archived and publicly available at https://doi.org/10.5281/zenodo.6600345 (Sarquis et al., 2022b). Result tables and code are stored at https://doi.org/10.5281/zenodo.7561189 (Sarquis and Sierra, 2023).

7 Author contribution

CAS supervised the study. CAS and AS conceptualized the study. AS curated the data and carried out the analysis.
335 AS wrote the original draft of the manuscript. CAS and AS revised and edited further versions of the manuscript.

8 Competing interests

The authors declare that they have no conflict of interest.

9 Acknowledgments

10 Financial support

11 Bibliography

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A. and Hegewisch, K. C.: TerraClimate, a high-resolution global

350 dataset of monthly climate and climatic water balance from 1958–2015, Sci. Data, 5(1), 170191, doi:10.1038/sdata.2017.191, 2018.

Adair, E. C., Parton, W. J., Del Grosso, S. J., Silver, W. L., Harmon, M. E., Hall, S. A., Burke, I. C. and Hart, S. C.: Simple three-pool model accurately describes patterns of long-term litter decomposition in diverse climates, Glob. Chang. Biol., 14, 2636–2660, doi:10.1111/j.1365-2486.2008.01674.x, 2008.

355 Adair, E. C., Hobbie, S. E. and Hobbie, R. K.: Single-pool exponential decomposition models: potential pitfalls in their use in ecological studies, Ecology, 91(4), 1225–1236, doi:10.1890/09-0430.1, 2010.

Adair, E. C., Parton, W. J., King, J. Y., Brandt, L. A. and Lin, Y.: Accounting for photodegradation dramatically improves prediction of carbon losses in dryland systems, Ecosphere, 8(7), e01892, doi:10.1002/ecs2.1892, 2017.

360 Austin, A. T.: Has water limited our imagination for aridland biogeochemistry, Trends Ecol. Evol., 26(5), 229–235, doi:10.1016/j.tree.2011.02.003, 2011.

Austin, A. T. and Vivanco, L.: Plant litter decomposition in a semi-arid ecosystem controlled by photodegradation, Nature, 442(7102), 555–558, doi:10.1038/nature05038, 2006.

Bradford, M. A., Veen, G. F. (Ciska), Bonis, A., Bradford, E. M., Classen, A. T., Cornelissen, J. H. C., Crowther, T.

365 W., De Long, J. R., Freschet, G. T., Kardol, P., Manrubia-Freixa, M., Maynard, D. S., Newman, G. S., Logtestijn, R. S. P., Viketoft, M., Wardle, D. A., Wieder, W. R., Wood, S. A. and Van Der Putten, W. H.: A test of the hierarchical model of litter decomposition, Nat. Ecol. Evol., doi:10.1038/s41559-017-0367-4, 2017.

Brun, R., Reichert, P. and Künsch, H. R.: Practical identifiability analysis of large environmental simulation models, Water Resour. Res., 37(4), 1015–1030, doi:10.1029/2000WR900350, 2001.

370 Burnham, K. P. and Anderson, D. R.: Model Selection and Multimodel Inference, edited by K. P. Burnham and D. R. Anderson, Springer New York, New York, NY., 2002.

Cornwell, W. K. and Weedon, J. T.: Decomposition trajectories of diverse litter types: a model selection analysis, edited by J. Oksanen, Methods Ecol. Evol., 5(2), 173–182, doi:10.1111/2041-210X.12138, 2014.

Cornwell, W. K., Cornelissen, J. H. C., Amatangelo, K., Dorrepaal, E., Eviner, V. T., Godoy, O., Hobbie, S. E.,

375 Hoorens, B., Kurokawa, H., Pérez-Harguindeguy, N., Quested, H. M., Santiago, L. S., Wardle, D. A., Wright, I. J., Aerts, R., Allison, S. D., van Bodegom, P., Brovkin, V., Chatain, A., Callaghan, T. V., Díaz, S., Garnier, E., Gurvich, D. E., Kazakou, E., Klein, J. A., Read, J., Reich, P. B., Soudzilovskaia, N. A., Vaieretti, M. V. and Westoby, M.: Plant species traits are the predominant control on litter decomposition rates within biomes worldwide, Ecol. Lett., 11(10), 1065–1071, doi:10.1111/j.1461-0248.2008.01219.x, 2008.

380 Cotrufo, M. F., Soong, J. L., Horton, A. J., Campbell, E. E., Haddix, M. L., Wall, D. H. and Parton, W. J.: Formation of soil organic matter via biochemical and physical pathways of litter mass loss, Nat. Geosci., 8(10), 776–779, doi:10.1038/ngeo2520, 2015.

D'Odorico, P., Porporato, A. and Runyan, C. W.: Dryland Ecohydrology, edited by P. D'Odorico, A. Porporato, and C. Wilkinson Runyan, Springer International Publishing, Cham., 2019.

385 Day, T. A. and Bliss, M. S.: Solar Photochemical Emission of CO2 From Leaf Litter: Sources and Significance to C Loss, Ecosystems, 23(7), 1344–1361, doi:10.1007/s10021-019-00473-8, 2020.

Evans, S., Todd-Brown, K. E. O., Jacobson, K. and Jacobson, P.: Non-rainfall Moisture: A Key Driver of Microbial Respiration from Standing Litter in Arid, Semiarid, and Mesic Grasslands, Ecosystems, 23(6), 1154–1169, doi:10.1007/s10021-019-00461-y, 2020.

390 Feng, S. and Fu, Q.: Expansion of global drylands under a warming climate, Atmos. Chem. Phys., 13(19), 10081–10094, doi:10.5194/acp-13-10081-2013, 2013.

García-Palacios, P., Maestre, F. T., Kattge, J. and Wall, D. H.: Climate and litter quality differently modulate the effects of soil fauna on litter decomposition across biomes, edited by J. Klironomos, Ecol. Lett., 16(8), 1045–1053, doi:10.1111/ele.12137, 2013.

395 Gholz, H. L., Wedin, D. A., Smitherman, S. M., Harmon, M. E. and Parton, W. J.: Long-term dynamics of pine and hardwood litter in contrasting environments: toward a global model of decomposition, Glob. Chang. Biol., 6(7), 751–765, doi:10.1046/j.1365-2486.2000.00349.x, 2000.

Grueber, C. E., Nakagawa, S., Laws, R. J. and Jamieson, I. G.: Multimodel inference in ecology and evolution: challenges and solutions, J. Evol. Biol., 24(4), 699–711, doi:10.1111/j.1420-9101.2010.02210.x, 2011.

400 Grünzweig, J. M., De Boeck, H. J., Rey, A., Santos, M. J., Adam, O., Bahn, M., Belnap, J., Deckmyn, G., Dekker, S. C., Flores, O., Gliksman, D., Helman, D., Hultine, K. R., Liu, L., Meron, E., Michael, Y., Sheffer, E., Throop, H. L., Tzuk, O. and Yakir, D.: Dryland mechanisms could widely control ecosystem functioning in a drier and warmer world, Nat. Ecol. Evol., 6(8), 1064–1076, doi:10.1038/s41559-022-01779-y, 2022.

Guttal, V. and Jayaprakash, C.: Self-organization and productivity in semi-arid ecosystems: Implications of

405 seasonality in rainfall, J. Theor. Biol., 248(3), 490–500, doi:10.1016/j.jtbi.2007.05.020, 2007.

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D. A., Aakala,

T., Abedi, M., Acosta, A. T. R., Adamidis, G. C., Adamson, K., Aiba, M., Albert, C. H., Alcántara, J. M., Alcázar C, C., Aleixo, I., Ali, H., Amiaud, B., Ammer, C., Amoroso, M. M., Anand, M., Anderson, C., Anten, N., Antos, J., Apgaua, D. M. G., Ashman, T., Asmara, D. H., Asner, G. P., Aspinwall, M., Atkin, O., Aubin, I., Baastrup-Spohr, L.,
410 Bahalkeh, K., Bahn, M., Baker, T., Baker, W. J., Bakker, J. P., Baldocchi, D., Baltzer, J., Banerjee, A., Baranger, A., Barlow, J., Barneche, D. R., Baruch, Z., Bastianelli, D., Battles, J., Bauerle, W., Bauters, M., Bazzato, E., Beckmann, M., Beeckman, H., Beierkuhnlein, C., Bekker, R., Belfry, G., Belluau, M., Beloiu, M., Benavides, R., Benomar, L., Berdugo-Lattke, M. L., Berenguer, E., Bergamin, R., Bergmann, J., Bergmann Carlucci, M., Berner, L., Bernhardt-Römermann, M., Bigler, C., Bjorkman, A. D., Blackman, C., Blanco, C., Blonder, B., Blumenthal, D.,
415 Bocanegra-González, K. T., Boeckx, P., Bohlman, S., Böhning-Gaese, K., Boisvert-Marsh, L., Bond, W., Bond-Lamberty, B., Boom, A., Boonman, C. C. F., Bordin, K., Boughton, E. H., Boukili, V., Bowman, D. M. J. S., Bravo, S., Brendel, M. R., Broadley, M. R., Brown, K. A., Bruelheide, H., Brumnich, F., Bruun, H. H., Bruy, D., Buchanan, S. W., Bucher, S. F., Buchmann, N., Buitenwerf, R., Bunker, D. E., et al.: TRY plant trait database – enhanced coverage and open access, Glob. Chang. Biol., 26(1), 119–188, doi:10.1111/gcb.14904, 2020.
420 Lu, X., Wang, Y.-P., Luo, Y. and Jiang, L.: Ecosystem carbon transit versus turnover times in response to climate warming and rising atmospheric CO_2 concentration, Biogeosciences, 15(21), 6559–6572, doi:10.5194/bg-15-6559-2018, 2018.
Lukacs, P. M., Burnham, K. P. and Anderson, D. R.: Model selection bias and Freedman's paradox, Ann. Inst. Stat. Math., 62(1), 117–125, doi:10.1007/s10463-009-0234-4, 2010.
425 Manzoni, S., Piñeiro, G., Jackson, R. B., Jobbágy, E. G., Kim, J. H. and Porporato, A.: Analytical models of soil and litter decomposition: Solutions for mass loss and time-dependent decay rates, Soil Biol. Biochem., 50, 66–76, doi:10.1016/j.soilbio.2012.02.029, 2012.
Meentemeyer, V.: Macroclimate and Lignin Control of Litter Decomposition Rates, Ecology, 59(3), 465–472, doi:10.2307/1936576, 1978.
430 Méndez, M. S., Ballaré, C. L. and Austin, A. T.: Dose–responses for solar radiation exposure reveal high sensitivity of microbial decomposition to changes in plant litter quality that occur during photodegradation, New Phytol., 235(5), 2022–2033, doi:10.1111/nph.18253, 2022.
Olson, J. S.: Energy Storage and the Balance of Producers and Decomposers in Ecological Systems, Ecology, 44(2), 322–331, doi:10.2307/1932179, 1963.
435 Parton, W. J., Schimel, D. S., Cole, C. V. and Ojima, D. S.: Analysis of Factors Controlling Soil Organic Matter Levels in Great Plains Grasslands, Soil Sci. Soc. Am. J., 51(5), 1173–1179, doi:10.2136/sssaj1987.03615995005100050015x, 1987.
Prescott, C. E. and Vesterdal, L.: Decomposition and transformations along the continuum from litter to soil organic matter in forest soils, For. Ecol. Manage., 498(July), 119522, doi:10.1016/j.foreco.2021.119522, 2021.
440 R Core Team: R: A language and environment for statistical computing., 2020.
Safriel, U. and Adeel, Z.: Dryland Systems, in Ecosystems and Human Well-being: Current State and Trends, Volume 1, edited by R. Hassan, R. Scholes, and N. Ash, pp. 623–662, Island Press, Washington., 2005.
Sarquis, A. and Sierra, C. A.: Supplementary Material for Sarquis & Sierra 2023 v1.0.0, doi:10.5281/zenodo.7561189, 2023.
445 Sarquis, A., Siebenhart, I. A., Austin, A. T. and Sierra, C. A.: Aridec : an open database of litter mass loss from aridlands worldwide with recommendations on suitable model applications, Earth Syst. Sci. Data, 14(7), 3471–3488, doi:10.5194/essd-14-3471-2022, 2022a.
Sarquis, A., Siebenhart, I. A., Austin, A. T. and Sierra, C. A.: aridec v1.0.2, doi:https://doi.org/10.5281/zenodo.6600345, 2022b.
450 Shumway, R. H. and Stoffer, D. S.: Time Series Analysis and Its Applications, Fourth., edited by R. DeVeaux, S. E. Fienberg, and I. Olkin, Springer International Publishing, Cham., 2017.
Sierra, C. A. and Müller, M.: A general mathematical framework for representing soil organic matter dynamics, Ecol. Monogr., 85(4), 505–524, doi:10.1890/15-0361.1, 2015.
Sierra, C. A., Müller, M. and Trumbore, S. E.: Models of soil organic matter decomposition: the SoilR package,
455 version 1.0, Geosci. Model Dev., 5(4), 1045–1060, doi:10.5194/gmd-5-1045-2012, 2012.
Sierra, C. A., Malghani, S. and Müller, M.: Model structure and parameter identification of soil organic matter models, Soil Biol. Biochem., 90, 197–203, doi:10.1016/j.soilbio.2015.08.012, 2015.
Sierra, C. A., Müller, M., Metzler, H., Manzoni, S. and Trumbore, S. E.: The muddle of ages, turnover, transit, and residence times in the carbon cycle, Glob. Chang. Biol., 23(5), 1763–1773, doi:10.1111/gcb.13556, 2017.
460 Sierra, C. A., Hoyt, A. M., He, Y. and Trumbore, S. E.: Soil Organic Matter Persistence as a Stochastic Process: Age and Transit Time Distributions of Carbon in Soils, Global Biogeochem. Cycles, 32(10), 1574–1588, doi:10.1029/2018GB005950, 2018.
Sinsabaugh, R. L., Antibus, R. K. and Linkins, A. E.: An enzymic approach to the analysis of microbial activity

during plant litter decomposition, Agric. Ecosyst. Environ., 34(1–4), 43–54, doi:10.1016/0167-8809(91)90092-
465   C, 1991.

Soetaert, K. and Petzoldt, T.: Inverse Modelling, Sensitivity and Monte Carlo Analysis in R Using Package FME, J. Stat. Softw., 33(3), 1–28, doi:10.18637/jss.v033.i03, 2010.

Tuomi, M., Thum, T., Järvinen, H., Fronzek, S., Berg, B., Harmon, M., Trofymow, J. A., Sevanto, S. and Liski, J.: Leaf litter decomposition-Estimates of global variability based on Yasso07 model, Ecol. Modell., 220(23), 3362–
470   3371, doi:10.1016/j.ecolmodel.2009.05.016, 2009.

Zanne, A. E., Flores-Moreno, H., Powell, J. R., Cornwell, W. K., Dalling, J. W., Austin, A. T., Classen, A. T., Eggleton, P., Okada, K., Parr, C. L., Adair, E. C., Adu-Bredu, S., Alam, M. A., Alvarez-Garzón, C., Apgaua, D., Aragón, R., Ardon, M., Arndt, S. K., Ashton, L. A., Barber, N. A., Beauchêne, J., Berg, M. P., Beringer, J., Boer, M. M., Bonet, J. A., Bunney, K., Burkhardt, T. J., Carvalho, D., Castillo-Figueroa, D., Cernusak, L. A., Cheesman, A.
475   W., Cirne-Silva, T. M., Cleverly, J. R., Cornelissen, J. H. C., Curran, T. J., D'Angioli, A. M., Dallstream, C., Eisenhauer, N., Evouna Ondo, F., Fajardo, A., Fernandez, R. D., Ferrer, A., Fontes, M. A. L., Galatowitsch, M. L., González, G., Gottschall, F., Grace, P. R., Granda, E., Griffiths, H. M., Guerra Lara, M., Hasegawa, M., Hefting, M. M., Hinko-Najera, N., Hutley, L. B., Jones, J., Kahl, A., Karan, M., Keuskamp, J. A., Lardner, T., Liddell, M., Macfarlane, C., Macinnis-Ng, C., Mariano, R. F., Méndez, M. S., Meyer, W. S., Mori, A. S., Moura, A. S.,
480   Northwood, M., Ogaya, R., Oliveira, R. S., Orgiazzi, A., Pardo, J., Peguero, G., Penuelas, J., Perez, L. I., Posada, J. M., Prada, C. M., Přívětivý, T., Prober, S. M., Prunier, J., Quansah, G. W., Resco de Dios, V., Richter, R., Robertson, M. P., Rocha, L. F., Rúa, M. A., Sarmiento, C., Silberstein, R. P., Silva, M. C., Siqueira, F. F., Stillwagon, M. G., Stol, J., Taylor, M. K., Teste, F. P., Tng, D. Y. P., Tucker, D., Türke, M., Ulyshen, M. D., Valverde-Barrantes, O. J., et al.: Termite sensitivity to temperature affects global wood decay rates, Science
485   (80-.)., 377(6613), 1440–1444, doi:10.1126/science.abo3856, 2022.

Zhang, X. and Wang, W.: Control of climate and litter quality on leaf litter decomposition in different climatic zones, J. Plant Res., 128(5), 791–802, doi:10.1007/s10265-015-0743-6, 2015.