**Noname manuscript No.**
(will be inserted by the editor)

# Kombucha: An Empirical Model

## Sugars, Acids, Cellulose and Alcohol

**Junling Ma · Dayten Sheffar**

Received: date / Accepted: date

**Abstract** This paper attempts to fill a gap in the existing literature on Kombucha fermentation by providing an empirical model for *open* systems that are either static (study 1) or agitated (study 2). A theoretical system of ordinary differential equations is established from underlying biology and chemical reactions. Using the Nelder-Mead method, the parameters of the model are optimized for two existing studies. We analyzed the conservation of carbon-mass and the data from study 1 suggests fructose is synthesized to gluconic acid (fig. 2) by Kombucha bacteria, which is not well documented in existing literature, study 2 was otherwise as predicted. Thus, from a more general theoretical model we derived then fit a predictive model per methodology. The resulting models predict the Kombucha fermentation process for up to 20 days, this is a time-frame consistent with and even beyond most commercial brewing applications. The models can therefore be adapted to further studies of Kombucha.

**Keywords** Kombucha · ODE · Model

**Mathematics Subject Classification (2010)** MSC 92D40 · MSC 92C42

## 1 Introduction

### 1.1 History

Kombucha is an ancient traditional beverage consisting of a sweetened tea base cultured with specific strains of bacteria in symbiosis with yeasts [7]. The result is a sour, sometimes vinegary beverage with naturally low levels of carbonation

Junling Ma
Department of Mathematics and Statistics, University of Victoria,
Victoria, BC V8W 2Y2, Canada
E-mail: junlingm@uvic.ca

Dayten Sheffar
Department of Mathematics and Statistics, University of Victoria,
Victoria, BC, V8W 2Y2, Canada
E-mail: dsheffar@uvic.ca

(a byproduct of the fermentation), and in commercial applications it is common to force carbonate the beverage to higher concentrations of $CO_2$. Kombucha also yields bacterial cellulose via specific strains of bacteria (*Komagataeibacter xylinus*), typically via a statically fermented batch.

The commercial product is typically prepared by mixing a determinate quantity of black or green tea leaves (a nitrogen source) with boiled water (80-100°C), and stirring in a sweetener (a carbon source) while still quite hot (usually sucrose but a plethora of options are available (molasses, grape juice, sour cherry juice, coconut water, *et cetera*)). This mixture is left to cool down (20-30°C) before adding a prepared innoculum of yeast and bacteria. The consortium can be prepared from a previously obtained culture which is placed into sweetened tea in sufficient quantities and left to ferment. During fermentation a cellulosic pellicle layer (otherwise referred to as the mother culture) is formed by the bacteria by metabolizing certain substrates. This small inoculating batch (prepared kombucha) can then be used for inoculating a new batch at 8-15% (v/v). It is typical to add both the solid cellulosic pellicle and the liquid sour broth to inoculate the present batch.

There are two general environments regarding fermenting Kombucha, the first being an *open* system which is left in a vessel to ferment with a porous cloth covering the vessel's opening. This allows the system to effectively breathe, as oxygen is required for the obligately aerobic bacteria (OAB). The second method is a *closed* system in which a sealed lid is affixed to the fermentation vessel and oxygen must be supplied at an appropriate rate to supply the OAB, else oxygen starvation and die-off occurs. An open ferment is generally *wild*, meaning the consortia of yeasts and bacteria is typically somewhat unknown, and are subject to vary in species and population size. The closed fermentation process can allow for control over species and provide reason for (and meaning to) quantifying the species present. This study focuses on open, wild ferments, of two general methodologies.

In either open or closed systems, there are options for either a static or agitated ferment. In a static ferment there is opportunity for a biofilm to develop (the aforementioned cellulosic pellicle layer). This symbiotic culture of yeast and bacteria (SCOBY) grows from the top down, pushing old layers into the broth as new layers form on top. However, a critical thickness of the SCOBY can occur, effectively starving the OAB of oxygen, a consequence of at least three factors.

First, the maximum transport distance of the oxygen [6] is much less than for sucrose, as oxygen is much less soluble in water (0.004 g/L at 25°C) than sucrose (2100 g/L). This presents a limiting factor for metabolic activity as the SCOBY thickens. Vershuchren *et al.* discovered that the saturated oxygen profiles started from 200 to 500$\mu$m depending on what was the initial broth of the pH, where all relevant cases ($pH_0 \leq 5$) reduced to under 100$\mu$m by day 3. The broths then maintained around 50$\mu$m during the ferment [6]. Moreover, those authors found cellulose forms exclusively in the 50 to 100 $\mu$m depths, which quickly become oxygen deprived.

Second, the amount of dissolved oxygen also critically depends on the interfacial area within the fermentation vessel [1], granting there is sufficient depth for saturation (not in too shallow a vessel). In the agitated case, if stirred frequently enough, the cellulose pellicle cannot form in a suspended matrix over the broth as the network of cells is broken and cannot form thick layers, hindering accelerated growth. Thus oxygen starvation is less of a concern in the open and agitated systems.

Third, a byproduct of fermentation is $CO_2$ release due to cellular respiration which can build up under the SCOBY and cause the partial pressure of oxygen to drop seriously enough to impact the bacteria and cause die-off [4]. This is only an issue in static cases where the SCOBY remains undisturbed, such that the $CO_2$ is not capable of escaping the critical metabolic zone within $50\mu$m of the surface, causing the oxygen deprived environment and effectively starving the OAB of a primary metabolic supplement.

The typical yeast present are of genus *Saccharomyces*, especially *Candida sp.*, and *Zygosaccharomyces*. [5]. The typical bacteria are *Acetobacter xylinum* (acetic acid bateria –AAB), *Gluconacetobacter* (glucose producing bacterium), and *Komagataeibacter xylinus*, which is the species best known for cellulose production. However, the intractability of modelling specific strains, and so instead modelling global behaviors per family, leads to referring to the consortia of yeasts as simply *the yeast*, and the various Kombucha bacteria as *the bacteria*.

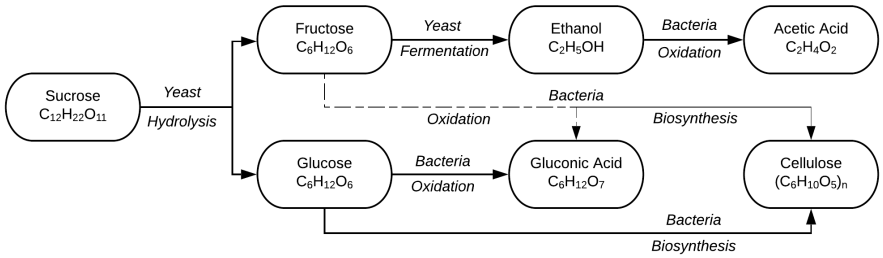


**Fig. 1** Substrate Compartment Flow; (Solid) Known mechanism; (Dashed) Postulated Mechanism in this paper

Regarding figure 1, the yeast plays a critical role through hydrolysis of the polysaccharide sucrose, producing two monosaccharides, fructose and glucose, by the invertsae enzyme. Fructose is fermented into ethanol by yeast, which in turn is oxidized into acetic acid by the *Acetobacter aceti*, this is the first chain of reactions to consider. The second is the conversion of glucose into gluconic acid by *Gluconobacter Oxydans* and *A. aceti* [2]. Finally, *A. aceti* and *K. xylinus* biosynthesize cellulose from various subtrates, however in our research we found glucose to be the predominant substrate for this process. The assumed compartment flow for our ordinary differential equations model is as in figure 1.

As per a study done by Sievers *et al.* in an open and agitated ferment, they found conservation of mass without evidence of significant anabolism [5]. The good stoichiometry they observed led them to conclude the major fermentation products (sucrose, fructose, ethanol, acetic acid, glucose, gluconic acid) have all been accounted for, with other substrates and species having negligible impact. The undertaking of this paper is to establish working models of both the static open ferment and the agitated open ferment — closed ferments are not considered at this time. The research comprises the data of two former studies, Sievers *et al.* [5] being one. The other paper by Chen and Liu [3] provides data for an open, static ferment. In the latter case, the mentioned substrates do not account for all the

chemistry, and we posit that in conjunction to those mentioned, considering also the formation of cellulose (as a sink taking carbon effectively out of the system) is a sufficient postulate for the models at hand, in order to main good conservation of carbon-mass.

## 2 Preamble to Formulation of the Models

### 2.1 Chemical Reactions

Here we turn to the various underlying chemical reactions to determine the flow of our system. Consider the reactants: sucrose, $C_{12}H_{22}O_{11}$; fructose and glucose, $C_6H_{12}O_6$; ethanol, $C_2H_5OH$; gluconic acid, $C_6H_{12}O_7$; and acetic acid, $C_2H_4O_2$. Sucrose is converted to fructose and glucose via the yeast enzyme invertase (hydrolysis)

$$C_{12}H_{22}O_{11} + H_2O \xrightarrow{Y} C_6H_{12}O_6 + C_6H_{12}O_6$$

One fructose molecule is fermented to two ethanol molecules, with carbon dioxide byproduct,

$$C_6H_{12}O_6 \xrightarrow{Y} 2C_5H_5OH + 2CO_2$$

and each resulting ethanol molecule is oxidized by bacteria into acetic acid (really a two step process, but herein simplified to one step)

$$C_2H_5OH + 2[O] \xrightarrow{B} C_2H_4O_2 + H_2O$$

and acetic acid dissociates into acetate ion and a hydrogen ion

$$C_2H_4O_2 \xrightarrow{\text{Dissociation}} C_2H_3O_2^- + H^+$$

Glucose is oxidized to gluconic acid by the bacteria

$$C_6H_{12}O_6 + [O] \xrightarrow{B} C_6H_{12}O_7$$

which dissociates to form a gluconate ion and a hydrogen ion

$$C_6H_{12}O_7 \xrightarrow{\text{Dissociation}} C_6H_{11}O_7^- + H^+$$

### 2.2 pH

The pH is defined as the base 10 logarithm of the reciprocal of the hydrogen ion concentration, or the negative base 10 logarithm of the hydrogen ion concentration

$$\text{pH} = \log\left(\frac{1}{H^+}\right) = -\log\left(H^+\right)$$

Also concerning pH, $CO_2$ reacts with $H_2O$ to form carbonic acid ($H_2CO_3$), with dissociates to release a hydrogen ion ($H^+$)and bicarbonate ions ($HCO_3^-$). The pH of the broth serves as a buffering system for the bacteria, halting their metabolic activity and rendering them dormant at some threshold. If the pH increases then the bacteria reinvigorate and continue metabolizing available substrates.

Accounting for molarity, we chart the chemical reactions

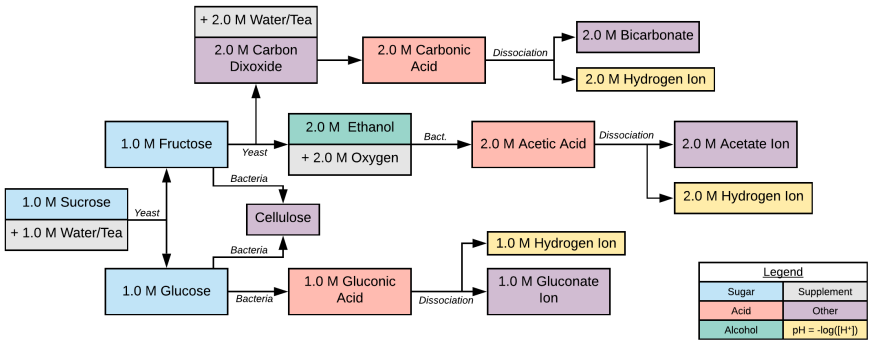


**Fig. 2** Chemical reactions accounting for molarity

## 3 Formulation of the Models

In this section we turn to the formulation of the two models. The studies are conducted over 60 days [5][3], but most commercial Kombucha for consumable products range from 10 to 20 days, before the pH of the beverage dips below 3.0, which is considered unsafe for the digestive tract upon human consumption. Therefore, we model only the first 20 days of each study for the current paper, but analyse their data over the whole time frame.

As mentioned, considering the ample consortia of yeasts and bacteria that may be present, in modeling the dynamics of their symbiosis over the substrates, we found modeling behaviours specifically by strain to be intractable. Switching to modeling genus behaviours was at least tractable and provided satisfactory results in optimizing model parameters to the various data. Examination of the involved chemical reactions provided a starting point for a model.

### 3.1 Theoretical Model

Let $S, F, E, A, G_s, G_a$ and $C$ be respectively the concentrations of sucrose, fructose, ethanol, acetic acid, glucose, gluconic acid and cellulose. Furthermore, let $Y$ denote the yeast and $B$ the bacteria, then along with parameters

$$a, b, c, d, e, f, \alpha, \beta, \gamma, \delta, \epsilon, \lambda, \rho, \text{ and } r$$

we derive the theoretical model through proportionality and substrate concentration limiting behavior, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}t}\text{Substrate} = \frac{\text{param}_1 \cdot \text{Substrate}}{1 + \text{param}_2 \cdot \text{Substrate}} \cdot \text{Catalytic Agent}$$

This, together with the chemical formulae, yields the theoretical model:

$$\mathrm{d}S = -\frac{aS}{1 + \alpha S} Y$$

$$
\begin{aligned}
\mathrm{d}F &= \frac{aS}{1+\alpha S}Y - \frac{bF}{1+\beta F}Y - \frac{cF}{1+\gamma F}B \\
\mathrm{d}E &= 2\frac{bF}{1+\beta F}Y - \frac{dE}{1+\delta E}B \\
\mathrm{d}A &= \frac{dE}{1+\delta E}B \\
\mathrm{d}G_s &= \frac{aS}{1+\alpha S}Y - \frac{eG_s}{1+\epsilon G_s}B - \frac{fG_s}{1+\lambda G_s}B \\
\mathrm{d}G_a &= \frac{eG_s}{1+\epsilon G_s}B \\
\mathrm{d}C &= \frac{cF}{1+\gamma F}B + \frac{fG_s}{1+\lambda G_s}B \\
\mathrm{d}Y &= \rho(\frac{aS}{1+\alpha S} + \frac{bF}{1+\beta F})Y \\
\mathrm{d}B &= r(\frac{cF}{1+\gamma F} + \frac{dE}{1+\delta E} + \frac{eG_s}{1+\epsilon G_s} + \frac{fG_s}{1+\lambda G_s}B)B
\end{aligned} \tag{1}
$$

### 3.2 Agitated Model

We fit this model to the research from Sievers *et al.* in statistical computing language R. Along with a few parameters going to zero there was a persistent lack of mass balance in the glucose parameter optimization. We were seeing that no matter the approach, there was always insufficient glucose, and thereby gluconic acid, in the fitted model. The fitting curves were consistently below the respective data points, indicating a mistaken assumption in the model that failed to correspond to what we're now positing is a flow from fructose to gluconic acid, which is not supported in current literature. This conclusion is supported weakly by the fact that it indeed works to improve the fit, stronger evidence however is provided by a analysis of the raw data.

In figure 3, the sucrose data points indicate the loss of sucrose $S_0 - S_t$, or equivalently, the gain of fructose and glucose over time. But as fructose and glucose are also decaying, we need to supplement their data values with their byproducts. As is clearly visible from the first and second graphs, glucose sits nearly to the required values until day 24, and adding the concentration of gluconic acid to it would bump the data values above the "sucrose" values, which cannot be the case. Both fructose and glucose data points ought to match the loss of sucrose (as one sucrose molecule hydrolyzes to one fructose and one glucose molecule). The problem here is that fructose is obviously low in the first graph of figure 3.

To remedy this, we added ethanol and acetic acid (the known byproducts in the fructose cycle) and still saw the fructose data points fell short of the "sucrose" data. Adding gluconic acid to the fructose cycle produces the second graph in figure 3, also up to day 24. After day 24 we see the fructose, ethanol, acetic acid and gluconic acid sum grows away from the sucrose loss values, this is symmetrically opposed to the glucose cycle, which decays away from the sucrose loss profile. Taking the fructose cycle, and average it with the glucose values, produces the third graph in figure 3, where the data falls nearly perfectly along the sucrose loss curve. This result gestures at the good conservation of carbon-mass claimed by Sievers *et al.* [5].

This analysis then proposes fructose must be oxidized into gluconic acid quite significantly. It is possible that around day 24 the bacteria switch to metabolizing

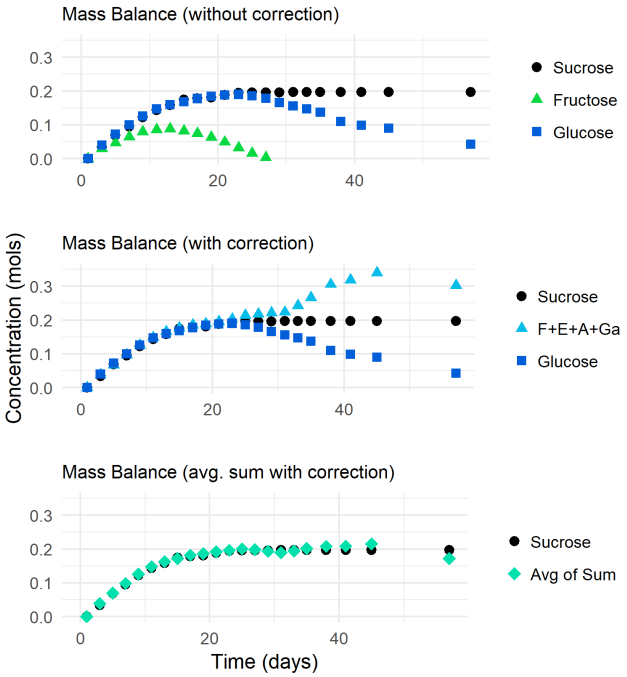


**Fig. 3** Data Analysis; Sucrose here is actually $S_0 - S_t$, hence the *loss* of sucrose. We plot to day 40 as the data beyond this is sparse and relevant behaviours are readily observed within this range.

glucose to gluconic acid (a well-supported process in existing literature). This proposition could fix the deviation after day 24 (in figure 3), but then adding gluconic acid to the glucose cycle, but before then the bacteria must be utilizing fructose for this process.

Forging ahead with this presupposition, plus factoring in that glucose *should* in essence be converted to gluconic acid at some point, and removing terms such that their main parameters go to zero during repetitive optimization, then we obtain the following model which provides a significant fit to the data. Additionally, we relabel some parameters to maintain a stylistic convention.

$$
\begin{aligned}
\mathrm{d}S &= -aSY \\
\mathrm{d}F &= aSY - \frac{bF}{1+\beta F}Y - \frac{cF}{1+\gamma F}B \\
\mathrm{d}E &= 2\frac{bF}{1+\beta F}Y - \frac{dE}{1+\delta E}B \\
\mathrm{d}A &= \frac{dE}{1+\delta E}B \\
\mathrm{d}G_s &= aSY - \frac{eG_s}{1+\alpha G_s}B \\
\mathrm{d}G_a &= \frac{cF}{1+\gamma F}B + \frac{eG_s}{1+\alpha G_s}B \\
\mathrm{d}C &= 0 \qquad \text{(left for explicitness)} \\
\mathrm{d}Y &= \rho\left(aSY + \frac{bF}{1+\beta F}\right)Y
\end{aligned}
$$

$$\mathrm{d}B = r\left(\frac{cF}{1+\gamma F} + \frac{eG_s}{1+\alpha G_s} + \frac{dE}{1+\delta E}\right)B \tag{2}$$

We leave in cellulose for explicitness as all parameters tended to zero during optimization. Any actual growth would be insignificant, reflecting the poor chance for a biofilm to grow between agitations, which, once stirred, would merely disintegrate or float in insignificant clumps along the surface. When the structural integrity of the biofilm is regularly compromised it's growth is seriously hindered.

We conclude that model 1 is a good selection for at least the first 20 days of Kombucha fermentation in an agitated and open ferment. Of notable interest is what we postulated regarding gluconic acid being a significant part of the fructose byproduction cycle for at least the first 20-odd days, without which the carbon-mass conservation is erroneous. Thus, running model 2 in R, using log-likelihood estimation by least squares regression analysis using the Nelder-Mead optimization algorithm for our scaling parameters, we obtain figure 4. It was considered briefly

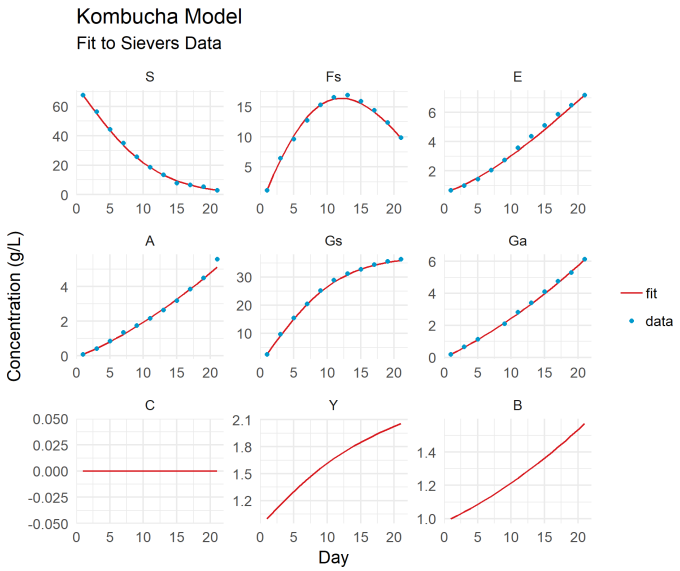


**Fig. 4** Agitated Kombucha model fit to Sievers *et al.* data over 20 days

**Table 1** Parameter Estimates for Sievers *et al.* data

| Parameter | Value | Parameter | Value | Parameter | Value |
|---|---|---|---|---|---|
| $a$ | 9.408889e-02 | $e$ | 7.149756e-04 | $\delta$ | 3.396073e+01 |
| $b$ | 3.531752e-01 | $\alpha$ | 7.645353e-04 | $\rho$ | 3.707392e+00 |
| $c$ | 1.154273e+00 | $\beta$ | 8.806522e+01 | $r$ | 2.134664e+00 |
| $d$ | 1.851705e-01 | $\gamma$ | 8.806696e+02 | | |

whether the original authors, Sievers *et al.*, switched their sugar labels in their paper's graph. However, a quick data analysis suggests that switching the fructose and glucose labels, then considering glucose plus gluconic acid on the one hand, and fructose, ethanol and acetic acid on the other, does not yield carbon-mass conservation (as per figure 3). That is, the sum in moles of the immediate glucose and fructose byproduction cycles results in compromised conservation, occurring as early day 5. Thus, against expectation, the data suggests gluconic acid is at least partly a product of fructose oxidation, for some period of time in the early stages of fermentation. This encourages an interesting potential line of research regarding which strains of relevant bacteria may produce gluconic acid from fructose, or else to disprove this hypothesis.

### 3.3 Static Model

We begin the determination of the static model for Kombucha fermentation from the Sievers *et al.* model *mutatis mutandis* for Chen and Liu's research [3]. In the static case we expect cellulose production, and so we cannot expect a similar carbon-mass conservation without the cellulose mass data, which is unfortunately not included in their research. That is, we expect cellulose to accumulate and thus remove carbon from the system. Performing a similar data analysis as in figure 5, we see this expected behaviour in the third graph, as the carbon balance tends away from the conservation equilibrium profile (loss of sucrose). From the theoretical

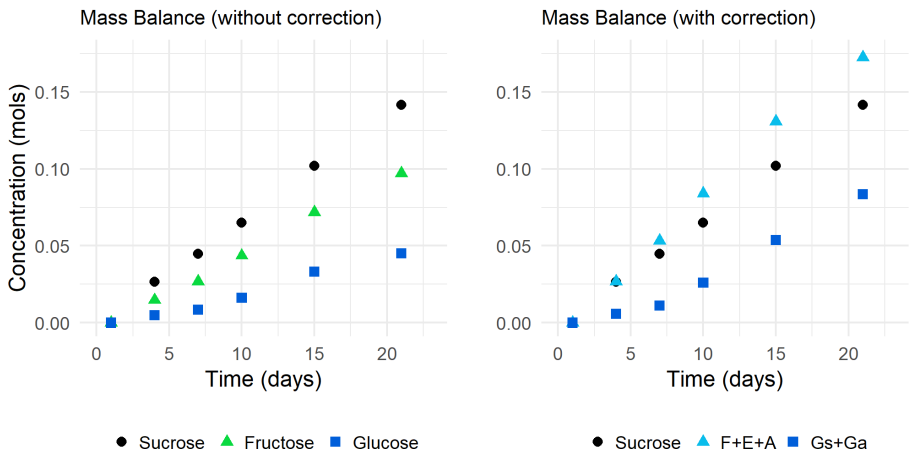


**Fig. 5** Data Analysis; Sucrose is actually $S_0 - S_t$, hence the *loss* of sucrose. We plot 30 days only as the data beyond this is sparse and relevant behaviours are readily observed within this range.

model 1, we disallow flow from fructose to gluconic acid. In the Chen & Liu data, it is apparent that acetic acid decays after some time. In their research paper they also note that accumulated $CO_2$, or decreasing partial pressure of $O_2$ ($pO_2$), under

the SCOBY, caused the bacteria to die. This presented no issue in the Sievers *et al.* study as there was a constant supply of oxygen at the surface due to the agitation i) preventing biofilm build-up which incidentally causes decreasing $pO_2$, and ii) oxygenating the broth which effectively increases $pO_2$. Due to these facts, we've introduce a die-off term for yeast ($\mu Y$) as well as for bacteria ($\nu B$), as modelling the $CO_2$ is, at present, intractable without a comprehensive data set of sucrose, fructose, ethanol, acetic acid, glucose, gluconic acid, yeast counts, bacterial counts, $CO_2$ concentrations, et cetera. Thus, instead, the *impact* of increasing $CO_2$ levels are modelled.

Now that bacterial cellulose (BC) production is a significant element of the system we'll spend a brief moment expanding on the biology. The bacteria biosynthesize cellulose from glucose and fructose. Sucrose is another known carbon source for bacterial sucrose synthesis, however, we neglect that here as there was no obvious benefit to including it in the optimization (no closer fit), also it doesn't seem to be a preferred carbon source for the bacteria. Favorable conditions are high levels of carbon sources and low levels of nitrogen sources (from steeping the tea). Sufficient oxygen $pO_2$, along with having ethanol present in the system, both supplement cellulose production. Acetic acid promotes, but gluconic acid hinders, BC production.

Similarly for this model, several parameters tend to zero and upon removing them and relabelling the remaining parameters, we obtain a static model

$$\begin{aligned}
\mathrm{d}S &= -\frac{aS}{1+\alpha S}Y \\
\mathrm{d}F &= \frac{aS}{1+\alpha S}Y - \frac{bF}{1+\beta F}Y \\
\mathrm{d}E &= 2\frac{bF}{1+\beta F}Y - dEB \\
\mathrm{d}A &= dEB \\
\mathrm{d}G_s &= \frac{aS}{1+\alpha S}Y - cG_sB - \frac{eG_s}{1+\gamma G_s}B \\
\mathrm{d}G_a &= cG_sB \\
\mathrm{d}C &= \frac{eG_s}{1+\gamma G_s}B \\
\mathrm{d}Y &= \rho\left(\frac{aS}{1+\alpha S} + \frac{bF}{1+\beta F}\right)Y - \mu y \\
\mathrm{d}B &= r\left(cG_s + dE + \frac{eG_s}{1+\gamma G_s}\right)B - \nu B
\end{aligned} \tag{3}$$

**Table 2** Parameter Estimates for Chen & Liu *et al.* data

| Parameter | Value | Parameter | Value | Parameter | Value |
|---|---|---|---|---|---|
| $a$ | 1.014047e+06 | $e$ | 2.708930e+05 | $\rho$ | 1.320359e+06 |
| $b$ | 9.960544e+05 | $\alpha$ | 1.337071e+08 | $r$ | 1.347670e+01 |
| $c$ | 1.080606e-01 | $\beta$ | 3.563466e+08 | $\mu$ | 1.370437e+04 |
| $d$ | 2.176345e-01 | $\gamma$ | 4.161985e+07 | $\nu$ | 2.750559e-01 |

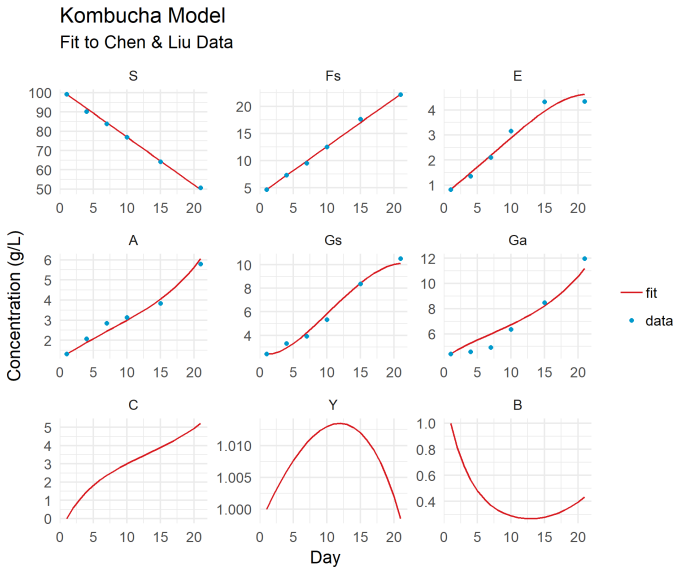


**Fig. 6** Static Kombucha model fit to Chen & Liu's data over 20 days. **Graph is placeholder until Static Model is at best known determination**

Considering that the Chen & Liu study was conducted on nine independent batches (different mother cultures) with 8 trials each, their data points represent the mean of all their samples. Therefore, an absolute fit is not as fundamentally necessary as with the research from Siever's *et al.*. Moreover, our fit is narrowly within the exact values of Chen and Liu's data (see figure 6), so any error is less meaningful for a narrow variance since it still falls well within the range of their data.

## 4 Computations

The model was coded in R along with a list of parameters and initial conditions. We wrote a log-likelihood function to compute a least squares regression model. Using the R package 'deSolve', we get a time series for each i) substrate, ii) byproduct, and iii) the yeast and bacteria. This requires converting the gathered data points (given in $gL^{-1}$) to moles, then running the model with a good first approximate parameterization through an ordinary differential equations solver from the 'deSolve' package, which calls the Runge-Kutta(4,5) ODE solving method. The solution matrix then needs to be converted back to $gL^{-1}$ in order to perform the least squares regression with the data. For sucrose, fructose, ethanol, acetic acid, glucose and gluconic acid, we code an objective function to optimize

$$F_{obj.} = \sum \frac{(\theta_{\text{predicted}} - \theta_{\text{observed}})^2}{|\theta_{\text{observed}}|} \tag{4}$$

which gets passed to an optimization function (*optim()*) provided in the base installation of R. This function is passed a list to optimize over, which for us

is the list of parameters that allows us to fit to the data. It is also passed the objective function $F_{obj.}$ from equation 4, whose value is minimized (in our case) such that the minimum value returned corresponds to the closest fit to the data. $F_{obj.}$ is an $n$-dimensional surface that the Nelder-Mead method searches along for smaller local minima using a simplex (generalization of a triangle or tetrahedron to $n$ dimensions), that ideally converges to a local minima. This local minima represents the parameter selection that minimizes $F_{obj.}$ as desired.

The values of the objective functions found for Sievers *et al.* and Chen & Liu are 1.380669 and 0.9560687 respectively.

## 5 Assessing Acidity

pH serves as a major quality control measure in brewing Kombucha as a commercial beverage. We've worked on extending our model to predict the pH of the beverage. To establish a connection to pH from our model, we had find a way to use our predictions of acetic acid and gluconic acid to calculate pH from the associated hydrogen ion concentrations. A quick look at the reactions in section 2.1 helps or the chart in Figure 2 may help to inspire the following discussion on acid dissociation.

Dissociation is a chemical reaction in the context of acid-base pairs, specifically acetic acid and gluconic acid reacting with $H_20$. The acids dissolve in water, meaning the acid molecules separate into cations and anions (positively and negatively charged atoms). This occurs when the covalent bond between an electro-negative atom and a hydrogen atom undergoes heterolysis. The result is a hydrogen ion $H^+$, and either an negative acetate ion ($C_2H_3O_2^-$) or a negative gluconate ion ($C_6H_{11}O_7^-$). We can express this as

$$\underset{acetic}{C_2H_4O_2 = C_2H_3O_2^- + H^+} \qquad \underset{gluconic}{C_6H_{12}O_7 = C_6H_{11}O_7^- + H^+} \tag{5}$$

First, one measure of the strength of an acid in solution is the acid dissociation constant $K_a$. It is the equilibrium constant of

$$HA \rightleftharpoons A^- + H^+$$

where equilibrium indicates the steady state of each concentration due to the forward and backward reactions occuring at the same rate. When the concentrations are at equilibrium, then the constant can be defined as

$$pK_a = \log_{10} \frac{[HA]}{[H^+][A^-]}$$

So one can find $K_a$ through

$$K_a = 10^{-pKa} = \frac{[H^+][A^-]}{[HA]}$$

Acetic acid has $pKa = 4.756$ and gluconic acid has $pKa = 3.86$, and so the equilibrium constants $k_A$ for acetic acid, and $k_{G_a}$ for gluconic acid, can be expressed as

$$k_A = \frac{[H^+][C_6H_3O_2^-]}{[C_2H_4O_2]} = 10^{-4.756} \qquad k_{G_a} = \frac{[H^+][C_6H_{11}O_7^-]}{[C_6H_{12}O_7]} = 10^{-3.86}$$

The active hydrogen ion concentrations can be expressed as

$$[H+] = [C_6H_3O_2^-] + [C_6H_{11}O_7^-]$$

We can then express the concentrations more succinctly, first we see for acetate ion

$$\begin{aligned}
[C_6H_3O_2^-] &= A(t)\frac{[C_6H_3O_2^-]}{[C_6H_3O_2^-] + [C_2H_4O_2]} \\
&= A(t)\frac{1}{1 + \frac{[C_2H_4O_2]}{[C_6H_3O_2^-]}} \\
&= A(t)\frac{1}{\frac{[H^+]}{k_A} + 1} \\
&= A(t)\frac{k_A}{[H^+] + k_A}
\end{aligned}$$

and similarly for gluconate ion

$$[C_6H_{11}O_7^-] = G_a(t)\frac{k_{G_a}}{[H^+] + k_{G_a}}$$

Wherefore, the concentration of hydrogen ions (which determine pH) are calculated by

$$\begin{aligned}
[H^+] &= [C_6H_{11}O_7^-] + [C_6H_3O_2^-] \\
&= \frac{G_a(t)k_{G_a}}{[H^+] + k_{G_a}} + \frac{A(t)k_{G_a}}{[H^+] + k_A}
\end{aligned} \tag{6}$$

Since $[H^+]$ is essentially strictly increasing while the right hand sight of 6 is decreasing, we expect one real positive root to indicate the hydrogen ion concentration at time $t$. However, our model was slightly inaccurate in predicting Sievers *et al.* pH data by several orders of magnitude in the logarithmic scale (7). It is too large an error in prediction to chalk up to instrument calibration, so must have been introduced in the process of analysis.

The error could have accrued since the initial data digitization, or possibly we are missing a key component for pH in the model. There are amino acids not considered and other organic acids can be present. The build up of carbon dioxide may play a crucial role in affecting the pH. First consider the agitated model, there is no cellulose build-up over the interfacial area, and so $CO_2$ is not buffered in the system as it can off-gas. As yeast can be both anaerobic or aerobic, it is possible for them to metabolize even in the lowest layers of the fluid where undissolved sucrose (a preferred substrate) may have fallen. Then any $CO_2$ particles rising through the fluid have ample opportunity to dissociate into carbonic acid and a hydrogen ion, effecting the pH while off gassing.

$$\begin{aligned}
\mathrm{pH}_{0_{exp.}} &= 3.75 & \mathrm{pH}_{f_{exp.}} &= 2.42 \\
\mathrm{pH}_{0_{obs.}} &= 3.52 & \mathrm{pH}_{f_{obs.}} &= 2.34
\end{aligned} \tag{7}$$

This indicates there must be a non-negligible quantity of substrate acting as a base, raising the pH.

One possible way to approach a solution could be to place as proxy whatever else *should* be adding to the pH via a term

$$\frac{Tk_T}{[H^+] + kT}$$

where $T$ is a constant body and $k_T$ is it's associated dissociation constant. Solving

$$[H^+] = \frac{G_a(t)k_{G_a}}{[H^+] + k_{G_a}} + \frac{A(t)k_A}{[H^+] + k_A} + \frac{Tk_T}{[H^+] + k_T}$$

would yield the indeterminate acidity if these elements remain unknown.

## 6 Discussion

Overall, our general theoretical model was able to fit well to two fairly different kombucha fermentation processes, with some coefficients optimizing out of the equations. The fit to Sievers *et al.* study data was quite adherent to the data points, with realistic yeast and bacteria growth curves, and no cellulose growth due to their constant stirring. Moreover, the analysis of the Sievers study suggests an interesting line of research into bacteria producing gluconic-acid from fructose, which is not reflected in current literature to the authors' awareness. Considering that the available data from Chen & Liu data was averaged data over multiple studies, the fit being less adherent to the data points is not as surprising, or at least not further determinable without access to the trial data underlying the averaged data. Lastly, the unimpressive pH predictions as discussed in the previous section could be accounted for with a proxy if in depth analysis of the contributors is

## References

1. Cvetković, Dragoljub, et al. "Specific Interfacial Area as a Key Variable in Scaling-up Kombucha Fermentation." Journal of Food Engineering, vol. 85, no. 3, 2008, pp. 387–392., doi:10.1016/j.jfoodeng.2007.07.021.
2. Gomes, Rodrigo José, et al. "Acetic Acid Bacteria in the Food Industry: Systematics, Characteristics and Applications." Food Technology and Biotechnology, vol. 56, no. 2, Jan. 2018, doi:10.17113/ftb.56.02.18.5593.
3. Chen, C., and B.y. Liu. "Changes in Major Components of Tea Fungus Metabolites during Prolonged Fermentation." Journal of Applied Microbiology, vol. 89, no. 5, 2000, pp. 834–839., doi:10.1046/j.1365-2672.2000.01188.x.
4. Kouda, Tohru, et al. "Inhibitory Effect of Carbon Dioxide on Bacterial Cellulose Production by Acetobacter in Agitated Culture." Journal of Fermentation and Bioengineering, vol. 85, no. 3, 1998, pp. 318-321., doi:10.1016/s0922-338x(97)85682-6.
5. Sievers, Martin, et al. "Microbiology and Fermentation Balance in a Kombucha Beverage Obtained from a Tea Fungus Fermentation." Systematic and Applied Microbiology, vol. 18, no. 4, 1995, pp. 590–594., doi:10.1016/s0723-2020(11)80420-0.
6. Verschuren, Peter G., et al. "Location and Limitation of Cellulose Production by Acetobacter Xylinum Established from Oxygen Profiles." Journal of Bioscience and Bioengineering, vol. 89, no. 5, 2000, pp. 414–419., doi:10.1016/s1389-1723(00)89089-1.
7. Villarreal-Soto, S.A.; Beaufort, S.; Bouajila, J.; Souchard, J.P.; Taillandier, P. Understanding kombucha tea fermentation: A review. J. Food Sci. 2018, 83, pp. 580–588, doi:10.1111/1750-3841.14068