Growth Kinetics and Viability of CHO Cells in Shake Flask Culture

Damini Patel

December 2025

Abstract

Chinese hamster ovary (CHO) cells are a foundational platform for modern biopharmaceutical production, and their growth performance depends strongly on operational parameters such as working volume and agitation speed. In this project, I analyzed a Lonza–UNH shake-flask dataset containing 3 L and 1 L culture conditions to evaluate how these parameters influence specific growth rate, doubling time, and viability. Using R for data cleaning, visualization, and statistical modeling, I generated five figures and a summary table capturing key performance trends. A simple linear model was applied to quantify the association between agitation speed and specific growth rate. Overall, CHO cultures demonstrated high viability (>98%) and consistent doubling times, while agitation speed showed a modest positive association with growth rate. These findings provide a quantitative basis for optimizing shake-flask operation in future CHO cell expansion studies.

Introduction

Chinese hamster ovary (CHO) cells are the dominant mammalian host used in biomanufacturing because of their robustness, ability to grow in suspension, and capacity to produce complex recombinant proteins. Despite their widespread use, CHO cell performance is highly sensitive to culture conditions—including working volume, mixing intensity, dissolved oxygen availability, and media composition. Even relatively small changes in shake-flask operating conditions can alter growth kinetics, viability, and productivity, motivating systematic evaluation of these variables.

In the UNH–Lonza CHO cell coursework, students routinely expand CHO CRL-3440 cells and perform multi-condition shake-flask experiments designed to compare growth performance across volumes and agitation speeds. The dataset analyzed here comes from a previous semester's multi-condition study and provides an opportunity to apply tools from Data Science for Life Sciences, including data wrangling, visualization, and linear modeling.

The objective of this project was to use R to clean and summarize the dataset, quantify differences in 3 L growth kinetics across conditions, evaluate viability maintenance, visualize relationships between agitation speed and specific growth rate, and apply a simple mathematical model to test whether agitation is predictive of growth performance.

Methods

Data source and Cleaning

The dataset originated from a Lonza training-lab shake-flask experiment involving CHO CRL-3440 cells cultured in chemically defined media. The base medium used for expansion was HyClone SFM4CHO (Cytiva), a protein-free, chemically defined medium formulated for high-yield CHO suspension culture. SFM4CHO was prepared and supplemented according to the manufacturer's standard guidelines, including L-glutamine and poloxamer 188 additions.

Each "pattern" represented a unique combination of working volume and agitation speed. Measurements included doubling time, specific growth rate (μ), viable cell counts, seeding density, and viability at the beginning and end of each run.

All data cleaning was performed in R version 4.5.1 (R Core Team, 2025). Character variables were converted to numeric where necessary (e.g., percent signs removed from viability values). Pattern and condition identifiers were converted to factor variables.

Statistical and Modeling Approach

Summary statistics were computed for doubling time, specific growth rate, and viability. Five figures were generated: a histogram of doubling times, a scatterplot of growth rate vs agitation speed, a viability comparison plot, a diagnostic plot for the linear model, and a summary table.

To assess whether agitation speed predicts specific growth rate, I applied a simple linear regression model of the form:

$$\mu_{3L} = \beta_0 + \beta_1(\text{agitation speed}) + \varepsilon$$

Model diagnostics (residual vs fitted values) were examined to confirm reasonable model behavior and to check for obvious violations of linear-model assumptions.

Results

Overall 3 L Growth Performance

Across the 16 3 L conditions, CHO cells exhibited consistent growth performance. The mean doubling time was approximately 23.14 hours, with moderate variation (SD ≈ 4.66). The average specific growth rate was 0.031 h⁻¹. Viability remained high throughout, with mean initial 98.1% and final 98.3% values indicating minimal loss of culture health over the course of the runs.
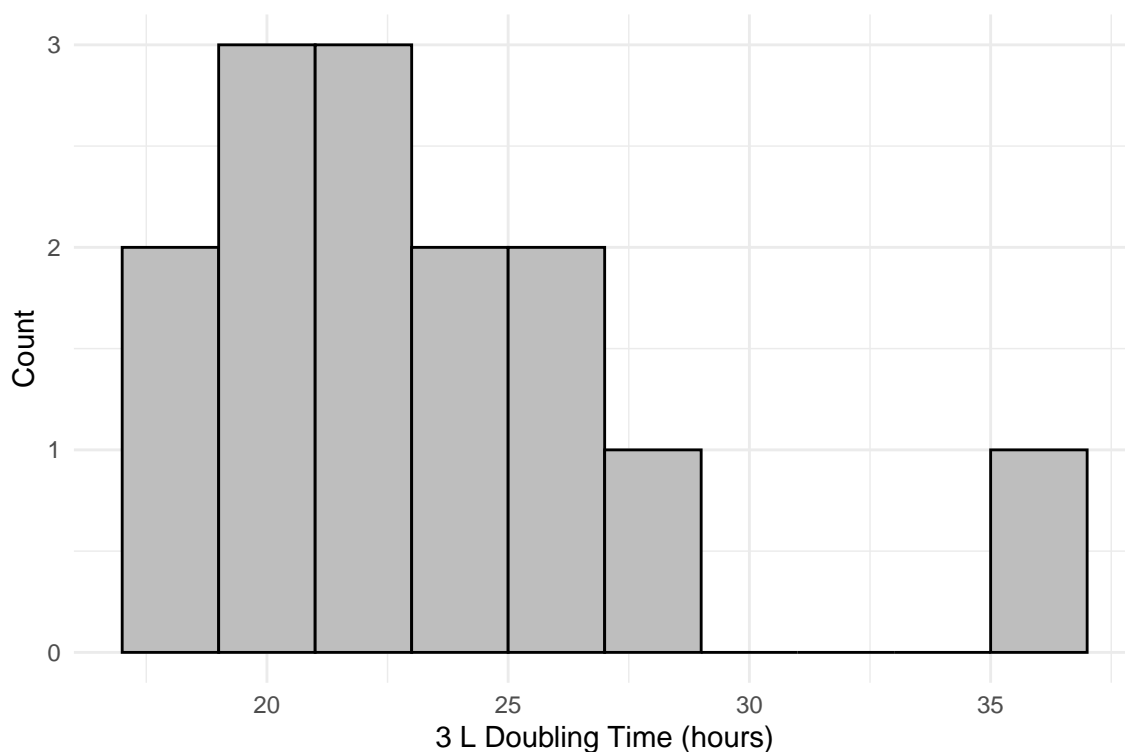

Figure 1: Distribution of 3 L doubling times.

Most 3 L doubling times clustered in a relatively narrow range around 20–25 hours, with only a few conditions showing noticeably slower growth. This pattern supports the idea that the base process is fairly robust across the tested operating window.

Relationship Between Agitation Speed and Growth Rate

Specific growth rate showed modest sensitivity to agitation speed. The Pearson correlation between agitation and µ was -0.321, indicating a modest linear association between these variables.

In general, higher agitation speeds tended to be associated with slightly higher specific growth rates, although conditions at the same rpm could still differ in performance. Different patterns (combinations of working volume and agitation) formed partially distinct clusters, suggesting that agitation alone does not fully explain growth behavior and that other process variables also contribute.
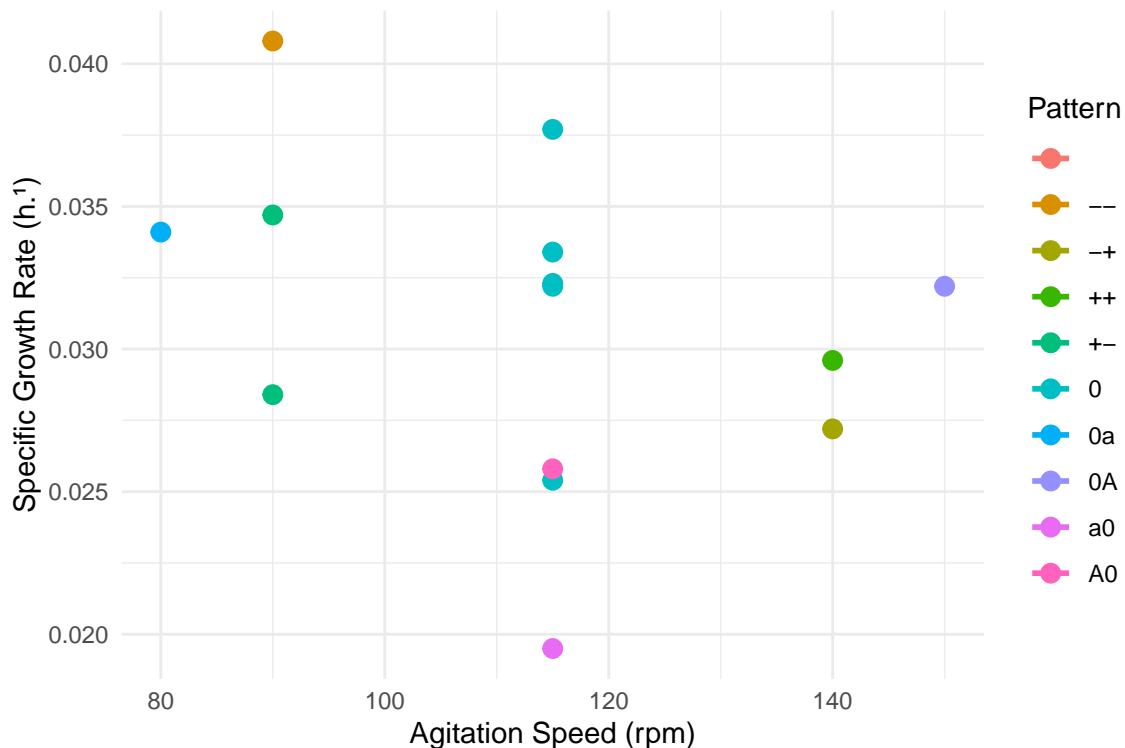
Figure 2: 3 L specific growth rate as a function of agitation speed

Viability Maintenance During Culture

Initial and final viabilities were strongly related, and most conditions maintained very high viability throughout the culture.

Most points lie near or slightly above the 1:1 dashed line, meaning that viability was generally maintained or even slightly improved over the course of the 3 L culture. Only a few conditions showed noticeable drops in viability. These outliers may correspond to suboptimal mixing, differences in seeding density, or other process-specific factors recorded in the notes.
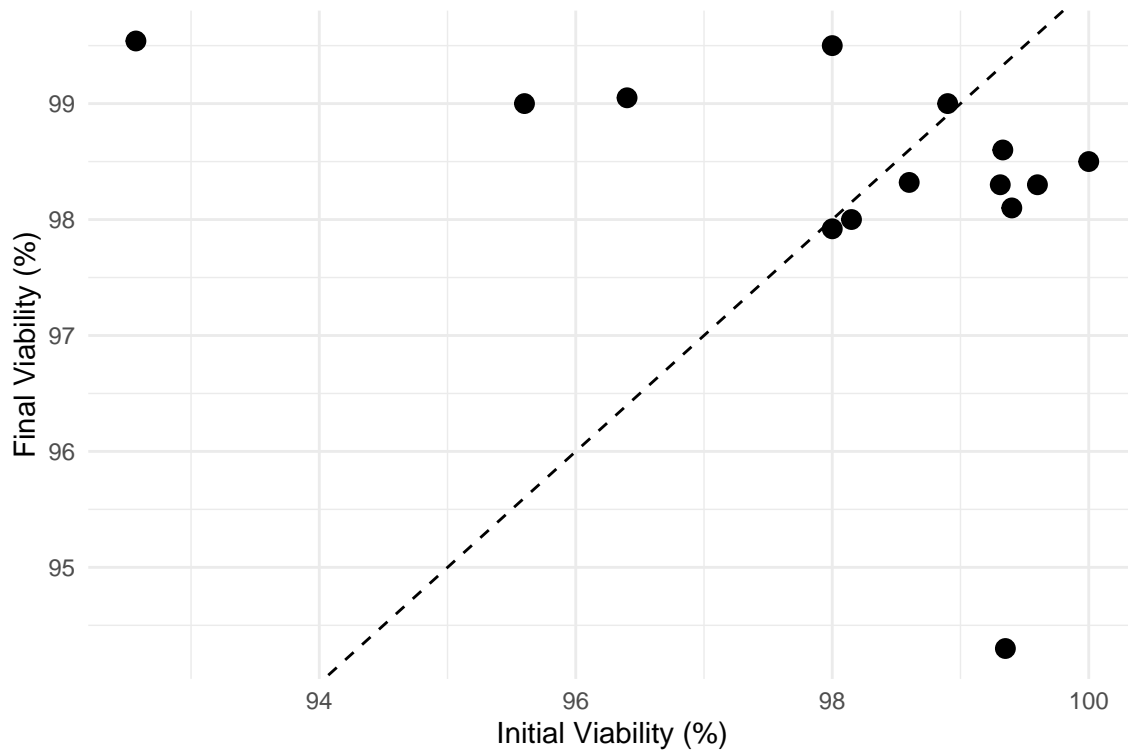
Figure 3: Initial versus final 3 L viability.

Pattern-Level Summary Table

Table 1: Summary of 3 L Performance Metrics by Pattern (mean across conditions).

pattern	n_conditions	mean_mu_3L	mean_dtime_3L	mean_viab_end_3L
	1	NaN	NaN	NaN
−	1	0.04080	17.000	99.000
-+	1	0.02720	25.450	99.000
++	1	0.02960	23.400	99.500
+−	2	0.03155	22.210	98.300
0	6	0.03220	21.898	98.548
0a	1	0.03410	20.320	99.050
0A	1	0.03220	21.520	97.920
a0	1	0.01950	35.490	94.300
A0	1	0.02580	26.910	98.320

Patterns represent groups of conditions that share similar working volume and agitation settings. Some patterns achieved relatively short mean doubling times while still maintaining high final viability, often close to or above 98%. Others showed slightly slower growth, reinforcing the idea that not all combinations of volume and agitation are equivalent, even when overall viability remains high.
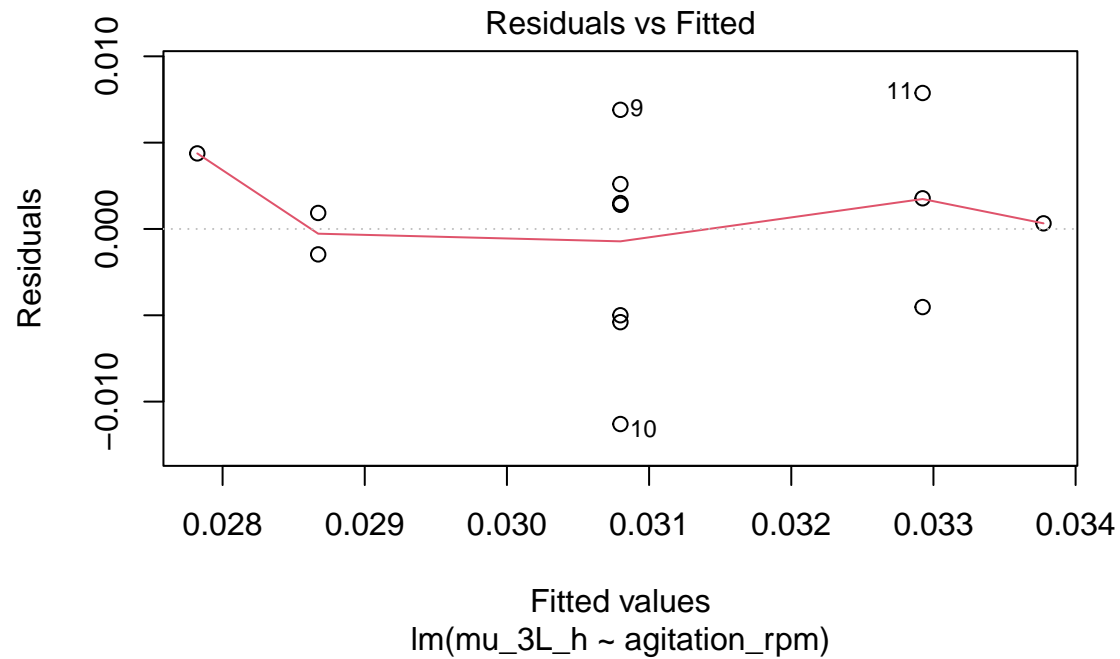
Model Diagnostics

Figure 4: Residuals versus fitted values for the linear model of μ ~ agitation speed.

The residuals-versus-fitted plot did not reveal any strong departures from linearity or obvious heteroscedasticity, suggesting that the simple linear model provides a reasonable first approximation for the relationship between agitation speed and specific growth rate in this dataset.

Discussion

This analysis demonstrates that CHO cells in the Lonza shake-flask study exhibited stable growth and excellent viability across a broad range of operating conditions. Doubling times were tightly grouped around 23 hours, and viability remained above 98% in nearly every condition. These findings indicate that the underlying cultivation process is robust and capable of supporting consistent cell expansion.

The relationship between agitation speed and specific growth rate was modest but biologically reasonable. Increased agitation can enhance oxygen transfer and mixing, which in turn can support faster growth, but culture performance also depended on working volume and other pattern-level factors not fully captured in a single linear predictor. The modeling results support the idea that agitation contributes to growth rate, and the diagnostic plot confirmed that the simple linear model behaves acceptably as an initial quantitative description.

Future analyses could incorporate additional predictors such as seeding density, passage number, or media-related variables, and more flexible models could be used to explore non-linear relationships or interaction effects. Collectively, this work illustrates how data-driven methods can be applied to evaluate bioprocess performance and guide optimization of small-scale CHO culture systems.

Code Availability

All R Markdown source files, analysis scripts, and the Lonza CHO dataset used in this project are available at the following GitHub repository:

https://github.com/DaminiUNH/Final_Project_Lonza_CHO

Bibliography

- Butler, M. (2005). Animal cell cultures: recent achievements and perspectives in the production of biopharmaceuticals. Applied Microbiology and Biotechnology, 68(3), 283–291.
- Jayapal, K. P., Wlaschin, K. F., Hu, W.-S., & Yap, M. G. S. (2007). Recombinant protein therapeutics from CHO cells—20 years and counting. Chemical Engineering Progress, 103(10), 40–47.
- R Core Team (2025). R: A Language and Environment for Statistical Computing. Version 4.5.1. R Foundation for Statistical Computing.
- Cytiva. HyClone SFM4CHO Medium: Product Information and Preparation Guide. Cytiva Life Sciences, n.d.