

# Optimized Decision Making About Library Shelf Space Through Simulation



DATA 604: Simulation and Modeling Techniques

Daina Bouquin, John DeBlase, James Topor

# Background

- *Shelf shifting*: free space in a library's collection gets divided up and redistributed amongst shelves to distribute books more evenly throughout the collection
  - Small or large scale shifts
  - Mitigated by *weeding* and *de-duplication*, but shelves still overfill
    - Shifting is time-consuming and therefore expensive

# Off-Site Storage

- Large library systems have adopted *off-site storage* models to deal with the overflow
  - portion of the physical collection is stored remotely and is accessed via courier
  - cost can be substantial



# The Research Question

What is the ideal amount of space to leave empty on a book shelf to minimize the need to shelf shift while accommodating the largest possible quantity of books?



If a library can improve its ability to predict when a shelf shifting initiative will be required, the library may be able to take steps to minimize the need for shelf shifting and more effectively plan and budget for any required shelf shifts when they do occur.

# **Proof of Concept:**

**Use simulation to answer our research  
question on a small virtual library collection**

Virtual library is based on Wolbach Library at the  
Harvard-Smithsonian Center for Astrophysics

<http://library.cfa.harvard.edu/>

# Assumptions

Assumptions	Details
Number of Books	Total of 1,000 books Represents a single, relatively small, collection within the broader library collection.
Types of Books	Textbooks and non-textbooks 5% of the collection is assumed to be master copies of textbooks Number of copies of a textbook determined by randomly sampling from a uniform distribution bounded by (1, 3) - Non-textbooks are not duplicated within the library.a uniform distribution bounded by (1, 3) - Non-textbooks are not duplicated within the library.
Book Widths	Varies between 1 and 2 inches according to a triangular approximation of a normal distribution. Based on architectural graphics standards used in library design to control for differences between hardcover and paperback books.
Shelves	Fifty (50) 36 inch shelves are available - typical amount of shelving designated for a small subcollection within Wolbach Library Books are assigned to the simulated shelves in a serial manner until a user-specified minimum amount of free space remains on a given shelf - when minimum amount is achieved, shelving continues on the next available shelf. Any volumes not shelvable due to space constraints are sent "off-site" for the purposes of the simulation.
Books in Circulation	10% of the collection has been checked out for use by library patrons prior to the simulation start - based on averages of monthly "COUNTER" statistics pulled from Harvard's Cognos reporting tool for Wolbach's non-periodical collection. Collection circulation rates tend to be seasonal in nature and are not typically tracked at the level of a subcollection. The width of any book that is in circulation is added back to the amount of free space available on the shelf.
Percentage of Books Weeded	2% of the simulated collection is flagged for weeding during each round of the simulation as per Wolbach's in-development weeding policy If a book is checked out at the time weeding it is considered exempt from weeding for that round of the simulation. This ensures that circulating books are not weeded.

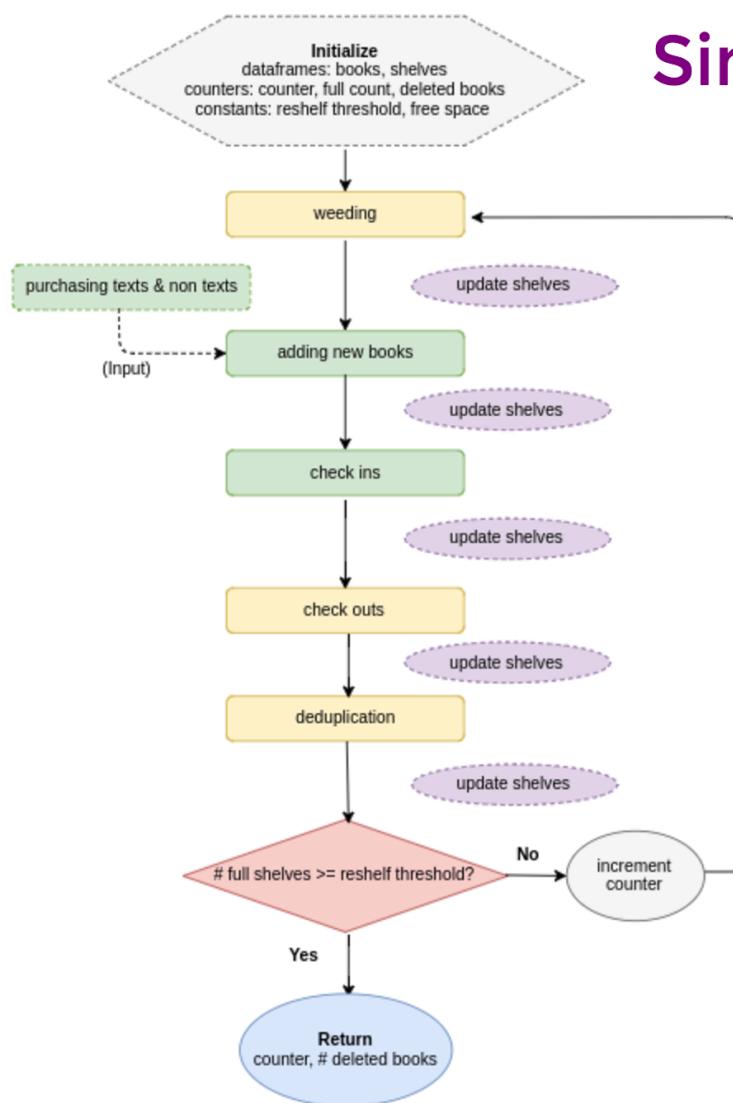
# Assumptions Cont.

Number of Textbooks De-Duplicated	We determine the number of textbooks to be de-duplicated during each round of the simulation by sampling from a uniform (1, 5) bound distribution - based on typical variability of new editions of textbooks being released.  A corresponding number of textbooks is then randomly selected from the simulated collection for de-duplication.  During de-duplication, all copies of the selected textbooks are removed from the collection. However, a master copy is retained.  If a copy of a textbook is flagged for de-duplication while it is in circulation, it is not immediately removed, instead it is flagged for removal when it is returned.
Percentage of Books Checked In	20% of books currently in circulation are returned to the library during each round of the simulation - based on average circulation rates at Wolbach (again these tend to be seasonal)  Books to be checked back in are selected via simple random sampling of the books currently in circulation.
Percentage of Books Checked Out	2% of books currently not in circulation are randomly selected for check out during each round of the simulation.
New Textbook Acquisitions	3 - 6 new textbooks and associated copies are added to the collection during each round of simulation, with the exact number determined via random sampling from a uniform (3, 6) bound distribution - Wolbach acquisition rates  The number of copies to be added to the master copy for each new textbook purchased is determined by sampling from a uniform (1,3) bound distribution.  Non-full shelves are randomly selected for each new textbook and associated copies are colocated on the same shelf.

# Assumptions Cont.

Amount of Initial Free Shelf Space	<p>The simulation examines a series of initial free shelf space values ranging from 3 inches up to 6 inches in 0.5 inch increments. (NOTE: Six inches was selected as the upper bound for the range to coincide with the "width of two human hands" shelf free space guideline espoused in many anecdotal library management discussions).</p> <p>The amount of initial free shelf space is used during shelving of the virtual library collection at the start of each round of the simulation, with each of the 50 shelves guaranteed to have no less than the indicated amount of free space.</p>
Shelf Filling	<p>A shelf is considered "full" if at least 95% of its available space has been consumed. When a shelf surpasses 95% usage, no further new textbook acquisitions may be added to that shelf.</p> <p>Books from such a shelf that are in circulation when the 95% usage boundary is surpassed are returned to the shelf whenever they are returned to the library despite the shelf having been deemed "full" for other purposes.</p>
Shelf Shift Required	A shelf shift is assumed to be required whenever at least 5 shelves achieve "full" status.

# Simulation Logic



# Results

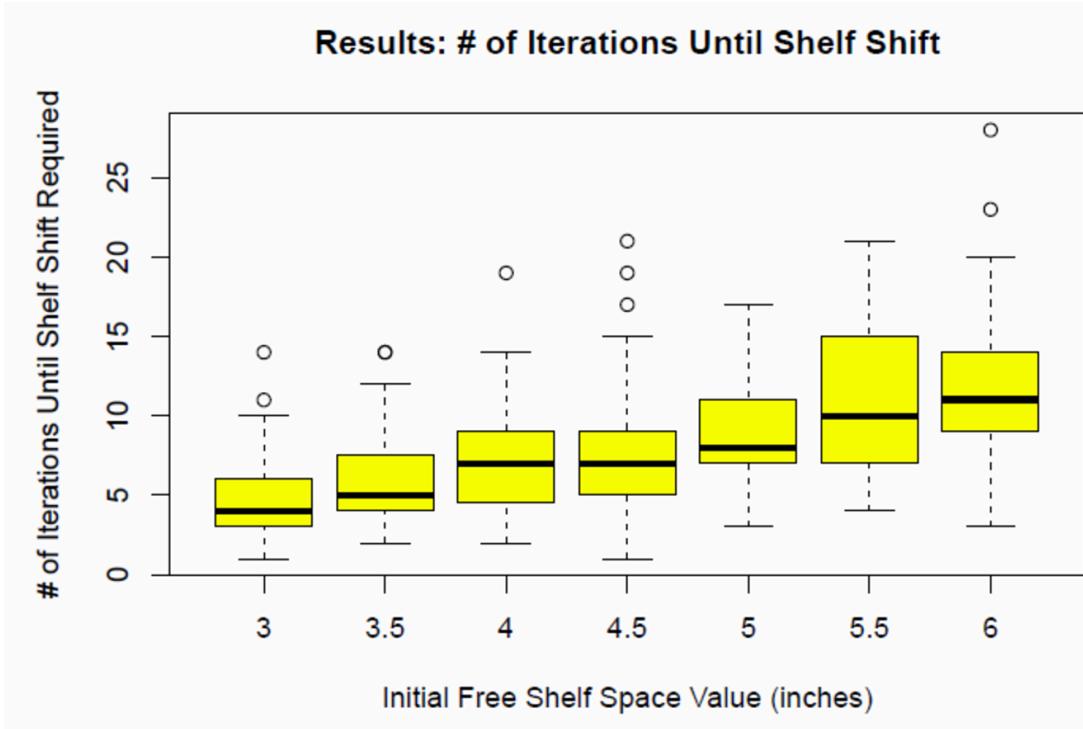


Table 1: Number of Iterations Until Shelf Shift Needed

Inches	mu	mdn	min	max	sd	c_int
3.0	4.77	4	1	14	2.411	(4.297, 5.243)
3.5	5.82	5	2	14	2.560	(5.318, 6.322)
4.0	7.05	7	2	19	3.141	(6.434, 7.666)
4.5	7.68	7	1	21	3.570	(6.980, 8.380)
5.0	9.17	8	3	17	3.172	(8.548, 9.792)
5.5	11.15	10	4	21	4.375	(10.293, 12.007)
6.0	11.17	11	3	28	4.120	(10.363, 11.977)

Table 2: Number of Books Sent Off Site

Inches	mu	mdn	min	max	sd	c_int
3.0	0.00	0	0	0	0.000	(0.000, 0.000)
3.5	0.00	0	0	0	0.000	(0.000, 0.000)
4.0	0.00	0	0	0	0.000	(0.000, 0.000)
4.5	0.00	0	0	0	0.000	(0.000, 0.000)
5.0	0.15	0	0	7	0.857	(-0.018, 0.318)
5.5	8.87	9	0	24	5.733	(7.746, 9.994)
6.0	25.15	25	10	37	5.511	(24.070, 26.230)

# Discussion

- The model has a high degree of face validity
- Flexible protocol that can be adapted to suit any library subcollection
- Informs librarians managing physical collections about what metrics will be needed to take advantage of such a simulation
- Novel tool and method previously unexplored in library science
- Informs decision making and planning capabilities and is accessible to libraries with very small budgets

# Future Work

- Scalability - method could be applied to larger and more complex library collections
  - Current simulation represents 1/60th of Wolbach's collection
    - Library collections are typically much larger than 1000 volumes
  - We were limited by processing power and granularity of collection development metrics
- Augment the model to reflect public library collections where textbooks are not a substantial portion of the collection
- Factor in seasonality of circulation metrics and decreases in physical acquisitions due to shifts toward ebook purchasing