Digital-twin to improve performance of a meal-kit delivery company

Dr Kean Dequeant, Senior Data Scientist, Gousto

Anylogic Conference, 2024

Summary

- Context: Gousto, a scale-up Tech company
- Why simulation models
- Self-serve simulation
- Use-case: improving Order Routing Algorithm

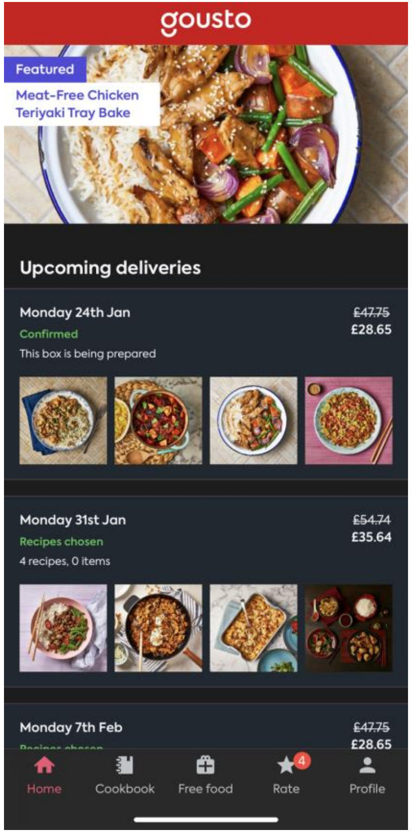
Context: Gousto, a scale-up Tech company

Context: Gousto, a scale-up Tech company

From start-up to scale-up

308 million

Revenue in 2023



Context: Gousto, a scale-up Tech company

What sets us apart?

Choice

250+ Monthly Recipes

Convenience

Recipes created, prepped and delivered

Waste

B-corporation as of 2021

Pricing

From £2.99 per portion

Tech

Extreme focus on data and artificial intelligence

Interface

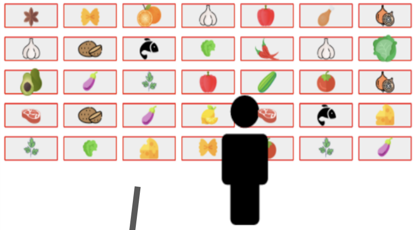
App + web products and dedicated CC team

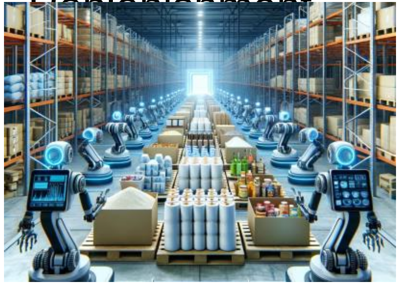
'We never stop obsessing over the customer'

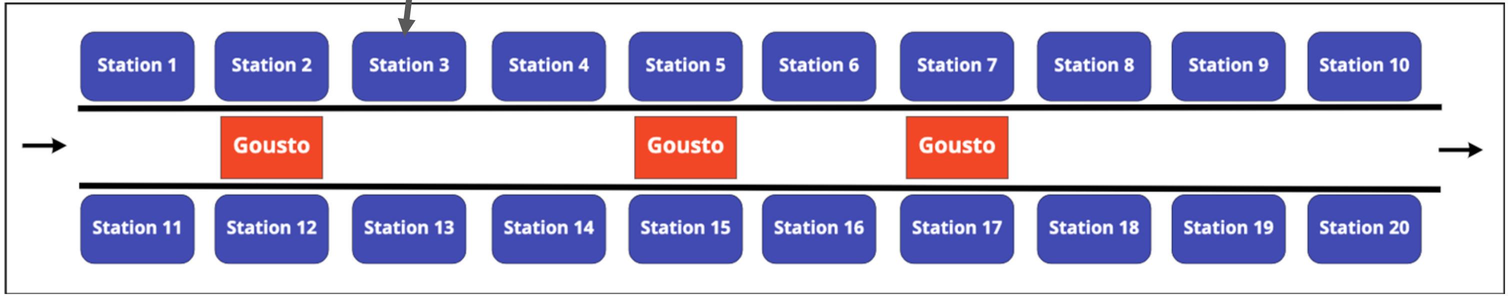
Context: Gousto, a scale-up Tech company

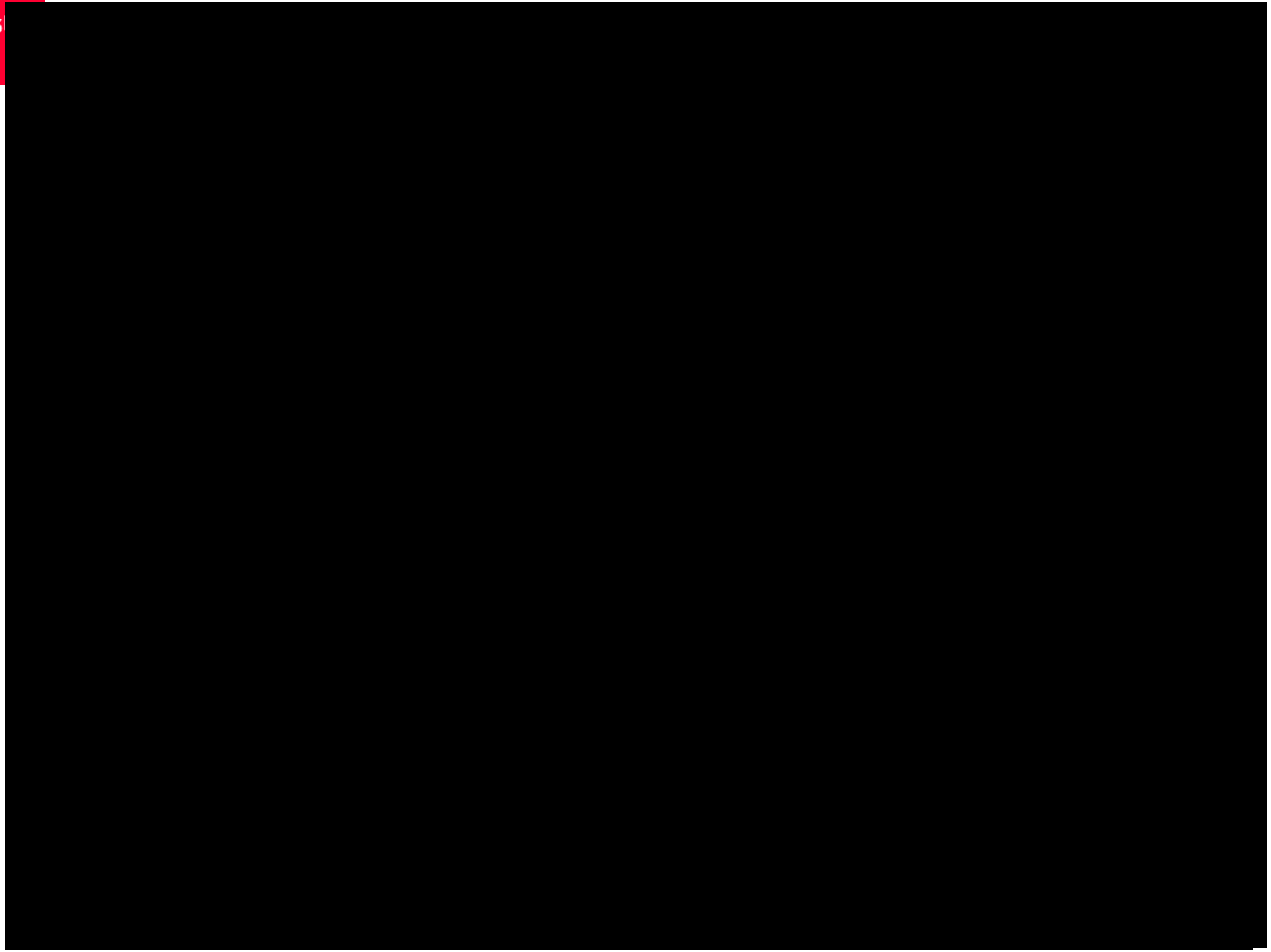
How do you make a recipe box?

line optimisations algorithms = Pick Face Optimisation + Dynamic Replenishment + Order Routing Algorithm







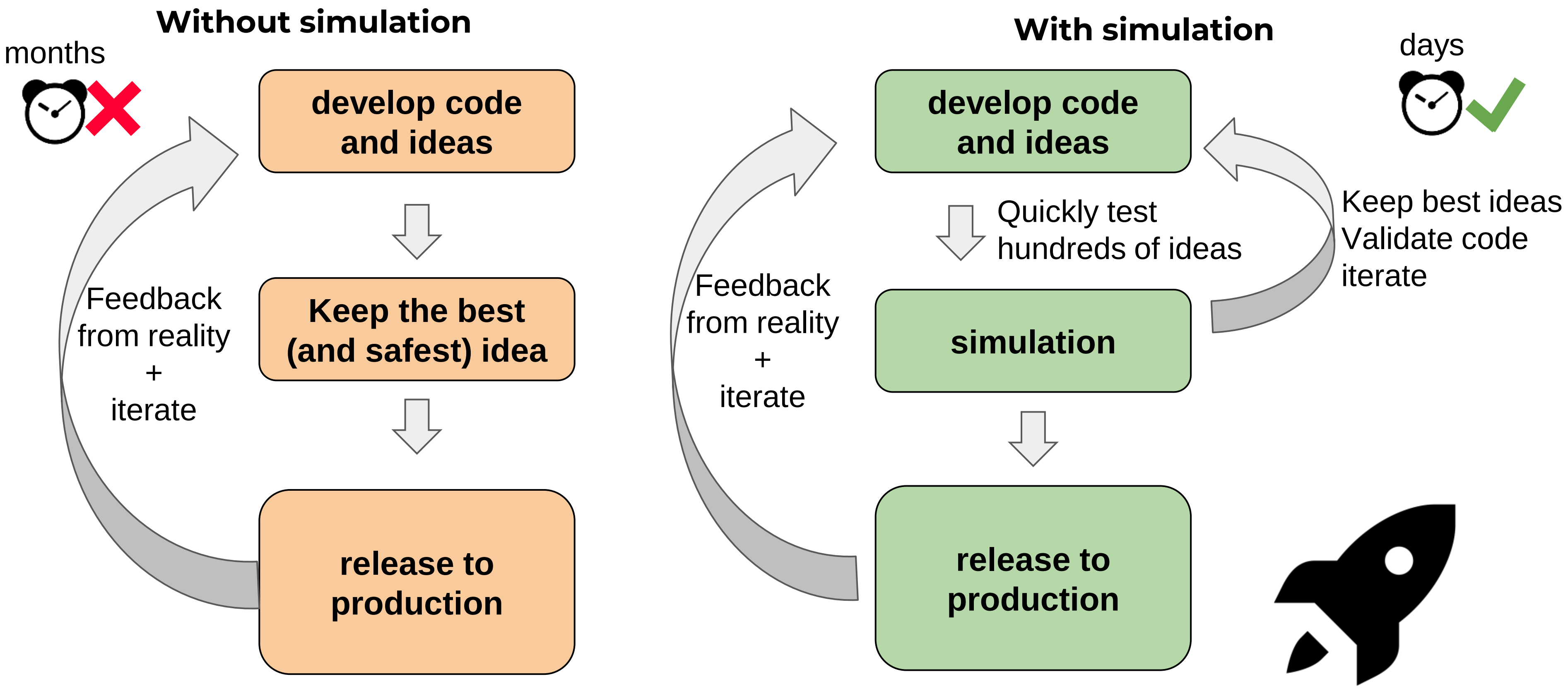


Pick Station front

Pick Station back

Why simulation models?

Why Factory simulation model?



Why Anylogic?

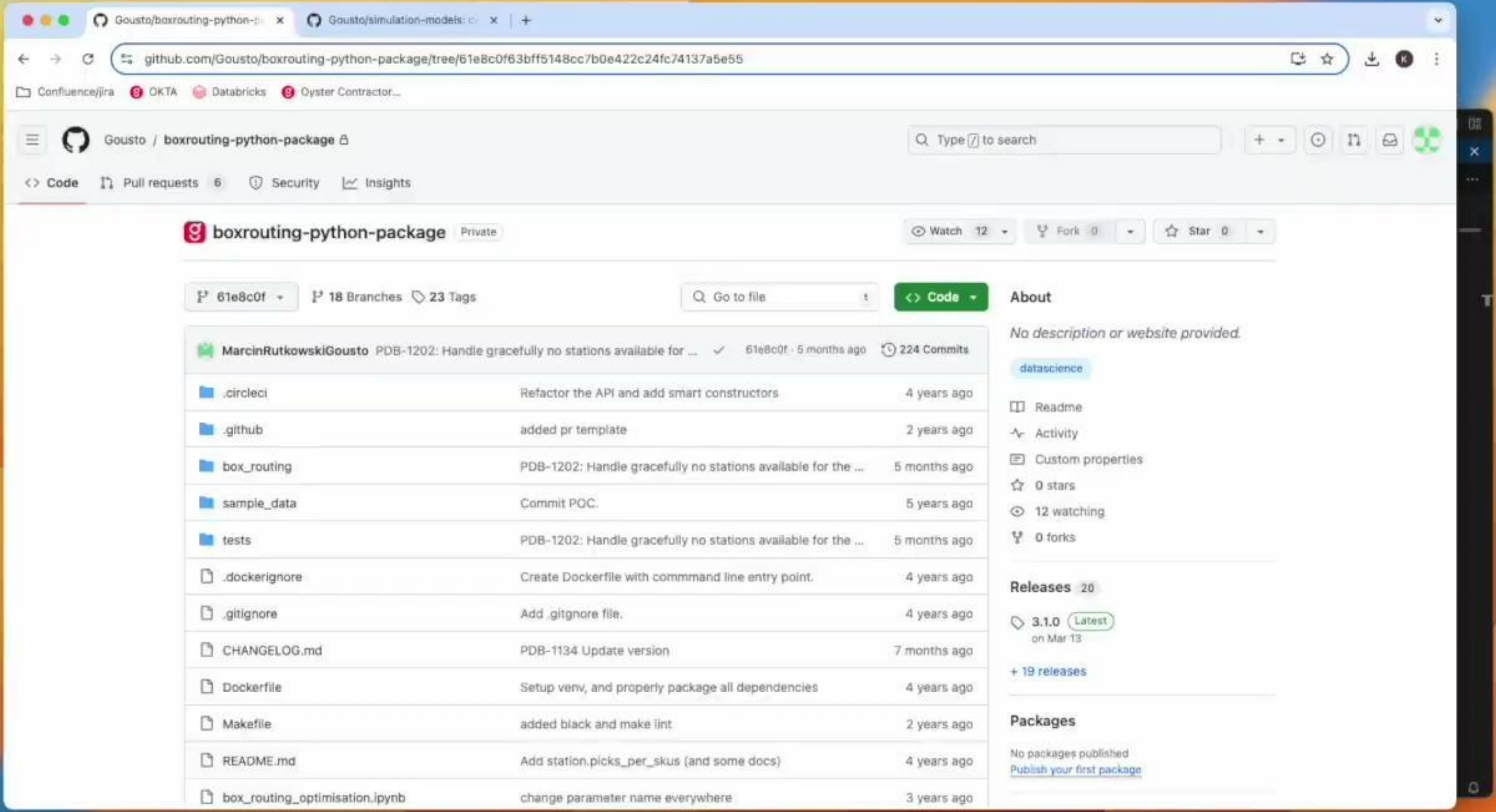
Our simulation requirements:

Self-serve models	Multiple operating systems	Integrating our python algorithms
↓	↓	↓
Anylogic “standalone model export”	Anylogic on : Mac + Windows + Linux	Anylogic Pypeline

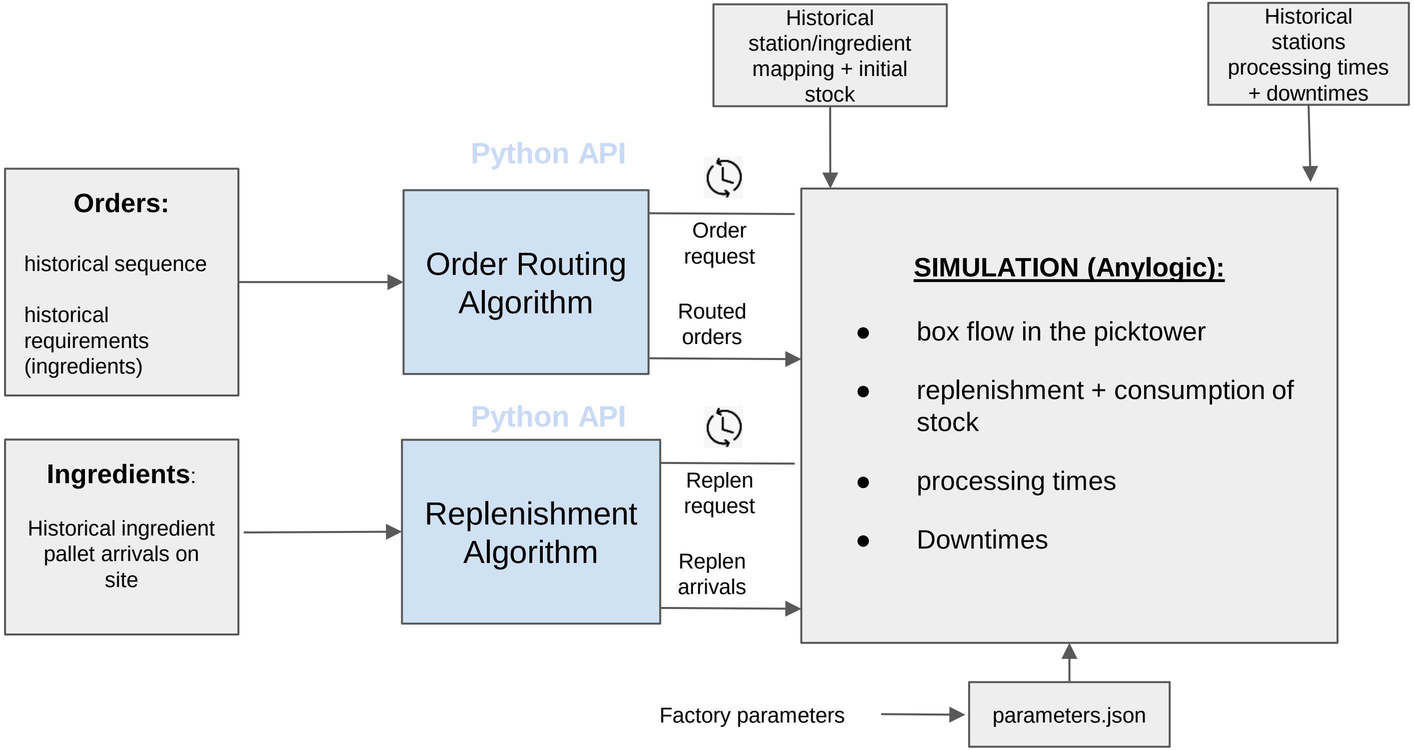
Self-serve simulation

End-to-end code integration

Code change → Update simulation → Run A/B experiment campaign → Prove the benefits & get approval → Release to production


Gousto/boxrouting-python-p
Gousto/simulation-models
github.com/Gousto/boxrouting-python-package/tree/61e8c0f63bff5148cc7b0e422c24fc74137a5e55
Confluence/jira
OKTA
Databricks
Oyster Contractor...
Gousto / boxrouting-python-package
Type / to search
Code
Pull requests 6
Security
Insights
boxrouting-python-package
Private
Watch 12
Fork 0
Star 0
61e8c0f
18 Branches
23 Tags
Go to file
Code
MarcinRutkowskiGousto PDB-1202: Handle gracefully no stations available for ...
61e8c0f · 5 months ago
224 Commits
.circleci
Refactor the API and add smart constructors
4 years ago
.github
added pr template
2 years ago
box_routing
PDB-1202: Handle gracefully no stations available for the ...
5 months ago
sample_data
Commit POC.
5 years ago
tests
PDB-1202: Handle gracefully no stations available for the ...
5 months ago
.dockerignore
Create Dockerfile with commmand line entry point.
4 years ago
.gitignore
Add .gitignore file.
4 years ago
CHANGELOG.md
PDB-1134 Update version
7 months ago
Dockerfile
Setup venv, and properly package all dependencies
4 years ago
Makefile
added black and make lint
2 years ago
README.md
Add station.picks_per_skus (and some docs)
4 years ago
box_routing_optimisation.ipynb
change parameter name everywhere
3 years ago
About
No description or website provided.
datascience
Readme
Activity
Custom properties
0 stars
12 watching
0 forks
Releases 20
3.1.0
Latest
on Mar 13
+ 19 releases
Packages
No packages published
Publish your first package

Latest factory simulation model



Simulation usage statistics

- First Anylogic licence bought in October 2022 (less than 2 years ago)
- Users: 3 squads
- 30+ use-cases
- Target of 10 use-cases this quarter
- 5000+ simulation runs
- 10 million+ boxes simulated
- 3 laptops destroyed

Factory simulation use-case: improving Order Routing Algorithm

What is ORA?

ORA = Order Routing Algorithm

Sets:

P = The set of all pick stations
S_o = The set of SKUs in order o
P_s = The set of pick stations containing sufficient stock of SKU s
P_{S_o} = The set of pick stations containing any SKU in S_o
G = The set of station groupings
L = The set of levels

Data:

w_p = The workload cost of pick station p
N_{s_o} = The number of units of SKU s required for order o
N_{s_p} = The number of units of SKU s available at pick station p
L_t = The target level
α = grouping cost weighting
β = level cost weighting

Variables:

$$x_p = \begin{cases} 1 & \text{if pick-station } p \text{ is visited} \\ 0 & \text{otherwise} \end{cases}$$

$$x_g = \begin{cases} 1 & \text{if a pick station in grouping } g \text{ is visited} \\ 0 & \text{otherwise} \end{cases}$$

$$x_l = \begin{cases} 1 & \text{if a pick station in level } l \text{ is visited} \\ 0 & \text{otherwise} \end{cases}$$

Minimise:

$$\sum_{p \in P_s} w_p x_p + \overline{w_p}\left(\alpha \sum_{g \in G} x_g + \beta \sum_{l \neq T_l} x_l\right)$$

Subject to:

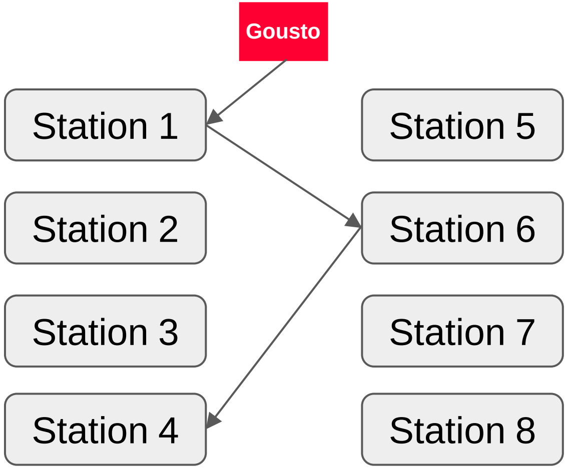
$$\sum_{p \in P_s} x_p \geq 1, \forall s \in S_o$$

$$\sum_{g \in G} x_g \geq 1$$

$$\sum_{l \in L} x_l \geq 1$$

Decision = For each box, which ingredients are picked from which stations.

Objective = maximise throughput



ORA V4 vs ORA V5

ORA V4	vs	ORA V5
Cold start (empty stations)		Hot start (initialise stations workload from production)
Batch size ~700		Batch size ~100
Assume perfect stock coverage		Live production stock coverage
Manually defined factory parameters		Parameter search through simulation

Simulation replay: benchmarking our model accuracy

Simulation replay Accuracy:

$$Accuracy = 1 - \frac{\int_{t_a}^{t_b} |f(t) - g(t)|\, dt}{\int_{t_a}^{t_b} f(t)\, dt}$$

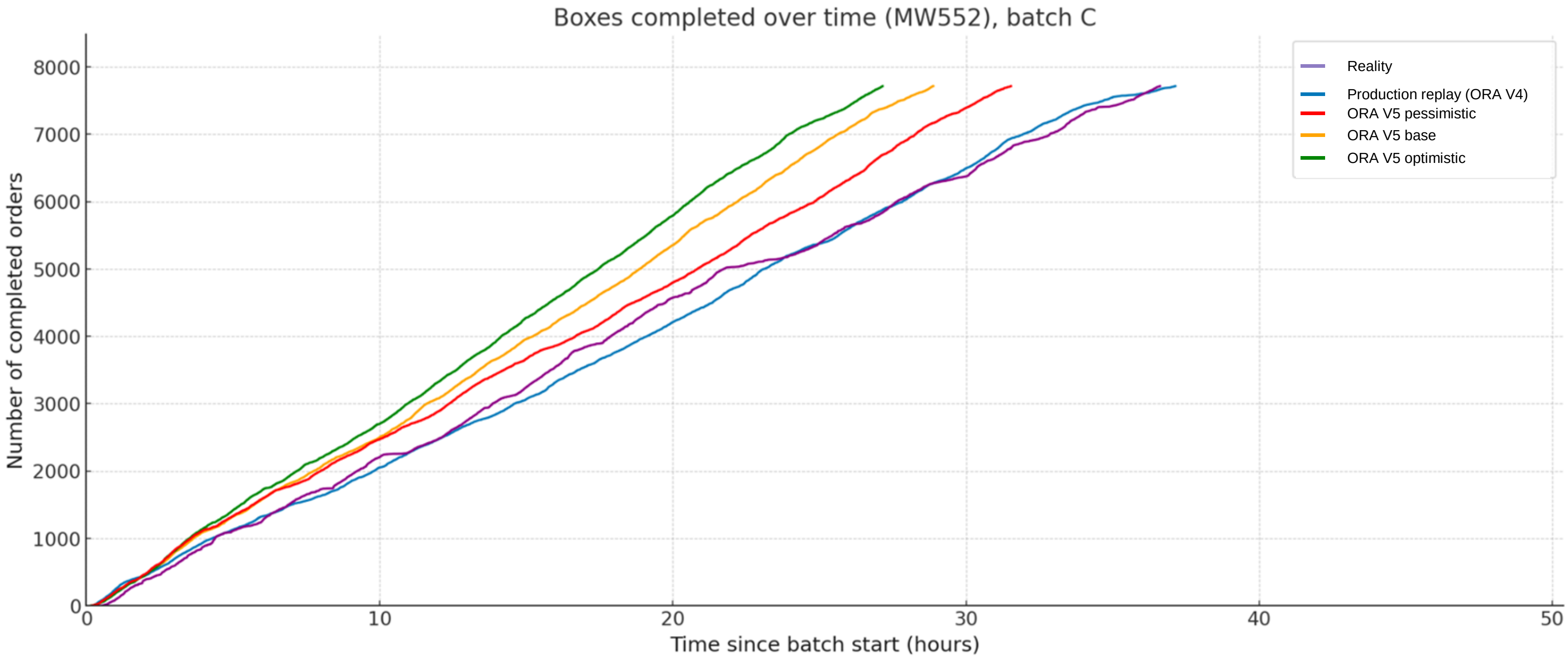
where $f(t)$ = *actual cumulative picks*
$g(t)$ = *simulation cumulative picks*

4 different scenarios

	Scenario 1: ORA V4, prod replay	Scenario 4: ORA V5, pessimistic	Scenario 2: ORA V5, base	Scenario 3: ORA V5, optimistic
Initialisation:	Cold start	Live data	Live data	Live data
Batch Size:	700	700	100	100
Factory parameters:	Unchanged	Unchanged	Unchanged	Simulation Optimised

Use-case results

Simulation V3 results: single run



Full results

Simulation results (5 consecutive weeks):

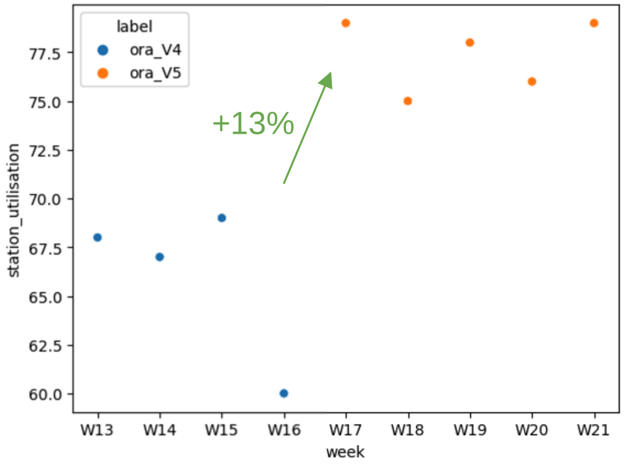
	Scenario 1: ORA V4, prod replay	Scenario 4: ORA V5, pessimistic	Scenario 2: ORA V5, base	Scenario 3: ORA V5, optimistic
Station utilisation	69%	76%	83%	88%
Normalised Throughput	100	107	117	126

Post deployment in production:

Scenario deployed: hybrid base/pessimistic

Station utilisation before: 67%

Station utilisation after: 76%



Perspectives

Extend the capabilities of our simulation model

- Extend 3D integration to end-to-end flow of stock and boxes
- Extend granularity of events

Extend the use of simulation

- Extend user base
- Real-time simulations (hot start)

Extend beyond single factory

- Simulation of second factory
- Network simulator

Thank you for your attention!