Week 8: Selecting and Revalidating Operating Conditions | Worked Companion

Cost reduction subject to specifications

Reject candidates that violate product specifications even when their cost is low. Recalculate candidates under identical reactor equations and specifications to compare acceptance and cost.

The baseline is 330 K and 6 min, giving B = 0.409091 and C = 0.245455 mol/L. It is reference-feasible and costs 76.80 $/h.

Retaining it during comparison preserves a checked option when a new model or solve fails.

Deriving the steady-state material balances

$$0 = \frac{C_{A0} - C_A}{\tau} - k_1 C_A$$

$$0 = -\frac{C_B}{\tau} + k_1 C_A - k_2 C_B$$

$$0 = -\frac{C_C}{\tau} + k_2 C_B$$

These are per-volume balances using $\tau = V/F$. Solve the first for A, the second for B, and the third for C to obtain the lecture's reference equations.

Summing the equations gives $C_A + C_B + C_C = C_{A0}$, under the equal-mole consecutive-reaction and constant-density assumptions.

Hand calculation of the baseline

At 330 K, $k_1 = 0.2$ and $k_2 = 0.1$ min^{-1}. Substitute $\tau = 6$ min.

$$C_A = 1.2 \,/\, \left(1 + .2 \times 6\right) = .545455$$
$$C_B = .2 \times 6 \times .545455 \,/\, \left(1 + .1 \times 6\right) = .409091$$
$$C_C = .1 \times 6 \times .409091 = .245455$$

This satisfies $B \geq 0.38$ and $C \leq 0.25$. Cost is $12 \times 6 + 0.24 \times 20 = 76.80$ \$/h.

Deriving one-sided margins

Define error as e = prediction − reference. If overprediction e_B is bounded by δ_B everywhere in the domain, then

$$C_B = \hat{C}_B - e_B \geq .38 + \delta_B - \delta_B = .38$$

For C, the required bound is on underprediction, $-e_C \leq \delta_C$. Then $\hat{C}_C \leq 0.25 - \delta_C$ gives $C_C = \hat{C}_C - e_C \leq 0.25$.

This argument assumes uniform bounds. Maxima on a finite calibration sample do not prove that assumption.

Calculating one-sided margins

Define $e=\text{prediction}-\text{reference}$. Positive e is adverse for B; negative e is adverse for C.

Quantity	B lower limit	C upper limit
Original specification	Reference $B \geq 0.38$	Reference $C \leq 0.25$
Applied margin	$\delta_B = 0.02130$	$\delta_C = 0.01305$
Required prediction	$\hat{B} \geq 0.40130$	$\hat{C} \leq 0.23695$
At the assumed error limit	$0.40130 - 0.02130 = 0.38$	$0.23695 + 0.01305 = 0.25$

The last row applies when the error limit holds at that input. The maximum over 961 calibration points does not prove a bound at every unseen point. Reevaluate the next candidate with the reference process even after adding margins.

Sample margins are $\delta_B = \max(0, \max_i e_{B,i})$ and $\delta_C = \max(0, \max_i(-e_{C,i}))$. B errors $[-0.003, 0.008, 0.002]$ give 0.008; C errors $[0.004, -0.006, -0.001]$ give 0.006. Unlike an absolute-error maximum, these select only the direction that can violate the specification.

Data partitions and reference-evaluation cost

Data	Size and role
Existing training	Week 6 network, 1,200 training points
Independent evaluation	997 inputs, seed 808, full-box prediction error
Calibration	$31 \times 31 = 961$, margin calculation
Local update	$3 \times 3 = 9$, affine residual fitting
Candidate revalidation	Fresh reference evaluation at each selected input
Comparison search	4,001 T values and local scalar refinement

Do not mix post-fit candidate revalidation data into independent test results. With expensive process models, calibration and benchmark costs also belong in the evaluation budget.

Error signs at a selected point

At the unadjusted candidate, predicted B is 0.380000 and reference B is 0.371157 mol/L. Thus e_B = +0.008843, the adverse overprediction direction.

Predicted C is 0.171228 and reference C is 0.178136 mol/L. Although e_C = −0.006908 is adverse underprediction, reference C remains below 0.25.

Error signs alone do not determine acceptance. Compare error magnitudes with the remaining margin to each constraint.

Local update coordinates and residual regression

The unadjusted candidate has scaled coordinates approximately (0.80148, 0.06556). Offsets ±0.06 correspond to ±2.4 K and ±0.48 min.

$$r_i = C^{ref}\left(x_i\right) - f_\theta\left(x_i\right)$$

$$\min_\beta \sum_{i=1}^{9} \left\| r_i - \left[1, x_{i1}, x_{i2}\right]\beta \right\|_2^2$$

Use the same input matrix for every output component. Adding an affine output term preserves the ReLU binary structure. Optimize the corrected model only inside the update-data box.

Local-data coordinates and the regression matrix

Candidate (342.0593 K,2.52444 min) has scaled coordinates [(342.0593−310)/40, (2.52444−2)/8]≈[0.80148,0.06556].

Axis	Scaled offset	Physical offset
Temperature	±0.06	±2.4 K
Residence time	±0.06	±0.48 min

```
design = np.column_stack((np.ones(len(x_local)), x_local))
residual = reference_y - network_y
beta = np.linalg.lstsq(design, residual, rcond=None)[0]
corrected_y = network_y + design @ beta
```

For nine inputs, design is (9,3), the three-component residual is (9,3), and beta is (3,3). Multiplying a row $[1, x_1, x_2]$ by beta gives three corrections. Restrict the next optimization to the fitted region and check remaining error at the new candidate.

Substituting the fitted correction coefficients

The recorded B correction coefficients are [−0.0160962605,0.0074357617,0.0823403944]. At local candidate (339.815193 K,2.870934 min), x_1=0.745379824 and x_2=0.108866748.

B correction term	Value [mol/L]
Intercept	−0.0160962605
Temperature-coordinate term	0.0074357617×0.745379824=0.0055424667
Residence-time-coordinate term	0.0823403944×0.108866748=0.0089641309
Sum	−0.0015896628

Adding this to the original network prediction B≈0.3835896628 gives corrected prediction 0.382000. The reference value at the same input is 0.3768201356, leaving error 0.0051798644.

Least squares reduces the summed squared errors of nine points. It does not require interpolation of every point or exact prediction at a new candidate. Correctly computed regression coefficients still require candidate revalidation.

Follow-up data from a failed correction

B error at the local-correction candidate is 0.005180 mol/L, exceeding the applied 0.002 margin. Although its prediction is 0.002 above the specification, its reference output falls short by 0.003180.

Add the failed point’s reference output to the update data and examine residual changes with temperature and residence time. A smaller region or larger margin are other candidate revisions.

Each revision defines a new model or problem. Revalidate against the same original specifications and retain a checked feasible candidate until a revision succeeds.

Calculating residence time from the product specification

At fixed T, setting C_B to target b gives a quadratic equation in τ.

$$C_B = \frac{1.2k_1\tau}{(1+k_1\tau)(1+k_2\tau)}$$

$$bk_1k_2\tau^2 + \left[b\left(k_1+k_2\right) - 1.2k_1\right]\tau + b = 0$$

With a nonnegative discriminant and two positive roots, B ≥ b between the roots. The τ bounds and C upper limit must also hold.

The smaller root is preferred because cost increases with τ at fixed T.

From the B specification to a quadratic in residence time

At fixed temperature, k_1 and k_2 are numbers. Set B to target b and multiply by the positive concentration denominator.

Step	Relation
Set B=b	$1.2k_1\tau = b(1+k_1\tau)(1+k_2\tau)$
Expand denominator	$1.2k_1\tau = b[1+(k_1+k_2)\tau+k_1k_2\tau^2]$
Collect terms	$bk_1k_2\tau^2+[b(k_1+k_2)-1.2k_1]\tau+b=0$

At 330 K and b=0.4, this is $0.008\tau^2-0.12\tau+0.4=0$. Dividing by 0.008 gives $(\tau-5)(\tau-10)=0$. $B\geq 0.4$ holds between the roots, [5,10].

At fixed temperature, cost increases with τ, so check 5 min first. It gives $C=0.1\times5\times0.4=0.2\leq0.25$. At 10 min, C=0.4 violates the cap; the B condition alone is insufficient.

The smaller root and a stable numerical formula

$$a = bk_1k_2, d = b\left(k_1 + k_2\right) - 1.2k_1$$
$$D = d^2 - 4ab$$
$$\tau_{low} = \frac{2b}{-d + \sqrt{D}}$$

For a > 0, d < 0, D ≥ 0, this computes the smaller positive root while avoiding subtraction of similar positive quantities in the standard formula.

Take max(2, τ_{low}), then check the larger root, the 10 min upper bound, and the C specification. No roots or an exceeded upper bound means no feasible candidate at that T.

Rationalizing the smaller-root formula

For $a\tau^2+d\tau+b=0$, define $D=d^2-4ab$. With $a>0, d<0, D\geq 0$, the smaller root is $(-d-\sqrt{D})/(2a)$. Multiply numerator and denominator by $-d+\sqrt{D}$.

$$\frac{-d-\sqrt{D}}{2a}=\frac{d^2-D}{2a(-d+\sqrt{D})}=\frac{2b}{-d+\sqrt{D}}$$

Since $d^2-D=4ab$, cancellation gives the last expression. Replacing subtraction of similar positive values $-d$ and $\sqrt{D}$ with addition reduces loss of significant digits.

At 330 K, $b=0.4$, $a=0.008$, $d=-0.12$, $D=0.0016$ and $\sqrt{D}=0.04$. The smaller root is $0.8/(0.12+0.04)=5$; the larger is $(0.12+0.04)/0.016=10$ min. Substitute roots back into the original B equation and C specification before accepting a candidate.

Calculating the reference-repaired candidate

Fix the local-correction candidate’s temperature at 339.8152 K and use b = 0.382. The smaller root is τ = 3.08559 min.

Reference outputs are A = 0.618470, B = 0.382000, and C = 0.199530 mol/L. They sum to 1.2 and satisfy the C upper bound.

Cost is 44.1828 $/h, which is 32.6172 $/h below the initial baseline. This point leaves the local optimization box, so report it separately as a repair checked directly with the reference equations.

Computing the smaller root and revalidating the candidate

At $T=339.8152$ K and $b=0.382$, calculate $a=bk_1k_2$ and $d=b(k_1+k_2)-1.2k_1$. A negative discriminant means there is no real residence time with $B=b$ at this temperature.

```
k1 = 0.2 * np.exp(40000 / 8.314 * (1/330 - 1/T))
k2 = 0.1 * np.exp(50000 / 8.314 * (1/330 - 1/T))
a = b * k1 * k2
d = b * (k1 + k2) - 1.2 * k1
disc = d*d - 4*a*b
tau = 2*b / (-d + np.sqrt(disc))
```

Here $k_1\approx0.3047$, $k_2\approx0.1693$/min and $\tau\approx3.08559$ min. This smaller-root formula uses $a>0, d<0, \text{disc}\geq0$. Check the original τ bounds and reference B,C for each candidate.

Reference $C\approx0.199530<0.25$ and cost is $12\times3.08559+0.24\times(339.8152-310)\approx44.1828$ \$/h. Savings from the original baseline are $76.8-44.1828=32.6172$ \$/h.

Benchmark differences and conditions for regret

$$\Delta J_{bench} = J\left(u_{candidate}\right) - J\left(u_{bench}\right)$$

Compute this when both points satisfy the original specifications. If the benchmark is only a found feasible solution, the difference does not certify regret relative to the global optimum. It could become negative for a better candidate.

The repaired candidate has difference 44.1828 − 41.3920 = 2.7908 $/h. Leave savings and benchmark differences blank for the infeasible unadjusted and local-correction candidates.

Predicted versus actual improvement

For minimization using an objective model m on a common feasible input domain, compare predicted and actual improvement between the current point and a candidate.

$$\rho = \frac{J(u_{old}) - J(u_{new})}{m(u_{old}) - m(u_{new})}$$

Use this ratio only with a positive denominator. Here the cost is an exact input function, so feasibility error is the issue rather than objective-model error. Do not use this ratio to accept infeasible candidates.

Alexandrov, Dennis, Lewis and Torczon, A trust-region framework for managing the use of approximation models in optimization (1998), author manuscript

If the true objective drops from 100 to 92 while the model predicts 100 to 90, actual improvement is 8, predicted improvement is 10, and ρ=0.8. If the model also predicts 100 at the new point, the denominator is zero and the ratio is not evaluated. Both candidates must first satisfy the original constraints.

Exercises

1. Calculate B violation at the unadjusted candidate and C slack at the baseline.
2. If the uniform overprediction bound for B is 0.01, what lower bound is needed on predicted B?
3. Does the maximum error on 961 calibration points guarantee the entire box? Explain.
4. State the acceptance rule when comparing an infeasible cost of 37.9875 with a feasible cost of 44.1828.
5. Propose a point after local-correction failure and explain why it should not be added to the independent test set.

Short answer checks

1. B shortfall is 0.008843 mol/L. Baseline C slack is $0.25 - 0.245455 = 0.004545$ mol/L.
2. Require predicted B $\geq$ 0.39 mol/L. The implication holds only where the uniform bound is valid.
3. Larger errors between sampled points remain possible. Reevaluate the new candidate.
4. Compare costs among specification-feasible candidates. The cost 37.9875 is not accepted savings.
5. The failed point (339.8152 K, 2.87093 min) and nearby boundary points are useful. If used for improvement, they become adaptively collected training data.

Final submission and supporting sources

Submit the candidate table, rejection reasons, feasible-cost comparison, model/data revision plan, and execution record. State the reference assumptions used in evaluation.

Use the weight hash recorded in the result file to check that the same fitted model was used. A solver’s optimal status applies to its formulation, not global optimality of the actual process.

Week 6 frozen model

Alexandrov, Dennis, Lewis and Torczon, A trust-region framework for managing the use of approximation models in optimization (1998), author manuscript