# Supplementary Methods and Results

*Explainable AI for Multiobjective Design Decisions in University Teaching Buildings*

Manuscript ID: buildings-4665314

This supplement documents the additional analyses prepared for the revision. The original 1000-record archive and primary reference checks are retained. New calculations and all full-precision tables are supplied in the updated Supplementary Data and Code under Revision/. Complete SQL files and annual simulation matrices remain in the authors' local archive and are available from the corresponding authors on reasonable request.

## S1 Energy accounting and assumed equipment efficiencies

The original EUI is an ideal-load energy index, comprising ideal heating and cooling thermal demand plus scheduled lighting and equipment electricity per gross model floor area. The additional electricity proxy divides heating and cooling demands by assumed heating/cooling COPs before adding the scheduled electrical loads. Cooling COPs 2.5/3.5/4.5 and heating COPs 2/3/4 define nine scenarios. These assumptions describe constant-efficiency sensitivity, without installed equipment, auxiliary consumption or part-load simulation.

**Table S1. Preference stability across the nine assumed-efficiency combinations. All use the original six reference-checked eligible designs. The pure-energy order can change even where the five weighted preference winners remain unchanged.**

| Heating COP | Cooling COP | Equal | Energy | Thermal | Daylight | View |
|---|---|---|---|---|---|---|
| 2.0 | 2.5 | D239 | D221 | D239 | D239 | D241 |
| 2.0 | 3.5 | D239 | D221 | D239 | D239 | D241 |
| 2.0 | 4.5 | D239 | D221 | D239 | D239 | D241 |
| 3.0 | 2.5 | D239 | D221 | D239 | D239 | D241 |
| 3.0 | 3.5 | D239 | D221 | D239 | D239 | D241 |
| 3.0 | 4.5 | D239 | D221 | D239 | D239 | D241 |
| 4.0 | 2.5 | D239 | D221 | D239 | D239 | D241 |
| 4.0 | 3.5 | D239 | D221 | D239 | D239 | D241 |
| 4.0 | 4.5 | D239 | D221 | D239 | D239 | D241 |

D102 reduces the electricity proxy by 8.74–11.33% relative to the same baseline; D221 by 1.92–3.31%. Full component values and scenario-specific changes are in Revision/Sensitivity/reference_energy_end_use_components.csv and hypothetical_COP_scenario_energy.csv.

## S2 Comfort and service-screen sensitivity

An absolute PMV boundary of 0.85 corresponds to approximately 20.25% instantaneous PPD under the PMV–PPD relation. This study tolerance is distinct from the annual fraction of occupied hours outside it. The separate 20% annual discomfort ceiling, 60% UDI floor and 5% full-sphere VP floor define the initial design brief. They do not establish regulatory compliance.

The PMV limits 0.50, 0.70 and 0.85 are crossed with annual discomfort ceilings of 15%, 20% and 25%, UDI floors of 55%, 60% and 65%, and VP floors of 4%, 5% and 6%. All 81 combinations are retained, including cases with no eligible design. Discomfort is computed directly from saved hourly PMV and scheduled-occupancy masks, with floor-area weights over 29 controlled rooms and 1940 occupied hours per room.

**Table S2. Area-weighted occupied discomfort percentage at three PMV tolerances. These are direct reaggregations of the original nine checked annual responses; the baseline is a comparator.**

| Design | \|PMV\| > 0.50 (%) | \|PMV\| > 0.70 (%) | \|PMV\| > 0.85 (%) |
|---|---|---|---|
| Baseline | 76.66 | 61.34 | 44.97 |
| D001 | 80.70 | 70.51 | 56.18 |
| D102 | 59.69 | 25.80 | 6.50 |
| D110 | 23.94 | 1.44 | 0.02 |
| D148 | 60.97 | 27.27 | 8.39 |
| D221 | 38.13 | 11.97 | 2.88 |
| D239 | 33.24 | 8.69 | 2.06 |
| D240 | 47.10 | 19.16 | 8.45 |
| D241 | 33.70 | 9.87 | 3.22 |
| D247 | 77.99 | 55.49 | 34.32 |

At the original service preferences, the 0.70 PMV boundary retains D221, D239, D240 and D241, with D241 the equal-preference winner. No original checked candidate meets the combined screen at the 0.50 boundary in any of the 27 tested service-threshold combinations. This statement concerns the checked candidates, not the feasibility of every point in the design domain. Complete eligibility and rankings are in Revision/Sensitivity/threshold_eligibility.csv and threshold_Topsis_rankings.csv.

## S3 Empirical prediction-error bands and conservative screening

Absolute development-fold residuals are collapsed from 804 records to 782 unique input groups. For each response, an empirical half-width is the sorted-residual order statistic ceil((782 + 1)p), with p = 0.90 or 0.95. Bounds are clipped to the physical response range. The 196 reserved-group predictions evaluate coverage without selecting the band width or refitting models. These cross-validation-residual bands are empirical diagnostics and do not carry a formal post-optimization coverage guarantee.

**Table S3. Development-derived 95% empirical error margins and observed reserved-test coverage. EUI units are kWh/(m²·year); DCH units are hours; UDI and VP units are percentage points.**

| Response | 95% half-width | Covered / 196 | Observed coverage (%) |
|---|---|---|---|
| EUI | 5.424 | 189 | 96.43 |
| DCH | 88.818 | 186 | 94.90 |
| UDI | 11.979 | 190 | 96.94 |
| VP | 0.983 | 191 | 97.45 |

Observed all-four-response coverage is 88.27% on the reserved test groups. On the nine optimization-selected reference cases, marginal coverage is 4/9 for EUI, 7/9 for DCH, 8/9 for UDI and 6/9 for VP. D110's UDI and VP errors are outside the corresponding margins. Selected-candidate coverage is therefore reported separately from independent test coverage.

**Table S4. Nominal and margin-adjusted conservative full-front priorities. The conservative rule uses upper DCH and lower UDI/VP bounds; TOPSIS scores the entire corresponding 256-vector set before filtering. Final checked-set ranking is a separate stage using direct simulation responses.**

| Rule | Eligible n | Equal | Energy | Thermal | Daylight | View |
|---|---|---|---|---|---|---|
| Nominal | 94 | D221 | D102 | D239 | D221 | D241 |
| 95% empirical margins | 68 | D221 | D148 | D239 | D221 | D241 |

All five conservatively screened priorities were included among the original nine direct checks, so no additional candidate was required solely to represent these five profiles. The complete vector-by-vector audit is in Revision/Sensitivity/front_empirical_error_band_screen.csv, and selected-candidate errors are in selected_candidate_error_band_audit.csv.

## S4 Controlled operating-setting ablation

Four selected designs retain their full original geometry, materials, shading, weather and schedules while occupied cooling/heating settings are reset to 26/20°C. For D102 and D221, the baseline envelope is also run at the original candidate settings. Each design consequently has a complete two-by-two comparison: baseline fabric/baseline operation, candidate fabric/baseline operation, baseline fabric/candidate operation and the original joint case. The interaction is the joint response minus the two single changes plus the baseline response.

**Table S5. Two-factor differences from the common baseline. Negative values indicate lower EUI or fewer discomfort hours. The interaction term is retained explicitly; the three component differences sum to the joint difference.**

| Design | Metric | Fabric Δ | Operation Δ | Interaction Δ | Joint Δ |
|---|---|---|---|---|---|
| D102 | EUI | -22.864 | 6.199 | -0.392 | -17.058 |
| D102 | DCH (h) | -268.578 | -482.923 | 5.162 | -746.339 |
| D221 | EUI | -15.115 | 12.306 | -1.085 | -3.894 |
| D221 | DCH (h) | -119.100 | -721.038 | 23.706 | -816.432 |

A symmetric allocation averages the two possible intervention orders, allocating half of the interaction to each factor. For D102 this gives fabric and operating DCH effects of −265.997 and −480.342 h; for D221, −107.247 and −709.185 h. The corresponding operating shares of the joint simulated discomfort reduction are 64.36% and 86.86%. These values describe the selected controlled simulations, separately from SHAP importance. Full energy-component decompositions are in Revision/Controlled/Factorial_effects_2x2.csv.

## S5 Linked catalogue-glazing checks

Product A uses Solarban 90 on Starphire/Starphire glass and product B Solarban 70 on clear/clear glass, each with 4 mm panes and a 12 mm cavity containing 90% argon. Manufacturer values are NFRC winter-night center-of-glass U-factors with linked SHGC and VT. The simplified-glazing model uses those tuples together; frames and a full installed assembly are not simulated. The source is Vitro Architectural Glass, Glass Performance Tables for International Markets (2022), available at https://www.vitroglazings.com/media/xmsegrq0/international-performance-charts.pdf, accessed 24 October 2026. The figures are manufacturer-published calculated comparison values, not this project's measurements.

**Table S6. Linked-product input and output checks. EUI is the ideal-load index; DCH is area-weighted discomfort hours. Full-precision inputs and all three PMV thresholds are included in Revision/Controlled/.**

| Case | U / SHGC / VT | EUI | DCH (h) | UDI (%) | VP (%) |
|---|---|---|---|---|---|
| D221–A | 1.36 / .23 / .54 | 118.654 | 72.440 | 78.262 | 8.980 |
| D239–B | 1.35 / .28 / .64 | 126.997 | 69.454 | 78.111 | 9.120 |

D221–A and D239–B meet the initial comfort/daylight/openness screen. Their energy indices are 2.17% and 9.36% above the common baseline. They are therefore reported as service-compatible product-linked alternatives with different energy consequences, not as reproductions of the original independent-tuple energy result.

## S6 Run verification and reproducibility

All eight additional EnergyPlus calculations completed with exit code zero, zero severe/fatal errors and accepted occupied temperature control. Geometry was compared object by object with the corresponding

source cases. Daylight and VP were reused only when their geometry/optical inputs were unchanged; both product cases have independently recalculated annual Radiance daylight, with geometric VP reused. The source model includes small louver planarity warnings, which are retained with the original error logs and quality records. They are not described as warning-free runs.

**Table S7. Additional calculation inventory and observed thermal-run duration. Durations reflect sequential local execution; the product daylight runs are additional and are documented in their quality files.**

| Case | Changed inputs | Thermal duration (min) | Accepted |
|---|---|---|---|
| D221 @26/20 | X16, X17 | 12.43 | Yes |
| D221–A | X4, X5, X6 | 13.04 | Yes |
| D102 @26/20 | X16, X17 | 10.75 | Yes |
| D239 @26/20 | X16, X17 | 8.70 | Yes |
| D241 @26/20 | X16, X17 | 10.31 | Yes |
| D239–B | X4, X5, X6 | 9.18 | Yes |
| Baseline @ D102 settings | X16, X17 | 7.59 | Yes |
| Baseline @ D221 settings | X16, X17 | 8.10 | Yes |

Revision/Sensitivity/reanalyze_sensitivity.py reproduces the 81 threshold scenarios, empirical bands and COP comparisons from the original archived outputs. Run it with explicit --source pointing to the extracted supplement root and --output pointing to a new independent output folder. Revision/Controlled/ contains compact inputs, hourly PMV/occupancy subsets, quality/error logs, daylight masks and summary/verification scripts for the additional cases. The full SQL files and dense annual daylight matrices are retained locally; the compact archive provides the numerical inputs and results needed to audit the reported checks without including those large intermediates.