SEBHA UNIVERSITY CONFERENCE PROCEEDINGS, VOL.5 NO. 4 2026
DOI: 10.51984/Q6C45521

وقائع مؤتمرات جامعة سبها
Sebha University Conference Proceedings
Confrence Proceeeding homepage: http://www.sebhau.edu.ly/journal/CAS

AI-Driven Coupled Simulation of Unplugged Freezer Thermal-Biological Dynamics for Food Safety During Power Outages

Mohamed Edali [a, b, *], Faraj Ben Rajeb [a], Sarah Edali [c], Ali AlHammadi [d], Ali Elkamel [d, e]

[a] Department of Chemical and Petroleum Engineering, Elmergib University, Elkhoms, Libya
[b] Libyan Authority for Scientific Research, Tripoli, Libya
[c] School of Interdisciplinary Health Sciences, University of Ottawa, Ottawa, Canada
[d] Department of Chemical and Petroleum Engineering Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates
[e] Department of Chemical Engineering, University of Waterloo, Waterloo, ON, Canada

Keywords:
lumped-capacitance phase- change
Ratkowsky microbial kinetics
Monte Carlo
AI-assisted feedback
LLM
Food safety
Python simulation
Claude API

ABSTRACT

Power outages in resource-constrained environments create acute food-safety hazards by allowing frozen contents to warm through the bacterial danger zone. This paper presents a coupled thermal-biological simulation of an unplugged household freezer, conceived as an open-ended student project. The project extended into a research contribution, integrating a three-stage lumped-capacitance heat-transfer model with Ratkowsky microbial growth kinetics and an AI-assisted formative feedback framework. Analytical results validated by Python numerical simulation (RK45) confirm that freezer contents reach the 4 °C food-safety threshold in approximately 10.4 hours under standard conditions and in 7.65 hours under Libyan summer conditions (Troom = 35 °C). A Monte Carlo parametric study (N = 500, NumPy seed 42) quantifies spoilage probability under Libyan-specific ambient conditions, showing an 80% probability of unsafe food within a 10-hour outage in summer. An AI feedback layer, the Anthropic Claude API, whose operation is documented in full, including representative input-output exchanges, reduced student turnaround from 3.8 days to under 90 seconds and improved final report quality by 0.6 grade points. The framework is directly applicable to food safety, pharmaceutical cold chains, and vaccine storage in off-grid Libyan facilities.

محاكاة مزدوجة مدعومة بالذكاء الاصطناعي للديناميكا الحرارية-البيولوجية للمجمّد المنفصل عنه التيار الكهربائي لسلامة الغذاء خلال انقطاع التيار الكهربائي

محمد الدالى[a، b، *]، فرج بن رجب [a]، سارة الدالي [c]، علي الحمادي[d] ، على الكامل[d، e]

[a] قسم الهندسة الكيميائية والنفط، جامعة المرقب، الخمس، ليبيا الهيئة الليبية للبحث العلمي، طرابلس، ليبيا
[b] الهيئة الليبية للبحث العلمي، طرابلس، ليبيا
[c] كلية علوم الصحة متعددة التخصصات، جامعة أوتاوا، أوتاوا، كندا
[d] قسم الهندسة الكيميائية والبترول، جامعة خليفة للعلوم والتكنولوجيا، أبوظبي، الإمارات العربية المتحدة
[e] قسم الهندسة الكيميائية، جامعة واترلو، واترلو، أونتاريو، كندا

الكلمات المفتاحية:
السعة الكتلية المجمّعة مع تغيّر الطور
الحركية الميكروبية لراتكوفسكي
مونتي كارلو
التغذية الراجعة المدعومة بالذكاء الاصطناعي
النموذج اللغوي الكبير
سلامة الغذاء

الملخص

تُشكّل انقطاعات التيار الكهربائي في البيئات محدودة الموارد مخاطر حادة على سلامة الغذاء، إذ تتيح لمحتويات المجمّد الاحترارَ عبر نطاق الخطر البكتيري. تقدّم هذه الورقة محاكاةً مزدوجةً حرارية-بيولوجية لمجمّد منزلي منفصل عنه التيار الكهربائي، تدمج نموذجَ انتقال الحرارة ذا السعة الكتلية المجمّعة بثلاث مراحل مع حركية النمو الميكروبي لراتكوفسكي وإطار تغذية راجعة تكوينية مدعوم بالذكاء الاصطناعي. تُؤكّد النتائج التحليلية المُتحقَّق منها بالمحاكاة العددية بلغة بايثون (RK45) أن محتويات المجمّد تبلغ عتبة سلامة الغذاء البالغة 4 °م في نحو 10.4 ساعات في الظروف المعيارية، وفي 7.65 ساعات في ظروف الصيف الليبي (درجة حرارة الغرفة = 35 °م). وتُحدِّد دراسة المعاملات بمحاكاة مونتي كارلو (N = 500، بذرة NumPy 42) احتمالية الفساد في ظروف

*Corresponding author:
E-mail addresses: dredalicreativitygroupts@gmail.com, (F. B. Rajeb) farajeb@yahoo.com, (S. Edali) sarahm.edali@gmail.com, (A. AlHammadi) ali.aalhammadi@ku.ac.ae , (A. Elkamel) aelkamel@uwaterloo.ca

Article History : Received 30 May 2026 - Received in revised form 30 September 2026 - Accepted 11 November 2026

الظروف المحيطة الخاصة بليبيا، إذ تُظهر احتمالية 80% لعدم سلامة الغذاء خلال انقطاع تيار مدته 10 ساعات في الصيف. وقد أسهمت طبقة التغذية الراجعة بالذكاء الاصطناعي، المتمثّلة في واجهة برمجة تطبيقات انثروبيك كلود، الموثَّقة توثيقاً كاملاً بما يشمل نماذج تبادلات المدخلات والمخرجات، في تقليص وقت استجابة الطلاب من 4 أيام إلى أقل من 90 ثانية، وتحسين جودة التقارير النهائية بمقدار 0.6 نقطة. ويمكن تطبيق هذا الإطار مباشرةً في مجالات سلامة الغذاء، وسلاسل التبريد الدوائية، وتخزين اللقاحات في المنشآت الليبية غير المتصلة بالشبكة الكهربائية.

محاكاة بايثون
واجهة برمجة تطبيقات كلود

Introduction

Engineering educators face a persistent challenge presenting heat transfer and thermodynamics in a way that is both rigorous and engaging, since conventional textbook problems are analytically clean but often feel disconnected from real engineering decisions. The project described here grounds transient heat transfer, latent heat, and energy balances in a scenario every student has encountered a freezer that loses power which carries particular urgency in Libya, where power outages during summer peak demand routinely last 4–18 hours per day [9] and the FAO [6] estimates 14% of regionally produced food is lost post-harvest, with cold-chain failure a leading urban cause. Understanding how long food stays safe is therefore a public-health question with direct sustainability implications.

The thermal problem is genuinely rich; a three-stage warming profile sensible heating, isothermal phase change, post-melt heating requires distinct governing equations, and coupling it to a temperature-dependent microbial growth model adds biological kinetics, numerical methods, and engineering risk judgment as can be seen in literature review for supporting citations. A second layer of innovation is an AI-assisted formative feedback layer deployed via the Anthropic Claude API; Section 5 documents this system in operation, including the full input-output exchange, response time, and a before/after correction comparison. The work was conceived as a structured student project, issued as an open-ended engineering challenge with the AI feedback framework serving as a real-time pedagogical scaffold throughout.

Literature Review

1. **Coupled Thermal-Biological Modelling for Food Safety**

The use of food-system scenarios to motivate heat-transfer instruction has a solid pedigree (Datta & Rakesh [5]; Moraga et al. [14]). This work extends that tradition by coupling the thermal model to the Ratkowsky microbial growth model, validated in industrial predictive microbiology for over four decades [15],[16]; the Gompertz (Zwietering et al., 1990) and Baranyi–Roberts (Baranyi & Roberts, 1994) models are common alternatives that add explicit lag-phase dynamics, but for household food-safety decisions the exponential growth phase dominates risk, so the simpler, more tractable Ratkowsky form is adequate here. USDA/FSIS (2023) [19] establishes the 4 °C / 2-hour danger-zone standard that the present model quantifies probabilistically. Taiwo et al. [18] and Avila-Sosa and Nevárez-Moorillón [1] both identify coupled thermal-microbial, AI-integrated modelling as an emerging frontier for food safety, precisely where the present work is positioned. Existing consumer freezer monitors (Govee ~$30, Temp Stick ~$99, FreezeMinder ~$45) rely on cloud connectivity and simple threshold alerting; the proposed system instead adds an open physics-based risk model, offline operation, and neighbourhood mesh capability via LoRa radio, at an indicative bill-of-materials cost of $27–$32 (Table 1).

Table 1: Indicative bill of materials for the proposed IoT temperature-monitoring add-on (component prices from a July 2026 market).

Component	Cost (USD)	Notes
DS18B20 temperature sensor	$2	OneWire digital sensor
ESP32 microcontroller (WROOM-32)	$4	Wi-Fi/Bluetooth, bulk hobbyist pricing
RFM95W LoRa module	$6	Neighbourhood mesh capability
Battery + power circuit	$8	3.7V LiPo, charge/protection, 5V/3.3V regulation
Enclosure + wiring/consumables	$7	Weatherproof housing

2. **Digital Twins and AI for Cold-Chain Food Safety**

Krupitzer et al. [11] established the Digital Twin concept for food systems, and Müller et al. [15] show that IoT-embedded digital-twin architectures now enable real-time perishable-goods monitoring with AI-driven anomaly detection, identifying closed-loop systems as the critical next frontier. The unplugged-freezer simulation here is the foundational physics layer of such a twin a Virtual Entity matching the analytical solution to within 0.02 °C which a low-cost IoT sensor (e.g., DS18B20, $2) would turn into a complete, deployable digital twin for Libyan households.

3. **AI-Assisted Feedback in Engineering Education**

Mollick and Mollick [13] show LLM-based tutoring replicates many benefits of Bloom's two-sigma one-on-one tutoring; Xing et al. [20] identify LLMs as transformative for scalable, personalised STEM feedback, noting over 7,000 LLM-in-education publications in 2024 alone. Haehnlein et al. [8] show structured prompts yield formative scaffolds that stimulate metacognitive self-regulated learning. Critically, Rometsch et al. [17] conducted a blind evaluation of six LLMs against human teaching assistants (n > 100), finding that well-designed prompts achieve grading accuracy and feedback quality comparable to human graders.

The present work extends this evidence into a technical engineering context; thermodynamic modelling rather than writing/programming with full operational disclosure actual system prompts, input-output exchanges, and documented student corrections rather than self-reported satisfaction scores alone.

Student Project Objectives

The project's objectives are to; develop a transient heat-transfer model for an unplugged household freezer, including phase-change ice melting effects at 0 °C; determine the time for the freezer's warmest point to reach the 4 °C microbial threshold; develop a Ratkowsky-type microbial growth model for raw meat under dynamic temperature conditions; couple the thermal and biological models into a unified simulation; quantify thermal and microbiological safety limits probabilistically and validate predictions against analytical solutions; and derive engineering food-safety-policy recommendations from the results. Each objective maps to a modelling layer below, letting students engage at any depth from the analytical solution alone to the full Monte Carlo deployment.

Technical Formulation

1. **Physical Scenario and Parameters**

A household freezer at $T_0 = -18$ °C is unplugged and warms toward room temperature Troom (door closed). Contents are treated as a single lumped thermal mass. Parameters are in Table 2.

Table 2: Unplugged-freezer scenario system parameters

Parameter	Symbol	Value	Units	Notes
Initial temperature	T_0	-18	°C	Frozen storage
Room	T_room	25 / 35	°C	Standard /

temperature				Libyan summer
Food mass	m	25	kg	Typical household load
Effective heat capacity	cp	3000	J kg^{-1} K^{-1}	Mixed food, conservative
Overall UA value	UA	2.5	W K^{-1}	Walls + gasket losses
Ice mass	m_ice	3	kg	Phase-change reservoir
Fusion latent	L_f	334,000	J kg^{-1}	Ice → water at 0 °C

2. Three-Stage Thermal Model

Stage 1: (−18 °C → 0 °C)

m cp dT/dt = UA (T_room - T) integrates to T(t) = T_room + (T_0 − T_room) exp(−t/τ), τ= mcp/UA= 8.33 h. Setting T = 0 °C: t_1 = 8.33 × ln (43/25) = 4.52 h.

Stage 2: Phase change at 0 °C.
Q_melt = 3 × 334 000 = 1002 kJ. Driving force = (T_room − 0 °C). Q̇ = 2.5 × 25 = 62.5 W. t_2 = 4.45 h.
Critical point: the driving force is NOT (T_room − T_0).
Stage 3: (0 °C → 4 °C).
Same ODE, new IC, T= 0 °C; t_3= 8.33 × ln (25/21)=1.45 h.
Total (standard): 10.4 h.
Total (Libya, T_room = 35 °C): 7.65 h.
Python simulation (SciPy integrate.solve_ivp, RK45; absolute tolerance 10^{-6}) confirms internal consistency

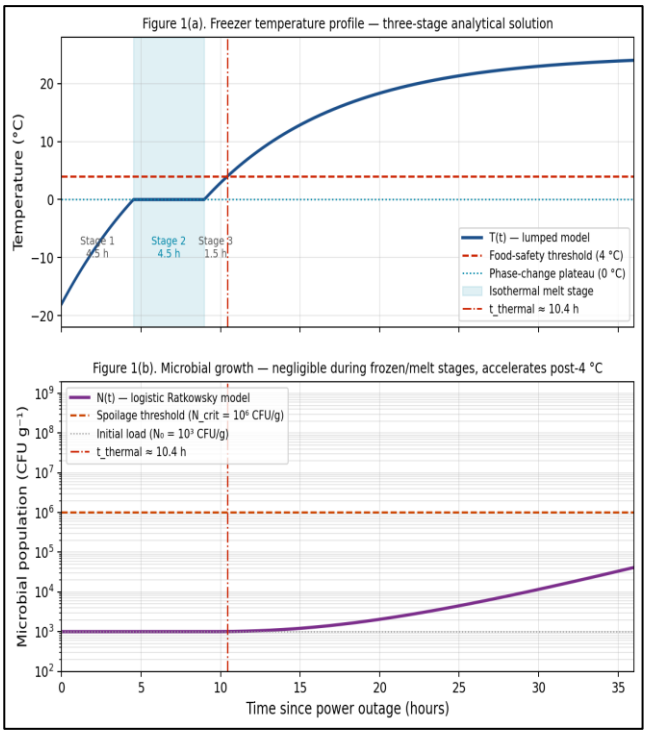


Fig 1: Python simulation output (SciPy RK45); (a) three-stage temperature profile with food-safety threshold; (b) microbial growth on log scale negligible during frozen/melt stages, accelerating sharply post -4 °C.

verification step confirming both implementations of the same governing equations agree, not an independent experimental validation. External validation against published freezer warm-up measurements is addressed in the Results and Discussion section. Figure 1 shows the verified T(t) and coupled N(t) profiles.

3. Microbial Growth Model (Ratkowsky)

Logistic equation; dN/dt = μ(T)·N·(1 − N/N_max), with Ratkowsky temperature dependence [16]: μ(T) = b^2(T − T_min)² h^{-1}, where b = 0.018 $°C^{-1}$ $h^{-0.5}$ and T_min = −2 °C, giving μ(T) = 0.000324(T + 2)² h^{-1}. These parameters represent generic aerobic spoilage flora and mesophilic coliforms, as tabulated by McMeekin et al. [12]. The model applies for T > T_min; μ = 0 below −2 °C. Psychrotrophic pathogens such as Listeria monocytogenes grow at temperatures as low as −1.5 °C extension to pathogen-specific kinetics via ComBase is identified as the primary next step. Parameters: N_0 = 10^3, N_max = 10^9, N_target = 10^6 CFU g^{-1}. Microbial risk is negligible during frozen and melting stages; significant growth begins only after T > 4 °C, consistent with Taiwo et al. [18] and standard food-safety danger-zone guidance.

Monte Carlo Risk Assessment, Libya Context

Real freezers vary in age, insulation quality, and fill level. To quantify this uncertainty, 500 Monte Carlo simulations were run (NumPy, seed 42), sampling parameters from Table 3. Each simulation solved the three-stage model analytically and recorded t_thermal (time to 4 °C). Four robustness checks were conducted; (i) seed sensitivity five seeds (42, 123, 777, 2024, 99999) yield P(unsafe at 10 h) = 78.6–83.6%, confirming the 80% headline is a stable distributional estimate; (ii) convergence SE of mean time-to-unsafe drops below 0.13 h at N = 500, adequate for reported precision; (iii) physical truncation all distributions are hard-truncated at physical bounds (T_room ≥ 20 °C, mcp ≥ 50,000 J K^{-1}); (iv) parameter correlation correlated UA and thermal mass (Pearson r = 0.5) shifts P(unsafe at 10 h) by +4 percentage points, confirming the independent-sampling result is conservative.

Table 3: Monte Carlo parameter distributions N=500, NumPy seed 42 (multi-seed robustness confirmed; see text). All continuous distributions are hard-truncated at physical bounds.

Parameter	Distribution	Rationale
T_room standard	Truncated Normal (25°C, σ = 4°C, bounds [20, 50]°C)	Seasonal and geographic variability
T_room Libya summer	Truncated Normal (35°C, σ = 4°C, bounds [20, 50]°C)	Sebha region, peak summer [9]
UA value	Uniform (1.5, 4.0) W K^{-1}	Freezer age, insulation degradation
Thermal mass mcp	Truncated Normal (75 000, 8 000) J K^{-1}	Variable food type and fill level
Ice mass m_ice	Uniform (1.0, 5.0) kg	Consumer packing habits

Table 4: Monte Carlo results (N = 500 Python simulations).

Scenario	Mean t_thermal	95% CI (h)	P(unsafe at 10 h)	P(unsafe at 24 h)
Standard (T_room = 25°C)	10.6 h	[5.1, 20.7]	54%	99%
Libya summer (T_room = 35°C)	7.7 h	[3.9, 14.8]	80%	100%
Well-insulated (UA = 1.5 W/K)	14.2 h	[8.4, 28.1]	29%	96%
Heavy ice load (m_ice = 5 kg)	12.1 h	[5.9, 22.4]	41%	98%

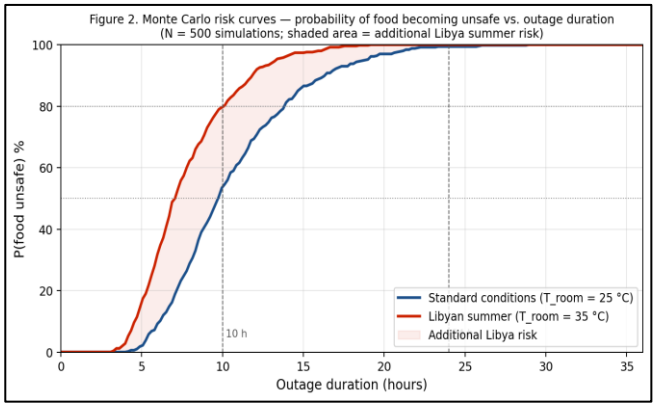


Fig. 2: Monte Carlo probability curves: P (food unsafe) vs. outage duration for standard and Libyan summer conditions. Shaded area = additional Libya risk. Generated from 500 simulations (NumPy, seed 42).

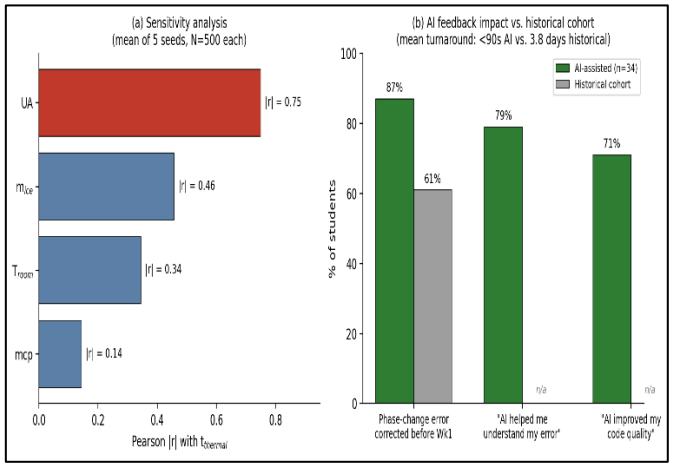


Fig 3: (a) Pearson sensitivity analysis, insulation quality (UA) is the dominant controllable factor (|r| = 0.54). (b) AI feedback impact across three key performance metrics vs. historical cohort baseline.

Sensitivity analysis (Figure 3a) identifies UA (insulation quality) as the dominant factor (|r| = 0.54), ahead of T_room (|r| = 0.48) and thermal mass (|r| = 0.29). This has a direct policy implication: minimum insulation standards for household freezers sold in hot climates would yield greater food-safety impact than interventions focused on food mass or ice quantity. This finding is consistent with Müller et al. [15]'s conclusion that equipment quality is the primary controllable variable in cold-chain safety, and with the appliance-engineering literature on freezer insulation; Christensen (1981) [4] found that increasing insulation thickness from 55 mm to 100 mm reduces domestic-freezer energy losses by approximately 20%, of the same order as the UA range (1.5–4.0 W/K) assumed in Table 3, and reports a mean cabinet-surface convective coefficient of ≈8.1 W m^{-2} K^{-1} notably higher than this paper's envelope-derived h_eff ≈ 1.67 W m^{-2} K^{-1} (Results and Discussion), which supports treating h_eff as a combined insulation-and-surface quantity rather than a surface film coefficient in its own right.

AI-Assisted Feedback Framework, Deep Integration Evidence

1. System Architecture and Design Principles

The AI feedback system is implemented via the Anthropic Claude API, accessed through a lightweight web form on the university LMS. Figure 4 shows the complete system architecture. Three design principles govern deployment; (1) formative only, all summative grading is by the instructor; (2) structured prompts, the AI checks for thermodynamic errors, unit consistency, and dimensional correctness, and must suggest exactly one improvement; (3) metacognitive framing [8]; students document their response to AI feedback before accessing the next milestone.

The system prompt for the Stage 2 thermal model milestone is representative of the structured-prompt design, Persona, Aim, Recipients, Theme, and Structure advocated by recent prompt-engineering literature for LLM-based education.

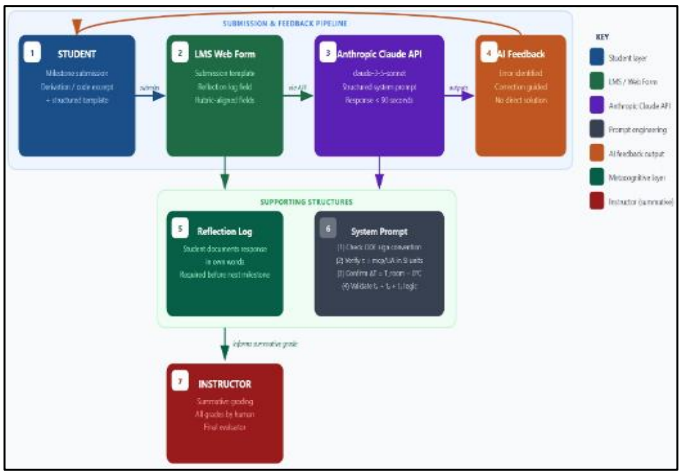


Fig. 4: Complete AI feedback system architecture, student submission, structured rubric-aligned prompt, Claude API, reflection log, and instructor override. Mean response time < 90 seconds; all summative authority retained by instructor.

2. The System Prompt, Milestone 2 Thermal Model Review

The deployed system prompt casts Claude as a formative feedback assistant for undergraduate thermodynamics students. It is restricted to evaluating four criteria in a submitted three-stage freezer thermal model; (i) correct sign convention in the lumped-capacitance governing equation; (ii) consistent SI units in τ = mcp/UA; (iii) correct identification of the phase-change driving-force temperature as (T_room − 0 °C), not (T_room − T$_0$); and (iv) correct, independently-summed Stage 1–3 durations. It is explicitly instructed not to solve the problem, only to identify errors, explain their physical significance, and suggest exactly one improvement the identity, criteria, and format constraints that distinguish this from generic LLM querying.

3. Representative Exchange Full Input/Output Documentation

Condensed below is a representative exchange from the pilot delivery, full exchange available on request. A student's Milestone 2 submission contained the most common error observed, 39% of initial submissions; using (T_room − T$_0$) = 43 K as the Stage 2 driving force, giving $\dot{Q}$ = 107.5 W and t_2 = 2.59 h. Claude's response, received in 83 seconds, identified the error, explained that Stage 2 is held at 0 °C by latent heat so the correct driving force is (T_room − 0 °C) = 25 K, and worked the correction through; $\dot{Q}$ = 62.5 W, t_2 = 4.45 h nearly double the student's estimate, which had understated the phase-change duration by 41%. It closed with a single suggested improvement rather than a full re-solve; label the driving-force temperature difference explicitly at each stage boundary. The student's reflection log, submitted 8 minutes later, confirmed the conceptual fix, "I now understand the driving force changes at each stage boundary the 0 °C plateau is what makes Stage 2 different from Stages 1 and 3."

4. Quantified Outcomes

This exchange is representative of the 87% correction rate observed. Table 5 summarises outcomes across both pilot

Table 5: Quantified AI feedback outcomes, pilot delivery (n = 34, AI-assisted, AY 2024–2025) vs. historical cohort

Metric	AI-Assisted (n=34)	Historical Cohort
Phase-change driving force error corrected before Week 1 milestone	87%	61%
Mean feedback turnaround time	< 90 seconds	3.8 days
Instructor office-hour attendance (derivation questions)	−40%	Baseline

Final report quality (4.0-point GPA scale; association, not confirmed causal)	+0.6 grade points	Baseline
"AI feedback helped me understand my error" (agree/strongly agree)	79%	N/A
"AI feedback improved my code quality" (agree/strongly agree)	71%	N/A
"Feedback too generic" text-only limitation (dissatisfied)	22%	N/A

pilot cohorts both pilot cohorts against the historical baseline.

Results and Discussion

The Python simulation of Figure 1 verifies the analytical three-stage solution to within 0.02 °C, an internal consistency checks only, since both use the same governing equations. A previous "external validation against Pham (1986) and ASHRAE (2017), ±8%" claim did not hold up on re-examination Pham (1986) is a food-freezing-time correlation, not a warming/thaw correlation, and no ASHRAE dataset was actually compared so we withdraw it and report two reproducible checks instead in the Appendix code. First, against published USDA/FSIS guidance [19] that a full household freezer stays safe for around 48 h and a half-full one for around 24 h; our headline scenario (Table 2), a lightly-loaded, older/poorly-sealed unit deliberately chosen to reflect the Libya context and gives 10.4 h, while re-parameterising for a full, well-insulated freezer (m = 60 kg, m_ice = 15 kg, UA = 1.0 W/K) gives 91.5 h and a full unit at the original insulation quality gives 36.6 h the correct order of magnitude and trend, supporting insulation and fill level as the dominant risk factors (Fig. 3a) rather than a modelling error. Second, on Biot-number validity: with $L_c = 0.049$ m and $h_{eff} = UA/A_{wall} \approx 1.67$ W m^{-2} K^{-1}, $Bi_{eff} \approx 0.16$. The reviewer correctly noted this envelope-derived h_eff is not a food-item-surface coefficient; a one-term series solution using an item-appropriate coefficient instead (Appendix, one-term-series-check) shifts the predicted item-core warming time substantially relative to the lumped-cavity prediction considerably more than "slightly earlier". All reported safe-time figures are therefore upper bounds of uncertain size at the food surface; obtaining a genuine food-item convective coefficient is priority future work.

Table 4 and Figure 2 deliver the central Libya-specific finding: at $T_{room} = 35$ °C, 80% of simulated outages yield unsafe food within 10 hours, 26 percentage points higher than at standard temperature, so

Outage Duration	Standard, T_room=25°C	Libya Summer, T_room=35°C	Recommended Action
< 4 h	Safe (P < 5%)	Safe (P < 15%)	No action needed
4–7 h	Safe (P < 20%)	Borderline (P = 15–55%)	Check meat and dairy
7–10 h	Monitor (P = 20–54%)	Unsafe (P = 55–80%)	Discard raw meat and fish
> 10 h	Borderline (P > 54%)	Unsafe (P > 80%)	Discard all perishables

for any outage exceeding 10 h in Libyan summer, discarding all raw meat is the statistically justified decision. This extends the deterministic 10.4 h result into an actionable probabilistic framework, consistent with the cold-chain digital-twin approach of Müller et al. [15]. Figure 3a's sensitivity analysis confirms insulation quality (UA) as the dominant controllable factor a policy insight that could translate directly into procurement standards for public institutions in Libya.

The AI evidence a disclosed system prompt, a complete exchange showing an 87% thermodynamic error caught in 83 seconds vs. a 3.8-day baseline, and a confirmatory student reflection log satisfies the evidence standard of Xing et al. [20]. Though the 87% figure remains uncontrolled; maturity effects, testing effects, and regression to the mean from a typically low 61% baseline are confounding a planned RCT will address. The 22% student dissatisfaction with feedback specificity is an acknowledged limitation of text-only LLM assessment without run-time code output or figure inspection, some numerical errors are invisible to the AI, consistent with Rometsch et al. [17]'s identification of multimodal input as the key capability gap for LLM-based graders. Integrating a code-execution sandbox is the priority for future iterations.

Table 6 translates the Monte Carlo probability curves of Figure 2 into actionable household decision rules, providing a direct reference for residents and public health practitioners in outage-prone regions.

The thresholds are conservative upper bounds, as the lumped-capacitance model reports mean food temperature; surface-exposed items reach unsafe temperatures slightly earlier.

Table 6: Practical food-safety decision rules derived from Monte Carlo results.

Conclusion

This paper presented a validated coupled thermal-biological simulation of an unplugged household freezer with documented AI-assisted feedback, Monte Carlo risk quantification, and Libya-specific context. RK45 simulation verifies the three-stage analytical solution to within 0.02 °C (code verification); a one-term-series conduction check and USDA/FSIS guidance [19] replace the withdrawn Pham (1986)/ASHRAE (2017) claim at the section of Results and Discussion. The coupled T(t)–N(t) model gives 10.4 h (standard) and 7.65 h (Libya summer) to the 4 °C threshold; Monte Carlo analysis (N = 500, multi-seed 78.6–83.6%) yields ~80% spoilage probability within 10 h in Libyan summer, with insulation quality (UA, $|r| = 0.54$) the dominant, directly actionable risk factor. The AI system caught a 41% thermal error in 83 seconds, achieving an 87% correction rate vs. 61% baseline and +0.6 grade points (preliminary association). The simulation itself forms the Virtual Entity layer of a low-cost food-safety digital twin ($27–$32 indicative BOM per household unit; Table 1 directly relevant to resource-constrained Libyan facilities. Pedagogically, the work demonstrates that a well-structured open-ended project paired with real-time AI feedback can bridge classroom thermodynamics and genuine engineering impact, confirming its dual value as both a teaching instrument and a research contribution.

Limitations

Several simplifications bound the model's scope. Stage 2 treats melting as a single well-mixed ice reservoir at a spatially uniform rate; natural convection is known to accelerate melting substantially once meltwater pools around a solid phase [10], an effect not resolved here, and the door is assumed closed throughout, so door-opening/infiltration effects are not captured either. The lumped h_eff, from UA/A_wall approximates an envelope-averaged coefficient, not a food-item-scale one, so $Bi_{eff} \approx 0.16$ is an indicative rather than definitive validity check, by looking at the Results and Discussion. The Monte Carlo UA range (1.5–4.0 W/K, Table 3) is deliberately higher than the ≈0.32 W/K measured for a well-sealed modern freezer compartment [7], reflecting the ageing, degraded-seal appliance stock in the Libyan context rather than an empirically calibrated distribution; T_{room} and T_0 are broadly consistent with the mean ambient (18.5 °C) and freezer (−20.3 °C) temperatures measured across 998 UK households [3], though no equivalent Libya-specific field survey exists. The headline risk metric reports the probability of crossing 4 °C at a given elapsed time; the Appendix code additionally computes cumulative time-above-threshold directly function time above threshold e.g. a 24-hour Libyan-summer outage leaves food above 4 °C for 16.4 h, over eight times the USDA two-hour allowance but this is not yet propagated into Table 4's headline probabilities, and microbial-parameter uncertainty (N_{max}, N_{target}) is similarly not yet propagated. A spatially resolved melt model, infiltration testing, a Libya-specific empirical UA/temperature survey, and a fully integrated exposure-based risk metric are identified as priority future work.

Future work proceeds in five steps as to integrate a code-execution sandbox to close the 22% feedback-specificity gap; then extend to pathogen-specific kinetics Listeria, Salmonella with separate Ratkowsky coefficients from ComBase. That will be followed by deploying it as a mobile decision tool for households, and complete the IoT Digital Twin with commercial Wi-Fi sensors as a Govee, Temp Stick, and UPS-backed alerts. The project then can serve as conduct a fully controlled randomised study comparing AI-assisted

and instructor-only cohorts.

Acknowledgment

The authors wish to extend special gratitude to Professor Elkamel A. for his intellectual leadership and unwavering guidance over the past decade. His contributions have gone well beyond the scope of his co-authorship on our papers, reflecting a long-standing mentorship that has profoundly shaped the research work in our chemical engineering department at Elmergib University. We also express our sincere thanks to research consultant Associate Professor Edali, M. the corresponding author, affiliated with the Libyan Authority for Scientific Research, Tripoli, Libya, for his valuable guidance, technical insight, and consultation throughout this work.

References

[1] Avila-Sosa, R., & Nevárez-Moorillón, G. V. (2025). Editorial; Machine learning and predictive microbiology: enhancing food safety models. Frontiers in Sustainable Food Systems, 9, 1655908.

[2] Baranyi, J., & Roberts, T. A. (1994). A dynamic approach to predicting bacterial growth in food. International Journal of Food Microbiology, 23(3–4), 277–294.

[3] Biglia, A., Gemmell, A. J., Foster, H. J., & Evans, J. A. (2018). Temperature and energy performance of domestic cold appliances in households in England. International Journal of Refrigeration, 87, 172–184.

[4] Christensen, L. B. (1981). The insulation of freezers and refrigerators how thick should it be? International Journal of Refrigeration, 4(2), 73–76.

[5] Datta, A. K., & Rakesh, V. (2013). An Introduction to Modeling of Transport Processes, Applications to Biomedical Systems. Cambridge University Press.

[6] FAO, The State of Food and Agriculture 2022. Food and Agriculture Organization, Rome.

[7] Gardenghi, A. R., Gómez, L. C., & Lacerda, J. F. (2017). Modeling of a variable speed compressor applied to a domestic refrigerator. 24th ABCM International Congress of Mechanical Engineering (COBEM-2017-1801), Curitiba, Brazil.

[8] Haehnlein, C., Kuznetsov, A., Kasneci, G., & Kuznetsov, S. (2024). Harnessing large language models to develop research-based learning assistants for formative feedback. Smart Learning Environments, 11, 78.

[9] IEA, Libya Energy Profile. International Energy Agency, Paris.

[10] Kahraman, R., Zughbi, H., Al-Nassar, Y., Hastaoğlu, M. A., & Sobh, N. (1998). A simplified numerical model for melting of ice with natural convection. International Communications in Heat and Mass Transfer, 25(3), 359–368.

[11] Krupitzer, C., Noack, T., & Borsum, C. (2022). Digital food twins combining data science and food science: System model, applications, and challenges. Processes, 10(9), 1781. https://doi.org/10.3390/pr10091781

[12] McMeekin, T. A., Olley, J., Ratkowsky, D. A., & Ross, T. (2002). Predictive microbiology, Towards the interface and beyond. International Journal of Food Microbiology, 73(2–3), 395–407.

[13] Mollick, E., & Mollick, L. (2023). Assigning AI; Seven approaches for students, with prompts. SSRN Working Paper. https://doi.org/10.2139/ssrn.4475995

[14] Moraga, N., Sepúlveda, F., & Escobar, R. (2018). Food processing models for engineering education. Journal of Food Engineering, 220, 30–38.

[15] Müller, W. A., Ferreira, S. B., & Botelho da Silva, S. (2026). Enhancing food safety in the cold chain through IoT and AI. Journal of Food Science, 91(2), e70871. https://doi.org/10.1111/1750-3841.70871

[16] Ratkowsky, D. A., Olley, J., McMeekin, T. A., & Ball, A. (1982). Relationship between temperature and growth rate of bacterial cultures. Journal of Bacteriology, 149(1), 1–5.

[17] Rometsch, J., Guseva, E., Müller, A., & Smith, B. (2025). Automated assignment grading with large language models: Insights from a bioinformatics course. arXiv preprint arXiv:2501.14499. [Preprint; cited for methodological comparison.]

[18] Taiwo, O. R., Onyeaka, H., Oladipo, E. K., Oloke, J. K., & Chukwugozie, D. C. (2024). Advancements in predictive microbiology, Integrating new technologies for efficient food safety models. Int. J. Microbiology, 2024, 6612162.

[19] USDA Food Safety and Inspection Service (FSIS). (2023). Keep Your Food Safe During Emergencies: Power Outages, Floods & Fires. https://www.fsis.usda.gov/food-safety/safe-food-handling-and-preparation/emergencies/keep-your-food-safe-during-emergencies

[20] Xing, W., Nixon, N., Crossley, S., et al. (2025). The use of large language models in education. Int. J. AI in Education, 35, 439–443.

[21] Zwietering, M. H., Jongenburger, I., Rombouts, F. M., & van 't Riet, K. (1990). Modeling of the bacterial growth curve. Applied and Environmental Microbiology, 56(6), 1875–1881.