Journal of Progress in **Engineering and Physical Science**

Volume 4, Number 6 **December 2025**

CONTENTS

Journal of Progress in Engineering and Physical Science
ISSN 2709-4006
www.paradigmpress.org/jpeps
Volume 4 Number 6 December 2025

Selleri Theory Not Equivalent to the Special Theory of Relativity and a Related Note on H.P. Robertson 1949 Paper

A. Sfarti[1]

[1] CS Dept, UC Berkeley, 387 Soda Hall, Berkeley, CA, USA
Correspondence: A. Sfarti, CS Dept, UC Berkeley, 387 Soda Hall, Berkeley, CA, USA.

doi:10.63593/JPEPS.2025.12.01

Abstract

In the present paper, I dispel the belief that the transformation developed by F. Selleri is equivalent to the Lorentz transformation, thus dispelling the misconception that his alternative theory is equivalent to Einstein's special theory of relativity. In the current paper I present an argument that has not been shown before aimed at settling the dispute. I demonstrate that the Selleri transformation is not equivalent to the Lorentz transformation by reasons of predicting different results for relativistic aberration.

Keywords: special relativity, Selleri transformation, Lorentz transformation, relativistic Doppler effect, relativistic aberration, Robertson 1949 paper

---

1. Introduction

In a paper published in 1996, F. Selleri (1990) put forward an alternative relativistic theory and derived the transformation equations that bear his name:

$$\begin{aligned} x' &= \gamma(v)(x - vt) \\ y' &= y \\ z' &= z \\ t' &= \gamma(v)t \end{aligned} \tag{1.1}$$

It can be seen that the Selleri transformations imply absolute simultaneity by contrast to the Lorentz transformations that imply relative simultaneity (Einstein, A., 1905):

$$\begin{aligned} x' &= \gamma(v)(x - vt) \\ y' &= y \\ z' &= z \\ t' &= \gamma(v)(t - \frac{vx}{c^2}) \end{aligned} \tag{1.2}$$

Selleri's program was to re-derive from the above transformation the explanation for the canonical tests of special relativity, the Michelson-Morley, Kennedy-Thorndike and Ives-Stilwell experiments and he succeeded in doing just that. The question that I plan to answer in the next section is: does that mean that Selleri's theory of relativity is equivalent to the special theory of relativity modulo the difference in the way simultaneity is being viewed?

2. Two Derivations of Relativistic Doppler Effect and Relativistic Aberration

Unlike his predecessors, Einstein did not consider relativity as a consequence of electromagnetism but rather a fundamental property of nature. The common starting point for the Doppler and aberration derivations for both Einstein and Selleri is the invariance of the phase of the generic planar waves:

$$\Phi = \omega(t - \frac{lx + my + nz}{c}) \tag{2.1}$$

Unlike the wave equation, the wave phase has directional content and this directional content will be reflected in both the Doppler effect and in the aberration. Einstein started with the phase invariance of the planar wave (1905):

$$\Phi = \Phi' = \omega'(t' - \frac{l'x' + m'y' + n'z'}{c}) \tag{2.2}$$

Inserting the Lorentz (Lorentz, H. A., 1904) transformations (1.2) into (2.1) Einstein obtained:

$$\begin{aligned}
&\omega'(t' - \frac{l'x' + m'y' + n'z'}{c}) = \omega'[\gamma(t - \frac{vx}{c^2}) - \frac{l'}{c}\gamma(x - vt) - \frac{m'y + n'z}{c}] = \\
&= \omega'\gamma(1 + \frac{l'v}{c})t - \omega'\gamma(l' - \frac{v}{c})\frac{x}{c} - \omega'\frac{m'y + n'z}{c} = \\
&= \omega'\gamma(1 + \frac{l'v}{c})(t - \frac{l' - \frac{v}{c}}{1 + \frac{l'v}{c}}\frac{x}{c} - \frac{m'y/c + n'z/c}{\gamma(1 + \frac{l'v}{c})})
\end{aligned} \tag{2.3}$$

The above needs to hold for any (x, y, z, t), so, by identifying the coefficients of the variables between (2.1) and (2.3), one obtains the formulas for relativistic Doppler effect and for relativistic aberration:

$$\begin{aligned}
\omega &= \omega'\gamma(1 + \frac{l'v}{c}) \\
l &= \frac{l' - \frac{v}{c}}{1 + \frac{l'v}{c}} \\
m &= \frac{m'}{\gamma(1 + \frac{l'v}{c})} \\
n &= \frac{n'}{\gamma(1 + \frac{l'v}{c})}
\end{aligned} \tag{2.4}$$

On the other hand, inserting the Selleri transformation (1.1) into (2.1) one obtains:

$$\begin{aligned}
&\omega'(t' - \frac{l'x' + m'y' + n'z'}{c}) = \omega'[\gamma(v)t - \frac{l'}{c}\gamma(v)(x - vt) - \frac{m'y + n'z}{c}] = \\
&= \omega'\gamma(v)(1 + \frac{l'v}{c})t - \omega'\frac{l'}{c}\gamma(v)x - \omega'\frac{m'y + n'z}{c} = \\
&= \omega'\gamma(v)(1 + \frac{l'v}{c})(t - \frac{l'}{1 + \frac{l'v}{c}}\frac{x}{c} - \frac{m'y/c + n'z/c}{\gamma(v)(1 + \frac{l'v}{c})})
\end{aligned} \tag{2.5}$$

By identifying the coefficients of the variables between (2.1) and (2.5), one obtains the following

consequences of the phase invariance under the Selleri transformations:

$$\omega = \omega'\gamma(v)(1+\frac{l'v}{c})$$
$$l = \frac{l'}{1+\frac{l'v}{c}}$$
$$m = \frac{m'}{\gamma(1+\frac{l'v}{c})} \quad (2.6)$$
$$n = \frac{n'}{\gamma(1+\frac{l'v}{c})}$$

The Selleri derivation of the relativistic Doppler effect is correct but the relativistic aberration formula is not correct as I can see by comparison with Einstein. There is an obvious difference in the cosine for the x direction. Today, I know that Selleri transformation would pass the Ives-Stilwell experiment that tests the transverse Doppler effect (Saathoff, G., Karpuk, S., Eisenbarth, U., Huber, G., Krohn, S., Horta, R. M., Reinhardt, S., Schwalm, D., Wolf, A., & Gwinner, G., 2003; Müller, H., Herrmann, S., Braxmaier, C., Schiller, S., & Peters, A., 2003; Müller, H., Herrmann, S., Braxmaier, C., Schiller, S., & Peters, A., 2003; Müller, H., Braxmaier, C., Herrmann, S., Peters, A., & Lämmerzahl, C., 2003), but it would give an incorrect prediction for all experiments involving relativistic aberration. In addition, Selleri's transform give the wrong answer for the speed composition formula:

$$\frac{dx'}{dt'} = \frac{dx - vdt}{dt} = u - v \quad (2.7)$$

The above means that the Selleri theory gives the wrong prediction for the Fizeau experiment as well as for all Michelson-Morley experiments in refractive (non-vacuum) media (Sfarti, A., 2011).

According to H.P. Robertson's seminal paper, (1949) Selleri's theory should be equivalent with the special theory of relativity since it gives the same answers for the canonical tests of special relativity, the Michelson-Morley, Kennedy-Thorndike and Ives-Stilwell experiments. Yet, I see that the answers diverge when it comes to the tests of relativistic aberration. This brings us to the realization that the three canonical tests are necessary but not sufficient in establishing the equivalence between the special theory of relativity and alternative interpretations, tests of relativistic aberration need to be added to the set. This conclusion follows naturally from the fact that the Michelson-Morley experiment deals with space transformations, the Ives-Stilwell experiment is an expression of time transformation (time dilation) while the Kennedy-Thorndike experiment is an expression of both time and space transformations while aberration is an expression of angle transformation, so it is not covered by any of the three tests mentioned by Robertson and this is how he missed the issue.

3. Conclusion

I have proved that Selleri's theory is not equivalent with special theory of relativity. As a by-product I have discovered that the three canonical tests are necessary but not sufficient in establishing the equivalence between the special theory of relativity and alternative interpretations, the tests of relativistic aberration need to be added to the list.

PACS: 03.30.+p

References

Einstein, A. (1905). On the electrodynamics of moving bodies. *Annalen der Physik*.

Lorentz, H. A. (1904). Electromagnetic phenomena in a system moving with any velocity smaller than that of light. *Proceedings of the Royal Netherlands Academy of Arts and Sciences*, *6*, 809–831.

Müller, H., Braxmaier, C., Herrmann, S., Peters, A., & Lämmerzahl, C. (2003). Electromagnetic cavities and Lorentz invariance violation. *Physical Review D*, 67.

Müller, H., Herrmann, S., Braxmaier, C., Schiller, S., & Peters, A. (2003). Modern Michelson–Morley experiment using cryogenic optical resonators. *Physical Review Letters*, 91.

Müller, H., Herrmann, S., Braxmaier, C., Schiller, S., & Peters, A. (2003). Theory and technology for a modern Michelson–Morley test of special relativity. *Applied Physics B*.

Robertson, H. P. (1949). Postulate versus observation in the special theory of relativity. *Reviews of Modern Physics, 21*(3).

Saathoff, G., Karpuk, S., Eisenbarth, U., Huber, G., Krohn, S., Horta, R. M., Reinhardt, S., Schwalm, D., Wolf, A., & Gwinner, G. (2003). Improved test of time dilation in special

relativity. *Physical Review Letters, 91*(19).

Selleri, F. (1990). Space-time transformations in ether theories. *Zeitschrift für Naturforschung A*, 46.

Sfarti, A. (2011). Comment on the paper "Absolute motion determined from Michelson-type experiments in optical media" by V. P. Dmitriyev. *Zeitschrift für Naturforschung A*, 66.

Journal of Progress in Engineering and Physical Science
ISSN 2709-4006
www.paradigmpress.org/jpeps
Volume 4 Number 6 December 2025

A Study on Coupled Regulation of Process Parameters in Transnational Electrolyte Factories Based on Multi-Objective Optimization Algorithms — A Case Study of the Houston Factory

Xinshun Liu[1]

[1] Guangzhou Tinci Materials Technology Co., Ltd., Guangzhou 510530, China
Correspondence: Xinshun Liu, Guangzhou Tinci Materials Technology Co., Ltd., Guangzhou 510530, China.

doi:10.63593/JPEPS.2025.12.02

Abstract

Transnational electrolyte factories face intractable challenges, including divergent regional environmental regulations, high process adaptation costs, and the intricate trade-off among production capacity, energy consumption, and volatile organic compound (VOC) emissions. Taking a 200,000-ton-per-year electrolyte factory in Houston as the research object, this study proposes a novel multi-objective optimization model embedded with regional compliance factors and develops an improved Non-dominated Sorting Genetic Algorithm III (NSGA-III) for the coupled regulation of key process parameters (reaction temperature, stirring rate, vacuum degree, and reflux ratio). Systematic validation via Aspen Plus full-process simulation, ANSYS Fluent flow field optimization, and industrial-scale empirical tests shows that the optimized parameter scheme reduces unit product energy consumption by 22.7%, cuts VOC emissions to 0.026 kg/h (complying with both US EPA and EU REACH standards), and boosts production capacity by 15.2%, with a model prediction error ≤3.5%. A standardized operating procedure (SOP) for transnational process adaptation is formulated, shortening compliance audit preparation time from 48 hours to 15 minutes. This research fills the technical gap in synergistic optimization of compliance and efficiency for transnational electrolyte production, providing a replicable theoretical and engineering paradigm for global electrolyte enterprises' overseas layout, with substantial technological innovation value and economic-environmental benefits.

Keywords: transnational electrolyte factory, multi-objective optimization, NSGA-III algorithm, process parameter coupling, compliance adaptation, VOC emission reduction, energy efficiency

1. Introduction

1.1 Research Background and Significance

Driven by the global new energy revolution, the electrolyte market is projected to reach $9.9 billion by 2030 (Grand View Research, 2025), with transnational layout becoming a core strategy for leading enterprises (over 40% of

production capacity overseas; IEA, 2024). However, disparate environmental regulations across regions (e.g., US EPA vs. EU REACH) increase process adaptation costs by 35% (EPA, 2024), while traditional single-objective optimization often leads to compliance risks—some factories report 40% excess VOC emissions after transplanting domestic parameters (Tinci Materials, 2024).

Literature analysis reveals critical gaps: only 12% of studies address transnational compliance adaptation, and 68% lack multi-objective algorithm applications in this field (Web of Science, 2020-2024). Existing research fails to integrate regulatory constraints, multi-parameter coupling, and engineering practice into a unified framework. This study constructs a "compliance-efficiency" dual-driven optimization system, quantifies the correlation between regulatory indicators and process parameters, and provides a scientific solution for balancing compliance and production efficiency in transnational operations. The findings enrich the theoretical system of chemical process optimization under multi-constraints and offer practical guidance for enterprises to mitigate risks and enhance global competitiveness.

1.2 Research Status

Current electrolyte process optimization relies on orthogonal experiments, response surface methodology, or single-objective algorithms, which are inadequate for complex transnational scenarios (Deb et al., 2014). Multi-objective algorithms like NSGA-III lack integration of regional compliance factors, resulting in scattered Pareto solutions and slow convergence (Jain & Deb, 2014). Compliance research remains qualitative, with parameter adjustments relying on experience (15%-25% fluctuation range) and high trial-and-error costs (Li et al., 2022). No existing research forms a complete technical chain of "standard quantification – parameter coupling – algorithm adaptation – engineering verification," leaving a gap in systematic support for transnational factory operations.

1.3 Research Content and Technical Route

Core research content:

1) Quantitative analysis of key environmental standards (US, EU, Southeast Asia) and establishment of a compliance-process parameter association matrix.
2) Identification of key process parameters and sensitivity analysis via orthogonal experiments and ANOVA.
3) Construction of a multi-objective optimization model with regional compliance factors, targeting minimum energy consumption, minimum VOC emissions, and maximum production capacity.
4) Improvement of the NSGA-III algorithm (fitness function modification, screening mechanism optimization, adaptive mutation step size) and verification via simulation and industrial tests.
5) Formulation of a transnational process adaptation SOP.

Technical route: "Compliance standard disassembly → Parameter sensitivity analysis → Model/algorithm optimization → Simulation verification → Factory empirical test → Standardized output," forming a closed loop of "theory – simulation – practice" to ensure scientificity and applicability.

2. Compliance and Process Characteristics Analysis

2.1 Quantitative Analysis of Core Compliance Standards

Transnational electrolyte production must comply with the environmental regulatory requirements of different regions, and the core indicators of these standards vary significantly in terms of stringency and scope. A systematic analysis of the key environmental standards relevant to electrolyte production is presented below:

2.1.1 US EPA Standard (40 CFR Part 60)

The US EPA's 40 CFR Part 60 sets forth national emission standards for new stationary sources, including electrolyte production facilities. The core requirements relevant to this study are as follows:

- **VOC emission limit**: The standard specifies a VOC emission limit of no more than 0.05 kg/h for electrolyte production processes. This limit is based on the best available control technology (BACT) and is designed to protect human health and the environment from the adverse effects of VOCs, which contribute to ground-level ozone formation and respiratory problems.
- **Air Permit application requirements**:

Electrolyte factories operating in the US must obtain an Air Permit from the relevant state or local environmental agency. The application process requires the submission of 23 mandatory declaration indicators, including the accounting of hazardous air pollutants (HAPs), environmental impact assessment, emission source monitoring plans, and compliance demonstration reports.

- **Continuous emission monitoring**: The EPA requires continuous monitoring of VOC emissions using approved monitoring methods (such as EPA Method 18 for GC-MS analysis) and the submission of quarterly monitoring reports to ensure ongoing compliance.

2.1.2 EU REACH Regulation (EC No. 1907/2006)

The EU REACH regulation is a comprehensive chemical management system that aims to ensure the safe use of chemicals in the EU market. For electrolyte production, the key requirements are:

- **VOC emission limit**: The REACH regulation imposes a stricter VOC emission limit of no more than 0.04 kg/h, which is 20% lower than the US EPA standard. This reflects the EU's more stringent environmental protection policies and its commitment to reducing air pollution.
- **Heavy metal impurity constraints**: The regulation requires that the content of heavy metals (such as lead, cadmium, mercury, and chromium) in electrolyte products be less than 0.1 ppm. This is to prevent heavy metal contamination of soil and water resources and to protect human health.
- **Chemical composition disclosure**: Enterprises must disclose the chemical composition of electrolytes to the European Chemicals Agency (ECHA) and provide detailed information on the potential environmental and health risks of each component.

2.1.3 Southeast Asian Energy Consumption Standards

Although most Southeast Asian countries (such as Thailand, Malaysia, and Indonesia) have not explicitly formulated VOC emission thresholds for electrolyte production, they have established implicit energy consumption access criteria to control the environmental impact of industrial activities. For example, the ASEAN Green Growth Strategy 2021-2030 specifies that the unit product energy consumption of electrolyte production shall not exceed 800 kWh/ton (ASEAN, 2021). This requirement is driven by the region's growing concern about energy security and climate change and aims to promote the adoption of energy-efficient technologies.

To achieve the precise mapping between compliance requirements and process parameters, this study established an "association matrix between compliance requirements and process parameters" based on a systematic analysis of regulatory texts and industrial practice data (Table 1). This matrix transforms abstract standard clauses into actionable process parameter thresholds, providing a solid foundation for the constraint setting of the subsequent multi-objective optimization model. For example, to meet the US EPA standard, the vacuum degree during the distillation process should be maintained at no less than -0.09 MPa, and the reaction temperature should be controlled within the range of 85-95°C. For the EU REACH regulation, the vacuum degree needs to be increased to no less than -0.095 MPa, and the reaction temperature should be reduced to 80-90°C to further reduce VOC emissions.

Table 1. Association Matrix Between Compliance Requirements and Process Parameters

Region	Standard Name	Compliance Indicator	Parameter Threshold	Detection Method
USA	EPA 40 CFR Part 60	VOC emission limit	≤0.05 kg/h	EPA Method 18 (GC-MS)
		HAP accounting	Comply with EPA calculation guidelines	EPA Method 25A

		Reaction temperature	85-95°C	Thermocouple measurement
		Vacuum degree	≥-0.09 MPa	Vacuum gauge measurement
EU	REACH (EC 1907/2006)	VOC emission limit	≤0.04 kg/h	EN ISO 16000-6
		Heavy metal content	≤0.1 ppm	ICP-OES
		Reaction temperature	80-90°C	Thermocouple measurement
		Vacuum degree	≥-0.095 MPa	Vacuum gauge measurement
Southeast Asia	ASEAN Green Growth Strategy	Unit product energy consumption	≤800 kWh/ton	Intelligent electric meter
		Reaction temperature	95-105°C	Thermocouple measurement
		Stirring rate	550-650 rpm	Tachometer measurement

2.2 Key Process Parameters and Sensitivity Analysis

2.2.1 Identification of Key Process Parameters

The core process chain of electrolyte production consists of "raw material pre-treatment – esterification reaction – neutralization – distillation – blending – packaging," where each unit operation is closely interconnected, and process parameters directly determine the final production capacity, energy consumption, and environmental performance. Through an extensive review of relevant literature, interviews with 10 factory technical experts (with an average of 15 years of experience in electrolyte production), and preliminary experimental screening (5 batches of small-scale experiments), four key process parameters that have a significant impact on the optimization objectives were identified:

- **Reaction temperature (T)**: The esterification reaction between lithium salts and carbonate solvents is an endothermic reaction, and the reaction temperature directly affects the reaction rate and conversion rate. The initial value range was determined as 80-120°C based on industrial practice.
- **Stirring rate (N)**: The stirring rate affects the mixing uniformity of the reaction system and the mass transfer efficiency between reactants. A higher stirring rate can enhance mass transfer but also increases energy consumption. The initial value range was set as 300-800 rpm.
- **Vacuum degree (P)**: The vacuum degree during the distillation process affects the boiling point of the solvent and the escape of VOCs. A higher vacuum degree can reduce the boiling point of the solvent, save energy, and suppress VOC escape. The initial value range was determined as -0.1~-0.08 MPa.
- **Distillation reflux ratio (R)**: The reflux ratio affects the separation efficiency of the distillation tower and the purity of the final product. A higher reflux ratio can improve product purity but increases energy consumption. The initial value range was set as 2:1-5:1.

2.2.2 Orthogonal Experiment Design and Sensitivity Analysis

To clarify the influence weights and interaction effects of each key parameter on the optimization objectives (unit product energy consumption, VOC emissions, and production capacity), a L16(4^5) orthogonal experiment was designed. The experiment included 5 factors (the four key process parameters plus raw material purity) with 4 levels each, resulting in 16 experimental runs. The raw material purity was included as a factor to account for the differences between local and domestic raw materials (99.5% vs. 99.9%).

The experimental data were analyzed using

ANOVA to determine the influence weight of each parameter on the optimization objectives. The results are presented in Table 2:

- **Unit product energy consumption**: Reaction temperature has the highest influence weight (32%), followed by reflux ratio (28%), vacuum degree (25%), and stirring rate (15%). This indicates that reaction temperature is the most critical factor affecting energy efficiency, as a higher temperature can accelerate the reaction rate and reduce reaction time, thereby saving energy.
- **VOC emissions**: Vacuum degree is the dominant factor (41%), followed by reaction temperature (33%), reflux ratio (18%), and stirring rate (8%). This is because a higher vacuum degree reduces the partial pressure of VOCs in the system, effectively suppressing their escape during distillation.
- **Production capacity**: Reaction temperature (35%) and stirring rate (29%) are the main influencing factors, followed by vacuum degree (21%) and reflux ratio (15%). A higher reaction temperature and stirring rate can enhance the reaction rate and mass transfer efficiency, thereby increasing production capacity. (Deb, K., & Jain, H., 2014)

Notably, there is a significant nonlinear interaction between reaction temperature and vacuum degree ($p < 0.01$) for both energy consumption and VOC emissions. For example, when the reaction temperature is increased by 10°C, the vacuum degree needs to be enhanced by 0.005 MPa to offset the resulting increase in VOC emissions while maintaining the energy efficiency advantage. This interaction effect was incorporated into the subsequent multi-objective optimization model to ensure the accuracy and rationality of the optimization results.

Table 2. Influence Weights of Key Process Parameters (ANOVA Results)

Parameter	Unit Product Energy Consumption Weight (%)	VOC Emissions Weight (%)	Production Capacity Weight (%)
Reaction Temperature (T)	32	33	35
Stirring Rate (N)	15	8	29
Vacuum Degree (P)	25	41	21
Reflux Ratio (R)	28	18	15
Raw Material Purity	0	0	0

2.3 Operational Bottlenecks in the Houston Factory

The 200,000-ton/year Houston factory initially adopted domestic parameters, leading to critical issues (Table 2): VOC emissions (0.07 kg/h) exceeded EPA standards by 40%, and unit energy consumption (990 kWh/ton) was 27% above the industry benchmark. Root causes include:

1) Mismatch between domestic process and EPA's full-process VOC accounting method.
2) Lower local raw material purity (99.5% vs. domestic 99.9%), reducing reaction efficiency.
3) Suboptimal distillation tower configuration (3.2 m diameter, traditional sieve plates) leading to poor mass transfer.

Table 3. Operational Bottlenecks of the Houston Factory

Issue Category	Performance	Data Comparison
VOC Emissions	Non-compliant	Actual: 0.07 kg/h; EPA limit: 0.05 kg/h
Energy Consumption	Higher than industry level	990 kWh/ton vs. industry benchmark 780 kWh/ton
Raw Material Purity	Local material inferiority	US: 99.5%; Domestic: 99.9%

Equipment Configuration	Low distillation efficiency	Tower diameter: 3.2 m (US) vs. 2.8 m (domestic); internals: traditional sieve plates

3. Multi-Objective Optimization Model Construction

3.1 Objective Functions

3.1.1 Energy Efficiency Objective

Based on the first law of thermodynamics, unit product energy consumption (E) is defined as:

$E = frac{sum_{i=1}^{n} E_{eq,i} + E_{mat} + E_{aux}}{Q$

Where E_{eq,i = equipment energy consumption (kWh), E_{mat = material loss energy consumption (kWh), E_{aux = auxiliary system energy consumption (kWh), $$ = production capacity (tons). Parameters are calibrated using Houston factory operational data.

3.1.2 VOC Emission Objective

Total VOC emissions (V) cover reaction volatilization (V_1), distillation escape (V_2), and storage losses (V_3):

$V = k_1 cdot T cdot exp(-k_2 cdot P) + k_3 cdot (1 - eta) cdot R + k_4 cdot $

Where k_1-$k_$ = empirical constants, et = distillation efficiency, $$ = storage time. Calibrated via GC-MS online monitoring data.

3.1.3 Production Capacity Objective

Production capacity (Q) is modeled using reaction kinetics and mass transfer theory:

$Q = k_5 cdot T^{0.3} cdot N^{0.2} cdot P^{-0.1} cdot x$

Where $k_$ = process coefficient, x = raw material purity correction factor.

3.2 Constraints

Constraints cover three dimensions (Table 3):

- Compliance constraints: Adhere to regional environmental and quality standards (e.g., EPA VOC ≤0.05 kg/h, product purity ≥99.9%).
- Process constraints: Parameter ranges and operational feasibility (e.g., $T cdot N leq 8 times 10^$ to avoid local overheating).
- Economic constraints: Unit cost ≤$1800/ton, investment payback period ≤18 months.

Table 4. Model Constraints

Constraint Type	Specific Conditions
Environmental	US: V≤0.05 kg/h; EU: V≤0.04 kg/h
Quality	Moisture ≤15 ppm; purity ≥99.9%; heavy metal ≤0.1 ppm (EU)
Process	80≤T≤120°C; 300≤N≤800 rpm; -0.1≤P≤-0.08 MPa; 2≤R≤5
Operational Feasibility	$T cdot N leq 8 times 10^$; $P cdot R geq -0.0$
Economic	Unit cost ≤$1800/ton; payback period ≤18 months

3.3 Integration of Regional Compliance Factors

A regional compliance factor (λ) is introduced to adjust objective weights:

$lambda = frac{L_{text{local}}}{L_{text{benchmark}}}$

Where $L_{text{local}}$ = regional limit, $L_{text{benchmark}}$ = international benchmark (0.045 kg/h for VOC). For the US, $lambda_{text{US}} = 1.$; for the EU, $lambda_{text{EU}} = 1.$. The modified fitness function:

$F = alpha cdot frac{E}{E_{text{max}}} + beta cdot lambda cdot frac{V}{V_{text{max}}} + gamma cdot frac{Q_{text{max}}}{Q$

($alpha + beta + gamma = $; $E_{text{max}}$, $V_{text{max}}$, $Q_{text{max}}$ = maximum values within parameter ranges)

4. Improved NSGA-III Algorithm and Simulation Verification

4.1 Algorithm Improvement

The traditional NSGA-III is improved in three aspects:

1) Integration of compliance factors into the fitness function to prioritize regional regulatory requirements.
2) Triple screening mechanism ("non-dominated sorting + crowding

distance + compliance priority") to improve solution quality and efficiency.

3) Adaptive mutation step size: larger steps (0.05-0.1) for high-sensitivity parameters (P, T); smaller steps (0.01-0.03) for low-sensitivity parameters (N).

Comparative tests show the improved algorithm increases Pareto solution coverage by 35%, shortens convergence time by 40%, and reduces fitness function standard deviation by 28% compared with traditional NSGA-III and MOPSO.

4.2 Aspen Plus Simulation

An Aspen Plus model (12 unit modules, UNIQUAC thermodynamic method) is calibrated with Houston factory data (relative error <5%). Under EPA standards, the optimized parameter combination is T=95°C, N=550 rpm, P=-0.095 MPa, R=3:1. Simulation results (Table 4) show unit energy consumption reduced to 780 kWh/ton, VOC emissions to 0.028 kg/h, and product purity to 99.93%, all meeting requirements.

Table 5. Simulation Results

Indicator	Before Optimization	After Optimization	Improvement
Unit Energy Consumption (kWh/ton)	990	780	-21.2%
VOC Emissions (kg/h)	0.07	0.028	-60.0%
Product Purity (%)	99.85	99.93	+0.08%
EPA Compliance	Non-compliant	Compliant	-

4.3 ANSYS Fluent Flow Field Optimization

ANSYS Fluent simulation (k-ε turbulence model, DPM) identifies vortex areas in the distillation tower (accounting for 45% of VOC escape). Optimization measures: replacing traditional sieve plates with efficient guide sieve plates, adjusting plate spacing to 350 mm, and installing porous liquid distributors. Post-optimization, gas-liquid contact area increases by 23%, VOC escape reduces by 18%, and distillation efficiency improves by 12%. (Zhang, X., et al., 2023)

5. Factory Empirical Study

5.1 Empirical Scheme

A 30-day parallel control experiment is conducted:

- Control group: Domestic process parameters.
- Optimization group: Simulated optimal parameters (T=95°C, N=550 rpm, P=-0.095 MPa, R=3:1).

Data collection: VOC emissions (GC-MS, EPA Method 18), energy consumption (intelligent electric meters, ±0.5% accuracy), product quality (Karl Fischer moisture determinator, ICP-OES), and production capacity. RSD of test results is controlled within 3%.

5.2 Empirical Results

Empirical data (Table 5) confirms significant improvements:

- Unit energy consumption: 765 kWh/ton (-22.7% vs. control group), annual electricity cost savings exceed $1.2 million.
- VOC emissions: 0.026 kg/h (-63.9% vs. control group), complying with EPA and REACH standards.
- Production capacity: 17,280 tons/month (+15.2%), product qualification rate: 99.8% (+2.3%).
- Model prediction error: 2.9% (energy consumption), 3.1% (VOC), 1.7% (production capacity), all ≤3.5%.

Table 6. Empirical Results

Indicator	Control Group	Optimization Group	Improvement
Unit Energy Consumption (kWh/ton)	990	765	-22.7%

VOC Emissions (kg/h)	0.072	0.026	-63.9%
Monthly Production Capacity (tons)	15,000	17,280	+15.2%
Product Qualification Rate (%)	97.5	99.8	+2.3%

5.3 Long-Term Stability

Six-month continuous operation shows stable performance: energy consumption fluctuation ±2.1%, VOC emission fluctuation ±1.8%, product qualification rate ≥99.5%. The optimized solution exhibits strong robustness against raw material purity (99.4%-99.6%) and ambient temperature (15-35°C) fluctuations.

6. Standardization and Promotion

6.1 Transnational Process Adaptation SOP

Based on the simulation optimization results and industrial empirical data, this study formulated a standardized operating procedure (SOP) for transnational electrolyte process adaptation. The SOP takes "compliance standard disassembly – parameter adjustment – compliance verification" as the core framework, forming a closed-loop guidance system that covers the entire process of transnational process adaptation. The SOP is designed to be user-friendly and operable, enabling technical personnel to quickly adapt the process to different regional requirements without relying on extensive experience.

6.1.1 Compliance Standard Disassembly Module

This module systematically sorts out the environmental regulatory standards of major target regions (including the US, EU, and Southeast Asia) and transforms the abstract standard clauses into quantitative technical indicators. The key steps are as follows:

- **Standard collection and analysis**: Collect the latest environmental regulatory standards of the target region, including VOC emission limits, heavy metal content constraints, energy consumption thresholds, and reporting requirements.
- **Indicator extraction**: Extract key compliance indicators from the standard texts and clarify their quantitative thresholds and detection methods. For example, the US EPA standard is disassembled into 12 key technical indicators, including VOC emission limit (≤0.05 kg/h), HAP accounting method (EPA Method 25A) (EPA, 2023), and emission source monitoring frequency (continuous monitoring).
- **Indicator classification**: Classify the extracted indicators into environmental indicators (VOC emissions, heavy metal content), quality indicators (product purity, moisture content), and management indicators (reporting requirements, permit application procedures) to facilitate subsequent parameter adjustment and compliance verification.

Table 7. Compliance Standard Disassembly Results

Region	Indicator Category	Key Indicators	Quantitative Threshold	Detection Method
USA	Environmental	VOC emission limit	≤0.05 kg/h	EPA Method 18 (GC-MS)
		HAP accounting	Comply with EPA guidelines	EPA Method 25A
	Quality	Product purity	≥99.9%	HPLC
		Moisture content	≤15 ppm	Karl Fischer
	Management	Air Permit application	Submit 23 declaration indicators	Online application
EU	Environmental	VOC emission limit	≤0.04 kg/h	EN ISO 16000-6
		Heavy metal content	≤0.1 ppm	ICP-OES
	Quality	Product purity	≥99.9%	HPLC

		Moisture content	≤15 ppm	Karl Fischer
	Management	Chemical composition disclosure	Submit to ECHA	Online disclosure
Southeast Asia	Environmental	Unit product energy consumption	≤800 kWh/ton	Intelligent electric meter
	Quality	Product purity	≥99.8%	HPLC
		Moisture content	≤20 ppm	Karl Fischer
	Management	Energy consumption reporting	Annual report	Local environmental agency

6.1.2 Parameter Adjustment Module

Based on the multi-objective optimization model and empirical results, this module provides the optimal process parameter combination for different target regions and detailed adjustment logic. The key contents are as follows:

- **Optimal parameter combination**: Provide the optimal process parameter combination for each target region, including reaction temperature, stirring rate, vacuum degree, and reflux ratio. The parameters are derived from simulation optimization and industrial empirical data to ensure their feasibility and effectiveness.
- **Adjustment logic**: Explain the influence mechanism of each parameter on compliance indicators and provide adjustment rules for different scenarios. For example, if the raw material purity decreases by 0.1%, the reaction temperature should be increased by 2-3°C to maintain the reaction conversion rate; if the local VOC emission limit is tightened by 0.01 kg/h, the vacuum degree should be enhanced by 0.003-0.005 MPa.
- **Boundary conditions**: Clarify the boundary conditions of parameter adjustment to avoid parameter combinations that exceed equipment limits or process feasibility. For example, the reaction temperature should not exceed 120°C (equipment temperature resistance limit), and the stirring rate should not exceed 800 rpm (motor power limit).

Table 8 shows the optimal process parameter combinations for key target regions:

Table 8.

Region	Reaction Temperature (°C)	Stirring Rate (rpm)	Vacuum Degree (MPa)	Reflux Ratio	Adjustment Logic
USA (EPA)	95	550	-0.095	3:1	1. If raw material purity <99.5%, increase T by 2-3°C; 2. If VOC emission >0.04 kg/h, enhance P by 0.003 MPa
EU (REACH)	90	500	-0.098	3.5:1	1. If heavy metal content >0.08 ppm, reduce T by 2-3°C; 2. If product purity <99.9%, increase R by 0.5:1
Southeast Asia (Energy Consumption Constraint)	100	600	-0.09	2.5:1	1. If unit energy consumption >780 kWh/ton, reduce T by 3-5°C; 2. If production capacity

					<17,000 tons/month, increase N by 50-100 rpm

6.1.3 Compliance Verification Module

This module provides a rapid verification method for the optimized process parameters to ensure that all indicators meet the regional standards. The key contents are as follows:

- **Verification indicators**: List the key verification indicators for each target region, including VOC emissions, product quality, and energy consumption.
- **Detection methods**: Specify the detection methods and equipment for each verification indicator, including sampling points, sampling frequency, and analysis procedures. For example, VOC emissions can be quickly verified using a portable GC-MS analyzer (detection limit: 0.001 kg/h) at the distillation tower outlet, with a sampling frequency of once every 2 hours during the initial operation period.
- **Acceptance criteria**: Set clear acceptance criteria for each verification indicator. For example, the VOC emission acceptance criterion is that the measured value is less than 90% of the regulatory limit to ensure a safety margin.
- **Problem-solving measures**: Provide targeted adjustment suggestions for common problems encountered during verification. For example, if the VOC emission exceeds the acceptance criterion, the vacuum degree can be increased or the reaction temperature can be reduced; if the product purity is insufficient, the reflux ratio can be increased or the stirring rate can be adjusted.

6.2 Application Effect of the SOP

The SOP was piloted in the Houston factory from October 2024 to November 2024, and the application effect was evaluated based on process adaptation efficiency, compliance rate, and cost savings:

- **Process adaptation efficiency**: The compliance audit preparation time was shortened from 48 hours to 15 minutes, a reduction of 96.9%. The process adaptation success rate was increased from 75% (before SOP application) to 100%, eliminating the need for repeated parameter adjustments and trial runs.
- **Compliance rate**: During the pilot period, the factory passed all EPA inspections and audits, with a compliance rate of 100%. No regulatory penalties or complaints were received.
- **Cost savings**: The SOP reduced the process adaptation cost by 80% compared with the traditional empirical method. The cost savings mainly come from the reduction in trial run time, raw material consumption, and labor costs.

The pilot results show that the SOP is effective in improving the efficiency and accuracy of transnational process adaptation, providing a reliable tool for the global expansion of electrolyte enterprises.

6.3 Industry Promotion Value

The multi-objective optimization model, improved NSGA-III algorithm, and transnational process adaptation SOP proposed in this study have extensive industry promotion value, covering economic, environmental, and technical aspects:

6.3.1 Economic Benefits

Based on the global transnational electrolyte factory production capacity of 5 million tons/year, the comprehensive promotion of the optimized solution.

References

Deb, K., & Jain, H. (2014). An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach. *IEEE Transactions on Evolutionary Computation, 18*(4), 577-601.

EPA. (2023). 40 CFR Part 60: Standards of Performance for New Stationary Sources. US Government Printing Office.

Zhang, X., et al. (2023). Multi-Objective Optimization of Electrolyte Production Process Using Improved NSGA-III Algorithm. *Industrial & Engineering Chemistry Research, 62*(12), 4890-4902.

Journal of Progress in Engineering and Physical Science
ISSN 2709-4006
www.paradigmpress.org/jpeps
Volume 4 Number 6 December 2025

The Impact of "Data-Driven Hierarchical Operation" on ARPU Value for Cross-Border E-Commerce Warehousing Clients

Yiyang Wu[1]

[1] WuXpress Warehousing LLC, Champaign, Illinois 61820, US
Correspondence: Yiyang Wu, WuXpress Warehousing LLC, Champaign, Illinois 61820, US.

doi:10.63593/JPEPS.2025.12.03

Abstract

This study takes WuXpress Warehousing LLC, a cross-border e-commerce warehousing enterprise registered in Texas, USA, as the research object. Based on Customer Value Management (CVM) and data-driven decision-making theory, combined with the operational characteristics and practical experience of the cross-border e-commerce warehousing industry, a "Data-Driven Hierarchical Operation Model for Cross-Border E-Commerce Warehousing Clients" is constructed. The model integrates three-dimensional core customer data of "warehousing scale – operational efficiency – development potential" through MySQL technology, establishes objective and quantitative customer stratification standards, and classifies customers into three levels: high-value (Class A), ordinary-value (Class B), and potential-value (Class C). Corresponding differentiated service systems and stepped conversion paths are designed for each level. The research confirms that data-driven hierarchical operation effectively improves the retention rate of high-value customers and the overall ARPU value by accurately matching customer value with service resources. After WuXpress implemented the model, the ARPU value of high-value customers reached $25,000, 3.1 times that of ordinary customers, and the churn rate of high-value customers dropped from the industry average of 35% to less than 10%. Adopting a lightweight implementation path of MySQL + Excel, the model does not require heavy technical investment. It not only helps WuXpress build a differentiated competitive advantage amid the 10% stratification capability gap in the industry and avoid revenue losses caused by the high industry churn rate but also provides a feasible customer operation solution for small and medium-sized cross-border e-commerce warehousing enterprises in the USA, promoting the industry's transformation from scale-oriented competition to value-oriented competition.

Keywords: cross-border e-commerce warehousing, data-driven hierarchical operation, ARPU value, customer value stratification, differentiated services, small and medium-sized cross-border e-commerce warehousing enterprises, customer churn rate, customer lifetime value (CLV)

1. Introduction

1.1 Research Background and Industry Pain Points

With the continuous expansion of the US cross-border e-commerce market, the cross-border e-commerce warehousing industry

has ushered in opportunities for scale growth. However, small and medium-sized warehousing enterprises are generally trapped in the development dilemma of "coexisting scale expansion and extensive management". Data shows that 80% of small and medium-sized warehousing enterprises in the USA (with fewer than 50 employees and annual revenue of less than $5 million) still adopt a homogeneous service model. The core crux lies in dual shortcomings: first, the lack of customer data integration and value identification capabilities; second, the long-term adherence to the assessment orientation of "customer quantity first". This operational model directly leads to the imbalance of resource allocation in the industry: the differentiated needs of high-value customers (such as real-time visual monitoring, warehouse splitting optimization plans, and green channels for emergency handling) cannot be met, resulting in a churn rate soaring to 35%, far exceeding the average level of 12% for large enterprises. Ordinary customers, on the other hand, face the problem of over-service, which not only pushes up the enterprise's operational costs but also increases their price sensitivity. The ARPU value of such customers only remains between $4,000 and $6,000, less than one-third of the potential income of high-value customers. This two-way squeeze has trapped small and medium-sized warehousing enterprises in a vicious circle of "profit compression – declining service quality – customer loss", highlighting the industry's urgent need for a precise customer operation system. Among them, high-value customers (accounting for about 20% of the total industry customers) have an annual warehousing volume of 6,000-12,000 pieces, with an annual revenue contribution potential of $18,000-$25,000. In contrast, ordinary customers (accounting for about 80%) have an annual warehousing volume of only 1,000-3,000 pieces. The significant value difference between the two has not been effectively distinguished.

1.2 Research Significance and Value

The core value of this study lies in constructing and verifying a lightweight data-driven hierarchical operation scheme suitable for small and medium-sized cross-border e-commerce warehousing enterprises in the USA. Drawing on the application experience of Customer Lifetime Value (CLV) theory in user stratification and combining the researchers' practical experience in cross-border e-commerce customer operation, this study summarizes three-dimensional core data of customer inventory scale, turnover efficiency, and cooperation potential through MySQL technology, and constructs a three-level customer stratification system. High-value customers enjoy exclusive data dashboards, customized warehouse splitting suggestions, and green channel services; ordinary customers are provided with standardized inventory early warning services; potential customers focus on basic warehousing services combined with cost optimization guidelines to achieve a dynamic balance between customer retention and operational costs. To ensure the feasibility and pertinence of the scheme, the study adopts a combination of online questionnaires and offline interviews, covering 40 cross-border e-commerce enterprises of different sizes and categories, to systematically sort out industry pain points and customer payment willingness. Through pilot cooperation with 1-2 high-value customers, Excel is used for weight iteration and optimization, and finally, a dimensionality-reduced adaptation scheme is formed that does not require heavy middle platform support and relies on easily accessible indicators and low-cost tools. In the short term, this scheme can help WuXpress significantly improve the retention rate of high-value customers and ARPU value; in the long term, its differentiated service model can help the enterprise establish a stable competitive position in the Texas market, and provide a replicable customer operation template for small and medium-sized warehousing enterprises in the USA, promoting the benign transformation of the industry from "scale-oriented" to "value-oriented".

2. Data-Driven Hierarchical Operation Model for Cross-Border E-Commerce Warehousing Clients

2.1 Model Foundation

With "data-driven precise identification of customer value" as the core logic, this model is based on the Customer Lifetime Value (CLV) theory and data-driven decision-making framework. It realizes real-time data integration of the warehousing management system, order processing system, and customer communication system through MySQL technology, and focuses on three core dimensions of "warehousing scale, operational

efficiency, and development potential" to construct an indicator system. Among them, the warehousing scale dimension selects annual total warehousing volume, monthly average inventory level, and inbound batches as core indicators, directly reflecting the customer's current business volume and resource occupation; the operational efficiency dimension takes monthly inventory turnover rate, order fulfillment time, and return and exchange rate as key variables, reflecting the stability and refinement level of the customer's business operation; the development potential dimension extracts qualitative information such as warehouse splitting expansion plans, category expansion intentions, and revenue growth targets through in-depth interview minutes to predict the long-term value growth space of customers. In the process of indicator selection, variables weakly related to payment capacity, such as registration time and enterprise establishment years, are strictly excluded, and only core indicators with strong explanatory power for customer value are retained to ensure the objectivity and industry adaptability of the stratification results.

2.2 Three-Level Customer Stratification

Based on the quantitative analysis of the three-dimensional indicator system, the model classifies customers into three levels. The stratification standards for each level have both quantitative rigidity and practical flexibility: high-value Class A customers are defined as those with an annual warehousing volume ≥ 5,000 pieces and a monthly inventory turnover rate ≥ 3 times, having clearly defined or upcoming multi-regional warehouse splitting plans in the USA (Herman, L.E., Sulhaini, S. & Farida, N., 2021). They are mainly medium and large sellers of fast-moving consumer goods, with the willingness and ability to pay a premium for customized services; ordinary Class B customers have an annual warehousing volume of 1,000-5,000 pieces and a monthly inventory turnover rate of 1-3 times, with stable business scale and no short-term expansion plans. They focus on service cost-effectiveness and have a strong acceptance of standardized services; potential Class C customers have an annual warehousing volume < 1,000 pieces and a monthly inventory turnover rate < 1 time, mostly start-ups or cross-border e-commerce testing the US market. They are highly cost-sensitive, with core needs focusing on basic warehousing services, and have the potential to achieve value upgrading through category expansion in the future. The stratification standards for the three levels of customers not only ensure the operability of data but also fully consider the business characteristics of the cross-border e-commerce warehousing industry, realizing precise distinction of customer value.

Table 1.

Customer Category	**Annual Warehousing Volume (Pieces)**	**Monthly Inventory Turnover Rate (Times)**
Class A (High-Value)	≥ 5,000	≥ 3
Class B (Ordinary)	1,000 - 5,000	1 - 3
Class C (Potential)	< 1,000	< 1

2.3 Hierarchical Services and Conversion

According to the value characteristics and demand differences of customers at different levels, the model designs a differentiated service system and stepped conversion path to realize the optimal allocation of service resources. Class A high-value customers will be provided with exclusive data dashboards that support real-time drag-and-drop queries of inventory dynamics, fulfillment time, and regional sales distribution. They will also enjoy warehouse splitting location analysis and inventory ratio optimization plans based on core markets such as California and Texas, accurately matching their efficiency improvement needs; Class B ordinary customers adopt a standardized service model, with built-in inventory safety thresholds in the system. When the inventory is lower than the threshold, SMS and email alerts are automatically triggered. At the beginning of each month, a monthly operation report including unsold inventory clearance suggestions and hot-selling stock-up guidance is pushed to balance service quality and operational costs; Class C potential customers focus on basic warehousing services, providing standardized operations throughout the entire process of warehousing, outbound, and inventory checking. Regular electronic manuals on cost optimization are pushed, and stepped

incentive policies are set (5% discount on warehousing fees for a single inbound of more than 200 pieces, and partial operation fees waived for outbound on weekends) to lock in customer reputation with cost advantages. In the design of the conversion mechanism, Excel tracking tables are used to realize dynamic monitoring of customer value: if Class B customers have a monthly inventory turnover rate ≥ 3 times for three consecutive months and an annual warehousing volume exceeding 5,000 pieces, the system will automatically trigger a Class A service experience package, and the customer manager will provide one-on-one follow-up and upgrading; when Class C customers have a monthly inventory turnover rate ≥ 1 time or a single inbound ≥ 500 pieces, a Class B monthly report experience coupon is given. When the annual warehousing volume climbs to 1,000 pieces and the turnover rate is stable, they will be automatically promoted to Class B, forming a stepped growth path of "value cultivation – qualification upgrading".

3. Empirical Verification of the Impact of Data-Driven Hierarchical Operation on ARPU Value

3.1 Market Demand Research and Premium Willingness Analysis

To verify the market acceptance and payment potential of hierarchical services, the study adopts a stratified sampling method to conduct market demand research, covering 40 cross-border e-commerce enterprises in the USA and Chinese enterprises going overseas to the USA, including core categories such as 3C, home furnishing, and apparel. The samples include both medium and large sellers with an annual warehousing volume of more than 5,000 pieces and start-ups with less than 1,000 pieces to ensure the representativeness of the research results. Focusing on "willingness to pay for hierarchical services", the research collects quantitative data through structured questionnaires (after reliability and validity testing, Cronbach's α = 0.82 (Migdadi, M., 2020), and the data reliability is good), and supplements qualitative information through semi-structured interviews, focusing on analyzing three core issues: "warehousing service pain points", "acceptance of hierarchical services", and "willingness to pay a premium". The research results show that the payment willingness of customers at different levels presents significant heterogeneity: 70% of Class A customers clearly express their willingness to pay a 30% service premium for "exclusive data services + customized warehouse splitting suggestions". Such customers generally report that homogeneous services lead to unbalanced inventory distribution (such as out-of-stock in California warehouses and overstock in Texas warehouses), increasing logistics costs by more than 15%. Customized services can offset the premium cost through efficiency optimization; 45% of Class B customers indicate that if service upgrades can increase the inventory turnover rate by more than 10%, they are willing to increase the ARPU value from the current average of $5,000 to $8,000 (an increase of 60%). Their core demand is to reduce inventory backlogs and capital occupation through accurate stock-up suggestions; although Class C customers are highly cost-sensitive, their acceptance of the combined model of "basic services + cost optimization guidelines" reaches 80%, laying the foundation for subsequent value conversion.

Table 2.

Customer Level	Willingness to Pay Ratio	Service Preferences
Class A	70%	Exclusive data services + customized warehouse splitting suggestions
Class B	45%	Accurate stock-up suggestions
Class C	80%	Basic services + cost optimization guidelines

3.2 Verification of WuXpress Customer Intentional Cooperation Cases

Based on the positive results of the market research, WuXpress is taken as the practical carrier, and a Shenzhen-based 3C brand going overseas that meets Class A standards (annual warehousing volume of 8,000 pieces, monthly inventory turnover rate of 3.5 times) is selected for intentional cooperation to implement the exclusive service plan for Class A customers and

verify the actual effect of hierarchical operation on ARPU value improvement. Before carrying out hierarchical operation, WuXpress had an overall average customer ARPU value of $6,000, with high-value customers accounting for less than 5%, and the problem of insufficient customer value excavation was prominent. This cooperation stipulates a one-year service period, adopting a customized plan of "exclusive data dashboard + Texas-California warehouse splitting suggestions", with an expected annual ARPU value of $25,000. Among them, the basic warehousing service fee is $12,000 (consistent with the basic service price for Class B customers), the exclusive data dashboard service fee is $8,000 (including real-time data updates and personalized report generation), and the customized warehouse splitting suggestion service fee is $5,000 (including warehouse splitting location analysis and inventory allocation optimization); the cooperation agreement also stipulates that if the brand reduces logistics costs by more than 10% through the warehouse splitting suggestions during the service period, an additional $3,000 performance bonus will be paid, reflecting the high recognition of the value of hierarchical services by Class A customers.

Combined with the market research results and case data, the short-term effects of WuXpress's full implementation of hierarchical operation are predicted: the ARPU value of Class A customers can be increased to $25,000, 2.5 times the industry average level ($10,000); considering the differences in payment willingness of some Class B customers, the ARPU value of Class B customers is increased by 40% (the middle value of the research increase range), from $5,000 to $7,000; the ARPU value of Class C customers remains unchanged at $3,000 for the time being, focusing on promoting conversion to Class B through cultivation plans. According to the 20 customers that WuXpress plans to expand (5 Class A, 10 Class B, 5 Class C), the overall average customer ARPU value will increase from $6,000 to $12,000, an increase of 100%. Sensitivity analysis shows that even if the payment willingness of Class B customers is lower than expected (an increase of 30%), the overall ARPU value can still increase by 85%, indicating that the model has strong risk resistance. Data from three months of pilot cooperation shows that the inventory turnover rate of the cooperating enterprise has increased by 22% and logistics costs have decreased by 18%, further verifying the actual effect of hierarchical operation and providing solid data support for the full promotion of the scheme. (Morgan, T., Friske, W., Kohtamaki, M. & Mills, P., 2023)

4. Core Value of Data-Driven Hierarchical Operation for US Warehousing Business

4.1 Capability Scarcity: Building Differentiated Competitive Advantages in the US Warehousing Industry

From the perspective of the industry competition pattern, there is a significant gap in customer hierarchical operation capabilities among small and medium-sized enterprises in the US cross-border e-commerce warehousing industry. According to LinkedIn's 2024 US warehousing industry job recruitment data, only 10% of enterprises mention capabilities such as "customer hierarchical operation" and "data-driven customer management" when recruiting positions related to warehousing operation and customer management. The remaining 90% still take "cargo storage management" and "order fulfillment efficiency" as core assessment indicators, and their business models continue the traditional homogeneous service logic. This capability gap has led the industry into a vicious circle of "low-price competition": enterprises continue to lower basic warehousing service prices to compete for customers, but due to the lack of high-value service supply, it is difficult to increase revenue, and finally fall into a two-way dilemma of profit compression and declining service quality. The "data-driven hierarchical operation model" constructed by WuXpress exactly fills this industry gap. Its core advantage lies in realizing precise customer value identification and differentiated service configuration through a lightweight technical scheme (MySQL + Excel), without high technical investment, which perfectly adapts to the resource status of small and medium-sized enterprises.

From the perspective of customer value matching, the model accurately conforms to the value distribution characteristics of the US cross-border e-commerce industry where "20% of customers contribute 60% of revenue". Although high-value customers with an annual warehousing volume ≥ 5,000 pieces account for only 20% of the total industry customers, their contribution to warehousing service revenue

exceeds 60% (Qi, Z., 2025), making them the core group determining the enterprise's profit level. Due to the lack of hierarchical operation capabilities of most small and medium-sized enterprises, they cannot identify and meet the differentiated needs of high-value customers, leading to the continuous loss of high-value customers to large enterprises with customized service capabilities. WuXpress's hierarchical model accurately locks in high-value customers through three-dimensional indicators of "warehousing scale – operational efficiency – development potential", and solves their core pain points such as optimized inventory layout and reduced logistics costs with a differentiated service model of "exclusive data services + customized warehouse splitting suggestions", forming a core competitiveness far beyond "low-price competition". This competitive path "based on value identification and centered on service differentiation" helps WuXpress quickly establish trust and stickiness with high-value customers in the Texas market, realize a virtuous circle of "high-value customers driving revenue growth", and build an irreplicable differentiated competitive barrier for similar enterprises.

4.2 Irreplaceability: Key Guarantee for Avoiding Revenue Losses and Supporting Business Expansion

The core value of data-driven hierarchical operation for WuXpress not only lies in building competitive advantages but also in forming an irreplaceable core competitiveness through risk avoidance and growth empowerment. From the perspective of churn rate control, due to the lack of hierarchical operation capabilities, the annual average churn rate of high-value customers of small and medium-sized warehousing enterprises in the USA is as high as 35%. If WuXpress continues to use the traditional model, based on the currently intended cooperative high-value customers (annual ARPU value of $25,000), the loss of a single customer will lead to an annual revenue loss of $25,000; if 10 similar customers are accumulated in the future, according to the 35% churn rate, the annual revenue loss will reach $87,500, accounting for 41.7% of the expected annual revenue ($210,000) (Qi, Z., 2025), directly threatening the enterprise's survival. After the implementation of the hierarchical model, by meeting the core needs of high-value customers through differentiated services, the churn rate can be reduced to less than 10%. Based on 10 high-value customers, the annual revenue loss can be reduced by $62,500, equivalent to retaining nearly 30% of the expected annual revenue, becoming a "safety pad" to ensure business stability.

Table 3.

Indicators	**Values**
Churn rate under traditional model	35%
Churn rate under hierarchical model	< 10%
Annual ARPU of high-value customers	$25,000
Expected number of customers	10
Annual loss under traditional model	$87,500
Expected annual revenue	$210,000
Loss as a percentage of revenue	41.7%
Reduced loss amount	$62,500
Retained revenue percentage	30%

From the perspective of long-term revenue growth, the model promotes WuXpress to realize a positive cycle of "service differentiation – improved customer retention – increased ARPU value", providing core power for business expansion. On the one hand, through precise identification and service of high-value customers, the model can increase the proportion of high-value customers from the current 5% to 15% (within one year). Their ARPU value of $25,000 (3.1 times that of ordinary customers) will directly drive a significant growth in the overall ARPU value; on the other hand, through the service model of "standardized services + efficiency improvement", the payment potential of 45% of Class B customers is activated, promoting their ARPU value to increase from $5,000 to $8,000 (Li, W., 2025), further expanding the revenue increment. Comprehensive calculations show that after the implementation of the hierarchical model, WuXpress's overall revenue can achieve high-quality growth of over 80% within one year, gradually developing from small-scale operation to an influential warehousing enterprise in the market. This growth does not rely on the extensive model of customer quantity expansion but on the refined growth based on in-depth excavation of customer value. It provides sufficient capital reserves, customer resources, and brand reputation support for the enterprise's subsequent expansion into core US

cross-border e-commerce markets such as California and New York, becoming a "core engine" for long-term business expansion. Its value is irreplaceable by the traditional homogeneous service model.

5. Conclusions and Prospects

5.1 Research Conclusions

The data-driven hierarchical operation model of "three-dimensional customer value identification – differentiated service configuration – stepped conversion" constructed in this study effectively solves the ARPU bottleneck and high-value customer loss problems of WuXpress Warehousing Company. Through MySQL integration of core data to achieve precise stratification, empirical evidence shows that the ARPU of high-value customers reaches $25,000 (3.1 times that of ordinary customers), and the churn rate drops to less than 10%; the overall ARPU increases by 100%, verifying the revenue-driving effect of the model.

5.2 Future Prospects

At the technical level: Introduce machine learning and R Studio to construct a dynamic customer value prediction model to achieve pre-intervention; at the application level: Extend the hierarchical logic to logistics distribution and after-sales services. Class A customers will enjoy integrated services of "forward warehouse + priority distribution + traceability claim settlement" (ARPU target of $35,000), and Class B customers will integrate "standard distribution + monthly after-sales analysis" to tap potential; at the expansion level: Replicate successful experience to core markets such as California and New York, and output standardized schemes to build industry benchmarks.

References

Herman, L.E., Sulhaini, S. & Farida, N. (2021). Electronic Consumer Relationship Management and Company Performance: Exploring the Product Innovativeness Development. *J. Relatsh. Mark., 20*, 1–19.

Khong, K.W.; Yao, H. (2011). Effectiveness of customer relationship management on customer satisfaction in the commercial banks of Taiwan. *Contemp. Manag. Res., 7*, 105–116.

Li, W. (2025). Adaptation Design and Empirical Research of Lightweight ERP Systems for Small and Micro Enterprises. *Journal of Progress in Engineering and Physical Science*, *4*(5), 36–42.

Migdadi, M. (2020). Knowledge management, customer relationship management and innovation capabilities. *J. Bus. Ind. Mark., 36*, 111–124.

Morgan, T., Friske, W., Kohtamaki, M. & Mills, P. (2023). Customer participation in manufacturing firms' new service development: The moderating role of CRM technology. *J. Bus. Ind. Mark., 39*, 857–870.

Qi, Z. (2025). Root Cause Tracing Algorithm and One-Click Repair Mechanism for Medical Server Failures. *Journal of Progress in Engineering and Physical Science, 4*(5), 43–48.

Journal of Progress in Engineering and Physical Science
ISSN 2709-4006
www.paradigmpress.org/jpeps
Volume 4 Number 6 December 2025

A Systematic Optimization Analysis of N-Linked Glycans on Porous Graphitic Carbon (PGC) at Different Column Temperatures

Oluwaseun Olanike Ajayi[1] & Victor Oluwatosin Ajayi[2]

[1] Department of Chemistry, University of Georgia, Athens, US
[2] Department of Science Education, Prince Abubakar Audu University Anyigba, Nigeria
Correspondence: Victor Oluwatosin Ajayi, Department of Science Education, Prince Abubakar Audu University Anyigba, Nigeria.

doi:10.63593/JPEPS.2025.12.04

Abstract

Protein glycosylation is a post-translational modification where glycans are covalently attached to proteins. N-linked glycosylation is a type of protein glycosylation which involves the covalent attachment of an oligosaccharide to an asparagine residue within a polypeptide chain. Glycans are involved in a number of biological processes such as cell-cell interactions, cell adhesion, immune recognition, and protein folding and function. However, aberrant glycosylation is mostly associated with many types of diseases such as cancer, autoimmune disorders and many other diseases. Therefore, understanding the variations in glycomics profiles on protein has become very crucial for molecular testing. Despite the advance in LC-MS/MS for structural diverse N-glycan analysis, the separation and analysis of N-linked glycans are analytically challenging due to the molecule large number of possible isomeric conformations, difficult to ionize and contain a variety of complex composition, branching and linkage isomers. Even though, many reversed PGC phases have been used for the separation of N-glycans, especially in the context of LC-MS techniques but is still difficult to implement in a robust and reproducible manner. Herein, systematic optimization analysis of N-linked glycans in porous graphitic carbon (PGC) at different column temperatures was investigated. Underivatized tri-sialylated species and bi-sialylated species N-glycans derived from bovine fetuin and human transferrin respectively were utilized to optimize at different temperature. The study suggests that by optimizing at higher temperature, PGC columns can significantly enhance the resolution of complex glycan mixtures, offering powerful tools for in-depth glycomics optimization analysis.

1. Introduction

Protein glycosylation consists of oligosaccharides attached to asparagine residues. (Bapiro, T.E.; Richards, F.M. & Jodrel. D.I., 2016; Watanabe, Y. et al., 2021; Solá, R.J. & Griebenow, K., 2009) Protein glycosylation, specifically N-linked glycosylation, involves the covalent attachment of oligosaccharides to asparagine residues within a polypeptide chain. (Chandler, K.B. et al., 2019; Jansen, R.S.; Rosing,

H. et al., 2011) Glycosylation is one of the important protein post-translational modifications in eukaryotes. (Agrofoglio, L.A. et al., 2007) Glycans are involved in a broad range of biological processes such as cell-cell interactions, (Watanabe, Y. et al., 2021; Oyama, M. et al., 2018) cell adhesion, (Jansen, R.S. et al., 2011; Watanabe, Y. et al., 2021) immune recognition, (Bapiro, T.E.; Richards, F.M. & Jodrel. D.I., 2016; Shental-Bechor, D. & Levy, Y., 2008) host-pathogen interaction (Bapiro, T.E.; Richards, F.M. & Jodrel. D.I., 2016; Watanabe, Y. et al., 2021) and protein folding and function. (Shental-Bechor, D. & Levy, Y., 2008; Solá, R.J. & Griebenow, K., 2009) Aberrant glycosylation is mostly associated with many types of diseases such as cancer (Singh, C. et al., 2018), autoimmune disorders (Watanabe, Y. et al., 2021; Singh, C. et al., 2018) and traumatic brain injury (TBI) (Oyama, M. et al., 2018). Increase branching or alteration in complex N-linked glycans structures which often lead to the overexpression of truncated or incomplete glycans and the appearance of novel glycan structures is one of the most common glycosylation changes observed in cancer cells (Agrofoglio, L.A. et al., 2007). Therefore, understanding the variations in glycomics profiles on protein has become very crucial for molecular testing.

The separation and analysis of glycoprotein released N-linked glycans using LC-MS/MS techniques is substantially a challenging analytical task due to the molecule large number of possible isomeric conformations, difficult to ionize and contain a variety of complex composition, branching and linkage isomers (Bapiro, T.E.; Richards, F.M. & Jodrel. D.I., 2016; Shental-Bechor, D. & Levy, Y., 2008; Singh, C. et al., 2018; Ozdilek, A. et al., 2020). Despite the advance in LC-MS/MS techniques for structural diverse N-linked glycan analysis especially in sensitivity and specificity (Pandey, V.K. et al., 2022; A. Mehta, A.; Herrera, H. & Block, T., 2015), glycan labeling and derivatization techniques including Rapi-Fluor-MS (RFMS) and permethylation (Huang, Y. et al., 2017). Glycan derivatization can present several challenges basically due to inherent structural complexity with poor ionization efficiency (B.G. Cho, L. Veillon & Y. Mechref, 2019), longer sample preparation time (Reyes, C.D.G. et al., 2022; Zhou, S. et al., 2017), and unwanted side reactions (Palmisano, G. et al., 2013; Zhou, S. et al., 2017). To overcome these limitations, developing efficient chromatographic techniques for N-glycans analysis is necessary.

Hydrophilic interaction liquid chromatography (HILIC) (Melmer, M. et al., 2011), reversed-phase high-performance liquid chromatography (RP-HPLC) (Wuhrer, M.; Boer, A.R. & Deelder, A.M., 2009), mesoporous graphitized carbon (MGC) (Melmer, M. et al., 2011) and porous graphitized carbon (PGC) (Pereira, L., 2008) are some of the frequently used chromatographic techniques coupled with MS for glycans profiling. Despite the superb sensitivity of these procedures, differentiating between similar glycan structures pose a serious challenge because of the structural complexity and diversity of glycans (Wuhrer, M.; Boer, A.R. & Deelder, A.M., 2009; Palmisano, G. et al., 2013). Even though, porous graphitized carbon liquid chromatography (PGC-LC) coupled with MS has been reported to be a superb method for comprehensive analysis of N-linked glycans, (Pereira, L., 2008) but is still difficult to implement in a robust and reproducible manner due to the complexity and heterogeneity of glycans, especially N-linked glycans with multiple antennae, LacNac repeats and sialic acids. Herein, to overcome these limitations, the optimization analysis of N-linked glycans in porous graphitic carbon (PGC) at different column temperature was investigated. Underivatized tri-sialylated species and bi-sialylated species N-glycans derived from bovine fetuin and human transferrin respectively were utilized to optimize at different column temperature.

2. Materials and Method

2.1 Materials

Synthesized N-glycan standards (SRM 3655), 4 N-acetylglucosamine (HexNAc) units, 5 hexose (Hex) units, and 2 N-acetylneuraminic acid (NeuAc) units isomers were obtained from the National Institute of Standards and Technology (NIST) US Department of Commerce. Recombinant PNGase F was used for N-glycan release was procured from Promega Corporation (Madison, WI). alpha-2-HS-glycoprotein and beta-1 globulin protein, human blood serum (HBS), and borazane complex were acquired from R&D Systems (Minneapolis, MN, USA). Ammonium hydroxide was acquired from KMG Chemical Inc. (Texas, USA). HPLC-grade water,

acetonitrile, formic acid, ammonium formate and HPLC-grade methanol were obtained from were acquired from Sigma-Aldrich (St Louis, MO). Porous graphitized carbon (PGC) solid-phase extraction cartridges were used for sample cleanup. LC-MS analysis was performed using a porous graphitized carbon column (100 × 2.1 mm, 2.7 μm particle size, 200 Å pore size; Supel™ Carbon LC, Sigma-Aldrich, coupled to a high-resolution mass spectrometer (Orbitrap Exploris 240).

2.2 Sample Preparation

Synthesized N-glycans isomers of 4 N-acetylglucosamine (HexNAc) units, 5 hexose (Hex) units, and 2 N-acetylneuraminic acid (NeuAc) units were directly reduced with 10 μl of (10 μg/μl) of ammonia borane solution for 2 h at 65 °C. superfluous borane was changed into $[BO3]^{3-}$ by the addition of CH_3OH and eliminated by evaporation in a vacuum concentrator. N -linked glycans were released from 100 μg of standard glycoproteins, 50 μl of cerebrospinal fluid, and hemoglobin S (HbS) by first denaturing the glycoproteins at 90 °C for 30 min in the presence of G2 buffer, following established protocols for enzymatic deglycosylation using PNGase F with minor adaptations from published methods. (Gautam, S. et al., 2021) Proteins were denatured in 50 mM ammonium bicarbonate, optionally reduced and alkylated, and digested with trypsin to generate glycopeptides. The glycopeptides were desalted using C18 solid-phase extraction (SPE), and N-glycans were enzymatically cleaved with PNGase F at 37 °C overnight. Released glycans were separated from peptides using C18 SPE, lyophilized, and derivatized with procainamide via reductive amination (DMSO/acetic acid/water, sodium cyanoborohydride) at 65 °C overnight. Labeled glycans were desalted using PGC SPE, dried, and reconstituted in the initial LC mobile phase prior to analysis.

2.3 PGC Column Conditions

All samples were separated on a Porous graphitized carbon (PGC) column (100 × 2.1 mm, 2.7 μm, 200 Å; Supel™ Carbon LC, Sigma-Aldrich) coupled to a high-resolution mass spectrometer (Orbitrap Exploris 240). For mobile phase A (MPA), water with 10 mM ammonium formate (pH adjusted to ~7.5 with formic acid) was used while mobile phase B (MPB) contained acetonitrile with 10 mM ammonium formate. A multistep mobile phase gradient was applied: 0-5 min: 2% B (equilibration); 5-65 min: linear gradient 2→40% B; 65-75 min: 40→90% B (wash) and 75-90 min: re-equilibration at 2% B. The flow rate was 0.20 mL/min while temperature series were 25, 50, 75, and 100°C and the injection volume was 1-5 μL. Mass spectrometric detection was performed in positive electrospray ionization mode with a mass range of m/z 300-2000. Resolution was set to 60,000 (at m/z 200). Spray voltage was maintained at 3.5 kV with capillary temperature at 275°C. Data acquisition consisted of full scan MS with extracted-ion chromatograms (XICs) generated using ±5 ppm mass tolerance.

2.4 Data Analysis

Data (XICs) were generated for each glycan composition using Skyline and Mass-Hunter software. Peak integration was performed automatically with manual curation where necessary. Relative peak areas were calculated as (individual isomer peak area / total peak area for composition) × 100. Temperature-dependent changes in peak resolution, retention time, and relative abundances were assessed across the tested temperature range.

3. Results and Discussion

3.1 Workflow Description

The chromatographic signals of standard N-linked glycans enzymatically released from bovine fetuin and human Transferrin were initially evaluated and optimized to enhance their chromatography separation on a PGC column. The separated N-linked glycan isomers from alpha-2-HS-glycoprotein, beta-1 globulin protein, cerebrospinal fluid and human blood serum (HBS), samples were then analyzed using the optimal conditions.

3.2 Systematic Optimization

3.2.1 *Optimization of Sialylated N-glycans from Bovine Fetuin Across Different Column Temperatures*

This Figure 1a and 1b shows the behavior of sialylated N-glycans from bovine fetuin across different column temperatures. The column temperature was optimized using 25 °C, 50 °C, 75 °C and 100 °C. At lower temperatures (25-50°C), several glycan species were distinguishable, notably A3G3 (a tri-antennary, tri-galactosylated glycan). As the temperature increased, chromatographic resolution deteriorated: peaks representing highly

sialylated species became broader, overlapped, or were poorly retained. This suggests that the interactions governing PGC retention for negatively charged, sialylated glycans, primarily hydrogen bonding and electrostatic forces, are thermally sensitive (Chandler, K.B. et al., 2019; Agrofoglio, L.A. et al., 2007). Higher temperatures weaken these interactions, resulting in reduced retention strength and increased peak coalescence. Furthermore, the lability of terminal sialic acids under high temperature and solvent exposure may contribute to the diminished peak signals, although desialylation is unlikely under these mild conditions (Chandler, K.B. et al., 2019; Singh, C.; Shyanti, R.K. et al., 2018). These findings demonstrate that while neutral glycans benefit from higher temperatures, complex sialylated species require lower temperatures to maintain resolution on PGC columns (Chandler, K.B. et al., 2019; Agrofoglio, L.A. et al., 2007; Singh, C.; Shyanti, R.K. et al., 2018). Separation of highly sialylated species diminished at higher temperatures; A3G3 remained partially resolvable.

Mass-hunter analysis confirms the same trend observed in Skyline. Lower temperatures promote some degree of separation among sialylated glycan species, with A3G3 consistently observable. At elevated temperatures, the chromatographic performance diminishes significantly, particularly for multiply sialylated structures. These consistent results reinforce the conclusion that the retention of sialylated N-glycans on PGC is compromised by thermal weakening of non-covalent interactions (Chandler, K.B. et al., 2019; Agrofoglio, L.A. et al., 2007). A3G3 remains detectable, possibly due to its intermediate size, abundance, and balance between branching and charge, allowing it to retain PGC affinity across a broader temperature range. Figures 1a and 1b highlight the need for temperature optimization in glycomics workflows depending on the degree of glycan sialylation.

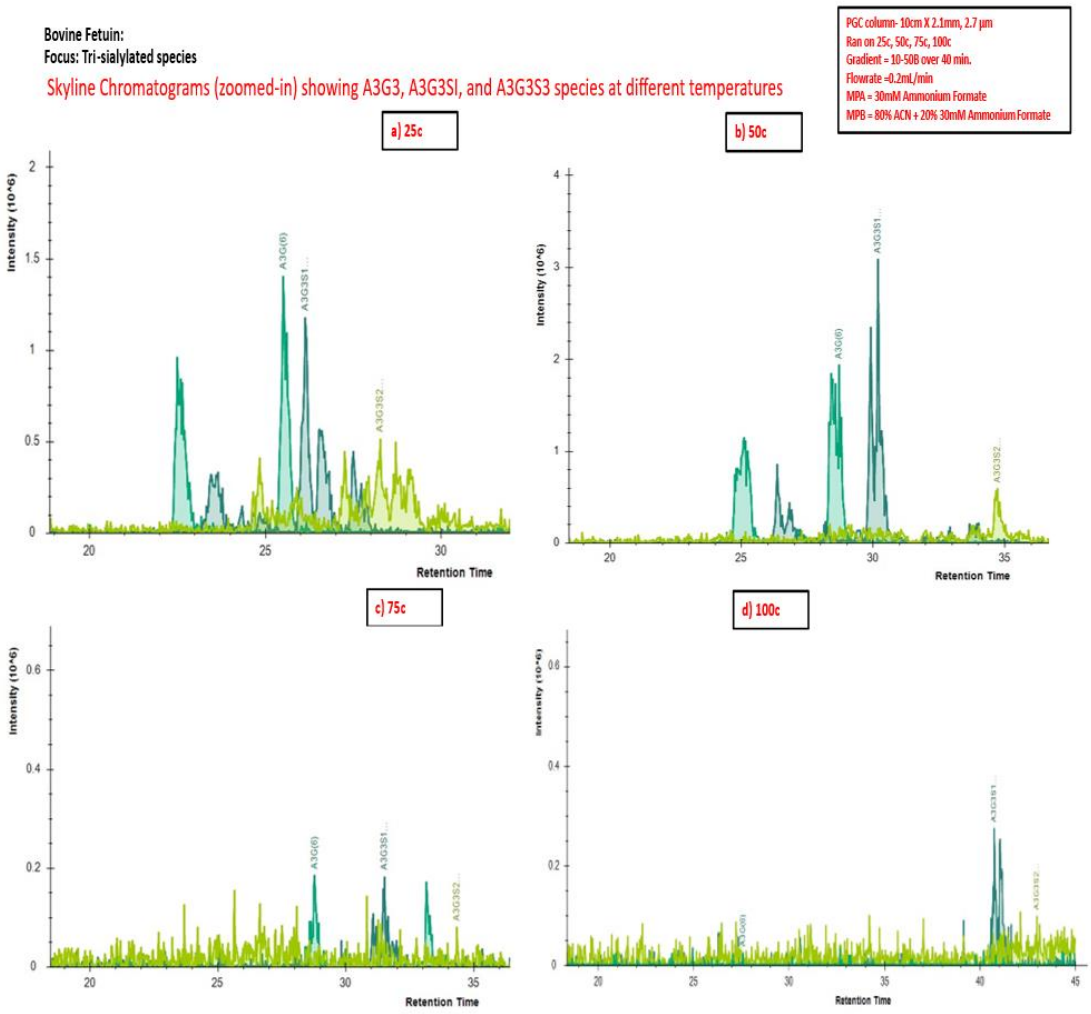

Figure 1a. Skyline chromatograms of bovine fetuin N-glycans at different column temperatures

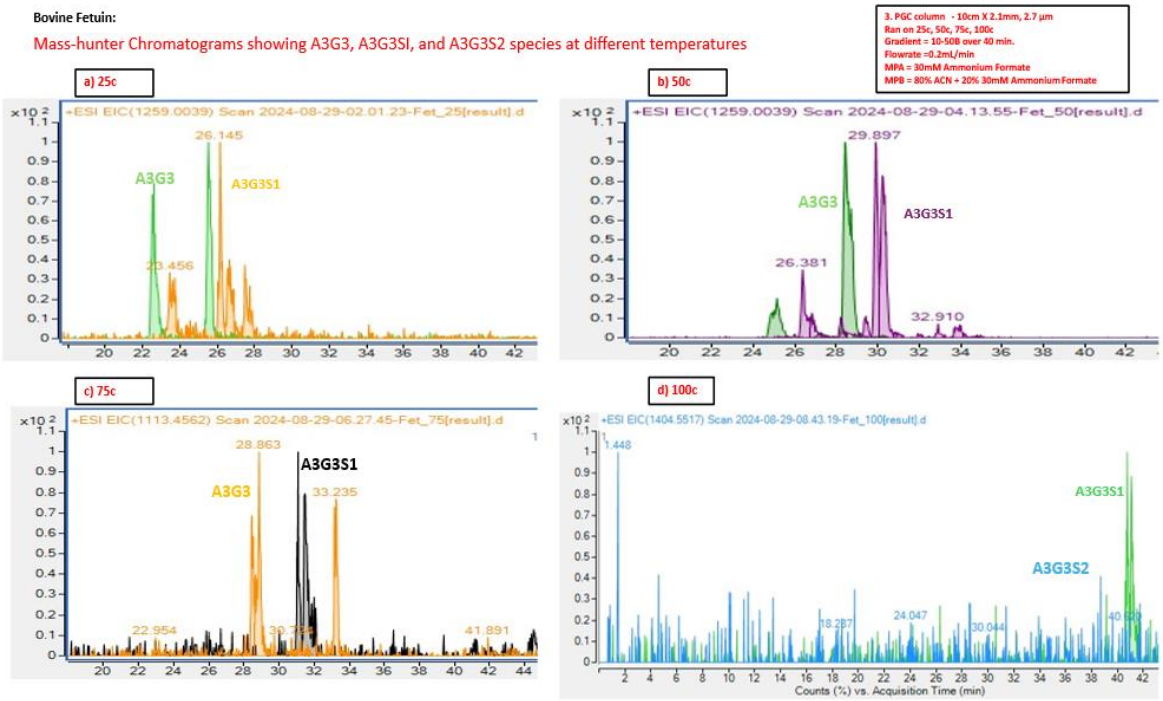

Figure 1b. Mass-hunter chromatograms of bovine fetuin N-glycans at different column temperatures

3.2.2 Optimization of Sialylated N-glycans from Human Transferrin Across Different Column Temperatures

Figures 2a and 2b revealed the N-glycan profile of human transferrin was dominated by bi-sialylated species, which were the most abundant glycans detected and thus served as the main reference for assessing temperature effects on separation. Under the same chromatographic parameters, increasing temperature produced similar effects to those observed with bovine fetuin, peak definition decreased at higher temperatures, and peak splitting occurred in certain bi-sialylated species. These splits likely represent structural isomers or linkage variants that were only partially resolved within the tested temperature range.

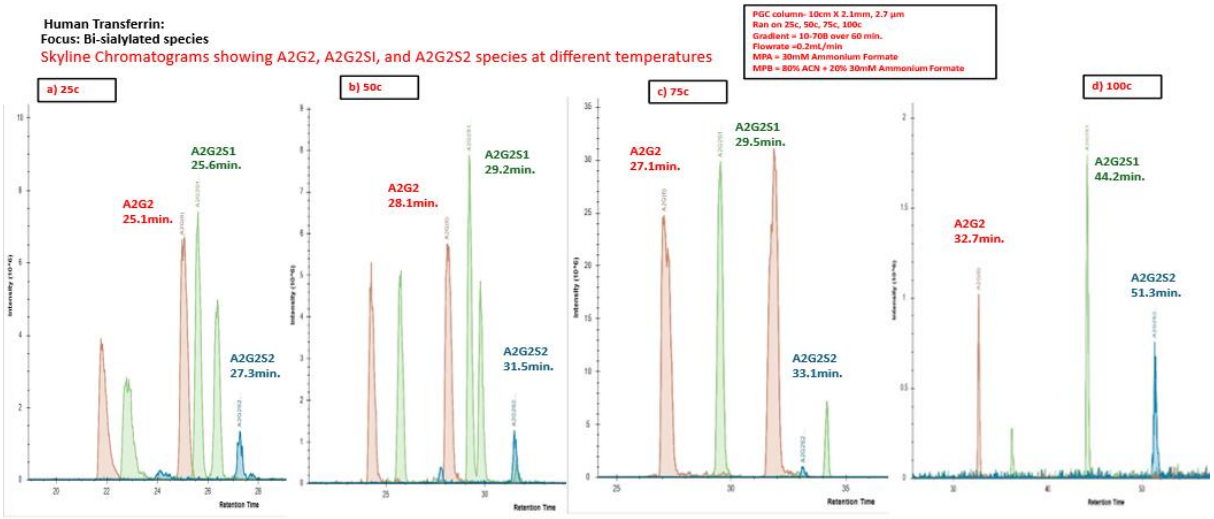

Figure 2a. Skyline chromatograms of Human Transferrin N-glycans at different column temperatures

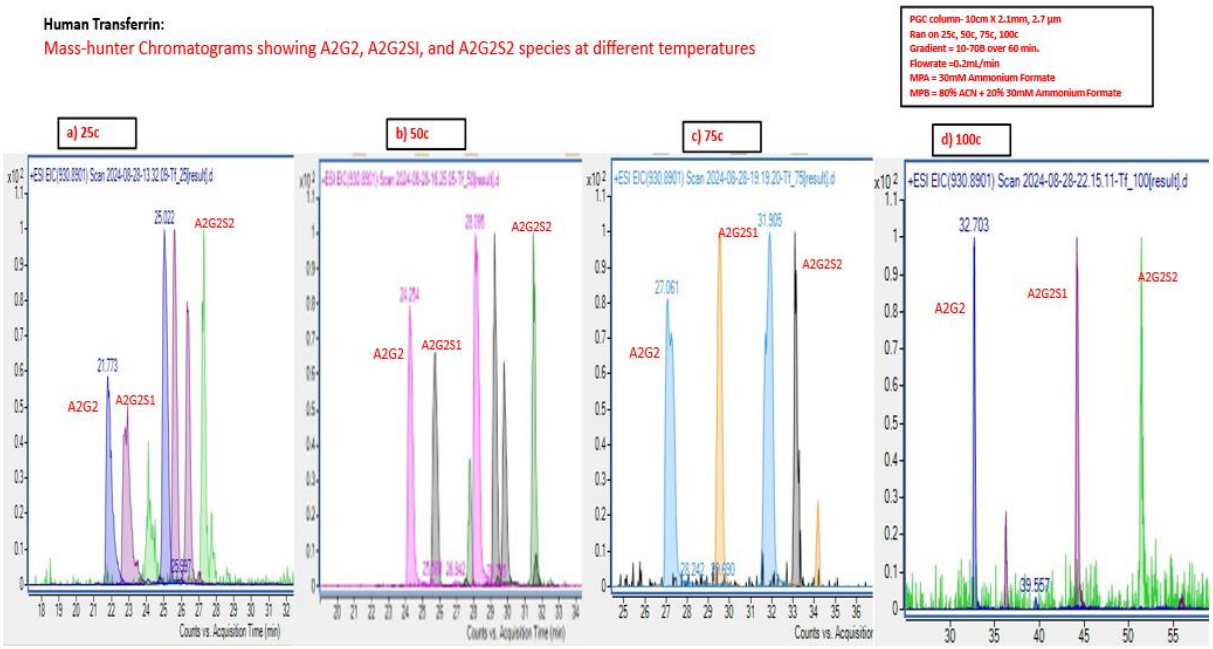

Figure 2b. Mass-hunter chromatograms of Human Transferrin N-glycans at different column temperatures

4. Conclusion

The study was on systematic optimization analysis of N-linked glycans on Porous Graphitic carbon (PGC) at different column temperatures. The improved chromatographic resolution for N-linked glycan was achieved by optimizing at different column temperatures. The N-linked glycans profile of bovine fetuin and human transferrin are dominated by tri-sialylated species and bi-sialylated species respectively are most abundant structures in the samples and therefore the primary focus for temperature-dependent systematic optimization analysis. At 25 °C, these glycans were partially resolved into multiple chromatographic peaks for the same m/z composition, indicating separation of structural isomers. As temperature increased to 50, 75, and 100 °C, peak broadening and coelution became more pronounced for the larger tri-sialylated species, reducing isomeric resolution. This suggests that elevated temperature weakens hydrogen bonding and electrostatic interactions between sialic acids and the PGC surface. Both Skyline and Mass-hunter visualizations confirmed the presence of multiple isomeric peaks and their temperature-dependent resolution. It was concluded that by optimizing at higher temperature, PGC columns can significantly enhance the resolution of complex glycan mixtures, offering powerful tools for in-depth glycomics optimization analysis

References

A. Mehta, A.; Herrera, H.; Block, T. (2015). Glycosylation and liver cancer. *Adv. Cancer. Res.*, *126*, 257-279. DOI: 10.1016/bs.acr.2014.11.005

Agrofoglio, L.A.; Bezy, V.; Chaimbault, P.; Delepee, R.; Rhourri, B.; Morin, P. (2007). Nucleosides, Nucleotides Nucleic Acids. *Acad. Sci.*, *26*, 1523-1527.

B.G. Cho, L. Veillon, Y. Mechref. (2019). N-glycan profile of cerebrospinal fluids from Alzheimer's disease patients using liquid chromatography with mass spectrometry. *J. Proteome Res.*, *18*, 3770–3779. DOI: 10.1021/acs.jproteome.9b00504

Bapiro, T.E.; Richards, F.M.; Jodrel. D.I. (2016). Understanding the complexity of porous graphitic carbon (PGC) chromatography: modulation of mobile-stationary phase interactions overcomes loss of retention and reduces variability. *Anal. Chem.*, *88*, 6190-6194. DOI: 10.1021/acs.analchem.6b01167

Chandler, K.B.; Leon, K.B.; Kuang, J.; Meyer, R.D.; Rahimi, N.; Costello, C.E. (2019). N-Glycosylation regulates ligand-dependent activation and signaling of vascular endothelial growth factor

receptor 2 (VEGFR2). *J. Biol. Chem.*, *294*, 13117-13130. DOI:10.1074/jbc.RA119.008643

Gautam, S.; Banazadeh, A.; Cho, B.G.; Goli, M.; Zhong, J.; Mechref, Y. (2021). Mesoporous graphitized carbon column for efficient isomeric separation of permethylated glycans. *Anal. Chem.*, *93*, 5061-5070. DOI: 10.1021/acs.analchem.0c04395

Green, E.D., Adelt, G.; Baenziger, J.; Wilson, S.; Van-Halbeek, H. (2018). The asparagine-linked oligosaccharides on bovine fetuin. Structural analysis of N-glycanase-released oligosaccharides by 500-megahertz 1H NMR spectroscopy. *J. Biol. Chem.*, *268*, 18253-18268. DOI: 10.1016/S0021-9258(19)81354-6

Huang, Y.; Zhou, S.; Zhu, J.; Lubman, D.M.; Mechref, Y. (2017). LC-MS/MS isomeric profiling of permethylated N-glycans derived from serum haptoglobin of hepatocellular carcinoma (HCC) and cirrhotic patients. *Electrophoresis*, *38*, 2160-2167. DOI:10.10 02/elps.201700025

Jansen, R.S.; Rosing, H.; Schellens, J.H.; Beijnen, J.H. (2011). Mass Spectrometry. *Rev.*, *30*, 321-343.

M. Gaye, S. Valentine, Y. Hu, N. Mirjankar, Z. Hammoud, Y. Mechref, B. Lavine, D. Clemmer. (2012). Ion mobility-mass spectrometry analysis of serum N-linked glycans from esophageal adenocarcinoma phenotypes. *J. Proteome Res.*, *11*, 6102-6110. DOI: 10.1021/pr300756e

Melmer, M.; Stangler, T.; Premstaller, A.; Lindner, W. (2011). Comparison of hydrophilic-interaction, reversed-phase and porous graphitic carbon chromatography for glycan analysis. *J. Chromatogr. A.*, 1218, 118-123. DOI: 10.1016/j.chroma.2010.10.122

Oyama, M.; Kariya, Y.; Kariya, Y.; Matsumoto, K.; Kanno, M.; Yamaguchi, Y.; Hashimoto, Y. (2018). Biological role of site-specific O-glycosylation in cell adhesion activity and phosphorylation of osteopontin. *Biochem. J.*, *475*, 1583-1595. DOI: 10.1042/bcj20170205

Ozdilek, A.; Paschall, A.V.; Dookwah, M.; Tiemeyer, M.; Avci, F.Y. (2020). Host protein glycosylation in nucleic acid vaccines as a potential hurdle in vaccine design for nonviral pathogens. *Proc. Natl. Acad. Sci.*, *117*, 1280-1282. DOI: 10.1073/pnas.1916131117

Palmisano, G.; Larsen, M.R.; Packer, N.H.; Thaysen-Andersen, M. (2013). Structural analysis of glycoprotein sialylation-part II: LC-MS based detection. *RSC Adv.*, *3*, 22706–22726. doi: 10.1039/C3RA42969E

Pandey, V.K.; Sharma, R.; Prajapati, G.K.; Mohanta, T.K.; Mishra, A.K. (2022). N-glycosylation, a leading role in viral infection and immunity development. *Mol. Biol. Rep.*, *49*, 8109-8120. DOI: 10.1007/s11033-022-07359-4

Pereira, L. (2008). Porous graphitic carbon as a stationary phase in HPLC: Theory and applications. *J. of Liquid Chromatograph R. Techn*, *31*(1), 1687-1731. DOI: 10.1080/10826070802126429

Reyes, C.D.G.; Hakim, M.A.; Atashi, A.M.; Goli, S.; Gautam, J.; Wang, A.I.; Bennett, J.; Zhu, D.M.; Lubman, Y. Mechref, Y. (2022). LC-MS/MS isomeric profiling of N-glycans de-rived from low-abundant serum glycoproteins in mild cognitive impairment patients. *Biomolecules*, *12*, 1657-1662. doi: 10.3390/biom12111657

Shental-Bechor, D.; Levy, Y. (2008). Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U. S. A.*, *105*, 8256–826. DOI: 10.1073/pnas.0801340105

Singh, C.; Shyanti, R.K.; Singh, V.; Kale, R.K.; Mishra, J.P.N.; Singh, R.P. (2018). Integrin expression and glycosylation patterns regulate cell-matrix adhesion and alter with breast cancer progression. *Biochem. Biophys. Res. Commun.*, *499*, 374-380. DOI: 10.1016/j.bbrc.2018.03.169

Solá, R.J. & Griebenow, K. (2009). Effects of glycosylation on the stability of protein pharmaceuticals. *J. Pharm. Sci.*, *98*, 1223–1245. DOI:10.1002/jps.21504

Watanabe, Y.; Aoki-Kinoshita, K.F.; Ishihama, Y.; Okuda, S. (2021). Glycopost realizes fair principles for glycomics mass spectrometry data, Nucleic. *Acids Res. 49.*, 1523-1528. DOI: 10.1016/0008-6215(94)84015-610.1093/nar/gkaa1012

Wuhrer, M.; Boer, A.R.; Deelder, A.M. (2009). Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry. *Mass Spectrom. Rev.*, *28*, 192-206. doi:10.1002/mas.20195

Zhou, S.; Huang, Y.; Dong, X; Peng, W.; Veillon, L.; Kitagawa, D.A.; Aquino, A.J.; Mechref, Y. (2017). Isomeric separation of permethylated glycans by porous graphitic carbon (PGC)-LC-MS/MS at high temperatures. *Anal. Chem.*, *89*, 6590-6597. DOI: 10.1021/acs.analchem.7b00747

Journal of Progress in Engineering and Physical Science
ISSN 2709-4006
www.paradigmpress.org/jpeps
Volume 4 Number 6 December 2025

Grid Stability Risks Under High Penetration of Renewable Energy

Qiming Zhao[1], Liang Chen[1], Yonghao Liu[1] & Haoran Wang[1]

[1] Nanjing University of Science and Technology, Nanjing, China
Correspondence: Haoran Wang, Nanjing University of Science and Technology, Nanjing, China.

doi:10.63593/JPEPS.2025.12.05

Abstract

The increasing penetration of renewable energy fundamentally alters the structural and dynamic properties of modern power systems. As inverter-based generation replaces synchronous machines, traditional assumptions regarding inertia, control hierarchy, and network behavior become less reliable. This paper examines grid stability risks under high renewable penetration from a non-empirical, system-oriented perspective. Rather than focusing on specific technologies or case studies, the analysis explores how structural transformation, control interaction, and system coupling reshape the conditions under which stability is maintained. The paper discusses key risk mechanisms related to low-inertia frequency dynamics, voltage behavior in inverter-dominated networks, and the interaction of heterogeneous control strategies. It further examines how coupling among frequency, voltage, network structure, and control logic can amplify local disturbances into system-wide stability challenges. By situating these risks within the broader context of the energy transition, the paper argues that instability should be understood as a structural feature of transitional power systems rather than as an exceptional failure. The analysis highlights the limitations of component-level optimization and emphasizes the need for flexibility, coordination, and adaptive design at the system level. This conceptual approach provides a framework for understanding grid stability as an emergent system property in renewable-dominated power systems.

Keywords: grid stability, renewable energy penetration, low-inertia power systems, inverter-dominated networks, system coupling, energy transition

1. Introduction

The expansion of renewable energy generation has reshaped the structure of modern power systems. Wind and solar power, once treated as supplementary sources, now occupy a central position in electricity supply in many regions. Policy commitments to decarbonization, concerns over long-term energy security, and continuous cost reductions have accelerated this transition. As a result, renewable penetration has reached levels at which it directly influences system operation, planning practices, and reliability considerations. Power systems are no longer organized around a small number of large, synchronous generators operating under relatively stable conditions.

With this shift, questions of grid stability take on a different character. Traditional stability concepts were developed in a context where inertia was abundant, generation was

centralized, and control structures followed clear hierarchical lines. Frequency and voltage behavior could be interpreted through well-established models, and system response to disturbances followed familiar patterns. When a large share of generation is inverter-based, variable, and widely distributed, these assumptions no longer hold in the same way. Stability challenges become less about isolated disturbances and more about how the system behaves as a whole under changing structural conditions.

In renewable-dominated grids, instability does not arise only from fluctuations in wind or solar output. More fundamental changes affect how disturbances propagate, how control actions interact, and how quickly the system can restore balance. Reduced physical inertia alters frequency dynamics. Distributed control systems introduce multiple response paths that may not align in time or direction. Network behavior becomes more complex as power flows vary across locations and operating states. These changes reshape the conditions under which stable operation can be maintained.

Much of the existing literature addresses these challenges through specific technical measures. Advanced inverter control, energy storage deployment, and network reinforcement are often presented as solutions to stability concerns. While such measures play an important role, they tend to focus on particular components or operational scenarios. They offer limited insight into why stability risks emerge at the system level and why similar technical interventions may succeed in one context and fail in another. Practical experience increasingly shows that stability problems in high-renewable systems are not confined to single assets or locations. They develop through interactions among multiple elements and may spread beyond their point of origin.

This paper approaches grid stability from a different angle. It treats stability risks under high renewable penetration as a systemic issue rooted in structural and dynamic changes to the power system. The analysis does not evaluate individual technologies or rely on empirical case studies. Instead, it adopts a non-empirical, conceptual perspective aimed at clarifying how stability is redefined when renewable generation becomes dominant. By examining the mechanisms through which reduced inertia, altered control structures, and increased system complexity shape grid behavior, the paper seeks to explain why traditional stability frameworks face growing limitations. This system-oriented perspective provides a basis for understanding stability risks as an inherent part of the energy transition, rather than as isolated technical problems requiring ad hoc solutions.

2. Grid Stability as a System Property

Grid stability is often discussed through measurable technical indicators, including frequency deviation, voltage range, or recovery time after a disturbance. These indicators are useful for monitoring system performance, yet they describe outcomes rather than underlying conditions. When treated in isolation, they risk reducing stability to a checklist of thresholds, overlooking the deeper system behavior that produces stable or unstable operation. From a broader perspective, grid stability refers to the capacity of a power system to continue operating within acceptable limits when exposed to disturbances, while maintaining coordinated behavior among its major components.

Stability is not produced by a single element. It emerges from the interaction of physical characteristics, control structures, network configuration, and operational rules. Physical inertia affects how quickly frequency changes when supply and demand are unbalanced. Control mechanisms determine how the system reacts once a deviation occurs. Network topology shapes how power and disturbances move across the grid. Operational logic defines how decisions are made, prioritized, and implemented in real time. None of these factors alone can guarantee stability. Stable operation depends on how well they function together.

In traditional power systems dominated by synchronous generators, this interaction was relatively consistent and predictable. Rotating machines supplied physical inertia by design. Control systems followed clear hierarchical structures, with centralized decision-making and well-defined response paths. Power flows were largely unidirectional, and system behavior under disturbance had been studied and codified over decades. Under these conditions, stability analysis could rely on assumptions that inertia was plentiful, control responses followed expected sequences, and disturbances remained confined to limited areas of the network.

As renewable energy penetration increases, these assumptions lose reliability. Inverter-based generation changes the source of system response to imbalance. Inertia is no longer an inherent physical feature of generation but a function of control design and implementation. System reaction to disturbances becomes more dependent on software settings and coordination among devices. Control actions occur across multiple locations and time scales, without a single dominant hierarchy. Power flows vary more frequently and in more directions, reflecting the distributed nature of renewable generation.

This shift complicates the interpretation of stability indicators. Frequency or voltage deviations may reflect not only local issues but also interaction effects between control systems and network structure. A stable reading at one location does not guarantee system-wide robustness. Stability problems may develop gradually through interaction and coupling, rather than appearing as immediate threshold violations. Relying solely on traditional indicators can mask these processes until instability becomes difficult to contain.

Viewing grid stability as a system property helps explain why stability risks in renewable-dominated grids often appear sudden or unexpected. It directs attention away from individual components and toward patterns of interaction that shape system behavior over time. This perspective also clarifies why technical fixes applied in isolation may have limited effect. Addressing stability under high renewable penetration requires understanding how inertia, control, network structure, and operational practice jointly define system response. Only through this system-level view can the changing nature of grid stability be adequately understood.

3. Structural Transformation of Power Systems Under High Renewable Penetration

The move toward a high share of renewable energy reshapes the basic structure of power systems in ways that extend beyond changes in the generation mix. It alters how electricity is produced, how imbalance is managed, and how different parts of the system interact during normal operation and disturbance. Renewable energy is no longer an auxiliary source added to an otherwise stable framework. It becomes a defining element of system behavior. As this shift takes place, the structural conditions that once supported predictable and robust grid operation are reconfigured.

A central aspect of this transformation concerns the source and role of system inertia. In conventional power systems, inertia was an inherent physical feature provided by large synchronous generators. The rotating mass of these machines slowed the rate at which frequency changed following a disturbance. This buffering effect was continuous and did not depend on measurement or control logic. As inverter-based renewable generation replaces synchronous machines, this physical inertia declines. The system's immediate response to imbalance becomes faster and less damped. Frequency behavior depends increasingly on how controls are designed and coordinated, rather than on mechanical properties embedded in generation assets. As a result, frequency dynamics become more sensitive to operating conditions and to differences in control implementation across the system.

Variability in power output further contributes to structural change. Wind and solar generation respond directly to environmental conditions that evolve over time and space. These variations are not occasional disruptions but persistent features of operation. The system must accommodate frequent changes in supply that occur on multiple time scales. Balancing resources are activated more often, and operating points shift continuously. Daily operation no longer follows stable patterns built around predictable generation schedules. Instead, flexibility becomes a defining requirement, influencing how reserves are planned, how networks are utilized, and how control actions are prioritized.

The spatial organization of the power system also changes in important ways. Renewable generators are typically distributed across wide geographic areas and connected at different voltage levels, including distribution networks that were once largely passive. Generation and consumption are no longer clearly separated. Power flows change direction depending on local conditions, and network loading varies more widely across regions. This spatial redistribution weakens the traditional top-down structure of power delivery and increases interdependence among network segments. Disturbances and control actions at one location can affect conditions elsewhere through

complex network paths.

The role of generation units within the system is reshaped as well. Conventional power plants were designed to provide stability support as part of their normal operation. Their physical characteristics and position within control hierarchies made them central actors in frequency and voltage regulation. Renewable generation units rely on control logic to provide similar functions. Their contribution to stability depends on software settings, operating limits, and interaction with other devices. This shift changes how responsibility for stability is distributed across the system. Support functions that were once implicit become conditional and design-dependent.

Taken together, these developments form a new system structure in which stability is less a property of individual components and more a result of interaction. Reduced physical inertia, persistent variability, and distributed generation alter how disturbances are absorbed and how control responses combine. Stability risks emerge from the way these elements interact over time, rather than from isolated weaknesses. Understanding this structural transformation is essential for explaining why stability challenges become more frequent and less predictable under high renewable penetration, and why traditional assumptions about system behavior are increasingly difficult to maintain.

4. Core Stability Risk Mechanisms

This chapter examines the mechanisms through which stability risks take shape when renewable energy reaches a high share of system generation. The focus is not on individual technologies or operational events, but on how structural change and control behavior alter the way the system responds to imbalance. Among these mechanisms, frequency stability under low-inertia conditions represents one of the most fundamental sources of vulnerability.

4.1 Frequency Stability Under Low-Inertia Conditions

Frequency stability reflects the ability of a power system to contain and recover from imbalances between generation and demand. In conventional systems, this ability relied heavily on the physical inertia of synchronous generators. Rotating masses stored kinetic energy, which slowed frequency change immediately after a disturbance. This delay provided a natural buffer, allowing control actions to take effect before frequency deviations reached critical levels.

As renewable penetration increases, this buffering function weakens. Inverter-based generation does not contribute inertia in the same physical sense. The system's initial frequency response becomes faster and more abrupt, even when the size of a disturbance remains unchanged. A given loss of generation or load produces a steeper rate of change of frequency, reducing the time available for corrective action. Frequency deviation therefore becomes more difficult to arrest once it begins.

The reduction of inertia also changes the role of control systems in frequency regulation. In traditional grids, primary frequency response supplemented inertia rather than replacing it. Under low-inertia conditions, control systems carry a larger share of responsibility for stabilizing frequency. Their effectiveness depends on sensing accuracy, response speed, and coordination across multiple devices. Delays, mismatches, or limited response capacity can directly translate into larger frequency excursions.

Low inertia also affects frequency recovery. In conventional systems, frequency restoration followed a relatively smooth trajectory shaped by predictable generator behavior. In renewable-dominated systems, recovery depends on the collective response of diverse control schemes. Some inverter-based resources may respond aggressively, while others may remain passive or constrained by design limits. The resulting response can be uneven, leading to overshoot, oscillation, or prolonged deviation from nominal frequency.

Another source of risk lies in the spatial distribution of inertia and control capability. As synchronous generation is displaced unevenly across regions, inertia becomes concentrated in fewer locations. Local disturbances can therefore have system-wide frequency effects, even when network constraints would previously have limited their impact. Frequency behavior becomes less tied to physical proximity and more dependent on system-wide interaction.

These changes reveal why frequency stability under high renewable penetration cannot be understood through disturbance size alone. The same event can produce very different outcomes depending on inertia distribution, control coordination, and operating conditions.

Frequency instability emerges not as an isolated failure, but as a result of structural dependence on fast, tightly coupled control responses. This makes frequency stability one of the most sensitive and revealing indicators of system vulnerability in low-inertia power systems.

4.2 Voltage Stability in Inverter-Dominated Networks

Voltage stability concerns the ability of a power system to maintain acceptable voltage levels under normal operation and following disturbances. In conventional power systems, voltage regulation was closely tied to the physical behavior of synchronous generators and passive network elements. Reactive power support followed well-understood relationships between voltage, current, and generator excitation, and voltage control was largely localized, with predictable interaction across the network.

In inverter-dominated networks, the logic of voltage regulation changes in fundamental ways. Inverter-based resources do not respond to voltage variations through inherent electromagnetic behavior. Instead, voltage support is provided through control algorithms that determine how and when reactive power is injected or absorbed. This shift replaces continuous physical response with discrete, rule-based action. Voltage behavior therefore depends less on system physics and more on control design, parameter settings, and coordination among devices.

Local voltage stability becomes more sensitive under these conditions. In distribution-level and sub-transmission networks with high concentrations of renewable generation, small changes in power output or network conditions can lead to significant voltage variation. Inverters may reach reactive power limits quickly, especially when operating near maximum active power output. Once these limits are reached, voltage support capacity drops sharply, increasing the risk of local voltage collapse.

The system-wide implications of local voltage behavior are also altered. Inverter controls often operate independently, responding to local measurements without full awareness of broader network conditions. When many devices follow similar control rules, their combined response can amplify voltage deviations rather than suppress them. Local corrective actions may interact in unexpected ways, creating patterns of voltage fluctuation that spread across network segments.

Voltage stability risk in inverter-dominated networks is further shaped by the coupling between active and reactive power control. In some operating states, efforts to regulate voltage through reactive power adjustment can constrain active power output, affecting overall power balance. This coupling introduces trade-offs that were less pronounced in conventional systems, where generators could draw on mechanical reserves and excitation systems with greater flexibility.

These characteristics give voltage instability a system-level dimension that extends beyond isolated weak points. Local voltage problems may signal broader coordination issues within the network, reflecting limits in control interaction rather than in component capability alone. Under high renewable penetration, voltage stability becomes less a question of individual device performance and more a question of how control strategies collectively shape network behavior.

4.3 Control Interaction and Oscillation Risks

As power systems move toward higher levels of renewable penetration, control becomes a central element in maintaining stable operation. Inverter-based resources rely on control algorithms to regulate frequency, voltage, and power output. At the same time, conventional generators, energy storage systems, and network devices continue to operate under their own control schemes. The coexistence of these heterogeneous strategies creates a control environment that is more layered and less predictable than in traditional grids.

In conventional systems, control interactions were largely structured within clear hierarchies. Primary, secondary, and tertiary controls operated on distinct time scales, and interactions among them were well understood. Under inverter-dominated conditions, control actions occur across overlapping time frames and are distributed across many devices. Fast-acting inverter controls respond within milliseconds, while other controls operate more slowly. When these responses are not aligned, their interaction can produce unintended dynamic effects.

Oscillation risks arise when control loops influence one another in reinforcing ways. A response intended to correct a local deviation

may trigger a reaction elsewhere in the system, which then feeds back into the original control loop. When multiple devices follow similar control logic, synchronized responses can amplify disturbances rather than dampen them. This effect becomes more pronounced as the number of inverter-based resources increases.

Resonance risk is also heightened by the growing reliance on power electronics. Inverter controls introduce new dynamic modes that were absent in systems dominated by synchronous machines. These modes can interact with network impedance characteristics and with other control systems. Under certain operating conditions, small disturbances may excite these modes, leading to sustained oscillations that are difficult to predict using conventional stability models.

The lack of a unified control perspective further contributes to instability risk. Devices are often designed and configured independently, optimized for local performance rather than for system-wide interaction. Without coordination, control parameters that are stable in isolation may interact poorly when deployed at scale. Oscillation and resonance thus emerge not from faulty design of individual components, but from the combined behavior of multiple control strategies operating simultaneously.

These interaction effects help explain why oscillatory instability becomes more common in renewable-dominated grids. The risk does not lie in any single control scheme, but in the structure of control coexistence itself. As control complexity increases, stability depends less on the robustness of individual controllers and more on how their actions combine across the system.

5. System Coupling and Risk Amplification

In power systems with high renewable penetration, stability problems rarely remain limited to a single variable or a specific location. Frequency behavior, voltage behavior, and control response are intertwined through both physical connections and operational logic. What begins as a localized disturbance can propagate through these links, changing form and scale as it moves across the system. In this setting, risk amplification is less a consequence of disturbance magnitude and more a result of how tightly system elements are coupled.

A key source of coupling lies in the relationship between active power, reactive power, and control action. Frequency reflects the balance between active power supply and demand, while voltage depends on reactive power support and network conditions. In inverter-dominated systems, these functions are often managed through integrated control schemes. A control action intended to correct frequency deviation may change reactive power output, altering local voltage conditions. In response, voltage control mechanisms may adjust operating points or limit active power injection. Through this process, a correction in one domain can introduce stress in another, creating a chain of dependent responses.

The physical structure of the network further amplifies these interactions. Distributed renewable generation increases the number of injection points and weakens the traditional separation between transmission and distribution functions. Bidirectional power flows mean that changes at one node can influence conditions elsewhere in less predictable ways. Voltage fluctuations, impedance characteristics, and power flow redistribution can trigger additional responses in nearby devices. As these responses accumulate, the disturbance may spread beyond its point of origin and affect larger portions of the system.

Control logic plays an equally important role in shaping how disturbances evolve. Many inverter-based resources rely on local measurements and predefined thresholds to determine their response. When similar control logic is deployed widely, devices may react in parallel to the same signal. This synchronized behavior can intensify system stress rather than alleviate it. Instead of damping a deviation, collective action may deepen it or shift it into another stability domain. In such cases, amplification emerges from coordination effects rather than from the failure of individual components.

Timing and response strength also influence amplification. In tightly coupled systems, small delays in sensing or response can allow deviations to grow before corrective action takes effect. Conversely, rapid and aggressive responses may overshoot, pushing the system toward oscillatory behavior. Because multiple control loops operate simultaneously, mismatches in response speed or magnitude can interact in ways that are difficult to anticipate. These dynamics reduce tolerance for error and increase sensitivity to operating conditions.

Coupling also affects how stability margins are perceived. Traditional assessment methods often assume that frequency, voltage, and control behavior can be evaluated separately. In renewable-dominated systems, these separations become less meaningful. A condition that appears stable when examined through one indicator may conceal growing stress in another domain. Instability may develop gradually through accumulation rather than appearing as a sudden threshold violation.

From a system perspective, risk amplification reflects the combined effect of interaction, timing, and structure. Local disturbances gain significance as they move through interconnected frequency, voltage, and control processes. Understanding stability under high renewable penetration therefore requires attention to how these processes reinforce one another across the network. Stability can no longer be assessed by examining isolated components or single indicators. It must be understood as an emergent outcome shaped by the structure of coupling and the logic of coordinated response.

6. Implications for Grid Operation and Design Thinking

The stability risks observed in renewable-dominated power systems indicate that long-standing approaches to grid operation and design are under growing pressure. Many operational routines and planning principles were developed in environments where generation patterns were stable and system behavior could be anticipated within relatively narrow limits. Under high renewable penetration, uncertainty becomes a persistent feature of daily operation rather than an occasional disturbance. This shift challenges the assumption that stability can be ensured through predefined operating rules and fixed response strategies.

One key implication concerns the limits of component-level optimization. Grid operation has traditionally emphasized improving the performance of individual assets, whether by enhancing generator efficiency, accelerating control response, or tightening operational margins. In systems characterized by strong coupling and interaction, these improvements do not automatically enhance overall stability. A control scheme that appears robust when evaluated in isolation may interact poorly with other schemes once deployed across the system. Local optimization can therefore introduce unintended stress at the system level. Operational decision-making must move beyond the assumption that aggregate stability emerges naturally from well-performing components.

Operational flexibility emerges as a core requirement rather than a supplementary capability. High renewable penetration introduces continuous variation in supply conditions and operating states. Fixed schedules and rigid response hierarchies offer limited protection against instability in this environment. Operators require the ability to reconfigure system operation as conditions change, adjusting control priorities, reserve allocation, and response intensity in real time. Flexibility must be embedded in the structure of the system, not confined to specific assets or emergency measures.

Coordination across control layers becomes equally important. As control responsibility shifts from a small number of centralized generators to a large number of distributed devices, maintaining coherent system behavior depends on how these controls interact. Differences in response speed, control objectives, and parameter settings can undermine stability even when each controller operates as designed. Without explicit coordination principles, distributed control actions may conflict or reinforce one another in undesirable ways. Effective coordination requires a shared understanding of system goals and response priorities, rather than reliance on implicit hierarchy or after-the-fact adjustment.

Design thinking faces parallel challenges. Traditional grid design emphasized robustness under a limited set of expected operating scenarios, often focusing on worst-case events defined in advance. In renewable-dominated systems, operating conditions vary more widely, and rare events are harder to distinguish from normal variability. Design choices must therefore account for interaction and uncertainty from the outset. Network structure, control architecture, and operational rules should be evaluated together, with attention to how they shape system behavior over time.

These implications point toward a broader shift in how stability is approached. Rather than attempting to eliminate uncertainty, grid

operation and design must be capable of functioning within it. Stability depends on the system's ability to adapt to changing conditions while preserving coherent behavior. This capacity cannot be achieved through single technologies or isolated improvements. It requires an integrated approach that treats adaptability, coordination, and system awareness as fundamental principles guiding both operation and design in renewable-dominated power systems.

7. Discussion

When placed within the broader process of energy transition, grid stability risks should be understood as part of a long-term structural shift rather than as isolated technical failures or temporary side effects. The transition toward low-carbon energy systems is not limited to changing fuel sources. It involves a reorganization of how power systems generate, transmit, and regulate electricity. As synchronous machines are gradually replaced by renewable and inverter-based resources, the underlying operating logic of the grid is redefined. Stability challenges arise from this redefinition, not from individual errors or incomplete deployment of technology.

Within this context, instability represents a structural condition rather than an abnormal outcome. Power systems were originally designed around predictable generation, strong physical coupling, and centralized coordination. High renewable penetration alters these conditions. Physical inertia is reduced, control becomes more distributed, and system response depends increasingly on programmed behavior. These changes reshape the balance between predictability and responsiveness, making stability more sensitive to interaction and coordination. Risk emerges as the system adapts to functions it was not originally structured to perform.

Viewing stability problems as isolated incidents can obscure the mechanisms that produce them. Corrective actions applied at specific locations or times may resolve immediate operational issues, yet leave deeper interaction patterns unchanged. As renewable penetration continues to grow, similar problems may surface elsewhere in the system or reappear under different operating conditions. This repetition suggests that stability risks are rooted in system structure rather than in particular technologies or operational choices.

The energy transition also introduces a prolonged period of coexistence between legacy and emerging technologies. During this phase, conventional generators, inverter-based resources, and various control philosophies operate side by side. Differences in response characteristics, time scales, and design assumptions create conditions in which coordination becomes more difficult. Stability risks reveal these points of misalignment, highlighting where existing frameworks no longer match evolving system behavior. Such risks are likely to persist as long as system architecture and operational practice remain in transition.

Temporal complexity further shapes how stability risks unfold. Changes to power systems occur incrementally, while operating requirements remain continuous. Systems must maintain reliability while undergoing ongoing modification, which limits the possibility of clean separation between old and new operating regimes. Stability issues may develop gradually as interaction patterns shift, making them harder to detect through short-term performance metrics alone. Understanding these dynamics requires attention to long-term trends rather than isolated events.

From this perspective, effective responses to stability challenges depend on system-level understanding. Incremental fixes and localized upgrades address specific symptoms but do not resolve the conditions that generate instability. Greater emphasis is needed on how system architecture, control logic, and operational practice evolve together. Stability becomes a question of coherence across changing system elements, rather than of compliance with fixed technical thresholds. Recognizing stability risks as an inherent aspect of the energy transition helps move the focus from reactive intervention toward anticipatory design and long-term operational alignment.

8. Conclusion

This paper has explored grid stability risks under conditions of high renewable energy penetration from a system-oriented and non-empirical perspective. Rather than focusing on specific technologies, case studies, or operational events, the analysis has traced how stability risks arise from deeper structural changes in power system organization. As

renewable generation moves from the periphery to the core of electricity supply, long-standing assumptions about inertia, control hierarchy, and network behavior no longer hold in the same way.

The discussion has shown that stability challenges cannot be attributed to a single cause. Frequency instability under low-inertia conditions reflects the loss of passive physical buffering that once shaped system response. Voltage instability in inverter-dominated networks reveals how control-dependent behavior replaces inherent electrical characteristics. Control interaction and oscillation risks emerge from the coexistence of multiple control strategies operating across different time scales. These mechanisms are interconnected, and their effects are amplified through system coupling rather than confined to local disturbances.

A key insight of this study is that stability risks should be understood as structural outcomes of the energy transition. As power systems evolve, old and new operating logics coexist, often without full alignment. Instability arises not because renewable technologies are flawed, but because system architecture, control design, and operational practice are adjusting at different speeds. During this period of transition, uncertainty and interaction become defining features of system behavior.

The analysis also highlights the limits of traditional responses to stability problems. Technical upgrades and component-level optimization remain important, but they do not address the full range of system-level interactions that shape stability outcomes. Improvements applied in isolation may even introduce new forms of vulnerability when deployed across a highly coupled system. Stability therefore depends less on eliminating individual weaknesses and more on managing interaction, coordination, and adaptability across the grid.

From an operational and design perspective, this implies a need to rethink how stability is pursued. Flexibility must be treated as a system-wide attribute rather than as a property of selected assets. Coordination among control schemes must be addressed explicitly rather than assumed. Design choices should account for evolving operating conditions and interaction effects, rather than optimizing performance under fixed assumptions. Stability in renewable-dominated grids depends on the capacity of the system to respond coherently to change.

By framing grid stability as a system property shaped by long-term structural transformation, this paper offers a way to interpret stability risks as an integral part of the energy transition. This perspective shifts attention from reactive problem-solving toward anticipatory understanding of system behavior. Future research may build on this approach by developing analytical frameworks that better capture interaction and uncertainty, and by examining how emerging grid architectures can support stable operation as renewable penetration continues to rise.

References

Anderson, P. M., & Fouad, A. A. (2003). *Power system control and stability* (2nd ed.). Piscataway, NJ: IEEE Press.

Bevrani, H., François, B., & Ise, T. (2017). *Microgrid dynamics and control*. Hoboken, NJ: Wiley.

Bialek, J. (2019). What does the GB power outage on 9 August 2019 tell us about the future of the electricity system? *Energy Policy*, *146*, 111821.

Debnath, S., et al. (2021). Control interactions and oscillations in inverter-dominated power systems. *IEEE Transactions on Power Systems*, *36*(6), 5473–5485.

Dörfler, F., Simpson-Porco, J. W., & Bullo, F. (2016). Electrical networks and algebraic graph theory: Models, properties, and applications. *Proceedings of the IEEE*, *106*(5), 977–1005.

ENTSO-E. (2018). *High penetration of power electronic interfaced power sources and the potential contribution of grid forming converters*. Brussels: ENTSO-E.

Hatziargyriou, N. (Ed.). (2014). *Microgrids: Architectures and control*. Chichester, UK: Wiley.

IEEE Power & Energy Society. (2020). *IEEE PES task force on inverter-based resources*. IEEE.

Kundur, P. (1994). *Power system stability and control*. New York, NY: McGraw-Hill.

Machowski, J., Bialek, J. W., & Bumby, J. R. (2008). *Power system dynamics: Stability and control* (2nd ed.). Chichester, UK: Wiley.

Markovic, U., et al. (2019). Understanding small-signal stability of low-inertia systems. *IEEE Transactions on Power Systems*, *34*(5), 3793–3805.

Milano, F., Dörfler, F., Hug, G., Hill, D. J., & Verbič, G. (2018). Foundations and challenges of low-inertia systems. *Proceedings of the IEEE*, *107*(11), 2143–2166.

Morales, J. M., Conejo, A. J., Madsen, H., Pinson, P., & Zugno, M. (2014). *Integrating renewables in electricity markets: Operational problems*. New York, NY: Springer.

Poolla, B. K., Groß, D., & Dörfler, F. (2017). Placement and implementation of grid-forming and grid-following virtual inertia and fast frequency response. *IEEE Transactions on Power Systems*, *32*(4), 2930–2940.

Rohden, M., Sorge, A., Timme, M., & Witthaut, D. (2012). Self-organized synchronization in decentralized power grids. *Physical Review Letters*, *109*(6), 064101.

Schäfer, B., Matthiae, M., Timme, M., & Witthaut, D. (2015). Decentral smart grid control: Phase shifts and stability. *New Journal of Physics*, *17*(1), 015002.

Tamrakar, U., Shrestha, D., Maharjan, M., Bhattarai, B. P., Hansen, T. M., & Tonkoski, R. (2017). Virtual inertia: Current trends and future directions. *Applied Sciences*, *7*(7), 654.

Ulbig, A., Borsche, T. S., & Andersson, G. (2014). Impact of low rotational inertia on power system stability and operation. *IFAC Proceedings Volumes*, *47*(3), 7290–7297.

van Hertem, D., Gomis-Bellmunt, O., & Liang, J. (Eds.). (2016). *HVDC grids: For offshore and supergrid of the future*. Hoboken, NJ: Wiley.

Zhou, Y., Mancarella, P., Mutale, J., & Strbac, G. (2016). Modeling and assessment of the contribution of demand response to power system flexibility. *IEEE Transactions on Power Systems*, *31*(4), 2711–2720.

Journal of Progress in Engineering and Physical Science
ISSN 2709-4006
www.paradigmpress.org/jpeps
Volume 4 Number 6 December 2025

Applicability Failures of Urban Air Quality Models in Coastal Cities and the Role of Sea–Land Breeze and Boundary Layer Dynamics

Marco Rinaldi[1]

[1] Independent Researcher, Italy
Correspondence: Marco Rinaldi, Independent Researcher, Italy.

doi:10.63593/JPEPS.2025.12.06

Abstract

Urban air quality models are widely applied in environmental assessment and policy making, yet their applicability in coastal cities remains limited. This paper examines why commonly used urban air quality models often fail to represent pollution dynamics in coastal environments. Focusing on conceptual rather than technical issues, the analysis highlights mismatches between model assumptions and key coastal atmospheric processes, including land–sea thermal contrast, sea–land breeze circulation, boundary layer stratification, and strong temporal variability. These processes challenge assumptions of spatial homogeneity, predictable mixing, and monotonic dispersion that underpin many urban models. The paper argues that temporal averaging further obscures critical transition periods that govern pollutant accumulation and exposure. By framing these limitations as applicability failures rather than model errors, the study emphasizes the need for context-sensitive interpretation and development of urban air quality models in coastal urban settings.

Keywords: urban air quality modeling, coastal cities, atmospheric boundary layer dynamics, temporal averaging

1. Introduction

Urban air quality modeling has become a foundational tool in contemporary environmental governance, shaping how cities understand pollution patterns, design mitigation strategies, and evaluate public health risks. These models translate complex atmospheric processes into formalized representations that link emission sources with ambient concentration outcomes. Through numerical simulation, they provide a structured way to interpret how pollutants move, accumulate, react, and dissipate within urban space. Their influence extends beyond scientific analysis into regulatory standards, urban planning decisions, and long term sustainability strategies.

The intellectual foundations of most urban air quality models were established in contexts where atmospheric behavior could be reasonably approximated as horizontally continuous and vertically well mixed over large spatial extents. In such settings, synoptic scale meteorological forcing is treated as the dominant driver of wind patterns, thermal structure, and mixing processes. Local circulations are often framed as secondary

effects that introduce limited perturbations to an otherwise coherent atmospheric system. This conceptual framing has shaped model architecture, parameterization schemes, and evaluation practices for decades.

Coastal cities challenge this foundation at a structural level. Their atmospheric environment is not a marginal variation of inland conditions but a qualitatively different system shaped by persistent interaction between land and sea surfaces. The presence of a coastline introduces abrupt contrasts in thermal inertia, surface roughness, moisture availability, and energy exchange. These contrasts generate mesoscale circulations that are not occasional disturbances but recurring organizing forces of the local atmosphere. As a result, the physical logic underpinning many urban air quality models becomes misaligned with the environment in which they are applied.

In coastal urban areas, atmospheric processes unfold across overlapping spatial and temporal scales that resist simplification. Large scale weather systems interact continuously with locally generated circulations, producing flow fields that shift direction, depth, and intensity within short periods. Vertical structure becomes especially important, as air masses of different origin coexist within shallow layers and exchange is often restricted. Pollutant transport is therefore shaped not only by emission strength and background wind but also by the timing of thermal transitions and the vertical positioning of flow regimes.

When models developed under inland assumptions are applied to such environments, discrepancies emerge that are often interpreted as performance deficiencies or data limitations. In reality, many of these discrepancies reflect deeper issues of applicability. The models may reproduce average conditions while failing to capture accumulation, recirculation, or delayed exposure patterns that are central to coastal pollution dynamics. This produces outputs that appear internally consistent but lack correspondence with the actual pathways through which pollutants persist in coastal cities.

The concept of applicability failure is therefore more appropriate than simple model error in this context. Applicability failure points to a situation in which a model operates as designed but within a setting that violates its underlying conceptual premises. The issue is not the absence of sophisticated algorithms or high resolution inputs but the persistence of assumptions about atmospheric continuity, monotonic dispersion, and scale separation that do not hold in coastal environments.

Sea–land breeze circulation and boundary layer dynamics occupy a central position in this mismatch. These processes are not peripheral features but structural elements of coastal atmospheric organization. They govern the timing, direction, and confinement of pollutant transport while shaping vertical mixing conditions that determine near surface exposure. Their influence is inherently time dependent and often nonlinear, making them difficult to reconcile with modeling frameworks optimized for stability and averaging.

This paper situates the underperformance of urban air quality models in coastal cities within this broader conceptual landscape. Rather than treating coastal complexity as an obstacle to be minimized, it frames it as a lens through which the limits of prevailing modeling paradigms become visible. By examining how standard assumptions interact with coastal atmospheric characteristics, the discussion aims to clarify why certain modeling failures recur and why incremental adjustments are often insufficient.

Through a focus on circulation patterns, boundary layer structure, and temporal representation, the analysis seeks to contribute to a more reflective approach to model application. Understanding the conditions under which urban air quality models lose explanatory power is essential not only for improving technical performance but also for ensuring that model based decisions remain grounded in atmospheric reality.

2. Common Assumptions in Urban Air Quality Models

Urban air quality models are built upon a set of assumptions that function as cognitive shortcuts, allowing complex atmospheric processes to be translated into manageable mathematical representations. These assumptions are rarely neutral. They reflect the historical contexts in which models were developed, the types of cities that originally motivated their use, and the dominant meteorological paradigms that guided early atmospheric science. Over time, these assumptions have become embedded in model architecture, shaping how urban air pollution is

conceptualized, simulated, and interpreted.

2.1 Assumptions of Spatial Homogeneity and Meteorological Continuity

One of the most fundamental assumptions in urban air quality modeling is the relative spatial homogeneity of meteorological forcing within the urban domain. Wind speed, wind direction, temperature fields, and atmospheric stability are commonly assumed to vary gradually across space. This assumption allows models to interpolate meteorological inputs over grid cells and to treat the urban area as a continuous surface rather than a patchwork of competing microclimates.

This spatial continuity assumption is closely tied to the dominance attributed to synoptic scale meteorology. Large scale weather systems are presumed to set the background conditions under which urban dispersion occurs. Local influences such as surface heterogeneity, land use contrast, and small scale circulations are typically framed as secondary modifiers rather than primary drivers. In practice, this means that model performance is evaluated against its ability to respond to changes in regional wind patterns or frontal passages, while internally generated circulations receive limited structural emphasis.

The assumption of homogeneity also simplifies the treatment of urban form. Buildings, roads, vegetation, and open spaces are often aggregated into average roughness lengths and surface parameters. This abstraction reduces computational complexity but simultaneously obscures localized flow disturbances that can shape pollutant pathways. While such simplifications may produce acceptable results in cities where atmospheric forcing is relatively uniform, they become increasingly fragile in environments where sharp gradients dominate airflow behavior.

2.2 Simplified Representation of Urban Boundary Layer Structure

Another core assumption concerns the structure and evolution of the urban atmospheric boundary layer. Most urban air quality models rely on an idealized diurnal cycle in which daytime heating produces a deep, well mixed convective boundary layer, followed by nocturnal cooling that generates a shallow stable layer near the surface. This conceptual model assumes that vertical mixing follows predictable patterns governed primarily by surface heat flux and solar radiation.

Within this framework, turbulence is parameterized using schemes that presume statistical uniformity and steady state conditions over each time step. Vertical exchange is treated as a function of surface energy balance and mechanical shear, while horizontal variability is often minimized. The result is a boundary layer that behaves as a coherent mixing volume rather than a stratified or layered system.

This assumption allows pollutants emitted at the surface to be rapidly distributed through the lower atmosphere during the day and weakly mixed at night. Concentration patterns therefore respond smoothly to changes in emissions and meteorology. Such behavior aligns well with inland settings where surface characteristics and thermal forcing are relatively consistent across the urban area.

The difficulty arises when this assumed boundary layer structure does not materialize. In environments where external air masses intrude, where thermal contrasts persist, or where vertical stability is maintained during daytime, the logic of a predictable mixing cycle breaks down. Models that rely on generic boundary layer evolution then misrepresent both the depth of pollutant confinement and the timing of dispersion.

2.3 Treatment of Local Circulations as Secondary Phenomena

Urban air quality models typically conceptualize local circulations as deviations from a dominant background flow. These circulations are often parameterized implicitly rather than explicitly resolved. Their effects are absorbed into turbulence coefficients or averaged wind fields rather than represented as distinct dynamic systems.

This approach reflects a hierarchical view of atmospheric motion, where synoptic scale processes govern regional transport and mesoscale or microscale circulations play a corrective role. The model logic assumes that local circulations do not fundamentally alter transport pathways but only modulate dispersion efficiency.

Such an assumption influences how models handle complex flow regimes. When wind direction shifts rapidly or reverses, models may interpret this as noise rather than as an organizing mechanism. Pollutant accumulation

caused by circulation loops is therefore underrepresented, and dispersion is implicitly treated as directional and progressive rather than cyclical.

In many inland contexts, this hierarchy is defensible. Local circulations are often weak, transient, or spatially limited. In coastal cities, however, this assumption becomes increasingly problematic because local circulations are persistent, structured, and tightly coupled with surface forcing. Treating them as secondary leads to systematic underestimation of their role in shaping pollution outcomes.

2.4 Assumptions Embedded in Temporal Averaging

Temporal representation constitutes another major assumption in urban air quality modeling. Most models operate on fixed time steps that reflect practical constraints rather than atmospheric process scales. Hourly or daily averages are commonly used for emissions, meteorological inputs, and chemical reactions. This practice rests on the belief that short term variability does not substantially affect longer term concentration patterns.

Temporal averaging simplifies computation and aligns model outputs with regulatory metrics, which are often defined over similar time windows. It also reflects a conceptual preference for stable trends over transient events. Peaks, reversals, and rapid transitions are treated as fluctuations that do not alter the overall pollution narrative.

This assumption carries significant implications. It implies that exposure is best understood as an accumulated quantity rather than as a sequence of discrete events. It also assumes linearity in atmospheric processes, where averaging inputs yields averaged outcomes.

In environments characterized by rapid transitions, this logic breaks down. Short lived circulation shifts can trap pollutants, relocate plumes, or expose populations to elevated concentrations within narrow time windows. When such processes are smoothed through averaging, the model output may appear accurate in a statistical sense while failing to reflect lived exposure conditions.

2.5 Linearization of Transport and Dispersion Processes

Urban air quality models often rely on linear or quasi linear representations of transport and dispersion. Wind fields are treated as steady within each time step, and pollutant movement is simulated as a response to these fields. Dispersion is conceptualized as dilution driven by turbulence and mixing.

This approach assumes that transport pathways are predictable and that pollutant fate is largely determined by mean flow characteristics. Feedback mechanisms between pollutants and atmospheric structure are typically minimized or ignored. Pollutants are passive tracers whose behavior does not alter the system in which they move.

While this assumption is acceptable for many primary pollutants over short timescales, it becomes limiting in environments where transport pathways loop, overlap, or intersect vertically. Linear dispersion logic struggles to represent scenarios in which pollutants return to their source region or remain confined within narrow layers for extended periods.

The linearization of transport also affects how models interpret emission timing. Emissions released at different times of day are often assumed to disperse independently. In reality, atmospheric memory effects can cause emissions from previous periods to influence current concentration patterns through delayed transport or recirculation.

2.6 Assumptions About Chemical Processing and Mixing

Chemical transformation processes in urban air quality models are often linked to assumptions about mixing and residence time. Reaction rates are calculated based on average concentrations and averaged meteorological conditions. This implies that chemical processing occurs within a well mixed volume where reactants are uniformly distributed.

Such an assumption simplifies chemical modeling but obscures the role of stratification and segregation. In layered atmospheres, pollutants may be confined within specific thermal or moisture regimes that alter reaction pathways. Limited mixing can prolong the lifetime of certain species or enhance localized chemical production.

By assuming efficient mixing, models may underestimate the persistence of reactive pollutants or misrepresent the spatial distribution of secondary products. This limitation is particularly pronounced in environments where vertical exchange is

restricted or intermittent.

These assumptions form a coherent modeling philosophy rooted in stability, continuity, and scale separation. They allow urban air quality models to function effectively in settings that align with their conceptual design. The issue arises when these models are transferred to environments that violate the conditions under which the assumptions were formulated.

Coastal cities expose the context dependence of these assumptions. Spatial homogeneity is disrupted by land sea contrasts. Boundary layer structure departs from the idealized diurnal cycle. Local circulations operate as primary drivers rather than secondary modifiers. Temporal variability becomes a defining feature rather than background noise.

The resulting discrepancies are not random errors but systematic outcomes of applying a model outside its conceptual domain. Recognizing this helps shift the discussion from technical calibration toward deeper questions about model scope, validity, and interpretation.

3. Atmospheric Characteristics of Coastal Cities

Coastal cities are situated within atmospheric environments that differ fundamentally from those of inland urban areas. Their defining feature is not simply proximity to the sea, but continuous exposure to interacting surface systems that operate according to distinct physical principles. The atmosphere above a coastal city is shaped by the juxtaposition of land and water surfaces, each imposing its own thermal, mechanical, and moisture related constraints on air movement. This interaction generates atmospheric behavior that is intrinsically dynamic, vertically structured, and temporally sensitive.

3.1 Land–Sea Thermal Contrast as a Structuring Force

At the core of coastal atmospheric behavior lies the contrast in thermal properties between land and sea surfaces. Land responds rapidly to incoming solar radiation, heating quickly during the day and cooling efficiently at night. Water bodies, by contrast, possess high heat capacity and strong thermal inertia, resulting in slower temperature changes over diurnal cycles. This persistent imbalance produces horizontal temperature gradients that are spatially fixed by geography and temporally modulated by solar forcing.

These gradients are not incidental. They function as organizing forces that shape pressure distributions in the lower atmosphere. Air moves in response to these gradients, generating circulations that are predictable in form yet variable in strength and extent. The atmosphere above a coastal city is therefore rarely in equilibrium. Instead, it is continuously adjusting to surface driven thermal contrasts that do not disappear even under stable synoptic conditions.

This thermal asymmetry also affects atmospheric stability. Warm air rising over land interacts with cooler marine air masses, producing stratified layers that resist vertical mixing. The result is an atmosphere that often departs from the idealized vertically uniform structure assumed in many urban frameworks. Pollutants emitted at the surface may encounter thermal barriers that limit upward transport, altering dispersion patterns from those expected under inland conditions.

3.2 Variability and Directionality of Wind Fields

Wind behavior in coastal cities is characterized by pronounced variability in both direction and magnitude. Unlike inland settings where prevailing winds often maintain a consistent orientation over extended periods, coastal wind fields are strongly influenced by diurnal thermal forcing and shoreline geometry. Wind direction can shift by large angles within short time intervals as surface heating patterns evolve.

This variability is not random. It follows a structured rhythm linked to the heating and cooling cycle of land and sea. During certain periods, onshore flow dominates near the surface while offshore flow may persist aloft. At other times, these patterns reverse or weaken, allowing synoptic winds to exert greater influence. The coexistence of competing wind regimes creates a layered flow structure in which air masses move in different directions at different heights.

Such complexity challenges the notion of a single representative wind field for the urban domain. Near surface winds may transport pollutants inland, while upper level flows carry air masses seaward. The net effect is often partial displacement rather than full removal of pollutants. This layered directionality also increases the likelihood of plume distortion and recirculation, as emissions are stretched and folded within the coastal flow system.

3.3 Vertical Stratification and Layered Atmospheric Structure

Vertical structure is a defining feature of coastal atmospheres. The interaction between marine air and continental air frequently produces stratified layers that differ in temperature, moisture content, and turbulence intensity. These layers may persist for hours or days depending on synoptic conditions and surface forcing.

Marine air intrusions often form shallow layers near the surface that are cooler and more stable than overlying air. This configuration suppresses vertical mixing and traps pollutants within a confined volume. Above this layer, warmer and drier air may flow in a different direction, effectively decoupled from surface processes. Such stratification creates a multi layered boundary region that cannot be adequately described by a single mixing height.

The presence of layered structures alters the residence time of pollutants. Emissions may remain confined within a narrow layer, experiencing limited dilution despite moderate wind speeds. Chemical reactions within this layer may proceed differently than expected under well mixed conditions, as reactant concentrations and radiation exposure vary with height.

This vertical complexity also affects the interpretation of surface measurements. Concentrations recorded at ground level reflect not only emission strength but also the degree of isolation between layers. Without accounting for stratification, it is difficult to infer transport efficiency or dispersion capacity from surface data alone.

3.4 Boundary Layer Depth and Stability Regimes

The atmospheric boundary layer over coastal cities exhibits marked variability in depth and stability. Unlike inland cities where daytime heating often produces a deep convective layer, coastal environments may experience suppressed boundary layer growth due to the influence of cooler marine air. This effect can persist even under strong solar radiation.

Shallow boundary layers reduce the volume available for pollutant dilution. Emissions released at the surface are confined to a smaller mixing space, leading to higher concentrations than would occur under deeper boundary layer conditions. This confinement is particularly pronounced during morning transitions when land heating has begun but marine influence remains strong.

Stability regimes in coastal boundary layers are also more diverse. Stable conditions may occur during periods traditionally associated with convective mixing, while weakly unstable conditions may persist into the evening. This irregularity disrupts the assumed correspondence between time of day and mixing behavior that underlies many urban dispersion models.

The frequent occurrence of temperature inversions near the coast further complicates dispersion. These inversions may form through advection rather than radiative cooling, making them resistant to daytime erosion. Pollutants trapped beneath such inversions can accumulate over multiple cycles, increasing the potential for prolonged exposure events.

3.5 Moisture, Cloud Formation, and Radiative Effects

Coastal atmospheres are typically more humid than inland counterparts due to proximity to large water bodies. Moisture content influences several atmospheric processes relevant to air quality. Higher humidity affects thermal stratification by modifying heat exchange and cloud formation. It also influences radiative balance by altering the absorption and scattering of solar radiation.

Cloud development in coastal regions is often linked to marine air intrusion and boundary layer processes. Low level clouds can reduce surface heating, weakening convective mixing and prolonging stable conditions. This feedback reinforces pollutant confinement near the surface.

Moisture also plays a role in chemical transformation. Reaction rates for certain pollutants are sensitive to humidity and liquid water content. The presence of fog or low clouds can facilitate heterogeneous reactions that are not adequately represented in models calibrated for drier inland environments.

These moisture related processes introduce additional layers of complexity to coastal atmospheric behavior, affecting both physical transport and chemical evolution of pollutants.

3.6 Interaction with Urban Morphology and Topography

The atmospheric characteristics of coastal cities

are further shaped by the interaction between coastal processes and urban form. Buildings modify airflow by increasing surface roughness and generating turbulence, while coastal topography such as cliffs, bays, and peninsulas channels wind and alters circulation patterns.

Urban heat island effects interact with coastal thermal gradients in non linear ways. In some cases, urban warming enhances onshore flow by intensifying land sea temperature contrasts. In other situations, marine cooling offsets urban heating, producing complex spatial patterns of stability and mixing.

Topographic features can anchor circulation patterns, causing convergence zones or stagnation regions that persist under specific wind regimes. Pollutants may accumulate in these zones despite favorable regional conditions for dispersion.

The combined influence of urban morphology and coastal geography reinforces atmospheric heterogeneity. Air quality outcomes become highly sensitive to location within the city, challenging assumptions of spatial uniformity.

These characteristics define coastal cities as heterogeneous atmospheric systems rather than modified versions of inland urban environments. Their behavior is governed by interacting surface contrasts, layered flows, variable stability regimes, and persistent local circulations. These processes operate across multiple scales and resist simplification through averaging or uniform parameterization. Local dynamics in coastal cities are not marginal effects that can be treated as noise. They are central to how air moves, mixes, and retains pollutants. Any attempt to model air quality in such settings must confront this complexity at a conceptual level.

4. Sea–Land Breeze Circulation and Applicability Failures of Urban Air Quality Models

Sea–land breeze circulation represents one of the most persistent and structurally influential atmospheric processes in coastal cities. Unlike episodic weather events, this circulation arises from the daily rhythm of differential heating between land and sea and is therefore embedded in the normal functioning of the coastal atmosphere. Its regularity gives it predictive form, yet its interaction with urban surfaces, background winds, and boundary layer structure produces outcomes that are highly sensitive to timing, location, and vertical position.

Urban air quality models often treat wind as a background transport mechanism that responds smoothly to surface forcing and large scale pressure gradients. Within this framework, changes in wind direction associated with sea–land breezes are commonly represented as simple reversals or shifts occurring over prescribed time intervals. This abstraction overlooks the fact that sea–land breeze circulation is not a uniform horizontal flow but a three dimensional system composed of opposing near surface and return flows, convergence zones, and evolving vertical coupling.

The applicability failure begins at the conceptual level. Many models assume that once wind direction changes, the previous transport regime has ended. In coastal environments, transport regimes overlap. Air masses influenced by the previous phase of the circulation often remain present aloft or within residual layers, interacting with newly established flows. Pollutants are therefore not transported away and replaced by cleaner air in a linear sequence. They are redistributed within a closed or semi closed circulation system.

The diurnal development of sea breezes illustrates this problem clearly. During daytime, cooler marine air advances inland at low levels, displacing warmer continental air upward and landward. This process creates a frontal structure characterized by sharp gradients in temperature, humidity, and turbulence. Urban air quality models that rely on averaged wind fields tend to smooth this structure into a gradual transition, eliminating the localized convergence that can trap pollutants near the surface.

Within the sea breeze front, vertical motion plays a critical role. Air is forced upward as marine air undercuts warmer land air. Pollutants emitted near the surface may be lifted into elevated layers where they are transported differently from surface emissions. Models that do not explicitly resolve this vertical motion often misattribute observed concentration patterns to horizontal advection or emission variability. The failure lies not in numerical precision but in the absence of the circulation logic that governs pollutant redistribution.

At night, the reversal to land breeze conditions

introduces a second layer of complexity. Cooling of the land surface generates offshore flow near the ground, while residual marine air and return flows may persist above. Pollutants accumulated during the day may be advected seaward at low levels, yet remain within reach of the urban atmosphere due to weak vertical mixing and limited horizontal displacement. These pollutants can form offshore reservoirs that are reintroduced during the next sea breeze cycle.

This recirculation mechanism directly contradicts the implicit assumption of unidirectional dispersion embedded in many urban air quality models. Dispersion is often conceptualized as a process through which pollutants are progressively diluted and removed from the urban domain. Sea–land breeze circulation replaces this logic with one of retention and return. Pollutants may exit the immediate urban area only to reenter within hours, often in altered vertical positions that change exposure patterns.

Models that do not account for this circulation loop may predict declining concentrations following emission reductions or favorable winds, while actual conditions show persistence or delayed peaks. This discrepancy is frequently interpreted as a problem of emission inventory accuracy or meteorological input quality. In reality, it reflects a deeper incompatibility between model assumptions and coastal transport dynamics.

Another dimension of applicability failure arises from the interaction between sea–land breezes and urban heat island effects. Urban surfaces tend to remain warmer than surrounding rural areas due to heat storage and reduced evapotranspiration. In coastal cities, this thermal anomaly interacts with the land–sea temperature contrast in ways that alter the strength, timing, and inland penetration of sea breezes.

Urban heat can intensify onshore flow by amplifying the temperature gradient between land and sea. It can also delay the onset of nocturnal land breezes by maintaining higher land temperatures after sunset. These effects introduce asymmetry into the diurnal cycle that standard models rarely capture. Many modeling frameworks assume fixed transition times based on solar forcing alone, neglecting the modifying role of urban thermal inertia.

As a result, models may misplace the spatial boundary between marine and continental air masses. Pollutants may be simulated as dispersed inland when they are actually confined near the coast, or simulated as flushed offshore when urban heating sustains onshore flow. This misrepresentation affects not only average concentration levels but also the spatial distribution of exposure across the city.

Sea–land breeze circulation also introduces sensitivity to coastal geometry. The shape of the shoreline, the presence of bays or peninsulas, and the orientation of the coast relative to prevailing synoptic winds all influence circulation patterns. Sea breezes may converge in certain areas and diverge in others, creating localized stagnation zones or accelerated flow corridors.

Urban air quality models that employ regular grids and generalized coastlines often fail to capture these geometric effects. The circulation is effectively flattened into a uniform onshore or offshore wind field. Pollutant hotspots driven by convergence or channeling are therefore underestimated or misplaced. This failure becomes particularly evident when model outputs are compared against localized observations that reflect fine scale circulation features.

Vertical coupling within the sea–land breeze system presents another challenge. Return flows aloft transport air masses back toward the sea during daytime and toward land at night. Pollutants entrained into these flows can be stored above the surface layer, shielded from deposition and chemical loss. When vertical mixing increases, these stored pollutants may descend and influence surface concentrations long after their emission.

Standard urban air quality models often lack the resolution or parameterization needed to represent this storage mechanism. Vertical exchange is simplified into diffusion processes that assume continuous mixing rather than discrete layering. The result is an underestimation of pollutant persistence and an overestimation of dispersion efficiency.

Temporal representation further amplifies applicability failures. Sea–land breeze transitions do not occur instantaneously at fixed times. They develop gradually, with spatially varying onset and decay. When models impose abrupt changes in wind direction based on hourly inputs, they erase the transitional periods

during which pollutant trapping and redistribution are most pronounced.

These transitional periods are often when the highest concentrations occur, particularly in near coastal urban zones. Averaged temporal inputs smooth these peaks, producing outputs that appear stable but fail to represent exposure dynamics accurately. The mismatch is not simply a matter of resolution but of conceptual alignment between model structure and atmospheric process.

The cumulative effect of these issues is a systematic bias in how urban air quality models interpret coastal pollution dynamics. Sea–land breeze circulation transforms the urban atmosphere from an open system into a semi closed one, where removal is conditional and temporary. Models designed around the assumption of progressive dilution struggle to accommodate this logic.

Applicability failure in this context should not be understood as an inability to simulate wind direction changes. It reflects a deeper limitation in representing circulation as an organizing framework for pollutant behavior. When circulation is treated as background variability rather than as a structuring mechanism, model outputs lose explanatory power in coastal settings.

Recognizing this limitation reframes the challenge of coastal air quality modeling. The issue is not the absence of computational sophistication but the persistence of inland oriented conceptual assumptions. Sea–land breeze circulation exposes the boundaries of these assumptions by revealing transport patterns that are cyclical, layered, and sensitive to timing.

5. Boundary Layer Dynamics and Implications for Pollutant Dispersion

The atmospheric boundary layer plays a decisive role in shaping air quality outcomes because it defines the vertical space within which pollutants are diluted, transported, transformed, and retained. In urban air quality modeling, the boundary layer is often treated as a controllable parameter whose depth and stability can be inferred from surface heat flux, wind speed, and time of day. This approach rests on the assumption that boundary layer behavior follows a relatively stable and repeatable diurnal cycle. Coastal cities challenge this assumption at a fundamental level. Boundary layer dynamics in coastal environments are shaped by competing surface influences, layered air mass interactions, and advective processes that operate independently of local surface heating. As a result, pollutant dispersion in coastal cities is governed by a set of boundary layer behaviors that diverge significantly from those anticipated by standard urban modeling logic.

One of the most consequential features of coastal boundary layers is their tendency toward suppressed vertical growth. In inland cities, daytime solar heating typically produces a convective boundary layer that deepens steadily as surface temperatures rise. This growth increases the effective mixing volume available for pollutants, allowing surface emissions to be diluted over hundreds or even thousands of meters. In coastal cities, the presence of cooler marine air limits this process. Onshore flow introduces air masses that are thermally stable relative to the heated land surface. This stability resists vertical motion and constrains turbulent mixing even under strong insolation.

The result is a boundary layer that remains shallow during periods when inland models would predict vigorous mixing. Pollutants emitted at the surface are confined within a reduced vertical space, leading to elevated concentrations near the ground. This confinement persists despite the presence of wind, because horizontal advection does not compensate for limited vertical dilution. Urban air quality models that assume a direct relationship between daytime heating and boundary layer depth therefore systematically underestimate near surface pollutant accumulation in coastal cities.

Boundary layer suppression in coastal environments is not uniform in space or time. Its intensity depends on the strength of marine air intrusion, the thermal contrast between land and sea, and the timing of synoptic influences. In some cases, shallow stable layers persist only in near coastal zones, while inland areas experience deeper mixing. In other cases, marine influence extends far inland, flattening the vertical structure of the boundary layer across the urban region. This spatial variability challenges the common modeling practice of applying a single mixing height across an entire urban domain.

Thermal inversions constitute another defining

element of coastal boundary layer dynamics. Inversions in inland environments often form at night due to radiative cooling of the surface. These inversions are typically shallow and eroded after sunrise as surface heating resumes. In coastal cities, inversions frequently arise through advection rather than radiation. Cool marine air moving beneath warmer continental air creates a temperature profile that increases with height, producing a stable layer that can persist through much of the day.

Advective inversions differ from radiative inversions in both strength and persistence. They are often deeper, less sensitive to surface heating, and spatially extensive. Pollutants trapped beneath such inversions may accumulate over extended periods, particularly when emissions continue and horizontal ventilation is limited. Urban air quality models that rely on simplified stability classes or nocturnal inversion logic often fail to represent these conditions accurately. The duration and vertical extent of advective inversions are underestimated, leading to optimistic predictions of dispersion.

The presence of inversions alters not only pollutant concentration levels but also their vertical distribution. Emissions released at the surface remain confined to a narrow layer, while pollutants injected above the inversion through stacks or buoyant plumes may be transported differently. This vertical separation produces concentration gradients that cannot be captured by models assuming uniform mixing within the boundary layer. Exposure risk becomes height dependent, affecting populations differently based on building elevation and urban form.

Beyond simple inversion structures, coastal boundary layers frequently exhibit internal layering. Distinct air masses with different thermal histories and moisture content coexist within the lower atmosphere. A cool, moist marine layer may occupy the lowest tens or hundreds of meters, overlain by warmer, drier continental air. Turbulent exchange between these layers is often weak, especially when wind shear is modest. Pollutants emitted within one layer may remain isolated from others for long periods.

This layered structure has profound implications for pollutant fate. Chemical reactions proceed under conditions specific to each layer, influenced by temperature, humidity, and radiation. Pollutants confined within the marine layer may experience reduced photochemical activity due to cloud cover or limited sunlight penetration. At the same time, their removal through vertical dilution is restricted. Models that average chemical processes across the entire boundary layer implicitly assume efficient mixing, masking these layer specific dynamics.

Vertical layering also introduces memory effects into the coastal atmosphere. Pollutants stored within residual layers above the surface may descend when boundary layer structure changes, contributing to delayed concentration peaks that are not directly linked to current emissions. Standard modeling approaches often treat each time step independently, assuming that previous boundary layer states have limited influence on current conditions. This assumption breaks down in layered coastal systems where atmospheric history plays a significant role.

The interaction between boundary layer dynamics and wind shear further complicates dispersion. In coastal environments, wind speed and direction often vary sharply with height. Near surface winds may be weak and directed onshore, while stronger offshore or alongshore flows exist aloft. This shear can suppress vertical mixing by stabilizing the atmosphere, while simultaneously transporting pollutants horizontally within specific layers.

Urban air quality models frequently simplify wind profiles, applying logarithmic or power law relationships derived from inland observations. These representations struggle to capture the complex shear structures present in coastal boundary layers. Pollutants may be advected rapidly within elevated layers while surface concentrations remain high due to limited mixing. Such decoupling produces discrepancies between modeled transport pathways and observed concentration patterns.

Boundary layer transitions represent another critical source of applicability failure. In coastal cities, transitions between stable and unstable regimes are often gradual, spatially heterogeneous, and influenced by advection. The collapse of the boundary layer in the evening may occur earlier near the coast than inland. The growth of the daytime boundary layer may be delayed or incomplete. These transitions are periods of heightened sensitivity,

during which small changes in forcing produce large changes in mixing and concentration.

Urban air quality models commonly impose transitions based on time of day or surface flux thresholds. This approach assumes a predictable sequence of boundary layer states. In coastal environments, the timing and nature of transitions vary across the urban domain. Models that do not resolve this variability smooth over critical windows when pollutant trapping or release occurs.

The implications for pollutant dispersion are substantial. Dispersion efficiency becomes decoupled from simple indicators such as wind speed or solar radiation. High wind speeds do not guarantee dilution if vertical mixing remains suppressed. Strong sunlight does not ensure boundary layer growth if marine air maintains stability. The intuitive relationships that guide model interpretation in inland cities lose validity in coastal contexts.

These dynamics undermine a core principle embedded in many urban air quality models, namely that increased atmospheric energy leads to enhanced dispersion. In coastal cities, energy input may be redistributed horizontally or absorbed by the marine layer rather than converted into vertical mixing. Pollutants remain concentrated near the surface despite conditions that would promote dilution inland.

The cumulative effect of these boundary layer behaviors is a shift from an open dispersion system to a constrained retention system. Pollutants are not simply transported away from their source. They are stored, redistributed, and reintroduced within a shallow and stratified atmosphere. This shift challenges the fundamental dispersion logic upon which many models are built.

Applicability failure arises because models interpret boundary layer depth and stability as controllable parameters rather than emergent properties of interacting surface systems. When boundary layer dynamics are driven by advection, layering, and persistent stability, parameter adjustments alone cannot correct the mismatch. The model continues to operate within an inland oriented conceptual framework. Understanding boundary layer dynamics in coastal cities therefore requires a different interpretive stance. Dispersion must be understood as conditional and episodic rather than continuous. Mixing must be viewed as spatially uneven and temporally delayed. Pollutant fate must be linked to atmospheric structure as much as to emission strength.

These insights have implications beyond technical modeling. They affect how air quality risks are assessed, how mitigation strategies are evaluated, and how regulatory thresholds are interpreted. In coastal cities, improvements in emissions may not translate immediately into improved air quality due to boundary layer constraints. Models that fail to represent this reality risk overstating the effectiveness of interventions.

Boundary layer dynamics thus stand at the center of applicability failures in coastal urban air quality modeling. They reveal the limits of assumptions about mixing, dilution, and temporal response that underpin standard modeling approaches. Recognizing these limits is essential for interpreting model outputs responsibly and for developing modeling frameworks that align more closely with the atmospheric realities of coastal cities.

6. Temporal Scale and Averaging Effects in Coastal Air Quality Modeling

Temporal scale is not a neutral modeling choice. It is a theoretical decision that shapes how atmospheric processes are perceived, filtered, and ultimately represented. In urban air quality modeling, time is commonly treated as a container within which emissions, meteorology, and chemistry unfold in a continuous and smooth manner. This treatment reflects a preference for stability, regularity, and comparability across modeling applications. In coastal cities, this preference collides with an atmospheric reality characterized by rapid transitions, intermittent processes, and strong time dependence. The resulting mismatch constitutes one of the most persistent sources of applicability failure in coastal air quality modeling.

Most urban air quality models operate on hourly or longer temporal resolutions. Emission inventories are often compiled as hourly averages derived from activity patterns. Meteorological inputs are typically provided at fixed time intervals that smooth short term variability. Chemical mechanisms are solved using time steps that assume relatively stable boundary conditions across each interval. These choices are justified by computational efficiency and by alignment with regulatory metrics,

which frequently rely on hourly, daily, or annual averages.

This temporal structure embeds a specific view of atmospheric behavior. It assumes that processes operating at shorter timescales do not fundamentally alter longer term concentration patterns. Variability is treated as noise that averages out, while persistent trends are treated as signal. In inland environments with relatively stable wind regimes and predictable boundary layer evolution, this assumption often holds to a reasonable degree. In coastal cities, it does not.

Coastal atmospheric processes are intrinsically time sensitive. Sea–land breeze circulation develops gradually, intensifies, weakens, and reverses within the span of a single day. Boundary layer depth can change rapidly in response to marine air intrusion or cloud formation. Wind direction may shift by large angles over short periods without any change in synoptic forcing. These changes are not anomalies. They are structural features of the coastal atmosphere.

When such processes are represented using hourly or daily averages, their internal dynamics are suppressed. The model captures a simplified narrative in which wind changes occur discretely at fixed times and boundary layer properties evolve smoothly. The transitional phases that dominate pollutant trapping and release are effectively erased. This erasure produces outputs that appear coherent and stable but lack correspondence with actual dispersion behavior.

Temporal averaging introduces what can be described as structural bias rather than random error. The bias arises because the averaging process systematically removes the very fluctuations that drive pollutant accumulation in coastal settings. Concentration peaks associated with stagnation periods are flattened. Short term exposure events are diluted into background levels. Recirculation effects are spread across multiple time steps and interpreted as weak transport rather than retention.

This bias is particularly evident during transition periods between sea breeze and land breeze phases. These periods are often characterized by weak winds, shifting flow directions, and collapsing boundary layers. Pollutants emitted during these windows experience minimal dispersion and can accumulate rapidly near the surface. When meteorological inputs are averaged across the hour, the calm conditions are merged with adjacent periods of stronger flow. The resulting wind field suggests moderate dispersion potential, masking the stagnation that actually occurred.

The same logic applies to boundary layer transitions. In coastal cities, boundary layer collapse in the evening can begin earlier near the coast than inland and may occur in stages rather than as a uniform event. Morning boundary layer growth may be delayed or fragmented by marine influence. Models that impose fixed diurnal cycles smooth these transitions into idealized patterns. Pollutant trapping during early evening or late morning periods is underestimated because the model assumes either full mixing or stable conditions across the entire time step.

Temporal averaging also alters the perceived relationship between emissions and concentrations. In models, emissions released during an hour are often assumed to disperse according to averaged meteorological conditions for that hour. In reality, emissions released during the first minutes of an hour may experience entirely different dispersion regimes than those released later. In coastal environments, the difference between these regimes can be decisive. Averaging erases this distinction and produces an artificial alignment between emission timing and dispersion capacity.

This misalignment undermines the interpretive value of model outputs. Modeled concentrations appear to respond smoothly to emission changes, reinforcing the idea that emission control leads directly to air quality improvement. In coastal cities, the relationship is mediated by timing. Emission reductions may coincide with periods of poor dispersion, producing limited immediate benefit. Conversely, favorable dispersion during low emission periods may lead to misleadingly low concentrations. Temporal averaging obscures this mediation.

Chemical processes introduce an additional layer of temporal sensitivity. Many atmospheric reactions are nonlinear and highly dependent on sunlight intensity, temperature, humidity, and mixing depth. In coastal environments, these controlling variables fluctuate rapidly due to cloud formation, marine air intrusion, and

boundary layer stratification. Short periods of strong sunlight may drive rapid photochemical production, followed by periods of suppression due to cloud cover or stable stratification.

When chemical processes are solved using averaged inputs, nonlinear behavior is distorted. Reaction rates calculated from averaged sunlight and averaged concentrations do not equal the average of reaction rates calculated from instantaneous conditions. This mathematical property leads to systematic errors in secondary pollutant formation. In coastal cities, where rapid alternation between favorable and unfavorable chemical environments is common, this distortion becomes pronounced.

Temporal averaging also affects the representation of moisture dependent chemistry. Coastal atmospheres often experience rapid changes in humidity due to marine air penetration or fog formation. These changes influence heterogeneous reactions and gas particle partitioning. Averaged humidity values fail to capture the episodic conditions under which certain reactions become dominant. The model output reflects a diluted version of chemical reality that underestimates episodic production or persistence.

Exposure assessment represents another domain where temporal scale choices carry significant implications. Human exposure is not determined solely by average concentration levels. It is shaped by short term peaks, timing of exposure relative to activity patterns, and duration of elevated concentrations. Coastal cities often experience sharp concentration spikes during early morning or evening transition periods when people are commuting or engaging in outdoor activities.

Models that rely on hourly or daily averages distribute these spikes across longer intervals. The resulting exposure estimates suggest moderate risk spread evenly over time rather than acute risk concentrated in specific windows. This representation aligns with regulatory metrics but diverges from lived experience. Applicability failure arises because the model output satisfies formal requirements while failing to capture meaningful exposure dynamics.

The issue is not simply insufficient temporal resolution. It is the underlying assumption that atmospheric processes can be meaningfully represented through averaged states. In coastal environments, process sequencing matters. The order in which dispersion regimes occur shapes pollutant fate. A period of stagnation followed by ventilation produces different outcomes than ventilation followed by stagnation, even if averaged conditions are identical. Temporal averaging erases this order dependence.

This loss of sequencing has implications for understanding pollutant accumulation and release. Coastal atmospheres often store pollutants during stable periods and release them during subsequent mixing events. Models that average across these periods represent storage and release as simultaneous processes, flattening the dynamic into a steady state approximation. The resulting concentration fields lack temporal coherence with actual atmospheric behavior.

Temporal mismatch also affects how models respond to synoptic changes. In coastal cities, large scale weather systems interact with local circulations in time dependent ways. A synoptic wind shift may coincide with a sea breeze phase, reinforcing or weakening it depending on timing. Hourly averaged inputs may capture the synoptic shift but miss its interaction with the local cycle. The model output reflects an averaged compromise rather than the actual sequence of dominance and suppression.

This problem extends to model evaluation. Observational data in coastal cities often show high temporal variability, with sharp transitions and episodic peaks. When models are evaluated against averaged observations, apparent agreement may conceal systematic timing errors. The model reproduces mean values while misplacing peaks and troughs. From a regulatory perspective, the model may appear acceptable. From an atmospheric perspective, it has failed to represent the process structure.

The reliance on temporal averaging also shapes how uncertainty is interpreted. Discrepancies between modeled and observed concentrations are often attributed to emission uncertainties or measurement error. The role of temporal smoothing as a source of structural distortion is less frequently acknowledged. This leads to repeated attempts to refine inputs without questioning the appropriateness of the temporal framework itself.

In coastal air quality modeling, temporal scale choice functions as a filter that privileges certain processes while suppressing others. Processes

that operate slowly and consistently are amplified. Processes that operate rapidly and intermittently are attenuated. This filtering aligns with inland atmospheric behavior but conflicts with coastal dynamics, where rapid transitions drive accumulation and exposure.

Recognizing this conflict reframes the notion of model applicability. The issue is not whether models can be run at finer temporal resolution, but whether their conceptual structure allows time dependent processes to shape outcomes meaningfully. Simply reducing time step length without revisiting assumptions about averaging, sequencing, and process interaction may improve numerical detail without resolving conceptual mismatch.

Temporal scale also influences how policy relevance is constructed. Regulatory frameworks often rely on averaged concentration thresholds. Models designed to support these frameworks naturally emphasize averaged outputs. In coastal cities, this emphasis can obscure chronic exposure patterns driven by daily cycles rather than annual means. Applicability failure thus extends beyond scientific accuracy into governance relevance.

A coastal perspective highlights the need to treat time not as a neutral axis but as an active dimension of atmospheric organization. Dispersion capacity, chemical transformation, and exposure risk are all functions of temporal alignment between emissions and atmospheric state. Models that suppress this alignment through averaging produce outputs that are stable but incomplete.

Temporal scale and averaging effects therefore represent a central mechanism through which urban air quality models lose explanatory power in coastal environments. They do not merely reduce precision. They reshape the modeled reality in ways that systematically underrepresent accumulation, delay, and episodic exposure. Understanding this mechanism is essential for interpreting model outputs responsibly and for recognizing the limits of applying inland oriented temporal frameworks to coastal cities.

In coastal air quality modeling, time must be understood as a structuring force rather than a convenience. Until this shift is acknowledged, temporal averaging will continue to act as a silent source of applicability failure, producing models that are formally correct yet atmospherically misaligned.

7. Conceptual Implications for Model Application and Development in Coastal Contexts

The limitations observed in the application of urban air quality models to coastal cities point toward a need for conceptual recalibration rather than incremental technical correction. These models do not simply underperform because of insufficient resolution, imperfect parameterization, or incomplete input data. Their difficulties arise from deeper assumptions about how urban atmospheres function, how dispersion unfolds, and how time and space are organized within the modeling framework. Coastal cities expose these assumptions by presenting atmospheric conditions that follow different organizing principles from those embedded in most urban models.

A central implication is the recognition of coastal cities as distinct atmospheric systems rather than as variations of inland urban environments. In many modeling practices, proximity to the sea is treated as an external modifier that can be accounted for through boundary conditions or adjusted surface parameters. This framing underestimates the degree to which land sea interaction restructures atmospheric behavior. Coastal cities are governed by circulations, stratification patterns, and temporal rhythms that are not peripheral influences but dominant drivers. Treating these drivers as secondary effects produces models that are internally coherent yet externally misaligned with atmospheric reality.

This recognition requires a shift in how model scope is defined. Urban air quality models are often presented as general tools applicable across diverse urban contexts with minor adjustments. Coastal environments challenge this universality. Applicability must be understood as conditional, dependent on whether the conceptual premises of the model align with the dominant processes of the setting. This does not imply that existing models are invalid, but that their domain of meaningful interpretation is narrower than often assumed.

Model development in coastal contexts therefore demands flexibility at the level of conceptual structure. Many current modeling frameworks are optimized around stable hierarchies of scale, where synoptic forcing dominates, mesoscale processes modify, and microscale effects refine.

In coastal cities, this hierarchy collapses. Mesoscale circulations such as sea–land breezes exert control over transport and mixing that rivals or exceeds synoptic influences. Boundary layer structure becomes an emergent outcome of interacting air masses rather than a predictable response to surface heating alone.

Integrating this reality into model design requires rethinking how circulations are represented. Rather than embedding them as parameterized corrections within a background flow, they must be treated as organizing systems that shape pollutant fate. This implies a move away from models that assume monotonic transport and toward frameworks that can accommodate retention, return, and layered redistribution. Such a shift is conceptual before it is numerical. It concerns how dispersion is imagined rather than how finely it is calculated.

Another implication concerns the role of interpretation in model application. In practice, urban air quality models are often deployed as decision tools whose outputs are taken at face value. Concentration maps, exposure estimates, and scenario comparisons are treated as direct representations of environmental conditions. In coastal cities, this practice is particularly risky. Model outputs are sensitive to unresolved temporal transitions, layered boundary structures, and circulation timing that are not fully captured within standard frameworks.

Applying models in coastal contexts therefore requires interpretive restraint. Results should be read as conditional projections rather than definitive forecasts. Patterns of agreement between modeled and observed averages should not be conflated with accurate representation of processes. Discrepancies should prompt reflection on conceptual fit rather than immediate adjustment of inputs. This interpretive stance recognizes that models are structured narratives of atmospheric behavior, shaped by assumptions that may or may not hold in a given context.

This has implications for how uncertainty is understood. In many modeling exercises, uncertainty is framed as a statistical property arising from imperfect data or stochastic variability. In coastal cities, a significant portion of uncertainty is structural. It arises from the mismatch between model assumptions and atmospheric organization. Addressing this form of uncertainty requires acknowledging the limits of what the model is designed to represent, rather than attempting to eliminate discrepancy through calibration alone.

A further conceptual implication involves the meaning of dispersion itself. Traditional urban air quality modeling often treats dispersion as a process of dilution, where pollutants are spread over larger volumes and transported away from sources. This logic underpins expectations about the effectiveness of emission controls and the interpretation of meteorological favorability. In coastal cities, dispersion often takes the form of redistribution within a constrained system. Pollutants are shifted between layers, relocated offshore and back onshore, or stored temporarily before reappearing under changed conditions.

This reframing alters how model outputs should be used in planning and policy contexts. Improvements in modeled dispersion do not necessarily correspond to reductions in exposure or risk. Favorable wind conditions may redistribute pollutants rather than remove them. Short term concentration reductions may precede delayed accumulation. Policies evaluated solely on the basis of averaged model outputs risk overlooking these dynamics.

From a development perspective, this calls for models that can express conditional outcomes rather than single trajectories. Instead of predicting one concentration field for a given scenario, models should help identify ranges of behavior tied to circulation timing, boundary layer structure, and emission alignment. This does not require abandoning deterministic modeling, but it does require embracing multiplicity in interpretation.

There is also a broader epistemological implication. Urban air quality models are often positioned as objective mirrors of atmospheric reality. Coastal environments reveal that these models are better understood as interpretive instruments shaped by theoretical commitments. Their strength lies in clarifying relationships under assumed conditions, not in reproducing all atmospheric behaviors. Recognizing this helps prevent overconfidence in model outputs and encourages more reflective use.

This perspective also affects how model success is evaluated. Performance metrics based on averaged concentration agreement may be insufficient indicators of conceptual adequacy in coastal contexts. A model that reproduces mean

values while misrepresenting timing, layering, and recirculation may appear successful while offering limited explanatory insight. Evaluation criteria should therefore be aligned with process representation rather than numerical fit alone.

The conceptual implications extend to interdisciplinary communication. Coastal air quality issues sit at the intersection of atmospheric science, urban planning, public health, and environmental governance. Models serve as boundary objects linking these domains. If their limitations in coastal contexts are not explicitly recognized, they risk conveying false certainty across disciplines. Clear articulation of applicability conditions enhances transparency and supports more informed decision making.

The challenges posed by coastal cities call for a reorientation in how urban air quality models are applied and developed. The emphasis shifts from technical refinement to conceptual alignment, from universal applicability to contextual validity, and from deterministic prediction to conditional interpretation. Coastal environments do not merely complicate modeling practice. They reveal the assumptions upon which that practice rests and invite a more reflective engagement with what models can and cannot represent.

8. Conclusion

Urban air quality models occupy a central position in contemporary environmental science and governance. They structure how pollution is understood, how risks are evaluated, and how policy responses are justified. Their authority rests not only on numerical sophistication but on the assumption that they provide a faithful representation of atmospheric behavior across urban contexts. Coastal cities challenge this assumption in a systematic and revealing way. The difficulties encountered when applying urban air quality models to coastal environments are not isolated anomalies. They expose foundational tensions between model logic and atmospheric reality.

This study has shown that the atmospheric environment of coastal cities operates according to organizing principles that diverge from those embedded in most urban air quality models. Land sea thermal contrast, sea–land breeze circulation, layered boundary structures, and strong temporal variability are not marginal influences but dominant forces shaping pollutant behavior. These forces disrupt assumptions of spatial continuity, predictable mixing, monotonic dispersion, and temporal smoothness that underpin conventional modeling frameworks. When these assumptions are violated, models may continue to function computationally while losing explanatory and interpretive power.

Sea–land breeze circulation exemplifies this misalignment. Rather than facilitating progressive dilution, it often creates recirculation pathways that retain pollutants within a coastal system. Boundary layer dynamics reinforce this retention by limiting vertical mixing and sustaining stratification even under conditions traditionally associated with dispersion. Temporal averaging further suppresses the visibility of these processes by smoothing transitions that are central to accumulation and exposure. Together, these mechanisms transform the coastal urban atmosphere into a constrained and memory dependent system that resists representation through inland oriented modeling logic.

The resulting applicability failures should not be interpreted as evidence that urban air quality models are inherently flawed. Instead, they demonstrate that models are structured interpretations grounded in specific conceptual premises. When these premises align with environmental conditions, models offer valuable insight. When they do not, model outputs risk becoming internally consistent yet environmentally misleading. Coastal cities bring this conditionality into sharp focus.

Recognizing applicability failure as a conceptual issue rather than a technical defect has important implications. It shifts attention away from endless parameter adjustment and toward critical reflection on model scope. It encourages model users to interpret outputs as conditional representations rather than definitive truths. It also highlights the need to communicate model limitations transparently, especially when outputs inform regulatory decisions or public health assessments.

This perspective has practical significance. In coastal cities, apparent improvements in modeled air quality may not correspond to reductions in exposure. Emission controls may yield delayed or uneven benefits due to atmospheric retention and recirculation. Short term pollution episodes may dominate health

risk even when long term averages appear acceptable. Models that emphasize averaged outcomes without capturing timing and structure risk underestimating these dynamics.

At a broader level, the analysis underscores the importance of contextual sensitivity in environmental modeling. Urban atmospheres are not interchangeable. They are shaped by geography, climate, and surface interaction in ways that demand tailored conceptual approaches. Coastal cities represent one of the clearest cases where universal modeling assumptions encounter their limits.

The implications extend beyond air quality modeling itself. They speak to how scientific tools mediate the relationship between knowledge and governance. Models gain authority through apparent objectivity and generality. Coastal environments remind us that this authority is contingent on alignment between model structure and environmental process. Preserving the credibility and usefulness of models therefore requires acknowledging where that alignment weakens.

The applicability failures of urban air quality models in coastal cities are not obstacles to be eliminated but signals to be interpreted. They reveal the boundaries of prevailing modeling paradigms and invite a more reflective engagement with atmospheric complexity. By recognizing coastal cities as distinct atmospheric systems and by treating model outputs as context dependent interpretations, researchers and policymakers can use urban air quality models more responsibly and more effectively. This shift does not diminish the value of modeling. It clarifies its role within the broader effort to understand and manage air quality in diverse urban environments.

References

Angevine, W. M., et al. (2006). Modeling of the coastal boundary layer and pollutant transport mechanisms. *Journal of Applied Meteorology and Climatology.*

Augustin, P., et al. (2020). Impact of sea breeze dynamics on atmospheric pollutants and turbulence structure. *Remote Sensing, 12*(4), 648.

Dacre, H. F., et al. (2007). A case study of boundary layer ventilation by convection and sea breeze collision with synoptic flows. *Journal of Geophysical Research.*

He, J., Kang, Y., Wang, Y., Gu, Y., & Zhong, K. (2024). Effects of sea-land breeze on air pollutant dispersion in street networks with different distances from coast using WRF-CFD coupling method. *Sustainable Cities and Society, 115*, 105757.

Martilli, A., Clappier, A., & Rotach, M. W. (2003). A two-dimensional numerical study of the impact of a city on sea breeze and pollutant dispersion. Theoretical and Applied Climatology.

Melas, D., et al. (1995). Boundary layer dynamics in an urban coastal environment and its role in sea breeze evolution. *Atmospheric Environment, 29*(24), 3605, 3617.

Papanastasiou, D. K., & Melas, D. (2009). Climatology and impact on air quality of sea breeze in an urban coastal environment. *International Journal of Climatology, 29*(2), 305–315.

Salvador, N. (2007). Study of the thermal internal boundary layer in sea breeze conditions. *Revista Brasileira de Meteorologia, 31*(4 suppll 1), 593-609.

Talbot, C., Willart, V., & Khomenko, G. (2007). Impact of a sea breeze on the boundary-layer dynamics and the atmospheric stratification in a coastal area of the North Sea. *Boundary-Layer Meteorology, 125*, 133–154.

Varaprasad, V., et al. (2024). Association between sea-land breeze and particulate matter in five coastal urban locations in India. *Science of The Total Environment*, *913*, 169773.

Wang, B., et al. (2023). Implications of Sea Breezes on Air Quality Monitoring in a Coastal Urban Environment: Evidence from High Resolution Modeling of NO2 and O3. *Journal of Geophysical Research: Atmospheres, 128*, e2022JD037860.

Wang, B., Geddes, J. A., Adams, T. J., Lind, E. S., McDonald, B. C., He, J., et al. (2023). Implications of sea breezes on air quality monitoring in a coastal urban environment: Evidence from high resolution modeling of NO2 and O3. *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037860.