Nigerian Research Journal of Chemical Sciences (ISSN: 2682-6054)
Volume 15, Issue 2, 2026

A Linear–Quadratic Differential-Game Framework for Reconciling Group-Contribution Predictions of the Acentric Factor of Halogenated Hydrocarbons: Formulation, Uniqueness and Cross-Validated Benchmarking

[1]Katule, A.I. and [2*]Otobrise, C.

[1]Department of Mathematics,

Federal University of Petroleum Resources, Effurun, Delta State, Nigeria.

[2]Department of Chemistry, Delta State University, Abraka, Nigeria.

*Corresponding Author: otobrisec@delsu.edu.ng

Accepted: *October 31, 2026* **Published Online***: November 4, 2026*

ABSTRACT

This study addresses large, inconsistent errors in estimating the acentric factor (ω) of halogenated hydrocarbons (HHCs) by Group-Contribution Methods (GCMs) when vapour-pressure data are unavailable. The reconciliation of two GCMs is reformulated as a two-player, non-zero-sum, linear-quadratic differential game over a structural complexity index $\tau \in [0, 1]$. Each player controls its additive correction to minimize a cost combining experimental fidelity, consensus penalty (weight κ) and smoothness penalty (weight θ). Pontryagin's principle yields coupled second-order Euler-Lagrange equations. The open-loop Nash equilibrium is proven to exist, be unique, with convergent best-response iteration, and the ensemble prediction is invariant to κ. Benchmarked on 23 HHCs with experimental ω and both GCM predictions using nested leave-one-out cross-validation, bootstrap CIs and Wilcoxon tests, the Nash ensemble reduced MAPE from 36.1% (best raw GCM) to 20.0% [95% CI 12.5–28.8%; p=0.048]. However, it did not beat static bias correction (20.2%), linear stacking (17.7%), training-mean null (16.3%) or a one-descriptor regression on τ (12.5%). Both GCM sets had negative predictive R^2, showing that input predictions, not the game framework, limit accuracy.

Keywords: Acentric factor, group contribution, halogenated hydrocarbons, differential game, nash equilibrium, cross-validation, model reconciliation

INTRODUCTION

The design of distillation columns, refrigeration cycles, heat pumps and organic Rankine cycles requires reliable thermophysical properties of the working fluids [1, 2]. Pitzer's acentric factor ω quantifies the departure of a fluid's vapour-pressure behaviour from that of a simple fluid of spherical molecules and enters almost every generalised corresponding-states correlation and cubic equation of state [2, 3].

It is defined as

$$\omega = -1 - log_{10}(P_r^{sat})|_{T_r=0.7} \quad (1)$$

where $P_r^{sat} = P^{sat}/P_c$ is the reduced saturation pressure and $T_r = T/T_c$ is the reduced temperature. Because ω requires both critical constants and a vapour-pressure point, it is frequently unavailable for halogenated hydrocarbons (HHCs), many of which are toxic, ozone-depleting or high in global-warming potential, which restricts experimental work [4, 5].

Group-contribution methods (GCMs) estimate ω and other properties from from molecular structure alone [6, 7]. The method of Constantinou, *et al.*, [8] combines first-order groups with second-order corrections through a non-linear transform of ω, while the method of Tahami, *et al.*, [9] uses first- and second-order groups with adjustable parameters that depend on the non-hydrogen atom types present. Higher-order and uncertainty-aware variants have since been developed [10, 11], as have machine-learning approaches [12]. For HHCs specifically, Otobrise and Eferurhobo [13] compared the two methods for 50 compounds, 23 of which have experimental ω in the compilation of Yaws [14], and reported deviations from −101.5 % to +147.3 % with no significant correlation between either method and experiment.

Static comparisons of this kind diagnose but do not correct errors. A natural question is whether two imperfect estimators can be reconciled in a principled way that (i) exploits systematic, structure-dependent error patterns, (ii) is transparent about how each estimator is adjusted and (iii) can be applied to compounds outside the training set. Differential games — dynamic optimisation problems in which several decision-makers with distinct objectives interact through a shared state [15–17] — provide one such framework: each estimator is treated as a player that adapts a correction along an ordered descriptor axis, while a coupling term formalises the incentive for the two estimators to agree.

The objectives of this study are (1) to formulate the reconciliation of two GCMs as a two-player LQ differential game in which the players' costs are genuinely coupled through a consensus term. (2) The open-loop Nash equilibrium via Pontryagin's minimum principle [18], proved its existence and uniqueness, and showed that the ensemble prediction was independent of the coupling weight. (3) to provide a reproducible numerical scheme and benchmark against static and statistical baselines under nested cross-validation, with uncertainty quantification and significance testing. (4) Through this benchmark, the accuracy of the published HHC predictions, rather than the reconciliation method, will show to be the binding constraint. What was needed for the framework to be competitive will be specified.

The overall workflow of this study is summarised in Figure 1.

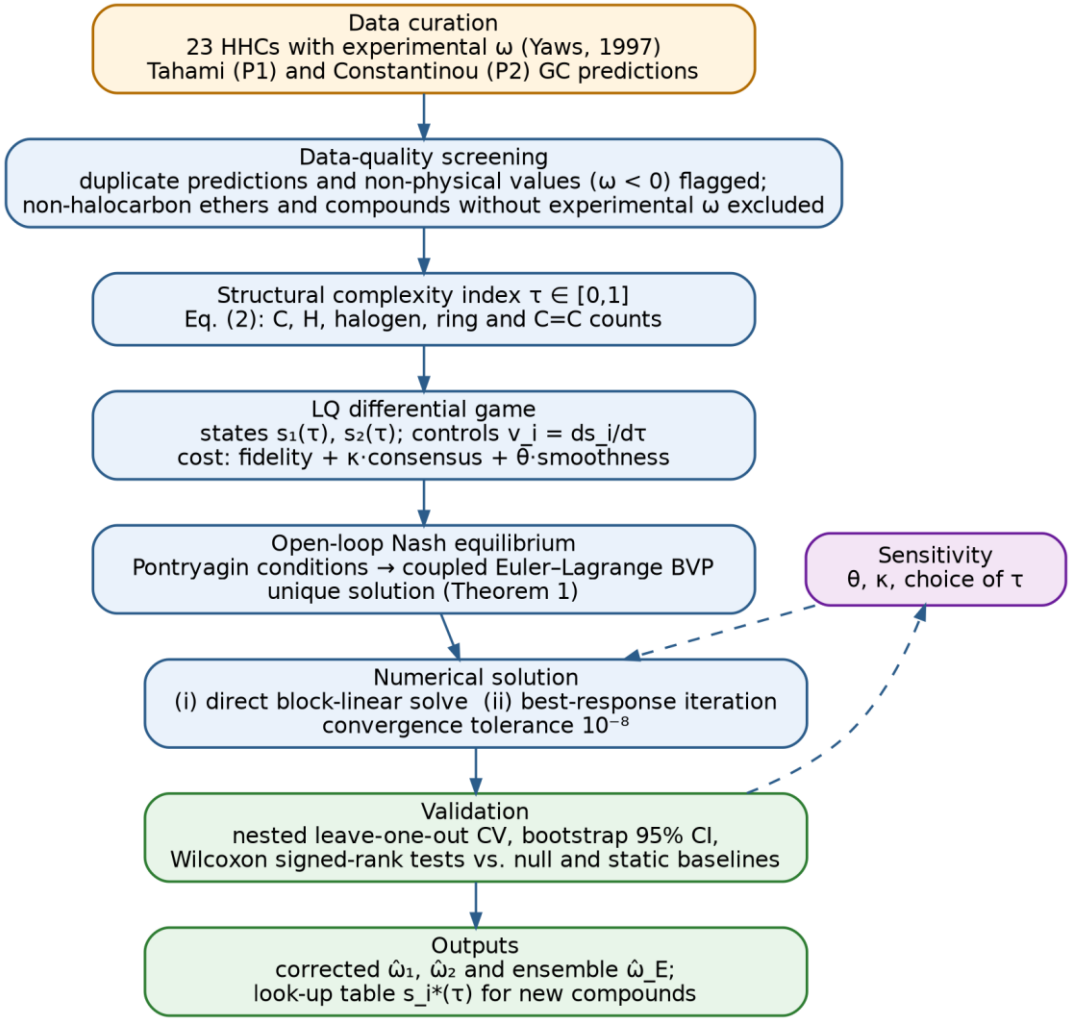


Fig. 1: Workflow of the proposed framework, from data curation and screening through the differential-game formulation, equilibrium computation and cross-validated benchmarking.

Data and quality screening

Source data

Predicted ω values for 50 compounds by both GCMs were taken from Table 1 of Otobrise and Eferurhobo [13]; experimental values for the 23 compounds that have them were taken from Table 2 of the same work, which cites the compilation of Yaws [14]. All 23 published percentage deviations were reproduced exactly from the tabulated values, and the published Pearson coefficients (Tahami vs Constantinou, $r = -0.152$, $p = 0.489$; experiment vs Constantinou, $r = 0.395$, $p = 0.062$; experiment vs Tahami, $r = 0.185$, $p = 0.398$) were also reproduced, confirming correct transcription.

Screening

Since only compounds with experimental ω can be used to fit or validate a correction, the modelling set comprises the 23 HHCs listed in Table 1. The remaining 27 compounds of Otobrise and Eferurhobo [13] lack experimental values and were not used; this set includes two oxygenated

compounds ($C_3H_3F_5O$ and $C_2H_2Cl_4O$), which are not halogenated hydrocarbons, and one physically implausible prediction ($\omega = 5.1668$ for $C_2H_2Cl_4O$ by the Constantinou method). Several anomalies in the source predictions were flagged but retained, because removing them would bias the benchmark: (i) a negative Tahami prediction for 1,1,1-trichloroethane ($\omega = -0.0032$); (ii) identical Constantinou predictions for chemically distinct compounds, e.g. 0.2410 for C_2F_6, C_3F_8, C_5F_{12} and $C_{10}F_{22}$, and 0.0084 for CCl_2F_2, C_2HClF_2, C_2F_4, C_4F_8 and C_6F_{12}; and (iii) identical predictions for 1,1-difluoroethane and 1,1-difluoroethylene by both methods. Since ω increases monotonically with chain length within a homologous series, such duplications indicate errors in group assignment or arithmetic in the source predictions; their consequences are examined in Section 7.

The experimental ω values range from 0.153 (methyl chloride) to 0.395 (hexafluorobenzene), with a mean of 0.241 and a coefficient of variation of 22.7%. This narrow range is important for interpreting percentage-error metrics.

Table 1: Modelling set: experimental acentric factors [12], GCM predictions [11] and structural complexity index τ (Eq. 2), ordered by τ.

Compound	**Formula**	**τ**	**ω_{exp}**	**ω_1 (Tahami)**	**ω_2 (Constantinou)**
Methyl chloride	CH_3Cl	0.000	0.153	0.0851	0.0568
Methyl fluoride	CH_3F	0.000	0.204	0.1365	0.1656
Dichloromethane	CH_2Cl_2	0.073	0.192	0.1595	0.0285
Difluoromethane	CH_2F_2	0.073	0.276	0.1779	0.0285
Bromoethane	C_2H_5Br	0.078	0.183	0.1007	0.1619
Ethyl chloride	C_2H_5Cl	0.078	0.204	0.0721	0.2025
Ethyl fluoride	C_2H_5F	0.078	0.209	0.1318	0.2067
Vinyl fluoride	C_2H_3F	0.112	0.189	0.1019	0.0568
Dichlorofluoromethane	$CHCl_2F$	0.145	0.207	0.0858	0.1047
Trifluoromethane	CHF_3	0.145	0.267	0.1629	0.1828
1,1-Dichloroethane	$C_2H_4Cl_2$	0.151	0.244	0.0447	0.2443
1,1-Difluoroethane	$C_2H_4F_2$	0.151	0.263	0.2373	0.2904
1,1,1-Trichloroethane	$C_2H_3Cl_3$	0.223	0.216	-0.0032	0.2552
1,1,1-Trifluoroethane	$C_2H_3F_3$	0.223	0.253	0.1346	0.1828
Tetrafluoroethylene	C_2F_4	0.330	0.226	0.1952	0.0084
1,1,2-Trichlorotrifluoroethane	$C_2Cl_3F_3$	0.441	0.255	0.0434	0.1835
1,2-Dibromotetrafluoroethane	$C_2Br_2F_4$	0.441	0.250	0.6183	0.1181
Hexafluoroethane	C_2F_6	0.441	0.245	0.2143	0.2410
Bromobenzene	C_6H_5Br	0.637	0.251	0.0551	0.1619
Fluorobenzene	C_6H_5F	0.637	0.247	0.0056	0.2067
Iodobenzene	C_6H_5I	0.637	0.247	0.0033	0.1498
Decafluorobutane	C_4F_{10}	0.888	0.372	0.1914	0.2081
Hexafluorobenzene	C_6F_6	1.000	0.395	0.1472	0.2410

Model formulation

Structural complexity index

Each compound is mapped to a point on a dimensionless "virtual-time" axis that orders molecules by structural complexity. We use a fixed, a priori linear descriptor as shown in Equation 2:

$$X = \alpha N_C + \beta N_H + \gamma N_X + \delta N_R + \varepsilon N_{C=C}, \quad \tau = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (2)$$

where N_C, N_H, N_X, N_R and $N_{C=C}$ are the numbers of carbon atoms, hydrogen atoms, halogen atoms (F + Cl + Br + I), rings and carbon–carbon double bonds (aromatic rings counted as three), and $\alpha = 0.10$, $\beta = 0.02$, $\gamma = 0.15$, $\delta = 0.30$ and $\varepsilon = 0.10$. The weights were fixed before any model fitting and were not tuned to the cross-validation error. τ is a static descriptor: it is computed directly from the molecular formula and therefore places any new compound on the same axis without reference to the training data. Because Eq. (2) weights all halogens equally, compounds such as CH_2Cl_2 and CH_2F_2 share the same τ; the 23 compounds occupy N = 12 distinct nodes.

Players, states and controls

Player 1 (P1) is the Tahami method and Player 2 (P2) the Constantinou method, with raw predictions $\omega_1(\tau)$ and $\omega_2(\tau)$. Each player $i \in \{1, 2\}$ maintains an additive correction state $s_i(\tau)$, so that its corrected prediction is

$$\hat{\omega_i}(\tau) = \omega_i(\tau) + s_i(\tau), \quad i \in \{1, 2\} \qquad (3)$$

and controls the rate at which this correction changes along the complexity axis:

$$\frac{ds_i}{d\tau} = v_i(\tau), \quad \tau \in [0, 1] \qquad (4)$$

Both initial and terminal states are free. Admissible controls are square-integrable functions on [0, 1].

Payoff functionals

Let τ_k be the index of training compound k with experimental value $\omega_{exp,k}$, and let $e_{ik} = \hat{\omega_i}(\tau_k) - \omega_{exp,k}$. Each player minimises

$$J_i = \sum_{k=1}^{n} [e_{ik}^2 + \kappa(\hat{\omega_{ik}} - \hat{\omega_{jk}})^2] + \theta \int_0^1 v_i^2(\tau)\, d\tau \qquad (5)$$

The first term rewards fidelity to experiment; the second, with weight $\kappa \geq 0$, penalises disagreement between the two corrected predictions; and the third, with weight $\theta > 0$, penalises rapid variation of the correction along τ. Because the consensus term depends on the rival's correction s_j, each player's optimal strategy depends on the other's: the problem is a genuine two-player, non-zero-sum

differential game rather than two independent optimal-control problems. With $\kappa = 0$ the coupling vanishes and the problem decomposes into two smoothing problems.

Open-loop Nash equilibrium

Necessary conditions

Writing the data terms as a running cost ℓ_i with point masses at the τ_k, the Hamiltonian of player i is

$$H_i = \ell_i(s,\tau) + \theta v_i^2 + \lambda_{ii} v_i + \lambda_{ij} v_j \quad (6)$$

where λ_{ii} and λ_{ij} are player i's costates for s_i and s_j. Minimising H_i over the player's own control gives

$$\frac{\partial H_i}{\partial v_i} = 2\theta v_i + \lambda_{ii} = 0 \Rightarrow v_i^* = -\frac{\lambda_{ii}}{2\theta} \quad (7)$$

and the costate equation with transversality conditions for free end-states is

$$\frac{d\lambda_{ii}}{d\tau} = -\frac{\partial H_i}{\partial s_i} = -2\sum_k [e_{ik} + \kappa(\hat{\omega}_{ik} - \hat{\omega}_{jk})]\,\delta(\tau - \tau_k), \quad \lambda_{ii}(0) = \lambda_{ii}(1) = 0 \quad (8)$$

Eliminating the costate yields the coupled Euler–Lagrange equations with natural boundary conditions:

$$\theta \frac{d^2 s_i}{d\tau^2} = \sum_k [e_{ik} + \kappa(\hat{\omega}_{ik} - \hat{\omega}_{jk})]\,\delta(\tau - \tau_k), \quad \dot{s}_i(0) = \dot{s}_i(1) = 0 \quad (9)$$

Between data points s_i is therefore linear in τ; at each data point its slope jumps in proportion to the local fidelity-plus-consensus residual. The equilibrium corrections are thus piecewise-linear functions of τ that are flat at both ends. Note that the rate of change of the correction is governed by a second-order condition: a first-order rule of the form $\dot{s}_i \propto -e_i$ does not satisfy the necessary conditions of the stated problem.

Discretisation

Let $\tilde{\tau}_1 < \ldots < \tilde{\tau}_N$ be the distinct node values, $h_m = \tilde{\tau}_{m+1} - \tilde{\tau}_m$, and $s_i \in \mathbb{R}^N$ the nodal corrections. For a piecewise-linear s_i the smoothness integral is exact:

$$\theta \int_0^1 v_i^2\, d\tau = \theta \sum_{m=1}^{N-1} \frac{(s_{i,m+1} - s_{i,m})^2}{h_m} = \theta\, s_i^T L s_i \quad (10)$$

where L is the weighted path-graph Laplacian with edge weights $1/h_m$. Let A be the diagonal matrix whose entry A_{mm} is the number of training compounds at node m, r_i the vector of nodal sums of residuals $\omega_{exp,k} - \omega_{ik}$, and $c_{12} = -c_{21}$ the vector of nodal sums of $\omega_{2k} - \omega_{1k}$. Setting $\partial J_i/\partial s_i = 0$ for each player gives the stationarity conditions

$$[(1 + \kappa)A + \theta L]\, s_i - \kappa A\, s_j = r_i + \kappa c_{ij}, \quad i \neq j \quad (11)$$

which, stacked for both players, form the $2N \times 2N$ linear system $K\,z = b$ with $K = [[M, -\kappa A], [-\kappa A, M]]$, $M = (1 + \kappa)A + \theta L$ and $z = (s_1, s_2)$.

Existence, uniqueness and structure

Theorem 1 (existence and uniqueness). If $\theta > 0$ and the training set is non-empty, the game (3)–(5) has a unique open-loop Nash equilibrium, given by the solution of Eq. (11). *Proof.* K is symmetric and, for any $z = (x, y)$, $z^T K z = x^T(A + \theta L)x + y^T(A + \theta L)y + \kappa(x - y)^T A(x - y)$. L is positive semidefinite with null space spanned by the constant vector 1, and $1^T A 1 = n > 0$, so $A + \theta L$ is positive definite; hence K is positive definite and Eq. (11) has a unique solution. Each J_i is strictly convex in s_i (Hessian $2[(1 + \kappa)A + \theta L] \succ 0$), so the stationarity conditions are sufficient for each player's best response, and the unique stationary pair is the unique Nash equilibrium.

Corollary (potential game and convergence). Because K is symmetric, the game is an exact potential game with potential $\Phi(z) = \tfrac{1}{2} z^T K z - b^T z$ [19]. Alternating best responses are block Gauss–Seidel iterations on a symmetric positive-definite system and therefore converge to the equilibrium from any starting point.

Proposition 2 (κ-invariance of the ensemble). Adding the two equations (11) eliminates κ: $(A + \theta L)\,\bar{s} = \bar{r}$, where $\bar{s} = (s_1 + s_2)/2$ and $\bar{r} = (r_1 + r_2)/2$. The ensemble prediction $\hat{\omega_E} = (\hat{\omega_1} + \hat{\omega_2})/2 = (\omega_1 + \omega_2)/2 + \bar{s}$ is therefore independent of κ. The consensus weight redistributes the correction between the players but cannot change the ensemble.

Proposition 3 (limits). As $\theta \to \infty$, $L\,s_i \to 0$ and the corrections become constants equal to the (consensus-weighted) mean residuals, i.e. a static bias correction. As $\theta \to 0$, the corrections interpolate the nodal mean residuals, which minimises training error but maximises variance. θ therefore controls a bias–variance trade-off and must be selected out-of-sample.

Numerical algorithm

The equilibrium was computed in two independent ways: (i) a direct solve of Eq. (11), and (ii) alternating best responses, as follows.

Step 1. Compute τ for all compounds (Eq. 2); identify the distinct nodes and assemble A, L, r_i and c_{ij} from the training compounds only.

Step 2. Initialise $s_1^{(0)} = s_2^{(0)} = 0$.

Step 3. For $\ell = 0, 1, 2, \ldots$: solve $M\,s_1^{(\ell+1)} = r_1 + \kappa c_{12} + \kappa A\,s_2^{(\ell)}$; then solve $M\,s_2^{(\ell+1)} = r_2 + \kappa c_{21} + \kappa A\,s_1^{(\ell+1)}$.

Step 4. Stop when $\max|s^{(\ell+1)} - s^{(\ell)}| < 10^{-8}$.

Step 5. For a compound not in the training set, compute its τ and evaluate $s_i^*(\tau)$ by linear interpolation between nodes (constant extrapolation beyond the end nodes, consistent with the natural boundary conditions).

All computations were carried out in Python 3 with NumPy and SciPy; the code and data are provided as Supplementary Material S1.

Validation protocol

Cross-validation

Because the corrections are fitted to experimental data, in-sample errors are optimistically biased. All reported errors for fitted models are therefore leave-one-out cross-validated (LOOCV): for each compound k, the equilibrium is computed with compound k removed from the data terms (its node is retained as an unweighted node if no other compound shares it) and $\hat{\omega}_i(\tau_k)$ is then predicted. To avoid optimism arising from hyper-parameter selection [20], θ was additionally selected by an inner LOOCV loop within each outer fold over the grid $\theta \in \{10^{-4}, 10^{-3.5}, \ldots, 10^{1}\}$, with $\kappa = 1$ (Proposition 2 makes the ensemble insensitive to κ). We refer to this as nested LOOCV.

Baselines

The Nash ensemble was compared with: the raw Tahami and Constantinou predictions and their arithmetic mean (no fitting); static bias-corrected versions of each method and of their mean, in which a constant shift equal to the training-set mean residual is added; linear stacking, i.e. an ordinary least-squares regression of ω_{exp} on ω_1 and ω_2 [21]; a null model that predicts the training-set mean of ω_{exp} for every compound; and a descriptor-only model, a least-squares regression of ω_{exp} on τ alone. All fitted baselines were evaluated by LOOCV.

Performance metrics and statistics

For predictions $\hat{\omega}_k$ we report the mean absolute percentage error

$$MAPE = \frac{100}{n}\sum_{k=1}^{n}\frac{|\hat{\omega}_k - \omega_{exp,k}|}{\omega_{exp,k}} \quad (12)$$

together with the average absolute deviation (AAD), root-mean-square error (RMSE), mean bias, the number of compounds predicted within ±20 %, and the predictive coefficient of determination

$$R^2_{pred} = 1 - \frac{\sum_k (\hat{\omega}_k - \omega_{exp,k})^2}{\sum_k (\omega_{exp,k} - \bar{\omega}_{exp})^2} \quad (13)$$

which is negative when a model predicts worse than the sample mean. Ninety-five per cent confidence intervals for MAPE were obtained from 5000 bootstrap resamples of compounds [22]. Paired differences in absolute error were tested with the two-sided Wilcoxon signed-rank test [23].

Agreement with experiment was further examined with Bland–Altman analysis (bias and 95 % limits of agreement) [24]. Because the experimental values span a narrow range (CV = 22.7 %), MAPE alone can be misleading; absolute metrics and the null-model comparison are therefore given equal weight.

RESULTS AND DISCUSSION

Structure of the raw prediction errors

Fig. 2 shows the relative deviations of the raw predictions as a function of molar mass and halogen count. Both methods under-predict ω for most compounds (mean bias −0.106 for Tahami and −0.081 for Constantinou). The Tahami errors are largest for the aromatic halides (−78 % to −99 %) and for 1,1,1-trichloroethane, with a single large over-prediction for 1,2-dibromotetrafluoroethane (+147 %). The Constantinou errors are largest for the small fluorinated and unsaturated compounds (difluoromethane −89.7 %, tetrafluoroethylene −96.3 %, dichloromethane −85.2 %). Neither method shows a monotonic trend with molar mass or halogen count, and both have negative R^2 relative to the sample mean (−8.38 and −3.27 respectively).

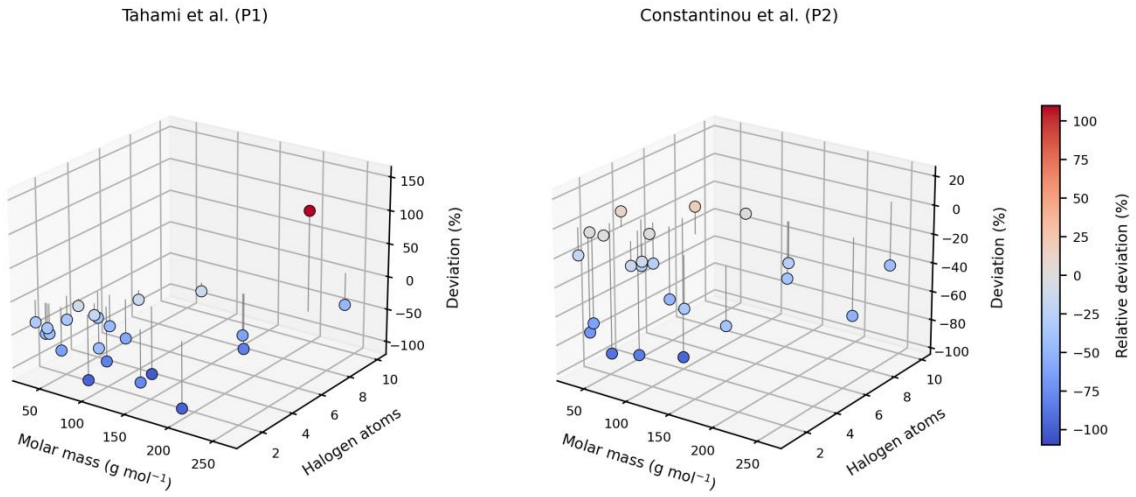


Fig. 2: Three-dimensional representation of the relative deviation of the raw GCM predictions from experiment as a function of molar mass and number of halogen atoms, for (left) the Tahami method and (right) the Constantinou method. Drop lines indicate the zero-deviation plane.

Equilibrium strategies

Fig. 3 shows the equilibrium corrections $s_1^*(\tau)$ and $s_2^*(\tau)$ fitted to all 23 compounds with θ = 0.1 and κ = 1, together with the compound-level residuals they smooth; the nodal values are listed in Table 2 and constitute the look-up table for new compounds. Both corrections are positive almost everywhere, reflecting the systematic under-prediction. The Tahami correction rises from about 0.07 for the C_1 halomethanes to 0.21–0.23 for the aromatic and perfluorinated compounds, except for a

dip to approximately zero at $\tau = 0.441$, where the large positive error of $C_2Br_2F_4$ offsets the negative errors of $C_2Cl_3F_3$ and C_2F_6. The Constantinou correction is flatter (0.03–0.15) and peaks at tetrafluoroethylene ($\tau = 0.330$).

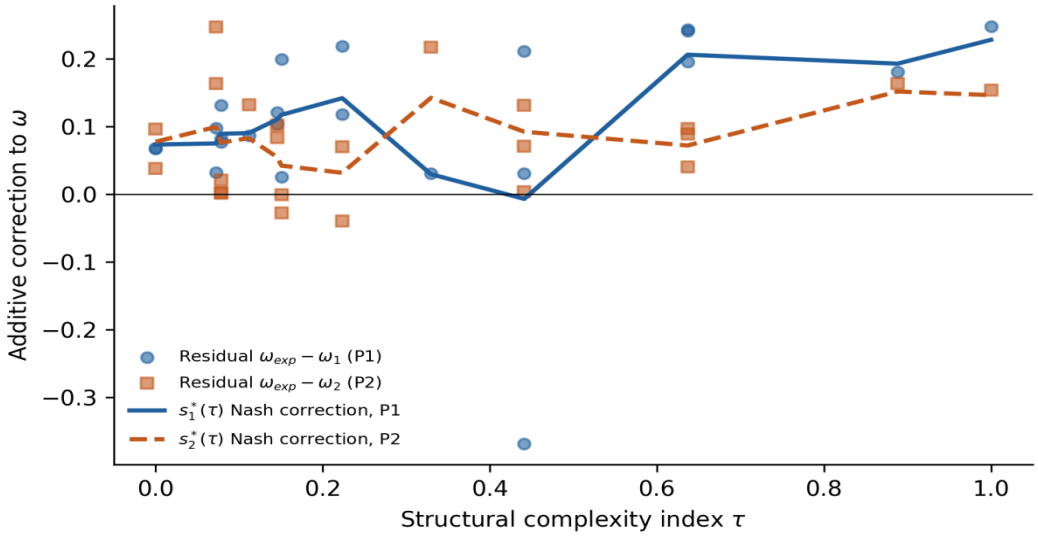


Fig. 3: Open-loop Nash equilibrium corrections $s_1^*(\tau)$ (Tahami, solid) and $s_2^*(\tau)$ (Constantinou, dashed) over the structural complexity index, fitted to all 23 compounds ($\theta = 0.1$, $\kappa = 1$). Symbols are the compound-level residuals $\omega_{exp} - \omega_i$.

Table 2: Nodal equilibrium corrections ($\theta = 0.1$, $\kappa = 1$; all 23 compounds) for use as a look-up table.

τ	s_1^* (Tahami)	s_2^* (Constantinou)	Compounds at node
0.000	0.0735	0.0777	CH_3Cl, CH_3F
0.073	0.0753	0.0996	CH_2Cl_2, CH_2F_2
0.078	0.0895	0.0766	C_2H_5Br, C_2H_5Cl, C_2H_5F
0.112	0.0906	0.0830	C_2H_3F
0.145	0.1105	0.0553	$CHCl_2F$, CHF_3
0.151	0.1175	0.0425	$C_2H_4Cl_2$, $C_2H_4F_2$
0.223	0.1421	0.0319	$C_2H_3Cl_3$, $C_2H_3F_3$
0.330	0.0298	0.1426	C_2F_4
0.441	-0.0067	0.0925	$C_2Cl_3F_3$, $C_2Br_2F_4$, C_2F_6
0.637	0.2063	0.0722	C_6H_5Br, C_6H_5F, C_6H_5I
0.888	0.1930	0.1519	C_4F_{10}
1.000	0.2283	0.1466	C_6F_6

Convergence

Best-response iteration (Fig. 4) reduced both player costs monotonically, with essentially all of the reduction occurring in the first two iterations, and reached the 10^{-8} tolerance after 13 iterations. The converged solution agreed with the direct solve of Eq. (11) to within 1.1×10^{-16}, i.e. to machine precision, confirming Theorem 1 numerically. Final costs were $J_1 = 0.595$ and $J_2 = 0.452$.

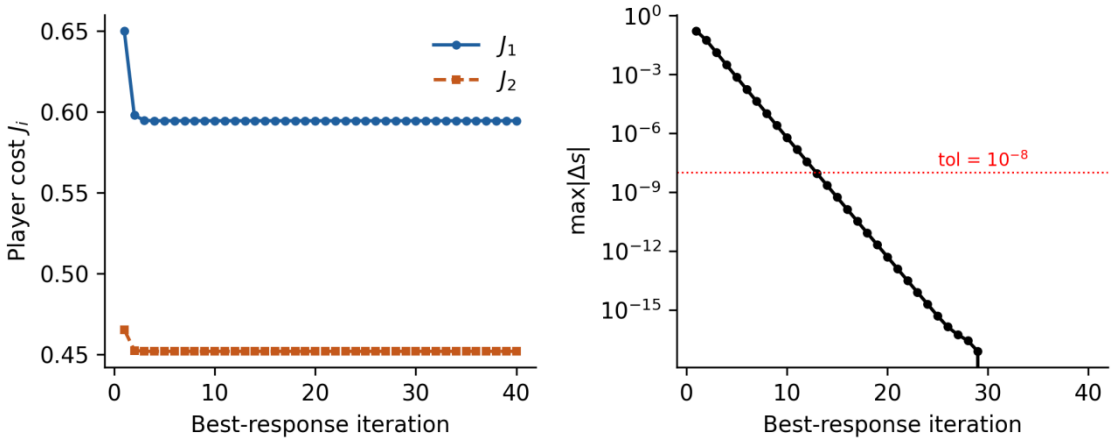


Fig. 4: Convergence of alternating best-response iteration: (left) player costs J_1 and J_2; (right) maximum change in the nodal corrections per iteration on a logarithmic scale, with the 10^{-8} tolerance.

Sensitivity to θ, κ and the complexity index

Fig. 5 shows the LOOCV MAPE as a function of $\log_{10}\theta$ and κ. For the individual players, the surface has a clear structure: P2 reaches its minimum of 22.9 % at small θ ($\approx 2 \times 10^{-3}$) and large κ, and P1 reaches 28.5 % at θ ≈ 0.46 and κ = 1.5. For the ensemble, the surface is flat in κ to within 4×10^{-13} percentage points, as required by Proposition 2, and has a shallow minimum of 18.2 % at θ = 0.1, rising to 19.8 % at both extremes of the grid. The range of ensemble MAPE across four decades of θ is therefore only 1.6 percentage points, indicating that the dynamic (τ-dependent) part of the correction contributes little beyond a constant shift. In the nested procedure, θ = 0.1 was selected in 17 of the 23 outer folds.

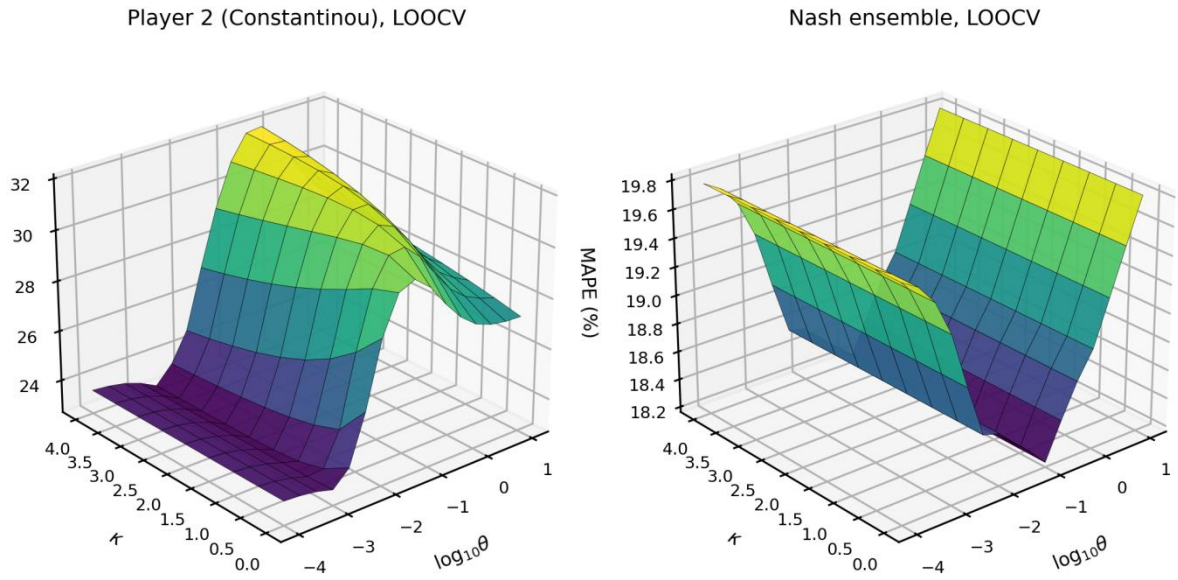


Fig. 5: LOOCV MAPE surfaces over the smoothness weight ($\log_{10}\theta$) and consensus weight (κ) for (left) Player 2 and (right) the Nash ensemble. The ensemble surface is invariant to κ (Proposition 2).

Increasing κ from 0 to 4 at θ = 0.1 changed the individual-player LOOCV MAPE from 28.9% to 31.2 % (P1) and from 29.7% to 30.9% (P2), while the ensemble remained at 18.2%. Coupling the players therefore did not improve their individual out-of-sample accuracy on these data. Replacing Eq. (2) with alternative indices (Table 3) removed the node ties but did not improve the ensemble: molar mass gave 20.3% (optimum at the largest θ, i.e. effectively a constant shift) and an atomic-polarisability sum gave 19.0%.

Table 3: Sensitivity of the Nash ensemble (LOOCV, κ = 1) to the choice of complexity index.

Complexity index	Distinct nodes	Optimal θ	Ensemble MAPE (%)
Eq. (2) structural index (this work)	12	0.100	18.2
Molar mass	23	10.0 (grid limit)	20.3
Sum of atomic polarisabilities	23	0.316	19.0

Predictive benchmark

Table 4 and Fig. 7 summarise the out-of-sample performance of all methods. The nested-LOOCV Nash ensemble achieved a MAPE of 20.0 % (95 % CI 12.5–28.8%), AAD 0.048 and RMSE 0.071, with near-zero bias (−0.0015) and 14 of 23 compounds within ±20%. This is a statistically significant improvement over the best raw method (Constantinou, 36.1 %; Wilcoxon $p = 0.048$). The parity and Bland–Altman plots (Figs. 6 and 8) show that the improvement arises chiefly from removing the systematic negative bias: the Bland–Altman bias falls from −0.081 to −0.0015, while the width of the 95% limits of agreement changes only from 0.302 to 0.283.

However, the Nash ensemble did not differ significantly from the static bias-corrected ensemble (20.2 %; $p = 0.94$), linear stacking (17.7%; $p = 0.66$) or the null model (16.3%; $p = 0.94$), and it was outperformed numerically by all three. The descriptor-only regression on τ, which ignores both GCMs entirely, gave the lowest error of all methods (12.5%, 95% CI 9.4–15.8%; $R^2_{pred} = 0.51$; 19 of 23 compounds within ±20%), although its difference from the Nash ensemble was not significant ($p = 0.26$). Comparing in-sample and LOOCV errors at θ = 0.1 (13.9 % vs 18.2%) quantifies the optimism that would have been reported without cross-validation.

Compound-level results

Table 5 lists compound-level deviations. The ensemble brought 10 compounds within ±10%, including bromobenzene (−1.8%), fluorobenzene (−0.9%), decafluorobutane (+0.1%) and hexafluorobenzene (−8.0%); iodobenzene improved from −98.7% (Tahami) and −39.4% (Constantinou) to −12.6%. The largest residual errors occur where the two players' raw errors have opposite signs at a shared node or where a single anomalous input dominates: 1,2-

dibromotetrafluoroethane (+89.3%), driven by the Tahami over-prediction; 1,1,2-trichlorotrifluoroethane (−51.4%), its node-mate; and 1,1-difluoroethane (+37.1%), whose node-mate 1,1-dichloroethane has a large Tahami under-prediction.

Table 4: Predictive performance on the 23 HHCs. Fitted models are evaluated by LOOCV; raw GCMs require no fitting. CI: bootstrap 95% confidence interval of MAPE.

Method	MAPE (%)	95 % CI	AAD	RMSE	Bias	R^2_{pred}	n (±20 %)
Tahami, raw	56.6	43.6–70.7	0.138	0.164	−0.106	−8.38	4
Constantinou, raw	36.1	25.2–47.9	0.087	0.111	−0.081	−3.27	9
Arithmetic mean, raw	42.8	35.5–49.5	0.104	0.116	−0.094	−3.66	3
Static bias-corrected Tahami	34.8	22.2–53.2	0.086	0.130	0.000	−4.94	9
Static bias-corrected Constantinou	27.2	19.8–35.2	0.064	0.079	0.000	−1.16	9
Static bias-corrected ensemble	20.2	13.5–28.7	0.051	0.071	0.000	−0.75	12
Linear stacking (OLS)	17.7	12.9–22.9	0.044	0.057	+0.003	−0.13	14
Nash P1 (θ = 0.1, κ = 1)	30.1	15.3–50.8	0.074	0.132	+0.001	−5.12	13
Nash P2 (θ = 0.1, κ = 1)	31.2	24.1–38.8	0.071	0.083	−0.004	−1.39	8
Nash ensemble (θ = 0.1)	18.2	11.2–26.9	0.044	0.066	−0.002	−0.52	16
Nash ensemble, nested LOOCV	20.0	12.5–28.8	0.048	0.071	−0.002	−0.74	14
Null model (training mean)	16.3	10.7–23.1	0.039	0.056	0.000	−0.09	17
Descriptor-only regression on τ	12.5	9.4–15.8	0.031	0.037	−0.001	+0.51	19

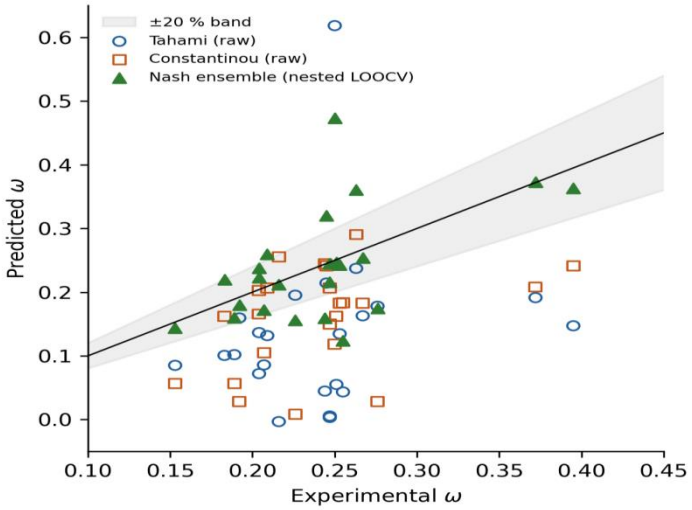


Fig. 6: Parity plot of predicted versus experimental ω for the raw Tahami and Constantinou predictions and for the Nash ensemble under nested LOOCV. The shaded band marks ±20%.

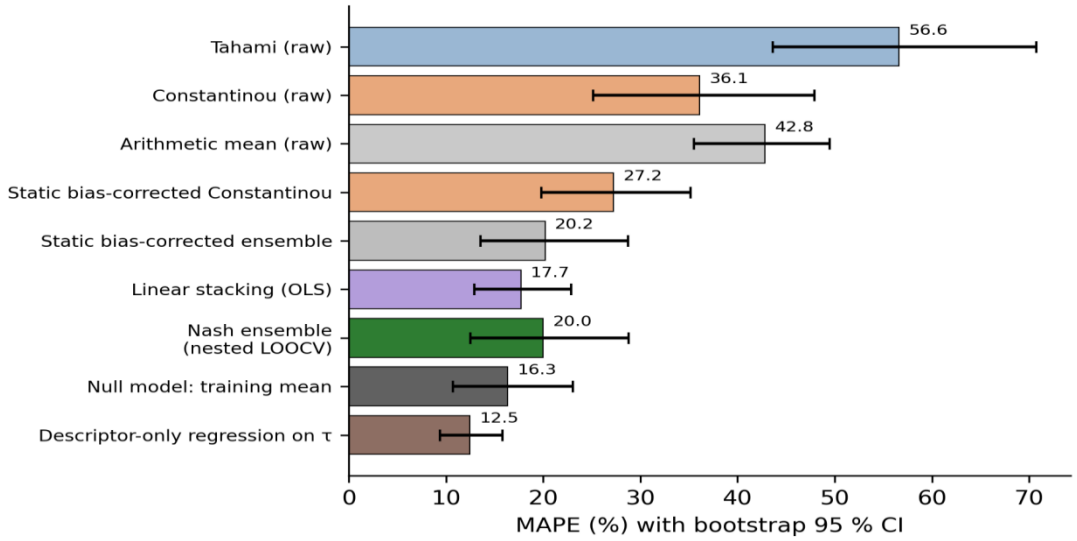


Fig. 7: Mean absolute percentage error of all methods with bootstrap 95 % confidence intervals (5000 resamples). Fitted models are evaluated by LOOCV.

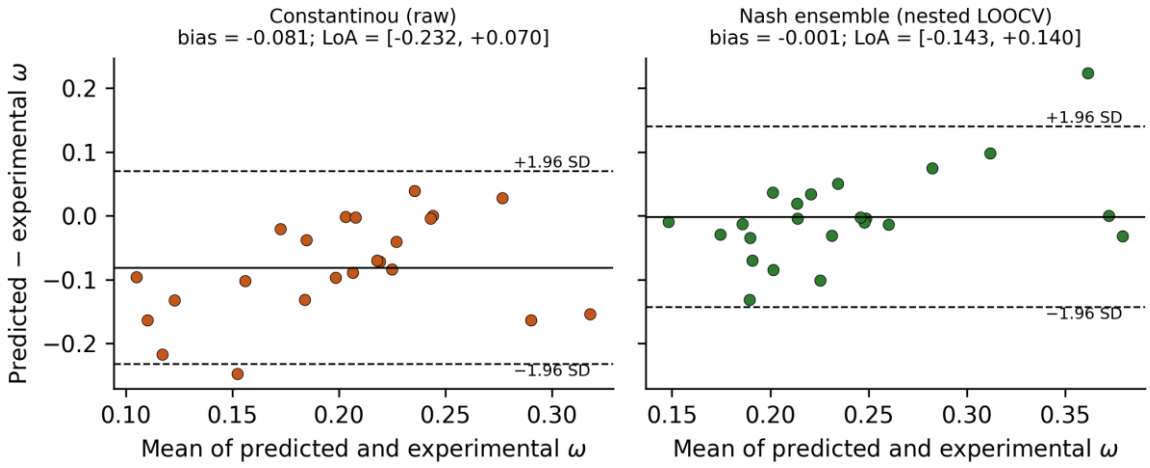


Fig. 8: Bland–Altman plots for (left) the raw Constantinou predictions and (right) the nested-LOOCV Nash ensemble. Solid lines show the mean bias; dashed lines show the 95% limits of agreement (LoA).

What the game contributes

The LQ game provides a transparent, well-posed mechanism for reconciling two estimators: its equilibrium exists and is unique, can be computed exactly or by best-response iteration, and yields interpretable correction profiles and a look-up table for new compounds. The consensus weight has a clear interpretation — it governs how much each method defers to the other — but Proposition 2 shows that it cannot change the averaged prediction. The practical role of κ is therefore limited to producing individually corrected predictions that are mutually consistent; for ensemble use it can be set to zero without loss.

Table 5: Compound-level relative deviations (%) of the raw GCMs and of the Nash ensemble under nested LOOCV.

Compound	ω_{exp}	Tahami (raw)	Constantinou (raw)	Nash ensemble $\hat{\omega}_E$	Nash ensemble dev.
Methyl chloride	0.153	-44.4	-62.9	0.144	-6.2
Methyl fluoride	0.204	-33.1	-18.8	0.237	+16.4
Dichloromethane	0.192	-16.9	-85.2	0.179	-6.5
Difluoromethane	0.276	-35.5	-89.7	0.175	-36.7
Bromoethane	0.183	-45.0	-11.5	0.220	+20.0
Ethyl chloride	0.204	-64.7	-0.7	0.223	+9.4
Ethyl fluoride	0.209	-36.9	-1.1	0.260	+24.2
Vinyl fluoride	0.189	-46.1	-69.9	0.160	-15.4
Dichlorofluoromethane	0.207	-58.6	-49.4	0.172	-16.7
Trifluoromethane	0.267	-39.0	-31.5	0.253	-5.1
1,1-Dichloroethane	0.244	-81.7	+0.1	0.159	-34.7
1,1-Difluoroethane	0.263	-9.8	+10.4	0.361	+37.1
1,1,1-Trichloroethane	0.216	-101.5	+18.1	0.212	-1.9
1,1,1-Trifluoroethane	0.253	-46.8	-27.7	0.243	-4.0
Tetrafluoroethylene	0.226	-13.6	-96.3	0.156	-31.1
1,1,2-Trichlorotrifluoroethane	0.255	-83.0	-28.0	0.124	-51.4
1,2-Dibromotetrafluoroethane	0.250	+147.3	-52.8	0.473	+89.3
Hexafluoroethane	0.245	-12.5	-1.6	0.320	+30.5
Bromobenzene	0.251	-78.0	-35.5	0.247	-1.8
Fluorobenzene	0.247	-97.7	-16.3	0.245	-0.9
Iodobenzene	0.247	-98.7	-39.4	0.216	-12.6
Decafluorobutane	0.372	-48.5	-44.1	0.372	+0.1
Hexafluorobenzene	0.395	-62.7	-39.0	0.363	-8.0

Why the improvement is mainly bias removal

The benchmark shows that almost all of the gain relative to the raw methods comes from removing a common negative bias: the static bias-corrected ensemble performs as well as the Nash ensemble, and the ensemble error varies by only 1.6 percentage points across four decades of θ. With 23 compounds on 12 nodes, there is little information with which to estimate a τ-dependent correction reliably, and LOOCV correctly penalises the variance that such a correction introduces. Any claim that the dynamic component improves accuracy must therefore be tested against a static bias correction, not only against the raw predictions.

The input predictions are the binding constraint

The most consequential finding is that a null model and a single-descriptor regression outperform every GCM-based estimate. On the present data, the published GCM predictions carry less information about ω than the compound's position on a simple structural index. This is not a general verdict on the two GCMs: both were developed and validated on large datasets in their original publications [8, 9], and higher-order GC methods routinely achieve far smaller errors for halocarbons

[10, 11]. Rather, it is strong evidence that the predictions tabulated in [13] contain implementation errors, as also suggested by the duplicated and non-physical values noted in Section 2.2 (for instance, a Constantinou prediction of 0.2410 for both C_2F_6 and $C_{10}F_{22}$, whereas ω increases substantially along the perfluoroalkane series). No correction layer can recover information that is absent from its inputs; the appropriate next step is to recompute both GCMs, verifying the implementation against worked examples in the original papers.

Chemical considerations

Two chemically motivated refinements follow from the results. First, Eq. (2) treats F, Cl, Br and I identically, so that, for example, CH_2F_2 and CH_2Cl_2 share a node although their acentric factors differ (0.276 vs 0.192). Weighting halogens by atomic polarisability removes these ties (Table 3), and a descriptor that also reflects molecular polarity would be physically better grounded, since highly fluorinated molecules combine low polarisability with strong bond dipoles. Second, the experimental ω values should be cross-checked against critically evaluated sources (e.g. NIST REFPROP [25] or the DIPPR 801 database) and reported with uncertainties, because uncertainties of a few per cent in ω are comparable with the residual errors of the better models.

Restoring the first-/second-order formulation

A natural extension, once both GCMs are recomputed at each level, is to let each player also control the weight $u_i(\tau) \in [0, 1]$ placed on its second-order contributions:

$$\hat{\omega_i} = (1 - u_i)\,\omega_i^{(1)} + u_i\,\omega_i^{(2)} + s_i \qquad (14)$$

where $\omega_i^{(1)}$ and $\omega_i^{(2)}$ are the first-order-only and first-plus-second-order predictions. Treating u_i as an additional state with a quadratic penalty on its rate of change keeps the problem linear–quadratic in (u_i, s_i), so Theorem 1 and Proposition 2 extend directly, with the box constraint on u_i handled by projection. The present data do not report the two levels separately, so this extension could not be evaluated here.

Limitations

1. The modelling set is small (23 compounds, 12 nodes) and narrow in ω (0.153–0.395); conclusions about τ-dependence are correspondingly uncertain, and no external test set was available.
2. The equilibrium is open-loop: strategies depend on τ only. A feedback (closed-loop) formulation would require solving coupled Riccati or Hamilton–Jacobi–Bellman equations.

3. The applicability domain is limited to the τ range of the training set; corrections for compounds outside [0, 1] are constant extrapolations and should be flagged.
4. All conclusions are conditional on the published GCM predictions.

CONCLUSIONS

We have formulated the reconciliation of two group-contribution estimators of the acentric factor as a coupled linear–quadratic differential game, derived its open-loop Nash equilibrium from Pontryagin's minimum principle, and proved that the equilibrium exists, is unique, is reached by best-response iteration and yields a κ-invariant ensemble. Under nested cross-validation on 23 halogenated hydrocarbons, the Nash ensemble reduced MAPE from 36.1 % to 20.0 % relative to the best raw method (p = 0.048), principally by removing systematic under-prediction. It did not outperform simpler static, statistical or descriptor-only baselines, and both published GCM prediction sets performed worse than the sample mean. The framework is therefore a rigorous and reproducible correction layer, but its value for halogenated hydrocarbons can only be established once the underlying GCM predictions are recomputed — preferably at both first- and second-order levels — and validated on a larger dataset with an external test set. We recommend that future reconciliation schemes for property estimators always be benchmarked against static bias-correction and null-model baselines under cross-validation.

REFERENCES

[1] Poling, B.E., Prausnitz, J.M. & O'Connell, J.P. (2001). The Properties of Gases and Liquids, 5th ed., McGraw-Hill, New York.

[2] Reid, R.C., Prausnitz, J.M. & Poling, B.E. (1987). The Properties of Gases and Liquids, 4th ed., McGraw-Hill, New York.

[3] Pitzer, K.S., Lippmann, D.Z., Curl, R.F., Huggins, C.M. & Petersen, D.E. (1955). The volumetric and thermodynamic properties of fluids. II. Compressibility factor, vapor pressure and entropy of vaporization, *Journal of American Chemical Society* 77 3433–3440. https://doi.org/10.1021/ja01618a002

[4] Pitzer, K.S. (1955). The volumetric and thermodynamic properties of fluids. I. Theoretical basis and virial coefficients, *Journal of American Chemical Society* 77 3427–3433. https://doi.org/10.1021/ja01618a001

[5] Mondejar, M.E., Cignitti, S., Abildskov, J., Woodley, J.M. & Haglind, F. (2017); Prediction of properties of new halogenated olefins using two group contribution approaches, Fluid Phase Equilib. 433 79–96. https://doi.org/10.1016/j.fluid.2016.10.020

[6] Otobrise, C.& Monago, K.O. (2015) 'New group contribution method for predicting the critical pressures of fatty acids and triglycerides', *International Journal of Computational and Theoretical Chemistry*, 3 (6) 45–49. https://doi.org10.11648/j.ijctc.20150306.11.

[7] Otobrise, C., Monago, K.O. & Ibezim-Ezeani, M.U. (2018). Group contribution method for the estimation of critical properties of some linear aliphatic dimethyl esters of dicarboxylic acids. *FUW Trends in Science & Technology Journal* 3 (2), 658 – 663.

[8] Constantinou, L., Gani, R. & O'Connell, J.P. (1995). Estimation of the acentric factor and the liquid molar volume at 298 K using a new group contribution method, *Fluid Phase Equilibria* 103 11–22. https://doi.org/10.1016/0378-3812(94)02593-P

[9] Tahami, S., Ghasemitabar, H. & Movagharnejad, K. (2019). Estimation of the acentric factor of organic compounds via a new group contribution method, *Fluid Phase Equilibria* 499 (2019) 112246. https://doi.org/10.1016/j.fluid.2019.112246

[10] Marrero, J. & Gani, R. (2001). Group-contribution based estimation of pure component properties, *Fluid Phase Equilibria.* 183–184 183–208. https://doi.org/10.1016/S0378-3812(01)00431-9

[11] Hukkerikar, A.S., Sarup, **,** Ten Kate, A., Abildskov, J., Sin, G. & Gani, R. (2012) Group-contribution+ (GC+) based estimation of properties of pure components: Improved property estimation and uncertainty analysis, *Fluid Phase Equilibria* 321 25–43. https://doi.org/10.1016/j.fluid.2012.02.010

[12] Biswas, S. Chung, Y., Ramirez, J., Wu, H. & Green, W.H. (2023) Predicting critical properties and acentric factors of fluids using multitask machine learning, *Journal of Chemical Information and Modeling* 63 4574–4588. https://doi.org/10.1021/acs.jcim.3c00546

[13] Otobrise, C. & Eferurhobo, G. (2024) Prediction of the acentric factor of some halogenated hydrocarbons via group contribution techniques, *Journal of the Nigerian Society of Physical. Sciences.* 6 2119. https://doi.org/10.46481/jnsps.2024.2119

[14] Yaws, C.L. (1997). Handbook of Chemical Compound Data for Process Safety, Gulf Publishing Company, Houston.

[15] Başar, T. & Olsder, G.J. (1999). Dynamic Noncooperative Game Theory, 2nd ed., SIAM, Philadelphia.

[16] Engwerda, J.C. (2005). LQ Dynamic Optimization and Differential Games, Wiley, Chichester.

[17] Dockner, E.J., Jørgensen, S., Van Long, N. & Sorger, G. (2000). Differential Games in Economics and Management Science, Cambridge University Press, Cambridge.

[18] Pontryagin, L.S. Boltyanskii, V.G., Gamkrelidze, R.V. & Mishchenko, E.F. (1962). The Mathematical Theory of Optimal Processes, Interscience, New York.

[19] Monderer, D. & Shapley, L.S. (1996). Potential games, *Games Econic Behavior* 14 124–143. https://doi.org/10.1006/game.1996.0044

[20] Varma, S. & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection, *BMC Bioinformatics* 7, 91. https://doi.org/10.1186/1471-2105-7-91

[21]Wolpert, D.W. (1992). Stacked generalization, *Neural Network* 5 241–259. https://doi.org/10.1016/S0893-6080(05)80023-1

[22] Efron, B. & Tibshirani, R.J. (1993) An Introduction to the Bootstrap, Chapman & Hall, New York, 1993.

[23] Wilcoxon, F. (1945). Individual comparisons by ranking methods, *Biometrics Bulletin 1* (1945) 80–83. https://doi.org/10.2307/3001968

[24] Bland, J.M. & Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement, *Lancet* 327 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8

[25] Lemmon, E, W., Bell, I.H., Huber, M.L. & McLinden, M.O. (2018). NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport Properties (REFPROP), Version 10.0, National Institute of Standards and Technology, Gaithersburg.