

Skeptical review: Constraint-Based Spatio-Temporal Equation Discovery via Balance Law Validation

Summary

This paper proposes a constraint-based, residual-driven framework for validating and refining candidate spatio-temporal balance laws when only a small number of temporal snapshots are available. Using 10 time slices of density ρ and velocity \mathbf{v} on a periodic 128^3 grid (Sec. 2.1, Sec. 3.1), it computes spatial derivatives via FFT/spectral methods (Sec. 2.2) and approximates temporal derivatives with first-order finite differences (Sec. 2.2). The method evaluates candidate equations by comparing observed temporal-change terms (e.g., $\Delta\rho/\Delta t$, $\Delta\mathbf{v}/\Delta t$) against spatially constructed source terms, forming residual fields and scalar error metrics (Sec. 2.3–2.5). Results show small, mostly unstructured residuals for continuity (Sec. 3.2), while an advection-only (pressureless) nonconservative momentum form yields large, structured residuals (Sec. 3.3), motivating a “missing-term” diagnosis via correlations with hypothesized forces (Sec. 2.4, Sec. 3.4). A pressure-gradient proxy under an isothermal closure $P \propto \rho$ correlates moderately with the momentum residuals, whereas a simple Laplacian viscosity proxy correlates weakly (Sec. 3.4), leading the manuscript to conclude compressible-Euler-like dynamics with minor viscosity (Sec. 3.5, Sec. 4). The overall workflow is clear and potentially useful as a practical validation/diagnostic tool under sparse-time sampling, but key elements—data provenance, temporal-difference error and scaling, momentum-form choice (conservative vs nonconservative), numerical spectral details (especially dealiasing), and the strength of the “Euler identification” claim based largely on correlation—need clarification and/or additional analyses to support the paper’s central conclusions and to properly position novelty relative to existing PDE discovery/validation methods.

Strengths

- Clear motivation and framing around validating balance laws under sparse temporal sampling while leveraging high-accuracy spatial derivatives (Introduction, Sec. 1; Sec. 2).
- Method is generally straightforward to follow: FFT-based spatial derivatives, finite-difference temporal terms, residual definitions, and scalar error metrics are laid out explicitly (Sec. 2.2–2.5).
- The continuity-vs-momentum contrast is an effective demonstration of the residual framework’s diagnostic value: continuity appears strongly satisfied while the simplified momentum model is clearly incomplete (Sec. 3.2–3.3).
- Residual-field visualizations combined with quantitative summaries (MAE/RMSE and correlation statistics) provide an interpretable path toward hypothesizing missing physics (Sec. 3.2–3.4).

- The paper is largely well organized, and the core balance-law statements and residual definitions are internally consistent (Eqs. (2)–(5) and related Results equations in Sec. 3).

Major issues

1. **Dataset provenance and ground truth are not sufficiently specified.** The dataset is described mainly as 10 time slices of ρ and \mathbf{v} on a periodic 128^3 grid (Sec. 2.1, Sec. 3.1), but it is unclear whether the data are simulated (and if so with what governing equations, solver, parameters, discretization, time step, and boundary conditions), experimental, or synthetic. Without this, it is difficult to interpret residual magnitudes, assess whether the inferred “Euler-like” model matches the true generator, and evaluate the significance/novelty of the identification claims (Sec. 3.5, Sec. 4).

Recommendation: Expand Sec. 2.1 / Sec. 3.1 to specify: (i) data origin (simulation/experiment/synthetic), underlying PDE(s) if known (Euler/Navier–Stokes/etc.), boundary conditions, numerical method, resolution, and actual Δt ; (ii) nondimensionalization/units and characteristic scales (e.g., reference ρ , velocity, length; Reynolds/Mach numbers if applicable); (iii) any preprocessing (filtering, normalization). If ground truth is known, state it explicitly and add a targeted comparison in Sec. 3.5 (what matches, what does not). If details are proprietary, provide as much generic information as possible and consider releasing a subset or a synthetic surrogate dataset for reproducibility.

2. **Positioning and novelty are currently overstated relative to existing PDE discovery/validation approaches.** Although the title/Abstract emphasize “equation discovery,” the presented workflow is primarily residual-based validation of a small set of hand-chosen candidate forms plus a correlation-based missing-term heuristic (Sec. 2.3–2.4, Sec. 3.2–3.4). This is conceptually close to established residual checking, PDE-FIND/SINDy-style libraries with derivative estimation, and weak-form/integral approaches designed for noisy derivatives, but the paper does not clearly articulate what is new (Introduction, Sec. 1).

Recommendation: In the Introduction (Sec. 1) add a focused related-work and positioning subsection that distinguishes your contribution from (i) PDE-FIND/SINDy (including weak-form variants), (ii) integral/constraint-based formulations, and (iii) PINN/adjoint identification. Explicitly state whether the novelty is (a) the sparse-time/high-space regime emphasis, (b) a practical spectral-derivative residual pipeline for 3D fields, (c) a diagnostic workflow for hypothesis testing rather than symbolic search, or (d) something else. If feasible, add a compact baseline comparison (Appendix or Sec. 3): run at least one established method under the same “10-slice” constraint and compare residuals/identified terms. If not feasible, narrow claims to “validation/refinement” and outline how the approach could plug into a broader discovery pipeline (Sec. 4).

3. **Temporal derivative estimation and time-step scaling are insufficiently analyzed.** The method relies on first-order forward differences over only nine intervals (Sec. 2.2), and the unstructured report notes an ambiguity around taking $\Delta t = 1$ “for relative comparison.” Temporal truncation error and unknown/implicit scaling can significantly affect both residual magnitudes (MAE/RMSE) and the apparent need for missing terms (Sec. 3.2–3.4).

Recommendation: Clarify in Sec. 2.2 whether Δt is physical/simulation time or an index step; if known, report it and propagate units consistently. Add a sensitivity analysis: (i) compare forward vs central differences for interior slices (and possibly a 2nd-order scheme), (ii) if higher-frequency data exist, subsample at multiple Δt to quantify how MAE/RMSE and correlations change with temporal sparsity, and/or (iii) include a controlled synthetic test (Appendix) where ground-truth ∂_t is known. Use this to bound how much of the momentum residual could plausibly arise from temporal differencing error versus missing physics (discuss in Sec. 3.5 and Sec. 4).

4. **Momentum-equation form being tested may be mismatched to compressible dynamics, weakening interpretation of residuals and “missing term” attribution.** The paper tests a nonconservative, pressureless velocity form $\partial_t \mathbf{v} \approx -(\mathbf{v} \cdot \nabla) \mathbf{v}$ (Sec. 3.3), while compressible momentum balance is naturally expressed in conservative form $\partial_t(\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v}) + \nabla P = \dots$. Testing $\partial_t \mathbf{v}$ rather than $\partial_t(\rho \mathbf{v})$ can entangle density-variation effects and may alter residual structure and the inferred “pressure term.”

Recommendation: Justify in Sec. 2.3 / Sec. 3.3 why the chosen nonconservative velocity form is the right diagnostic for this dataset (e.g., quantify ‘weak compressibility’ beyond density range, or show density fluctuations are negligible in the momentum budget). Strongly consider adding a parallel validation in Sec. 3.3 of the conservative momentum equation using available ρ and \mathbf{v} : compute $\Delta(\rho \mathbf{v})/\Delta t$ and $-\nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v})$, then analyze residuals and missing-term correlations/regressions for $-\nabla P$ (and viscous stress divergence if applicable). This would substantially strengthen the physical interpretability of Sec. 3.4–3.5.

5. **Pressure/closure identification is not yet supported at the level claimed.** The conclusion that the system is “accurately governed by compressible Euler” with a pressure gradient inferred from an isothermal proxy rests mainly on moderate Pearson correlations ($r \approx 0.60$ – 0.64) with $F_P \propto -(1/\rho)\nabla \rho$ (Sec. 2.4, Sec. 3.4.1–3.4.2), without estimating coefficients, testing competing EOS forms, addressing collinearity among candidate terms, or demonstrating out-of-sample predictive skill (Sec. 3.5, Sec. 4). Correlation is suggestive but not identification.

Recommendation: Tone down definitive language in the Abstract/Sec. 3.5/Sec. 4 to “consistent with” unless stronger evidence is added. Strengthen Sec. 3.4 by moving from correlation-only to coefficient estimation: (i) fit $\mathbf{R}_V \approx c F_P$ (or the conservative-form analogue) and report residual reduction / explained variance; (ii) test EOS vari-

ants $P \propto \rho^\gamma$ (e.g., $\gamma \in 1, 1.4$) and/or fit γ over a small grid; (iii) report statistics across all nine time intervals (mean \pm std) rather than a single representative interval (Table 2 / Sec. 3.4). If feasible, add a short forward-prediction or held-out-interval check to demonstrate that the inferred term(s) improve predictive agreement beyond the analyzed interval.

6. **Viscosity and dissipative effects are not tested with a physically appropriate compressible operator, so concluding “viscous effects are minor” is premature.** The manuscript correlates residuals with $\nabla^2 \mathbf{v}$ (Sec. 2.4, Sec. 3.4.2), but the compressible Navier–Stokes viscous term is the divergence of a stress tensor and typically includes both $\nabla^2 \mathbf{v}$ and $\nabla(\nabla \cdot \mathbf{v})$ contributions (and possibly density/temperature-dependent viscosity). Weak correlation with $\nabla^2 \mathbf{v}$ alone does not rule out viscous/bulk-viscous effects.

Recommendation: In Sec. 2.4 / Sec. 3.4.2, expand the dissipative candidate library to include at least $\nabla(\nabla \cdot \mathbf{v})$ and (if feasible) a combined operator mimicking constant-viscosity compressible stress divergence (up to unknown coefficients). Then perform regression (preferred) or correlation on this expanded set and report whether dissipative terms materially reduce residual variance. If the data were generated by an inviscid solver, state that explicitly in Sec. 2.1 and reframe the viscosity test accordingly.

7. **Spectral-derivative implementation details (and aliasing control) are underspecified, potentially impacting nonlinear terms and residual structure.** For pseudo-spectral evaluation of nonlinear products like $\nabla \cdot (\rho \mathbf{v})$ and $(\mathbf{v} \cdot \nabla) \mathbf{v}$, dealiasing/filtering choices (e.g., 2/3-rule) and treatment of Nyquist modes can significantly affect computed derivatives and thus residuals/correlations (Sec. 2.2–2.4).

Recommendation: Augment Sec. 2.2–2.4 with concrete implementation details: FFT conventions/normalization, how derivatives are taken in Fourier space, how nonlinear products are formed (physical vs spectral), whether dealiasing or spectral filtering is used (and parameters), and how Nyquist modes are handled. If no dealiasing is used, add a short sensitivity check (Appendix acceptable) comparing residuals/correlations with and without standard dealiasing/filtering to demonstrate robustness.

8. **Evidence base is narrow (single periodic fully observed dataset; limited robustness tests), which limits the generality of conclusions and the method’s practical scope (Sec. 3; Sec. 4).** The framework’s behavior under noise, coarser spatial resolution, non-periodic boundaries, partial observability, and different dynamics is not evaluated.

Recommendation: Add at least one additional controlled experiment (Appendix acceptable) where ground truth is known and you can vary temporal sparsity/noise (e.g., 1D Burgers or 2D/3D synthetic flow) to show how residual-based validation and missing-term inference degrade. If additional experiments are infeasible, expand Sec. 4 with a sharper limitations section that explicitly enumerates assumptions (periodicity-

ty/FFT, full-field access, sparse candidate library) and outlines concrete paths to extension (finite-volume derivatives for nonperiodic domains, weak-form constraints for noisy data, handling sparse sensors).

Minor issues

1. Error metrics need scale-aware reporting; MRE is pathological when target terms are near zero (Sec. 2.5, Sec. 3.2; Table 1). The manuscript notes the issue for continuity but still reports MRE values that are misleadingly enormous, and the paper lacks normalized residual measures that make MAE/RMSE interpretable across equations/terms.

Recommendation: Revise Sec. 2.5 and Table 1 to either drop MRE or compute it only where $|T|$ exceeds a threshold (explicitly stated). Add normalized measures such as $|R|_2/|T|_2$, $|R|_2/|S|_2$, or variance-explained for (regression-based) term fits. Use these normalized metrics consistently in Sec. 3.2–3.4 so statements like “MAE \approx 0.035 is small” are quantitatively grounded.

2. Correlation statistics are not clearly aggregated over time. Table 2 / Sec. 3.4 appear to emphasize a single representative interval (e.g., $t = 4 \rightarrow 5$), leaving uncertainty about stability across the nine intervals (Sec. 2.4, Sec. 3.4).

Recommendation: Report correlation/regression results across all nine intervals (mean \pm std; optionally min/max). If behavior is nonstationary, discuss why (e.g., flow regime changes) and whether term identification is interval-dependent. Consider adding an alternative robustness check (e.g., Spearman correlation) if correlation remains a central diagnostic.

3. Vector/scalar conventions are ambiguous for residual magnitudes, metrics, and correlations (Sec. 2.5; Sec. 3.3–3.4; Table 1/Table 2). The momentum residual \mathbf{R}_V is a vector, but MAE/RMSE/MRE are defined for scalar R_i and the text alternates between componentwise and magnitude-based discussion.

Recommendation: In Sec. 2.5, explicitly define how metrics are computed for vectors (per-component then averaged vs norms such as $|\mathbf{R}|$ or $|\mathbf{R}|_2$). Similarly specify whether correlations are computed per component (x/y/z) or on vector magnitudes, whether fields are mean-subtracted, and how samples N are counted (grid points \times intervals \times components). Ensure Table 1/Table 2 captions match these conventions.

4. Figures/tables have actionability issues: multiple captions/content mismatches, unclear time-interval/slice specification, inconsistent color normalization, and missing units/nondimensionalization cues (notably across Figs. 1–4 and 6–10; Sec. 3.2–3.4). This makes it harder to connect qualitative panels to Table 1/Table 2 and to reproduce comparisons across time.

Recommendation: Systematically revise figure captions to specify: (i) which time slice(s) or interval(s) are shown, (ii) which spatial slice/plane and its location (e.g., $z = L/2$), (iii) whether colorbars are shared across panels, and (iv) units or nondimensionalization. Add explicit links in text/captions between tables and the corresponding figures (e.g., “Table 1 averages over all nine intervals; Fig. X shows interval $t = i \rightarrow i + 1$ ”).

5. Workflow presentation is somewhat scattered between Methods and Results (Sec. 2 vs Sec. 3), making it harder for readers to apply the approach to other systems.

Recommendation: Add a short algorithmic summary (end of Sec. 2 or start of Sec. 3): data/preprocess \rightarrow spectral derivatives \rightarrow temporal differencing \rightarrow construct candidate balance laws \rightarrow compute residuals/metrics \rightarrow residual-structure inspection \rightarrow missing-term testing (correlation/regression) \rightarrow model refinement/validation.

6. Interpretation of “weakly compressible” is not tied to standard nondimensional measures (Sec. 3.1). Using only a narrow density range as justification is heuristic without Mach number/sound speed or other context.

Recommendation: Either (i) provide relevant nondimensional parameters from the dataset (Mach number, sound speed proxy, etc.), or (ii) explicitly label the statement as heuristic based on relative density fluctuations and avoid stronger compressibility claims.

7. Closure statement in the proposed governing system is ambiguous. Sec. 3.5 proposes a momentum equation involving P but earlier analysis uses an assumed proxy $P \propto \rho$ (Sec. 3.4.1). As written, the final system introduces an unobserved pressure field without specifying closure or how it would be determined from available observables.

Recommendation: In Sec. 3.5, either (i) explicitly state the assumed closure (e.g., $P = c^2\rho$ or $P = c\rho^\gamma$ with fitted/assumed parameters), or (ii) clearly state that the momentum equation is identified only up to an unknown pressure field requiring additional state variables/closure not present in the dataset.

Very minor issues

1. Typographic/notation consistency: mixed vector formatting (bold vs nonbold), inconsistent magnitude notation, and minor LaTeX/unit formatting issues (Sec. 2.2–2.5, Sec. 3.1–3.5; Table 1).

Recommendation: Standardize vector notation (e.g., consistently use \mathbf{v}), magnitude/norm notation, and unit formatting (e.g., “length/time²” rather than informal strings). Ensure equations/tables/captions use consistent symbols.

2. Grid-size typo/formatting: “1283 periodic spatial grid” appears where 128³ is intended (Abstract; Sec. 2.1; Sec. 3.1).

Recommendation: Replace “1283” with “128³” (and ensure $\Delta x = L/128$ is consistent with the stated grid).

3. Dropped proportionality constants in proxy terms can confuse interpretation even if they do not affect Pearson correlation (Sec. 2.4; Sec. 3.4.1). For example, from $P \propto \rho$ the acceleration term is $-c^2(1/\rho)\nabla\rho$, but c^2 is omitted in F_P . Similarly, viscosity is introduced as $\nu\nabla^2\mathbf{v}$ but evaluated as $\nabla^2\mathbf{v}$.

Recommendation: Explicitly state that multiplicative constants are omitted for pattern-correlation purposes, and (if moving to regression) reintroduce coefficients to be estimated (e.g., c^2 , ν , bulk-viscosity coefficient).

4. Section/heading style inconsistencies and minor line-break/hyphenation artifacts (e.g., inconsistent capitalization; awkward word splits) reduce polish.

Recommendation: Do a final formatting pass to standardize heading styles/capitalization and remove awkward line breaks/hyphenation.

Key statements and references

- • Assuming a simple equation of state characteristic of an isothermal ideal gas, where pressure is proportional to density ($P \propto \rho$), the hypothesized pressure-gradient acceleration term $F_P = -(1/\rho)\nabla\rho$ exhibits a strong positive Pearson correlation with the momentum residual components R_V at $t = 4$, with correlation coefficients of approximately 0.610, 0.641, and 0.606 for the x, y, and z components respectively, indicating that a pressure gradient force is a major component of the missing physics in the simplified momentum equation.
- *Reference(s):* $P \propto \rho$
- • When validating the mass conservation law expressed as $\partial\rho/\partial t = -\nabla \cdot (\rho v)$ against the dataset of ten time slices on a 128^3 periodic grid, the residuals between the observed temporal density change $T_M = \Delta\rho/\Delta t$ and the spatial divergence term $S_M = -\nabla \cdot (\rho v)$ remain consistently small across all nine time intervals, with an average Mean Absolute Error (MAE) of approximately 0.035, demonstrating that the continuity equation accurately describes the system’s mass evolution despite first-order temporal differencing.
- *Reference(s):* (none)
- • Validation of the simplified Euler-like momentum equation $\partial v/\partial t = -(v \cdot \nabla)v$ against the same dataset shows that the residual vector $R_V = T_V - S_V$ has an average Mean Absolute Error (MAE) of 1.717, comparable to the magnitudes of the balanced terms (typically 2–4), and exhibits coherent,

structured spatial patterns rather than noise, thereby indicating that advection alone is insufficient and that substantial additional forces must be included in the governing momentum law.

- *Reference(s)*: (none)

Mathematical consistency audit

This section audits **symbolic/analytic** mathematical consistency (algebra, derivations, dimensional/unit checks, definition consistency).

Maths relevance: substantial

The paper’s mathematics centers on validating candidate balance-law PDEs by comparing a forward-difference approximation of time derivatives to spatial terms computed via spectral differentiation, defining residuals, and correlating momentum residuals with hypothesized missing forces (pressure gradient and viscosity). The analytic content is mostly definitions and consistency of PDE forms rather than multi-step derivations.

Checked items

1. ✘ **Grid size and spacing relation** (Sec. 2.1, p.3 (also Abstract p.1; Sec. 3.1 p.6))

- **Claim:** Fields are on a “1283” periodic grid with $L = 1$ and $\Delta x = \Delta y = \Delta z = L/128$.
- **Checks:** definition-consistency, notation-consistency
- **Verdict:** FAIL; confidence: high; impact: minor
- **Assumptions/inputs:** Uniform Cartesian grid in each spatial dimension, Periodic domain
- **Notes:** “1283” suggests 1283 points in one dimension, but $\Delta x = L/128$ implies 128 points per dimension (i.e., 128^3 total). This reads as a typography/formatting issue but is internally inconsistent as written.

2. ✔ **Spectral first-derivative rule** (Sec. 2.2 (spatial derivatives), p.3)

- **Claim:** Compute $\partial\phi/\partial x_i$ by FFT, multiply by ik_i in Fourier space, then inverse FFT.
- **Checks:** symbolic-correctness, assumption-consistency
- **Verdict:** PASS; confidence: medium; impact: minor
- **Assumptions/inputs:** Periodic boundary conditions, Standard Fourier differentiation
- **Notes:** The differentiation rule is symbolically correct. The exact definition of k_i (e.g., inclusion of $2\pi/L$ factors) is not given, but that is a scaling/detail omission rather than an internal algebra contradiction.

3. ✔ **Forward difference temporal derivative** (Eq. (1), Sec. 2.2 (temporal derivatives), p.3)

- **Claim:** Approximate $\partial\phi/\partial t$ by $\Delta\phi/\Delta t = (\phi(t + \Delta t, x) - \phi(t, x))/\Delta t$.
 - **Checks:** algebra, definition-consistency
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** Forward Euler / forward difference in time, Uniform time step Δt
 - **Notes:** Equation is correct for a first-order forward difference.
4. ✓ **Continuity equation form** (Eq. (2) Sec. 2.3 p.4; repeated Eq. (9) Sec. 3.2 p.7)
- **Claim:** Mass conservation is $\partial\rho/\partial t = -\nabla \cdot (\rho v)$.
 - **Checks:** symbolic-correctness, sign-consistency
 - **Verdict:** PASS; confidence: high; impact: critical
 - **Assumptions/inputs:** ρ is density, v is velocity, No sources/sinks
 - **Notes:** The form and sign convention are consistent across Methods and Results.
5. ✓ **Mass residual definition** (Eq. (3), Sec. 2.3, p.4)
- **Claim:** Residual $R_M = T_M - S_M$ where $T_M = \Delta\rho/\Delta t$ and $S_M = -\nabla \cdot (\rho v)$.
 - **Checks:** algebra, definition-consistency
 - **Verdict:** PASS; confidence: high; impact: moderate
 - **Assumptions/inputs:** T_M and S_M evaluated on the same spatio-temporal support for comparison
 - **Notes:** If Eq. (2) holds exactly, then $T_M = S_M$ and $R_M = 0$; algebra is consistent.
6. ✓ **Advective acceleration expansion** (Sec. 2.2 (spatial derivatives), p.3)
- **Claim:** $(v \cdot \nabla)v$ is computed component-wise as $v_x\partial v/\partial x + v_y\partial v/\partial y + v_z\partial v/\partial z$.
 - **Checks:** symbolic-correctness, notation-clarity
 - **Verdict:** PASS; confidence: medium; impact: minor
 - **Assumptions/inputs:** $v = (v_x, v_y, v_z)$, $\partial v/\partial x$ denotes the vector of partial derivatives of each component with respect to x
 - **Notes:** Expression is the standard vector identity; wording could clarify that $\partial v/\partial x$ is a vector, but no internal contradiction.
7. ✓ **Simplified momentum (advective) hypothesis** (Eq. (4) Sec. 2.3 p.4; repeated Eq. (10) Sec. 3.3 p.10)
- **Claim:** Hypothesized momentum equation: $\partial v/\partial t = -(v \cdot \nabla)v$.
 - **Checks:** symbolic-correctness, sign-consistency
 - **Verdict:** PASS; confidence: high; impact: critical
 - **Assumptions/inputs:** Acceleration modeled solely by advection (no pressure/viscosity/body forces)

- **Notes:** As a candidate model, the equation is self-consistent and matches later residual construction.
8. ✓ **Momentum residual definition and implied missing force sign** (Eq. (5) Sec. 2.3 p.4; discussion Sec. 3.3 p.10)
- **Claim:** Define $R_V = T_V - S_V$ with $T_V = \Delta v / \Delta t$ and $S_V = -(v \cdot \nabla)v$.
 - **Checks:** algebra, sign-consistency, model-consistency
 - **Verdict:** PASS; confidence: high; impact: critical
 - **Assumptions/inputs:** T_V and S_V evaluated compatibly in time (e.g., both at t , using the same interval), Residual interpreted as missing acceleration
 - **Notes:** $R_V = \Delta v / \Delta t + (v \cdot \nabla)v$. If the true equation is $\Delta v / \Delta t = -(v \cdot \nabla)v - (1/\rho)\nabla P$, then $R_V = -(1/\rho)\nabla P$, matching the later correlation target sign.
9. ✓ **Pressure-gradient term under $P \propto \rho$** (Sec. 2.4 p.5; Sec. 3.4.1 p.12)
- **Claim:** Assuming $P \propto \rho$, the pressure acceleration $-(1/\rho)\nabla P$ becomes proportional to $-(1/\rho)\nabla \rho$; define $F_P = -(1/\rho)\nabla \rho$.
 - **Checks:** algebra, dimensional-consistency, definition-consistency
 - **Verdict:** PASS; confidence: high; impact: moderate
 - **Assumptions/inputs:** $P = C\rho$ for some constant C (not specified), $\rho > 0$ to avoid division by zero
 - **Notes:** From $P = C\rho$, $\nabla P = C\nabla \rho$, hence $-(1/\rho)\nabla P = -C(1/\rho)\nabla \rho$. The paper correctly states “proportional” but then drops C in the definition of F_P ; that is acceptable for correlation but should be stated as a normalization choice.
10. △ **Viscous term definition vs notation** (Sec. 2.4 p.5; Sec. 3.4.2 p.13)
- **Claim:** Viscous force is often $\nu \nabla^2 v$, but computed hypothesized term is $F_\nu = \nabla^2 v$.
 - **Checks:** definition-consistency, dimensional-consistency
 - **Verdict:** UNCERTAIN; confidence: high; impact: minor
 - **Assumptions/inputs:** Testing spatial pattern of viscosity up to scaling, ν is constant (if included)
 - **Notes:** Not internally contradictory if ν is intentionally omitted (equivalent to setting $\nu = 1$ or ignoring multiplicative constants for correlation), but this intent is not stated explicitly and could confuse readers about units/magnitudes.
11. ✓ **MAE definition** (Eq. (6), Sec. 2.5, p.5)
- **Claim:** $MAE = (1/N) \sum |R_i|$.
 - **Checks:** algebra, definition-consistency

- **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** R_i are scalar residual samples
 - **Notes:** Correct scalar MAE definition.
12. ✓ **RMSE definition** (Eq. (7), Sec. 2.5, p.5)
- **Claim:** $RMSE = \sqrt{(1/N) \sum R_i^2}$.
 - **Checks:** algebra, definition-consistency
 - **Verdict:** PASS; confidence: high; impact: minor
 - **Assumptions/inputs:** R_i are scalar residual samples
 - **Notes:** Correct scalar RMSE definition.
13. △ **MRE definition and interpretation** (Eq. (8) Sec. 2.5 p.5; discussion Sec. 3.2 p.9)
- **Claim:** $MRE = (1/N) \sum |R_i|/(|T_i| + \epsilon)$, with $\epsilon = 1 \times 10^{-8}$ to avoid division by zero; large MRE can occur where $T_i \approx 0$.
 - **Checks:** algebra, limiting-case-sanity, definition-consistency
 - **Verdict:** UNCERTAIN; confidence: medium; impact: minor
 - **Assumptions/inputs:** T_i denotes the observed temporal derivative term used as denominator, R_i and T_i are compatible scalars (or magnitudes/components if vector-valued)
 - **Notes:** Scalar formula is correct and the explanation about inflation when $|T_i|$ is small is mathematically sound. However, for momentum T_i is a vector (T_V), and the paper does not specify whether $|T_i|$ denotes magnitude or componentwise absolute value in the metric computation.
14. ✓ **Proposed momentum equation consistency with residual correlation** (Sec. 3.5, p.13–14)
- **Claim:** Final proposed momentum law: $\partial v / \partial t + (v \cdot \nabla)v = -(1/\rho)\nabla P$.
 - **Checks:** sign-consistency, notation-consistency
 - **Verdict:** PASS; confidence: high; impact: moderate
 - **Assumptions/inputs:** Inviscid (no viscosity) and no other body forces, Same sign conventions as in Eq. (10) and Eq. (5)
 - **Notes:** This form matches the earlier residual identity $R_V = \Delta v / \Delta t + (v \cdot \nabla)v$, so R_V should align with $-(1/\rho)\nabla P$, consistent with correlating R_V to F_P .

Limitations

- Only the PDF text provided was audited; figures were not numerically checked and no computations were reproduced (per scope).
- Several implementation-level details needed to fully verify temporal/spatial term alignment are not specified (e.g., whether spatial terms are evaluated at t or $t + \Delta t$ when paired with forward differences); these omissions limit verification of some com-

parisons but do not create direct symbolic contradictions.

- The audit does not assess whether the chosen candidate PDEs are physically appropriate; it only checks internal mathematical consistency.

Numerical results audit

This section audits **numerical/empirical** consistency: reported metrics, experimental design, baseline comparisons, statistical evidence, leakage risks, and reproducibility.

A set of 13 numerical consistency checks were specified (arithmetic recomputation, count consistency, exact cross-references, inequality checks, and range/approximate-match checks), but no check results were produced due to an execution error. Three additional numerically-relevant claims tied to figures/underlying data remain unverified per scope limitations.

Checked items

1. [△ C1_grid_spacing_L_over_128](#) (Page 3 (Methods §2.1 Dataset))
 - **Claim:** Domain size is $L = 1$ on a 128^3 grid; therefore grid spacing is $\Delta x = \Delta y = \Delta z = L/128$.
 - **Checks:** `recompute_from_given_values`
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified as exact arithmetic ($\Delta = 1/128 = 0.0078125$ and equality across axes), but no execution result is available.
2. [△ C2_time_intervals_from_time_slices](#) (Page 3 (Methods §2.2 temporal derivatives))
 - **Claim:** Given ten time slices, nine temporal intervals were available for analysis.
 - **Checks:** `count_consistency`
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified as exact integer relationship ($\text{intervals} = \text{slices} - 1$), but no execution result is available.
3. [△ C3_unity_time_step_assumption](#) (Page 4 (Methods §2.2 temporal derivatives))
 - **Claim:** Time step Δt was assumed to be unity for relative comparison of terms.
 - **Checks:** `constant_propagation_check`
 - **Verdict:** UNCERTAIN
 - **Notes:** Candidate describes a heuristic unit/label consistency scan; no execution result is available.
4. [△ C4_epsilon_value_in_MRE](#) (Page 5 (Methods §2.5 Evaluation metrics, MRE definition))

- **Claim:** MRE uses epsilon $\epsilon = 10^{-8}$ to prevent division by zero.
 - **Checks:** numeric_literal_check
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified as exact parse (including Unicode minus handling), but no execution result is available.
5. **C5_table1_rmse_ge_mae_mass** (Page 8 (Results §3.2, Table 1))
- **Claim:** Table 1 lists Avg. MAE=0.035 and Avg. RMSE=0.048 for Mass Conservation.
 - **Checks:** inequality_consistency
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified as $RMSE \geq MAE$, but no execution result is available.
6. **C6_table1_rmse_ge_mae_momentum** (Page 8 (Results §3.2, Table 1))
- **Claim:** Table 1 lists Avg. MAE=1.717 and Avg. RMSE=2.102 for Momentum (Euler-like).
 - **Checks:** inequality_consistency
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified as $RMSE \geq MAE$, but no execution result is available.
7. **C7_table1_values_match_abstract_mass** (Page 1 (Abstract) vs Page 8 (Table 1))
- **Claim:** Abstract states mass conservation residual average MAE of 0.035; Table 1 lists Avg. MAE 0.035 for Mass Conservation.
 - **Checks:** cross_reference_exact_match
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified as exact numeric match ignoring unit label differences, but no execution result is available.
8. **C8_table1_values_match_abstract_momentum** (Page 1 (Abstract) vs Page 8 (Table 1))
- **Claim:** Abstract states simplified momentum residual average MAE of 1.717; Table 1 lists Avg. MAE 1.717 for Momentum (Euler-like).
 - **Checks:** cross_reference_exact_match
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified as exact numeric match, but no execution result is available.
9. **C9_correlation_range_vs_table2** (Page 1 (Abstract) / Page 2 (Intro) / Page 13 (Table 2))

- **Claim:** Paper states Pearson coefficients 0.60–0.64 between momentum residuals and pressure-gradient proxy; Table 2 lists $x=0.610$, $y=0.641$, $z=0.606$.
 - **Checks:** range_includes_listed_values
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified as inclusion within narrative range allowing rounding tolerance, but no execution result is available.
10. \triangle **C10_viscous_corr_around_0p14** (Page 13 (Table 2) vs Page 15 (Conclusions))
- **Claim:** Conclusions say viscous term shows negligible correlation (Pearson coefficients around 0.14); Table 2 lists 0.140, 0.142, 0.143.
 - **Checks:** approximate_match_to_claim
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified as closeness to 0.14 within ± 0.01 , but no execution result is available.
11. \triangle **C11_density_range_width** (Page 7 (Results §3.1, Figure 3 description))
- **Claim:** Density range is approximately 0.987 to 1.007.
 - **Checks:** cheap_derived_quantity
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified to compute width and relative width/mean from stated endpoints; no execution result is available.
12. \triangle **C12_MRE_units_and_magnitude_sanity** (Page 8 (Table 1) and Page 9 (discussion of MRE artifact))
- **Claim:** Table 1 reports Mass Conservation Avg. MRE = 38591.7, and text explains it is inflated due to small denominators with epsilon 1×10^{-8} .
 - **Checks:** order_of_magnitude_consistency
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified as a heuristic magnitude sanity test ($\text{mass_avg_mre} > 1000$ and $> \text{mom_avg_mre}=121.7$); no execution result is available.
13. \triangle **C13_table1_momentum_mre_value** (Page 8 (Table 1))
- **Claim:** Table 1 lists Momentum (Euler-like) Avg. MRE = 121.7.
 - **Checks:** cross_reference_repeated_value
 - **Verdict:** UNCERTAIN
 - **Notes:** Check specified as exact match between table and narrative mentions of 121.7; no execution result is available.

Limitations

- Only parsed text was available; no access to underlying numerical dataset or computed residual fields, so most metric recomputation checks cannot be performed.
- Figures are present but numeric values in colorbars/axes are not reliably extractable without image-based reading; such checks are excluded by scope.
- Many statements are qualitative (e.g., 'order of magnitude', 'remarkable similarity') and cannot be numerically audited without additional data.
- Execution error prevented running the specified checks: "Sandbox policy violation: call to 'compile' is not allowed".