GAP FILLING OF CLOUD-OBSCURED SATELLITE NO_2 DATA AS GEOMETRIC NAVIGATION ON A RECONSTRUCTED ATTRACTOR: A CONCEPTUAL FRAMEWORK

Plamen Trenchev

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: ptrenchev@space.bas.bg

Keywords: *TROPOMI NO_2; gap filling; shadow manifold reconstruction; Stark embedding theorem; convergent cross mapping; analogue inference; forced nonlinear systems; missing not at random*

Abstract: *Cloud and quality screening removes approximately 65% of daily TROPOMI tropospheric NO_2 pixels. These gaps are not random: they coincide with the meteorological conditions that generate pollution episodes. We present a conceptual framework that treats urban NO_2 as a forced nonlinear dynamical system and replaces statistical interpolation with analogue inference in a reconstructed phase space. Continuous ground-based NO_2 and O_3 measurements, together with ERA5 meteorology, form a multivariate shadow manifold grounded in Stark's embedding theorem for forced systems. This manifold tracks the atmospheric state even when the satellite view is obscured by clouds. Missing values are reconstructed by transferring historical TROPOMI values from dynamically equivalent past states; no surface-to-column regression is trained. We develop two conceptual elements and their realisation. First, photolytic and boundary-layer coordinates act as forcing coordinates that prevent the mixing of incompatible regimes; an ablation on a synthetic forced system shows that removing the forcing coordinate doubles the phase separation between a target state and its retrieved analogues. Second, convergent cross mapping serves as a precondition on dynamical coupling rather than as post-hoc validation. We then set out the eight-stage realisation of the framework and the qualitative properties that distinguish it from pointwise-loss estimators. The framework was applied to ten European urban stations in Trenchev et al. (2026), where cross-map convergence was obtained for all 60 ablation configurations ($p < 0.001$). The present paper develops the reasoning behind that result.*

РЕКОНСТРУКЦИЯ НА ПРОПУСКИТЕ В САТЕЛИТНИ ДАННИ ЗА NO_2, ПРИЧИНЕНИ ОТ ОБЛАЧНО ЕКРАНИРАНЕ, ЧРЕЗ ГЕОМЕТРИЧНА НАВИГАЦИЯ ВЪРХУ РЕКОНСТРУИРАН АТРАКТОР: КОНЦЕПТУАЛНА РАМКА

Пламен Тренчев

Институт за космически изследвания и технологии – Българска академия на науките
e-mail: ptrenchev@space.bas.bg

Ключови думи: *TROPOMI NO_2; попълване на пропуски; реконструкция на сенчести многообразия; теорема на Старк за вграждане; сходимо кръстосано отображение; аналогово заключение; принудителни нелинейни системи; липсващи данни, които не са случайни*

Резюме: *Премахването на облаците и проверката за качество елиминират приблизително 65% от ежедневните пиксели на TROPOMI за тропосферния NO_2. Тези пропуски не са случайни: те съвпадат с метеорологичните условия, които предизвикват епизоди на замърсяване. Представяме концептуална рамка, която разглежда градския NO_2 като принудена нелинейна динамична система и замества статистическата интерполация с аналогово заключение в реконструирано фазово пространство. Непрекъснатите наземни измервания на NO_2 и O_3, заедно с метеорологичните данни от ERA5, образуват многомерна сянкова многообразие, основана на теоремата на Старк за вграждане на принудени системи. Това многообразие проследява състоянието на атмосферата дори когато сателитният изглед е затъмнен от облаци. Липсващите стойности се реконструират чрез прехвърляне на исторически стойности от TROPOMI от динамично еквивалентни състояния от миналото; не се обучава регресия от повърхността към колоната. Разработваме два концептуални елемента и тяхната реализация. Първо, фотолитичните и граничнослоевите координати действат като координати на въздействието, които предотвратяват смесването на несъвместими режими;*

аблация върху синтетична система под въздействие показва, че премахването на координатата на въздействието удвоява фазовото разделяне между целевото състояние и възстановените му аналози. Второ, конвергентното кръстосано картографиране служи като предварително условие за динамичното свързване, а не като валидиране пост-хок. След това представяме осеметапната реализация на рамката и качествените свойства, които я отличават от оценяващите функции с точкова загуба. Рамката беше приложена към десет европейски градски станции в Trenchev et al. (2026), където беше постигната конвергенция на кръстосаното картографиране за всички 60 конфигурации на аблация ($p < 0,001$). Настоящата статия развива разсъжденията, стоящи зад този резултат.

1. Introduction

Satellite observations of tropospheric nitrogen dioxide from TROPOMI/Sentinel-5P [1] have become a primary source of information on urban air quality, offering daily global coverage at 3.5 × 5.5 km². The instrument is a passive sensor operating in the ultraviolet and visible range, and its tropospheric column retrieval is meaningful only when the light path between the sun, the boundary layer and the sensor is unobstructed. Cloud screening, surface albedo constraints and quality-assurance filtering [2, 3] together remove approximately 65% of daily pixels over a typical European urban station. For any application requiring a continuous record – trend estimation, exposure assessment, emission inversion, assimilation into a chemical transport model – that record must first be completed. The standard response is interpolation. Spatial and temporal kriging, climatological composites [4], and increasingly machine-learning regression of the satellite column on co-located surface and meteorological predictors [5] are all applied to this task. They differ in sophistication but share a structural assumption: that the relationship learned from the observed subset of the record transfers to the unobserved subset. Under random missingness this assumption is innocuous. Under the missingness regime that actually governs satellite air-quality data it is not.

Cloud cover is not a random sampling of atmospheric states. Low cloud, high humidity, frontal passage, shallow boundary layer and weak ventilation co-occur, and they co-occur precisely with the synoptic conditions under which urban NO_2 accumulates. The satellite is therefore systematically blind during the episodes that matter most. In the terminology of missing-data theory this is missing not at random [6]: the probability that a value is absent depends on the value itself.

This paper argues that the problem has been framed incorrectly. If urban NO_2 is treated not as a variable to be estimated but as the observable of a forced nonlinear dynamical system, the question changes form. It is no longer what value should be assigned to this timestamp, but which region of the system's state space was the atmosphere occupying at that moment, and has that region been visited before under compatible forcing. The first is a question of statistical estimation; the second is a question of geometry. Our central claim is that under non-ignorable missingness only the second is answerable from the data.

The empirical foundation for this argument is the station-level gap-filling framework published by the present authors in Trenchev et al. [7], which applies multivariate shadow manifold reconstruction to ten European urban stations and reports its configuration space and performance. The present contribution is different in kind. It is a conceptual synthesis: it extracts the theoretical reason why the approach behaves as it does, sets out the three structural elements that distinguish it from regression, describes the eight-stage processing architecture through which those elements are realised, and identifies the conditions under which the framing holds and the conditions under which it must fail. Performance results are not restated; the reader is referred to [7] for the full quantitative assessment.

2. Cloud Gaps Are Not Missing at Random

Rubin's taxonomy distinguishes three missingness regimes [6, 8]. Data are missing completely at random when the mechanism is independent of both observed and unobserved values; missing at random when it depends only on observed covariates; and missing not at random when it depends on the unobserved value itself. The three carry different licences. Under the first, complete-case analysis is unbiased. Under the second, conditioning on observed covariates restores unbiasedness. Under the third, no conditioning on observed quantities is sufficient, and any method proceeding as if the missingness were ignorable transfers the structure of the observed subset onto a population it does not describe.

Satellite NO_2 gaps sit in the third regime, and the mechanism is physical rather than instrumental. Retrieval requires sufficient photon return from below the cloud top, so cloud fraction and cloud radiance fraction above threshold cause rejection [3]. Cloudiness in the mid-latitude urban environment accompanies elevated humidity, reduced incoming shortwave radiation, suppressed convective mixing, and under winter anticyclonic–stratus conditions a shallow, stable boundary layer.

The same conditions reduce ventilation of the urban source and raise surface concentration. Missingness is thus correlated with the magnitude of the quantity being retrieved, through couplings well documented in boundary-layer meteorology [9].

Two consequences follow, and they differ in severity. The first is bias of the mean. If high-NO_2 days are preferentially absent, the clear-sky mean underestimates the true mean. This is widely recognised and partially correctable if the sampling probability can be modelled. The second is loss of the tail, and it is not correctable within the regression framing. A model fitted by minimisation of a pointwise loss on a clear-sky training set produces conditional expectations, and conditional expectations contract variance by construction: the predicted trajectory occupies a smaller volume of state space than the observed one. When the missing states are precisely the extreme states, the model must extrapolate into the region where its contraction is most severe, using a mapping never exposed to that region. The result is a completed record that is smooth, plausible, and well-behaved under conventional error metrics – and systematically devoid of the episodes for which it was completed.

This is the point at which the problem ceases to be one of estimator choice. No reweighting of the loss function and no increase in model capacity removes a deficiency that originates in the geometry of the training sample. What is required is a method whose inference does not pass through a fitted conditional mean at all.

3. Gap Filling as Geometric Navigation

3.1. The shadow manifold and Stark's embedding theorem

Consider the urban atmosphere above a monitoring station as a deterministic dynamical system driven by external forcing, with state evolving on a manifold M of unknown dimension. Neither state nor manifold is observable. What is observable is a set of scalar time series: ground-level NO_2 and O_3 from the continuous monitoring network, and meteorological variables from reanalysis [10]. Takens' theorem [11] and its generalisation by Sauer et al. [12] establish that for an autonomous system a delay-coordinate map built from a single generic observable is, for sufficiently large embedding dimension, a diffeomorphic image of M – a shadow manifold preserving its topology and neighbourhood structure.

The atmosphere is not autonomous. It is driven by the diurnal and seasonal solar cycle and by synoptic forcing external to the station-scale system. The relevant result is therefore not Takens' but Stark's extension to forced systems [13,14], which shows that the embedding property is retained for a skew-product system provided the forcing is itself represented in the reconstruction – through its own coordinates, or through observables carrying its signature. This requirement is not a technical refinement. It is the condition under which the reconstruction is a manifold at all, rather than a projection in which dynamically unrelated states are superimposed.

The multivariate form of the reconstruction [15] admits mixed coordinates drawn from several observables rather than lags of one. The resulting embedding – ground NO_2 and O_3 lags, meteorological coordinates, and forcing coordinates – defines a trajectory that continues to be traced during cloudy periods, because its constituent measurements are ground-based and continuous. The satellite record is absent on those days; the state estimate is not. This asymmetry is the operational core of the framework (Fig. 1).

Gap filling then proceeds as a transfer, not a fit. For a target time t at which an overpass occurred but was rejected by screening, the k nearest neighbours of the reconstructed state x(t) are located on the library manifold, restricted to past states at which a valid column was retrieved. The estimate is the exponentially distance-weighted mean of the historical column values attached to those neighbours,

$$(1) \qquad \hat{V} = \frac{\sum_{i=1}^{N} \omega_i V(t_i)}{\sum_{i=1}^{N} \omega_i}, \; \omega_i = exp\left(-\frac{d_i}{d_1}\right)$$

with d_i the distance to the i-th neighbour in the embedding and d_1 that to the nearest. No functional relationship between surface concentration and tropospheric column is estimated, and none is assumed to exist. The inference is that two atmospheres occupying the same region of state space under the same forcing carry the same column. The estimate is a convex combination of observed historical values and is therefore bounded by the support the library has sampled; it does not extrapolate beyond that support, as a fitted mapping can. It is not, however, variance-preserving by construction – a distance-weighted mean of k analogues compresses variance relative to the true trajectory, which is why [7] applies an explicit variance-restoration step under random and covariate-dependent missingness.

This is analogue inference in the sense of nonlinear forecasting from reconstructed state space [16], and it differs from classical analogue prediction in three respects that matter here: neighbours are sought in a reconstructed multivariate space rather than in the space of raw observations; forcing coordinates constrain which neighbours are admissible (Section 3.2); and the existence of a shared dynamical substrate is tested before any transfer is performed (Section 3.3).

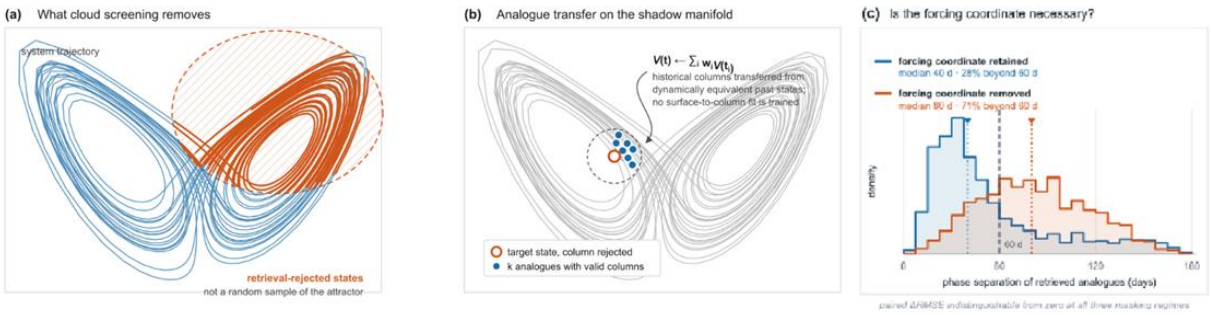


Fig. 1. Gap filling framed geometrically. (a) Cloud and quality screening does not sample the system's state space uniformly: the rejected states occupy a connected region of the attractor rather than a random subset, which is what makes the missingness non-ignorable. (b) In the shadow-manifold formulation the state at a rejected overpass remains reconstructible from continuous ground observables, and the column estimate is obtained by transferring historical values from the k dynamically nearest past states at which a valid retrieval exists; no surface-to-column relationship is estimated. (c) The forcing-coordinate requirement tested directly on a synthetic forced Lorenz-96 system: removing the forcing coordinate doubles the median circular phase separation between a target state and its retrieved analogues and raises the fraction of targets whose analogues come from a different part of the annual cycle from 28% to 71%. The panel shows one representative replicate (medians 40 d and 80 d); medians over the 20 replicates are 40 d and 79 d. Accuracy is unchanged, for the reason given in Section 3.4. The attractor silhouettes in (a) and (b) are schematic and carry no data

3.2. Forcing coordinates as regime separators

Two coordinates carry particular weight: a photolytic regime coordinate derived from solar geometry, and boundary-layer height.

The photolytic coordinate encodes the deterministic astronomical forcing. Without it, a December morning state and a June evening state with comparable surface NO_2 and comparable wind are geometrically adjacent, and the neighbour search treats them as dynamically equivalent. They are not: photochemical regime, mixing depth and the air-mass factor of the retrieval all differ.

Boundary-layer height enters for a different reason. It is the physical quantity linking surface measurement to column: a given surface concentration corresponds to a small column under deep mixing and a large one under a shallow inversion. It is therefore the coordinate along which the surface-column correspondence itself varies, and augmenting the embedding with it was found in [7] to be near-universally beneficial.

Forcing coordinates do not improve a fit; they restrict which library states may serve as analogues. Their role is exclusion, not explanation. Whether that exclusion also improves accuracy is a separate question, and Section 3.4 answers it directly.

3.3. Convergent cross mapping as a precondition

The framework rests on an assumption testable before any value is filled: that ground observables and satellite column are observables of a common dynamical system. If they are not – if station and overlying column are dynamically decoupled at the relevant scales – then neighbourhood in the reconstructed space carries no information about the column, and the transfer is unjustified however well it scores.

Convergent cross mapping [17] provides the test. Cross-map skill ρ(L) is computed as a function of library length L; a genuine shared attractor produces monotone increase and saturation, whereas a spurious or purely forcing-driven association produces a flat profile. Significance is assessed against surrogate ensembles [18]. The essential point is procedural: here CCM is applied before the filling step, so that a configuration failing to converge is declared inadmissible rather than reported with degraded accuracy. The test is graded rather than binary – cross-map skill ranks with coupling strength, and the rejection threshold is set by the surrogate null. No configuration has yet been rejected at the sites examined in [7], so the test has to date served as a verified precondition rather than as an active filter.

4. Realisation: An Eight-Stage Pipeline

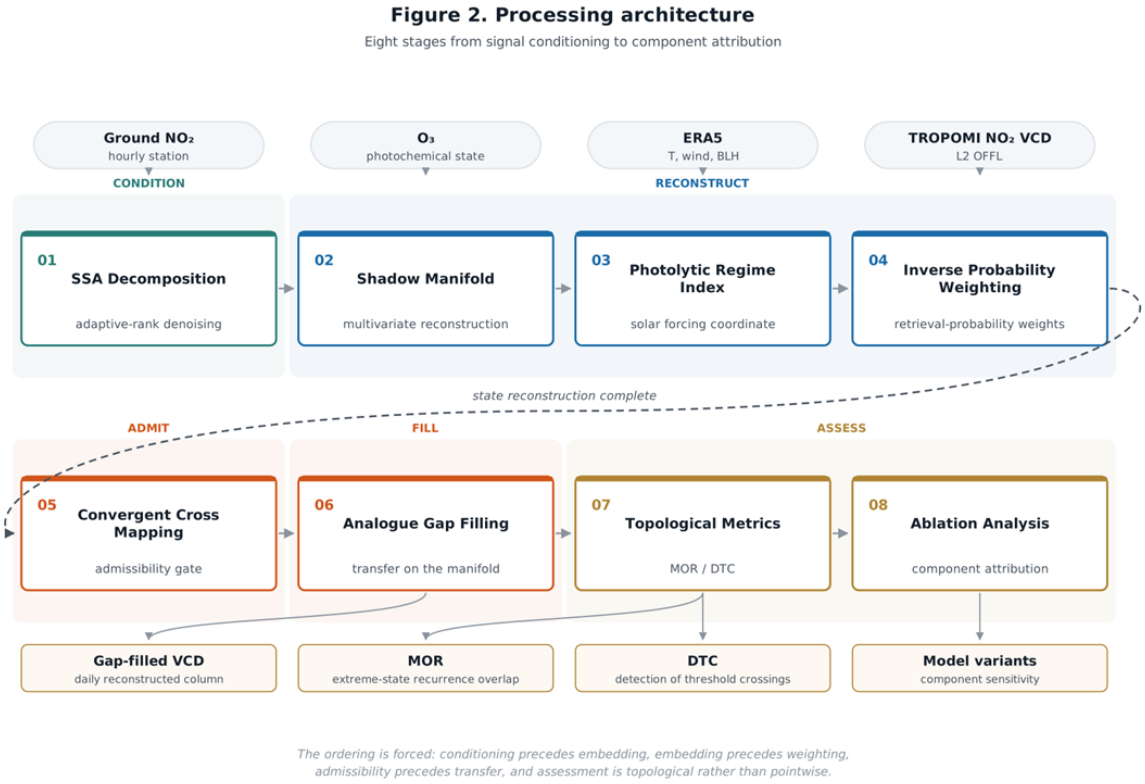


Fig. 2. Processing architecture. Inputs enter at the top and are transformed through eight stages into gap-filled columns and topology-aware validation metrics. The upper row reconstructs the atmospheric state: signal conditioning separates the deterministic component with the retained rank selected by prediction skill rather than by a variance threshold; multivariate shadow manifold reconstruction selects and ensembles candidate embeddings; the photolytic regime index and boundary-layer height enter as forcing coordinates; and inverse-probability weighting prevents the library from being dominated by the fair-weather regime. The lower row acts on that state: convergent cross mapping establishes admissibility before any value is filled, analogue transfer produces the estimate, topological metrics assess structural rather than pointwise fidelity, and ablation attributes each reported property to a specific component. MOR denotes the Manifold Overlap Ratio and DTC Dynamic Trend Capture

The three conceptual elements of Section 3 are realised in a sequence of eight stages (Fig. 2), each of which answers one well-posed question. The architecture is described here because the conceptual argument is only as good as its operational commitments: each stage corresponds to a decision that a reader could make differently, and naming them makes the framework falsifiable rather than merely plausible.

Signal conditioning. The continuous ground record carries instrument noise that would corrupt the neighbourhood structure of the embedding. Singular spectrum analysis separates the deterministic component, with the retained rank selected by a leave-one-out simplex-projection skill criterion rather than by a fixed variance threshold - the rank that best supports short-term prediction, not the one that explains most variance.

Manifold construction. Multivariate shadow manifold reconstruction searches a space of candidate embeddings, each a selection of source series and lags, and retains those with the best out-of-sample skill. Prediction is formed from an ensemble of top-ranked views rather than a single embedding, which both stabilises the estimate and avoids committing to one arbitrary reconstruction.

Forcing coordinates. The photolytic regime index and boundary-layer height enter here, in the role set out in Section 3.2.

Observation weighting. Because retrieval availability is itself state-dependent, library states are weighted by the inverse of their estimated retrieval probability, so that the library is not silently dominated by the fair-weather regime.

Admissibility. Convergent cross mapping against surrogate ensembles, applied as a gate (Section 3.3).

Transfer. Analogue retrieval and distance-weighted transfer of historical columns, as in Section 3.1.

Topological assessment. Two metrics address what conventional error measures do not. The Manifold Overlap Ratio (MOR) quantifies the overlap of recurrence structure between observed and reconstructed records in the extreme regime. Dynamic Trend Capture (DTC) quantifies agreement on exceedances above the 95th percentile, weighted by amplitude fidelity and by the ratio of predicted to observed spread. Both answer whether the completed record has the same dynamical structure as the original, not whether it is close to it pointwise. This distinction matters because a record can be accurate on average while being useless for the episodes it was completed to represent.

Ablation. Each configuration is re-run with components removed, so that every reported property can be attributed to a specific element rather than to the ensemble as a whole.

The ordering is not arbitrary. Conditioning must precede embedding, because a noisy series yields false neighbours; embedding must precede weighting, because weights are defined on reconstructed states; admissibility must precede transfer, because a transfer across an unverified coupling is not interpretable; and assessment must be topological rather than pointwise, because the framework's claim is about structure. A pipeline in which these stages were reordered would not implement the same argument.

5. Qualitative Properties of the Geometric Formulation

Four properties follow from the formulation rather than from tuning, and each was observed across the ten stations examined in [7].

Abstention in place of spurious skill. Where no dynamically compatible analogue exists in the library, the method returns no estimate. At one station in [7] this led to a reported absence of results rather than a number, which was treated as the correct output. A regression returns a numerically well-formed value under identical circumstances, and nothing in its output distinguishes that case from an informed one.

Availability at every overpass. Lag-feature regression requires a complete predictor vector and silently declines to predict when preceding observations are themselves missing; in the winter test subset of [7] a substantial fraction of candidate overpasses was lost for this reason. The geometric formulation requires only that the current state be reconstructible and a compatible neighbourhood exist, and produced an estimate at every overpass. Under non-ignorable missingness the days on which lag-based predictors fail are not a random subset either.

Retention of extreme structure. Across all stations, ablations and missingness regimes examined in [7] – more than 3600 evaluations - no non-dynamical reference method (random forest on full and meteorology-only predictor sets, persistence, climatology, linear interpolation) attained a Manifold Overlap Ratio above 0.104, while the dynamical method never fell below 0.59. The two distributions are separated by a factor exceeding five at their closest approach, with no overlap.

Bias as a diagnostic rather than a defect. Where the library spans the dynamical regimes present in the evaluation period, bias is negligible. Where the system has drifted outside it – through a change in emission baseline, or a combination of heating regime and circulation absent from the historical record – a systematic bias appears whose sign indicates the direction of the drift. For an analogue method this is the correct behaviour: the bias is a statement that the attractor sampled by the library is no longer the attractor being sampled by the atmosphere.

6. Limits and Conditions of Validity

The framework substitutes a continuously observed shadow manifold for the interrupted satellite view. Where no such ground record exists, the state during the gap is unknown and there is nothing to navigate. This restricts the approach to station-level application and to networks with high temporal completeness. It is not a spatial gap-filling method and should not be presented as one.

The deeper constraint is library coverage. A transfer-based estimate exists only if a dynamically equivalent past state has been observed under valid retrieval conditions. The method cannot extrapolate; it can only report that no admissible analogue exists. This is a limitation, but a transparent one.

An attempt to mitigate regime drift by temporal down-weighting of older library states was tested in [7] and did not succeed: no weighting achieved a meaningful bias reduction without a disproportionate accuracy penalty. The negative result is instructive. The library is a geometric object whose value lies in its coverage of state space, and preferentially discarding its older regions reduces that coverage faster than it reduces the regime mismatch.

Finally, the framework inherits the predictability limits of the underlying system. Where local dynamics are low-dimensional and strongly constrained, neighbourhoods are tight and transfer is accurate. Where they are high-dimensional and weakly constrained, neighbourhoods are diffuse and the estimate degrades. The method does not promise uniform skill and should not be evaluated as if it did.

7. Conclusions

Cloud-induced gaps in satellite NO_2 records are missing not at random, and the missingness mechanism is physically coupled to the pollution episodes the record is used to study. Under this regime the question of what value to assign to a missing observation is not answerable by estimation from the observed subset, because the observed subset does not sample the relevant region of state space.

We have argued that the question should be reformulated geometrically: whether the atmospheric state at the missing timestamp is reachable from the observed history, and if so, which historical states it is dynamically equivalent to. This reformulation rests on Stark's embedding theorem for forced systems, requires forcing coordinates to prevent inadmissible analogue matches, admits an operational precondition on dynamical coupling through convergent cross mapping, and is realised in an eight-stage architecture whose ordering follows from the argument itself.

The framework's principal limitation, dependence on library coverage, is also its principal epistemic advantage: it abstains where it cannot know.

References:

1. Veefkind, J.P.; Aben, I.; McMullan, K.; Förster, H.; de Vries, J.; Otter, G.; Claas, J.; Eskes, H.J.; de Haan, J.F.; Kleipool, Q.; et al. TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. Remote Sens. Environ. 2012, 120, 70–83.
2. Eskes, H.J.; van Geffen, J.; Boersma, K.F.; Eichmann, K.-U.; Apituley, A.; Pedergnana, M.; Sneep, M.; Veefkind, J.P.; Loyola, D. Sentinel-5 Precursor/TROPOMI Level 2 Product User Manual Nitrogen Dioxide; Document S5P-KNMI-L2-0021-MA; KNMI: De Bilt, The Netherlands, 2022.
3. van Geffen, J.; Eskes, H.; Compernolle, S.; Pinardi, G.; Verhoelst, T.; Lambert, J.-C.; Sneep, M.; ter Linden, M.; Ludewig, A.; Boersma, K.F.; Veefkind, J.P. Sentinel-5P TROPOMI NO_2 retrieval: impact of version v2.2 improvements and comparisons with OMI and ground-based data. Atmos. Meas. Tech. 2022, 15(7), 2037–2060.
4. Gerber, F.; de Jong, R.; Schaepman, M.E.; Schaepman-Strub, G.; Furrer, R. Predicting missing values in spatio-temporal remote sensing data. IEEE Trans. Geosci. Remote Sens. 2018, 56(5), 2841–2853.
5. Lary, D.J.; Alavi, A.H.; Gandomi, A.H.; Walker, A.L. Machine learning in geosciences and remote sensing. Geosci. Front. 2016, 7(1), 3–10.
6. Rubin, D.B. Inference and missing data. Biometrika 1976, 63(3), 581–592.
7. Trenchev, P.; Avetisyan, D.; Dimitrova, M.; Trencheva, E. Station-level gap filling of TROPOMI NO_2 via physics-informed shadow manifold reconstruction. Remote Sens. 2026, 18(14), 2387. https://doi.org/10.3390/rs18142387
8. Little, R.J.A.; Rubin, D.B. Statistical Analysis with Missing Data, 3rd ed.; Wiley: Hoboken, NJ, USA, 2019.
9. Seibert, P.; Beyrich, F.; Gryning, S.-E.; Joffre, S.; Rasmussen, A.; Tercier, P. Review and intercomparison of operational methods for the determination of the mixing height. Atmos. Environ. 2000, 34(7), 1001–1027.
10. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 2020, 146(730), 1999–2049.
11. Takens, F. Detecting strange attractors in turbulence. In Dynamical Systems and Turbulence, Warwick 1980; Rand, D., Young, L.-S., Eds.; Lecture Notes in Mathematics 898; Springer: Berlin, Germany, 1981; pp. 366–381.
12. Sauer, T.; Yorke, J.A.; Casdagli, M. Embedology. J. Stat. Phys. 1991, 65(3–4), 579–616.
13. Stark, J. Delay embeddings for forced systems. I. Deterministic forcing. J. Nonlinear Sci. 1999, 9(3), 255–332.
14. Stark, J.; Broomhead, D.S.; Davies, M.E.; Huke, J. Delay embeddings for forced systems. II. Stochastic forcing. J. Nonlinear Sci. 2003, 13(6), 519–577.
15. Deyle, E.R.; Sugihara, G. Generalized theorems for nonlinear state space reconstruction. PLoS ONE 2011, 6(3), e18295.
16. Sugihara, G.; May, R.M. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature 1990, 344, 734–741.
17. Sugihara, G.; May, R.; Ye, H.; Hsieh, C.-h.; Deyle, E.; Fogarty, M.; Munch, S. Detecting causality in complex ecosystems. Science 2012, 338(6106), 496–500.
18. Theiler, J.; Eubank, S.; Longtin, A.; Galdrikian, B.; Farmer, J.D. Testing for nonlinearity in time series: the method of surrogate data. Physica D 1992, 58(1–4), 77–94.