

Spatial-temporal tensor completion methods with applications to space weather

Yang Chen

This talk is based on joint work with Hu Sun, Shasha Zou, Zuofeng Shang, Zihan Wang, Anthea Coster, Jingyang Li, Jiuqian Shang

Department of Statistics, University of Michigan



Outline

- 1 Motivating Science Topic: Space Weather Monitoring
- 2 **Imputation**: Spatio-Temporal Matrix Completion Method (**VISTA**)
- 3 Data Product: Constructing A Space Weather Database with VISTA
- 4 **Completion**: FLOST

Outline

- 1 Motivating Science Topic: Space Weather Monitoring
- 2 **Imputation**: Spatio-Temporal Matrix Completion Method (**VISTA**)
- 3 Data Product: Constructing A Space Weather Database with VISTA
- 4 **Completion**: FLOST

Motivating Science Topic: Space Weather Monitoring

Weather within our atmosphere: rain, snow, heat, and wind. Space weather: radio blackouts, solar radiation storms, and geomagnetic storms caused by disturbances from the Sun.

Space Weather

Space weather refers to the variable conditions on the Sun and in the space environment that can influence the performance and reliability of space-based and ground-based technological systems, as well as endanger life on health. Just like weather on Earth, space weather has 85 seasons, with solar activity rising and falling over an approximate 11 year cycle.

Sunspots
Temporary dark spots on the Sun's surface, caused by intense magnetic activity.

Coronal Mass Ejections (CMEs)
Large portions of the corona, or outer atmosphere of the Sun, can be ejected from it into space, sending billions of tons of plasma and superheated gas. Earth's magnetic field can deflect some of this material, but some can reach Earth's magnetic field, resulting in geomagnetic storms. But Ejecta of these CMEs can reach Earth in under a day with the devastating force 3 days in transit.

Solar Wind
The solar wind is a constant outflow of electron and proton from the Sun, always present and building Earth's magnetic field. The density of solar wind from the Sun is approximately one million miles per hour!

Solar Flares
Disruption of the magnetic field on the surface of the Sun allows the biggest explosions in our solar system. These solar flares release tremendous amounts of energy and result in electromagnetic emissions spanning the spectrum from gamma rays to radio waves. Traveling at the speed of light, these emissions reach the Earth within eight to ten minutes.

Earth's Magnetic Field
Earth's magnetic field, largely that of a bar magnet, gives the Earth's outer protection from the effects of the Sun. Earth's magnetic field is constantly in contact with the solar wind and stretched on the right side by the ever-present solar wind. During geomagnetic storms, the disturbance to Earth's magnetic field can become intense. In addition to some being lost by the atmosphere, the field disturbance may be due to the charged particles of a radiation storm.

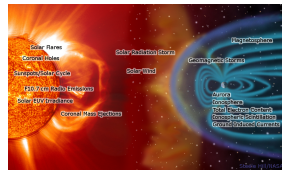
Sun's Magnetic Field
Strong and ever-changing magnetic fields drive the life of the Sun and accelerate its storms. These strong magnetic fields are the energy source for space weather and their heating, blowing, and reconnection lead to solar flares.

Solar Radiation Storms
Charged particles, including electrons and protons, can be accelerated by coronal mass ejections and solar flares. These particles become and space their way through space, mostly following the magnetic field lines and ultimately being blocked by Earth's ionosphere. The danger of these particles can affect Earth's ionosphere after a solar flare.

Geomagnetic Storms
A geomagnetic storm is a temporary disturbance of Earth's magnetic field typically associated with fluctuations in the solar wind. These storms are caused when the solar wind and its magnetic field interacts with Earth's magnetic field. The primary source of geomagnetic storms is CMEs which stretch the magnetosphere on the nightside causing it to release energy through magnetic reconnection. Disturbances in the ionosphere in regions of Earth upper atmosphere are usually associated with geomagnetic storms.

Effects on Earth: Auroral displays, Radio blackouts, GPS, Satellite navigation, Power grid disturbances.

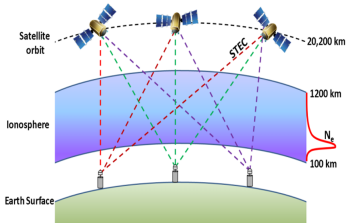
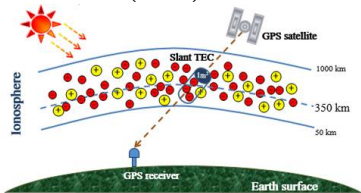
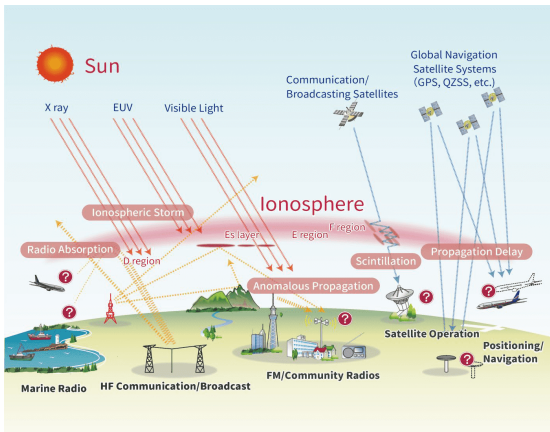
Source: NASA, NOAA



<p>GEOPHYSICAL MONITORING</p> <p>The science program and instruments located in space and ground-based systems for monitoring space weather.</p>	<p>IN-ORBIT MONITORING</p> <p>Space-based systems with communication and navigation capabilities.</p>	<p>AIR TRAFFIC MANAGEMENT</p> <p>Global navigation satellite systems (GNSS) and communication satellite systems for air traffic management.</p>
<p>POWER GRID</p> <p>Strong geomagnetic storms can affect power grids through the ionosphere, which can cause power outages and blackouts.</p>	<p>SPACE WEATHER AND GPS SYSTEMS</p> <p>The ionosphere and other layers of the atmosphere can affect GPS signals, which are used for navigation and timing.</p>	<p>SPACE WEATHER AND POWER</p> <p>Geomagnetic storms can cause power outages and blackouts, which can affect critical infrastructure.</p>

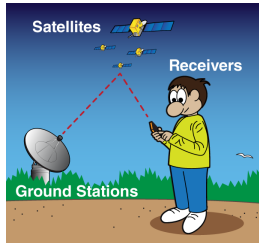
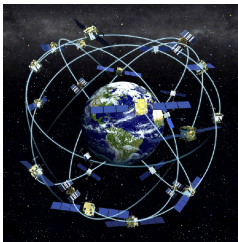
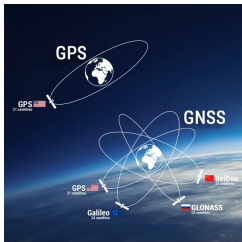
Motivating Case in Space Weather: The TEC

Earth's Ionosphere: part of the upper atmosphere ionized by extreme ultraviolet (EUV) and x-ray solar radiation. Figures: Ionospheric effects on radio propagation and measurement of the total electron content (TEC).



Importance of Monitoring Total electron content (TEC)

- The change in the path and velocity of radio waves in the ionosphere greatly impacts the accuracy of satellite navigation systems such as GPS/GNSS.
- Neglecting changes in the ionospheric TEC can introduce tens of meters of error in the position calculations. The more electrons in the path of the radio wave, the more the radio signal will be affected.
- For ground-to-satellite communication and satellite navigation, TEC is a good parameter to monitor for possible space weather impacts.



Motivating Case in Space Weather: The TEC Data

Global **total electron content** (TEC) time series: TEC is an essential quantity for space weather monitoring but is sparsely observed, especially in the oceanic regions due to the lack of ground-based receivers.

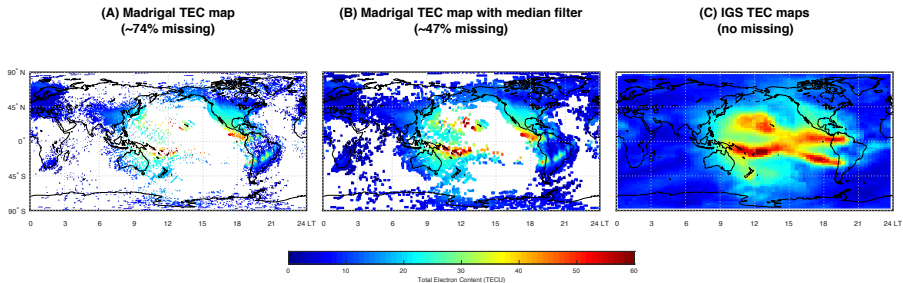


Figure 1.1: (A) Global Total Electron Content (TEC) Map, white pixels are missing data; (B) Median-Filtered TEC Map. (C) The popularly used imputation of the TEC in the geoscience community (spherical Harmonics) is typically overly smoothed.

Statistical Problems with TEC Data

Spatio-temporal tensor data expressed as a 3-mode tensor data $\mathcal{X} \in \mathbb{R}^{m \times n \times T}$:

- \mathcal{X} consists of a time series of spatial data. Let $\mathcal{X}_{::t} := \mathbf{X}_t \in \mathbb{R}^{m \times n}$.
- Each \mathbf{X}_t is spatial data collected on an $m \times n$ grid, and is only partially observed. For \mathbf{X}_t , binary matrix $\mathbf{\Omega}_t$ denotes the missingness:

$$\mathbf{\Omega}_t(i, j) = \begin{cases} 1, & \text{if } \mathbf{X}_t(i, j) \text{ is observed,} \\ 0, & \text{if } \mathbf{X}_t(i, j) \text{ is missing.} \end{cases}$$

- Statistical questions:
 - 1 **Video Completion**: given the observations in $\mathbf{X}_1, \dots, \mathbf{X}_T$ (indexed by $\mathbf{\Omega}_1, \dots, \mathbf{\Omega}_T$), can we impute the missing values?
 - 2 **Map Prediction**: given the observations in $\mathbf{X}_1, \dots, \mathbf{X}_{t-1}$, and other (global) vector time series predictors, can we predict \mathbf{X}_t ?
 - 3 Proper uncertainty quantification for the two problems above.

Outline

- 1 Motivating Science Topic: Space Weather Monitoring
- 2 **Imputation**: Spatio-Temporal Matrix Completion Method (**VISTA**)
- 3 Data Product: Constructing A Space Weather Database with VISTA
- 4 **Completion**: FLOST

Imputation Problem Formulated as Matrix Completion

The Matrix Completion Problem

- Partially observed $m \times n$ matrix: \mathbf{X} .
- Binary indicator of observed entries: Ω .
- Projection operator: $P_{\Omega}(\cdot)$, where

$$P_{\Omega}(\mathbf{X}) = \begin{cases} \mathbf{X}(i, j) & \text{if } \Omega(i, j) = 1 \\ 0 & \text{if } \Omega(i, j) = 0 \end{cases}$$

The matrix completion problem seeks a matrix $\hat{\mathbf{X}} \in \mathbb{R}^{m \times n}$ such that:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{M} \in \mathbb{R}^{m \times n}} \underbrace{\|P_{\Omega}(\mathbf{X} - \mathbf{M})\|_{\mathbb{F}}^2}_{\text{goodness-of-fit on the observed entries}} + \underbrace{\lambda g(\mathbf{M})}_{\text{penalty of matrix complexity}}.$$

Matrix Completion for Spatial Data: Does it Work?

The soft-Impute algorithm in [Hastie et al., 2015] gives:

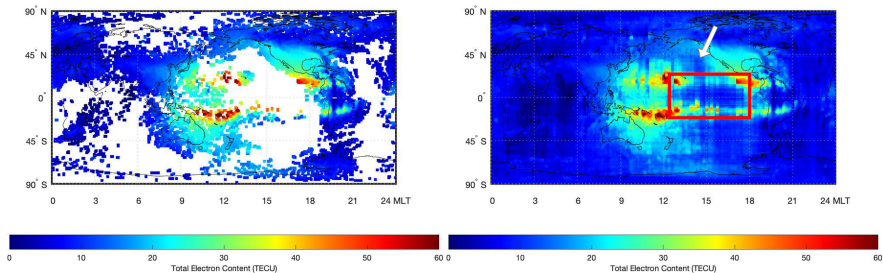


Figure 2.1: Left: Source matrix. Right: Imputed matrix based on Soft-Impute.

The result is *problematic scientifically* because the completed maps are:

- not spatially smooth (the blue bands); and
- not informative when missingness are spatially clustered (the red box).

Alternative: Kernel Smoothing for Matrix Completion

Recall the matrix completion problem:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{M} \in \mathbb{R}^{m \times n}} \|\mathbf{P}_{\Omega}(\mathbf{X} - \mathbf{M})\|_{\mathbb{F}}^2 + \lambda g(\mathbf{M}).$$

In a kernel smoothing method, one assumes that \mathbf{M} is a discrete evaluation of an unknown function f on the matrix grid. Specifically,

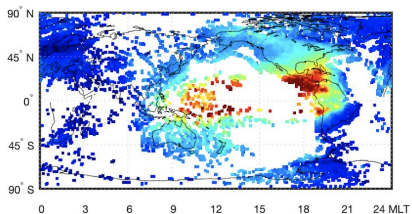
$$\mathbf{M}(i, j) = f(s_{ij})$$

where $s_{ij} \in \mathbb{R}^2$ is the spatial location of the (i, j) -th entry. And $g(\mathbf{M})$ is a penalty on the functional parameter g , e.g., the functional norm in a Hilbert space.

Widely adopted kernel in space weather: spherical Harmonics.

Alternative: Kernel Smoothing for Matrix Completion

(A) Original Map



(B) SH fitting Map

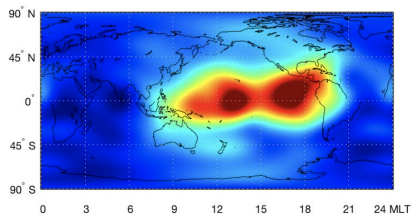


Figure 2.2: Example of Spherical Harmonics Kernel Smoothing

The result from Spherical Harmonics fitting is:

- poorly fitted over the observed entries; and
- overly smooth locally but **capture the global trend**.

Proposed Method: The VISTA Approach

We aim to develop a method that:

- simultaneously imputes all $\mathbf{X}_1, \dots, \mathbf{X}_T$;
- fits well on the scattered missing entries in $\mathbf{X}_1, \dots, \mathbf{X}_T$, and
- fits well over the large-patch missingness.

A novel tensor completion approach, named VISTA, bridges

- Soft-Impute ALS: low-rank matrix completion method, and
- Spherical Harmonics: kernel smoothing method, while
- accounting for temporal continuity & temporally varying missingness.

Proposed Method: The VISTA Approach

VISTA Video Imputation Method

We impute $\mathcal{X} \in \mathbb{R}^{m \times n \times T}$ by finding the minimizer of:

$$\min_{\mathbf{A}_t \in \mathbb{R}^{m \times r}, \mathbf{B}_t \in \mathbb{R}^{n \times r}, \forall t \in [T]} \frac{1}{2} \sum_{t=1}^T \|\mathbf{P}_{\Omega_t} (\mathbf{X}_t - \mathbf{A}_t \mathbf{B}_t^\top)\|_{\mathbb{F}}^2 + g_\lambda(\mathbf{A}_1, \dots, \mathbf{A}_T, \mathbf{B}_1, \dots, \mathbf{B}_T)$$

where $g_\lambda(\cdot)$ has the following three terms:

- $\frac{\lambda_1}{2} \sum_{t=1}^T (\|\mathbf{A}_t\|_{\mathbb{F}}^2 + \|\mathbf{B}_t\|_{\mathbb{F}}^2)$: the **S**oft-impute penalty;
- $\frac{\lambda_2}{2} \sum_{t=2}^T \|\mathbf{A}_t \mathbf{B}_t^\top - \mathbf{A}_{t-1} \mathbf{B}_{t-1}^\top\|_{\mathbb{F}}^2$: the **T**emporal smoothness penalty;
- $\frac{\lambda_3}{2} \|\mathbf{Y}_t - \mathbf{A}_t \mathbf{B}_t^\top\|_{\mathbb{F}}^2$: the **A**uxiliary algorithm (data) penalty.

We coined the term VISTA since the method is a **V**ideo **I**mputation method with **S**oft-impute, **T**emporal-smoothing and **A**uxiliary data.

Simulation Experiment - Design

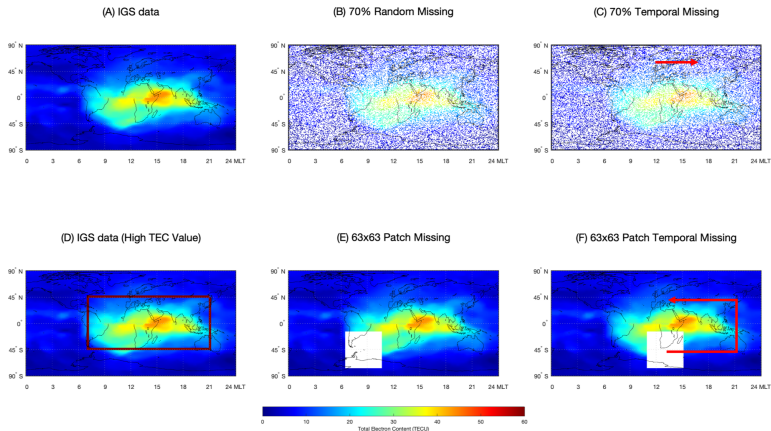


Figure 2.3: (A) Ground Truth; (B) **random+non-temporal+70%**; (C) **random+temporal+70%**; (E) **patch+non-temporal+63 × 63**; (F) **patch+temporal+63 × 63**.

Simulation Experiment - Result

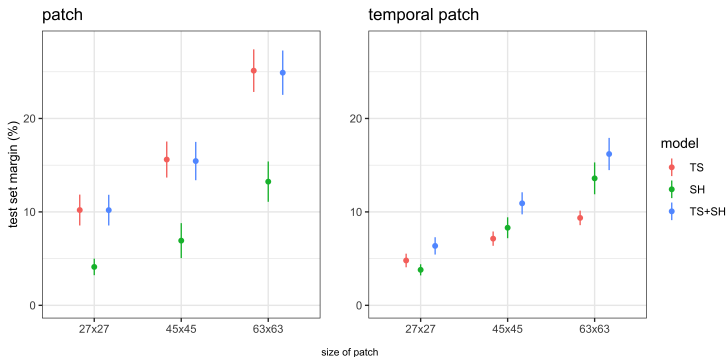


Figure 2.5: **Patch**+non-temporal missing (left) and **Patch**+temporal patch missing (right) results. The scatter points show the average test set RSE margin over the baseline softImpute method, **positive means performance better than softImpute**. Error bar gives the 95% confidence interval.

Real Data Implementation - Result

Applying the VISTA method to the high-resolution global TEC data:

Storm Day				
Model	test RSE	test MSE	# matrices better than softImpute	# matrices worse than Full model
softImpute ($\lambda_1 = 0.9$)	10.895%	2.675	/	285 (98.96%)
TS ($\lambda_1 = 0.9, \lambda_2 = 0.2$)	9.643%	2.106	284 (98.62%)	267 (92.71%)
SH ($\lambda_1 = 0.9, \lambda_3 = 0.021$)	9.936%	2.227	287 (99.65%)	274 (95.14%)
Full ($\lambda_1 = 0.9, \lambda_2 = 0.2, \lambda_3 = 0.021$)	9.357%	1.983	285 (98.96%)	/
Directly use Spherical Harmonics	17.354%	6.720	0 (0%)	288 (100%)
Non-Storm Day				
Model	test RSE	test MSE	# matrices better than softImpute	# matrices worse than Full model
softImpute ($\lambda_1 = 0.9$)	10.424%	1.324	/	283 (98.26%)
TS ($\lambda_1 = 0.9, \lambda_2 = 0.31$)	8.880%	0.958	281 (97.57%)	235 (81.60%)
SH ($\lambda_1 = 0.9, \lambda_3 = 0.03$)	9.231%	1.032	287 (99.65%)	278 (96.53%)
Full ($\lambda_1 = 0.9, \lambda_2 = 0.31, \lambda_3 = 0.03$)	8.592%	0.895	283 (98.26%)	/
Directly use Spherical Harmonics	15.732%	2.893	0 (0%)	288 (100%)

Table 1: Empirical study results from the madrigal database.

Figure 2.6: Storm day is Sept-08, 2017 and non-storm day is Sept-03, 2017.

Summary

In this project, we:


- propose a spatiotemporal video completion method **VISTA** by modifying classical matrix completion method with:
 - ① temporal variation regularization;
 - ② regularization based on an auxiliary data/algorithm (kernel smoothing)
- establish the convergence guarantee of the algorithm;
- validate the method with extensive simulation & real data experiments.

Our paper: Sun, H., Hua, Z., Ren, J., Zou, S., Sun, Y., & **Chen, Y.** (2022). Matrix Completion Methods for the Total Electron Content Video Reconstruction. *The Annals of Applied Statistics*, 16(3), 1333-1358.

Outline

- 1 Motivating Science Topic: Space Weather Monitoring
- 2 **Imputation**: Spatio-Temporal Matrix Completion Method (**VISTA**)
- 3 Data Product: Constructing A Space Weather Database with VISTA
- 4 **Completion**: FLOST

The VISTA TEC Database



The banner features two global maps of Total Electron Content (TEC) on the left and right, showing high concentrations (red/yellow) over the equatorial region. In the center, the text 'VISTA TEC' is displayed in a large, bold, black font, with a small globe icon above the letter 'I'.

Complete Global Total Electron Content Database based on the VISTA Algorithm [Open Access](#)

Complete Global Total Electron Content Database based on the VISTA Algorithm

Collection Details	
Total items	2
Resource type	Collection
Creator	Sun, Hu
Keyword	Total Electron Content Matrix Completion VISTA Spherical Harmonics Spatial-Temporal Smoothing
Language	English
Last Updated	2023-03-26

Figure 3.1: The VISTA TEC Open Source Database [[link](#)].

The VISTA TEC Workflow

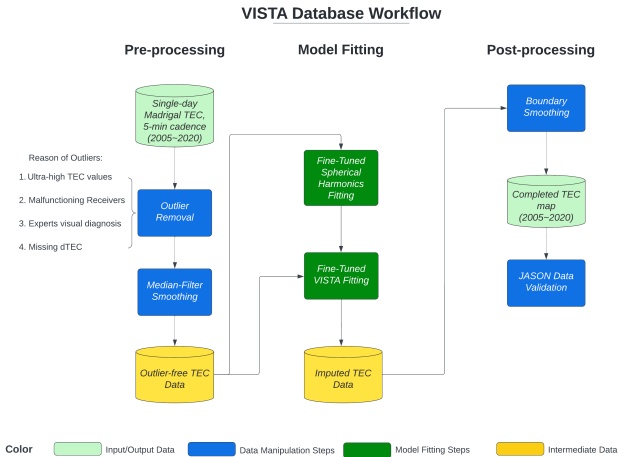


Figure 3.2: The Database Generating Workflow. Image source: Sun et al. [2023].

Sample of the VISTA TEC Database

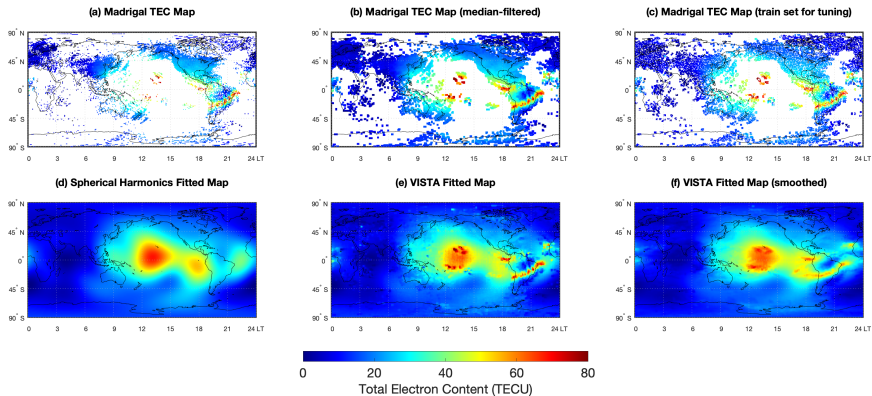


Figure 3.3: TEC maps throughout the workflow. Image source: [Sun et al. \[2023\]](#).

Summary

In this project, we:

- create a 16-year spatiotemporal database for the global TEC with VISTA;
- achieve higher imputation accuracy and higher spatiotemporal resolution as compared to previous open source databases;
- validate the imputed entries in the VISTA TEC maps with an independent data source.

Our paper: Sun, H., **Chen, Y.**, Zou, S., Ren, J., Chang, Y., Wang, Z., & Coster, A. (2023). Complete Global Total Electron Content Map Dataset based on a Video Imputation Algorithm VISTA. *Scientific Data*, 10(1), 236.

Outline

- 1 Motivating Science Topic: Space Weather Monitoring
- 2 **Imputation**: Spatio-Temporal Matrix Completion Method (**VISTA**)
- 3 Data Product: Constructing A Space Weather Database with VISTA
- 4 **Completion**: FLOST

Motivation

- Restating the TEC Completion problem as a tensor completion problem:
 $\mathcal{T} = [\mathbf{X}_1, \dots, \mathbf{X}_T] \in \mathbb{R}^{M \times N \times T}$ with missingness given by
 $\Omega = \{(i, j, k), \omega_{ijk} = 1\}$, where $\omega_{ijk} \sim \text{Bernoulli}(p)$.
- Existing approach: low-rank structures given by CP [Jain and Oh \[2014\]](#); [Cai et al. \[2022a\]](#), Tucker [Xia and Yuan \[2019\]](#); [Xia et al. \[2021\]](#); [Li et al. \[2023\]](#), Tensor-Train (TT) [Cai et al. \[2022c,b\]](#), and tubal-rank [Zhang and Aeron \[2016\]](#); [Liu et al. \[2019\]](#); [Jiang and Ng \[2019\]](#); [Song et al. \[2023\]](#).
- Unique challenge:
 - stable/slow-varying TEC for many regions
 - rapid fluctuations for certain regions (geomagnetic storm)

Definition

Definition (Fourier Low-rank and Sparse Tensor)

We say a tensor $\mathcal{T} \in \mathbb{R}^{M \times N \times T}$ is (\mathbf{r}, K, s) -FLOST for $\mathbf{r} = (r_1, \dots, r_K)$, if

$$\text{rank}(\mathcal{T} \times_3 \mathbf{f}_l^\top) \leq r_l, \quad l = 1, \dots, K, \quad \|\mathcal{T} \times_3 \mathbf{F}_1^\top\|_0 \leq s,$$

where $\mathbf{F}_1 = [\mathbf{f}_{K+1}, \dots, \mathbf{f}_{\lceil \frac{T+1}{2} \rceil}]$ and $K \leq \lceil \frac{T+1}{2} \rceil$.

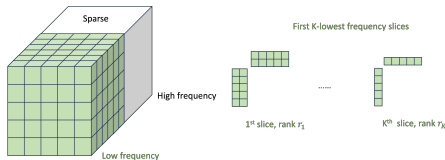


Figure 4.1: Structure of FLOST in the frequency domain.

Here, $\mathcal{T} \times_3 \mathbf{f}_l^\top$ extracts the l -th frontal slice of the tensor in the frequency domain, and $\mathcal{T} \times_3 \mathbf{F}_1^\top$ corresponds to the collection of high-frequency slices.

Model and Loss

Let $\mathcal{T}_0 \in \mathbb{R}^{M \times N \times T}$ be a tensor that is (\mathbf{r}, K, s) -FLOST. We can obtain a collection of partially observed and noisy entries where the sampling indicator $\omega_{ijt} = 1$:

$$\mathcal{Y}_{ijt} = [\mathcal{T}_0]_{ijt} + \mathcal{E}_{ijt}.$$

Here \mathcal{E} represents the random noise independent of $\{\omega_{ijt}\}_{ijt}$. We propose:

$$\hat{\mathcal{T}} := \arg \min_{\mathcal{T} \in \mathbb{R}^{M \times N \times T}} \frac{1}{2} \|\mathcal{T} - \rho^{-1} \mathcal{P}_{\Omega}(\mathcal{Y})\|_{\mathbb{F}}^2 + \sum_{l=1}^K \lambda_{1,l} \|\mathcal{T} \times_3 \mathbf{f}_l^{\top}\|_* + \lambda_2 \|\mathcal{T} \times_3 \mathbf{F}_1^{\top}\|_{\ell_1}. \quad (1)$$

Here the nuclear norm and ℓ_1 norm are the regularizations for the low-rank and sparse structures, respectively.

Computation

Parallel Implementation:

Since $\|\mathcal{T} - p^{-1}\mathcal{P}_\Omega(\mathcal{Y})\|_{\mathbb{F}}^2 = \sum_{t=1}^T \|(\mathcal{T} - p^{-1}\mathcal{P}_\Omega(\mathcal{Y})) \times_3 \mathbf{f}_t^\top\|_{\mathbb{F}}^2$, the optimization problem in (1) can be equivalently written as $K + 1$ independent problems in the Fourier domain:

$$\hat{\mathbf{T}}_l = \arg \min_{\mathbf{T}} \frac{1}{2} \|\mathbf{T} - p^{-1}\mathcal{P}_\Omega(\mathcal{Y}) \times_3 \mathbf{f}_l^\top\|_{\mathbb{F}}^2 + \lambda_{1,l} \|\mathbf{T}_l\|_*, \quad l = 1, \dots, K \quad (2a)$$

$$\hat{\mathcal{T}}_1 = \arg \min_{\mathcal{T}} \frac{1}{2} \|\mathcal{T} - p^{-1}\mathcal{P}_\Omega(\mathcal{Y}) \times_3 \mathbf{F}_1^\top\|_{\mathbb{F}}^2 + \lambda_2 \|\mathcal{T}\|_{\ell_1}. \quad (2b)$$

Once we obtain $\hat{\mathbf{T}}_l, \hat{\mathcal{T}}_1$, we first concatenate these matrices/tensor, then apply the inverse Fourier transform to recover the estimator for \mathcal{T}_0 .

Theoretical Guarantee

Corollary

Assume the noise \mathcal{E}_{ijt} are independently distributed centered sub-Gaussian random variables with proxy variance σ^2 , and $\max_{ijt} |[\mathcal{T}_0]_{ijt}| \leq \gamma$. Choose

$$\lambda_{1,l} \geq C_1(\sigma \vee \gamma) \left(\frac{\sqrt{(N \vee M) \log(M \vee N)}}{\sqrt{p}} + \frac{\sqrt{\log^3(M \vee N)}}{p\sqrt{T}} \right),$$

$$\lambda_2 \geq C_2(\sigma \vee \gamma) \left(\frac{\sqrt{\log(M \vee N \vee T)}}{\sqrt{p}} + \frac{\log(M \vee N \vee T)}{\sqrt{T}p} \right).$$

Assume $n \geq \max \left\{ (M \wedge N), \min \left\{ (M \wedge N) \frac{s}{\sum_l r_l}, MN \right\} \right\}$, then

$$\frac{1}{MNT} \|\hat{\mathcal{T}} - \mathcal{T}_0\|_{\text{F}}^2 \leq \tilde{C} \frac{\sum_{l=1}^K (M \vee N) r_l + s}{n} (\sigma \vee \gamma)^2$$

holds with probability exceeding $1 - 2(M \vee N \vee T)^{-10}$.

Numerical Performance

Method	$T = 100$		$T = 500$		$T = 1000$	
	Test RMSE	Time	Test RMSE	Time	Test RMSE	Time
FLoST	0.5183	0.06	0.5216	0.31	0.5198	0.69
	± 0.0123	± 0.00	± 0.0132	± 0.02	± 0.0147	± 0.05
Ablation	0.5318	0.16	0.5298	0.89	0.5292	2.03
	± 0.0004	± 0.01	± 0.0002	± 0.08	± 0.0001	± 0.16
RCGD	0.5785	0.47	0.5753	2.81	0.5743	6.59
	± 0.0005	± 0.01	± 0.0002	± 0.18	± 0.0002	± 0.47
Alt-Min-LS	0.5423	8.88	0.5406	47.76	0.5399	95.24
	± 0.0006	± 0.10	± 0.0003	± 2.85	± 0.0002	± 4.17
LRTC-TNN	0.5591	9.91	0.5543	53.14	0.5522	121.33
	± 0.0005	± 0.52	± 0.0002	± 4.84	± 0.0002	± 10.52
PRGD	0.5496	6.18	0.5475	33.15	0.5469	67.37
	± 0.0004	± 0.36	± 0.0002	± 1.65	± 0.0001	± 4.83

Table 1: Ablation: FLoST with no sparsity. LRTC-TNN [Lu et al., 2019], Alt-Min-LS [Liu et al., 2019] PRGD [Zhang et al., 2025].

Conclusions

- VISTA TEC Database construction [[Sun et al., 2022, 2023](#)].
- FLoST Tensor Completion Algorithm [[Li et al., 2025](#)].
- Tensor completion with uncertainty quantification (UQ):
 - Conformalized tensor completion [[Sun and Chen, 2024](#)];
 - UQ for low-tubal rank tensor completion [[Shang et al., 2026](#)].
- Video/tensor prediction [[Wu et al., 2026](#); [Sun et al., 2024](#)].
- Check out our newly developed spaceweatherlab.org

References I

- C. Cai, G. Li, H. V. Poor, and Y. Chen. Nonconvex low-rank tensor completion from noisy data. *Operations Research*, 70(2):1219–1237, 2022a.
- J.-F. Cai, W. Huang, H. Wang, and K. Wei. Tensor completion via tensor train based low-rank quotient geometry under a preconditioned metric. *arXiv preprint arXiv:2209.04786*, 2022b.
- J.-F. Cai, J. Li, and D. Xia. Provable tensor-train format tensor completion by riemannian optimization. *Journal of Machine Learning Research*, 23(123): 1–77, 2022c.
- T. Hastie, R. Mazumder, J. D. Lee, and R. Zadeh. Matrix completion and low-rank svd via fast alternating least squares. *The Journal of Machine Learning Research*, 16(1):3367–3402, 2015.
- P. Jain and S. Oh. Provable tensor factorization with missing data. *Advances in Neural Information Processing Systems*, 27, 2014.
- Q. Jiang and M. Ng. Robust low-tubal-rank tensor completion via convex optimization. In *IJCAI*, pages 2649–2655, 2019.

References II

- J. Li, J.-F. Cai, Y. Chen, and D. Xia. Online tensor learning: Computational and statistical trade-offs, adaptivity and optimal regret. *arXiv preprint arXiv:2306.03372*, 2023.
- J. Li, J. Shang, and Y. Chen. Fourier low-rank and sparse tensor for efficient tensor completion. *arXiv preprint arXiv:2505.11261*, 2025.
- X.-Y. Liu, S. Aeron, V. Aggarwal, and X. Wang. Low-tubal-rank tensor completion using alternating minimization. *IEEE Transactions on Information Theory*, 66(3):1714–1737, 2019.
- C. Lu, X. Peng, and Y. Wei. Low-rank tensor completion with a new tensor nuclear norm induced by invertible linear transforms. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5996–6004, 2019.
- J. Shang, J. Li, and Y. Chen. Uncertainty quantification for noisy low-tubal-rank tensor completion. *arXiv preprint arXiv:2604.10353*, 2026.

References III

- G.-J. Song, X.-Z. Wang, and M. K. Ng. Riemannian conjugate gradient descent method for fixed multi rank third-order tensor completion. *Journal of Computational and Applied Mathematics*, 421:114866, 2023.
- H. Sun and Y. Chen. Conformalized tensor completion with riemannian optimization. *arXiv preprint arXiv:2405.00581*, 2024.
- H. Sun, Z. Hua, J. Ren, S. Zou, Y. Sun, and Y. Chen. Matrix completion methods for the total electron content video reconstruction. *The Annals of Applied Statistics*, 16(3):1333–1358, 2022.
- H. Sun, Y. Chen, S. Zou, J. Ren, Y. Chang, Z. Wang, and A. Coster. Complete global total electron content map dataset based on a video imputation algorithm vista. *Scientific Data*, 10(1):236, 2023.
- H. Sun, Z. Shang, and Y. Chen. Matrix autoregressive model with vector time series covariates for spatio-temporal data. *arXiv preprint arXiv:2305.15671*, 2024.

References IV

- D. Wu, K. Jin, and Y. Chen. Tensor tree regression with additional boosting for tensor prediction with auxiliary covariates, 2026.
- D. Xia and M. Yuan. On polynomial time methods for exact low-rank tensor completion. *Foundations of Computational Mathematics*, 19(6):1265–1313, 2019.
- D. Xia, M. Yuan, and C.-H. Zhang. Statistically optimal and computationally efficient low rank tensor completion from noisy entries. *The Annals of Statistics*, 49(1), 2021.
- Y. Zhang, F. Bian, X. Zhang, and J.-F. Cai. Preconditioned riemannian gradient descent algorithm for low-multilinear-rank tensor completion. In *Forty-second International Conference on Machine Learning*, 2025.
- Z. Zhang and S. Aeron. Exact tensor completion using t-svd. *IEEE Transactions on Signal Processing*, 65(6):1511–1526, 2016.