# Spatio-Temporal Tensor Completion with Auxiliary Information: Application to TEC Video Reconstruction

Hu Sun [1] Zhijun Hua [3] Jiaen Ren [2] Shasha Zou [2] Yuekai Sun [1] Yang Chen [1]

[1]Dept. of Statistics, U-M [2]Dept. of Climate and Space Sciences and Engineering, U-M [3]SC Johnson College of Business, Cornell University

## Introduction

**Problem Setting**: We investigate a special type of tensor completion problem where the data tensor $\mathcal{X} \in \mathbb{R}^{m\times n\times T}$ is a matrix time-series $\{X_t\}_{t=1}^T$, where $X_t \in \mathbb{R}^{m\times n}$ and is sparsely observed on a 2D spatial grid. We aim at "filling in" the missing values of $\mathcal{X}$ with *high accuracy*.

**Motivating Application**: Our motivating dataset is the global total electron content (TEC) time-series, where 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:

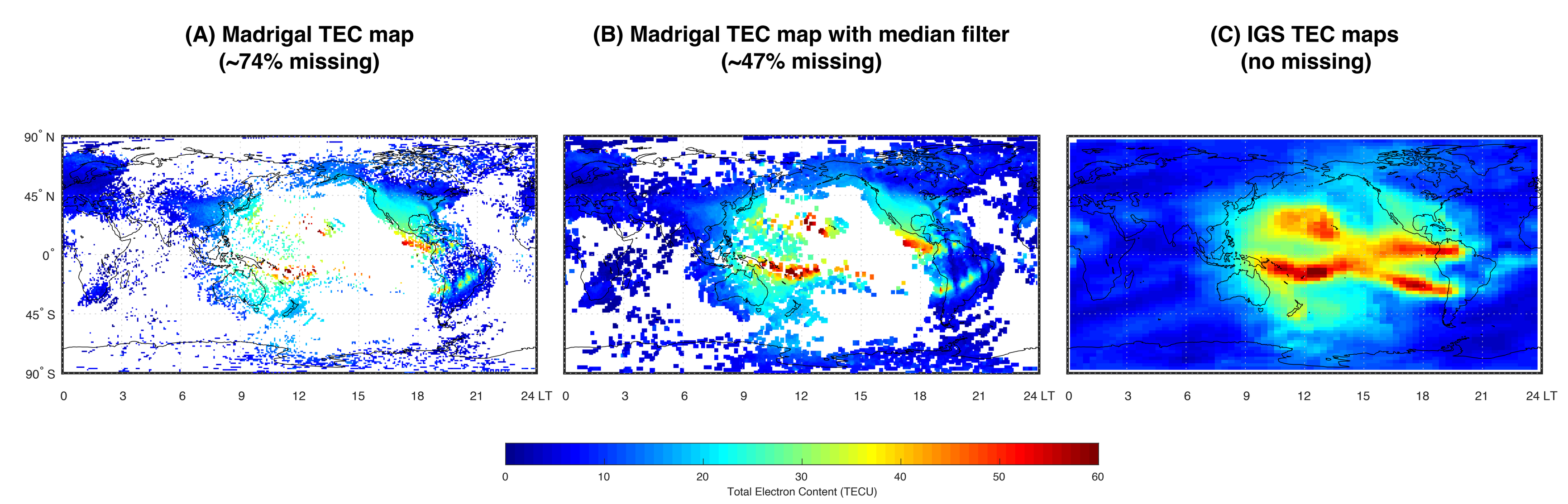


Figure 1. (A) Global Total Electron Content (TEC) Map; (B) Median-Filtered TEC Map. (C) The popularly-used imputation of the TEC in the geoscience community, which is typically overly smoothed.

The popular method of imputing the TEC map (Fig.1 (C)) is the Spherical Harmonics (SH) algorithm, which estimates each matrix $X_t$ as a linear combination of a series of basis functions:

$$X_t(i,j) \approx \sum_{l=0}^{L}\sum_{v=-l}^{l} \widehat{a_{lv}} \times Y_l^v(\theta_i,\phi_j) \quad (1)$$

where $(\theta_i,\phi_j)$ is the spatial coordinate of the $(i,j)$-th entry. This typically *overly* smoothed.

## Classic Matrix Completion Method: Will it work?

We consider the Soft-Impute (JMLR, 2015) formulation to approximate $X_t$ as $X_t \approx A_tB_t^\top$ by:

$$\min_{A_t,B_t}\left\{F(A_t,B_t) := \underbrace{\frac{1}{2}\|P_{\Omega_t}(X_t - A_tB_t^\top)\|_F^2}_{\text{Reconstruction Error}} + \overbrace{\frac{\lambda_1}{2}(\|A_t\|_F^2+\|B_t\|_F^2)}^{\ell_2\text{-regularization}}\right\}, \quad (2)$$

where $\Omega_t$ is a binary indicator matrix with $\Omega_t(i,j) = \mathrm{I}_{\{X_t(i,j)\neq \text{NaN}\}}$, and $P_{\Omega_t}(M) = M \odot \mathrm{I}_{\{\Omega_t=1\}}$, for any arbitrary matrix $M \in \mathbb{R}^{m\times n}$ and $\odot$ is element-wise multiplication.

The algorithm yields *under*-smoothed result, especially when data are missing in big patches:

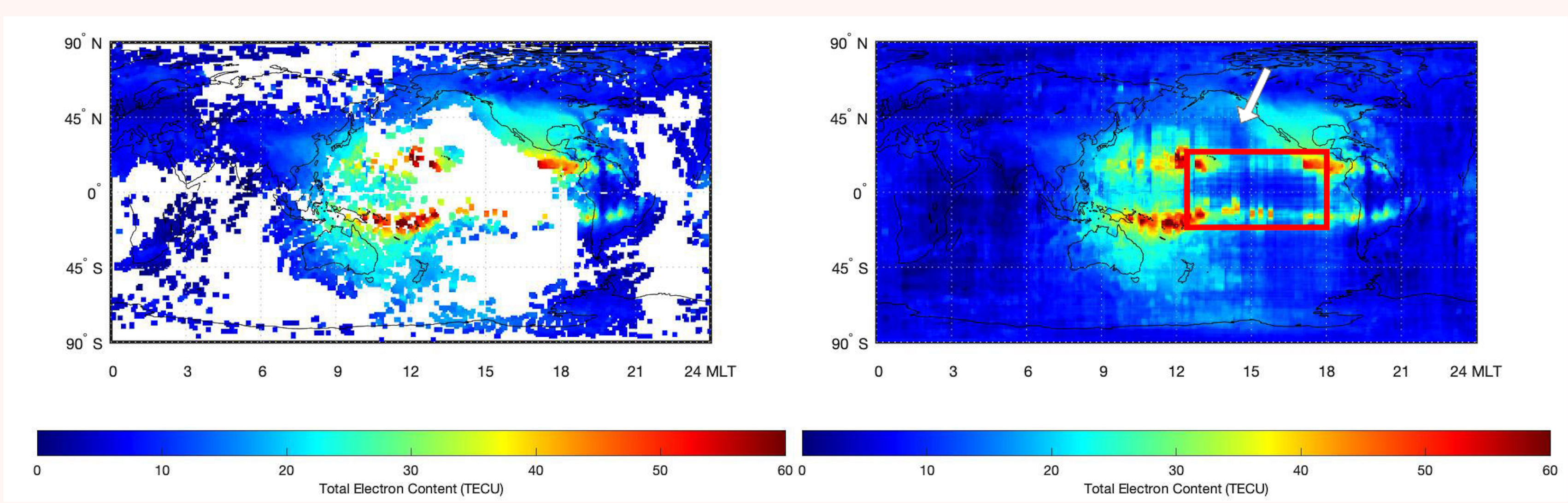

Figure 2. Left: Source matrix. Right: Imputed matrix based on Soft-Impute.

## Balancing Over-smooth & Under-smooth: The VISTA Approach

The previous methods have multiple issues:

- Soft-Impute (as well as other matrix completion method) fails at imputing the regions with high missingness (e.g. oceanic regions).
- Spherical Harmonics (SH) yields an overly-smoothed map with unwanted physical structure.
- Both methods do not account for the temporal dimension, but fit the model *in parallel*.

We propose our **V**ideo **I**mputation with **S**oft-Impute, **T**emporal Smoothing and **A**uxiliary Data (**VISTA**) approach for imputing spatio-temporal tensor data as:

$$\min_{A_{1:T},B_{1:T}}\Big\{F(A_{1:T},B_{1:T}) := \overbrace{\frac{1}{2}\sum_{t=1}^T\|P_{\Omega_t}(X_t - A_tB_t^\top)\|_F^2}^{\text{Reconstruction Error}} + \overbrace{\frac{\lambda_1}{2}\sum_{t=1}^T(\|A_t\|_F^2+\|B_t\|_F^2)}^{\ell_2\text{-regularization}} + \underbrace{\frac{\lambda_2}{2}\sum_{t=2}^T\|A_tB_t^\top - A_{t-1}B_{t-1}^\top\|_F^2}_{\text{Temporal Consistency}} + \underbrace{\frac{\lambda_3}{2}\sum_{t=1}^T\|Y_t - A_tB_t^\top\|_F^2}_{\text{Auxiliary regularization}}\Big\}, \quad (3)$$

where $Y_{1:T}$ are $m\times n$ auxiliary data with no missing values, and in TEC imputation, we use the output of the SH algorithm as $Y_{1:T}$. Essentially, **we use the output of an overly-smoothed algorithm (SH) to regularize an under-smoothed algorithm (matrix completion)**.

## Estimating Algorithm: Alternating Majorization-Minimization

We solve the non-convex optimization problem in (3) by cyclically updating the parameters, *one at a time*, in the order of: $A_1 \to A_2 \to \cdots \to A_T \to B_1 \to B_2 \to \cdots \to B_T \to A_1 \to A_2 \to \ldots$. In iteration $k+1$, when updating $A_t$:

- **Step 1: Majorization**

$$\|P_{\Omega_t}(X_t - A_t(B_t^{(k)})^\top)\|_F^2 \leq \|\overbrace{P_{\Omega_t}(X_t) + P_{\Omega_t^\perp}(A_t^{(k)}(B_t^{(k)})^\top)}^{\text{A "filled-in" proposal of } X_t} - A_t(B_t^{(k)})^\top\|_F^2. \quad (4)$$

using the majorization bound above in (3) yields a convex relaxation of (3), denoted as $\widetilde{Q}$.

- **Step 2: Minimization**

$$A_t^{(k+1)} = \arg\min \widetilde{Q}(A_t|A_{1:t-1}^{(k+1)}, A_{t:T}^{(k)}, B_{1:T}^{(k)}) \quad (5)$$

This leads to an exact minimization solution of $A_t^{(k+1)}$ as:

$$A_t^{(k+1)} = \left[(1+\lambda_2(\mathbf{I}_{\{t<T\}}+\mathbf{I}_{\{t>1\}})+\lambda_3)(B_t^{(k)})^\top B_t^{(k)} + \lambda_1 \mathrm{I}\right]^{-1} Z_t^{(k)} B_t^{(k)}. \quad (6)$$

where $Z_t^{(k)}$:

$$Z_t^{(k)} = \underbrace{P_{\Omega_t}(X_t) + P_{\Omega_t^\perp}(A_t^{(k)}(B_t^{(k)})^\top)}_{\text{A "filled-in" proposal of } X_t} + \lambda_2\underbrace{\left(\mathbf{I}_{\{t>1\}}A_{t-1}^{(k+1)}(B_{t-1}^{(k)})^\top + \mathbf{I}_{\{t<T\}}A_{t+1}^{(k)}(B_{t+1}^{(k)})^\top\right)}_{\text{Neighboring timestamps' current imputation}} + \lambda_3 Y_t$$

Compared to the original regression response $X_t$, the synthetic response $Z_t^{(k)}$ is a *weighted* linear combination of:

1. (Red part) The original data matrix $X_t$ with its missing entries replaced by the current imputation $A_t^{(k)}(B_t^{(k)})^\top$.
2. (Blue part) The sum of the neighboring timestamps' (i.e. $t-1$ and $t+1$) imputation.
3. (Purple part) The auxiliary data, i.e. output of the Spherical Harmonics algorithm.

We proved [1] that this alternating Majorization-Minimization algorithm converges to a stationary point at the rate of $\mathcal{O}(1/K)$.

## Numerical Simulation Study & Real Data Validation

We simulate 4 patterns of missingness on the IGS TEC map:

| Missing Pattern | Description |
|---|---|
| Random | Randomly drop 30%/50%/70% of the pixels of each matrix. |
| Temporal | Similar to random, but the missing pattern has a temporal drift. |
| Random + Patch | Randomly drop a $27\times 27$ or $45\times 45$ or $63\times 63$ patch as missing. |
| Temporal + Patch | Similar to random patch, but the missing patch has a fixed drift trajectory. |

Table 1. Four Simulated data-missing patterns on the IGS TEC data.

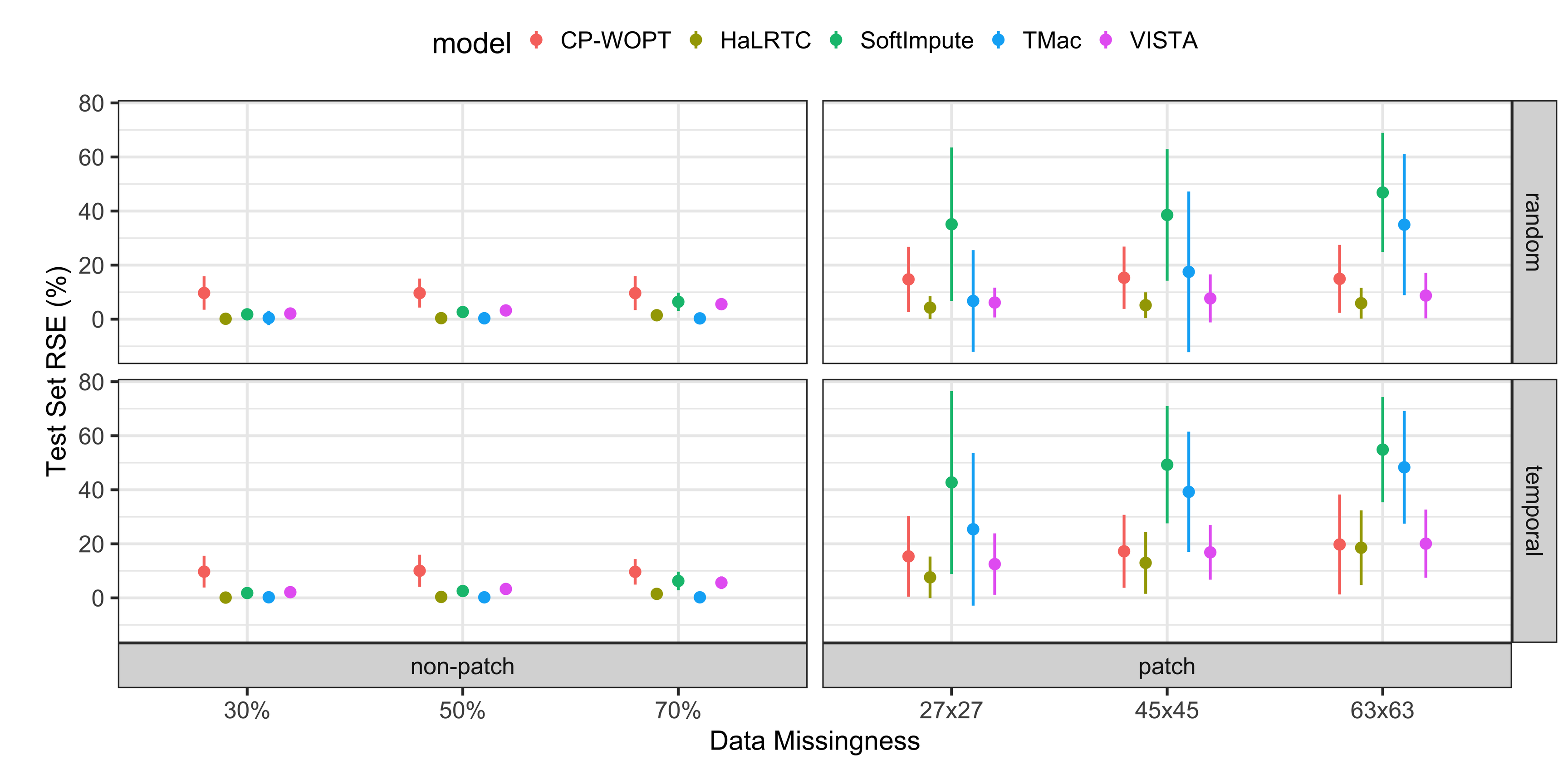


Figure 3. Method comparisons with CP-WOPT, HaLRTC, TMac, Soft-Impute and our VISTA.

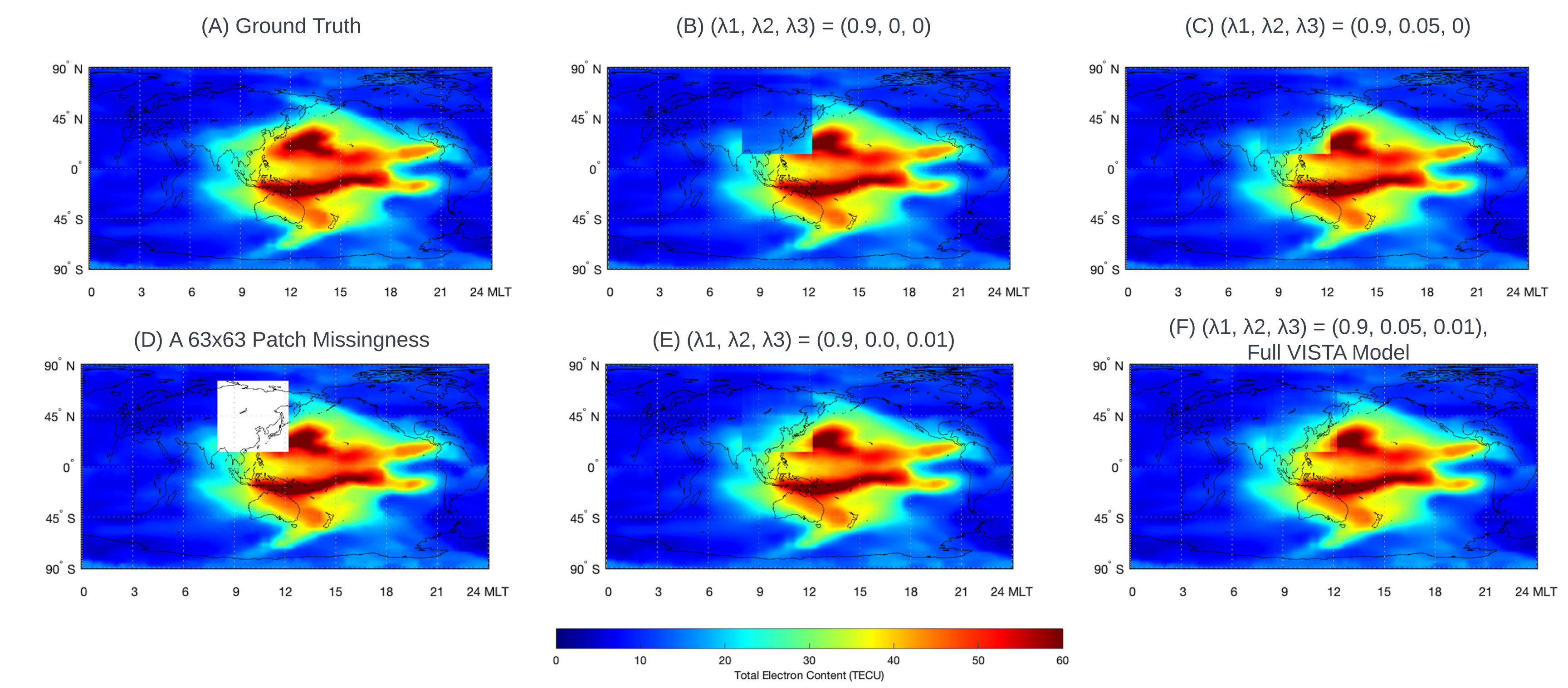


Figure 4. Compare VISTA with different tuning parameter combinations in (3).

## References

[1] **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.

[2] **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*, in press.