Cut-n-Reveal: Time-Series Segmentations with Explanations

NIKHIL MURALIDHAR, Virginia Tech, USA
ANIKA TABASSUM, Virginia Tech, USA
LIANGZHE CHEN, Pinterest, USA
SUPRIYA CHINTHAVALI, Oak Ridge National Laboratory, USA
NAREN RAMAKRISHNAN, Virginia Tech, USA
B. ADITYA PRAKASH, Georgia Institute of Technology, USA

Recent hurricane events have caused unprecedented amounts of damage on critical infrastructure systems and have severely threatened our public safety and economic health. The most observable (and severe) impact of these hurricanes is the loss of electric power in many regions, which causes breakdowns in essential public services. Understanding power outages and how they evolve during a hurricane provides insights on how to reduce outages in the future, and how to improve the robustness of the underlying critical infrastructure systems. In this paper, we propose a novel scalable segmentation with explanations framework to help experts understand such datasets. Our method, CnR (Cut-n-Reveal), first finds a segmentation of the outage sequences based on the temporal variations of the power outage failure process so as to capture major pattern changes. This temporal segmentation procedure is capable of accounting for both the spatial and temporal correlations of the underlying power outage process. We then propose a novel explanation optimization formulation to find an intuitive explanation of the segmentation, such that the explanation highlights the *culprit* time-series of the change in each segment. Through extensive experiments, we show that our method consistently outperforms competitors in multiple real datasets with ground truth. We further study real county-level power outage data from several recent hurricanes (Matthew, Harvey, Irma) and show that CnR recovers important, non-trivial and actionable patterns for domain experts, while baselines typically do not give meaningful results.

CCS Concepts: • **Networks** → *Sensor networks*; • **Information systems** → **Data mining**; **Spatial-temporal systems**.

Additional Key Words and Phrases: Multivariate Time-series, Spatio-temporal Segmentation

ACM Reference Format:
Nikhil Muralidhar, Anika Tabassum, Liangzhe Chen, Supriya Chinthavali, Naren Ramakrishnan, and B. Aditya Prakash. 2019. Cut-n-Reveal: Time-Series Segmentations with Explanations. *ACM Trans. Web* 9, 4, Article 39 (January 2019), 26 pages. https://doi.org/0000001.0000001

This document has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

Authors' addresses: Nikhil Muralidhar, Virginia Tech, USA, nik90@vt.edu; Anika Tabassum, Virginia Tech, USA, anikat1@vt.edu; Liangzhe Chen, Pinterest, USA, liangzhechen@pinterest.com; Supriya Chinthavali, Oak Ridge National Laboratory, USA, chinthavalis@ornl.gov; Naren Ramakrishnan, Virginia Tech, USA, naren@cs.vt.edu; B. Aditya Prakash, Georgia Institute of Technology, USA, badityap@cc.gatech.edu.

1559-1131/2019/1-ART39 $15.00
https://doi.org/0000001.0000001

1 INTRODUCTION

Power outages during several recent hurricanes have had a severe impact on our national security, economy, and public safety. The 2017 hurricane season was the most expensive in U.S. history resulting in huge economic losses (greater than $250 billion). Hurricane Irma caused one of the largest power outages which reportedly knocked out power to 4.5 million of the 4.9 million Florida Power & Light customers. Hence, better understanding of power outages and how they evolve during hurricanes is a very important task for damage prevention and control.

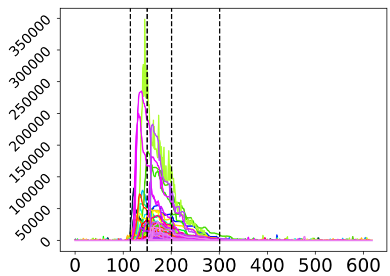
Domain experts in critical infrastructure systems (CIS) constantly seek solutions and ideas on how to reduce power outages during hurricanes. For example, Oak Ridge National Laboratory's (ORNL) Energy Awareness and Resiliency Standardized Services (EARSS) project developed a fully automated procedure to take wind speed and location estimates provided by hurricane monitoring experts and provide a geo-spatial estimate on the impact to the electric grid in terms of outage areas and projected duration of outages [12].

Retrospectively identifying 'cut-points' with a sudden change in the number of outages in historical data can help in many aspects like identifying phases and causes through inter-dependency analysis. This helps disaster management personnel learn from past events and be better prepared for future contingencies. For example, a retrospective analysis of hurricane Sandy highlighted the underlying causes due to inter-dependencies with communication, oil and natural gas infrastructures [8]. Further, pinpointing 'culprit' counties responsible for each such cut-point helps domain experts localize points of failure and analyze restorative periods [21]. Hence such analysis can be used to shorten restoration periods of vulnerable points in the grid thereby improving grid resiliency to future disasters.

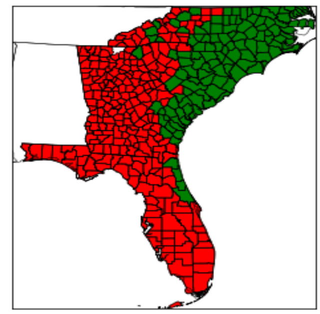
The above analysis goals may be addressed using the time-series mining task of 'segmentation'. However computing interpretable 'culprits' for each cut-point is a task which has not been studied before. In this paper we address this issue via a novel segmentation-with-explanations approach. Our main contributions are:

- We propose a novel problem and algorithm CnR for computing segments of power outage data. CnR captures temporal and spatial relationships between counties experiencing power outages, modeling the power failure process as a segmentation problem. We also propose a novel explanation algorithm that identifies the *culprit* counties for each segmentation cut-point.
- Our proposed formulation uses low dimensional latent factor models and achieves significant speed up.
- Experiments were performed with CnR and other popular segmentation algorithms on synthetic and real datasets including historical hurricane power outage data. The other segmentation procedures perform significantly worse relative to CnR on the real hurricane data, due to their inability to model complex spatial dynamics of the failure process. Although CnR is developed for power outage data, it can be applied to any multivariate time series.

The rest of the paper is organized as follows. In Section 2 we formally state the segmentation and explanation problems. Section 3 introduces our spatially agnostic CnR-V segmentation and explanation model. We then introduce the novel spatio-temporal CnR-UV model in Section 4 designed to incorporate extensive spatial information for segmenting and explaining failure process dynamics in a multivariate time series. Section 5 showcases the performance of CnR-UV with respect to other state-of-the-art algorithms, and on real-world problems. Section 6 provides a brief review of related literature and Section 7 explores avenues for future work. For lack of space, we defer some additional experiments to the appendix [6]. All codes and datasets are made public[7].



(a) Temporal Segmentation to identify different failure phases of the power grid across all counties affected by Hurricane Matthew.



(b) Spatial Clustering revealed by CnR of counties in Florida that experienced extensive damage (in green) to power infrastructure due to Hurricane Matthew.

Fig. 1. An example of the holistic spatial and temporal analysis results from our novel CnR model to analyze the damage to the power grid during Hurricane Matthew. Fig. 1(a) depicts the overall temporal segmentation over a dataset where each time series indicates the total number of households that lost power in a single county over the course of Hurricane Matthew (household count per county is recorded every hour). Fig. 1(b) represents the spatial clustering of all counties that experienced significant damage to power infrastructure through the course of Hurricane Matthew, essentially representing the spatial span of damage.

2 FOCUS AND SET-UP

Motivation: Large power grids usually contain thousands of generators, hundreds of thousands of transmission lines and millions of consumers. Grid components have strong inter-dependencies like in the transmission grid where multiple paths exist between generators and consumers and these paths typically are arranged in a mesh grid manner. Hence, if one path or line fails, the electricity instantaneously follows an alternate path governed by Kirchhoff's voltage and current laws. If the alternate path however cannot handle the overload in flow, it in-turn fails and this failure cascades to neighboring components. Due to the well studied property of cascading failures and small-world properties in the power grid [26, 28], a few initial points of failure due to a hurricane quickly cause network instability in a region potentially causing millions of people to suffer the effects of brownouts or blackouts. Natural disasters like hurricanes exhibit multiple phases of varied intensity along their path causing failures with different levels of severity at different regions. We model the progression of this grid failure process as a temporal segmentation problem. Modeling this failure process over time, across different regions (e.g. counties) affected by a hurricane, is essential for improving the resilience of critical infrastructure to disasters.
Focus: We characterize the severity of this grid failure process by measuring the number of people in a hurricane affected region (a county in our case) without power over the entire time period of the hurricane. Three critical questions need to be answered for characterization of this process:

- How can we identify different phases of a hurricane as a function of severity of the damage to critical infrastructure like the power grid, using sparse customer power loss data?
- Which counties are most important for characterizing each phase?
- How can counties be grouped together based on their overall failure dynamics during the hurricane?

Our main goal is to help domain experts answer the aforementioned questions.

Notation: We assume we are given a set of time series $X = \{x_1, x_2, ..., x_n\}$, where each time series $x_i = [x_i(t_1), x_i(t_2), ..., x_i(t_m)]$, and $x_i(t_j)$ represents the value at time stamp t_j for the i^{th} time series. We also assume there is a known underlying graph structure G that captures the relationship among these time series $\{x_i\}$. For example, in critical infrastructure systems, the number of electric outages in all disaster affected counties form a set of time series, and the relationship among these counties can be based on their geographical proximity.

Definitions: Our algorithm CnR (Cut-n-Reveal) contains two parts: detecting a good *segmentation* of the outage data to capture the main changes; and finding the corresponding *explanations* (subset of important counties) per segment. With this knowledge of the segmentation and the explanations for each segment, the expert has a holistic picture of the different phases of the failure process as well as the specific time series that contributed significantly to each phase change. We now formally state our definition of a segmentation and an explanation.

Definition 1 (Segmentation S). *A segmentation of X contains a set of distinct temporal cut-points $S = \{c_1, c_2, ..., c_k\}$, where $c_i \in \{t_1, t_2, ..., t_m\}$.*

Definition 2 (Explanations E). *$E = \{e_1, e_2, ..., e_k\}$, where e_i is an n by 1 non-negative explanation vector. $||e_i||_1 = 1$ and e_{ij} represents the importance of time series j for explaining the cut-point c_i.*

Set-up: The cut-points of S naturally divide the entire time period in the dataset into a set of disjoint time segments. The i^{th} time segment is denoted as a set of contiguous time steps $s_i = [c_{i-1}, c_i)$ with $i \in \{1, 2, .., k+1\}$, $c_0 = t_1$, and $c_{k+1} = t_m$. Two sets / segments s_i, s_j are said to be neighboring segments if $s_i = \{t_l, .., t_{l+\Delta l}\}$ and s_j contains t_{l-1} or $t_{l+\Delta l+1}$.

Assuming we are given a segmentation S of X, containing a set of cut-points $\{c_i\}$ and corresponding segments, $\{s_i\}$, a desired explanation of the segmentation should be simple yet effective enough to guide efforts to prevent or curtail the effects of critical infrastructure failure in future disasters.

To this end, we introduce an explanation vector e_i for each cut-point c_i in S. Each e_i is an $n \times 1$ vector where n represents the number of counties and e_{ij} represents the importance of the j^{th} time series/county in explaining the cut-point. Intuitively, if time series x_j shows very different patterns before and after the cut-point c_i, we consider it important in explaining why c_i is a good cut-point. On the other hand, if x_j remains constant/unchanged across c_i, it does not provide useful information in terms of the cut-point c_i and should have low values in e_i. In the hurricane outage data where there are hundreds of time series/counties, such explanation vectors are able to highlight the "culprit" time series/counties

We propose two versions of the CnR algorithm, namely CnR-V (Temporal Cut-n-Reveal) and CnR-UV (Spatio-temporal Cut-n-Reveal). Section 3.1, Section 3.2 discuss the segmentation and explanation formulations and corresponding solutions using CnR-V. Section 4.1, Section 4.2 detail our segmentation and explanation solutions using CnR-UV.

3 TEMPORAL CUT-N-REVEAL

In this section we propose our CnR-V model which performs segmentation on a multivariate time series only using temporal information and also yields explanations for each segmentation cut-point.

3.1 CnR-V Segmentation

The segmentation problem addressed by CnR-V is stated as follows.

Problem 1. ***Given** a set of time series X and a number k, **find** the k-segmentation of S that captures the main pattern changes in X.*

3.1.1 Overview of our approach. Through the segmentation model, we wish to isolate temporal sequences into discrete segments such that the properties of the failure process in each segment differ from neighboring segments. The process of manually or algorithmically picking reasonable segments is non-trivial as segments that are too small fail to capture significant properties of the failure process and picking segments that are too large although capturing all failure process characteristics, do not highlight the differences between the various phases of the process. Since the failure process is highly dynamic and the failure dataset is sparse in nature, methods based on capturing long-term correlation [25] or invariant learning [37] from the data will be unable to perform adequately.

Table 1. Definitions

$X \in \mathbb{R}^{n \times m}$	The data matrix consisting of n time series each with m time steps.
$D \in \mathbb{R}^{n \times n}$	Depicts a degree Matrix
$A \in \mathbb{R}^{n \times n}$	Depicts an adjacency Matrix
$L = D - A$	Represents the Laplacian Matrix of A.
$U \in \mathbb{R}^{n \times l}$	Spatial feature matrix with l latent features.
$V \in \mathbb{R}^{l \times m}$	Temporal feature matrix with l latent features. In Eq. 1 i.e in CnR-V, $l = m$, otherwise, i.e. in CnR-UV, $l << m$ (Eq. 5).
$R \in \mathbb{R}^{m \times m-1}$	A lower triangular matrix with -1's on the primary diagonal and 1's on the second diagonal
$E \in \mathbb{R}^{n \times k}$	Represents the explanation matrix to quantify importance of each time series in explaining each cut-point.
$\mathbf{1} \in \mathbb{R}^{n \times 1}$	Denotes a vector of ones
$\lambda, \alpha, \lambda_i, \gamma_i$	Scalar hyper-parameters used in the segmentation and explanation formulations.

3.1.2 Formulation. We consider the different phases of the failure process in the power grid during natural disasters as a collection of disjoint segments $\{s_1,..,s_{k+1}\}$. We wish to discover a collection of k cut-points S that minimizes similarity between any two neighboring segments s_i, s_j. Hence, each segment s_i would capture a different pattern from its neighboring segments (s_{i-1}, s_{i+1}), thus the segmentation S captures pattern changes in the time series. We employ the normalized cut framework which has been shown to work well in subspace clustering and segmentation tasks [50]. Our goal now, is to represent each time step allowing for effective similarity calculation between time steps so that the continuous evolution of the failure process is captured by the inter-time-step similarity. In an effort to find a principled approach to capture the similarity between different time steps in the failure process, we adopt the formulation provided by Tierney et al. [52] for video scene segmentation for our purposes of modeling the hurricane failure process. The model represents each time step in the data X, as a function of other important time steps. It is through this latent representation V that we attempt to capture the dynamics in the data X.

$$\begin{aligned} &\min_V \frac{1}{2}||X - XV||_F^2 + \lambda_1||V||_1 + \lambda_2||VR||_{1,2} \\ &\text{subject to} \qquad \mathbf{diag}(V) = 0 \end{aligned} \tag{1}$$

In Eq. 1, V is an m by m matrix whose i^{th} column can be considered the latent representation of time step i in terms of all the other time steps. The first term in Eq. 1 calculates the reconstruction error between X and XV while, the second term introduces sparsity into the latent representation, enforcing that each time step be explained as a function of a small subset of other important time steps. The term VR (R is defined in Table 1) calculates the difference of each time step with its previous time step in the latent V space. This term essentially serves as a smoothness constraint penalizing the dissimilarity of neighboring time steps. The $l_{1,2}$ norm term forces whole column similarity between two columns of V, i.e. between neighboring time steps in V as opposed to just element-wise similarity in the case of a simpler l_1 norm on VR. The solution to Eq. 1 can be obtained by applying the alternating direction method of multipliers (ADMM) [15].

Solving Eq. 1 yields a temporal weight matrix $V \in R^{m \times m}$ from which we derive an affinity matrix $W = VV^T$. The affinity matrix is then segmented using the normalized cuts procedure to obtain the set of cut-points S.

3.2 CnR-V Explanation

Recently, there has been a push toward making complex machine learning model outputs quantifiable, explainable and simple [45]. Despite the sparsity of our segmentation procedure, the output is complex and it is often not possible to identify the cause for each segment due to many simultaneously changing time series. A domain expert may want to know simply which time series are changing and which are behaving anomalously at the sudden outage changes (at each segment). This can help them make decisions about which counties they can use for retrospective analysis by localizing points of failure. Finding out these “culprit" time series using just the temporal and spatial segmentation from V and U matrices seems difficult. Existing time series segmentation algorithms do not provide any explanation of the result in an automatic principled way. Hence, to design good explanations specifically for hurricane outage data, we consider the characteristics of the data, as well as the requirements from the domain experts to propose an optimization problem CnR-V as follows:

Problem 2. ***Given*** *a set of time series X, the Laplacian matrix L of the underlying network, a number k, and the k-segmentation of S,* ***find*** *the associated explanations E, that capture the main pattern changes in X.*

3.2.1 Overview of our approach. We formulate an optimization problem that automatically learns explanations considering the underlying geographical relation between counties, revealing to domain experts, a small number of truly important "culprit" counties per cut-point.

3.2.2 Formulation. We aim to design an optimization problem that automatically finds a good set of explanation vectors $\{e_i\}$. Assume that we have a function $d(S, i)$, which takes a segmentation S and a cut-point index i as inputs, and returns an n by 1 vector which captures the difference of each time series before and after the i^{th} cut-point c_i in S. We want e_i to assign higher weights to time series with higher $d(S, i)_j$ values (therefore higher difference across cut-point c_i) where $d(S, i)_j$ represents the importance of county j at cut-point i and is defined in Eq. 3. The formulation also needs to capture the effects of spatial proximity of counties i.e adjacent counties should have similar importance, due to the continuous trajectory of a hurricane affecting neighboring counties at the same time. The explanation needs to be 'simple' in the sense of highlighting only a few culprit counties. With these considerations in mind, the optimization problem we solve to obtain simple explanations considering the county geography is shown below.
Given: A set of time series X,L, a segmentation S, α, λ.
Find: $E = \{e_i\}$ such that

$$\begin{aligned} &\arg\max_{E} \sum_{i=1}^{k} [e_i^T d(S,i) - \alpha e_i^T L e_i] - \lambda \sum_{i=1}^{k} ||e_i||_1 \\ &\text{subject to} \quad 0 \le e_{ij} \le 1, ||e_i||_1 = 1 \end{aligned} \tag{2}$$

The geographical smoothness is introduced in the second term using the Laplacian matrix L (obtained from the underlying county-county network). This term minimizes the difference of e_i for adjacent counties. The third term is an l_1 norm regularization on e_i, which introduces sparsity in e_i leading to simpler explanations due to only a few important counties having non-zero values in e_i to explain cut c_i.

The distance function $d(S, i)$ captures the difference across a cut-point c_i, by considering a time window before the cut-point c_i and a time window after c_i. The difference of these two time windows is calculated as the difference of the time series across c_i. Let w_{ij}^- represent the sub-sequence of x_j in the time window before c_i, and w_{ij}^+ represents the sub-sequence in the time window after c_i. The distance function then calculates the difference of w_{ij}^- and w_{ij}^+ using simple, standard time series features: the mean value (f_1), the standard deviation (f_2), the maximum value (f_3) and the minimum value (f_4).

$$d(S,i)_j = \frac{1}{4} \sum_{z=1}^{4} |f_z(w_{ij}^-) - f_z(w_{ij}^+)| \tag{3}$$

As a preprocessing step which we do not elaborate on in the equation, we perform a min-max normalization of $|f_z(w_{ij}^-) - f_z(w_{ij}^+)|$ across all time series to make the scales uniform. As both w_{ij}^- and w_{ij}^+ are of a short length (a deliberate setting since the pattern changes that justify the choice of a particular cut-point usually lie in the local area), these simple features are enough to capture the main pattern difference.

Finally, to solve Eq. 2, we optimize each e_i separately. For each e_i, the optimization can be re-written as a Quadratic Programming (QP) problem in the following way.

$$\begin{aligned} &\arg\min_{e_i} \quad \alpha e_i^T L e_i - [d(S,i)^T - \lambda \mathbf{1}^T] e_i \\ &\text{subject to} \quad 0 \le e_{ij} \le 1, ||e_i||_1 = 1 \end{aligned} \tag{4}$$

The QP problem is well studied in the literature, and it is NP-hard in its general form. In our case, where the QP is convex in e_i, it can be solved in polynomial time using an Interior Point method [57], and we use the existing Matlab function (quadprog) to solve the problem.

4 SPATIO-TEMPORAL CUT-N-REVEAL

We now augment our CnR methodology to incorporate spatial relationships into the temporal segmentation and explanation phases and propose our novel CnR-UV model .

4.1 CnR-UV Segmentation

Modeling the power outage failure process using CnR-V presents a few drawbacks. Firstly, the segmentation process in Eq. 1 does not account for or attempt to model spatial relationships between entities (counties in our case) over which the failure process (power outage) occurs. However, phenomena like cascading failures indicate the existence of strong spatial interactions between components in the power grid and incorporating spatial relationships can aid in more effective modeling of the power outage process. We update problem 1 stated in Section 3.1 as follows:

Problem 3. ***Given*** *a set of time series X, the Laplacian matrix L of the underlying spatial network, and a number k,* ***find*** *a spatial weight matrix U, a temporal weight matrix V, the k-segmentation of S derived from V that captures the main temporal pattern changes in X.*

4.1.1 Overview of our approach. For effective spatio-temporal modeling, we allow temporal latent matrix V to consider the underlying spatial relationships between counties. To this end, we introduce a spatial weight matrix U, jointly learnt with V. U, V are latent weight matrices, and the latent factor modeling approach is used for our segmentation model because of its success in similar sparse settings like recommendation systems [30, 38].

4.1.2 Formulation. We develop a temporal segmentation formulation influenced by spatial constraints, where the failure process at each time step is represented in a rich low-dimensional latent space l such that $V \in \mathbb{R}^{l \times m}$, $U \in \mathbb{R}^{n \times l}$ and $l << m$, $l << n$. Let $v_i \in \mathbb{R}^{l \times 1}$ and $v_j \in \mathbb{R}^{l \times 1}$ represent the i^{th} and j^{th} column vectors of V respectively where $i, j \in \{1, .., m\}$ and $i \neq j$. We can also consider v_i and v_j to be the latent representation for the i^{th} and j^{th} time steps respectively. The goal of this formulation is that the similarity of v_i, v_j is not solely influenced by the temporal proximity of v_i to v_j (a constraint strongly enforced by the $||VR||_{1,2}$ term in Eq. 1) but also by the underlying spatial behavior of the counties at time steps i and j. To achieve this goal, we formulate a novel temporal segmentation model in Eq. 5 to jointly model spatial and temporal characteristics of power outage during natural disasters.

$$\begin{aligned} \min_{U,V} \ & \frac{1}{2}||X - UV||_F^2 + \lambda_1||U||_1 + \frac{\beta_1}{2} Tr(U^T LU) \\ & + \lambda_2||V||_1 + \lambda_3||VR||_{1,2} \\ \text{subject to} \quad & U \geq 0, V \geq 0 \end{aligned} \tag{5}$$

In Eq. 5, the matrix U learns the latent representation for each of the n counties. The first term calculates the re-construction error wherein the original failure behavior observed in X, is re-created as a combination of U (latent county outage characteristics), and V (latent temporal outage characteristics). The l_1 norm terms on U and V ensure sparsity, in line with our goal of designing simple interpretable explanations of our segmentation. The $||VR||_{1,2}$ term has the same effect as described in section 3.1. The term $Tr(U^T LU)$ represents the Laplacian regularization constraining the matrix U to be influenced by the underlying geographic layout of the counties in X. Here, the matrix L represents the Laplacian of the county-county adjacency matrix A.

We once again employ the ADMM method to solve Eq. 5 using the Lagrangian formulation as represented in Eq. 6. To separate each term in Eq. 5, we assign $J = U$, $K = V$, $P = KR$.

$$\begin{aligned} \mathcal{L}(U, V, J, K, P) = & \frac{1}{2}||X - JK||_F^2 + \lambda_1||U||_1 + \frac{\beta_1}{2}\mathbf{Tr}(J^T LJ) \\ & + \lambda_2||V||_1 + \lambda_3||P||_{1,2} + \langle G, V - K\rangle + \frac{\gamma_1}{2}||V - K||_F^2 \\ & + \langle H, U - J\rangle + \frac{\gamma_2}{2}||U - J||_F^2 + \langle F, P - KR\rangle \\ & + \frac{\gamma_3}{2}||P - KR||_F^2 \end{aligned} \tag{6}$$

We solve for U, V, J, K, P in an alternating manner. The update steps for each term in Eq. 6 are discussed below

(1) **Update V:**
   (a) Fixing U, J, K, P we solve for V.

$$\min_{V} \quad \lambda_2||V||_1 + \langle G, V - K\rangle + \frac{\gamma_1}{2}||V - K||_F^2$$

This can be restated as follows:

$$\min_{V} \quad \lambda_2||V||_1 + \frac{\gamma_1}{2}||V - (K - \frac{G}{\gamma_1})||_F^2 \tag{7}$$

Equation 7 has element-level closed form solutions that can be obtained using the soft thresholding operator [10, 34, 52]. The element-level closed form solution is as defined in Eq. 8

$$V = sign\left(K - \frac{G}{\gamma_1}\right) max\left(\left|K - \frac{G}{\gamma_1}\right| - \frac{\lambda_2}{\gamma_1}\right) \tag{8}$$

(b) Fixing U, V, J, P we solve for K.

$$\begin{aligned}\min_{K} \frac{1}{2}||X - JK||_F^2 + \langle G, V - K\rangle + \frac{\gamma_1}{2}||V - K||_F^2 \\ + \langle F, P - KR\rangle + \frac{\gamma_3}{2}||P - KR||_F^2\end{aligned} \tag{9}$$

Differentiating Eq. 9 w.r.t K and setting the derivative to zero yields:

$$\begin{aligned}(X^T J^T JX + \gamma_1 I)K + \gamma_3 KRR^T = \\ X^T JX + G + \gamma_1 V + FR^T + \gamma_3 PR^T\end{aligned} \tag{10}$$

If we set,

(i) $A = (X^T J^T JX + \gamma_1 I)$

(ii) $B = \gamma_3 RR^T$

(iii) $C = X^T JX + G + \gamma_1 V + FR^T + \gamma_3 PR^T$

then Eq. 10 takes the form of a sylvester equation.

$$AK + KB = C \tag{11}$$

The solution to Eq. 11 is a well studied problem. [13, 24].

(c) Fixing U, V, K, J we solve for P.

$$\min_{P} \lambda_3||P||_{1,2} + \langle F, P - KR\rangle + \frac{\gamma_3}{2}||P - KR||_F^2 \tag{12}$$

this is equivalent to

$$\min_{P} \lambda_3||P||_{1,2} + \frac{\gamma_3}{2}||P - \left(KR - \frac{F}{\gamma_3}\right)||_F^2 \tag{13}$$

Let $M = KR - \frac{F}{\gamma_3}$ then Eq. 13 has the following closed form solution [52]

$$P(:, i) = \begin{cases} \frac{||M(:,i)|| - \frac{\lambda_3}{\gamma_3}}{||M(:,i)||} M(:, i) & if\ ||M(:, i)|| > \frac{\lambda_3}{\gamma_3} \\ 0 & otherwise \end{cases} \tag{14}$$

(2) **Update U:**

(a) Fixing V, J, K, P we solve for U. We can follow a similar procedure to the V update step above and obtain the following element-wise closed form solution for U.

$$U = sign\left(J - \frac{H}{\gamma_2}\right) max\left(\left|J - \frac{H}{\gamma_2}\right| - \frac{\lambda_1}{\gamma_2}\right) \tag{15}$$

(b) Fixing U, V, K, P we solve for J.

$$\min_{J} \frac{1}{2}||X - JK||_F^2 + \frac{\beta_1}{2} Tr(J^T L J) + \langle H, U - J \rangle + \frac{\gamma_2}{2}||U - J||_F^2 \tag{16}$$

Differentiating Eq. 16 w.r.t J and setting the derivative to 0 similar to the K update procedure, the expression in Eq. 17 is obtained.

$$\beta_1 LJ + J(XKK^T X + \gamma_2 I) = XKX^T + H + \gamma_2 U \tag{17}$$

We once again solve Eq. 17 by reduction to a Sylvester equation,

(i) $A = \beta_1 L$

(ii) $B = XKK^T X + \gamma_2 I$

(iii) $C = XKX^T + H + \gamma_2 U$

(3) **Update G:**

$$G = G^{old} + \gamma_1(V - K)$$

(4) **Update H:**

$$H = H^{old} + \gamma_2(U - J)$$

(5) **Update F:**

$$F = F^{old} + \gamma_3(P - KR)$$

(6) **Update $\gamma_1, \gamma_2, \gamma_3$:**

$$\gamma_1 = \rho\gamma_1^{old}; \gamma_2 = \rho\gamma_2^{old}; \gamma_3 = \rho\gamma_3^{old}$$

Solving Eq. 6 yields an $l \times m$ temporal weight matrix V and an $n \times l$ spatial weight matrix U. We construct an affinity matrix $W = V^T V$ which is passed to the normalized cuts algorithm to obtain the temporal segmentation S from V. As an additional step, we also construct a separate spatial affinity matrix $W_U = UU^T$ which represents county similarity, and obtain a spatial clustering of W_U using the normalized cuts procedure. In addition to the temporal segmentation and the explanation of each temporal segment, we believe the spatial clustering of counties provides an additional level of insight about aggregate county behavior during natural disasters in a given region.

4.2 CnR-UV Explanation

The explanation procedure for model CnR-V proposed in Section 3.2, incorporates spatial constraints through the Laplacian of the county adjacency matrix. In our experience, this approach overly emphasizes spatial locality as a factor for learning the explanation vector e_i for a particular cut-point c_i. It need not always be the case that the effects of a power outage in a particular county are felt only in the neighboring counties or that only counties directly affected by a hurricane experience outages. As outlined in [27], the influence graph for county power outage need not necessarily be exactly similar to the grid topology or geographic county layout. Intuitively this means that outage in a county in one part of a state can have far-reaching effects leading to outages or in counties located in a different part of the state in an instantaneous or delayed manner. Accommodating for such effects in our explanation methodology requires a smoother, less stringent spatial constraint. This leads to better explanations for domain experts as well.

4.2.1 Overview of our approach. In addition to the V matrix, the segmentation formulation of the CnR-UV model also learns a rich latent factor representation of each disaster-affected county U. Due to the richness of the latent factor representation and flexibility of design, the U matrix in addition to local spatial effects, is also able to capture far-reaching effects of counties on each other. Hence, the affinity matrix UU^T would capture far-reaching county-county similarities extending significantly beyond the immediate neighborhood of a county. We utilize this property and construct $W_U = UU^T$ which can be considered the adjacency matrix of a weighted undirected graph of counties whose degree matrix D is a diagonal matrix where each D_{jj} represents the weight of county j and is calculated as the sum of row j of W_U. Matrix L_U is the Laplacian calculated as $L_U = D - W_U$.

PROBLEM 4. ***Given*** *a set of time series X, a number k, the k-segmentations of S, and the Laplacian L_U derived from the spatial latent factor matrix U,* ***find*** *the associated explanations E, that capture the main pattern changes in X.*

4.2.2 Formulation. The explanation formulation employed by CnR-UV is defined in Eq. 18. It is similar to Eq. 2 but for a smoother Laplacian matrix L_U in the regularization term which allows e_i to consider counties in a larger spatial radius as opposed to in the explanation step of CnR-V wherein a strict spatial constraint based on the county graph is imposed through the Laplacian matrix L.

$$\begin{aligned} &\arg\max_{E} \sum_{i=1}^{k} [{e_i}^T d(S,i) + \alpha e_i^T L_U e_i] - \lambda \sum_{i=1}^{k} ||e_i||_1 \\ &\text{subject to} \quad 0 \le e_{ij} \le 1, ||e_i||_1 = 1 \end{aligned} \tag{18}$$

The function $d(S, i)$ returns an explanation vector $e_i \in \mathbb{R}^{n \times 1}$ and is defined in Eq. 3. Equation 18 can be solved by optimizing each e_i as a separate QP problem convex in e_i similar to the explanation formulation for CnR-V in Section 3.2. The complete pseudo-code for CnR-UV is given next (CnR-V is similar, using Eq. 1, Eq.4 instead of Eq. 6, Eq. 18 respectively).

ALGORITHM 1: CnR-UV Segmentation with Explanation

Input: X: Hurricane Power Outage Data, G: Spatial Graph, l: Num. Latent Features
Result: $S = \{c_1, ..., c_k\}$, Temporal Segmentation
$E = \{e_{c_1}, .., e_{c_k}\}$, Temporal Explanation
Init: U = V = 0
while *not converged* **do**
  Estimate, U, V using Eq. 6
  Retrieve S using Normalized Cuts
  Estimate explanation vectors $e_{c_1}, .., e_{c_k}$ using Eq. 18
end

REMARK 1. *CnR-UV takes worst-case time $O(\#iter l^2(X + m^2 + n^2))$i.e quadratic in the number of time steps and time series. In practice we found QP to be very fast, and total time to be sub-quadratic on the dataset size. CnR-V takes worst-case time $\sim O(\#iter(m^{2.3} + n^2))$ (using the best matrix multiplication exponent) and is hence sub-cubic in the number of time steps and quadratic in the number of time series. The space complexity of CnR-UV is near-linear $O(X + nl + ml)$ and that of CnR-V is $O(X + m^2 + n^2)$.*

5 EMPIRICAL STUDY

We implement CnR in Python and Matlab. Our experiments were conducted on a 4 Xeon E7-4850 CPU with 512 GB of 1066Mhz main memory.

5.1 Set-up

5.1.1 Datasets. We collect datasets from different domains with the ground truth segmentations to quantitatively evaluate our performance. For efficiency purposes, we perform a standard rolling average as a pre-processing step to all the data. The final statistics are in Table 2.

ChickenDance: The *ChickenDance* datasets, *ChickenDance1* and *ChickenDance2* [17] are recorded as motion capture sequences of 4-dimensional data points with ground-truth segmentation [35] and is originally from CMU motion capture database [2]. The ground-truth segmentation is based on different motions in the chicken dance.

WalkJog: We used two variants of the *WalkJog* datasets, *WalkJog1* where we uniformly sampled 1000 data points from the *WalkJog* dataset used in [23] and *WalkJog2* used in [18]. These datasets adapted from the REALDISP Activity recognition Dataset [11] have recordings of walking and jogging motions with segments between different motions.

GrandMal Seizures: [23] has 3-min recordings of neural activity (pre-seizure, seizure and post-seizure) of a subject, recorded using a scalp electrode.

Synthetic Data: We also generated synthetic data consisting of 4 time-series sampled from normal distributions with different means and standard deviations. Time-series were perturbed at different times to cause segments and the goal is to identify these segments.

NILM: Non Intrusive Load Monitoring dataset. This dataset consists of real power measurements for various household appliances like lamps, laptops, and refrigerators, recorded through the use of MAU (Measurement and Actuation Units) connected between the device and the wall-socket (more details are in [44]). We use a 24-hr hour snapshot of the NILM data from 2012-01-17 00:00:00 to 2012-01-17 23:59:59 sampled at two minute intervals, and use the time when a device switches states as the ground truth cut-points.

Hurricane Outage data: ORNL has developed several grid situational awareness products over the last decade such as VERDE, EARSS and EAGLE-I [19] for different stakeholders like DOE and FEMA, primarily for emergency management. For example, the National Outage Map within EAGLE-I collects distribution outage data of all the customers from utility websites every 15 minutes. Due to the recent coverage expansion (with more utilities exposing data from their Outage Management Systems), in this paper, we consider the more recent hurricane outage data namely for Matthew, Harvey and Irma since it covers nearly 90% of the population in the hurricane affected areas.

5.1.2 Baselines. We wish to evaluate the *segmentation* and *explanation* parts of our CnR-UV algorithm. We first start with evaluating the performance of the CnR-UV segmentation procedure and later detail the CnR-UV explanation evaluation.

5.1.3 Segmentation Baselines. First, we compare the segmentation of CnR-UV with several state-of-the-art multivariate time series segmentation algorithms.

Autoplait [35] is a hidden markov model (HMM) based algorithm which discovers different regimes in co-evolving time series. Each regime can be thought of as the segments for our problem.

TICC [25] is a recent algorithm for multivariate time series to discover repeated patterns. It clusters time stamps into segments using their model.

Dynammo [32] learns a dynamical system (Kalman Filter) and segments the time-series wherever the reconstruction error becomes high.

Floss [23] is an unsupervised semantic segmentation algorithm which learns the segmentation from the local minimas obtained in the matrix profile.

5.1.4 Explanation Baselines. To the best of our knowledge, there is no method that retrieves explanations for each segment the way CnR-UV does. Hence, we are unable to compare CnR-UV

Table 2. Datasets Used.

Dataset	#Time stamps	#Time series	Ground Truth
Synthetic	1000	4	✓
NILM	721	17	✓
ChickenDance 1	1590	4	✓
ChickenDance 2	322	4	✓
WalkJog 1	1000	2	✓
WalkJog 2	303	2	✓
GrandMal	1000	2	✓
Harvey	264	250	
Irma	169	271	
Matthew	252	369	

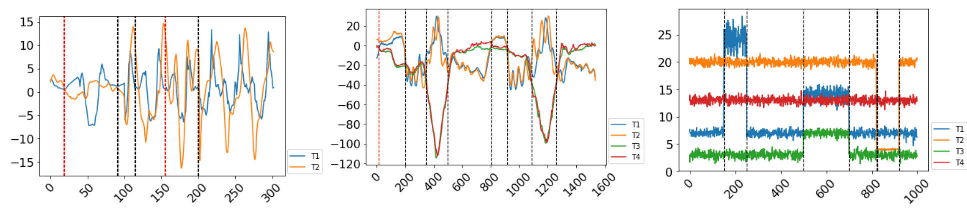
Table 3. Evaluation of segmentation (seg) and explanation (exp) on Ground Truth Datasets based on F1-score.

Dataset \ Method	CnR-UV seg	CnR-UV exp	CnR-V	Auto Plait	TICC	Dyn.	Floss
Synthetic	**1.0**	1.0	0.58	0.5	**1.0**	0.52	0.85
NILM	**0.83**	1.0	0.56	0.4	0.82	0.71	0.73
Chicken1	**0.93**	1.0	0.63	0.85	0.92	0.54	0.53
Chicken2	**0.85**	1.0	0.73	0.73	0.5	0.75	0.71
WalkJog1	0.22	1.0	0.57	0	0.86	0.54	**1.0**
WalkJog2	0.75	1.0	**1.0**	0	0.33	0	0
GrandMal	0.86	1.0	0.58	0.5	**1.0**	0.36	0.5

explanations with those of other state-of-the-art algorithms. We do however evaluate explanation performance for the aforementioned datasets with ground-truth segments. The evaluation procedures for segmentation and explanation are detailed in section 5.2.

5.2 Quantitative Evaluation

5.2.1 Segmentation. We compare CnR-UV performance with several segmentation baselines on datasets with ground truth segmentations: *NILM, ChickenDance, Synthetic, GrandMal* and *WalkJog*. We evaluate the detected cut-points by calculating the F1 score based on the ground truth cut-points (as in [35]). Higher F1 scores indicate better segmentation. For all our experiments with CnR-UV, we set $l = 2$ (our algorithm was robust to varying latent factor dimensions) and chose the hyperparameters using *gridsearch*. We show the results in Table 3 where we observe that CnR-UV outperforms all methods on most datasets except *GrandMal, WalkJog1*. To visually inspect the CnR-UV segmentation, we depict segmentation results in Fig. 2 for *ChickenDance1*, *WalkJog2*, and *Synthetic* datasets where CnR-UV performs the best. For *ChickenDance1* in Fig. 2(b) CnR-UV is able to isolate all the different data trends successfully. It correctly identifies all of the 7 ground truth cut-points, the remaining cut-point at time step 13 (red dashed line) is a false positive. In Fig. 2(a), for *WalkJog2*, we see that CnR-UV correctly separates the sequences of data generated due to walking from those due to jogging. Time series segmentation models are less affected by the number of time series in the model and more by the degree, frequency of perturbation of the time



(a) *WalkJog.* (b) *ChickenDance.* (c) *Synthetic.*

Fig. 2. CnR-UV segmentation results (vertical dashed lines) for the *Walkjog2*, *ChickenDance1*, and *Synthetic* datasets. True positive segments are colored black and false positives are colored red. We consider all the segments within tolerance window (5% of total timesteps) of the ground truth segments as true positive.

series. We have shown in Table 3 that CnR-UV performs well in cases where the number of time series is high e.g NILM. Each cut-point discovered by our method lies in a 5% cut-point location tolerance window with respect to the ground truth cut-point, we adopt this practice from previous literature [17, 35].

5.2.2 Explanation. CnR-UV is also able to retrieve reasonable explanations for proposed segmentations in each case. Since there is no existing literature performing explanation in an automatic and principled way, we were unable to compare our explanation algorithm with other baselines. Also, since we did not have any ground truth for explanations, we created a ground truth dataset by manually generating explanations for each cut-point. We did this by identifying a subset k_i of the n time series in a dataset that experienced perturbation across a cut-point c_i. This subset k_i of time series can be considered the ground truth explanation for cut-point c_i. We then compare the top $|k_i|$ (cardinality of set k_i) values in the explanation vector e_i against the ground truth explanations using the F1 score. This is repeated for all cut-points and an average F1 score is calculated for explanations on the dataset. Results of this procedure have been outlined for each dataset in Table. 3 (CnR-UV exp). Explanations were evaluated on *ChickenDance*, *WalkJog*, *NILM*, *Synthetic* and *GrandMal* datasets. We only consider true positive segments identified by CnR-UV while calculating explanation F1 scores. For *WalkJog1* dataset, even though CnR-UV segmentation is low (F1 score = 0.22) our explanation performs well (F1 score = 1.0). This is because we only use the true positive segments and calculate explanations for those segments because of the availability of ground truth explanation data only for true positive segments.

5.2.3 Discussion of CnR-UV compared to other baselines. For the best performance, CnR-UV models should be provided with the spatial graph relating the time series being modeled. CnR-UV was so designed, bearing in mind the goal of modeling power system failure processes during hurricanes using real-world data. However, in order to holistically evaluate temporal segmentation performance, we compared CnR-UV to many state-of-the-art baselines on several datasets. In this context, all time series were considered spatially independent (as we did not have prior knowledge of spatial inter-dependencies). Due to this lack of spatial information, CnR-UV underperformed as the 'U' matrix was unable to learn the best possible representation. Despite this, our model matches or outperforms strong baselines like TICC, Dynammo, Autoplait and Floss in 5 out of 7 datasets. It must be noted that despite the lack of spatial information, CnR-UV also outperforms the CnR-V model on 5 out of 7 datasets (Table 3) indicating that the low-dimensional latent factor U and V matrices in CnR-UV are indeed able to learn rich representations of the failure process compared to the large sparse square V matrix as in the case of CnR-V.In the datasets where spatial information is missing, we treat all time-series as independent. The reason for the underperformance of CnR-UV in Walkjog and GrandMal datasets w.r.t TICC may have to do with this time series independence assumption being sub-optimal.

Dynammo performs segmentation based on re-construction error w.r.t a tolerance threshold specified by the user. We found that the segmentation was sensitive to this tolerance threshold parameter which directly governs the number of segments allowed. In most cases, Dynammo was found to over-segment or under-segment depending on the error tolerance.

In the case of the *Auto-plait model*, it is found to perform better on datasets with less spiky (sudden) changes in time series. For example, if we observe Fig. 2, the *WalkJog* and *Synthetic* dataset time series have a much more spiky and sudden changing behavior than the *ChickenDance* dataset. However, in the case of the *ChickenDance* dataset, although Fig. 2(b) shows spikes around time step 400 and 1200, the rest of the patterns are either increasing or decreasing trends that are relatively non-spiky in nature.

Floss performs segmentation in multivariate timeseries by finding local minima on the average CAC curve [23]. Thus, for *GrandMal* and *ChickenDance1* data where multiple groups of time series exhibit large and sometimes spiky changes at different time steps, CAC curve of all time series do not exhibit local minima at the same or close timesteps. This results in the average CAC curve not yielding local minima at all the groundtruth cut points which causes the low F1 scores for Floss in Table 3. Also, we must note that the *ChickenDance2* dataset is smoother than *ChickenDance1* (i.e changes are smaller / more gradual than *ChickenDance1*), and we immediately see a significant performance improvement in this scenario in the F1 score of Floss.

We will now present real-world applications of CnR-UV on several hurricane power outage datasets as case studies. We characterize both CnR-UV segmentation and explanation procedures on all the hurricane datasets.

5.3 Case Studies: Hurricanes

In Section 2, we set out to design a model to address the following goals

(1) Identify phases of a hurricane as a function of severity of damage to critical infrastructure like the power grid.
(2) Identify the most important counties that characterize each phase (i.e "explain" each phase).
(3) Group counties together based on their overall failure dynamics through the hurricane, to allow for overall assessment of spatial span of the damage.

With the aforementioned goals in mind, we ran CnR-UV for power outage failure data from three recent hurricanes. We show that CnR-UV can find meaningful pattern changes and insightful associated explanations. Specifically, the current culprit definition can be used to distinguish between regimes while also applicable to the hurricane failure setting where the failure process of each county follows a typical increasing, peak, decreasing trend pattern. This is because multivariate hurricane failure time series are highly complex. Although the failure pattern (increasing failure rate, peak, decreasing failure rate) is consistent across counties, the time of failure and rate of failure differs widely across counties requiring a formulation as defined by us in Eq. 5 to capture the complexities (and local changes around cut-points) of this spatio-temporal multivariate failure process. The effectiveness of our model in capturing complex patterns in multiple hurricanes will be demonstrated in the following section. For the segmentation model we set the number of latent dimensions (l) to 5 and for each cut-point, we consider counties with explanation weight > 0.1 as important.

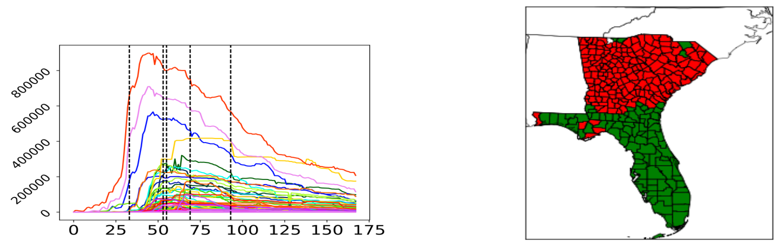
5.3.1 Hurricane Irma. We show the results in Fig. 3. Fig. 3(a) represents the overall segmentation that CnR-UV yields for hurricane Irma, while Fig. 3(c) to Fig. 3(g) show the explanations yielded by CnR-UV across each cut-point in the segmentation. All explanation figures (except those for the first and last cut-point) consist of three cut-points i.e the cut-point being explained, along with the previous and the next cut-point. Each explanation figure is accompanied with a spatial visualization of the important counties highlighted by the explanation of the cut-point. Fig. 3(h) to Fig. 3(l) are spatial depictions of the explanations in Fig. 3(c) to Fig. 3(g) respectively. Finally, Fig. 3(b) represents clustering results of the spatial matrix U. All the following hurricane result visualizations are organized in the same manner.

The first cut-point (showcased in Fig. 3(c)) at time step 35 (around September 10^{th}) shows hurricane landfall when the outages of a few counties seem to rise sharply. Indeed, Fig. 3(h) shows these counties at the southern tip of Florida indicating the location of landfall of hurricane Irma. The second and the third cut-points in Fig. 3(a) might seem redundant owing to their close proximity. However, the second cut-point (showcased in Fig. 3(d)) is capturing a small rising trend of county power outages, for the counties highlighted in Fig. 3(i). On the other hand, the third cut-point captures fluctuations and plateaus in a different set of counties. The fourth cut-point c_4 (and the corresponding explanation e_4) (Figs. 3(f) and 3(k)) is interesting: first, it captures a short rising outage trend (of smaller magnitude) at Dekalb, Fulton and Gwinnett counties in the North-West. Reports [3] suggest this is due to a separate tropical storm. At the same time, it also captures the start of the decrease in outages at Miami and Boward counties, both of which rise at the beginning in the first cut-point. Thus, CnR-UV can correctly capture the power restoration period of these counties (Miami, Boward) automatically. The last cut-point c_5 (and corresponding explanation e_5) at time step 93 (around September 12^{th}) captures the date when hurricane Irma was downgraded to a category 2 storm and the outages of the counties started to decrease. Note that these cut-points and explanations are non-trivial, and are successfully modeled since CnR-UV is able to capture the diverse trends in power outages of different magnitudes including in far away counties which do not follow the hurricane trajectory. As mentioned in Section 1, retrospective analysis of hurricanes through CnR-UV helps capture failure and restorative phases (ex: Miami, Boward counties) through *segmentation*, that can help experts understand grid resilience and restoration patterns. At the same time, CnR-UV *explanations* in addition to pin-pointing hurricane affected regions that incurred major power outages, can also uncover subtle trends in regions where consequential events occur, ex: like the tropical storm at Dekalb, Fulton and Gwinnett which was caught by CnR-UV explanations. Such insights can alert grid maintainers about the potential for such situations in future.

Spatial Clusters: Fig. 3(b) shows the spatial clustering of all counties affected by hurricane Irma (i.e. the clustering based on W_U where $W_U = UU^T$ is the spatial affinity matrix as explained in Section 4.2.1). It turns out that the green cluster contains counties most affected by power outages, whereas the red cluster shows the counties whose power outages were comparatively lower. This extent of the green cluster (towards the Western/North-Western part) is challenging to estimate by hand or through physical surveys, but has been uncovered by CnR-UV solely based on time-series dynamics and spatial constraints. This ultimately can aid disaster management experts and power companies to plan recovery for future hurricanes [16, 39].

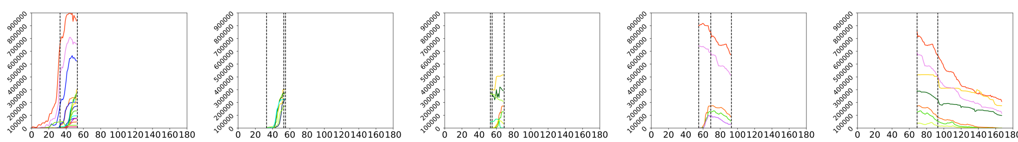
5.3.2 Hurricane Harvey. The CnR-UV results for hurricane Harvey are depicted in Fig. 4. The spatial depiction of explanations in Fig 4(g) - 4(j) broadly trace the trajectory of the hurricane along the eastern coast of Texas, with a few additional non-coastal counties also being highlighted as important in the Northern and North-Western parts of Texas.

In Fig. 4(a), the first and the second cut-point might seem redundant owing to their close proximity, and their both capturing increasing outage trends. However, upon closer investigation, we find that the first cut-point is detected when there is a sharp spike in El Paso and Howard counties at the very beginning of the hurricane. As no other counties have begun to experience outages at this



(a) Irma CnR-UV segmentation

(b) Spatial cluster



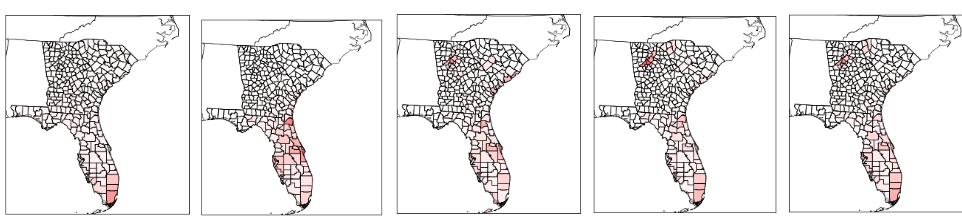
(c) Important timeseries for c_1

(d) Important timeseries for c_2

(e) Important timeseries for c_3

(f) Important timeseries for c_4

(g) Important timeseries for c_5



(h) Explanation e_1

(i) Explanation e_2

(j) Explanation e_3

(k) Explanation e_4

(l) Explanation e_5

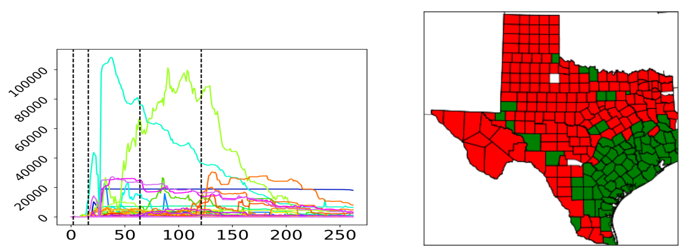
Fig. 3. Segmentation and the corresponding explanations for Irma. (a) shows all the counties having grid failures during Hurricane Irma. Each county is represented by a timeseries with an individual color in solid line. The vertical dashed lines are the cutpoints obtained by CnR-UV. (b) Spatial clustering result showing the spatial span of grid failure, based on spatial proximity of counties and similarity in failure patterns of their time series. (h)-(l) e_i visualizations in geographic space for each cut-point. Counties with higher e_i values (higher values represented by darker red) are more important for the cut-point, and are marked with a color closer to red. (c)-(g) The most important time series for each cut-point in the segmentation obtained from e_i whose explanation weight > 0.1.

point, even the relatively low absolute values of power outages (around 1600 homes) are captured by our segmentation model, leading to the first cut-point.

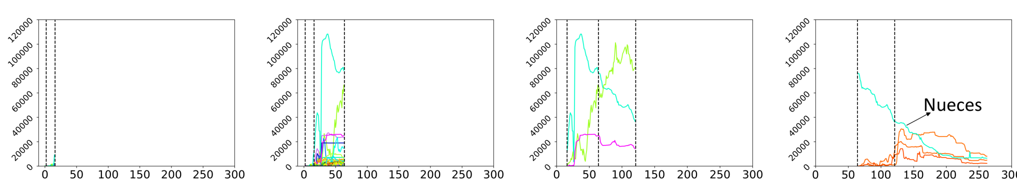
The spatial explanation Fig. 4(g), depicts as important a few disconnected counties in the Northern and North-Western part of Texas which might seem counter-intuitive at first. However, reports [1, 4] suggest that these counties were hit by floods as a result of hurricane Harvey causing major damage.

Similarly, the explanation of the second cut-point (Fig. 4(d)) highlights the spike in outages at Nueces and Aransas (around $100,000$ homes), but also captures Fort Bend, Brazoria and Harris (Fig. 4(h) group of three counties highlighted in the Northern part of the East coast of Texas) as important. Although their outages are low (around 2000) compared to Nueces, they have a very sharp peak at this cut-point. Report [5] states that the sudden rise of this peak is due to an EF1 level tornado on August 26, which caused major damage at Fort Bend also potentially affecting the surrounding counties.

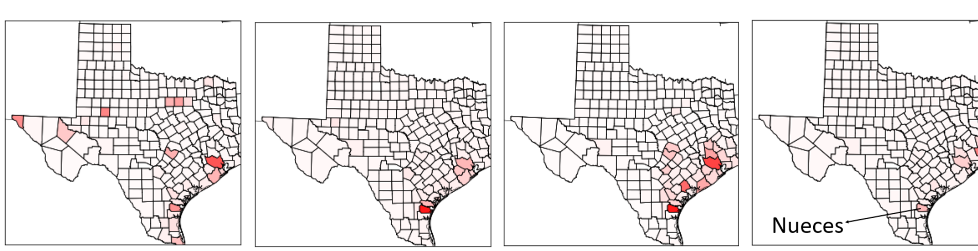
The explanation for the third cut-point (Fig. 4(e)) captures two different patterns, the high outage spike of Harris county (green line) and the declining trend at Victoria and Nueces counties. Finally, for the last cut-point (Fig. 4(f)) while the outages in many counties are decreasing, our algorithm correctly highlights a sudden rise of outages in counties Orange, Jefferson, Hardin. The main reason for this increase is due to the rising water of the Neches river, which causes the city to lose service from its major pump stations. As in case of Irma, spatial clustering results for hurricane Harvey (Fig. 4(b)) also help us glean the overall picture of the spread of damage due to the hurricane. This explanation is beneficial for doing an inter-dependency study. Note that CnR-UV captures long range county dependencies even if the counties are not geographically close to each other; such information of subtle county relationships is often buried deep in the original set of hundreds of time series and cannot be uncovered through simple models or through rudimentary visual or statistical analysis of the original data.



(a) Harvey CnR-UV segmentation (b) Spatial cluster



(c) Important timeseries for c_1 (d) Important timeseries for c_2 (e) Important timeseries for c_3 (f) Important timeseries for c_4



(g) Explanation e_1 (h) Explanation e_2 (i) Explanation e_3 (j) Explanation e_4

Fig. 4. Segmentation and the corresponding explanations and spatial clustering for Hurricane Harvey obtained by CnR-UV analogous to Fig 3. See detailed discussions in Section 5.3.2.

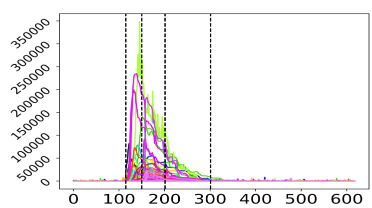
5.3.3 Hurricane Matthew. Similar to previous results, CnR-UV is able to extract insightful cut-points and explanations of all the major regimes of hurricane Matthew.

Cut-point 1 (Fig. 5(c)): captures the phase of hurricane landfall (Oct 2). However, CnR-V does not capture the bottom most southern county depicted in Fig. 5(g), whereas CnR-UV successfully captures this thereby yielding a more holistic explanation of the cut-point.

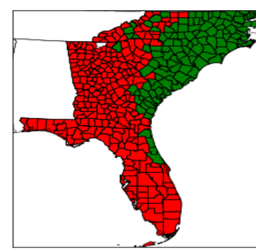
Cut-point 2 (Fig. 5(d)): This cut-point is detected because of the high rise of peak of outage in Chatham (Oct 4). The spatial representation of the explanation in Fig. 5(h) highlights counties along the trajectory of the hurricane.

Cut-point 3 (Fig. 5(e)) : This cut-point captures high decrease of outages which captures the restoration of Chatham, Duval etc. At the same time this cut-point is capturing sudden rise of outages of Horry county which is colored as bright red (in Fig. 5(i)) on top right. This county is influenced in the previous cut-point and has now severely affected (after Oct 4) and the influence has spread to nearby counties as well.

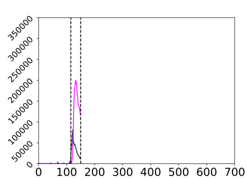
Cut-point 4 (Fig. 5(f)): This cut-point was captured when power outages of the counties started to abate. The explanation results Fig. 5(j) show the important counties whose outages started to decrease at this cut-point.



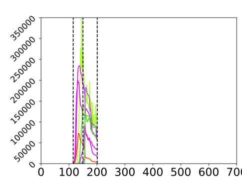
(a) Matthew CnR-UV segmentation



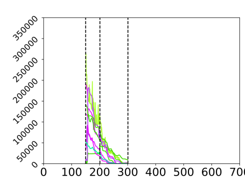
(b) Spatial cluster



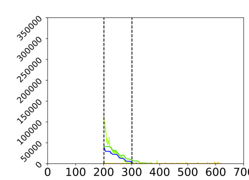
(c) Important timeseries for c_1



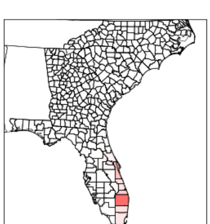
(d) Important timeseries for c_2



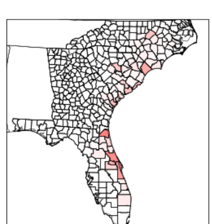
(e) Important timeseries for c_3



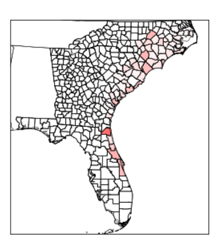
(f) Important timeseries for c_4



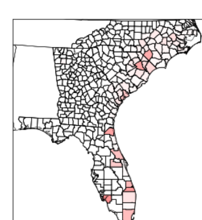
(g) Explanation e_1



(h) Explanation e_2



(i) Explanation e_3



(j) Explanation e_4

Fig. 5. Segmentations and the corresponding explanations and spatial cluster results for Hurricane Matthew obtained by CnR-UV analogous to Fig 3.

In addition, interestingly, although both hurricanes Irma and Matthew have similar geographic trajectories, CnR-UV learns very different spatial clusters which captures counties with variable dynamics (Fig. 5(b)).

5.3.4 Details of baseline algorithms on hurricane data. In contrast to the CnR-UV performance, the baseline algorithms all consistently either fail to converge or under-segment giving low quality unexplainable cut-points. *TICC, Autoplait* under segment on some and fail to converge in the case of other hurricane datasets, while *Dynammo* yields over-segmented results. *Floss*, while avoiding convergence and over-segmentation problems yields segments that only capture the initial rise and final fall of the time series in the case of all the hurricanes, completely missing out phases of power failure in-between.

5.4 Comparison with CnR-V

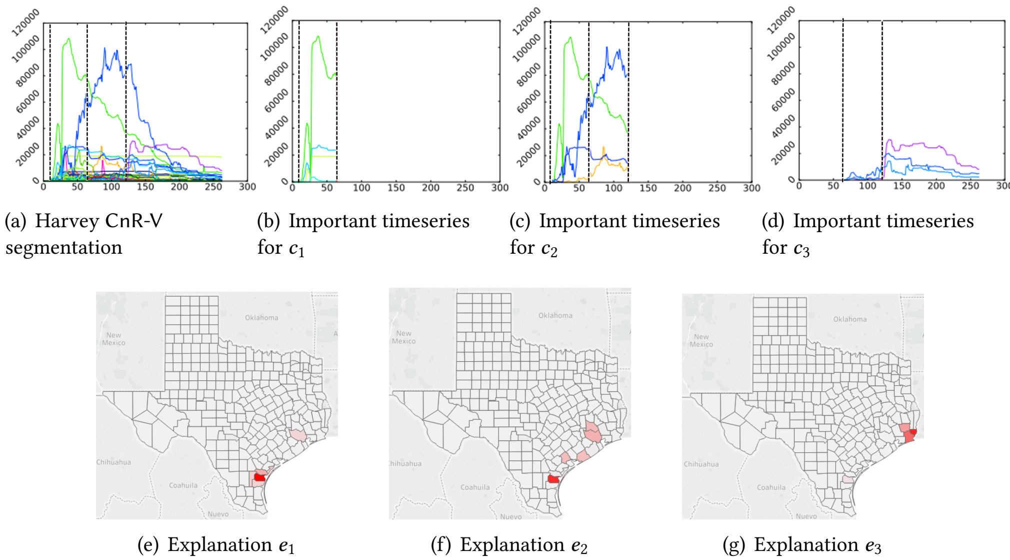
We characterized the performance of CnR-V on hurricane Harvey, Irma, Matthew power outage data, where there are long-range spatial dependencies, and found that CnR-V gives lower quality cut-points and explanations as expected. As an example, see Fig. 4(e) and Fig. 4(i): while CnR-V is able to capture this cut-point (Fig. 6(d) and Fig. 6(g)), its explanations only point to the sudden rise in a small cluster of spatially close counties; it fails to capture the large decrease in Nueces county (cyan line) (which CnR-UV is able to) because Nueces is not geographically close to the other ones.

5.4.1 Comparison results of hurricane Harvey. Cut-point 1 (Fig. 4(c)) is only captured by CnR-UV and not captured by CnR-V. In this cut-point, El Paso, Howard, and other counties in central Texas were considered important in Fig. 4(g) because they were flood affected due to hurricane Harvey. Cut-point 2 (Fig. 4(d)) captured by CnR-UV is similar to the first cut-point of Fig. 6(b) captured by CnR-V. However, some counties Fort Bend, Harris, and Brazoria are also considered important (from reports, it was found some major damage occurred in those spots due to a tornado see details in Sec 5.3.2) by CnR-UV but not captured by CnR-V. Cut-point 3 (Fig. 4(e)) by CnR-UV is similar to the cut-point depicted in Fig. 6(c) discovered by CnR-V. Montgomery (orange line) is also considered as an important county in CnR-V but not highlighted in Fig. 4(e) by CnR-UV. Upon further investigation, we found that this county did not face any major damage around this time, but it was showed important by CnR-V only because it is geographically close to Harris. Cut-point 4 (Fig. 4(f)) discovered by CnR-UV is similar to the cut-point depicted in Fig. 6(d) discovered by the CnR-V model. However Fig. 6(d) does not capture Nueces (green line) which has a high decrease of outage. Hence, CnR-V only captures increase of outages at this cut-point while CnR-UV captures both increasing and decreasing trends simultaneously.

5.4.2 Comparison results of hurricane Irma. For better understanding the figures of CnR-V are shown in Fig 7 for segmentation and explanation. CnR-V could not capture the small rising trend in Fig 3(d) and fluctuation of outages in Fig 3(e) separately as CnR-UV. CnR-V only identifed a cut-point near 50 (Fig 7(c)) which did not explain the counties which had fluctuation of outages (see Sec 5.3.1 for detail description). Moreover if we compare the geographic explanation of counties of CnR-V Fig 7(f)- 7(i) with CnR-UV in Fig 3(h)- 3(l) we observe CnR-V could not capture the long range spatial dependencies of counties and they were sparse.

5.4.3 Comparison results of hurricane Matthew. We were unable to run the entire Matthew dataset on CnR-V, and hence considered a sub-sampled version to obtain cut-points using CnR-V. We notice that CnR-V is unable to capture cut-point 1 (Fig 5), where it fails to capture a few important counties (Fig. 5(g)) on the southern tip of Florida which are captured by CnR-UV. It must be noted that in the case of each hurricane, there is no notion of spatial clustering in the case of CnR-V and spatial clusters similar to those represented in Fig. 3(b), 4(b) and 5(b) are obtained only by CnR-UV.

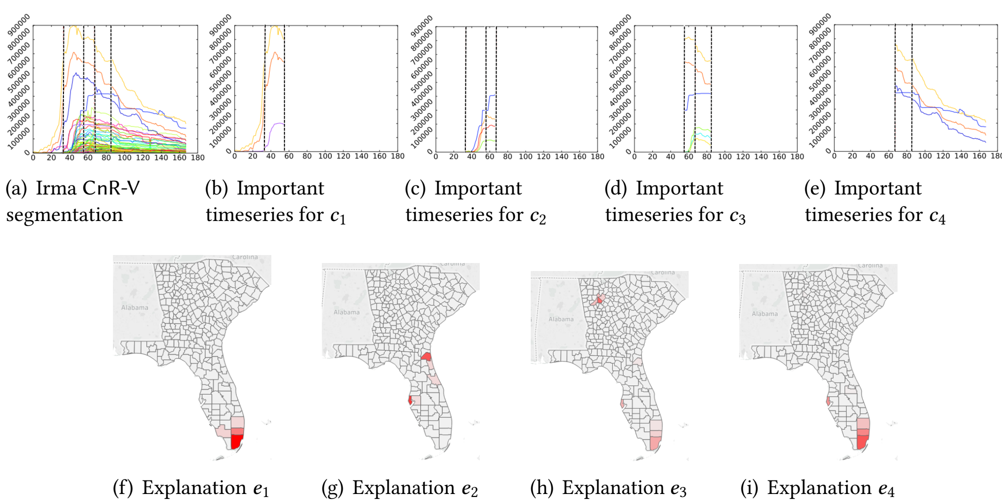
5.4.4 Scalability Comparison. We also recorded the run-times (in seconds) for CnR-V and CnR-UV after varying both the number of time-steps and time-series in the dataset. We get better scaling in practice (from our worst-case complexities): in both cases CnR-UV (due to the low-dimensional



(a) Harvey CnR-V segmentation (b) Important timeseries for c_1 (c) Important timeseries for c_2 (d) Important timeseries for c_3

(e) Explanation e_1 (f) Explanation e_2 (g) Explanation e_3

Fig. 6. Segmentations and explanations results of CnR-V for hurricane Harvey analogous to Fig 3.



(a) Irma CnR-V segmentation (b) Important timeseries for c_1 (c) Important timeseries for c_2 (d) Important timeseries for c_3 (e) Important timeseries for c_4

(f) Explanation e_1 (g) Explanation e_2 (h) Explanation e_3 (i) Explanation e_4

Fig. 7. Segmentations and explanations results of CnR-V for hurricane Irma analogous to Fig 3.

latent factor representation), scales quadratically, while CnR-V is significantly more expensive (sub-cubic in the number of time steps). We performed two kinds of experiments, one wherein the number of time-steps in the dataset was maintained constant (720 timesteps) while the number of time-series were varied (Table 4) and the other where we maintained the number of time-series constant (15 time-series) and the number of time-steps were varied (Table 5). The results indicate that in both cases CnR-UV is more scalable with increasing number of time-series and increasing number of time-steps. This can be attributed to our choice of representing U and V , in Eq. 5 as low dimensional latent factor weight matrices instead of full square matrices where $U \in \mathbb{R}^{n \times n}$ and $V \in \mathbb{R}^{m \times m}$ as in the case of CnR-V. We have also included scalability comparisons of CnR-UV with other state-of-the-art baselines in the appendix [6]. In these comparisons, we noticed that CnR-UV scaled equally as well as the FLOSS model and better than the TICC, Dynammo models with increasing number of time series. CnR-UV scales quadratically in the number of timesteps; future work will be aimed at caching and smart computational strategies to scale it to larger datasets.

Table 4. Scalability Experiment varying no. of timeseries keeping timesteps constant at 720. (Wall clock time seconds.)

No. of Timeseries	CnR-V	CnR-UV
15	734.37	179.07
30	733.39	180.35
60	822.99	173.44
120	835.65	184.50
240	863.03	190.47
480	952.81	210.45

Table 5. Scalability Experiment varying no. of timesteps keeping timeseries constant at 15. (Wall clock time seconds.)

No. of Timesteps	CnR-V	CnR-UV
500	285.19	64.63
1500	7388.49	115.38
2500	25755.02	3660.25

5.5 Summary of Observations

(1) CnR-UV consistently outperforms the baseline algorithms (upto *0.79x*) in all datasets (including with ground truth) for both time series segmentations and explanations (for hurricanes, the baselines heavily undersegment or do not even finish).
(2) For hurricane datasets, CnR-UV discovers non-trivial cut-points capturing the overall trajectory, as well as subtle anomalies like a combination of sudden increasing and/or decreasing cluster of outage trends and plateaus across regions. CnR-UV also discovers useful spatial clusters of counties based on their location and outages.
(3) Most importantly, we are also able to identify an informative set of culprit time series for each cut-point, providing valuable insights to the domain experts aiding management, recovery and resource allocation efforts.
(4) CnR-UV scales quadratically with the number of time-steps as opposed to CnR-V which scales sub-cubically.

6 RELATED WORK

We will now review lines of research that have attempted to answer questions similar to our goals in this paper.

Time Series Segmentation. There has been an abundance of work on time series segmentation based on monitoring changing temporal patterns, such as modeling co-evolving time series and segmenting them using multi-level HMMs [35] on motion capture data, discovering patterns in data streams [46] using distributed video data, developing online algorithms for frequent sequence mining [36] on different application domain, i.e., robotics, wild life, and health monitoring, and time series segmentation using temporal mixture model and Bayesian information criterion on railway data [48] and Kalman filters on motion capture sequences and chlorine measurement data [32]. GOALIE [42], is another algorithm applied in the context of biological process data that produces segmentations of multivariate time series but its focus is on finding cutpoints where significant shifts of clusters (of time series) occur, in contrast, to CnR-UV wherein the focus is both on recovering the major segments and explaining the segments while leveraging the underlying spatial structure of the data.

Change point detection has also been a popular topic in the climate sciences [43, 49]. Characterizing the dynamics of natural disasters like hurricanes lends itself naturally to a change point detection approach but there has been little work conducted in this regard. Zhao and Chu [58] propose a hierarchical Bayesian framework for detecting shifts in annual hurricane counts while Ruggieri [47] introduces a Bayesian change point detection algorithm to detect changes in temperature using climate data records. The other line of work on modeling failure cascading on CIS [18] does not explicitly segment the time series. Despite the extensive research conducted in time series segmentation and change point detection, we have found that there exists little prior work in leveraging them to characterize the dynamic effect natural disasters have on critical infrastructure systems.

Two limitations of previous work in temporal segmentation are that not many of them easily incorporate spatial information into temporal segmentation and none of the existing models provide any explanation framework wherein ‘culprit‘ counties (at each segment of a hurricane failure process) can be identified in space AND time. Most change point work is focused only on identifying temporal cut points [9] with a few applications in computer vision and video analysis modeling spatial relationships [14, 51] but no work has identified important time series (i.e spatial) thereby providing an explanation of each identified temporal segment. Such a model that identifies ’culprit’ counties is helpful to experts involved in maintenance of cyber-physical infrastructure and teams responsible for disaster management and planning.

Another line of related work in time series corresponds to subspace clustering based techniques. Many applications in multivariate time series analysis exist wherein the temporal data is drawn from multiple spaces and hence exhibits multi-segment behavior. It is often useful to develop techniques to represent the data in a subspace to capture richer temporal relationships and apply clustering to explicitly demarcate these multiple segments. This approach called subspace clustering has been applied to video and image segmentation [33, 52, 55], image compression [29] and spatio-temporal action segmentation [20, 31]. A comprehensive review about the different types of subspace clustering methods is provided by Vidal in [53]. There has been extensive work in subspace clustering in data mining [40] but to the best of our knowledge, it has not been applied on the hurricane outages data for finding the temporal relation among time steps. Further, these subspace clustering techniques also do not provide explanations of the results.

Simple Interpretable Models. There has recently been a push towards quantifying model uncertainty [22] and making machine learning model outputs quantifiable, explainable and simple [45]. Sangdeh et al. presented literature where they designed several quantitative and qualitative experiments to investigate the impact of features and model transparency on model prediction, a measure of trust and explainability [41]. These models and their explanations are specific to the underlying machine learning models and cannot be applied to our segmentation problem. We find that temporal segmentation is inherently unsupervised and the intuition behind the segments might not be readily apparent or explainable in certain applications. To the best of our knowledge, our explanation optimization problem is the first attempt at designing simple explanations for time series segments. In such cases, producing interpretable, simple segmentation results are effective in addressing the explainability problem.

Spatio-temporal Models. In Yao et al.[56], the authors develop Spatio-Temporal Dynamic Network (STDN) for traffic flow prediction. Our model CnR-UV is designed to recover explainable segments of the major failure phases in data containing bursty time series that don’t contain periodic, cyclical effects like traffic flow patterns. In Wu et al.[54], the authors propose an Urban Anomaly PreDiction (UAPD) model with a change point detection facet to detect evolving anomaly patterns. In contrast, CnR-UV, in addition to detecting the major change points (i.e segments) of the data, is also able to return sparse explanations about each retrieved change point yielding a holistic representation of the change point.

In this paper, we propose a dual-objective segmentation framework designed to provide spatio-temporal segments of the data and simple explanations of the generated segments. Our proposed framework optimizes the segmentation and explanation to obtain simple interpretable and sparse segmentations of the data. We demonstrate our model on the dynamic degradation of critical infrastructure during natural calamities. To the best of our knowledge this is the first attempt toward designing simple explainable segments of time series data.

7 CONCLUSIONS

In this paper, we have developed a novel effective and scalable combined framework CnR for providing segmentations and simple interpretable explanations for multivariate time-series like outage data. We evaluated the performance of our methodology against state of the art segmentation and time series clustering procedures on open ground truth datasets. We have also conducted an extensive analysis on the failure of the power grid during three hurricane events. There are many avenues for future work. Methodologically, we can explore performing a joint learning of segmentations and explanations, and more complex explanations. We can also explore integrating CnR with existing analysis tools, such as the URBANNET toolkit [18] in use at national labs.

Acknowledgements: This article is based on work partially supported by the NSF (Expeditions CCF-1918770, CAREER IIS-1750407, DGE-1545362, IIS-1633363), the NEH (HG-229283-15), ORNL, (H98230-14-C-0127), and a Facebook faculty gift. We also thank all the reviewers, whose comments and suggestions helped to improve the manuscript.

REFERENCES

[1] [n. d.]. Boston Herald Hurricane Harvey. http://www.bostonherald.com/news/national/2017/08/harvey_lashes_texas_coast_with_high_wind_torrential_rain.
[2] 2014. CMU Graphics Lab Motion Capture Database. http://mocap.cs.cmu.edu.
[3] 2017. Georgia Power Working to Restore Service to 161,000 in DeKalb. https://www.dekalbcountyga.gov/news/georgia-power-working-restore-service-161000-dekalb.
[4] 2017. Major Storm System Impacting Holiday Travel Through Friday. https://www.weather.gov/crp/hurricane_harvey.
[5] 2017. Severe Weather Event Review For Saturday August 26 2017. http://www.spc.noaa.gov/exper/archive/event.php?date=20170826.
[6] 2019. Appendix. https://bit.ly/2JVt8GP.
[7] 2020. Code and Datasets. https://github.com/anikat1/cnr-tist.
[8] M Allen, S Fernandez, O Omitaomu, and K Walker. 2014. Application of hybrid geo-spatially granular fragility curves to improve power outage predictions. *J Geogr Nat Disast* 4, 127 (2014), 2167–0587.
[9] Samaneh Aminikhanghahi and Diane J Cook. 2017. A survey of methods for time series change point detection. *Knowledge and information systems* 51, 2 (2017), 339–367.
[10] Francis Bach, Rodolphe Jenatton, Julien Mairal, Guillaume Obozinski, et al. 2011. Convex optimization with sparsity-inducing norms. *Optimization for Machine Learning* 5 (2011), 19–53.
[11] Oresti Banos, Mate Attila Toth, Miguel Damas, Hector Pomares, and Ignacio Rojas. 2014. Dealing with the effects of sensor displacement in wearable activity recognition. *Sensors* 14, 6 (2014), 9995–10023.
[12] Alan M Barker, Eva B Freer, Olufemi A Omitaomu, Steven J Fernandez, Supriya Chinthavali, and Jeffrey B Kodysh. 2013. Automating natural disaster impact analysis: An open resource to visually estimate a hurricane's impact on the electric grid. In *Southeastcon*. IEEE, 1–3.
[13] Richard H. Bartels and George W Stewart. 1972. Solution of the matrix equation AX+ XB= C [F4]. *Commun. ACM* 15, 9 (1972), 820–826.
[14] Durell Bouchard. 2006. Automated time series segmentation for human motion analysis. *Center for Human Modeling and Simulation,University of Pennsylvania: Pennsylvania, PA, USA* (2006).
[15] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning* 3,

1 (2011), 1–122.

[16] Liang Chang and Zhigang Wu. 2011. Performance and reliability of electrical power grids under cascading failures. *International Journal of Electrical Power & Energy Systems* 33, 8 (2011), 1410–1419.

[17] Liangzhe Chen, Sorour E Amiri, and B Aditya Prakash. 2018. Automatic Segmentation of Data Sequences. AAAI.

[18] Liangzhe Chen, Xinfeng Xu, Sangkeun Lee, Sisi Duan, Alfonso G Tarditi, Supriya Chinthavali, and B Aditya Prakash. 2017. HotSpots: Failure Cascades on Heterogeneous Critical Infrastructure Networks. In *CIKM*. ACM, 1599–1607.

[19] Eagle. 2012. Eagle-I. https://eagle-i.doe.gov/.

[20] Ehsan Elhamifar and René Vidal. 2009. Sparse subspace clustering. In *CVPR*. IEEE, 2790–2797.

[21] Rozhin Eskandarpour and Amin Khodaei. 2017. Machine learning based power grid outage prediction in response to extreme events. *IEEE Transactions on Power Systems* 32, 4 (2017), 3315–3316.

[22] Zoubin Ghahramani. 2015. Probabilistic machine learning and artificial intelligence. *Nature* 521, 7553 (2015), 452.

[23] Shaghayegh Gharghabi, Yifei Ding, Chin-Chia Michael Yeh, Kaveh Kamgar, Liudmila Ulanova, and Eamonn Keogh. 2017. Matrix Profile VIII: Domain Agnostic Online Semantic Segmentation at Superhuman Performance Levels. In *ICDM*. IEEE, 117–126.

[24] Gene Golub, Stephen Nash, and Charles Van Loan. 1979. A Hessenberg-Schur method for the problem AX+ XB= C. *IEEE Trans. Automat. Control* 24, 6 (1979), 909–913.

[25] David Hallac, Sagar Vare, Stephen Boyd, and Jure Leskovec. 2017. Toeplitz inverse covariance-based clustering of multivariate time series data. In *KDD*. ACM, 215–223.

[26] Paul Hines, Karthikeyan Balasubramaniam, and Eduardo Cotilla Sanchez. 2009. Cascading failures in power grids. *IEEE Potentials* 28, 5 (2009).

[27] Paul DH Hines, Ian Dobson, and Pooya Rezaei. 2017. Cascading power outages propagate locally in an influence graph that is not the actual grid topology. *IEEE Transactions on Power Systems* 32, 2 (2017), 958–967.

[28] Åke J Holmgren. 2006. Using graph models to analyze the vulnerability of electric power networks. *Risk analysis* 26, 4 (2006), 955–969.

[29] Wei Hong, John Wright, Kun Huang, and Yi Ma. 2006. Multiscale hybrid linear models for lossy image representation. *IEEE Transactions on Image Processing* 15, 12 (2006), 3655–3671.

[30] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* 42, 8 (2009).

[31] Colin Lea, Austin Reiter, René Vidal, and Gregory D Hager. 2016. Segmental spatiotemporal cnns for fine-grained action segmentation. In *ECCV*. Springer, 36–52.

[32] Lei Li, James McCann, Nancy S Pollard, and Christos Faloutsos. 2009. Dynammo: Mining and summarization of coevolving sequences with missing values. In *KDD*. ACM, 507–516.

[33] Junmin Liu, Yijun Chen, Jiangshe Zhang, and Zongben Xu. 2014. Enhancing low-rank subspace clustering by manifold regularization. *IEEE Transactions on Image Processing* 23, 9 (2014), 4022–4030.

[34] Jun Liu and Jieping Ye. 2010. Efficient l1/lq norm regularization. *arXiv preprint arXiv:1009.4766* (2010).

[35] Yasuko Matsubara, Yasushi Sakurai, and Christos Faloutsos. 2014. Autoplait: Automatic mining of co-evolving time sequences. In *SIGMOD*. ACM, 193–204.

[36] Abdullah Mueen and Eamonn Keogh. 2010. Online Discovery and Maintenance of Time Series Motifs *(KDD '10)*. 1089–1098.

[37] Nikhil Muralidhar, Chen Wang, Nathan Self, Marjan Momtazpour, Kiyoshi Nakayama, Ratnesh Sharma, and Naren Ramakrishnan. 2018. illiad: InteLLigent Invariant and Anomaly Detection in Cyber-Physical Systems. *ACM TIST* 9, 3 (2018), 35.

[38] Xia Ning and George Karypis. 2011. Slim: Sparse linear methods for top-n recommender systems. In *ICDM*. IEEE, 497–506.

[39] Min Ouyang. 2014. Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability engineering & System safety* 121 (2014), 43–60.

[40] Lance Parsons, Ehtesham Haque, and Huan Liu. 2004. Subspace clustering for high dimensional data: a review. *Acm Sigkdd Explorations Newsletter* 6, 1 (2004), 90–105.

[41] Forough Poursabzi-Sangdeh, Daniel G Goldstein, Jake M Hofman, Jennifer Wortman Vaughan, and Hanna Wallach. 2018. Manipulating and measuring model interpretability. *arXiv preprint arXiv:1802.07810* (2018).

[42] Naren Ramakrishnan, Satish Tadepalli, Layne T Watson, Richard F Helm, Marco Antoniotti, and Bud Mishra. 2010. Reverse engineering dynamic temporal models of biological processes and their relationships. *Proceedings of the National Academy of Sciences* 107, 28 (2010), 12511–12516.

[43] Jaxk Reeves, Jien Chen, Xiaolan L Wang, Robert Lund, and Qi Qi Lu. 2007. A review and comparison of changepoint detection techniques for climate data. *Journal of Applied Meteorology and Climatology* 46, 6 (2007), 900–915.

[44] Andreas Reinhardt, Paul Baumann, Daniel Burgstahler, Matthias Hollick, Hristo Chonov, Marc Werner, and Ralf Steinmetz. 2012. On the accuracy of appliance identification based on distributed load metering data. In *SustainIT*

(2012). IEEE, 1–9.

[45] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. Why should i trust you?: Explaining the predictions of any classifier. In *KDD*. ACM, 1135–1144.

[46] Guy Rosman, Mikhail Volkov, Dan Feldman, John W Fisher III, and Daniela Rus. 2014. Coresets for k-segmentation of streaming data. In *NIPS*. 559–567.

[47] Eric Ruggieri. 2013. A Bayesian approach to detecting change points in climatic records. *International Journal of Climatology* 33, 2 (2013), 520–528.

[48] Allou Samé and Gérard Govaert. 2012. Online Time Series Segmentation Using Temporal Mixture Models and Bayesian Model Selection. *ICMLA* 1 (2012), 602–605.

[49] Rishu Saxena, Layne T Watson, Randolph H Wynne, Evan B Brooks, Valerie A Thomas, Yang Zhiqiang, and Robert E Kennedy. 2018. Towards a polyalgorithm for land use change detection. *ISPRS journal of photogrammetry and remote sensing* 144 (2018), 217–234.

[50] Jianbo Shi and Jitendra Malik. 2000. Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence* 22, 8 (2000), 888–905.

[51] Yuliya Tarabalka, Guillaume Charpiat, Ludovic Brucker, and Bjoern H Menze. 2014. Spatio-temporal video segmentation with shape growth or shrinkage constraint. *IEEE Transactions on Image Processing* 23, 9 (2014), 3829–3840.

[52] Stephen Tierney, Junbin Gao, and Yi Guo. 2014. Subspace clustering for sequential data. In *IEEE CVPR*. 1019–1026.

[53] René Vidal. 2011. Subspace clustering. *IEEE Signal Processing Magazine* 28, 2 (2011), 52–68.

[54] Xian Wu, Yuxiao Dong, Chao Huang, Jian Xu, Dong Wang, and Nitesh V Chawla. 2017. Uapd: Predicting urban anomalies from spatial-temporal data. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 622–638.

[55] Allen Y Yang, John Wright, Yi Ma, and S Shankar Sastry. 2008. Unsupervised segmentation of natural images via lossy data compression. *Computer Vision and Image Understanding* 110, 2 (2008), 212–225.

[56] Huaxiu Yao, Xianfeng Tang, Hua Wei, Guanjie Zheng, and Zhenhui Li. 2019. Revisiting spatial-temporal similarity: A deep learning framework for traffic prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 5668–5675.

[57] Yinyu Ye and Edison Tse. 1989. An extension of Karmarkar's projective algorithm for convex quadratic programming. *Mathematical programming* 44, 1-3 (1989), 157–179.

[58] Xin Zhao and Pao-Shin Chu. 2006. Bayesian multiple changepoint analysis of hurricane activity in the eastern North Pacific: A Markov chain Monte Carlo approach. *Journal of climate* 19, 4 (2006), 564–578.

Cut-n-Reveal: Time-Series Segmentations with Explanations

NIKHIL MURALIDHAR, Virginia Tech, USA
ANIKA TABASSUM, Virginia Tech, USA
LIANGZHE CHEN, Pinterest, USA
SUPRIYA CHINTHAVALI, Oak Ridge National Laboratory, USA
NAREN RAMAKRISHNAN, Virginia Tech, USA
B. ADITYA PRAKASH, Georgia Institute of Technology, USA

CCS Concepts: • **Networks** → *Sensor networks*; • **Information systems** → **Data mining**; **Spatial-temporal systems**.

Additional Key Words and Phrases: Multivariate Time-series, Spatio-temporal Segmentation

ACM Reference Format:
NIKHIL MURALIDHAR, ANIKA TABASSUM, LIANGZHE CHEN, SUPRIYA CHINTHAVALI, NAREN RAMAKRISHNAN, and B. ADITYA PRAKASH. 2019. Cut-n-Reveal: Time-Series Segmentations with Explanations. *ACM Trans. Web* 9, 4, Article 39 (January 2019), 3 pages. https://doi.org/0000001.0000001

1 APPENDIX

1.1 Additional Discussion About CnR-UV Explanation Formulation

Our goal with learning each e_i vector is to learn a *local* explanation of 'culprit' time series at a particular cut-point i in the segmentation. Hence, our design is a deliberate attempt to uncover local explanations per cut-point, free of untoward global temporal influence from other cut-points. In addition, optimizing each e_i separately also enables parallelization aiding in scalability of the explanation formulation.

1.2 CnR-UV Explanation Formulation Compared with Attention Mechanisms

We may assume that the explanation step (that identifies the 'culprit' time series) in our case, represents a kind of attention mechanism (popular in encoder-decoder architectures) over each time series for each cut point. This is because the explanation model essentially can be considered as a learned scoring function that given a cut point c and a set of time series (similar to a set of encoder hidden states $\mathbf{H}$), learns a scoring function $\mathbf{S}$ that assigns scores to each of the time series in $\mathbf{H}$ considering their behavior around the cutpoint c_i. However, our mechanism is slightly more sophisticated than traditional attention mechanisms as we are also able to jointly model spatial constraints between counties, while learning the explanation weights as opposed to traditional

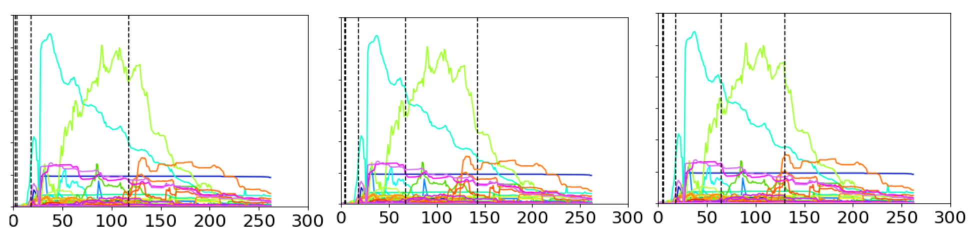
Authors' addresses: NIKHIL MURALIDHAR, Virginia Tech, USA, nik90@vt.edu; ANIKA TABASSUM, Virginia Tech, USA, anikat1@vt.edu; LIANGZHE CHEN, Pinterest, USA, liangzhechen@pinterest.com; SUPRIYA CHINTHAVALI, Oak Ridge National Laboratory, USA, chinthavalis@ornl.gov; NAREN RAMAKRISHNAN, Virginia Tech, USA, naren@cs.vt.edu; B. ADITYA PRAKASH, Georgia Institute of Technology, USA, badityap@cc.gatech.edu.

1559-1131/2019/1-ART39 $15.00
https://doi.org/0000001.0000001

attention mechanisms which use straight forward similarity functions like cosine similarity to calculate attention energies.

1.3 Effect of Changing Latent Dimension on CnR-UV Segmentation

In Fig. 1 we show CnR-UV segmentation on Hurricane Harvey with varying values of latent dimension l. The segmentation results don't vary much with varying l, keeping other hyperparameters constant. This indicates that the CnR-UV segmentation model is not overly sensitive to variation in the latent dimension. Having higher dimension of l gives the model more expressive power which could lead to model overfitting. However, in this case, greater sparsity can be achieved by controlling the appropriate regularization terms in CnR-UV. It must be noted that the latent factor dimension across all our experiments is fairly small, as we have employed $l \leq 5$.



(a) l=2 (b) l=3 (c) l=4

(d) l=5 (e) l=6 (f) l=8

Fig. 1. CnR-UV segmentation results (vertical dashed lines) for the Hurricane *Harvey* varying l values to check robustness. The segmentation performance indicates that the CnR-UV model is not overly sensitive to changes in l. For the actual experimental results on all hurricanes we used $l = 5$.

1.4 Additional Scalability Comparison of CnR-UV with Baseline Models

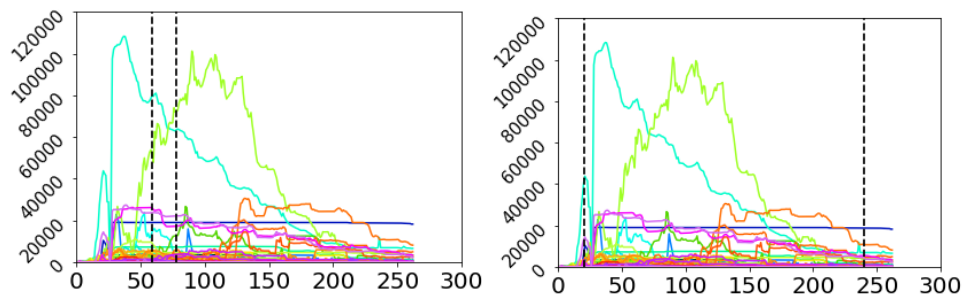
We recorded the running time of the baselines TICC, Dynammo, and Floss in Table 1 varying number of timeseries. Dynammo did not converge when number of timeseries ≥ 30.

Table 1. Scalibility Experiment on Baselines Varying Number of Timeseries

Number of timeseries	CnR-UV	TICC	Floss	Dynammo
15	179.07	6.69	5.63	44.28
30	180.35	25.97	10.66	–
60	173.44	99.88	21.07	–
120	184.50	342.81	41.97	–
240	190.47	771.20	84.12	–
480	210.45	3855.13	167.03	–

1.5 Baseline Results on Hurricane Dataset

We show the results of two best performing baseline models TICC and Floss on segmenting the hurricane Harvey power failure data in Fig. 2. TICC did not converge for hurricane Irma, and Matthew while Floss gives similar segments as in Harvey for both the hurricanes, i.e., one at the very beginning, other at the very end.



(a) TICC segments (b) Floss segments

Fig. 2. TICC and FLOSS segmentation results (vertical dashed lines) for the Hurricane *Harvey*. We noticed that both these models do not capture many important phases of the hurricane failure process like the CnR-UV model does (see Fig. 1).

1.6 Hyperparameter Values of Baselines

We note all the baseline hyperparameter values in Table 2. For TICC, we used window-size= 5,λ = $11e - 2$, $\beta = 5$, threshold= $2e^{-5}$

Table 2. Baseline Hyperparameter Values

Dataset	**TICC** Number of Clusters	**Floss** Sub-Sequence Length	**Dynammo** Error Threshold
Synthetic	5	4	50
NILM	8	4	45
ChickenDance 1	8	5	73
ChickenDance 2	8	4	35
WalkJog 1	3	4	95
WalkJog 2	3	4	65
GrandMal	2	10	85

"The supplementary material contains the appendix where we have additional results of CnR-UV performance with respect to baseline models primarily focused on scalability comparison. We also evaluate the segmentation performance of popular segmentation baselines and finally evaluate the effect of latent dimension on segmentation performance of CnR-UV."