



Entropi Ölçülerinin Kestirim Başarımının İncelenmesi

Performance Investigation of the Entropy Measures Estimations

Özge Canlı^{1*}, **Serkan Günel¹**

¹ Dokuz Eylül Üniversitesi Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği, İzmir, TÜRKİYE

Sorumlu Yazar / Corresponding Author *: ozge.canli@deu.edu.tr

Geliş Tarihi / Received: 06.05.2020

Kabul Tarihi / Accepted: 05.02.2021

Araştırma Makalesi/Research Article

DOI:10.21205/deufmd.2021236808

Atıf şekli/How to cite: CANLI Ö., GÜNEL S. (2021). Entropi Ölçülerinin Kestirim Başarımının İncelenmesi. DEÜ FMD 23(68), 435-457.

Öz

Sistemler arasındaki bilgi akışını sistem durumları kullanılarak entropiden türetilen ölçüler yardımı ile nicelendirebiliriz. Bu çalışmada birbirine bağlanmış Gauss süreçlerinin ortak bilgi ölçüsü, aktarım entropisi ve iki durumlu sistemin entropisi bu amaçla kestirilmiştir. Sistemlerin bilgi ölçüleri histogram, çekirdek yoğunluk kestirimcisi, k-en yakın komşuluklu entropi kestirimcisi ve kpN entropi kestirimcisinin farklı parametreleri için elde edilmiştir. Ölçülerin kestirim başarımını belirlemek için normalize olmuş kestirim yanlılık ve normalize olmuş standart hata değerleri hesaplanmıştır. Kestirimcilerin en iyi parametreleri bu ölçüler için saptanmıştır. k-en yakın komşuluklu kestirimcisinin birbirine bağlanmış Gauss süreçleri için en iyi performansla sahip olduğu gösterilmiştir. Buna ek olarak, doğrusal olmayan iki durumlu sistemin entropisi kestirilmiştir. Çekirdek yoğunluk kestirimcisi ve kpN entropi kestirimcisi bu sistem için entropi kestiriminde başarılı olduğu, histogram ve k-en yakın komşuluklu entropi kestirimcilerinin normalize olmuş kestirim yanlılığının ve normalize olmuş standart hata oranlarının yüksek olduğu gösterilmiştir. Entropi kestirimlerinin bulunan en iyi parametreleri kullanıldığında birbirine bağlanmış dinamik sistemlerin entropi kestirimleri yüksek doğrulukla kestirilir. Aktarım entropisi bahsedilen entropi yöntemleri kullanılarak gerçek bir veri kümesi için kestirilmiş ve hipotez testleri yardımı ile aktarım entropisini kestiren yöntemlerin istatistiksel önem dereceleri incelenmiştir. Bu çalışma sistemlerin belirsizliği ve sistemler arasındaki bilgi akışını temsil eden bilgi ölçülerinin parametrelerini uygun şekilde seçmemiz için bir yol sunar.

Anahtar Kelimeler: Entropi, Ortak bilgi ölçüsü, Aktarım entropisi, Entropi kestirimcileri, Gauss süreçleri, İki durumlu sistem

Abstract

Information flow between the systems can be quantified by using the measures that are derived from entropies by observing the state variables. In this study, the mutual information and the transfer entropy of the coupled Gaussian processes and the entropy of the bistable system have been estimated for this purpose. Information measures of the systems have been obtained by the histogram, kernel density estimator, k-nearest neighbor entropy estimator and kpN entropy estimator for the different parameters. Normalized bias and normalized standard error have been calculated to determine the performance of the estimation of the entropy measures. The optimal parameters of the entropy estimators of information measures have been obtained. It is shown that

k-nearest neighbor entropy estimator has the best performance for the coupled Gaussian processes. Additionally, the entropy of a nonlinear system i.e. the bistable system has been estimated. While kernel density estimator and k-nearest neighbor entropy estimator have succeeded in achieving the estimation of the entropy in the bistable system, it is shown that normalized bias and normalized standard error of histogram and k-nearest neighbor method have high ratios. The entropy estimations of the coupled dynamical systems have been estimated with considerable accuracy when the optimal parameters of the entropy estimations have been used. The transfer entropy of a real dataset has also been estimated via these methods. We have discussed the statistical significance of the estimation methods of transfer entropy by using a hypothesis test. The study provides a way to determine information measures that quantify the uncertainty of the systems and information flows between the systems, accurately.

Keywords: Entropy, Mutual information, Transfer entropy, Gaussian processes, Entropy estimators, Bistable system

1. Giriş

Shannon bilgi kuramına göre entropi rasgele bir değişkeni gözlemleyerek elde edilen bilgi miktarını veya kaldırılan belirsizliğini ölçer [1]. Entropi veri sıkıştırma ve kodlama ya da kanal kapasitesi hesaplama alanlarında sıklıkla kullanılır. Literatürde, entropi temel alınarak türetilen pek çok ölçü vardır. Bu ölçülerden biri olan ortak bilgi ölçüsü (mutual information), rasgele bir değişkeni gözlemlenmesi ile ilişkili diğer bir rasgele değişkenin belirsizliğindeki azalmayı saptar [2]. Bir diğeri, aktarım entropisi (transfer entropy), iki rasgele değişkenin arasındaki bilgi akışını sistemin dinamik olasılıklarını da göz önüne alarak saptayan bilgi ölçüsüdür [3]. Bu ölçüler kullanılarak özellikle doğrusal olmayan sistemler arasındaki bilgi akışının nicelendirebildiği bilinmektedir [4]–[6].

Aktarım entropisi endüstriyel süreçlerin bilgi akışını ölçmede ve sistemlerin birbirine neden sonuç ilişkisi ile bağlı olup olmadığını ayırtmada kullanılır [4]. Bu ölçü sayesinde sinirbiliminde elektroensefalografi (EEG) gibi ölçüm teknikleri kullanarak nöronlar arasındaki bilgi akışı tespit edilmeye çalışılmıştır [5]. Son yıllarda, aktarım entropisi sinirbilimindeki nöronların iletimini sağlayan alt süreçlerin veri akışını belirlemek için de kullanılan bir ölçü olmuştur [6]. Başta aktarım entropisi olmak üzere bilgi ölçülerini doğru kestirmek sistemler arasındaki bilgi akışını doğru ölçmek için önemli bir problemdir. Pratikte entropi kestirimcisinin seçiminde veriye uygun ölçünün seçilmesi, kullanılacak kestirimci tipinin belirlenmesi, veri kümesindeki doğrusal olmayan bileşenler ve veri kümesindeki uzun zamanlı korelasyon gibi belirsizliklerin saptanması önemlidir [7].

Entropi ve ondan türetilen ölçüler kestirilirken genellikle sonlu sayıda gözlemden elde edilen olasılık yoğunluk fonksiyonları kullanılır. Histogram ya da çekirdek yoğunluk kestirimi gibi parametrik olmayan kestirimci tiplerinde öncelikle olasılık yoğunluk fonksiyonları kestirilerek ölçüler hesaplanır. Parametrik olmayan kestirimci tiplerinde gözlemlenen veri düşük boyutta ise ve yeterli sayıda gözlem var ise, bilgi kuramı ölçüleri yüksek doğrulukla kestirilir. Ancak bu tip kestirimcilerde gözlemlenen veri sayısının yeterli sayıda olmaması durumunda ya da boyut sayısı arttıkça kestirim başarımının düştüğü bilinmektedir [9].

Alternatif olarak, komşuluk tabanlı entropi kestirim yöntemlerinde, sonlu sayıda nokta doğrudan kullanılarak entropi kestirilir [8]. Bunlardan en temel olanı k-en yakın komşuluk entropi kestirimcisi, verinin yerel olarak düzgün dağıldığını varsaymaktadır. Veri uzayında belirli bir yönlülük olması halinde bu varsayım gerçekliğini yitirmektedir. Bu kısıtlamanın gevşetilmesi için, $k_p N$ entropi kestirimcisi verinin yerel olarak Gauss dağılımına sahip olduğunu varsayar. Bu tip kestirimci ile bir veri noktasının civarındaki yerel dağılım Gauss çekirdek fonksiyonu ile kestirilirken, verinin genel karakteri k-en yakın komşuluk kestirimi ile belirlenmektedir. Genel olarak bakıldığında, komşuluk tabanlı kestirimci tipleri yerel olarak düzgün dağılımlı veriler için yüksek doğrulukla kestirilir ve parametrik olmayan kestirimci tiplerine göre entropi ve benzeri ölçüleri daha hızlı bir şekilde kestirir [9]. Öte yandan gözlemlenen veride bir yönlülük var ise komşuluk tabanlı kestirimcilerin başarımının düşük olduğu bilinmektedir.

Bilgi kuramı ölçüleri kestirilirken gözlemlenen nokta sayısı, olasılıkları hesaplamak için kullanılan aralık sayısı, komşuluk sayısı ya da çekirdek yoğunluk kestiriminin tipi gibi parametreler önemli rol oynar ve bu değişkenleri uygun aralıklarda ya da tiplerde belirlemek kestirimin doğruluğunu artırır. Bilgi akışını nicelendiren bu ölçülerin doğruluğunun yüksek olması önemlidir ve doğruluğu yüksek olan kestirimcinin seçilmesi gereklidir.

Literatürde özellikle aktarım entropisi ölçüsünün doğru kestiriminin zorluğu vurgulanmıştır. Her kestirimcinin kendine ait ince ayarları olduğundan herhangi bir veri kümesi için kestirilen aktarım entropisinin kestirim yönteminin alternatif en iyi çözümü olmadığından bahsedilmiştir. Bu amaçla; histogram, uyarlamalı aralıklı ve genişlikli histogram kestirimcisi ve çekirdek yoğunluk kestirimcileri kullanılarak birbirine bağlanmış Gauss süreçleri ve Lorenz dinamik sistemleri için aktarım entropisi kestirimleri yapılmıştır [10]. Xiong vd. [7], doğrusal, çekirdek yoğunluk ve k-en yakın komşuluklu entropi kestirimcilerin performanslarını entropi, koşullu entropi (conditional entropy) ve entropiden türetilen bilgi depolama (information storage) ölçüleri için karşılaştırılmıştır. Kestirimcilerin sapması ve sınırları, birbirine bağlanmış Gauss süreçleri için detaylıca incelenmiştir [7]. Son yıllarda aktarım entropisi Perron-Frobenius işlemleri kullanılarak da kestirilmiştir [11]. Olasılık ve olasılık yoğunluk fonksiyonlarını gözlemlenen noktalardan kestirmek yerine veriden Perron-Frobenius işlemlerinin değişmeyen dağılımları çıkarılmış ve aktarım entropisi hesaplanmıştır. Birbirine bağlanmış ayrık zamanlı ve sürekli zamanlı sistemlerin farklı gürültü seviyeleri için aktarım entropileri hesaplanmıştır. Önerilen yöntem k-en yakın komşuluklu entropi kestirimcisi ve çekirdek yoğunluk kestirimcisi ile karşılaştırılmıştır [11].

Bu çalışmanın amacı, doğrusal ve doğrusal olmayan sistemlerde bilgi ölçülerini kestirmek için kullanılan teknikleri kıyaslamak, bu sayede kestirimcilerin en iyi parametrelerini belirlemek ve bilgi ölçülerini en yüksek doğruluk oranı ile elde etmektir. Bu doğrultuda entropi, ortak bilgi ölçüsü ve aktarım entropisi ölçüleri; histogram, çekirdek yoğunluk fonksiyonu kestirimi yönteminin yanı sıra, k-en yakın komşuluk ve son yıllarda önerilen kpN entropi yöntemi yardımıyla kestirilmiştir. Histogramın

kestirilmesi yoluyla bilgi ölçülerinin kestirimi literatürde sıklıkla kullanılmaktadır [10, 19, 20]. Çekirdek yoğunluk kestirimi yöntemi ile entropinin hesaplanması parametrik olmayan kestirim yaklaşımına bir örnektir. Bu nedenle bu çalışmada iki farklı çekirdek yoğunluk fonksiyonu seçildiğinde kestirim başarımının nasıl değişeceği incelenmiştir. k-en yakın komşuluk yöntemi ile bilgi ölçüleri çok hızlı bir şekilde hesaplanır. kpN yöntemi çekirdek yoğunluk fonksiyonu kestirimi ve komşuluk tabanlı bir entropi yöntemini içeren kestirim yaklaşımına bir örnektir. Komşuluk tabanlı ve çekirdek yoğunluk-komşuluk birleşiminden oluşan yöntemlerin kıyaslamasını yapabilmek için k-en yakın komşuluk ve kpN yöntemleri incelenmiştir.

Kestirimcilerin doğruluğunu test etmek için bilgi ölçüleri teorik olarak bilinen birbirine bağlanmış Gauss süreçleri kullanılmış ve kestirimlerin parametrelere bağlı nasıl değişeceği detaylıca incelenmiştir. Buna ek olarak, doğrusal olmayan iki durumlu sistemin entropi kestirimleri yapılmıştır. Kestirimcilerin başarımları normalize olmuş kestirim yanlılığı ve normalize olmuş standart hata kriterleri ile ölçülecektir. Kestirilen bilgi ölçüleri için kestirim parametrelerinin doğru seçilmesiyle başarımların yüksek entropi kestirimi yapılabileceği tespit edilmiştir. Doğrusal ya da doğrusal olmayan sistemlerde en iyi performansa sahip entropi kestirimcileri gösterilmiştir. Son olarak, alta yatan dinamiğin bilinmediği gerçek bir veri kümesi için bahsedilen entropi kestirim yöntemleri kullanılarak aktarım ölçüsü kestirilmiştir. Bu durum için kestirim başarımların analizleri hipotez testleri yardımı ile yapılmıştır.

2. Bilgi Kuramı Ölçüleri

Verilen bir $S \subset \mathbb{R}^d$ üzerinde tanımlanan sürekli rasgele değişken X için diferansiyel entropi aşağıdaki şekilde tanımlanır [2]:

$$h(X) = - \int_S f_x(x) \log f_x(x) dx . \quad (1)$$

Burada, $f_x(x)$, X değişkeninin olasılık yoğunluk fonksiyonudur. Ayrıca $f_x(x)$ fonksiyonunun türevinin sonlu olduğu kabul edilmektedir ($|f'_x(x)| \leq L, \forall L \in \mathbb{R}^+$).

İki sürekli rasgele değişken için X_1, X_2 tanımlanan ortak bilgi ölçüsü, değişkenlerden

birini gözlemleyerek diğeri hakkında belirsizliği ölçmemize yarar [2]:

$$I(\mathbf{X}_1, \mathbf{X}_2) = h(\mathbf{X}_1) + h(\mathbf{X}_2) - h(\mathbf{X}_1, \mathbf{X}_2) \quad (2)$$

Burada $h(\mathbf{X}_1, \mathbf{X}_2)$, $(\mathbf{X}_1, \mathbf{X}_2)$ değişkeninin bileşik entropi tanımını ifade eder. Aktarım entropisi, iki rasgele değişken arasındaki bilgi akışını dinamik olasılıkları da göz önüne bulundurarak saptayan bilgi ölçüsüdür [3]. Gözlenen \mathbf{X}_1 çıkışı örneklenirse, \mathbf{X}_1 'in n anından $n - k$ ve $n - \ell$ anına kadar olan gösterim sırası

$$\mathbf{X}_{1,n}^{(k)} = [X_{1,n}, X_{1,n-1}, \dots, X_{1,n-k+1}],$$

$\mathbf{X}_{2,n}^{(\ell)} = [X_{2,n}, X_{2,n-1}, \dots, X_{2,n-\ell+1}]$ ile ifade edilir. Buna göre aktarım entropisi,

$$TE_{\mathbf{X}_1 \rightarrow \mathbf{X}_2} = h(\mathbf{X}_{2,n+1}, \mathbf{X}_{2,n}^{(\ell)}) + h(\mathbf{X}_{2,n}^{(\ell)}, \mathbf{X}_{1,n}^{(k)}) - h(\mathbf{X}_{2,n+1}, \mathbf{X}_{2,n}^{(\ell)}, \mathbf{X}_{1,n}^{(k)}) - h(\mathbf{X}_{2,n}^{(\ell)}) \quad (3)$$

marjinal ve bileşik entropi terimlerinin toplamı cinsinden yazılır. Burada, $X_{2,n+1}$ \mathbf{X}_2 değişkenin $n + 1$ anındaki gözlemini ifade eder. Ortak bilgi ölçüsü simetrik bir ölçüdür ($I(\mathbf{X}_1, \mathbf{X}_2) = I(\mathbf{X}_2, \mathbf{X}_1)$). Bilginin akışını yönlü olarak ifade etmez. Aktarım entropisi ise değişkenlerden birinin geçmiş değerlerini gözlemleyerek diğeri hakkında ne miktarda bilgi edinebileceğimizi ölçmeye yarayan pozitif yönlü bir bilgi ölçüsüdür ($TE_{\mathbf{X}_1 \rightarrow \mathbf{X}_2} > 0, TE_{\mathbf{X}_2 \rightarrow \mathbf{X}_1} > 0$). Genel olarak simetrik değildir ($TE_{\mathbf{X}_1 \rightarrow \mathbf{X}_2} \neq TE_{\mathbf{X}_2 \rightarrow \mathbf{X}_1}$). Sistemler arasındaki ortak bilgi ölçüsü ve aktarım entropisi ölçüleri, histogram kestirimcisi ve kpN entropi kestirimcisi yönteminde marjinal ve bileşik entropi terimlerinin toplamı kestirilerek elde edilecektir. Çekirdek yoğunluk kestirimcisi ve k-en yakın komşuluklu entropi kestirimcisinde ise, ortak bilgi ölçüsü ve aktarım entropisi ölçüsü doğrudan kestirilecektir.

3. Kestirimci Yöntemleri

θ teorik değeri bilinen ya da bilinmeyen kestirilmek istenen değişkeni ifade etsin. Sonlu sayıda özdeşçe dağıtılmış bağımsız X_1, X_2, \dots, X_n ölçümlerinden θ için bir kestirim yapılmak istensin. $T(\cdot)$ kestirimcinin fonksiyonunu göstermek üzere, $\hat{\theta} = T(X_1, X_2, \dots, X_n)$, θ değişkenine yakın olması beklenen nokta kestirimidir. $b = E[\hat{\theta}] - \theta = 0$ ise, $\hat{\theta}$, θ 'nın yansız bir kestirimcisidir. $b \neq 0$ ise kestirimci

yanlıdır denir. $\sigma_{\hat{\theta}}$ kestirilen $\hat{\theta}$ değerlerinin standart sapmasını gösterebilir. $\hat{\theta}$ kestirimcisinin ortalama karesel hatası (MSE)

$$MSE(\hat{\theta}) = b^2 + \sigma_{\hat{\theta}}^2, \quad \theta \neq 0 \quad (4)$$

ile tanımlanır. Açık ki, kestirimci yansız ise, $MSE(\hat{\theta}) = \sigma_{\hat{\theta}}^2$ 'dir. MSE özdeşçe dağıtılmış bağımsız örneklerle yapılan bir kestirimin başarımını ölçmek için standarttır. Başarımları ölçmek için, normalize olmuş kestirim yanlılığı (NKY), normalize olmuş standart hata (NSH) ve normalize olmuş ortalama karesel hatası (NMSE)

$$NKY(\hat{\theta}) = \frac{E[\hat{\theta}] - \theta}{\theta}, \quad \theta \neq 0 \quad (5)$$

$$NSH(\hat{\theta}) = \frac{\sigma_{\hat{\theta}}}{\theta}, \quad \theta \neq 0 \quad (6)$$

$$NMSE(\hat{\theta}) = \frac{E[(\hat{\theta} - \theta)^2]}{\theta^2}, \quad \theta \neq 0 \quad (7)$$

değerleri kullanılmaktadır [17]. Eğer kestirimci yansız ise, bu ölçüler farklı kestirimcileri kıyaslamak için kullanılabilir. NKY değerlerinin düşük olduğu kestirim parametreleri için $NSH(\hat{\theta}) \approx \sqrt{NMSE(\hat{\theta})} / \theta$ 'dir dolayısıyla kestirimcinin başarımlarını hakkında doğrudan bilgi verir. İki farklı kestirimcinin tam bir kıyaslamasını yapabilmek için NKY değerlerinin düşük olduğu parametrelerde NSH değerleri kıyaslanmalıdır.

Kestirilmek istenen olasılık yoğunluk fonksiyonu hakkında herhangi bir ön bilgi yoksa parametrik olmayan entropi kestirimcileri yardımıyla öncelikle olasılık yoğunluk fonksiyonu kestirilir. Kestirilen olasılık yoğunluk fonksiyonu kullanılarak entropi ve ondan türetilen ölçüler hesaplanır. Buna karşılık doğrudan kestirim yöntemlerinde, entropi ve entropiden türetilen ölçüler, gözlem noktalarını doğrudan kullanan k-en yakın komşuluk kestirimcisi ile elde edilir. Hem noktaları doğrudan kullanarak hem de yerel veriden Gauss çekirdek fonksiyonları kestirerek entropiyi elde etmek de mümkündür. Bu çalışmada parametrik olmayan entropi kestirimcilerinden sıklıkla kullanılan histogram ve çekirdek yoğunluk kestirimcileri ile k-en yakın komşuluklu entropi kestirimcisi ve kpN

entropi kestirimcileri aynı veri üstündeki kestirim başarımları açısından kıyaslamıştır.

3.1. Histogram kestirimcisi

$\mathcal{S} \subset \mathbb{R}^d$ olmak üzere ve \mathcal{S} kümesi ε çaplı birbiriyle örtüşmeyen $B_m \in M, m = 1, 2, \dots, M$ bölüntülerinden oluşsun. Bilinmeyen $f_x(\mathbf{x})$ olasılık yoğunluk fonksiyonunu N adet örnekten histogram yöntemi ile kestirmek istenirse

$$f_x(\mathbf{x}, B_m) = \sum_{B_m \in M} \frac{B_m \text{ içerisindeki nokta sayısı}}{N \text{ vol}(B_m)} \times I(\mathbf{x}_i \in B_m) \quad i = 1, 2, \dots, N \quad (8)$$

ile hesaplanır. Burada, $I(\cdot)$ belirteç fonksiyonudur, $\text{vol}(\cdot)$ ise B_m bölüntüsünün hacmini gösterir. Histogram kestirimcisi kullanılarak olasılık yoğunluk fonksiyonu kestirilirse aralık genişliği ve aralık sınırlarının doğru saptanması önemli bir rol oynar [12]. Kestirilen olasılık yoğunluk fonksiyonu $f_x(\mathbf{x}, B_m)$, denklem 8'te yerine koyularak entropi kestirilecektir:

$$\hat{h}(\mathbf{X}) = - \sum_{m=1}^M f_x(\mathbf{x}, B_m) \log f_x(\mathbf{x}, B_m). \quad (9)$$

3.2 Çekirdek yoğunluk kestirimcisi

Bu yöntemde olasılık yoğunluk fonksiyonu daha önceden tanımlanmış baz fonksiyonlarının ağırlıklı toplamlarından oluşan çekirdek yoğunluk fonksiyonları kullanılarak kestirilir. $\mathbf{X} \in \mathbb{R}^d$ rasgele değişkeninden N adet örnek alalım. Bu örneklerden çekirdek yoğunluk fonksiyonları yardımıyla olasılık yoğunluk fonksiyonu kestirilir. En genel halde d -boyutlu çekirdek yoğunluk kestirimcisi

$$f_x(\mathbf{x}, \mathbf{H}) = \frac{1}{N} \sum_{i=1}^N K_H(\mathbf{x} - \mathbf{x}_i) \quad (10)$$

ile tanımlanır [12]. Burada, \mathbf{H} $d \times d$ boyutlu kesin pozitif simetrik genişlik matrisini gösterir [12]. $K_H(\mathbf{x}) = |\mathbf{H}|^{-\frac{1}{2}} K\left(\mathbf{H}^{-\frac{1}{2}} \mathbf{x}\right)$, K fonksiyonundan oluşan çekirdek yoğunluk fonksiyonudur ve K fonksiyonu $\int K(\mathbf{x}) d\mathbf{x} = 1$ şartını sağlar. Gauss çekirdek tipi kestirimcinin çekirdek fonksiyonu

$$K(\mathbf{x}) = \left(\frac{1}{2\pi}\right)^{\frac{d}{2}} e^{-\left(\frac{1}{2} \mathbf{x}^T \mathbf{x}\right)} \quad (11)$$

ile Epanechnikov çekirdek yoğunluk kestirimcisi ise

$$K(\mathbf{x}) = \left(\frac{3}{4}\right)^d \prod_{j=1}^d (1 - \mathbf{x}_j^2) \times [u(\mathbf{x}_j + 1) - u(\mathbf{x}_j - 1)], \quad (12)$$

$$\mathbf{x}_j = [x_1, x_2, \dots, x_d]^T$$

ile verilir. Burada $u(\cdot)$ birim basamak fonksiyonunu göstermektedir. Kestirilen $f_x(\mathbf{x}, \mathbf{H})$, denklem 1'de yerine konularak entropi kestirilecektir. d -boyuttan oluşan çekirdek yoğunluk fonksiyonları kullanıldığında genişlik matrisinin \mathbf{H} ve çekirdek yoğunluk fonksiyonun tipinin seçilmesi kestirimci performansını belirler. Çekirdek yoğunluk kestirimcisi gözlemlenen veri düşük boyutlu ve yeterli sayıda gözlem noktası varsa doğru sonuç verir. Ancak boyut sayısı arttıkça çekirdek yoğunluk kestirimcisi veriye uygun bir çekirdek yoğunluk fonksiyonuna yakınsamayabilir.

3.3 k-en yakın komşuluklu entropi kestirimcisi

Denklem 1'de tanımlanan $\mathbf{X} \in \mathbb{R}^d$ rasgele değişkenin entropisini kestirmek için olasılık yoğunluk fonksiyonundan $f_x(\mathbf{x})$ olan süreçten sonlu sayıda N adet örnek aldığımızı düşünelim. Bu örneklerden $\mathbf{x}_i \in \mathbb{R}^d, i = 1, 2, \dots, N$ entropi kestirmek istenirse,

$$\hat{h}(\mathbf{X}) = -E[\log f_x(\mathbf{x}_i)]$$

$$\hat{h}(\mathbf{X}) = -\frac{1}{N} \sum_{i=1}^N \log f_x(\mathbf{x}_i) \quad (13)$$

yazabiliriz [13]. Burada $f_x(\mathbf{x}_i)$ bilinmeyen olasılık yoğunluk fonksiyonudur. Kozachenko ve Leonenko'ya göre, denklem 13'de entropiyi elde etmek için önce $f_x(\mathbf{x}_i)$ kestirilmelidir [13]. Bu entropi kestirimcisinde amaç, her bir \mathbf{x}_i noktası için k -en yakın komşularının olasılıklarını yazarak $f_x(\mathbf{x}_i)$ olasılık yoğunluk fonksiyonunu elde etmektir. \mathbf{x}_i noktasının ε -küresi içerisinde olma olasılığı,

$$f_k(\epsilon)d\epsilon = k \binom{N-1}{k} \frac{d F_i(\epsilon)}{d \epsilon} d \epsilon (F_i(\epsilon))^{k-1} \times (1 - F_i(\epsilon))^{N-k-1} \quad (14)$$

ile verilir. Burada, $f_k(\epsilon)$ x_i noktasının k. en yakın komşu noktasına olan uzaklığının ϵ 'a göre olan olasılık yoğunluğudur. $F_i(\epsilon)$ ϵ -küresinin olasılık kütle fonksiyonudur. Olasılık kütle fonksiyonunun logaritmasının ortalama değerinin

$$E[\log F_i(\epsilon)] = \psi(k) - \psi(N) \quad (15)$$

olduğu gösterilebilir. Burada, $\psi(\cdot)$ digamma fonksiyonudur. Bununla birlikte olasılık kütle yoğunluğu

$$F_i(\epsilon) = \int_{B(\epsilon, x_i)} f_x(x) dx \quad (16)$$

ifadesiyle yazılır. Eğer $f_x(x_i)$ dağılımı yerel olarak düzgün ise, olasılık kütle yoğunluğu için,

$$F_i(\epsilon) \approx v f_x(x_i) . \quad (17)$$

yazılabilir. Burada $v = c_d \epsilon_i^d$ d-boyutlu ϵ çaplı kürenin hacmidir. c_d metriğe bağlı sayısal çarpanı temsil eder. Olasılık yoğunluk fonksiyonu $f_x(x_i) = F_i(\epsilon) / c_d \epsilon_i^d$ şeklinde düzenlenerek yazılabilir. Olasılık yoğunluk fonksiyonunun logaritması alınıp ortalama değeri hesaplanırsa

$$E[\log f_x(x_i)] = \psi(k) - \psi(N) - E[\log c_d] - d E[\log \epsilon_i] \quad (18)$$

elde edilir. Entropi kestirimi denklem 18 kullanılarak

$$\hat{h}(X) = -E[\log f_x(x_i)] = -\psi(k) + \psi(N) + \log c_d + \frac{d}{N} \sum_{i=1}^N \log \epsilon_i \quad (19)$$

şeklinde yazılabilir. Burada, ϵ_i x_i noktası ile k. en yakın komşusunun arasındaki uzaklıktır. Benzer bir kurgu ile, bileşik entropi aşağıdaki denklemle ifade edilir:

$$\hat{h}(X_1, X_2) = -\psi(k) + \psi(N) + \dots + \log c_{d_{x_1}} c_{d_{x_2}} + \frac{d_{x_1} + d_{x_2}}{N} \sum_{i=1}^N \log \epsilon_i. \quad (20)$$

Ortak bilgi ölçüsü ise benzer bir yaklaşımla Kraskov vd. tarafından kestirilmiştir [8]:

$$\hat{I}(X_1, X_2) = \psi(k) + \psi(N) + \dots - \frac{1}{N} \sum_{i=1}^N [\psi(\eta_{x_1}(i)) + \psi(\eta_{x_2}(i))]. \quad (21)$$

$\eta_{x_1}(i)$ ve $\eta_{x_2}(i)$ sırasıyla ϵ_{x_1} ve ϵ_{x_2} çaplı küreler içerisindeki noktaların sayısıdır. Ayrıca, aktarım entropisi k-en yakın komşuluklu entropi kestirimcisi yardımıyla ifade edilmiştir [14]:

$$\widehat{T}E_{x_1 \rightarrow x_2} = \psi(k) + \frac{1}{N} \sum_{i=1}^N [\psi(\eta_{x_{2,n}}(i) + 1) \dots - \psi(\eta_{(x_{1,n}, x_{2,n})}(i) + 1) \dots - \psi(\eta_{(x_{2,n+1}, x_{2,n})}(i) + 1)] \quad (22)$$

Burada, $\eta_{x_{2,n}}(i)$, $\eta_{(x_{1,n}, x_{2,n})}(i)$ ve $\eta_{(x_{2,n+1}, x_{2,n})}(i)$ sırasıyla $\epsilon_{x_{2,n}}$, $\epsilon_{(x_{1,n}, x_{2,n})}$ ve $\epsilon_{(x_{2,n+1}, x_{2,n})}$ çaplı küreler içindeki nokta sayısını ifade eder. Dağılımdaki gözlem sayısı N ve komşuluk sayısı k bu kestirimci tipi kullanıldığında belirlenmesi gereken parametrelerdir.

k-en yakın komşuluklu entropi kestirimcisinde entropi, ortak bilgi ölçüsü ve aktarım entropisi gözlem noktaları doğrudan kullanılarak kestirilir. Eğer veri yerel olarak düzgün dağılımlı ise ölçüler yüksek doğrulukla kestirilecektir. Bu kestirimcinin avantajı, veri yüksek boyutlu olsa bile entropi ve benzeri ölçülerin çok hızlı bir şekilde hesaplanabilmesidir. Fakat, kaotik sistemlerin garip çekerleri gibi verinin yerel olarak düzgün dağılmadığı ve uzayda yönlülük gösterdiği durumlarda küçük veri kümeleri için bu tip kestirimcinin yüksek doğrulukla sonuçlar verilmesi beklenmemelidir.

3.4 kpN entropi kestirimcisi

k-en yakın entropi kestirimcisinde noktaların seçilen bir veri noktası civarında düzgün dağıldığı varsayılarak entropi kestirilmiştir. Lombardi ve Pant'in önerdiği kpN entropi kestirimcisinde ise hem çekirdek yoğunluk hem

de k-en yakın komşuluklu entropi kestirimcisi yardımıyla entropi kestirilir [9]. Her bir nokta için olasılık kütle fonksiyonu yerel Gauss çekirdek yoğunluk fonksiyonları kullanılarak hesaplanır. Kestirimci her x_i noktasının en yakın p_k noktası için en iyi yerel çekirdek fonksiyonunu hesaplar. Olasılık yoğunluk fonksiyonları k-en yakın komşuluklu entropi kestirimcisinde olduğu gibi, x_i noktası ve noktanın k. en yakın komşusu arasındaki uzaklığa bağlı olarak hesaplanır. $B(\epsilon, x_i)$ küresi içerisindeki olasılık yoğunluk fonksiyonun Gauss dağılımlı olduğu kabul edilmiştir. Bilinmeyen olasılık yoğunluk fonksiyonu

$$f_x(\mathbf{x}) \approx \frac{1}{(2\pi)^d |\Gamma|^d} e^{(-0.5(\mathbf{x}-\boldsymbol{\mu})^T \Gamma^{-1} (\mathbf{x}-\boldsymbol{\mu}))} \quad (23)$$

ile ifade edilir. Burada, $\boldsymbol{\mu}$ ortalama vektörünü ve Γ kovaryans matrisidir. $B(\epsilon, x_i)$ küresi içerisindeki olasılık yoğunluk fonksiyonu yerel Gauss fonksiyonları cinsinden yazılabilir:

$$f_x(\mathbf{x}) \approx f_x(x_i) \frac{g(\mathbf{x})}{g(x_i)} \quad (24)$$

$$g(\mathbf{x}) = e^{-0.5(\mathbf{x}-\boldsymbol{\mu})^T \Gamma^{-1} (\mathbf{x}-\boldsymbol{\mu})},$$

$g(x_i)$ denklemindeki Γ , x_i noktasının p_k komşuluğunun kovaryans matrisidir. Olasılık kütle yoğunluğu ise

$$F_i = f_x(x_i) \frac{G_i}{g(x_i)}, \quad (25)$$

$$G_i = \int_{B(\epsilon, x_i)} g(\mathbf{x}) d\mathbf{x}$$

ifadesine eşit olur. Olasılık yoğunluk fonksiyonu çekilip $\log f_x(x_i)$ 'nin beklenen değeri hesaplanırsa

$$\hat{h}(\mathbf{X}) = -E[\log f_x(x_i)] = \psi(k) + \dots \quad (26)$$

$$\dots - \psi(N) - \frac{1}{N} \sum_{i=1}^N \log g(x_i) + \frac{1}{N} \sum_{i=1}^N \log G_i$$

kestirimcisi elde edilir. Benzer bir yaklaşımla, dağılımdaki gözlem sayısı N , komşuluk sayısı k ve çekirdek yoğunluk kestirimcisinde kullanılan nokta sayısı p_k gibi parametrelerinin seçimi bu kestirimci tipi için incelenecektir. Her bir noktanın olasılık kütle fonksiyonunun p_k

komşuluğunu yerel Gauss çekirdek yoğunluk fonksiyonları ile kestiren kpN entropi kestirimcisi, yerel olarak düzgün dağılımlı olmayan veriler için entropiyi daha doğru kestirmektedir. Üstelik verinin olasılık dağılımında kuyrukları var ise, bu kestirimci Gauss tipi çekirdek yoğunluk fonksiyonu da kullandığından verinin kuyruklarını daha iyi temsil ederek entropiyi daha doğru kestirir.

4. Benzetim ve Sonuçlar

Bu kısımda entropi, ortak bilgi ölçüsü ve aktarım entropisi kestirimlerine ilişkin benzetimler Intel® Xeon® E5630 işlemcili 16 çekirdekli 48 GB hafızalı bir iş istasyonunda paralel benzetimlerle ve Python programlama dili kullanılarak yapılmıştır.

Birbirini Gauss gürültü ile süren doğrusal sistemlerin bilgi ölçülerini kestirmenin yanı sıra, doğrusal olmayan ve beyaz gürültü artışlı gürültü terimine sahip rasgele bir sistemin entropi ölçüsü de kestirilmiştir. Gauss dağılımlar için iyi sonuç vermesi beklenen çekirdek yoğunluk kestirimcisi ve kpN entropi kestirimcisinin doğrusal olmayan bir sistemdeki entropi kestirimlerinin başarımının nasıl olacağı tartışılacaktır. Ayrıca doğrusal olmayan bir sistemde yerel olarak düzgün dağılımlı olasılık yoğunluk fonksiyonları için iyi çalışan k-en komşuluklu entropi kestirimcisinin başarımı da incelenmek istenmiştir. İki durumlu sistemin entropisi kestirimi için en iyi parametreler elde edilecektir. Ayrıca kestirimcilerin başarımını kıyaslayabilmek için, aktarım entropisi kestiriminin sentetik olmayan bir veri üstünde incelenen kestirimcileri kullanarak istatistiksel önem analizi sunulmuştur.

4.1 Birinci dereceden vektörel ardışık bağlantı modelleri

Gauss beyaz gürültü artırımı ile sürülmüş d adet doğrusal ardışık bağlantı modelleri ile birbirine bağlanan sistemler tanımlansın:

$$\mathbf{X}_{n+1} = \mathbf{A} \mathbf{X}_n + \boldsymbol{\eta}_n. \quad (27)$$

Burada, $\mathbf{X}_n = [X_{1,n}, X_{2,n}, \dots, X_{d,n}]^T \in \mathbb{R}^d$ n anındaki durum vektörünü, $\mathbf{A} \in \mathbb{R}^{d \times d}$ gerçel elemanlardan oluşan kararlı bağlantı matrisini ve $\boldsymbol{\eta}_n = [\eta_{1,n}, \eta_{2,n}, \dots, \eta_{d,n}]^T$ Gauss gürültü vektörünü temsil eder. Shannon bilgi kuramına göre ortalaması $\boldsymbol{\mu}$ vektörü ve kovaryans matrisi Γ olan çok değişkenli Gauss dağılımının entropisi

$$h(\mathbf{X}) = 0.5 \log((2\pi e)^n \det[\mathbf{\Gamma}]) \text{ nats} \quad (28)$$

cinsinden ifade edilir [2]. $\det[\mathbf{\Gamma}]$, $\mathbf{\Gamma}$ matrisinin determinantıdır. Denklem 2 kullanılarak ortak bilgi ölçüsü ise

$$\begin{aligned} I(\mathbf{X}_1, \mathbf{X}_2) &= h(\mathbf{X}_{1,n}) + h(\mathbf{X}_{2,n}) + \dots \\ &\dots - h(\mathbf{X}_{1,n}, \mathbf{X}_{2,n}) \quad (29) \\ &= 0.5 \log \frac{\det[\mathbf{\Gamma}(\mathbf{X}_{1,n})] \det[\mathbf{\Gamma}(\mathbf{X}_{2,n})]}{\det[\mathbf{\Gamma}(\mathbf{X}_{1,n}, \mathbf{X}_{2,n})]} \end{aligned}$$

olur. Birbirine bağlanmış Gauss süreçlerinin aktarım entropisi ise,

$$\begin{aligned} TE_{X_1 \rightarrow X_2} &= h(\mathbf{X}_{2,n+1}, \mathbf{X}_{2,n}^{(\ell)}) - h(\mathbf{X}_{2,n}^{(\ell)}) + \dots \\ &\dots + h(\mathbf{X}_{2,n}^{(\ell)}, \mathbf{X}_{1,n}^{(k)}) - h(\mathbf{X}_{2,n+1}, \mathbf{X}_{2,n}^{(\ell)}, \mathbf{X}_{1,n}^{(k)}) \\ &= 0.5 \log \frac{\det[\mathbf{\Gamma}(\mathbf{X}_{2,n+1}, \mathbf{X}_{2,n}^{(\ell)})] \det[\mathbf{\Gamma}(\mathbf{X}_{2,n}^{(\ell)}, \mathbf{X}_{1,n}^{(k)})]}{\det[\mathbf{\Gamma}(\mathbf{X}_{2,n+1}, \mathbf{X}_{2,n}^{(\ell)}, \mathbf{X}_{1,n}^{(k)})] \det[\mathbf{\Gamma}(\mathbf{X}_{2,n}^{(\ell)})]} \quad (30) \end{aligned}$$

ifade edilir [15]. Denklem 29'da ve 30'da kovaryans matrisinin elemanları hesaplanarak sistemler arasındaki ortak bilgi ölçüsü ve aktarım entropisinin teorik değeri elde edilir.

Denklem 27 kullanılarak oluşturulan 3 adet sistemi ele alalım ve \mathbf{A} matrisi, $\mathbf{A} = [0,5 \ 0 \ 0; 0,5 \ 0,6 \ 0; 0,2 \ 0,7 \ 0,3]$ seçilsin. Denklem 27'da tanımlanan Gauss gürültü vektörlerinin ortalaması $\boldsymbol{\mu} = [0, 0, 0]^T$ ve kovaryans matrisi $\mathbf{\Gamma} = \text{diag}\{1, 1, 1\}$ alınmıştır. Bu sistem için teorik olarak denklem 29 kullanılarak $I(\mathbf{X}_1, \mathbf{X}_2) = 0.025 \text{ nats}$, $I(\mathbf{X}_1, \mathbf{X}_3) = 0.011 \text{ nats}$, $I(\mathbf{X}_2, \mathbf{X}_3) = 0.188 \text{ nats}$ 'tır. Denklem 29'da $k = \ell = 1$ seçilirse, $TE_{X_1 \rightarrow X_2} = 0.092 \text{ nats}$, $TE_{X_1 \rightarrow X_3} = 0.038 \text{ nats}$, $TE_{X_2 \rightarrow X_3} = 0.282 \text{ nats}$, $TE_{X_2 \rightarrow X_1} = 0 \text{ nats}$ hesaplanmıştır. Histogram, çekirdek yoğunluk ve k-en yakın komşuluklu entropi kestirimci tiplerinde $N = 256$ 'dan $N = 2048$ 'ye kadar farklı gözlem sayısı için kestirilmiştir. kpN entropi kestirimcisi, gözlem sayısı N , k-komşuluk sayısı ve p_k komşuluk sayısı gibi 3 farklı parametreye sahiptir. Gözlem sayısı arttıkça bu kestirimcinin performansının arttığı bilinmektedir [9]. Bu yüzden gözlem sayısı sabit tutularak ve yeterince büyük bir gözlem sayısı için ($N = 8192$), kestirimler hesaplanacaktır. Bilgi ölçüleri kpN entropi kestirimcisinin farklı k ve p_k komşuluk değerleri için kestirilmiştir.

Parametreler için kestirimler 1000 kez tekrarlanmıştır.

Bilgi ölçülerinin teorik değerleri bilindiğinden, kestirim hatasının ortalaması ve varyansı 0.05 ve altında olduğu seviyeler kabul edilebilir varsayılmıştır. Herhangi bir kestirimci tipi kullanılarak elde edilen bilgi ölçülerinin kestirim hataları kabul edilebilir seviyenin altında ise kestirimcinin başarımı yüksektir.

İlk olarak ortak bilgi ölçüsü ve aktarım ölçüsü değişen aralık çapları ε ve N için histogram yöntemi ile denklem 8 kullanılarak kestirilmiştir. Aralıklar arasında örtüşme yoktur ve $2^{-3} < \varepsilon < 2^1$ 'dir. Şekil 1 ve 2'de ortak bilgi ölçüsünün NKY ve NSH değerleri gösterilmiştir. Nokta sayısı ve aralık çapı arttıkça sistemler arasındaki ortak bilgi ölçüsünün NKY değerlerinin azaldığı gözlenmiştir.

Ayrıca Şekil 2 incelendiğinde ortak bilgi ölçüsünün NSH değerlerinin de aralık çapı ve nokta sayısı arttıkça azaldığı görülmektedir. Şekil 1'de gösterilmeyen $\varepsilon \gg 2$ çap değerleri için NKY değerleri hesaplanmıştır. Histogram kestirimcisi ile hesaplanan bilgi ölçülerinin aralık çapının sürekli olarak artması NKY'nın sürekli azalmasına neden olmaz. Aralık çapı, gözlemlenen olası tüm durumların aralığından daha büyük ise, histogram kestirimcisi her bir boyutta tek bir aralığa sahip olur. Bu durumda ancak veri uzayda düzgün dağılımlı ise, histogram kestirimcisi başarılı olabilir. Aksi takdirde, Şekil 1'de gösterilmeyen $\varepsilon \gg 2$ çap değerleri için, histogram kestirimcisi ile kestirilen ortak bilgi ölçüsünün NKY değerlerinin arttığı gözlenmiştir. Histogram kestirimcisi kullanılarak kestirilen ortak bilgi ölçüsünde $N = 2^{11}$ ve $\varepsilon = 1.68$ olduğunda minimum NKY ve NSH değerleri elde edilir.

Şekil 3 ve 4 incelendiğinde, benzer bir sonuçla, çap ve nokta sayısı arttıkça sistemler arasındaki aktarım entropisinin NKY değerleri azalır. NKY değerleri tüm N değerleri ve $\varepsilon > 1$ olduğunda kabul edilebilir olduğu görülmektedir. Şekil 4'te aktarım entropisinin NSH değerlerine bakıldığında $N > 2^{10}$ ve $\varepsilon > 1$ 'den kestirimlerin kabul edilebilir olduğu tespit edilmiştir.

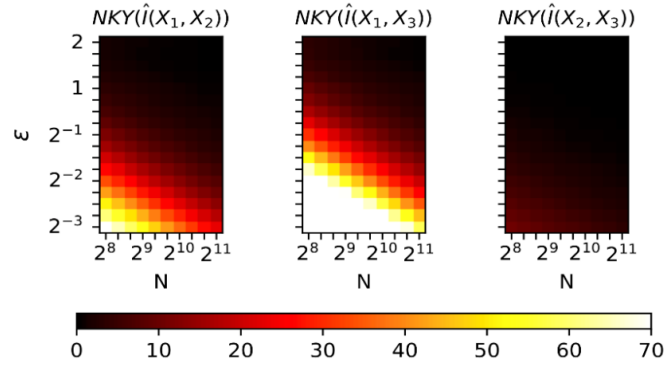
Denklem 27'deki sistemlerin aktarım entropisi histogram kestirimlerinin minimum NKY ve NSH değerleri $N = 2^{11}$ ve $\varepsilon = 2$ olduğu Şekil 3'te ve 4'te açıkça görülmektedir. Birbiri ile doğrudan bağlı olmayan \mathbf{X}_1 ve \mathbf{X}_3 sistemlerinin ortak bilgi ölçüsü ve aktarım entropisi ölçülerinin yanlılık

ve standart hata değerleri diğer sistemler arasındaki ortak bilgi ölçüsü ve aktarım entropisi ölçüsüne göre daha yüksek değer aldığı gözlenmiştir.

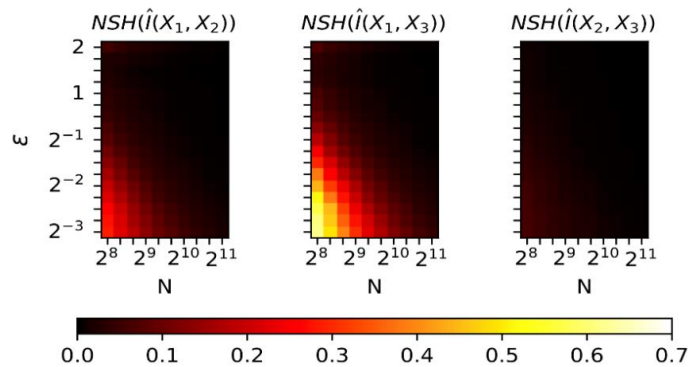
Ortak bilgi ölçüsü çekirdek yoğunluk kestirimcisi ile kestirilirken denklem 11 ve 12’de tanımlanan Gauss ve Epanechnikov çekirdek kullanılmıştır. Genişlik matrisinin \mathbf{H} diyagonal elemanların uzunluğu $0.1 \leq |h| \leq 1$ ve $|h| = |h_{11}| = |h_{22}|$ olacak şekilde seçilmiştir. Ortak bilgi ölçüsünde 2 boyutlu çekirdek yoğunluk fonksiyonu ve aktarım entropisinde 3 boyutlu çekirdek yoğunluk fonksiyonu kullanılmıştır. Şekil 5, 6, 7 ve 8’de ortak bilgi ölçüsünün kestirimlerinin NKY ve NSH değerleri ve Şekil 9, 10, 11 ve 12’de aktarım entropisinin kestirimlerinin NKY ve NSH değerleri sunulmuştur.

Şekil 5 incelendiğinde birbirine bağlanmış Gauss süreçlerinin Gauss çekirdek yoğunluk fonksiyonu ile kestirildiğinde ortak bilgi

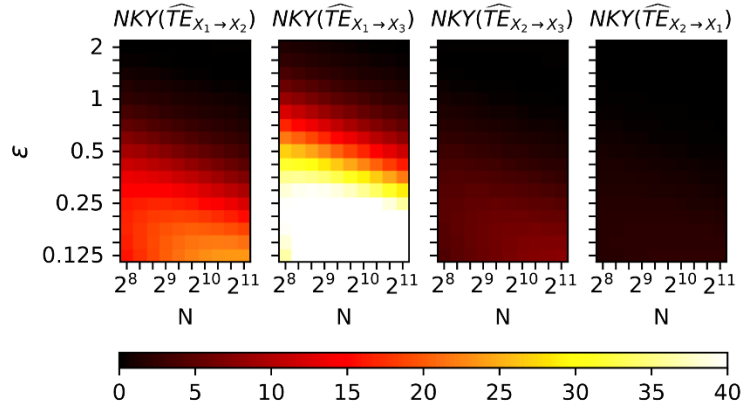
ölçüsünün NKY değerleri N' den bağımsız olarak ve $|h| > 0.3$ için kabul edilebilir seviyededir. Şekil 6’ya bakıldığında ortak bilgi ölçüsünün NSH değerleri kayda değer ölçüde düşüktür. Beklenildiği üzere Gauss gürültü ile sürülmüş sistem çıkışlarına ilişkin Gauss çekirdek yoğunluk fonksiyonu ile kestirilirse başarımları yüksektir. Epanechnikov yoğunluk kestirimcisi seçildiğinde N' den bağımsız olarak ve $|h| > 0.6$ ise ortak bilgi ölçüsünün kestirimlerinin NKY değerleri kabul edilir aralıktadır (Şekil 7). NSH değerleri sabit bant genişliğinde örnek sayısı arttıkça fazla değişmez. Fakat bant genişliğinin artması ortak bilgi ölçüsü kestirimlerinin NSH değerini azaltır (Şekil 8). N' den bağımsız olarak ve $|h| = 0.9$ seçilmesi daha uygundur. Şekil 6 ve 8 karşılaştırıldığında NSH değerlerinin Gauss çekirdek kestirimcisi ile elde edilen NSH değerlerine göre daha yüksek olduğu gözlenmiştir.



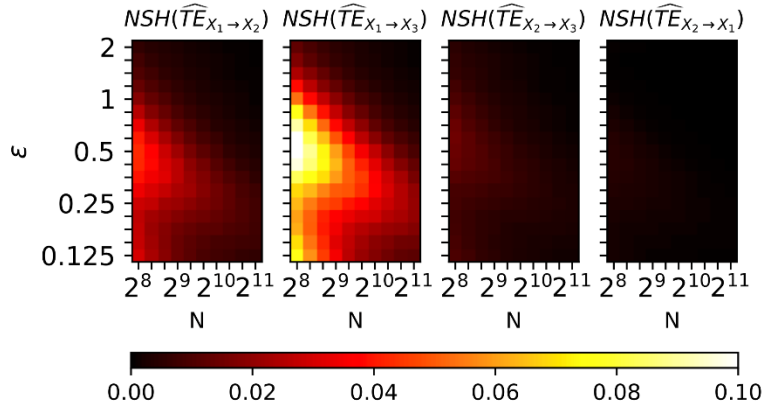
Şekil 1. Sistemler arasındaki ortak bilgi ölçüsünün histogram kestirimcisi kullanılarak farklı (N, ϵ) parametreleri için NKY değerleri



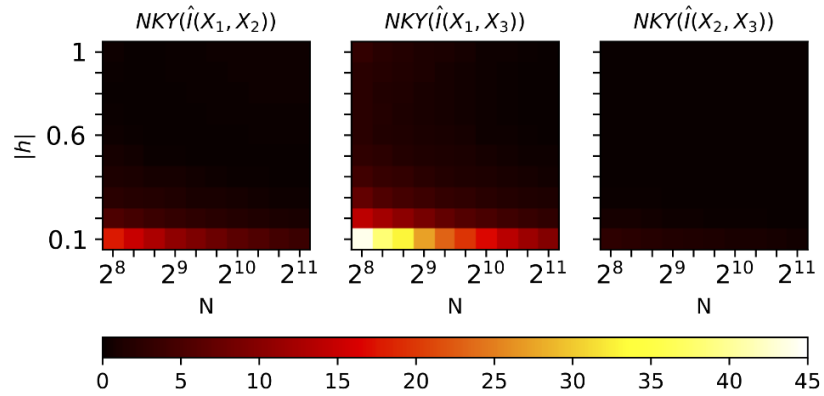
Şekil 2. Sistemler arasındaki ortak bilgi ölçüsünün histogram kestirimcisi kullanılarak farklı (N, ϵ) parametreleri için NSH değerleri



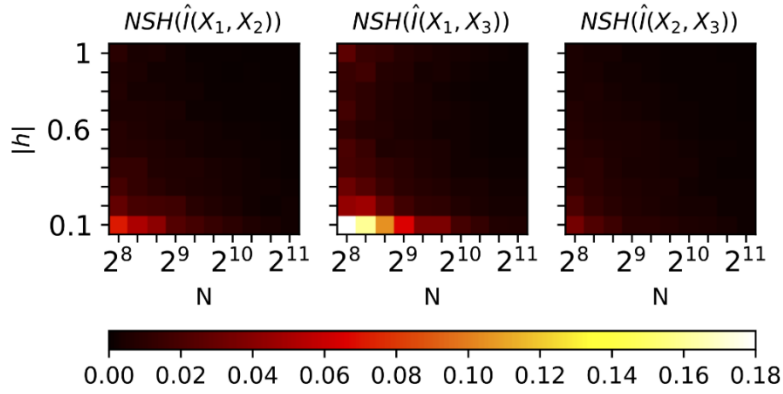
Şekil 3. Sistemler arasındaki aktarım entropisinin histogram kestirimcisi kullanılarak farklı (N, ϵ) parametreleri için NKY değerleri



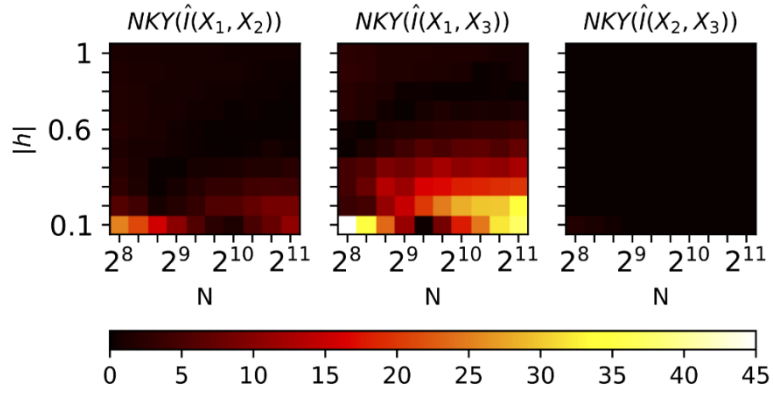
Şekil 4. Sistemler arasındaki aktarım entropisinin histogram kestirimcisi kullanılarak farklı (N, ϵ) parametreleri için NSH değerleri



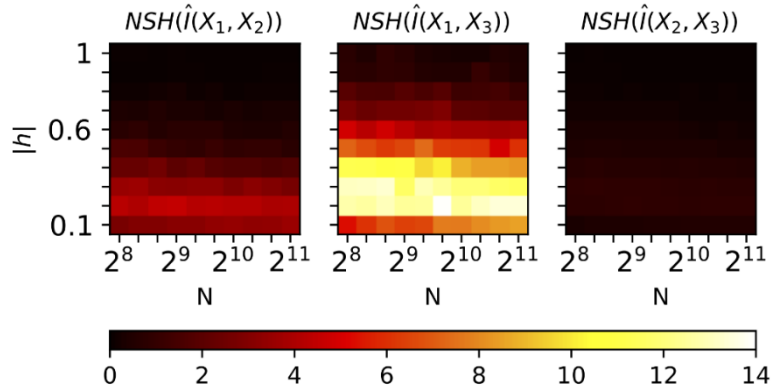
Şekil 5. Sistemler arasındaki ortak bilgi ölçüsünün Gauss çekirdek yoğunluk kestirimcisi kullanılarak farklı $(N, |h|)$ parametreleri için NKY değerleri



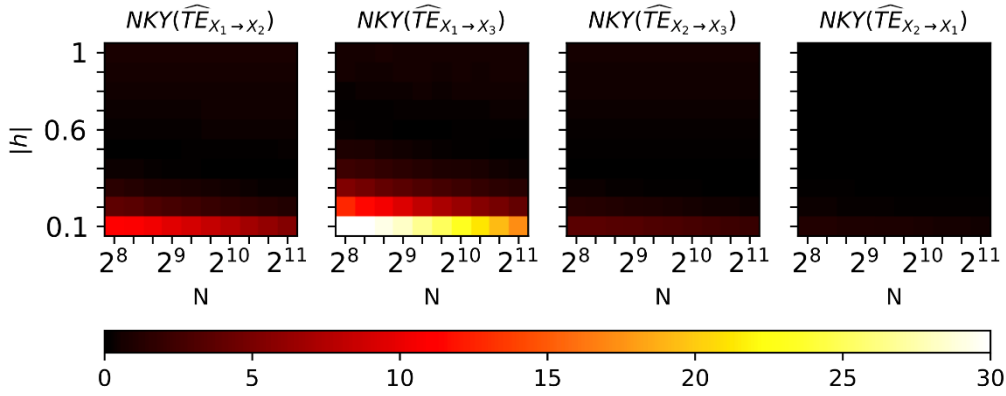
Şekil 6. Sistemler arasındaki ortak bilgi ölçüsünün Gauss çekirdek yoğunluk kestirimcisi kullanılarak farklı $(N, |h|)$ parametreleri için NSH değerleri



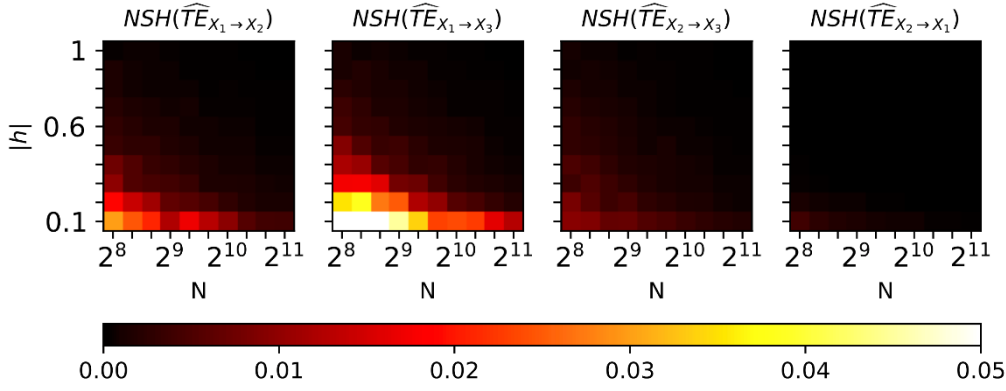
Şekil 7. Sistemler arasındaki ortak bilgi ölçüsünün Epanechnikov çekirdek yoğunluk kestirimcisi kullanılarak farklı $(N, |h|)$ parametreleri için NKY değerleri



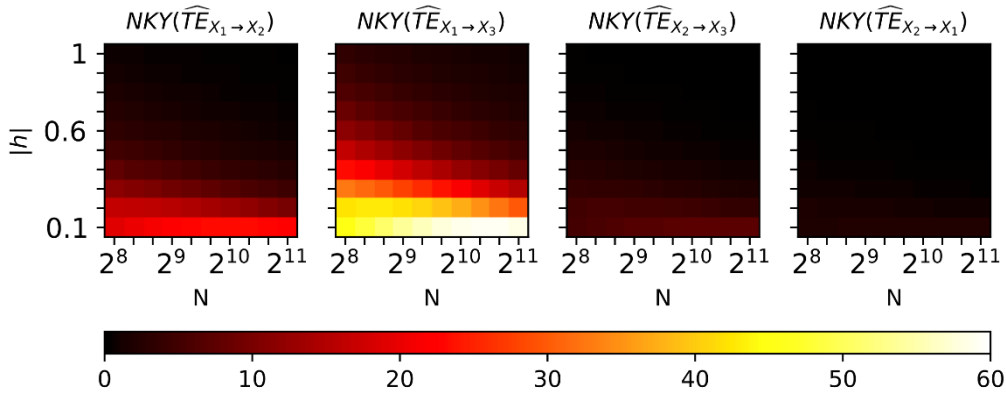
Şekil 8. Sistemler arasındaki ortak bilgi ölçüsünün Epanechnikov çekirdek yoğunluk kestirimcisi kullanılarak farklı $(N, |h|)$ parametreleri için NSH değerleri



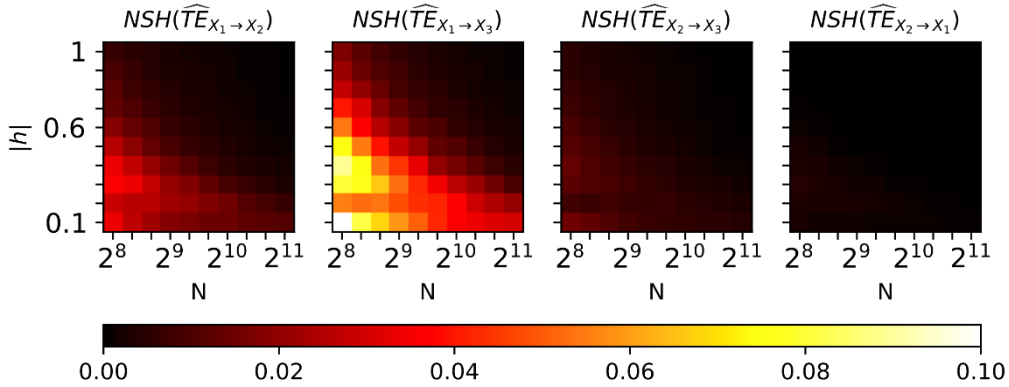
Şekil 9. Sistemler arasındaki aktarım entropisinin Gauss çekirdek yoğunluk kestirimcisi kullanılarak farklı $(N, |h|)$ parametreleri için NKY değerleri



Şekil 10. Sistemler arasındaki aktarım entropisinin Gauss çekirdek yoğunluk kestirimcisi kullanılarak farklı $(N, |h|)$ parametreleri için NSH değerleri



Şekil 11. Sistemler arasındaki aktarım entropisinin Epanechnikov çekirdek yoğunluk kestirimcisi kullanılarak farklı $(N, |h|)$ parametreleri için NKY değerleri



Şekil 12. Sistemler arasındaki aktarım entropisinin Epanechnikov çekirdek yoğunluk kestirimcisi kullanılarak farklı $(N, |h|)$ parametreleri için NSH değerleri

Benzer bir şekilde, aktarım entropisi Gauss çekirdek yoğunluk fonksiyonu ile kestirilirse NKY değerleri N 'den bağımsız olarak ve $|h| > 0.4$ için kabul edilebilirdir (Şekil 9). Şekil 10'da NSH değerleri 0.05'den küçüktür. Epanechnikov çekirdek kestirimcisi kullanıldığında $N > 2^{10}$ $|h| > 0.9$ ise NKY ve NSH değerleri kabul edilebilir seviyededir (Şekil 11 ve 12).

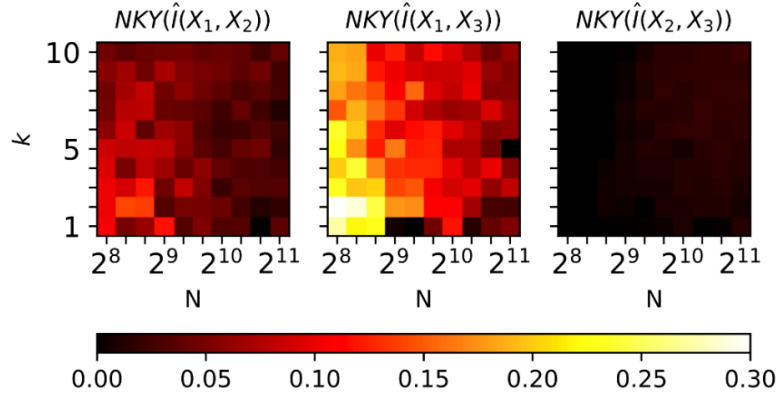
Gauss gürültü ile sürülmüş sistem çıkışlarının entropi ölçülerinin sonsuz kuyruklu Gauss tipi çekirdek tipi kestirilmesi NKY ve NSH değerlerini düşürür. Şekil 7 ve 11 karşılaştırıldığında ortak bilgi ölçüsü ve aktarım entropisi kestirilirken sırasıyla 2-boyutlu ve 3-boyutlu bir Epanechnikov çekirdek yoğunluk fonksiyonu kullanıldığında boyut sayısı arttıkça NSH değerinin arttığı gözlenmiştir.

k -en yakın komşuluklu entropi kestirimcisi için değişen örnek ve komşuluk sayısına göre kestirimler yapılmıştır. Ayrıca, bu kestirimci için komşuluk sayısı $1 \leq k \leq 10$ seçilmiştir. Şekil 13 ve 14'te ortak bilgi ölçülerinin NKY ve NSH değerleri, Şekil 15 ve 16'da aktarım entropilerinin NKY ve NSH değerleri gösterilmiştir. Şekil 13'te ortak bilgi ölçüsünün NKY değerleri incelendiğinde değerleri $N > 2^{10}$ ve $k > 7$ için yapılan kestirimler kabul edilebilirdir. $N > 2^{10}$ ve $k > 7$ seçilirse ortak bilgi ölçüsü kestirimlerinin NSH değerleri 0.02'in altında kalmaktadır (Şekil 14). Şekil 15 incelendiğinde denklem 22 kullanılarak yapılan kestirimlerde NKY değerleri $N = 2^{10}$ ve $k > 7$ için oldukça düşüktür. Şekil 16'da ortak bilgi ölçüsüne paralel bir şekilde, NSH değerleri

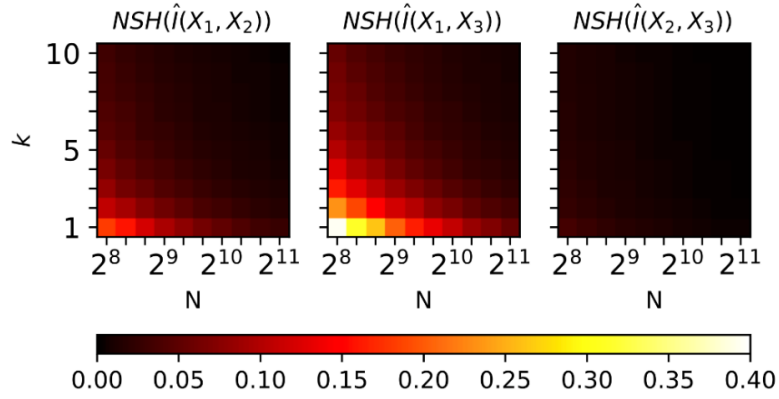
için 0.02'den küçüktür. Öte yandan genel olarak k -en yakın komşuluklu entropi kestirimcisinin NKY ve NSH değerleri, histogram ve çekirdek yoğunluk kestirimcisine kıyasla daha küçük kalmıştır. Verinin dağılımı yerel olarak düzgünse k -en yakın komşuluklu entropi kestirimcisi yeterince büyük N ve komşuluk sayısı k için gürbüzdür. (Kestirimci parametrelerinin değişiminden fazlaca etkilenmez.)

kpN entropi kestirimcisi için sistemler arası ortak bilgi ölçüsü ve aktarım entropisi kestirilirken $0.01 \leq p_k/N \leq 0.1$ ve $1 \leq k \leq 10$ olacak şekilde seçilmiştir. Kestirimlerin NKY ve NSH değerleri Şekil 17, 18, 19, 20'de gösterilmiştir. N arttığında, ortak bilgi ölçüsünün ve aktarım entropisinin kpN kestirimcisi için NKY değerleri azalmıştır. Başka bir deyişle kestirimcinin *tutarlı* olduğu görülmüştür. p_k/N oranı 0.05 ve tüm k değerleri için ortak bilgi ölçüsü kestirimlerinin NKY değerleri kabul edildir (Şekil 17). $\frac{p_k}{N} = 0.1$ ve $k = 10$ ise NSH değerleri minimum değer alır (Şekil 18). $\frac{p_k}{N}$ oranının değişmesi NSH değerlerini büyük ölçüde değiştirmezken, k değerleri arttıkça kestirimlerin NSH değerleri azalır. Şekil 19'da, $\frac{p_k}{N}$ oranı arttıkça ve tüm k için NKY değerleri kabul edilebilir seviyededir. Benzer bir şekilde, NSH değerleri tüm $(p_k/N, k)$ değerleri için düşük seviyededir (Şekil 20).

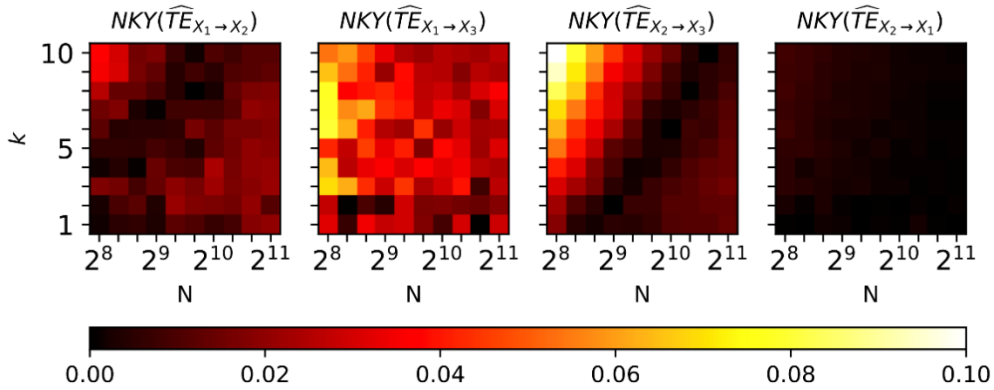
Denklem 27'in doğası gereği, yerel dağılımlar Gauss olduğundan, kpN entropi kestirimcisinin başarımı yüksektir.



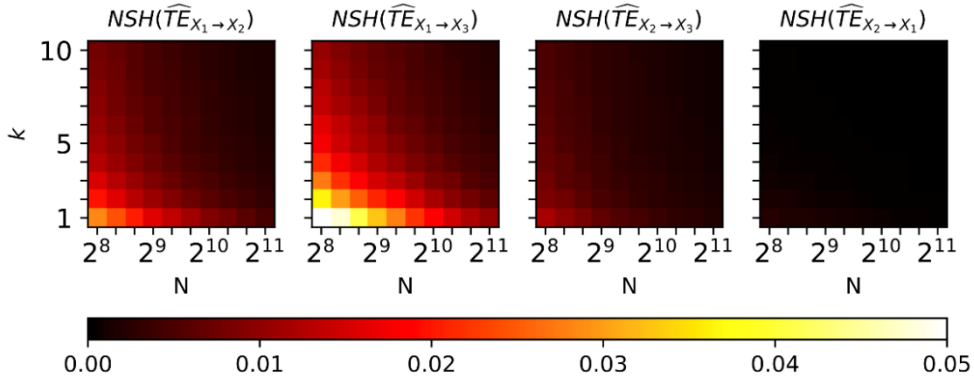
Şekil 13. Sistemler arasındaki ortak bilgi ölçüsünün k -en yakın komşuluklu entropi kestirimcisi kullanılarak farklı (N, k) parametreleri için NKY değerleri



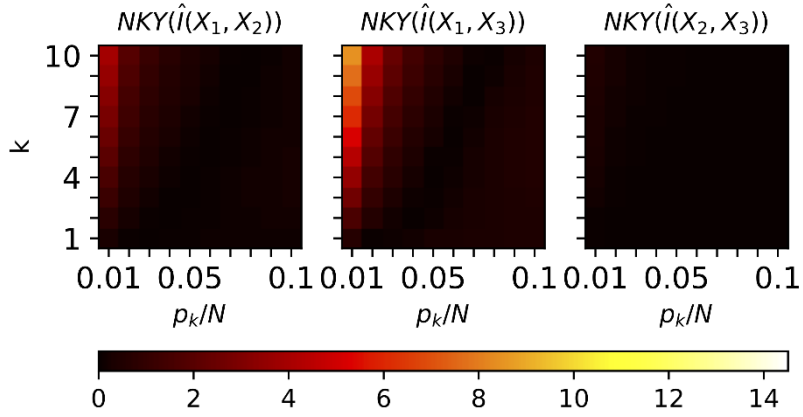
Şekil 14. Sistemler arasındaki ortak bilgi ölçüsünün k -en yakın komşuluklu entropi kestirimcisi kullanılarak farklı (N, k) parametreleri için NSH değerleri



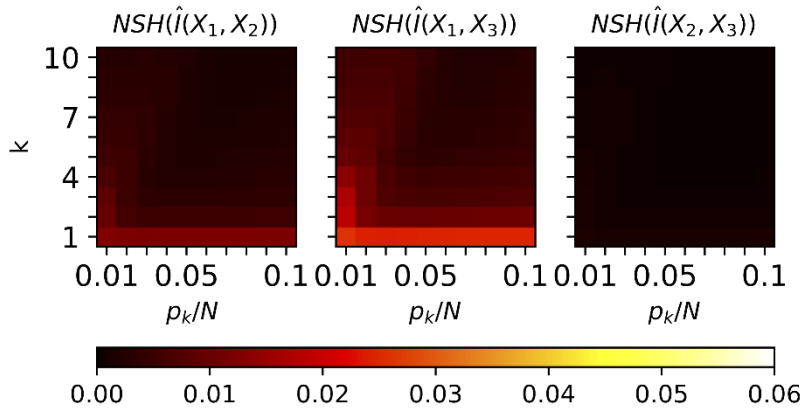
Şekil 15. Sistemler arasındaki aktarım entropisinin k -en yakın komşuluklu entropi kestirimcisi kullanılarak farklı (N, k) parametreleri için NKY değerleri



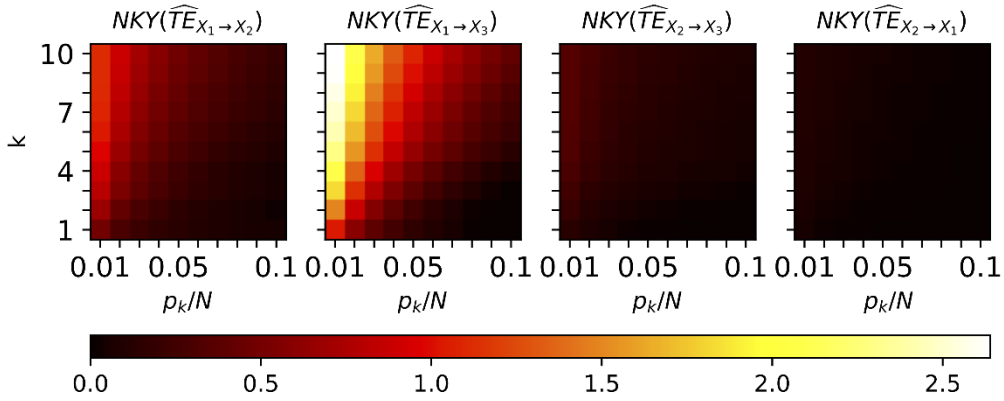
Şekil 16. Sistemler arasındaki aktarım entropisinin k-en yakın komşuluklu entropi kestirimcisi kullanılarak farklı (N, k) parametreleri için NSH değerleri



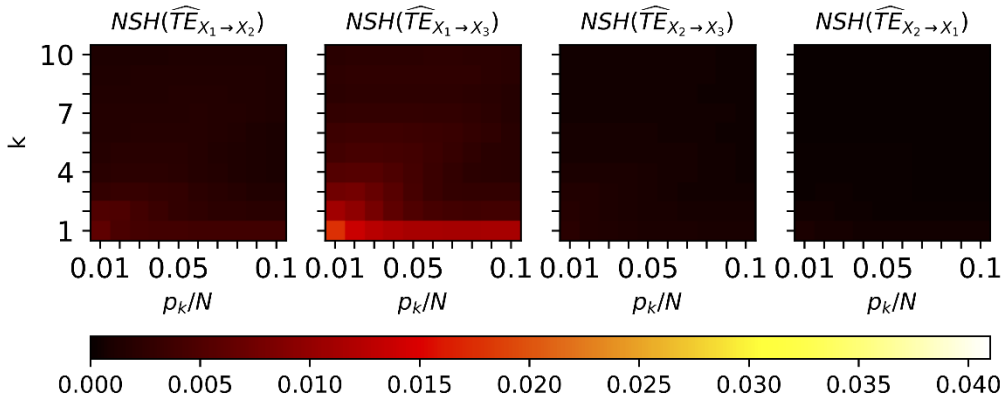
Şekil 17. Sistemler arasındaki ortak bilgi ölçüsünün $k p N$ entropi kestirimcisi kullanılarak örnek sayısı $N = 8192$ sabit tutularak farklı $(p_k/N, k)$ parametreleri için NKY değerleri



Şekil 18. Sistemler arasındaki ortak bilgi ölçüsünün $k p N$ entropi kestirimcisi kullanılarak örnek sayısı $N = 8192$ sabit tutularak farklı $(p_k/N, k)$ parametreleri için NSH değerleri



Şekil 19. Sistemler arasındaki aktarım entropisinin $k p_N$ entropi kestiricisi kullanılarak örnek sayısı $N = 8192$ sabit tutularak farklı $(p_k/N, k)$ parametreleri için NKY değerleri



Şekil 20. Sistemler arasındaki aktarım entropisinin $k p_N$ entropi kestiricisi kullanılarak örnek sayısı $N = 8192$ sabit tutularak farklı $(p_k/N, k)$ parametreleri için NSH değerleri

Genel olarak bakıldığında, tüm kestirimci tipleri için N arttıkça, kestirim performanslarının başarımlı beklenildiği gibi artmıştır. Histogram kestirimcisinde kestirimlerin daha az NKY ve NSH değerinde olması için ϵ 'nin doğru seçilmesi önemlidir. Gauss dağılımlı gürültüye sahip birbirine bağlı sistemlerin bilgi ölçüleri kestirilirken Gauss çekirdek yoğunluk fonksiyonunu kullanmak beklenildiği üzere avantajlıdır. Gauss dağılıma sahip veri ile farklı bir çekirdek yoğunluk fonksiyonu kullanıldığında NKY ve NSH değerleri Gauss çekirdek tipine göre daha yüksek gelecektir. Epanechnikov çekirdek fonksiyonunun boyutu arttıkça NKY değerlerinin arttığı görülmüştür. Yeterince örnek ile kestirildiğinde k -en yakın komşuluklu entropi kestirimcileri kullanılırsa en

az hataya sahip kestirimler elde edilir. $k p_N$ kestirimcisinde, k ve p_k komşuluklarının doğru seçimi kestirimlerin NKY değerini azaltırken, p_k komşuluk değerinin değişmesi NSH değerini değiştirmez. NSH komşuluk sayısının artması ile azalır. Hem çekirdek hem k -en yakın komşuluklu entropi kestirimcisi yöntemlerini kullanan $k p_N$ entropi kestirimcisi kullanılırsa Gauss dağılımlar için kabul edilir kestirimler elde ederiz.

Ayrıca denklem 27 ile ifade edilen sistemlerde birbirini doğrudan etkilemeyen sistemlerin $I(\mathbf{X}_1, \mathbf{X}_3)$ ve $TE_{X_1 \rightarrow X_3}$ kestirimlerinin NKY ve NSH değerleri birbirini doğrudan etkileyen sistemlere göre daha yüksek değerler almıştır. Dağılımın yerel olarak düzgün olduğu durumlarda, yeterli sayıda gözlem sayısı için, k

en yakın komşuluklu entropi kestirimcisi kullanılarak entropiden türetilen ölçülerin başarılı olarak kestirildiği gözlenmiştir. Bu sistemler için durumların dağılımları Gauss olduğundan Gauss çekirdek kestirimcisi kullanıldığında başarımın yüksek olacağı açıktır.

4.2 İki durumlu sistem

İki durumlu sistemin denklemi

$$dx = 0.5(x - x^3)dt + \sqrt{\alpha} dW(t). \quad (31)$$

ile verilsin. Burada, α sabit bir değişkeni, $W(t)$ Gauss artışı gürültü terimini ifade eder. Bu sistemin durumlarının durağan haldeki olasılık yoğunluk fonksiyonu $f(x)$ Fokker-Planck denklemi kullanılarak

$$f(x) = \frac{2}{\pi \left(I_{-\frac{1}{4}}\left(\frac{1}{8\alpha}\right) + I_{\frac{1}{4}}\left(\frac{1}{8\alpha}\right) \right)} e^{-\frac{\frac{1}{4}x^2 + x^4}{2\alpha}} \quad (32)$$

formunda elde edilir (Bakınız Ek) [16]. $I_n(x)$ değiştirilmiş birinci tür Bessel fonksiyonudur. İki durumlu sistemin entropisinin teorik değeri denklem 1 kullanılarak hesaplanabilir.

Doğrusal olmayan bir sisteme ilişkin kestirimci performansları iki durumlu sistemin çıkışları kullanılarak test edilmiştir. Denklem 31, farklı gözlem sayıları için $2256 < N < 4048$ $\Delta t = 0.1 sn$ alınarak Euler-Maruyama çözücüsü ile çözülmüştür [22]. İlk 2000 gözlem sayısı geçici hal cevabı sayılarak entropi kestiriminde dikkate alınmamıştır. $\alpha = 1$ ve farklı başlangıç koşulları için benzetimler 1000 kez tekrarlanmıştır. Denklem 1 ve 32 kullanılarak sayısal olarak hesaplanan entropi değeri $h(X) = 1.351 nats'$ 'dir.

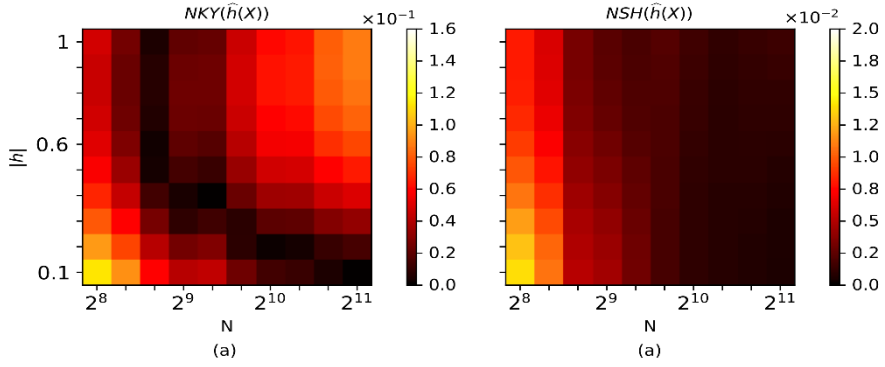
Histogram yöntemi ile kestirilen sonuçlarda $\varepsilon < 1$ için entropinin NKY değeri ε arttıkça azalır. N' 'den bağımsız olarak $\varepsilon = 1$ için entropi kestirimlerinin NKY değerleri düşük seviyededir ve tüm parametreler için NSH 0.01'den küçüktür. Şekil 21'de entropi Gauss çekirdek yoğunluk

fonsiyonu ile kestirilmiştir. NKY değerlerini minimize eden farklı $N, |h|$ parametreleri olduğu görülmektedir. NSH değerleri tüm kestirim parametreleri için kabul edilebilir seviyedir (Şekil 21).

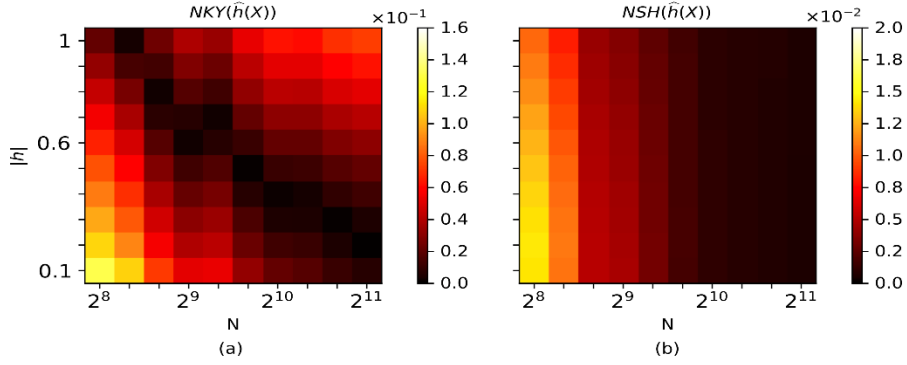
İki durumlu sistemin entropisinin Epanechnikov çekirdek fonksiyonu kullanılarak kestirimleri yapılmış, entropinin NKY ve NSH değerleri Şekil 22'de sunulmuştur. Epanechnikov yoğunluk kestirimcisinde geniş bir $(N, |h|)$ parametre aralığı için NKY değerleri 0.02'den düşüktür ve NSH değerleri 0.01'den küçüktür. Epanechnikov çekirdek fonksiyonu sınırlı bant genişliğine sahip olduğundan iki durumlu sistemin entropisini daha yüksek doğrulukla kestirmiştir. Çünkü sistemin durumları sınırlı bir aralıkta yoğunlaşmaktadır, olasılık yoğunluk fonksiyonunun etekleri Gauss'a göre daha hızlı azalır. Şekil 23'te benzetimlerde kullanılan (N, k) parametreleri için k-en yakın komşuluklu entropi kestirimcisinin NKY ve NSH değerleri sırasıyla sunulmuştur. Entropinin NKY değerinin yüksek (entropi kestiriminin yanlı) olduğu gözlenir ve entropi kestiriminin NSH değeri oldukça düşüktür (Şekil 23). Şekil 24'te kpN kestirimcisinin entropi kestirim NKY ve NSH değerleri sunulmuştur. kpN kestirimcisinde NKY ve NSH değerleri 0.02'den küçüktür. kpN entropi kestirimcisi iki durumlu sistemin entropisini yüksek başarımla kestirmiştir.

Şekil 25'te α' 'ın artan değerleri için histogram, Gauss çekirdek yoğunluk, Epanechnikov çekirdek ve kpN entropi yöntemleri ile kestirilen entropi kestirimlerinin ortalaması ve normalize olmuş ortalama karesel hatası gösterilmiştir.

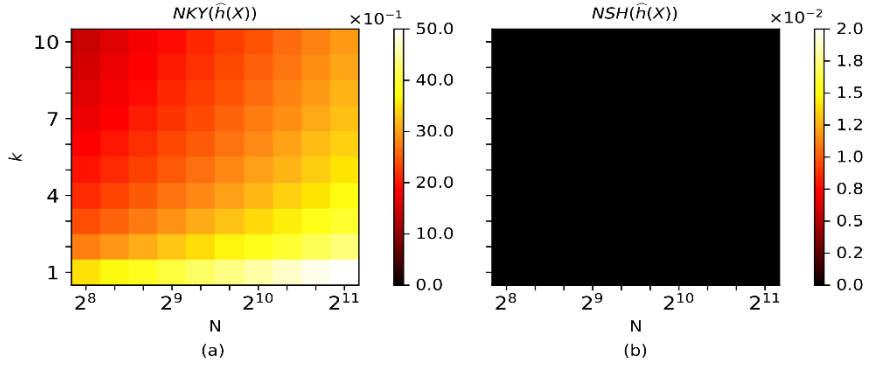
Bu benzetimde $N = 2048$ için, kullanılan kestirimcilerin parametreleri α' 'ın her değeri için sabit tutulmuştur. Histogram için $\varepsilon = 1$, Gauss çekirdek fonksiyonu için $|h| = 0.1$, Epanechnikov çekirdek fonksiyonu için $|h| = 0.2$, k-en yakın komşuluk sayısı $k = 10$, kpN entropi kestirimcisi için $k = 10$, $p_k/N = 0.1$ olarak belirlenmiştir.



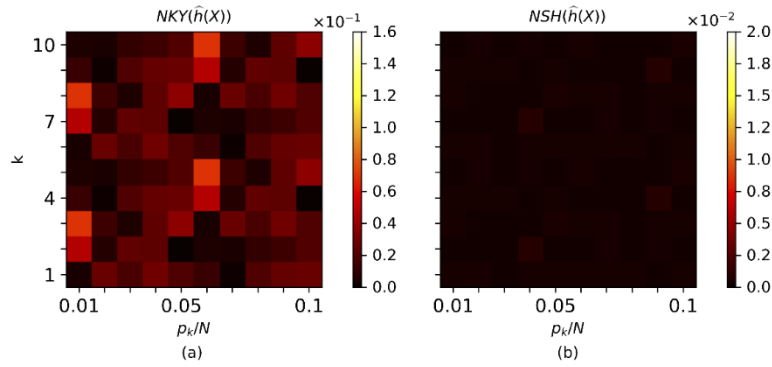
Şekil 21. İki durumlu sistemin entropisinin Gauss çekirdek yoğunluk kestirimcisi kullanılarak farklı $(N, |h|)$ parametreleri için (a) NKY değerleri ve (b) NSH değerleri



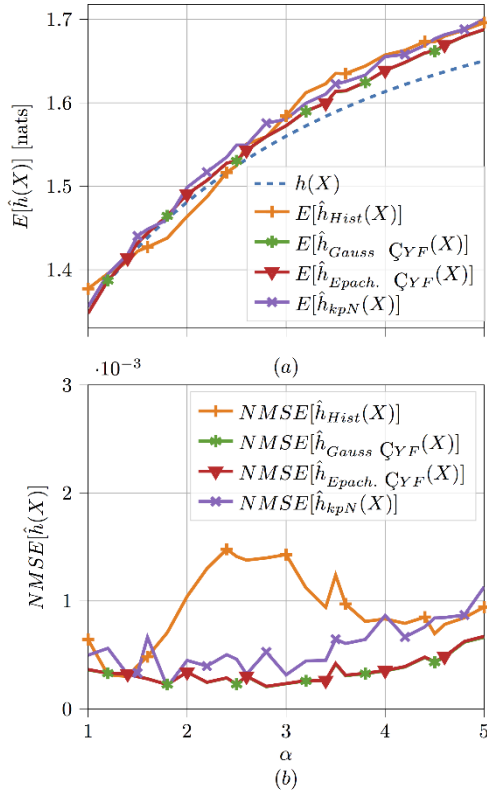
Şekil 22. İki durumlu sistemin entropisinin Epanechnikov çekirdek yoğunluk kestirimcisi kullanılarak farklı $(N, |h|)$ parametreleri için (a) NKY değerleri ve (b) NSH değerleri



Şekil 23. İki durumlu sistemin entropisinin k-en yakın komşuluklu entropi kestirimcisi kullanılarak farklı (N, k) parametreleri için (a) NKY değerleri ve (b) NSH değerleri



Şekil 24. İki durumlu sistemin entropisinin kpN entropi kestirimcisi kullanılarak örnek sayısı $N = 2048$ farklı $(p_k/N, k)$ sabit tutularak parametreleri için (a) NKY değerleri ve (b) NSH değerleri



Şekil 25. (a) İki durumlu sistemde artan gürültü seviyelerine göre histogram, Gauss çekirdek yoğunluk fonksiyonu, Epanechnikov çekirdek yoğunluk fonksiyonu, kpN yöntemi kullanılarak elde edilen (a) entropi kestirimlerinin ortalaması. Grafikte kesikli çizgi entropinin teorik değerini göstermektedir. (b) entropi kestirimlerinin normalize olmuş ortalama karesel hatası.

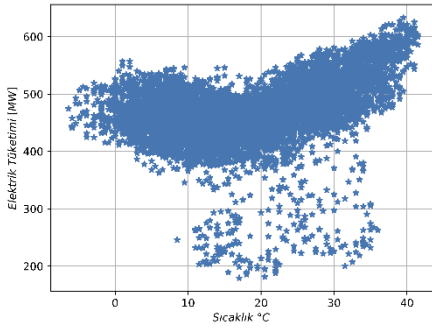
Şekil 25 incelendiğinde, gürültü arttıkça kestirim yanlılığının tüm kestirimcilerde yaklaşık olarak aynı seviyede olduğu görülebilir. Histogram kestirimcisinde gürültü arttıkça kestirim yanlılığının artıp azaldığı göze çarpmaktadır. NMSE değerleri Gauss çekirdek, Epanechnikov çekirdek yoğunluk, kpN kestirimcileri için tüm gürültü seviyelerinde oldukça düşüktür. Histogram kestirimcisinde etkinliğini gürültü seviyesine bağlı olarak değişmektedir. Bu açıdan bakıldığında histogram kestirimcisinin diğerine kıyasla daha az gürbüz olduğu söylenebilir.

Özet olarak, histogram tabanlı kestirimcide doğrusal olmayan sistemler arasındaki bilgi ölçüleri kestirilirken, kestirimlerin başarımları aralık çapına ve nokta sayısına bağlıdır. Sonlu bant genişliğine sahip Epanechnikov çekirdek yoğunluk kestirimcisinin, durum olasılık fonksiyonlarının kuyrukları Gauss dağılımına göre daha hızlı azalan iki durumlu sistem için kullanılması halinde beklenildiği gibi yanlılığın azaldığı gözlenmiştir. Bu nedenle Gauss çekirdek kestirimcisinin NKY değerleri Epanechnikov çekirdek kestirimcisi ile karşılaştırıldığında daha yüksektir. Veri dağılımı yerel olarak düzgün olan sistemlerde, boyut sayısı ne olursa olsun hızlı bir şekilde entropi bazlı ölçüleri yüksek doğrulukla kestiren k -en yakın komşuluklu entropi kestirimcisinin, doğrusal olmayan iki durumlu sistem için entropiyi yanı olarak kestirdiği saptanmıştır. kpN entropi kestirimcisi yerel olarak Gauss çekirdek fonksiyonunu ve global olarak da k -en yakın komşuluklu entropi kestirimcisini kullandığından doğrusal olmayan sistemlerde de entropiyi yüksek doğrulukla kestirmiştir. Entropi tabanlı ölçüler hızlı bir şekilde kestirilmek isteniyorsa hesaplama yükü

açısından k-en yakın komşuluklu entropi yöntemleri kullanılmalıdır. Buna karşılık, kestirimlerin yanlı olabileceği dikkate alınmalıdır. kpN yöntemi entropiyi yüksek doğrulukla kestirirken hesaplama yükü açısından kıyaslanan diğer yöntemlere göre daha ağırdır.

4.3 Sıcaklık-Elektrik Tüketimi Veri Kümesi

Kestirimcilerin başarımlarını analizleri gerçek bir veri kümesi üzerinde de test edilmiştir. Şekil 26'da gösterilen veri kümesi Türkiye'de bölgesel bir enerji dağıtıcısından alınmıştır. Hava (dış ortam) sıcaklığı ve elektrik tüketimi arasındaki nedensellik ilişkisi incelenmek istemiştir. Bu veri kümesi için sıcaklık artışının elektrik tüketimi artışına neden olduğu, başka bir deyişle, *Sıcaklık °C → Elektrik Tüketimi [MW]* nedensellik ilişkisi bilinmektedir [18].



Şekil 26. *Sıcaklık °C → Elektrik Tüketimi [MW]* ilişkisini gösteren gerçek bir veri kümesi.

Bu durumda kestirilen entropi vb. ölçülerin teorik değerleri bilinmediğinden kestirimcilerin başarımlarını analizinde veri kümesinden vekil veri kümeleri türetilecek ve hipotez testleri kullanılarak bu ilişkiye ait istatistiksel önem analizi yapılabilir [21]. Gözlemlenen iki farklı zaman serisi için kestirilen aktarım entropisi, bu iki değişkenin birbirine bağımlı ya da bağımsız olduğunu istatistiksel olarak ölçer.

Bu durumda aktarım entropisinin kestiriminde hata dağılımı bilinmediğinden,

" \mathcal{H}_0 : Sıcaklıktaki değişimlerin enerji tüketimi üzerinde bir etkisi yoktur."

sıfır hipotezi kestirim sürecinden bağımsız olarak reddedilebilmelidir. Nedensellik ilişkisini bozmak için elektrik tüketimi dizisi, E, sabit

tutulurken, sıcaklığa ilişkin veri dizisi S zamanda rastgele yeniden dizilerek elde edilen vekil seri dizisi (S^*) için, sıfır hipotezi

$$\mathcal{H}_0: TE_{S^* \rightarrow E} > TE_{S \rightarrow E}$$

ile denktir [23,24].

\mathcal{H}_0 'ın testi için 10000 adet vekil veri dizisi S^* için aktarım entropisi $TE_{S^* \rightarrow E}$ incelenen kestirimciler kullanılarak hesaplanmıştır. Sıfır hipotezinin reddi için istatistiksel önem sabiti $\alpha_{\mathcal{H}_0} = 0.05$ seçilmiştir. $p_{\mathcal{H}_0}$ hipotez testi ile elde edilen sonuçların anlamlılığını ölçer.

Tablo 1'de seçilen parametreler için tüm kestirimci tiplerinde $p_{\mathcal{H}_0}$ değerleri değişen gözlem sayıları için verilmiştir ($p_{\mathcal{H}_0} < \alpha_{\mathcal{H}_0}$ olduğunda sıfır hipotezi reddedilir).

Tablo 1'e bakıldığında, gözlem sayısı arttıkça kestirimcilerin sıfır hipotezini reddettiği görülmüştür. Histogram ve Gauss çekirdek yoğunluk yöntemlerinde $N \geq 2048$; Epanechnikov ve kpN yöntemlerinde $N \geq 4096$ için sıfır hipotezi doğru şekilde reddedilir. Histogram, Gauss ve Epanechnikov, kpN kestirim yöntemlerinde altta yatan olasılık yoğunluk fonksiyonlarının doğru temsil edilebilmesi için yüksek gözlem sayıları gerekir. Bu nedenle, nedenselliğin belirlenmesinde gözlem sayısının yüksek olması gerekliliği Tablo 1'den açıkça görülmektedir. k-yakın komşuluk yönteminde Tablo 1'deki tüm gözlem sayıları için sıfır hipotezi reddedilir. Bu durumda, yanlılığına rağmen k-yakın komşuluk yöntemi hızlı bir şekilde aktarım entropisinin kestirimi aracılığıyla yukarıdakine benzer hipotezlerin reddinde, başka bir deyişle süreçler arasında nedensellik ilişkisinin belirlenmesinde, etkin olarak kullanılabilir.

5. Sonuç

Bu makalede bir ya da birden fazla sistemin çıkışları kullanılarak entropi, ortak bilgi ölçüsü ve aktarım entropisi ölçüsü kestirilmiş ve ölçülerin farklı kestirim yöntemlerinin başarımlarını incelenmiştir. Entropi ve ondan türetilmiş ölçülerin kestiriminde, çekirdek yoğunluk kestirimcileri için parametre seçimi gözlemlenen örnek sayısı, bant genişliği ve çekirdek fonksiyonu tipi dikkate alınarak yapılmalıdır. Ayrıca, k-en yakın komşuluklu entropi ve kpN entropi kestirimcileri için gözlemlenen örnek sayısı ve komşuluk sayısının belirlenmesi ve histogram kestirimcisinde

Tablo 1: Farklı gözlem sayıları için yapılan \mathcal{H}_0 testi ve $p_{\mathcal{H}_0}$ sonuçları. Koyu değerler testin reddedildiği durumları göstermektedir.

Kestirimci	Kestirimci	$N = 1024$	$N = 2048$	$N = 4096$
Yöntemleri	Parametreleri	$p_{\mathcal{H}_0}$	$p_{\mathcal{H}_0}$	$p_{\mathcal{H}_0}$
Epanechnikov ÇYF	$ h = 4.5$	0.493	0.248	0.026
Gauss ÇYF	$ h = 2.9$	0.244	0.041	0.001
Histogram	$\varepsilon = 4$	0.209	0.036	0.002
kpN	$k = 3, p_k/N = 0.1$	0.172	0.103	0.028
kNN	$k = 3$	0.037	0.003	0.002

gözlemlenen örnek sayısı, bölüntülerin çaplarının saptanması önemlidir. Etkileşim halindeki rassal sistemler doğrusal ise k-en yakın komşuluklu entropi kestirimcisi, doğrusal değilse Epanechnikov çekirdek yoğunluk ve kpN entropi kestirimcisi kestirim başarımı açısından etkindir. Çekirdek yoğunluk ve kpN entropi kestirimcileri gözlemlerin yerel dağılımına çok bağlıdır. Doğrusal sistemler için Gauss çekirdek kestirimcisi uygunken, sınırlı bir küme içine dağılmış veri üreten sistemler için Epanechnikov kestirimcisi benzeri sınırlı bir küme üzerinde tanımlı çekirdeklere dayanan kestirimciler uygun olmaktadır. Doğrusal sistemlerde gözlem sayısı genellikle kestirim başarımının artması için en önemli parametredir. Bant genişliğini kontrol eden ε , $|h|$, p_k/N parametreleri belirgin bir en iyi değer göstermemiştir. Oysa doğrusal olmayan sistemlerde ε , $|h|$, p_k/N belirgin bir en iyi değere işaret etmekte, daha az veri ve daha doğru bir bant genişliği seçimi ile yüksek başarımlar mümkün olabilmektedir. Birbiri ile bağlı sistemlerin entropi ölçülerinin kestiriminin başarımı, sistemlerin birbirine nasıl bağlı oldukları ile de ilişkilidir. Daha uzak komşuluklu sistemler arasında ortak bilgi ölçüsü ve aktarım entropisi kestirimlerinin başarımları daha düşük gözlenmiştir.

Gerçek veri kümelerinde nedensellik ilişkisinin belirlenebilmesi için, aktarım entropisinin ilişkiyi koruyacak şekilde kestirilebilmesi önemlidir. İncelenen kestirimcilerin bu anlamdaki başarımlarını kullanılan gözlem sayısına sıkı şekilde bağlıdır. k-en yakın komşuluk yönteminin az sayıda gözlem ile nedensellik ilişkisini istatistiksel önem açısından doğru şekilde belirleyebildiği gösterilmiştir.

Etkileşim halindeki karmaşık modellerin arasındaki bilgi akışını en doğru şekilde

kestirmemiz ekonomiden, biyolojiye, mekanikten elektrige her alanda karşılaşılan karmaşık sistemlerin davranışını anlamamıza ışık tutacaktır.

Ek: İki durumlu Sistemin Olasılık Yoğunluk Fonksiyonu

$$\begin{aligned} dx &= \mathcal{F}(x, t)dt + \mathcal{G}(x, t)dW(t) \\ x(0) &= x_0 \quad x \in \mathbb{R}^n, x_0 \sim f_0(x_0) \end{aligned} \quad (33)$$

rassal diferansiyel denklemler ile tanımlanmış bir sistem olsun. Burada, $\mathcal{F}(x, t)$ sistemin sürüklenme dinamiğini, $\mathcal{G}(x, t)$ ise difüzyon terimidir. Başlangıç koşulunun dağılımı $f(x_0, 0) = f_0(x_0)$ olmak üzere, bu formda tanımlanmış bir sistemin olasılık yoğunluk fonksiyonu $f(x, t)$,

$$\begin{aligned} \frac{\partial f(x, t)}{\partial t} &= \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} (a_{ij}f) + \dots \\ &\dots - \sum_{i=1}^n \frac{\partial}{\partial x_i} (\mathcal{F}_i f) \quad \forall t > 0, x \in \mathbb{R}^n \end{aligned} \quad (34)$$

denklemini sağlar. En genel formda Fokker-Planck (FP) işleci,

$$\begin{aligned} \mathcal{L}f &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} (a_{ij}(x)(\cdot)) \\ &\quad - \sum_{i=1}^n \frac{\partial}{\partial x_i} (\mathcal{F}_i(x)(\cdot)) - \frac{\partial}{\partial t} (\cdot) \end{aligned} \quad (35)$$

ile verilir [16]. $f_{st}(x)$, sürekli haldeki olasılık yoğunluk fonksiyonunu gösterebilir. Bu durumda

$$\begin{aligned} \mathcal{L}f_{st} &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} (a_{ij}(x) f_{st}) \\ &\quad - \sum_{i=1}^n \frac{\partial}{\partial x_i} (\mathcal{F}_i(x) f_{st}) = 0 \end{aligned} \quad (36)$$

olur. Denklem 36 \mathcal{L}_i $i = 1, 2, \dots, N$ kısmi diferansiyel operatörleri cinsinden yazılabiliyorsa,

$$\begin{aligned} \mathcal{L}f_{st} &= \sum_{i=1}^n \mathcal{L}_i [(\beta_i(x) f_{st}(x) \\ &\quad + \gamma_i(x) \frac{\partial f_{st}(x)}{\partial x})] = 0 \end{aligned} \quad (37)$$

$f_{st}(x)$, $\beta_i(x) f_{st}(x) + \gamma_i(x) \frac{\partial f_{st}(x)}{\partial x} = 0$, $i=1, 2, \dots, N$ denklemlerini sağlar. Diferansiyel denklemlerin çözümü

$$f_{st}(x) = c \prod_{i=1}^n \exp\left(- \int_0^{x_i} \frac{\beta_i(s)}{\gamma_i(s)} ds\right) \quad (38)$$

ile verilir. Denklem 31 için denklem 36'daki operatör

$$\mathcal{L}f_{st} = \frac{\alpha}{2} \frac{\partial^2 f_{st}}{\partial x^2} - \frac{\partial}{\partial x} [0.5(x - x^3) f_{st}] \quad (39)$$

formundadır. Denklem 39 düzenlenerek,

$$\frac{\partial}{\partial x} \left(\frac{-1}{2} (x - x^3) f_{st}(x) + \frac{\alpha}{2} \frac{\partial f_{st}(x)}{\partial x} \right) = 0 \quad (40)$$

elde edilir. İki durumlu sistem için $\mathcal{L}_1 = \frac{\partial}{\partial x}$, $\beta_1(x) = \frac{-1}{2}(x - x^3)$ ve $\gamma_1(x) = \frac{\alpha}{2}$ olur. Denklem 38 iki durumlu sistem için hesaplandığında

$$f_{st}(x) = c_1 \exp\left(- \frac{1}{\alpha} \left[-\frac{x^2}{2} + \frac{x^4}{4} \right] \right) \quad (41)$$

elde edilir. $\int f_{st}(x) dx = 1$ olduğu için,

$$c_1 \int_{-\infty}^{\infty} \exp\left(- \frac{1}{\alpha} \left[-\frac{x^2}{2} + \frac{x^4}{4} \right] \right) dx - 1 = 0 \quad (42)$$

denklemini sağlayan normalizasyon katsayısı

$$c_1 = \frac{2}{\pi \left(I_{-\frac{1}{4}} \left(\frac{1}{8\alpha} \right) + I_{\frac{1}{4}} \left(\frac{1}{8\alpha} \right) \right)} e^{\frac{-1}{8\alpha}} \quad (43)$$

bulunur. Denklem 43'teki c_1 denklem 41'de yerine konularak düzenlendiğinde denklem 32 elde edilir.

Kaynakça

- [1] Shannon, C. E. 1948. A mathematical theory of communication. Bell System Technical Journal, Cilt. 27 (3), s. 379-423. DOI: 10.1002/j.1538-7305.1948.tb01338.x
- [2] Cover, T. M. & Thomas, J. A., 2012. Elements of information theory, John Wiley & Sons. 2nd Edition, 748s.
- [3] Schreiber, T., 2000. Measuring information transfer. Physical review letters, Cilt 85 (2), s. 461-464. DOI: 10.1103/PhysRevLett.85.46
- [4] Duan, P., Yang, F., Chen, T., Shah, S. L., 2013. Direct causality detection via the transfer entropy approach. IEEE transactions on control systems technology, Cilt. 21(6), s. 2052-2066. DOI: 10.1109/TCST.2012.2233476
- [5] Wibral, M., Vicente, R., Lindner, M., 2014. Directed Information in Neuroscience Transfer entropy in neuroscience. ss 3-36. Wibral, M., Vicente, R., Lindner, M., ed. 2014. Directed information measures in neuroscience, Springer, Berlin, Heidelberg, 224 s.
- [6] Spinney, R. E., Prokopenko, M., Lizier, J. T., 2017. Transfer entropy in continuous time, with applications to jump and neural spiking processes. Physical Review E, Cilt. 95(3), 032319. DOI: 10.1103/PhysRevE.95.032319
- [7] Xiong, W., Faes, L., Ivanov, P. C., 2017. Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations. Physical Review E, Cilt. 95(6), 062114. DOI: 10.1103/PhysRevE.95.062114
- [8] Kraskov, A., Stögbauer, H., Grassberger, P., 2004. Estimating mutual information. Physical review E, Cilt. 69(6), 066138. DOI: 10.1103/PhysRevE.69.066138
- [9] Lombardi, D., & Pant, S., 2016. Nonparametric k-nearest-neighbor entropy estimator. Physical Review E, Cilt. 93(1), 013310. DOI: 10.1103/PhysRevE.93.013310
- [10] Gencaga, D., Knuth, K. H., Rossow, W. B., 2015. A recipe for the estimation of information flow in a dynamical system. Entropy, Cilt. 17(1), s. 438-470. DOI: 10.3390/e17010438
- [11] Diego, D., Haaga, K. A., Hannisdal, B., 2019. Transfer entropy computation using the Perron-Frobenius operator. Physical Review E, Cilt. 99(4), 042212. DOI: 10.1103/PhysRevE.99.042212
- [12] Wand, M. P. ve Jones, M. C., 1994. Kernel smoothing, Chapman and Hall/CRC, 224 s.

- [13] Kozachenko, L. F. ve Leonenko, N. N., 1987. Sample estimate of the entropy of a random vector. *Problemy Peredachi Informatsii, Cilt 23(2)*, s. 9-16.
- [14] Zhu, J., Bellanger, J. J., Shu, H., Le Bouquin Jeannès, R., 2015. Contribution to transfer entropy estimation via the k-nearest-neighbors approach. *Entropy, Cilt. 17(6)*, s. 4173-4201. DOI: 10.3390/e17064173
- [15] Kaiser, A. ve Schreiber, T., 2002. Information transfer in continuous processes. *Physica D: Nonlinear Phenomena, Cilt. 166(1-2)*, s. 43-62. DOI: 10.1016/S0167-2789(02)00432-3
- [16] Soong T. T., 1973. *Random differential equations in science and engineering*, New York, U.S.A.: Academic Press, 333 s.
- [17] Shanmugan K. S., Breipohl, A. M., 1988. *Random Signals: Detection, Estimation and Analysis*, 688 s.
- [18] Mooij, J. M., Peters, J., Janzing, D., Zscheischler, J., & Schölkopf, B., 2016. Distinguishing cause from effect using observational data: methods and benchmarks. *The Journal of Machine Learning Research, Cilt. 17(1)*, s. 1103-1204. DOI: 10.5555/2946645.2946677
- [19] Daub, C. O., Steuer, R., Selbig, J., & Kloska, S., 2004. Estimating mutual information using B-spline functions—an improved similarity measure for analysing gene expression data. *BMC bioinformatics, Cilt. 5(1)*, 118. DOI:10.1186/1471-2105-5-118
- [20] Sun, J., & Bollt, E. M., 2014. Causation entropy identifies indirect influences, dominance of neighbors and anticipatory couplings. *Physica D: Nonlinear Phenomena, Cilt. 267*, s. 49-57. DOI: 10.1016/j.physd.2013.07.001
- [21] Bossomaier, T., Barnett, L., Harré, M., & Lizier, J. T., 2016. *An introduction to transfer entropy*. Cham: Springer International Publishing, s. 65-95.
- [22] Kloeden, P. E., & Platen, E., 2013. *Numerical solution of stochastic differential equations*. Cilt. 23. Springer Science & Business Media, 636 s.
- [23] Vicente, R., Wibral, M., Lindner, M., & Pipa, G., 2011. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *Journal of computational neuroscience, Cilt. 30(1)*, s. 45-67. Doi: 10.1007/s10827-010-0262-3
- [24] Lizier, J. T., Heinzle, J., Horstmann, A., Haynes, J. D., & Prokopenko, M., 2011. Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fMRI connectivity. *Journal of computational neuroscience, Cilt. 30(1)*, s. 85-107. Doi: 10.1007/s10827-010-0271-2