

# Hatalı Güneş Enerjisi Üretim Ölçümlerinin Tespit Edilmesi

## Identification of Bad Solar Power Generation Data

Deniz ŞENGÜL, Erencan DUYMAZ, Murat GÖL

Elektrik Elektronik Mühendisliği Bölümü

Orta Doğu Teknik Üniversitesi

Ankara, Türkiye

deniz.sengul@metu.edu.tr, erencan.duyamaz@metu.edu.tr, mgol@metu.edu.tr

**Özetçe—** Bu bildiride, güneş enerjisi ölçüm verilerinin kurtarılması sürecinde kullanılmak üzere yanlış verilerin tespit edilmesi için özgün bir yöntem sunulmaktadır. Güneş enerjisi ölçüm kayıt cihazlarında yaşanan bir arızadan dolayı, uzun bir süreyi kapsayan güneş enerjisi üretim verilerinin kurtarılması gerekli olabilmektedir. Kurtarılan bu bilgilerin doğru olduğundan emin olmak için, günlük güneş enerjisi üretim eğrisinin incelenmesi gerekmektedir. Veri miktarının fazlalığından dolayı bu işlemin otomatik olarak gerçekleştirilmesi gereklidir. Günlük üretim eğrisinin kosinüs işaretine benzerliği yapılan incelemelerde gözlenmiştir. Bu çalışmada kurtarılan veriler referans kosinüs işareti ile karşılaştırılarak, hatalı olduğundan şüphelenilen veri kümesinin tespit edilmesi amaçlanmaktadır. Bu bağlamda düzgelenmiş kalıntılar testi temel alınmıştır.

**Anahtar Kelimeler —** Yanlış Verilerin Belirlenmesi, Güneş Enerjisi, Günlük Üretim Eğrisi, Veri Kurtarma

**Abstract—** In this paper, a novel method has been presented for the bad data identification of the data recovery process from a solar power measurement data set. Due to failures in solar power measurement logging devices, a long period of solar power generation data might necessary to be recovered. In order to ensure that the recovered data is the correct one, the daily solar generation curve has to be investigated. Due to the high data amount, this control process has to be done automatically. It is observed that the daily solar generation curve is much similar to cosine wave. In this study, the recovered data is compared to reference cosine wave and it is aimed to identify the suspicious data group. In this context, this work is based on the normalized residuals test.

**Keywords —** Bad Data Identification, Solar Power, Daily Generation Curve, Data Recovery

### I. GİRİŞ

Dünya çapında yenilenebilir enerji kaynaklarına verilen önem ve bunun sonucunda kurulan yenilenebilir enerji santralleri her geçen gün artmaktadır. Yenilenebilir enerji alanına yapılan yatırımların geri dönüş süresi alanda yaşanan teknolojik gelişmelerden dolayı düşmeye devam etmektedir. Hatta maliyetteki düşüşler, bu alana yapılan teşviklerin kaldırılması tartışmalarını da beraberinde getirmiştir. Lisanslı ve lisanssız Güneş Enerjisi Santrallerinin (GES) yanında çatı

üstü güneş enerji sistemleri de yaygın olarak kurulmaktadır. ABD’de ev ve iş yerlerine kurulan 3.4 GW kapasite ile küçük çaptaki fotovoltaik market 2016 yılında bir rekora imza atmıştır [1]. Bu yüzden özellikle küçük ölçekli güneş enerji sistemlerinin üretim bilgilerinin kayıt altına alınması, dağıtık üretimin sisteme etkilerinin incelenmesi açısından büyük önem taşımaktadır.

Piyasadaki ticari cihazların kayıt altına aldığı bu verilere, cihaz çalıştığı sürece uygun veri formatında erişilebilmektedir. Ancak cihazın arızası durumunda, cihaz hafızasından erişilen bilgiler kullanıcının veri formatını çözmesi ile kullanılabilir bir hale getirilebilir. Format bilinse bile, bilinmeyen sebeplerden ötürü bazı dosyalarda doğru veriye ulaşılamayabileceği deneyimler ile gözlenmiştir. Uzun süreler boyunca alınmış verilerin doğruluğunun tek tek kontrol edilmesi mümkün olmadığı için yanlış bilgi içeren veri dosyalarını tek seferde belirleyecek bir yöntemle ihtiyaç duyulmuş ve önerilen yöntem geliştirilmiştir.

Güç sistemlerinde kullanılan ölçümlerin her zaman hata içerdiği bilinen bir gerçektir. Ancak bu hataların birbirinden bağımsız olması ve Gauss dağılımına sahip olması kullanılan ölçümlerden hangisinin hatalı olduğunu bulmaya yardımcı olmaktadır. Ağırlıklandırılmış En Küçük Kareler (AEKK) kestirimi ile sisteme ait durum kestirimlerinin elde edilmesi mümkündür [2] – [3]. Kestirim sonuçları kullanılarak literatürde düzgelenmiş (normalize edilmiş) kalıntılar testi (DKT) [4] olarak bilinen yöntem sayesinde hatalı ölçümlerin tespit edilmesi mümkündür. Bu bildiride DKT temel alınarak hatalı güneş enerjisi üretim dosyalarının tespiti amaçlanmıştır.

Güneş enerjisi günlük üretim eğrisi yılın farklı aylarında güneşlenmeye göre değişiklikler göstermektedir. Tamamen güneşli bir günde kesintisiz bir şekilde sahipken, bulutlu günlerde kesikli bir şekilde seyir etmektedir. Özellikle kış aylarına ait bulutlu günlerdeki eğrilerin doğruluğundan emin olmak için karşılaştırma yapılabilecek bir referans sinyali gerekmektedir. Elde edilen günlük üretim eğrileri, günün en yüksek üretim gücüne göre düzgelendiğinde, bir kosinüs işaretine benzediği gözlenmiştir. Bu bildiride sunulan yöntemde ise, günlük üretim eğrileri bu referans kosinüs işareti karşılaştırılarak tespit süreci gerçekleştirilmiştir.

Kurtarılan verilerden yanlış olanlar incelendiğinde iki tür yanlış veriye rastlanmıştır. Birinci tip yanlış veride, elde edilen günlük üretim eğrisinin bir kısmının yanlış olduğu, ikinci tip veride ise tamamının yanlış olduğu gözlenmiştir. Dolayısıyla karşılaştırma sonucunda her iki tip yanlış veri referans işaretten büyük miktarda sapmaktadır. Bulutlu günlerde meydana gelen sapmalarına yöntem tarafından başarıyla ayırt edildiği gözlenmiştir.

Bildirinin içeriği şu şekildedir; 2. Kısımda önerilen yöntem ayrıntıları ile verilmiştir. 3. Kısım gerçek verilerle yöntemi doğrularken 4. Kısım sonuç ve değerlendirmeleri içermektedir.

## II. ÖNERİLEN YÖNTEM

Önerilen yöntem ile hatalı olduğundan şüphelenilen veri kümelerinin tespit edilmesi amaçlanmaktadır. Bu bağlamda DKT [4] temel alınmıştır. Bu bölümde öncelikle DKT açıklanacak, sonrasında önerilen yöntem ile şüpheli veri dosyalarının tespiti ayrıntılarıyla verilecektir.

### A. Düzgelenmiş Kalıntılar Testi

Lineer bir sistemde ölçümler ile durumlar arasındaki ilişki aşağıdaki gibi ifade edilebilir.

$$z = Hx + e \quad (1)$$

(1)'de  $z$  ( $m \times 1$ ) ölçüm vektörü olarak kullanılırken,  $x$  ( $n \times 1$ ) durum vektörünü ifade etmektedir.  $H$  ( $m \times n$ ) ise durumlar ve ölçümler arasındaki lineer bağıntıyı ifade etmektedir. Son olarak  $e$  ( $m \times 1$ ) ölçüm hatası vektörü olarak kullanılmıştır.

Ölçüm hatalarının birbirinden bağımsız ve Gauss olduğu kabul edilirse, durum kestirimleri ağırlıklandırılmış en küçük kareler (weighted least squares – AEKK) kestirimcisi kullanılarak bulunabilir. AEKK'nin en önemli özelliği, ölçümlerde sadece Gauss hata olması durumunda en iyi doğrusal yansız kestirimci (Best Linear Unbiased Estimator - BLUE) olmasıdır [5]. Bu bağlamda durum kestirimi vektörü,  $\hat{x}$ , aşağıdaki ilişkinin çözümüne eşittir.

$$\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} z \quad (2)$$

(2)'de  $R$  ( $m \times m$ ) ölçüm kovaryans matrisini ifade etmektedir. (2) göz önüne alınarak kalıntılar,  $r$ , ile ölçüm hataları arasındaki bağıntı aşağıdaki gibi bulunabilir.

$$\begin{aligned} r &= z - H\hat{x} \\ r &= (Hx + e) - H(H^T R^{-1} H)^{-1} H^T R^{-1} (Hx + e) \\ r &= e - H(H^T R^{-1} H)^{-1} H^T R^{-1} e \\ r &= (I - H(H^T R^{-1} H)^{-1} H^T R^{-1}) e \rightarrow r = Se \end{aligned} \quad (3)$$

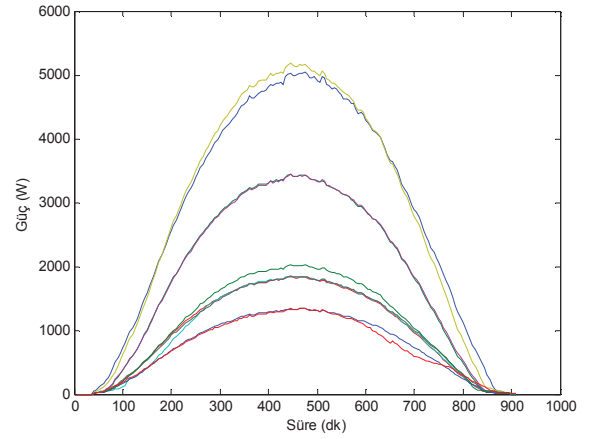
Kalıntıların beklenen değerleri ve kovaryans matrisi aşağıdaki gibi tanımlanır.

$$\begin{aligned} E[r] &= E[Se] = SE[e] = 0 \\ \text{cov}(r) &= E[rr^T] = SE[ee^T] S^T = SRS^T = \Omega \end{aligned} \quad (4)$$

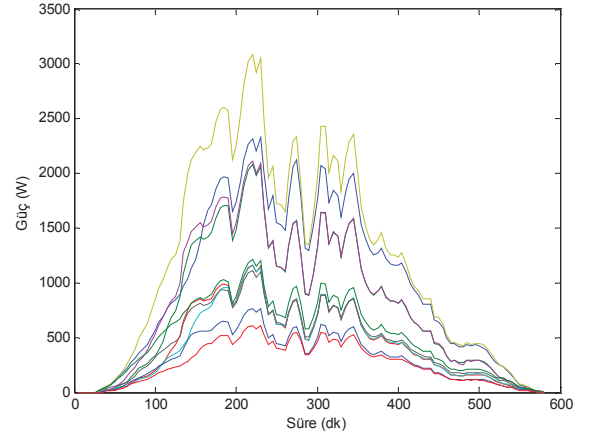
Kalıntı kovaryans matrisinin,  $\Omega$ , tersi alınamaz. Bu yüzden de kullanıldığı analizlerde sadece diyagonal elemanları (kalıntı varyansları) hesaba katılır [2]. Her bir kalıntı, kendi varyansının karekökü ile düzgelendiğinde, belirli bir eşik değerini aşan değerlerin bulunması hatalı verilerin varlığına işaret eder. Düzgelenmiş kalıntıların en büyüğü ise hatalı veri olarak işaretlenir.

### B. Şüpheli Dosyaların Tespiti

Bu çalışmada, DKT temel alınarak, güneş enerjisi üretim verilerindeki hata şüphesi taşıyan dosyaların tespit edilmesi amaçlanmıştır. İlgilenilen sistem üniversite kampüsünde kurulu olan, paralel bağlanmış 10 çeviricinin verilerini kaydetmektedir. Her bir dosya belirli bir güne ait güneş enerjisi üretimini içermektedir. Şekil-1'de, çeşitli mevsimsel koşullarda kaydedilmiş üretim verileri görülmektedir.



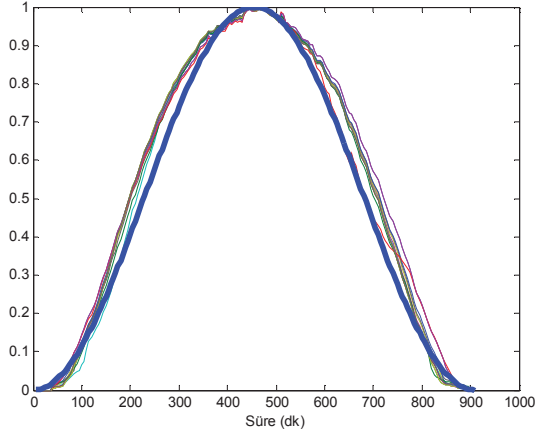
(a) Güneşli bir güne ait üretim verileri.



(b) Bulutlu bir güne ait üretim verileri.

Şekil-1. Çeşitli mevsimsel koşullarda kaydedilmiş üretim verileri.

Yapılan incelemeler sonunda, güneşli bir günde üretimin zamana göre değişiminin Dünya'nın kendi etrafındaki hareketinden ötürü bir kosinüs fonksiyona benzer olduğu görülmüştür. Şekil-2'de, tepe noktasına göre düzgelenmiş üretim eğrilerinin kosinüs fonksiyonuyla karşılaştırması görülmektedir.



Şekil-2. Tepe noktasına göre düzeltilmiş üretim eğrilerinin kosinüs fonksiyonuyla karşılaştırması (kosinüs işareti kalın olarak çizilmiştir).

Şekil-2’de görülen benzerlikten yola çıkılarak, beklenen güneş enerjisine dayalı üretim ile alınan ölçümler arasındaki matematiksel ilişki aşağıdaki gibi modellenmiştir.

$$z = x + v \quad (5)$$

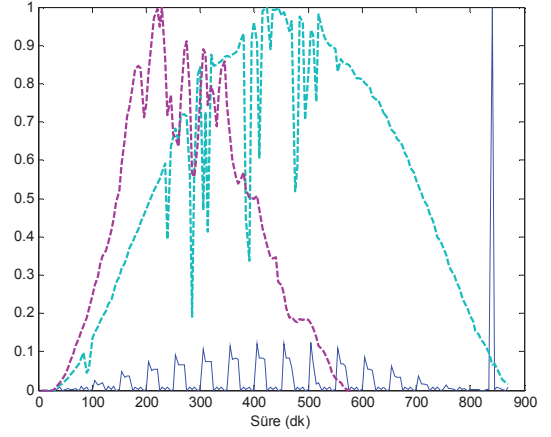
(5)’te  $z$  alınan ölçümleri,  $v$  ise hava durumundaki değişikliklerden ötürü oluşan enerji üretimindeki sapmaları ifade etmektedir. Bu sapmalar, ihmal edilebilir büyüklükte olabilecekleri gibi, Şekil-2’de görüleceği üzere beklenen kosinüs karakteristiğinin bozulmasına da sebep olabilmektedir.

Hava durumundan kaynaklı sapmalar ne kadar ciddi boyutta olursa olsun, hatalı verilerle kıyaslanamayacakları Şekil-3’te verilen karşılaştırmada görülmektedir. Bu bağlamda, her bir dosyaya ait kalıntıların, ortalama kalıntı varyansına oranının belirli bir eşik değeri ile karşılaştırılmasıyla hata şüphesi taşınan veri dosyalarının tespit edilebileceği görülmektedir. Ortalama kalıntı varyansı aşağıdaki gibi hesaplanabilir.

$$\sigma_i^2 = \frac{1}{N} \sum_{k=1}^N [z(k) - x(k)]^2 \quad (6)$$

$$\sigma_{ort}^2 = \frac{1}{T} \sum_{i=1}^T \sigma_i^2$$

(6)’da  $\sigma_{ort}^2$ , ortalama kalıntı varyansını ifade etmektedir.  $N$  bir dosyadaki toplam veri sayısıdır, ve her günün güneşlenme süresi farklı olduğu için dosyadan dosyaya değişmektedir.  $T$  toplam dosya sayısıdır, ve ölçüm alınan gün sayısına denk gelmektedir. Daha önce de bahsedildiği gibi, ilgilenilen sistemde birden fazla çevirici olduğundan, her dosyada her bir çeviriciye ait varyans hesaplanmakta, bunlardan en büyüğü o dosyaya ait varyans olarak alınmaktadır.  $\sigma_{ort}^2$  hesaplanırken hatalı veri içermeyen ve hem bulutlu hem güneşli günleri içeren bir dosya grubu kullanılmıştır.



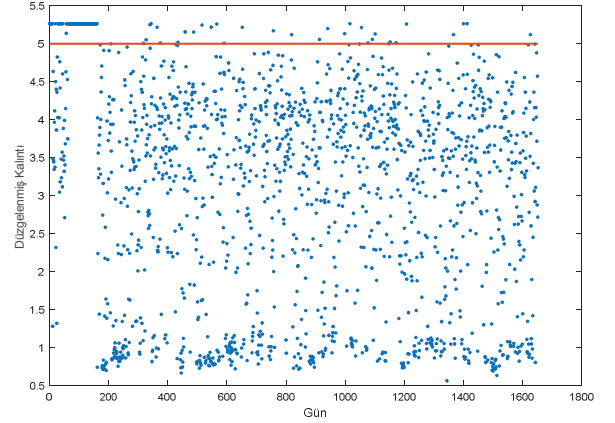
Şekil-3. Tepe noktasına göre düzeltilmiş, bulutlu iki güne ait üretim eğrilerinin hatalı verilerle karşılaştırması (kesikli çizgiler bulutlu günlere ait verileri ifade etmektedir).

### III. ÖNERİLEN YÖNTEMİN DOĞRULANMASI

Doğrulama için kullanılan veriler üniversite kampüsünde kurulu olan toplam anma değeri 50 kW olan 10 ayrı çeviriciden alınmıştır.

Veri dosyaları 01.01.2012 tarihinden 31.12.2016 tarihine kadar sistemin çalıştığı 1652 günü içermektedir. Bu çalışmada şüphelenme eşik değeri 5 olarak seçilmiştir.  $\sigma_{ort}^2$  ise 0.19 olarak hesaplanmıştır.

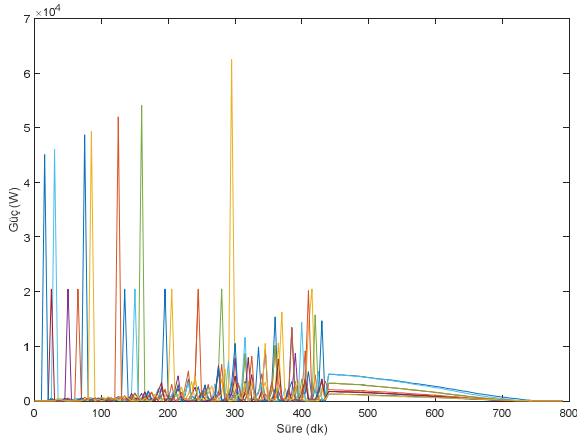
Şekil-4’te her bir dosya için hesaplanmış düzeltilmiş kalıntılar gösterilmektedir. 1652 günden 158’ine ait dosyalar önerilen yöntem vasıtasıyla hatalı veri olarak işaretlenmiştir.



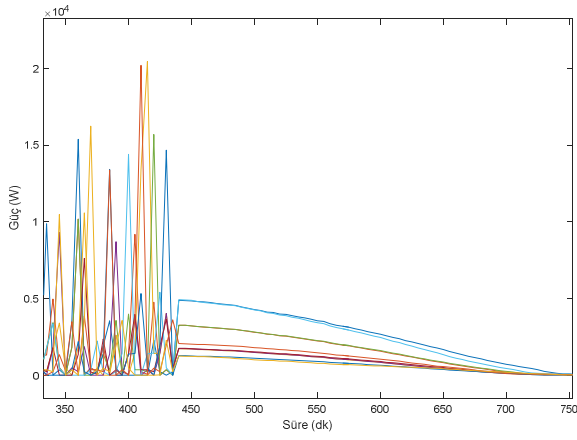
Şekil-4. Veri Toplanan her bir güne ait düzeltilmiş kalıntılar.

Hatalı olarak işaretlenmiş dosyalara örnekler Şekil-5’te ve Şekil-6’da verilmiştir. Şekil-5 21.06.2012 tarihinde alınmış verileri tam ölçekli ve yakınlştırılmış olarak göstermektedir. Eğrinin ilk kısmı verilerin yanlış dönüşümünden kaynaklı olduğu tahmin edilen hatalı verilerle başlamış, sonlarına doğru veriler düzelse de dosyanın bir bütün olarak kullanılamayacak durumda olduğu gözlemlenmiştir. Şekil-6’da ise kayıttan kaynaklı olduğu düşünülen, hatalı olduğu eğrilerin şekillerinden ve güç değerinden de belli olan bir örnek verilmiştir. Şekil-1.b’de verilen bulutlu güne ait üretim eğrisi, üretimdeki salınımlara

rağmen olması gerektiği gibi yöntem tarafından hatalı olarak işaretlenmemiştir.



(a) Tam ölçek.



(b) Yakınlaştırılmış.

Şekil-5. 21.06.2012 tarihli dosyaya ait veriler.

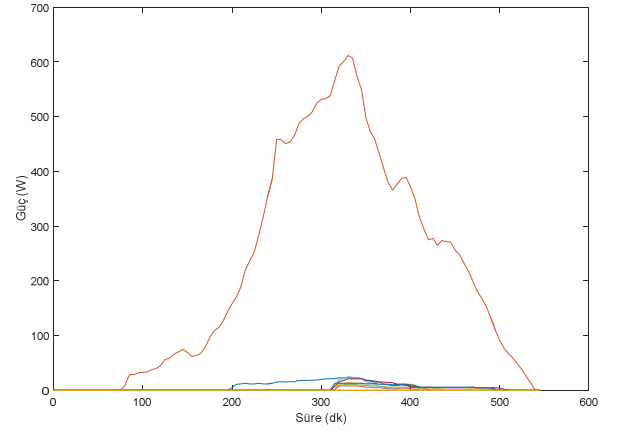
Son olarak hata eşiği olan 5 civarındaki düzelenmiş kalıntılara sahip 367 veri dosyası elle incelenmiştir. Yapılan incelemenin sonucunda önerilen yöntemin hatalı olarak işaretlediği veri dosyalarının hepsinin hatalı olduğu gözlenmiştir. Sadece 9 dosya hatalı veri içermesine rağmen hatasız olarak işaretlenmiştir. Bu bağlamda önerilen yöntemin yüksek hassasiyete sahip olduğu sonucuna varılmıştır. Eşik değerinin eniyilenmesiyle performansın artırılacağı öngörülmektedir.

#### IV. SONUÇ

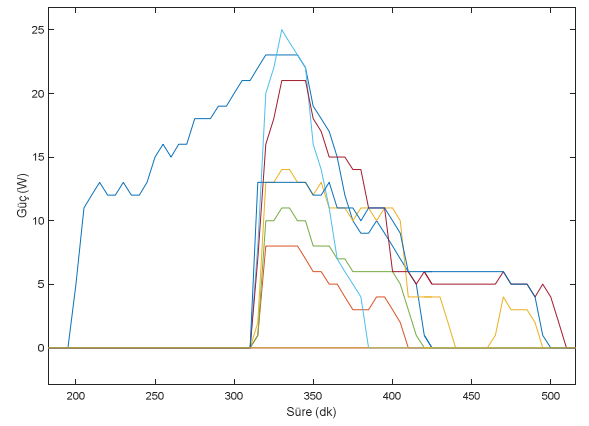
Bu çalışmada, karşılaşılan bir probleme çözüm üretmek amacıyla DKT'ne dayanan bir yöntem ile güneş üretimine ait hatalı verilerin tespiti amaçlanmıştır. Bu bağlamda alınan ölçümlerin, beklenen değer olarak kabul edilen kosinüs işaretinden sapması kalıntı olarak ele alınmıştır. Geleneksel olarak DKT tarafından 3 kabul edilen hata şüphe eşiğinin 5 olarak atanması gerektiği görülmüştür. Bunun başlıca sebebi, bulutlu günlerde üretim eğrisinin, beklenen değerlerden büyük miktarlarda sapabilmesidir.

Gerçek verilerle yapılan testler sonucunda, yöntemin bulutlu günlere ait dosyalar ile hatalı verileri birbirinden

ayırabildiği görülmüştür. Dolayısıyla, yöntemin başarıyla çalıştığı sonucuna ulaşılmıştır.



(a) Tam ölçek.



(b) Yakınlaştırılmış.

Şekil-6. 08.01.2013 tarihli dosyaya ait veriler.

#### KAYNAKÇA

- [1] Bloomberg New Energy Finance, "Sustainable Energy in America 2017 Factbook", 2017.
- [2] A. Abur and A. Gomez-Exposito, "Power System State Estimation: Theory and Implementation", book, Marcel Dekker, 2004.
- [3] A. Monticelli, "Electric Power System State Estimation," Proceedings of the IEEE, vol. 88, no. 2, pp. 262-282, Feb. 2000.
- [4] A. Monticelli and A. Garcia, "Reliable Bad Data Processing for Real-Time State Estimation", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-102, No. 5, May 1983, pp. 1126-1139.
- [5] A. C. Aitken, "On Least Squares and Linear Combinations of Observations", Proc. Royal Society of Edinburg, 1935, vol. 35, pp. 42-48.