

tepaV

The Economic Policy Research Foundation of Turkey

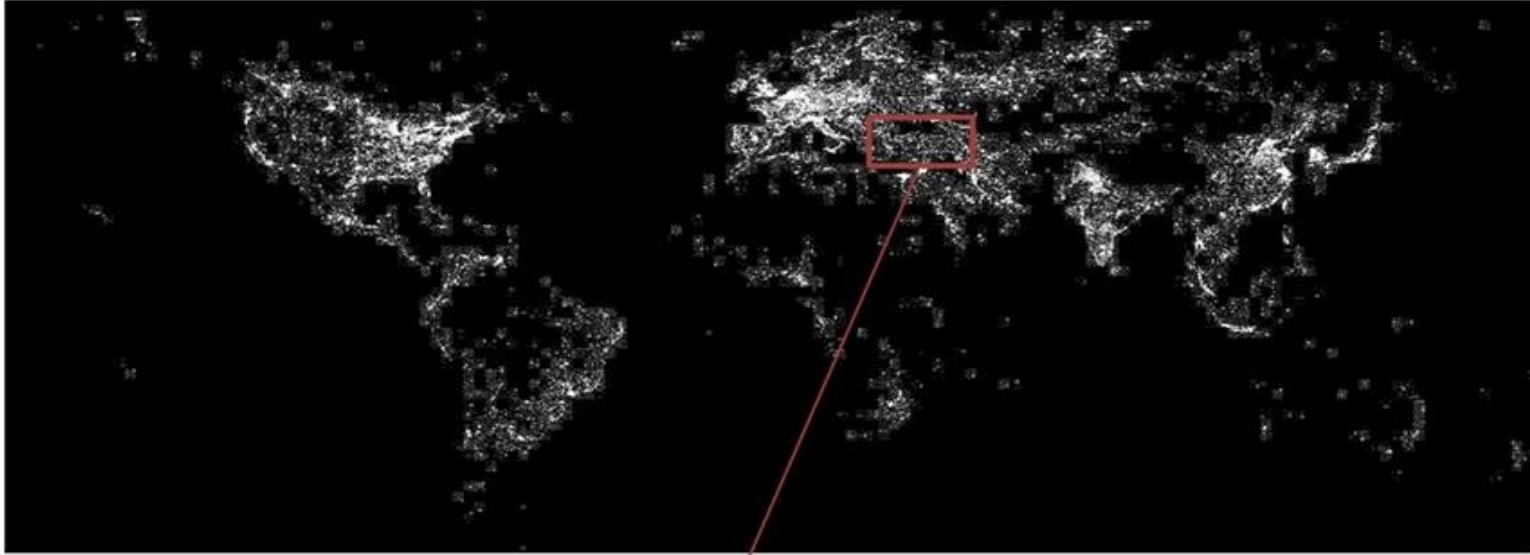
Gelişmişlik Göstergesi Olarak Gece Işıkları: İl Bazında GSYH Tahmini

Seda Başhoş

Ankara

Kasım 30, 2016

Gece Işıkları Nedir?

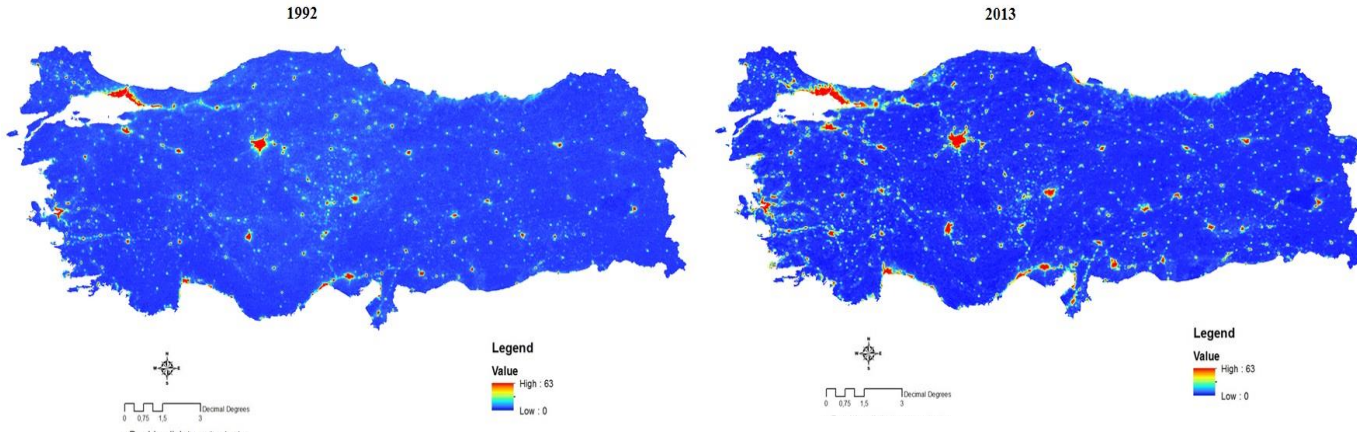


Temsili Bir Gösterge...

- Ekonomik büyüklük
- Ekonomik Gelişim endeksi
 - Bölgesel karşılaştırmalar
- Şehirleşme ve şehir dinamikleri
- Kayıtdışı ekonominin ölçülmesi
- ...

Numerik Verinin Elde Edilmesi

1. Yüksek çözünürlüklü fotoğrafların maskelenmesi
2. 0'dan 255' kadar ölçeklendirilmiş dijital sayıların RGB formatında 0'dan 63'e yeniden ölçeklendirilmesi



3. Ölçek aralıklarına göre piksel sayılarının ağırlıklandırılması

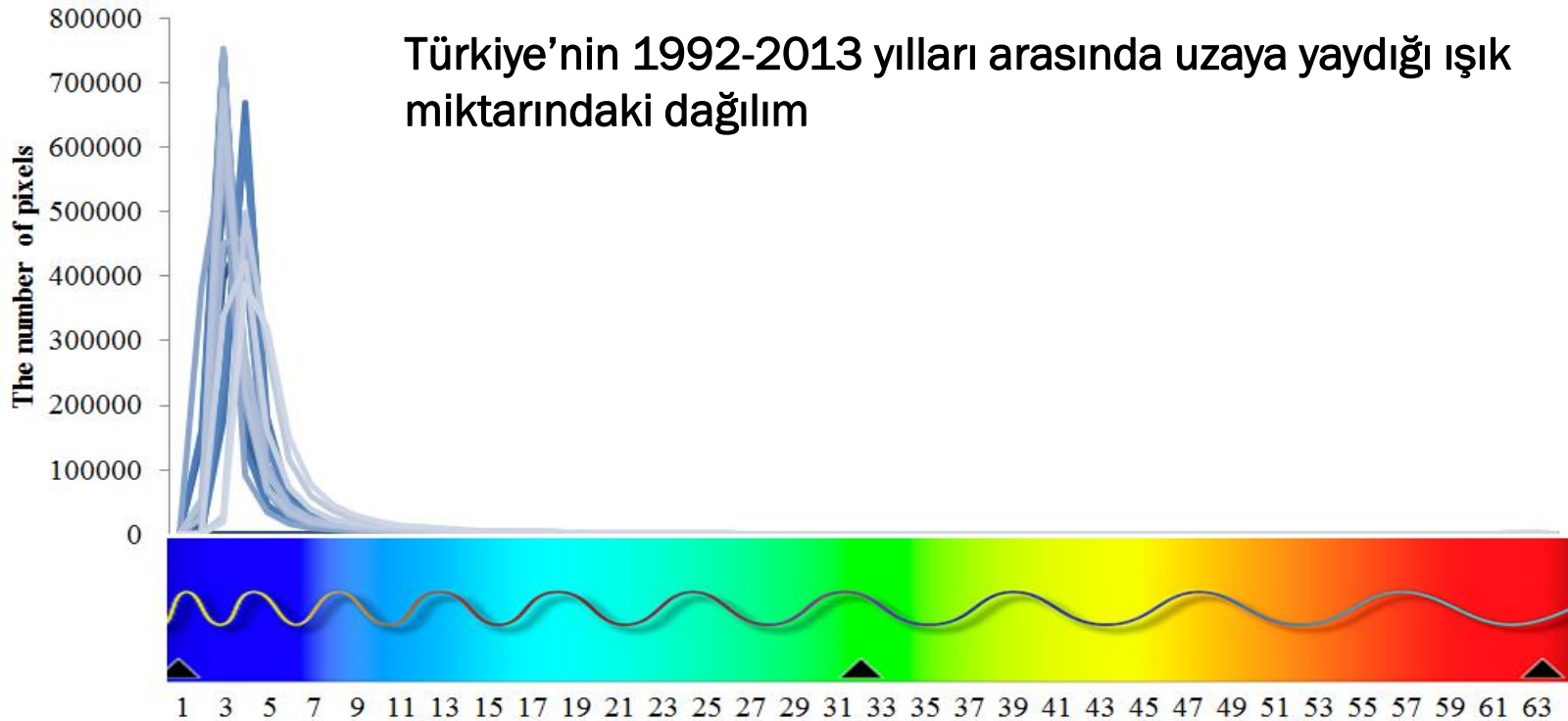


$$X_t = \sum_{i=0}^{63} i * (\# \text{ piksel sayısı});$$

$DN=i$ ve X_t , t zamanda ağırlıklandırılmış gece ışıkları miktarı

Numerik Verinin Elde Edilmesi: Türkiye

- 0-63 aralığındaki ışık dağılımı, Türkiye'nin ışık yoğunluğunun yaklaşık 4. değerde olduğunu gösteriyor.
 - Gelişmiş ülkelerde bu değer 9 ile 15 aralığında değişmektedir.



GYSH'nin Hesaplanması: Yöntem

- Gece ışıkları serileri, periyodik olmayan kayıtlar, meteorolojik koşullar ve bazı yıllarda uydu teknolojisindeki değişim nedeniyle bozulmaya uğrayabilir.
 - Bu durum, verinin kalitesini düşürmekte ve ekonomik faaliyetlerin ölçümünü engellemektedir
- 1. Bozulmaları gidermek için sinyal işlemeyle (FIR filtresi) gece ışıkları serisini filtreledik. Böylelikle, gece ışıklarının frekansı $[0,1]$ aralığında senkronize edilmiş oldu.
- 2. GSYH'yi tahmin etmek için doğrusal olmayan, otoregresif bir model oluşturduk. Kalibrasyon için, insan beyninin nöronları arasındaki ilişkiye dayanılarak geliştirilen Sinir Ağı algoritmasını kullandık.

GYSH'nin Hesaplanması: Yöntem

■ Neden Sinir Ağı Modeli Kullanıyoruz?

1. Regresyona sokulabilecek diğer ekonomik değişkenlerin il bazında verisinin olmaması
2. Kalibre edeceğimiz il bazında GSYH verisinin sadece sınırlı zaman aralığında mevcut olması, 1992-2001
3. Gece ışıkları serisinin *non-smooth* olması ve ekonomik değişkenlerle doğrusal olmayan ilişkisi

Basit Model

- Model, sistemi katmanlar olarak görür. Bağımsız, gizli değişkenlerden ve ışıklar verisinden, $X_0, X_1, X_2 \dots, X_n$, oluşan sinirsel birimler (nöron) girdi katmanını oluşturmaktadır
 - $t = 1$ ise doğrusal model

$$GDP = \sum_{j=1}^j X_j \alpha_j \quad \alpha > \Phi$$

2

- α_j , bilgi aktarımında nöronun aktivasyonunu belirleyen ağ parametresidir.

* Aslında modelimiz, otoregresif olduğu için $GDP_t = \sum_{j=1}^j X_{t,j} \alpha_{t,j} + GDP_{(t-1)} \eta_t$ olacak şekilde nöron aktivasyonu $\eta > \hat{\eta}$ kritik değerini sağlamalıdır. Basit olması açısından $t=1$ zamanda gösteriyoruz.

Basit Model

- Bilgi aktarımı için nöron aktivasyonu, bir eşik fonksiyonu üzerinden gerçekleşmektedir.

→ Bu fonksiyon logit (sigmoid) olacaktır.

$$f(\mu) = \frac{1}{1 + e^{-\mu}}$$

3

- Dolayısıyla doğrusal yapının gizli katman yoluyla doğrusal olmayan bir ilişkiye girmesine izin veriyoruz.

→ Böylece, giriş nöronlarından alınan bilgiler önce gizli bir katmanda işlenir ve daha sonra çıktı katmanına iletilir

$$H = f \left[\sum_{j=0}^j X_j \alpha_j \right]$$

4

→ θ_j gizli nöronları girdi nöronlarına bağlayacak bir ağırlık/parametre olacak şekilde

$$GDP = \sum_{j=0}^j H_j \theta_j \quad \longrightarrow \quad GDP = h \left[\left(\sum_{k=1}^k \alpha_k \right) f \left(\sum_{j=0}^j \theta_{ik} X_j \right) \right].$$

5

Basit Model

Kalibrasyon

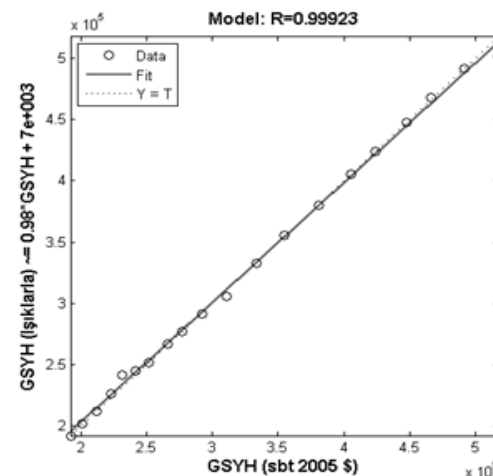
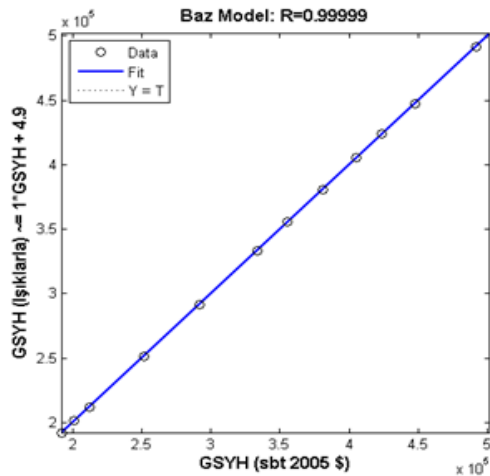
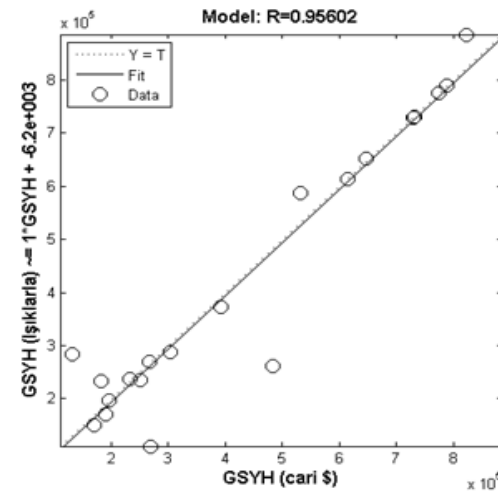
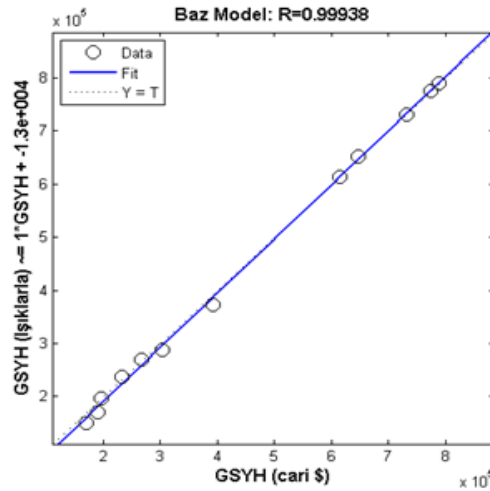
$$\min_{\alpha_k \theta_{jk}} SSD = \sum_{t=1}^T \left[GDP_t - h \left(\sum_{k=1}^k \alpha_k f \left(\sum_{j=0}^j \theta_{ik} X_{jt} \right) \right) \right]^2$$

6

- Kritik değerlerin, α_k ve θ_{jk} , geçilmesi durumunda yakınsama gerçekleşmektedir.
- Sınır ağında kalibrasyon için tipik olarak üç farklı veri setine gerek durulmaktadır;
 1. Parametrelerin başlangıç tahmini için *training* örnekleme gereklidir.
 - Modelimizde, ağın eğitim için belirlenen verilerin yüzde 70'ini kullanmasına izin veriyoruz (temel model).
 2. Eğitilmiş modelden elde edilen tahminin doğruluğunu doğrulamak için *testing* örnekleme...
 - Gizli nöronların sayısı, test örneği modelinin performansına (saturated model) göre belirlenir.
 3. Modelin tahminde kullanması için, tahmin örnekleme (target) gereklidir.

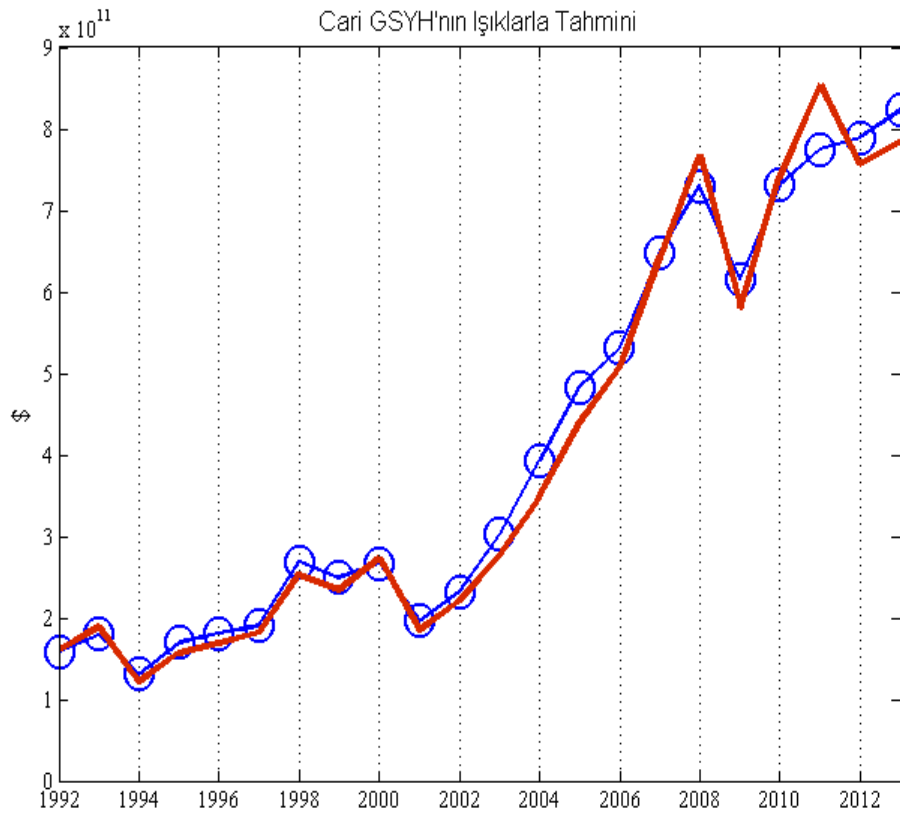
Sonuçlar

Temel model ile tahmin edilen model arasındaki korelasyon, Türkiye'nin reel ve nominal GSYİH

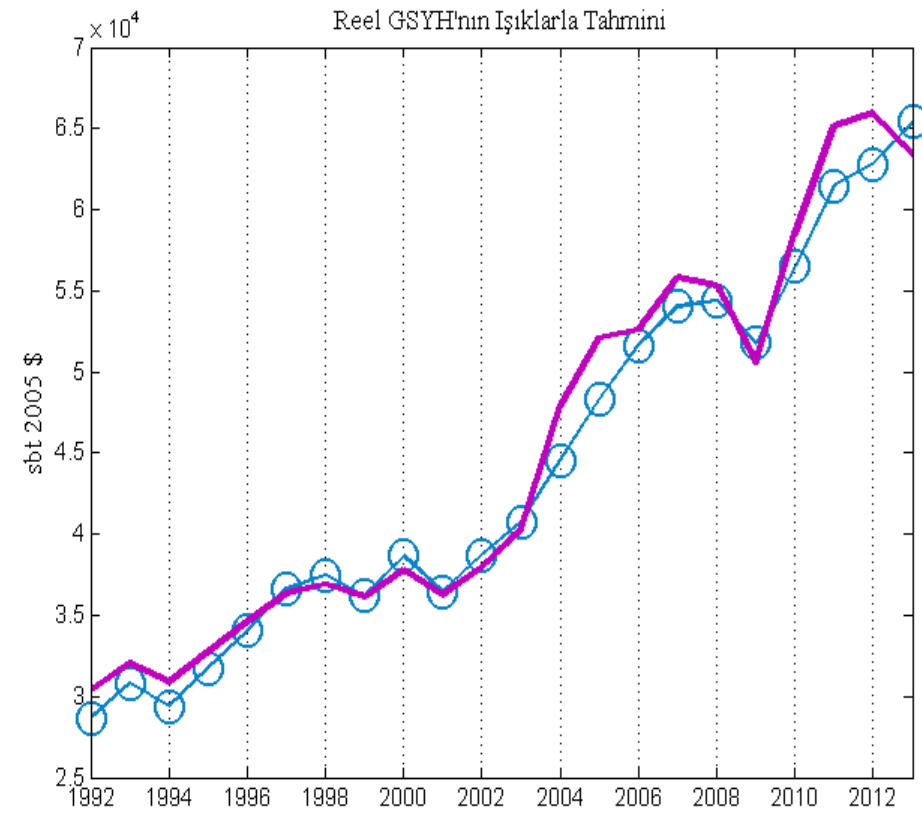


Sonuçlar

Gece Işıklarıyla Tahmin Edilen cari (nominal) ve reel GSYH, 1992-2013



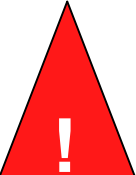
—○ Açıklanan GSYH (cari \$) — Işıklarla Tahmin Edilen GSYH



—○ Açıklanan GSYH (2005 \$) — Işıklarla Tahmin Edilen GSYH

İl Bazında GSYH Hesaplanması

- Modelin yapısal geçerliliğini kontrol etmek için Türkiye'nin reel ve nominal GSYH'sini tahmin ettikten sonra, aynı yöntemle il bazında Gayrisafi Yurtiçi Hasıla(GPP) tahmin ediyoruz
 - GSYH hesaplamasında kullandığımız kritik değerleri, α_k ve θ_{jk} , kullanıyoruz.
- 1. Modele ikinci bir kontrol daha!
 - 2001 yılı il bazında GSYH'yi baz sene olarak alıyoruz.
 - 2001 baz verisini 1992-2013 yılları arasındaki 100 normalize edilmiş ışıkların oransal büyümesi ile çarpıyoruz
 - Reel GSYH'nin tahmininden elde edilen doğrusal olmayan ağ parametreleri (başlangıç parametreleri) ile serileri yeniden tahmin ediyoruz



Her il ekonomik ve coğrafi koşullar bakımından aynı özelliklere sahipmiş gibi davranamayız. Bu yaklaşım doğrultusunda, gizli nöronların sayısını değiştirmeden ekstra gizli bir katman koyduk.

Temel Bulgular

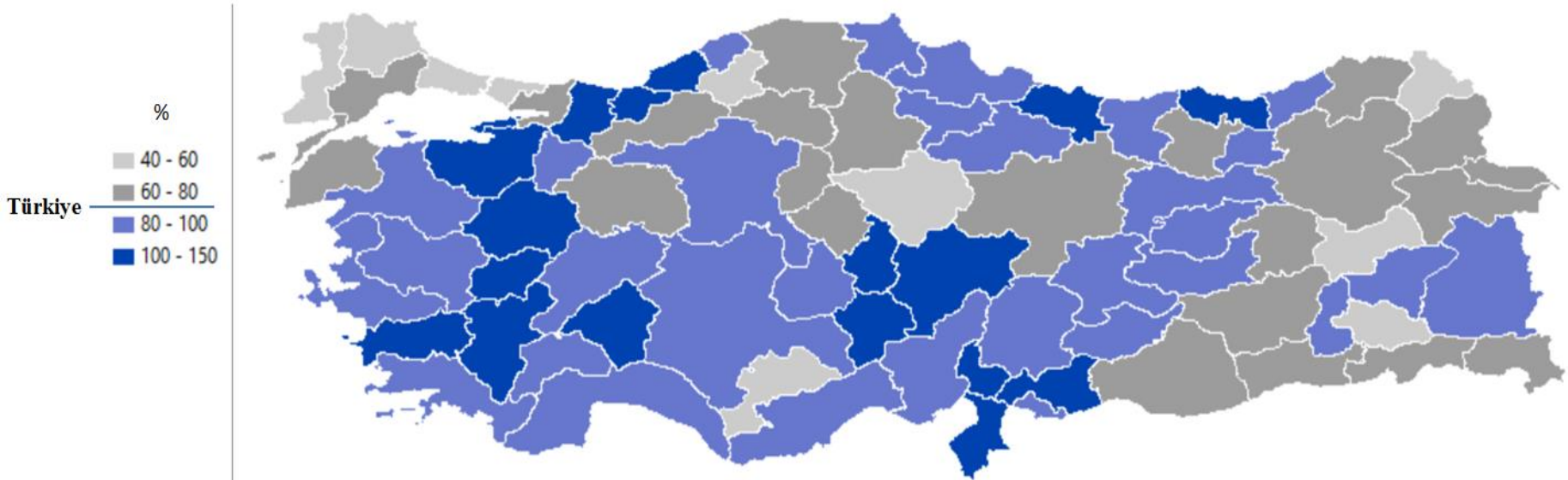
- 2013 yılında en fazla hasılaya 117 milyar dolar (sabit 2005) ile İstanbul ili sahiptir. İstanbul'u yaklaşık 52 milyar dolar (sabit 2005) ile İzmir ve Ankara takip etmektedir.
 - İstanbul ili 2013 yılında tek başına Türkiye'nin GSYH'nin 6'da 1'ni oluşturmaktadır. İlk üç sıradaki İstanbul, Ankara ve İzmir ise aynı yıl için Türkiye GSYH'nin yaklaşık 10'da 4'nü oluşturmaktadır.
- En düşük GSYH değerine sahip olan ilimiz Ardahan'dır (300 milyon dolar). Ardahan'ı yaklaşık benzer GSYH düzeyinde sırasıyla Bayburt, Tunceli, Hakkâri ve Iğdır izlemektedir.

* TÜİK'in yayınladığı verilerdeki ulusal GSYH ile bütün illerin toplamından elde edilen GSYH arasındaki %34,64'lük hata payı vardır. Bu hata, bizim hesaplamamız ile giderilmiştir.

Temel Bulgular

- Genel olarak illerin 2001 ile 2013 yılları arasındaki büyüme trendi Türkiye'nin büyüme trendiyle oldukça benzerdir. İllerin çoğu 2008 bunalımından sonra hızlı büyüme performansı (*recovery*) sergilemiştir.
 - Diğer yandan 2008 küresel krizinden en fazla etkilenen illerimizden biri, İstanbul'dur. İstanbul'un GSYH'si 2008'den 2009'a %6,8 daralmıştır. 2009'daki dipten sonra 2010 yılındaki en yavaş toparlanma yine İstanbul'da gerçekleşmiştir (%5,7 büyüme). 2008 krizinden az etkilenen ilimiz ise Hatay'dır. Hatay ekonomisi 2009'den 2008'e sadece %0,4 daralmıştır

İllerin 2001-2013 Yılları Arası Dönemlik Büyümesi (%)



İleri Bulgular: İl Bazında Kişi Başına Düşen GSYH

■ Türkiye kapsayıcı bir büyüme sergilerken, illerin büyüme oranlarında bir asimetri gözlemlenmektedir.

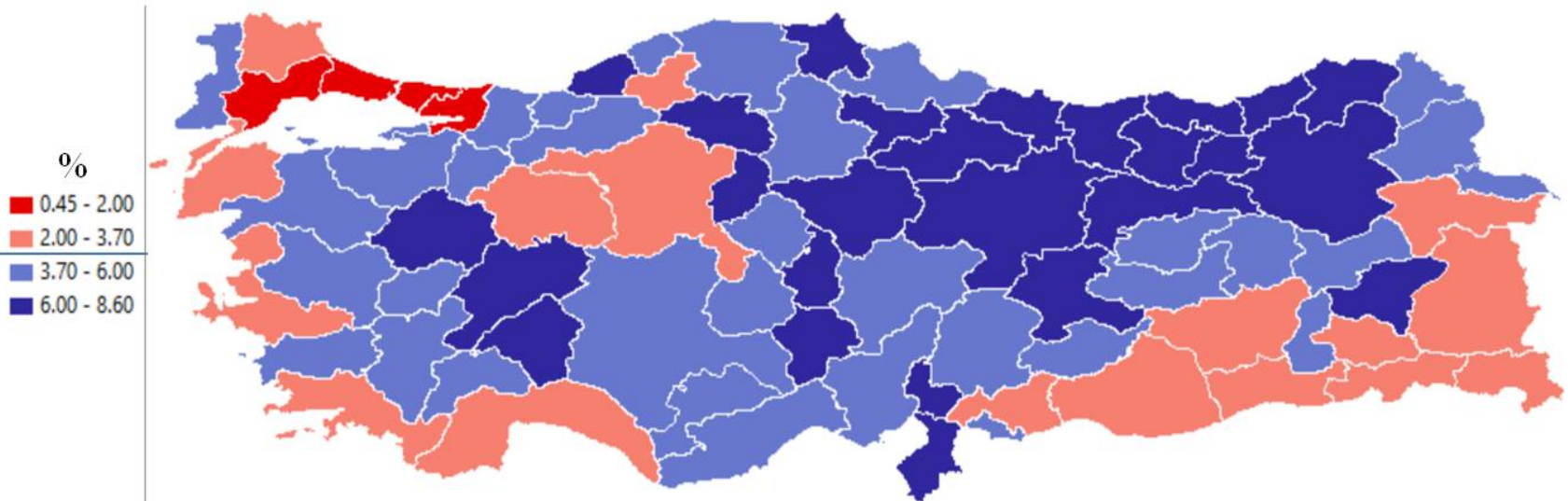
- 2001 yılında, 13.851 dolar ile Kocaeli en yüksek kişi başı gelire sahipken, 2013 yılında 19.478 dolar ile Yalova, Kocaeli'nin yerini almıştır. 2001 yılında 1.100 dolar ile en düşük kişi başı gelire Ağrı sahipken, 2013 yılında Şırnak 1.587 dolar ile en düşük kişi başına gelire sahip il olmuştur .

En Yüksek (2001)		En Düşük (2001)		En Yüksek (2013)		En Düşük (2013)	
İL	Kişi başı GSYH (sbt 2005 \$)	İL	Kişi başı GSYH (sbt 2005 \$)	İL	Kişi başı GSYH (sbt 2005 \$)	İL	Kişi başı GSYH (sbt 2005 \$)
Kocaeli	13851	Kars	1947	Yalova	19478	Kars	3459
Yalova	11087	Iğdır	1745	Kocaeli	16837	Ardahan	2917
Bilecik	9140	Van	1702	Bilecik	15654	Iğdır	2750
Karabük	9067	Bingöl	1585	Bolu	13831	Bitlis	2716
Kırklareli	8773	Ardahan	1577	Manisa	13673	Van	2574
Bolu	8556	Bitlis	1231	Karabük	13553	Bingöl	2468
İzmir	8249	Şırnak	1231	Denizli	12956	Muş	2029
Tekirdağ	7917	Muş	1200	İzmir	12696	Ağrı	1699
İstanbul	7817	Hakkari	1115	Kırklareli	12669	Hakkari	1640
Manisa	7742	Ağrı	1100	Zonguldak	12382	Şırnak	1587

İleri Bulgular: İl Bazında Kişi Başına Düşen GSYH

- Türkiye'nin en fazla GSYH değerlerine sahip olan üç büyük ili; İstanbul, İzmir ve Ankara'da kişi başı gelir artışı yavaşlamıştır.
- 2001'den 2013'e illerin büyük bir kısmının kişi başı geliri, Türkiye ortalamasının üstünde büyümüştür.

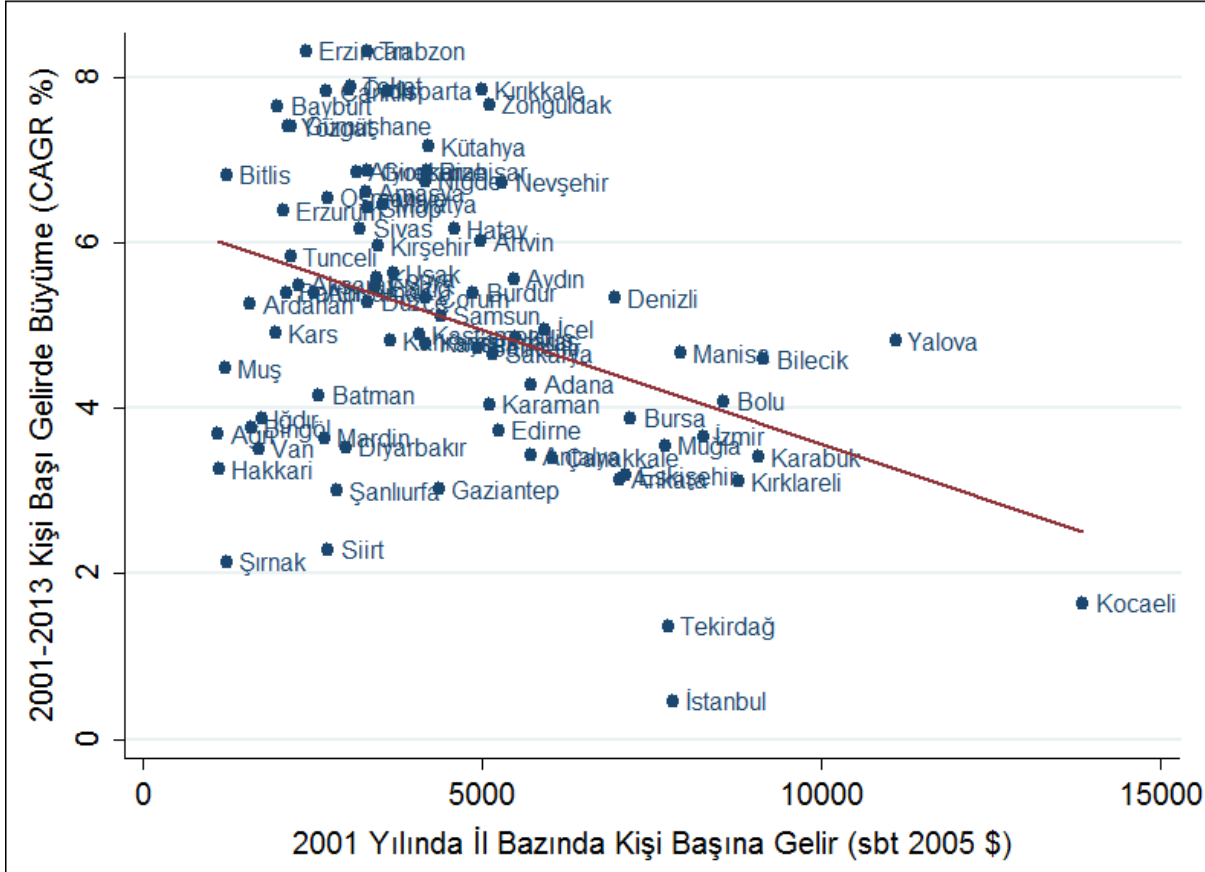
İllerin kişi başı gelirindeki ortalama yıllık büyüme hızları, 2001-2013, CAGR



İleri Bulgular: İl Bazında Kişi Başına Düşen GSYH

- Kişi başı gelir incelendiğinde, (i) tüm illerde artış ve (ii) iller arası gelir eşitsizliğinde bir azalma görülmektedir.

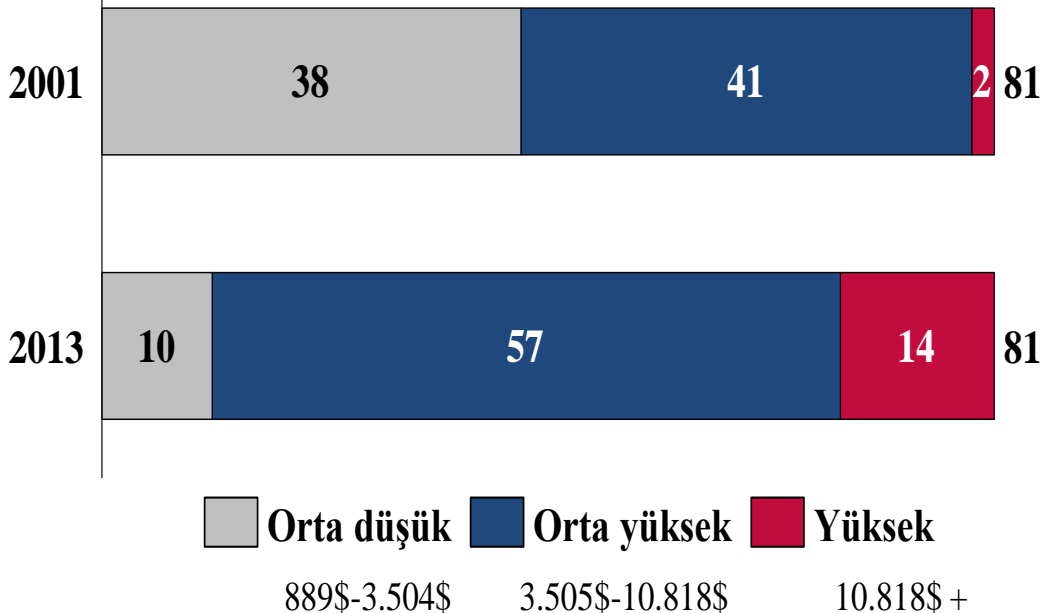
Kişi başına gelirden iller arasında yakınsama



İleri Bulgular: İl Bazında Kişi Başına Düşen GSYH

- Türkiye’de kişi başı gelir bakımından orta sınıf illerin sayısının arttığı ve illerin refah seviyesinde bir dönüşüm yaşandığı gözlemlenmiştir.
 - 28 il orta düşük gelir grubundan orta yüksek gelir grubuna geçmiştir

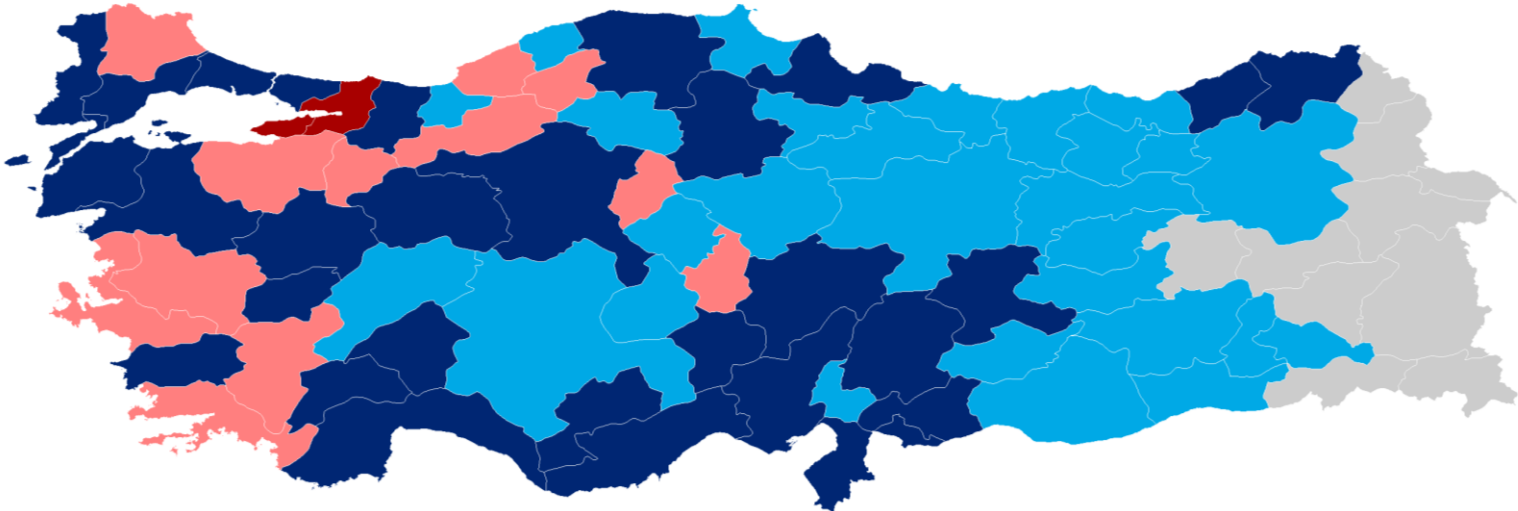
Gelir gruplarına göre illerin dağılımı, 2001 ve 2013



İleri Bulgular: İl Bazında Kişi Başına Düşen GSYH

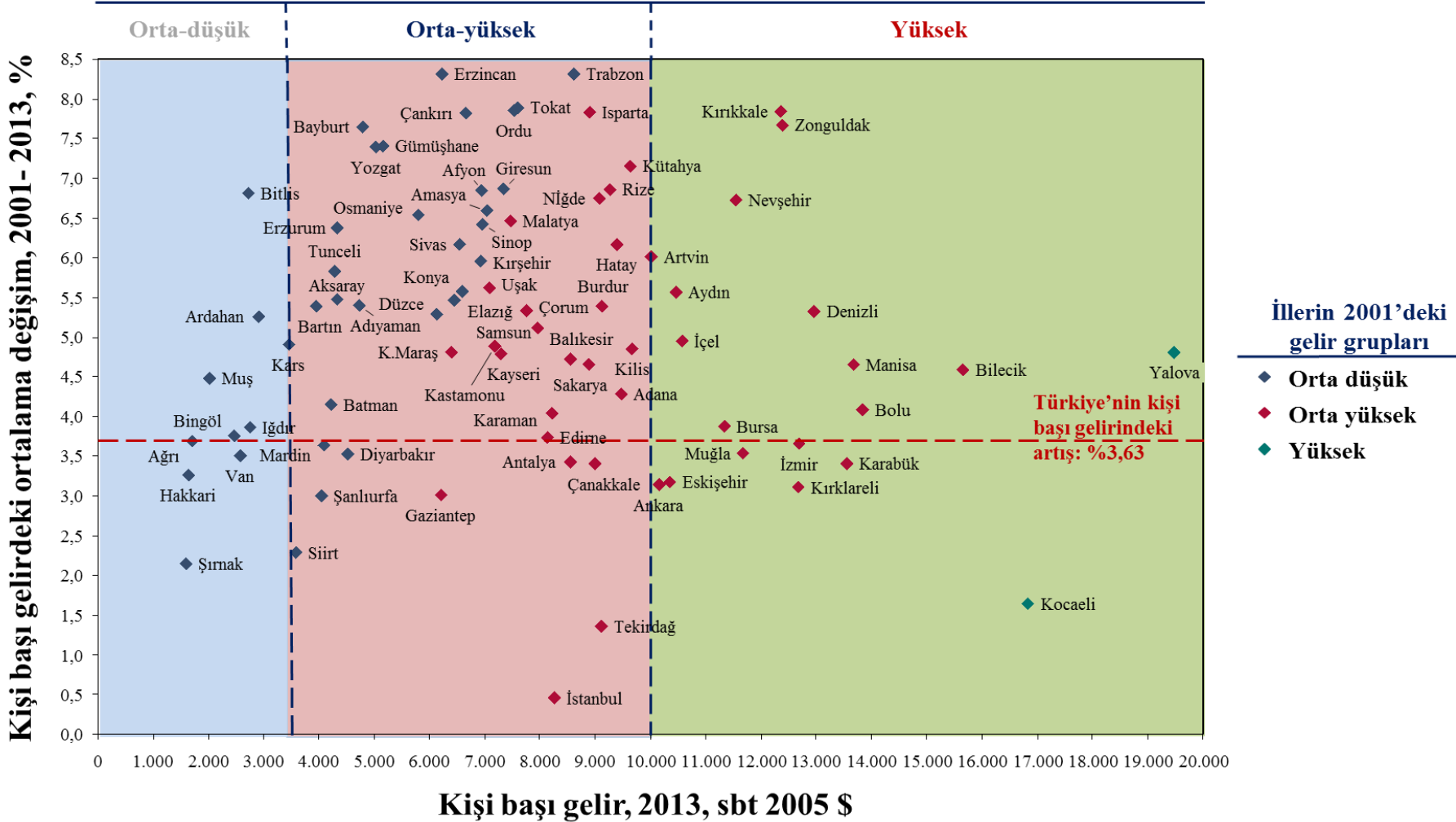
- Gelir sınıfı arasındaki geçiş hem sosyal uyum hem sürdürülebilir büyüme açısından bireyler adına umut oluşturmaktadır. Ancak, bu geçirgenliğin tüm illeri kapsamadığı gözükmemektedir

2001 yılına göre değişimler ile illerin 2013'teki gelir grupları



İleri Bulgular: İl Bazında Kişi Başına Düşen GSYH

İllerin 2013'teki gelir grupları



EK: Yöntem

■ *Machine Learning* temel sorunlardan biri *overfitting*'dir.

- Değişkenler arasında genelleştirilmiş bir temel ilişki aranırken *overfitting* büyük bir sorun oluşturmaktadır. Bununla birlikte, gece ışığı ile GSYİH arasındaki genel ilişkiyi açığa çıkarmamak yerine, il düzeyinde GSYH'yi tahmin etmek için bu değişkeni bir araç olarak (refree variable) kullanıyoruz.
- Sistemin tolere edebileceği kompleksite (tanıtılan değişken ve parametre sayısı) ile değişkenlerin *nonlinear* hareketleri arasında bir *trade-off* mevcuttur. Bu yüzden
↓
tanımlanan gizli değişken sayısı önemlidir.

■ Çok fazla parametre tanıtıldığında, sinir ağı modeli patikayı öğrenmek yerine ezberleyecektir. Bu da tahmin gücünü düşürür.

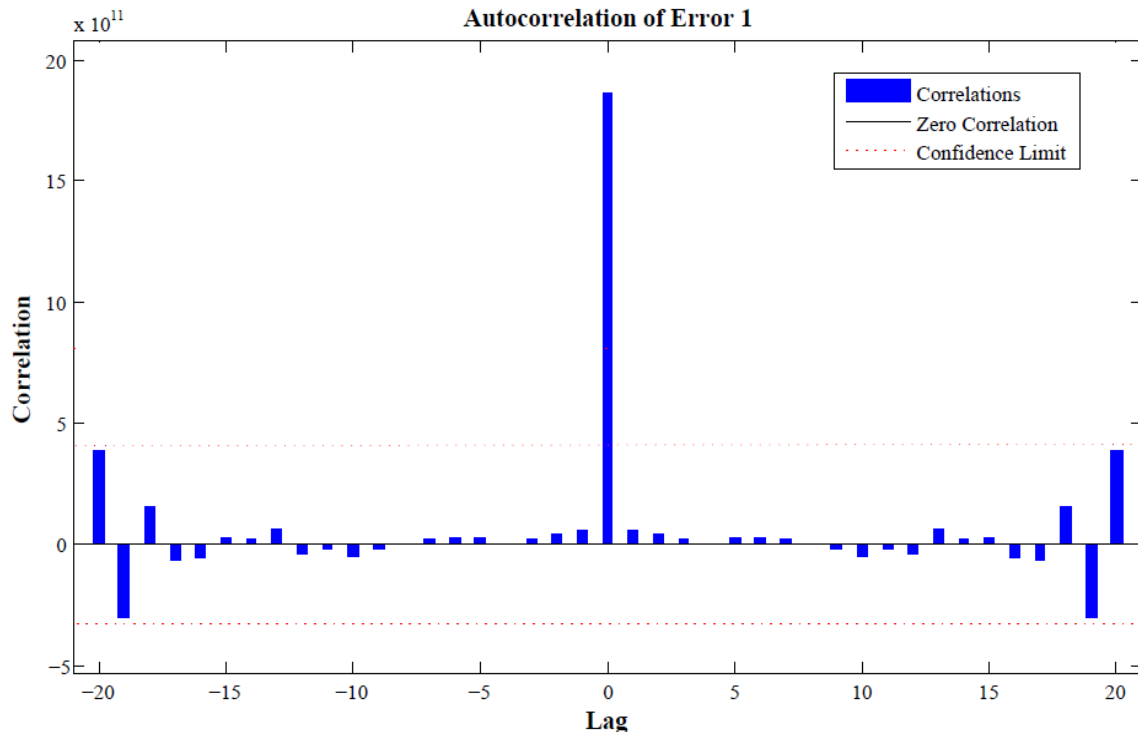
- **Çözüm:** Sistemin öğrenmeye olanak tanıyacak şekilde training örneklemin geniş tutulması!

EK: Yöntem

■ Temel varsayım:

GSYH ile ışıklar verisinin hata terimleri arasında korelasyon yoktur.

$$\text{Cov}(\zeta_{mt}, \zeta_{nt}) = 0 \text{ for } m \neq n \in \{1,2,3\}.$$



Test:

Modelimiz için, hataların terimlerimin tamamen birbiriyle ilişkisiz olduğunu söyleyemeyiz. Bununla birlikte güven sınırı içindeki hatalar arasında anlamlı bir korelasyon olmadığını göstermektedir.

Teşekkürler...

Katkılarından dolayı Prof. Dr. Güven Sak'a, Prof. Dr. Erol Taymaz'a, TEPAV Şehir Çalışmaları Proje Direktörü Emre Koyuncu'ya, Jeomorfolog İlker Alan'a (Meteoroloji Genel Müdürlüğü), TEPAV Politika Analistleri Ayşegül Düşündere, Can İtez ve Yasemin Satır Çilingir'e teşekkür ederim.