

## SMRGT yöntemi ile bulanıklaştırılmış veriler için bulanık doğrusal regresyon

\* Seçil YALAZ<sup>1</sup>, Arife ATAY<sup>1</sup>, Z. Fuat TOPRAK<sup>2</sup>

<sup>1</sup> Dicle Üniversitesi, Fen Fakültesi Matematik Bölümü,

<sup>2</sup> Dicle Üniversitesi, Mühendislik Fakültesi İnşaat Mühendisliği Bölümü

### ÖZET

Klasik regresyon analizinde oluşturulan modellerin tahminleri ile gözlenen bağımlı değişken arasında genellikle gözlem hatasından kaynaklanan sabit varyanslı ve 0 (sıfır) ortalamalı normal dağılıma sahip olduğu varsayılan bir fark oluşur. Bulanık regresyon analizinde bu fark model yapısının bulanıklığı olarak düşünülmektedir. Belirli kısıtlamalar nedeniyle klasik doğrusal regresyon uygulanması yerine bulanık doğrusal regresyon uygulanması daha uygun olan veriler için, veri bulanık değişkenler içermiyorsa olasılıksal modelin kullanılması zorunlu hale gelebilmektedir. Oysaki olasılıksal modelin de belli kısıtlamalar nedeniyle kullanılmasının uygun olmadığı durumlar vardır. Burada değişkenleri bulanık olan modellerin kullanılması için, bulanık olmayan değişkenler, yeni bir yöntem olan SMRGT ile bulanıklaştırılmış, elde edilen değişkenler için bulanık En Küçük Kareler (EKK) modeli ile bulanık doğrusal regresyon denklemi oluşturulmuştur. Sonuç olarak klasik doğrusal regresyon ve bulanık doğrusal regresyon bir veri setine uygulanmış ve bu iki yaklaşımın performansları çeşitli ölçüm kriterleri kullanılarak karşılaştırılmıştır.

### Anahtar Kelimeler:

Bulanık Regresyon, Olasılıksal Model, Bulanık EKK Modeli, Diamond Modeli, SMRGT Metodu

## Fuzzy linear regression for the data which is fuzzified with SMGRT method

### ABSTRACT

In classical regression analysis a difference is occurred between the observed dependent variable and estimates of the generated models. This difference is generally caused by the assumption of that the observed dependent variable has to be normally distributed with a constant variance and zero (0) average. In fuzzy regression analysis this difference is considered to be the blurring of the model structure. for the data which is appropriate for the fuzzy regression analysis instead of classical regression analysis due to certain restrictions, if the data does not include fuzzy variables, using probabilistic models may become mandatory. However, because of certain restrictions there are cases using of probabilistic model is not appropriate. We generate a fuzzy linear regression equation with Fuzzy Least Square Regression (FLSR) model. Here nonfuzzy variables are fuzzified with a new method SMRGT to use models have fuzzy variables. In conclusion, classical linear regression and fuzzy linear regression were applied a data set, and these two approaches performances were compared by using different measures.

### Key Words:

Fuzzy Regression, Probabilistic Model, FLSR Model, Diamond Model, SMRGT Method

## 1. Giriş

İnsan tahmininin etkili olduğu bazı sistemlerin modellenmesinde bulanık bir yapıyla karşılaşılabılır. Bu bulanık yapı parametreleri bulanık kümelerden elde edilen bir bulanık doğrusal fonksiyon olarak sunulabilir. Bulanık doğrusal fonksiyonlar Zadeh'in genişleme kuramı [1] ile tanımlanır. Son 30 yıla kadar sürekli kullanılan belirgin mantık ve özellikle de yine 2000 yılına kadar insanların basitçe kullandıkları Aristo mantığı yerine, belirsizlik içeren ve bulanıklık (fuzzy) denilen bir mantık yapısı, değişik teknolojik cihazlarda kullanılır hale gelmiştir. Burada, günlük konuşma dilinde geçen sözel belirsizlikleri de, modelleme veya hesap yapılırken işlemlere katılması mümkün olmuştur. Belirsizliğin nesnel olarak incelenmesi için bazı teknik, yöntem algoritma ve yaklaşımlar ile alışlagelmiş belirgin matematik dışında, belirsizliği yakalayabilecek ve onu sayılaştrabilecek kuralları olan bazı bilimsel yöntemlere ihtiyaç duyulmaktadır. Bu türlü konular için geçerli olabilecek olasılık, istatistik, stokastik, fraktal, kaotik, kuantum gibi değişik belirsizlik yöntemlerinin en eskisi ve diğerlerinin anlaşılabilmesi için gerekli olanı olasılıktır [2]. İstatistik biliminin en önemli konularından birisini regresyon analizi oluşturmaktadır. Regresyon analizi genellikle değişkenler arasındaki ilişkilerin incelenmesinde [3] ve bir fonksiyonun uygun değerleri için katsayılarının belirlenmesinde kullanılmaktadır [4]. Sistemin bulanık yapısının modellenmesi için bulanık bir doğrusal fonksiyon düşünüldüğünde bir bulanık doğrusal regresyon analizi formüle edilir.

Klasik regresyon analizinde aşağıdaki durumlarda karşılaşılan problemler bulanık regresyon analiziyle çözülebilmektedir:

- Gözlenen veri sayıları yetersiz ise,
- Analizin uygulanması için gereken varsayımların doğruluğunun testleri yapılırken güçlüklerle karşılaşılıyor ise,
- Bağımlı ve bağımsız değişkenler arasındaki bağlantı belirsiz ise,
- Gözlenen olayların frekans sıklığı belirsiz ise,
- Veriler normal dağılımlı veya değişkenler arasındaki ilişki doğrusal değil ise [5].

Regresyon analizi uygulamalarında bağımsız değişken veya değişkenlerin bağımlı değişkeni ne oranda etkilediğine katsayılar yardımıyla bakılmaktadır. Doğrusal regresyon modeli  $y = \alpha^t x$  olmak üzere; regresyon analizi sonucunda gözlenen değerlerle hesaplanan değerler arasında belirli bir sapma meydana gelmektedir;  $y_i - y_i^* = \epsilon_i, i = 1, \dots, N$ . Sapmaları minimuma indirmek için bulanık regresyon yöntemleri geliştirilmiştir.

Bulanık regresyon modeli inşasında başlıca iki öneri vardır; Tanaka vd. (1982) tarafından önerilen bulanık doğrusal regresyon ve onun varyasyonu, Diamond (1988) [6] ve Celmins (1987 a, b) [7-8] tarafından geliştirilen bulanık en küçük kareler regresyon önerisidir.

Bulanık modellenen ilk doğrusal regresyon analizi 1982'de Tanaka, Uejima ve Asai tarafından ortaya konulmuştur [9]. Bu metoda göre; gözlenen değerlerle hesaplanan değerler arasındaki

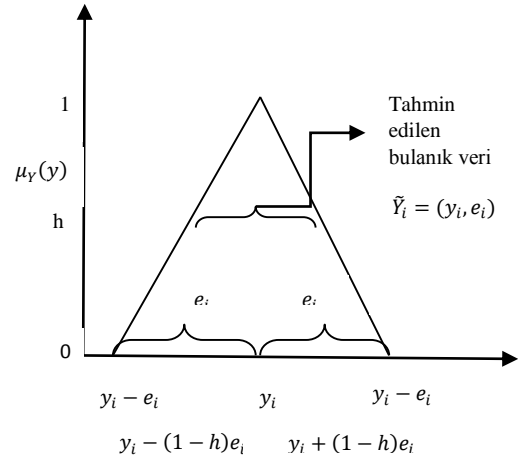
sapmalar, istatistiksel regresyondaki gibi ölçüm ve gözlem hatalarından değil, sistem parametrelerinin bulanıklığından veya regresyon katsayılarının bulanıklığından kaynaklanır. Regresyon katsayıları bulanık sayılar olduğu için, tahmin edilen bağımlı değişken de bulanık bir sayıdır [10]. Bu yöntemin en önemli avantajı her bir gözlemin üyelik derecesine göre katsayı tahminine katılmasıdır. Bu yaklaşımda tahmin edilen bulanık sayısı, simetrik üçgensel üyelik fonksiyonuna sahip  $\tilde{Y}_i = (y_i, e_i)$  şeklinde gösterilmektedir.

Bir bulanık regresyon modeli; değişken ilişkilerinin bulanık olduğu modeller ve değişkenleri bulanık olan modeller şeklinde iki farklı şekilde oluşturulabilir. Oluşturulan modelin uyumu için, bulanık katsayıların toplam yayılımını minimize ederek modelin bulanıklığını minimize eden olasılıksal model ve merkezlerine ve yayılımlarına bağlı olarak modelin çıktısı ve gözlenen çıktı arasındaki uzaklığı minimize eden EKK modeli kullanılabilir.

## 2. Materyal ve metot

### 2.1 Olasılıksal Model (Tanaka Modeli)

Bir çeşit olasılıksal model olan Tanaka yaklaşımında bulanık elemanların üçgensel bulanık sayılar olduğu varsayılır.



Şekil 1.  $h$ -düzye kümesine sahip bir bulanık sayının simetrik üçgensel bulanık çıktı fonksiyonu

Burada;  $y_i$  ; bulanık merkez,  $e_i$  ; bulanık genişliktir.  $h$  faktörü alt ve üst regresyon doğrularının arasındaki doğrudur. Simetrik üçgensel bulanık sayılar varsa  $h = 1$  olur ve tam orta noktadadır. Gözlemlenen değerlerin yaptığımız tahminlerin içinde kalabilmesi için  $h$  değerinin 1'e çok yakın olmaması gerekir.  $h$  değeri 0 seviyesine yaklaştıkça bu da güven seviyesinin azalmasına ve üyelik değerinin düşmesine neden olur.  $h$  faktörü,  $\alpha$  kesim seviyesinin tam tersidir.  $h$  faktörü destek bölgesini büyütürken  $\alpha$  kesim seviyesi küçültür. Güven seviyesi ( $h$ ) en uygun modelin elde edilmesini sağlayacak şekilde seçilmelidir [11]. Bu bilgilerle beraber  $h$ 'ın seçiminde genellikle tavsiye edilen değer 0,5'dir. Bu değer duruma göre değişiklik gösterebilir. Farklı noktalarda analizler yapıp karar vermek gerekir [12]. " $h$ " ne kadar farklı değer alırsa alsın bulanık ortalama değerler değişmez [13]. Bulanık regresyonda  $h$  değerini belirlemek analiste bırakılmıştır.  $h$ 'ın seçimi konusunda belirli bir ölçüt yoktur. [14].

Simetrik üçgensel biçime sahip olan  $y_i$ 'nin üyelik fonksiyonu aşağıdaki şekilde tanımlanabilir. Burada  $\mu$  üyelik fonksiyonudur. Bir bulanık doğrusal regresyon modelinin formülize edilmesi için aşağıdaki adımlar izlenmektedir:

1. Veri, bir bulanık doğrusal model ile temsil edilir.

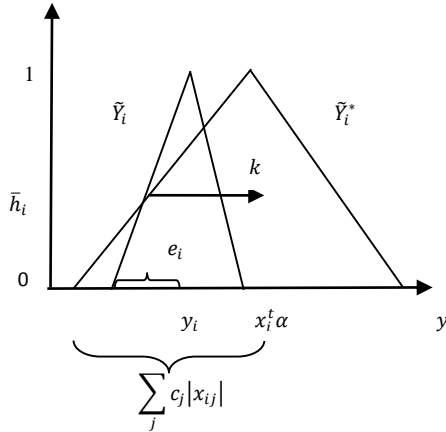
$$Y_i^* = A_1^* x_{i1} + \dots + A_n^* x_{in} = A^* x_i$$

Buradaki  $A_i$  bulanık parametrelerinin yazılışı denklem (1)'de verilmiştir.  $x_i, Y_i^*$ 'ler aşağıdaki denklemden elde edilir.  $A = \{\alpha, c\}$  parametresinde  $\alpha$ ; üçgenin merkezi,  $c$ ; yayılım miktarı olmak üzere;

$$\mu_Y(y) = \begin{cases} 1 - \frac{|y - x^t \alpha|}{c^t |x|}, & x \neq 0 \\ 1, & x = 0, y = 0 \\ 0, & x = 0, y \neq 0 \end{cases} \quad (2)$$

Burada;  $|x| = (|x_1|, \dots, |x_n|)^t$  ve  $c^t |x| \leq |y - x^t \alpha|$  iken  $\mu_Y(y) = 0$ 'dır.

2.  $Y_i = (y_i, e_i)$  şeklindeki verinin  $Y_i^* = A^* x_i$  ile belirtilen tahmin edilen bulanık doğrusal modelinin uyum derecesi,  $Y_i^h \subset Y_i^{*h}$  iken  $h$ 'ı en büyükleen  $\bar{h}_i$  tarafından ölçülür.  $Y_i^h = \{y | \mu_{Y_i}(y) \geq h\}$  ve  $Y_i^{*h} = \{y | \mu_{Y_i^*}(y) \geq h\}$ ,  $h$ -düzey setleridir. Tüm  $Y_1, \dots, Y_N$  verisine uyan bulanık doğrusal modelinin uyum derecesi  $\min_j [\bar{h}_i]$  ile tanımlanır.



Şekil 2.  $Y_i^*$ 'ın verilen  $Y_i$  bulanık verisine uyum derecesi

3. Bulanık doğrusal modelinin kesinsizliği;  $J = c_1 + \dots + c_n$  şeklinde tanımlanır.

Problem,  $\forall i$  için  $\bar{h}_i \geq H$  iken  $J$ 'yi en küçükleyen  $A_i^*$  bulanık parametrelerini elde etmek olarak açıklanabilir. Burada  $H$ 'ın seçimi, bulanık doğrusal modelinin uyum derecesi gibi analiste bırakılmıştır.  $\bar{h}_i = 1 - \frac{|y_i - x_i^t \alpha|}{\sum_j c_j |x_{ij}| - e_i}$  şeklinde elde edilebilir.

Buradaki denklem  $k = |y_i - x_i^t \alpha| + e_i(1 - \bar{h}_i)$  olması durumunda  $1: (1 - \bar{h}_i) = (\sum_j c_j |x_{ij}|): k$  olmasından gelmektedir.

Daha spesifik olarak problemimizi, aşağıdaki doğrusal programlama probleminin çözümleri olan  $A_i^* = (\alpha_i, c_i)$  bulanık parametrelerini bulmak olarak tanımlayabiliriz:

$$\begin{aligned} \min_{\alpha, c} J &= c_1 + \dots + c_n; \\ c &\geq 0, \\ \alpha^t x_i + (1 - h) \sum_j c_j |x_{ij}| &\geq y_i + (1 - h)e_i, \\ -\alpha^t x_i + (1 - h) \sum_j c_j |x_{ij}| &\geq -y_i + (1 - h)e_i, \end{aligned}$$

$$\mu_Y(y) = \begin{cases} 1 - \frac{|y - y_i|}{e_i}, & y_i - e_i \leq y_i \leq y_i + e_i \\ 0, & \text{aksi taktirde} \end{cases} \quad (1)$$

Genellikle kısıt sayısı ( $2N$ ), değişken sayısından ( $n$ ) çok daha fazla olduğundan (3) probleminin primal problemi yerine dual problemini çözmek daha kolaydır.  $\alpha$  değişkenleri ister istemez negatif olacağından,  $d^t = (d, \dots, d)$  için  $d$  yeterince küçük negatif sayılar olarak varsayılırsa,  $\alpha = \alpha' + d$  olmak üzere,  $\alpha' \geq 0$  değişkenleri her zaman pozitif olacaktır.

$\beta = (\beta_1, \dots, \beta_N)^t$  ve  $D = (D_1, \dots, D_N)^t$  dual değişkenler olarak kullanılırsa, (3)'ün dual problemi aşağıdaki formda yazılabilir:

$$\begin{aligned} \max_{\beta, D} J &= \sum_{i=1}^N \beta_i \left[ y_i + (1 - h)e_i - d \sum_{j=1}^n x_{ij} \right] \\ &\quad - \sum_{i=1}^N D_i \left[ y_i - (1 - h)e_i - d \sum_{j=1}^n x_{ij} \right]; \\ \beta &\geq 0, D \geq 0, \end{aligned}$$

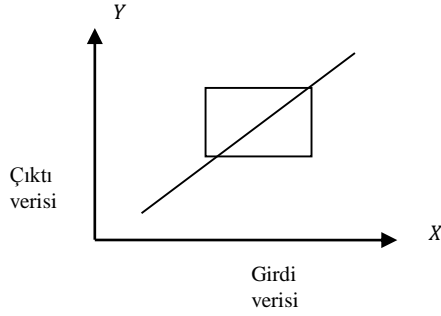
$$\begin{aligned} (1 - h) \sum_{i=1}^N \beta_i |x_{ij}| + (1 - h) \sum_{i=1}^N D_i |x_{ij}| &\leq 1, \\ \sum_{i=1}^N \beta_i x_{ij} - \sum_{i=1}^N D_i x_{ij} &\leq 0, j = 1, \dots, n \end{aligned}$$

#### i. Olasılıksal modelin eksiklikleri

- 2) Tanaka görünüşte lineer regresyona benzer bir model geliştirmek için doğrusal programlama tekniklerini kullanmış ama en küçük kareler konsepti ile ilişkisinin ne olduğunu açıklamamış ve artıklar tarafından sağlanan en iyi uyumun herhangi bir ölçümünü vermemiştir [6].
- 3) Orijinal Tanaka modeli sapan değerlere karşı oldukça hassastır [15].
- 4) Bulanık regresyon aralığı hakkında uygun bir yorum bulunmamaktadır [16].
- 5) Tahminin sonucu önceden belirtilmelidir [17].
- 6) Bulanık doğrusal regresyon, bağımsız değişken sayısı atıkça çoklu bağlantılı olma eğiliminde olabilir [18].
- 7) Eğer ilk olarak, bağımsız değişkenlerin ortalamasını  $x_j$  yerine  $(x_j - \bar{x}_j)$ 'yi kullanarak çıkartırsak, tahmin edilen fonksiyon çok farklı olacağından çözüm,  $x_j$  referans sorununa bağımlı olacaktır [19].

#### a. Bulanık EKK Modeli

Bulanık regresyon modelini istatistiksel regresyon ile aynı çizgiye taşımak için açık bir yol, bulanık regresyonu aynı doğrular boyunca modellemektir. Tek bir açıklayıcı değişken halinde standart doğrusal regresyon modeli olan  $i = 1, \dots, n$  için  $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$  ile başlanır [20]. Buna karşılık olarak bulanık regresyon modeli  $\tilde{Y}_i = \beta_0 + \beta_1 \tilde{X}_i + \tilde{\varepsilon}_i$ ,  $i = 1, \dots, n$  formunu alabilir. Kavramsal olarak  $i$ -inci bulanık yanıt ve açıklayıcı değişkenler aşağıdaki şekilde gösterilebilir:



Şekil 3. Bulanık doğrusal regresyonda değişkenler arası ilişki

$$i = 1, \dots, N \quad (3)$$

Denklem yeniden düzenlenirse;  $\tilde{\varepsilon}_i = \tilde{Y}_i - \beta_0 - \beta_1 \tilde{X}_i$ ,  $i = 1, \dots, n$  bulunur. En küçük kareler bakış açısıyla problem  $\min \sum_{i=1}^n (\tilde{Y}_i - \beta_0 - \beta_1 \tilde{X}_i)^2$  şekline dönüşür.

Bulanık EKK Modeline örnek olarak Diamond Modeli verilebilir.

### 2.2.1 Diamond Modeli

Bulanık en küçük kareler regresyonunu uygulamak için kullanılan yollardan birisi uzaklık ölçümlerini kullanmaktır. Diamond (1988) ilk olarak uzaklık ölçümlerini kullanarak bulanık en küçük kareler regresyonunu uygulayan kişidir ve onun metodu en çok kullanılan metottur [21]. Temelde iki üçgensel bulanık sayı arasında bir  $L^2$  metriği

$$\begin{aligned} d(\langle m_1, l_1, r_1 \rangle, \langle m_2, l_2, r_2 \rangle)^2 \\ = (m_1 - m_2)^2 + ((m_1 - l_1) - (m_2 - l_2))^2 \\ + ((m_1 + r_1) - (m_2 + r_2))^2 \end{aligned}$$

ile tanımlamıştır [8].

Verilen üçgensel bulanık sayıların sol ve sağ yayılımları ayrı ayrı, merkezlerine bağlı olarak iki bulanık sayı arasındaki uzaklığın bir ölçümünü sağlar.

Model  $\tilde{Y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{x}_i + \tilde{\varepsilon}_i$ ,  $i = 1, \dots, n$  formunu alır ve  $S(\beta_0, \beta_1) = \sum_{i=1}^n d^2(\tilde{y}_i, \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{x}_i)$  aşağıdaki şekilde optimize edilir:

$$\sum_{i=1}^n (y_{im} - \beta_{0m} - \beta_{1m} x_{im})^2 + (y_{il} - \beta_{0l} - \beta_{1l} x_{il})^2 + (y_{ir} - \beta_{0r} - \beta_{1r} x_{ir})^2$$

Burada,  $\tilde{\beta}_1$  ve  $\tilde{\beta}_0$  bulanık parametrelerinin tahminlerini bulmak için  $\beta_{0l}, \beta_{0m}, \beta_{0r}, \beta_{1l}, \beta_{1m}, \beta_{1r}$ 'ye göre türevler alınır;

$$\hat{\beta}_1 = (\hat{\beta}_{1l}, \hat{\beta}_{1m}, \hat{\beta}_{1r}), \hat{\beta}_0 = (\hat{\beta}_{0l}, \hat{\beta}_{0m}, \hat{\beta}_{0r}) \text{ iken}$$

$$\bar{x}_l = \frac{\sum_{i=1}^n x_{il}}{n}, \bar{x}_m = \frac{\sum_{i=1}^n x_{im}}{n}, \bar{x}_r = \frac{\sum_{i=1}^n x_{ir}}{n} \text{ ve}$$

$$\bar{y}_l = \frac{\sum_{i=1}^n y_{il}}{n}, \bar{y}_m = \frac{\sum_{i=1}^n y_{im}}{n}, \bar{y}_r = \frac{\sum_{i=1}^n y_{ir}}{n} \text{ olmak üzere;}$$

$$\hat{\beta}_{1l} = \frac{\sum_{i=1}^n x_{il} y_{il} - n \bar{x}_l \bar{y}_l}{\sum_{i=1}^n x_{il}^2 - n \bar{x}_l^2}, \hat{\beta}_{0l} = \bar{y}_l - \hat{\beta}_{1l} \bar{x}_l,$$

$$\hat{\beta}_{1m} = \frac{\sum_{i=1}^n x_{im} y_{im} - n \bar{x}_m \bar{y}_m}{\sum_{i=1}^n x_{im}^2 - n \bar{x}_m^2}, \hat{\beta}_{0m} = \bar{y}_m - \hat{\beta}_{1m} \bar{x}_m$$

$$\hat{\beta}_{1r} = \frac{\sum_{i=1}^n x_{ir} y_{ir} - n \bar{x}_r \bar{y}_r}{\sum_{i=1}^n x_{ir}^2 - n \bar{x}_r^2}, \hat{\beta}_{0r} = \bar{y}_r - \hat{\beta}_{1r} \bar{x}_r$$

olarak tanımlanır [22].

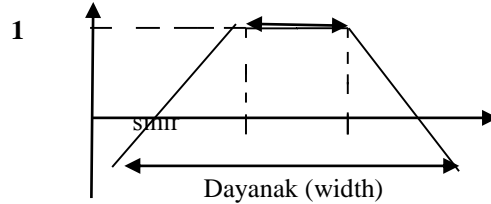
## 2.2 Basit Üyelik Fonksiyonu Ve Bulanıklık Kuralı Üretim Tekniği (SMRGT)

Verilen bir bulanık sistemdeki temel soru, sistemdeki en iyi sonuçları veren üyelik fonksiyonu (MFs) ve bulanıklık kurallarının (FRs) nasıl oluşturulacağıdır. Bulanık bir sistemdeki MFs ve FRs'lerin sayısına, şekline ve mantığına karar vermedeki zorlukların üstesinden gelmek için SMRGT geliştirilmiştir [23]. Bu yeni bulanık teknik sadece girdi ve çıktı değişkenlerinin tüm MFs'leri için üretilen bazı anahtar sayılarına dayanmaktadır. Anahtar sayıları MFs'lerin şekline (üçgen, yamuk, vb.) ve durulaştırma metoduna (cenetroid, MF SMRGT metodu aşağıdaki adımları takip etmektedir.

- Bağımlı ve bağımsız değişkenlere karar verilir. Burada bağımsız değişkenler sistemin girdilerini, bağımsız değişken ise çıktısını belirtmektedir.
- Her bir değişken için (anlamlı bölgede) maksimum ( $V_{max}$ ) ve minimum ( $V_{min}$ ) değerleri tanımlanır.
- MNFs'lerin şekline karar verilir (üçgensel).
- Her bir bağımsız değişken için MFs sayısına minimum 3 olacak şekilde karar verilir (MFs sayısı 5).
- Her bir bağımsız değişken için anahtar değeri ( $K_1, K_5$ ) ile üyelik fonksiyonlarının öz ( $C_{i-1}, C_i, C_{i+1}$ ) ve dayanakları belirlenir (Burada anahtar değer sayısı ile üyelik fonksiyon sayısı eşit olacaktır).

$$\begin{aligned} V_R &= V_{max} - V_{min}, \\ UW &= \frac{V_R}{n_u}, \quad EUW = \frac{3UW}{2}, \\ K_1 &= V_{min} + \frac{EUW}{3}, \quad K_5 = V_{max} - \frac{EUW}{3}, \\ C_i &= K_i = \frac{V_R}{2} + V_{min}, \\ C_{i+1} &= K_{i+1} = V_{max} - \left( \frac{V_{max} - K_i}{2} \right), \\ C_{i-1} &= K_{i-1} = \frac{K_i - V_{min}}{2} + V_{min}. \end{aligned}$$

Öz (care)



Şekil 4. Bulanık küme kısımları

Artık bu anahtar değerler ( $K_1, K_2, K_3, K_4, K_5$ ) bulanık modelin girdileri olacaktır.

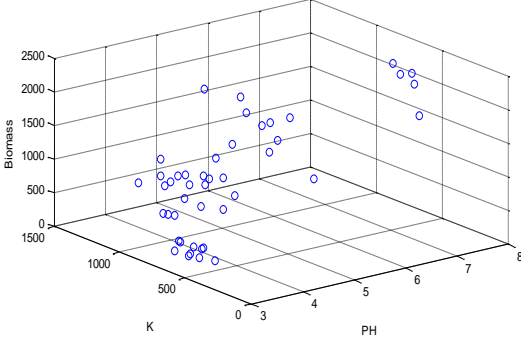
Her bir değişken için ilk ve son MFs'lerin anahtar değerleri arasında dağılan veri için bulanık modelimiz geçerli olacaktır. Tablo 1'deki gibi tablo hazırlanır.

## 3. Bulgular

Uygulamada, toprağın biyo-kütlesi ( $Y$ ) ve buna etki eden topraktaki PH ( $X_1$ ) ve Potasyum ( $X_2$ ) değerlerini içeren bir veri kullanılmış, serpilme diyagramı Şekil 5'te verilmiştir [24].

Bağımlı ve bağımsız değişkenler, SMRGT Metodu kullanılarak bulanıklaştırılmıştır.

$X_1$  ve  $X_2$ 'nin beşer üçgensel üyelik fonksiyonuna ayrılmasıyla elde edilen  $\tilde{X}_1$  ve  $\tilde{X}_2$ ,  $Y$ 'nin yirmi beşlik üçgensel üyelik fonksiyonuna ayrılmasıyla, MATLAB [25] programı yardımı ile elde edilen  $\tilde{Y}$  bulanık değerleri aşağıdaki tablolarda (degree) göre seçilmiştir.



Şekil 5. Bağımlı ve bağımsız değişkenler için serpilme diyagramı

$X_1$ (PH) için;

$$V_{min} = 3, V_{max} = 8$$

$$V_R = 5$$

$$UW = 0.625$$

$$EUW = 0.9375$$

$$K_1 = 3.3125, K_5 = 7.6875$$

$$K_i = 5.5, K_{i+1} = 6.75, K_{i-1} = 4.25$$

gösterilmektedir.

$X_2$ (potasyum) için;

$$V_{min} = 345, V_{max} = 1445$$

$$V_R = 1100$$

$$UW = 137.5$$

$$EUW = 206.25$$

$$K_1 = 413.75, K_5 = 1376.25$$

$$K_i = 895, K_{i+1} = 1170, K_{i-1} = 620$$

$Y$ (Biyo-kütle) için;

$$V_{min} = 230, V_{max} = 2440$$

$$V_R = 2210$$

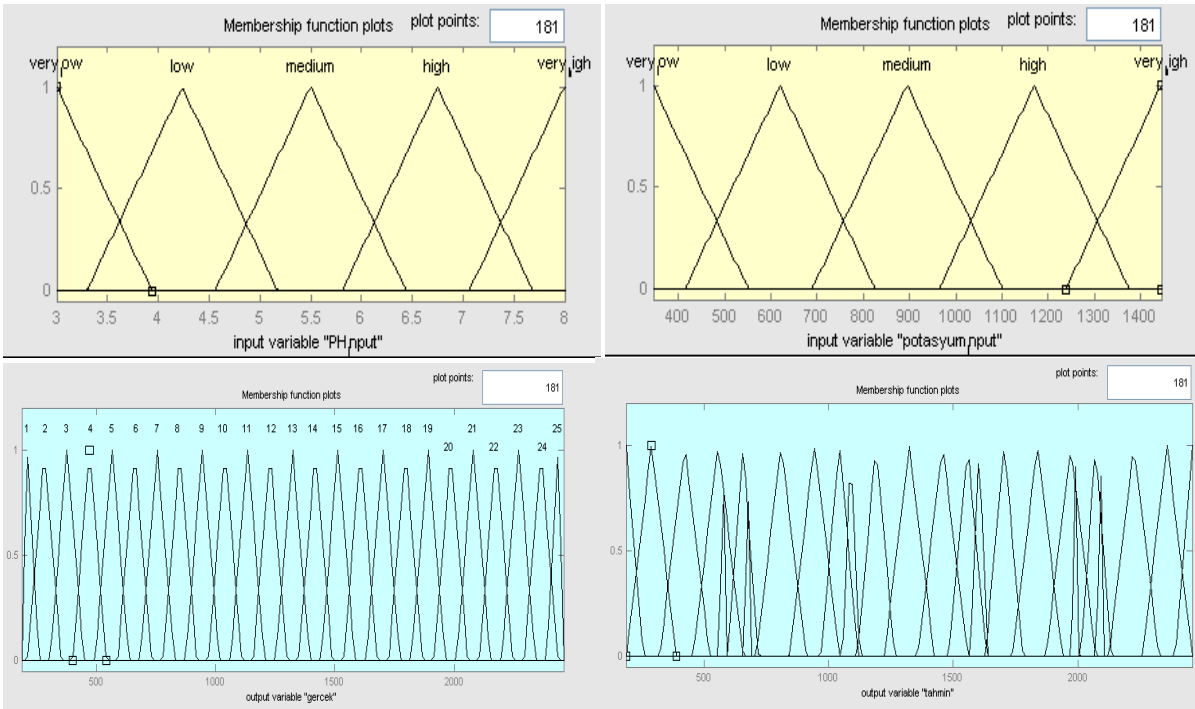
$$UW = 276.25$$

$$EUW = 414.375$$

$$K_1 = 368.125, K_5 = 2301.875$$

$$K_i = 1335, K_{i+1} = 1887.5, K_{i-1} = 782.5$$

$\hat{Y}$ 'nin elde edilmesinde Diamond'un yaklaşımı SMRGT Modeli için denenmiş ve Tablo 2'deki sonuçlar bulunmuştur.



Şekil 6. MATLAB programı ile oluşturulan MFs'ler

FRs Sayısı	X <sub>1</sub> (PH)		X <sub>2</sub> (potasyum)			Y Sözel Değer
	Sayısal Değer	Sözel Değer	Sayısal Değer	Sözel Değer	Sayısal Değer	
1	3.31	Çok az	413.75	Çok az	655.1824	1
2	3.31	Çok az	620	Az	554.718	2
3	3.31	Çok az	895	Orta	420.7655	3
4	3.31	Çok az	1170	Yüksek	286.813	4
5	3.31	Çok az	1376.25	Çok Yüksek	186.3486	5
6	4.25	Az	413.75	Çok az	1042.462	6
7	4.25	Az	620	Az	941.998	7
8	4.25	Az	895	Orta	808.0455	8
9	4.25	Az	1170	Yüksek	674.093	9
10	4.25	Az	1376.25	Çok Yüksek	573.6286	10
11	5.5	Orta	413.75	Çok az	1557.462	11
12	5.5	Orta	620	Az	1456.998	12
13	5.5	Orta	895	Orta	1323.046	13
14	5.5	Orta	1170	Yüksek	1189.093	14
15	5.5	Orta	1376.25	Çok Yüksek	1088.629	15
16	6.75	Yüksek	413.75	Çok az	2072.462	16
17	6.75	Yüksek	620	Az	1971.998	17
18	6.75	Yüksek	895	Orta	1838.046	18
19	6.75	Yüksek	1170	Yüksek	1704.093	19
20	6.75	Yüksek	1376.25	Çok Yüksek	1603.629	20
21	7.69	Çok Yüksek	413.75	Çok az	2459.742	21
22	7.69	Çok Yüksek	620	Az	2359.278	22
23	7.69	Çok Yüksek	895	Orta	2225.326	23
24	7.69	Çok Yüksek	1170	Yüksek	2091.373	24
25	7.69	Çok Yüksek	1376.25	Çok Yüksek	1990.909	25

Şekil 7. Bağımlı ve bağımsız değişkenler için oluşturulan bulanıklık kuralları

Tablo 2. Klasik Doğrusal Regresyon Analizi Sonucu (K.D.R.A.S) ve Bulanık Doğrusal Regresyon Analizi Sonucu (B.D.R.A.S:)

<i>i</i>	Gerçek (Y)	K.D.R.A.S ( $\hat{Y}$ )	Bulanık $\tilde{Y}$	B.D.R.A.S $\hat{\tilde{Y}}$	<i>i</i>	Gerçek (Y)	K.D.R.A.S ( $\hat{Y}$ )	Bulanık $\tilde{Y}$	B.D.R.A.S $\hat{\tilde{Y}}$
1	676	996,8925	208,698	186,3486	24	236	489,271	2365,21	2359,278
2	516	817,1646	279,792	286,813	25	340	545,7957	2436,302	2459,742
3	1052	661,1551	374,58	420,7655	26	2436	2162,156		
4	868	796,7074	469,38	554,718	27	2216	2246,899		
5	1008	1525,519	564,17	573,6286	28	2096	2319,844		
6	436	953,512	658,96	655,1824	29	1660	2339,123		
7	544	588,1929	753,75	674,093	30	2272	2299,102		
8	680	715,6351	848,54	808,0455	31	824	1035,148		
9	640	844,7052	943,33	941,998	32	1196	952,3478		
10	492	763,2751	1038,12	1042,462	33	1960	1176,864		
11	984	912,4976	1132,92	1088,629	34	2080	944,0054		
12	1400	673,4121	1227,71	1189,093	35	1764	1197,96		
13	1276	657,7129	1322,5	1323,046	36	412	885,4343		
14	1736	903,9977	1417,29	1456,998	37	416	657,0315		
15	1004	735,9515	1512,08	1557,462	38	504	559,6088		
16	396	517,0168	1606,87	1603,629	39	492	587,2279		
17	352	622,816	1701,67	1704,093	40	636	738,053		
18	328	640,5594	1796,46	1838,046	41	1756	1449,137		
19	392	631,4571	1891,25	1971,998	42	1232	1316,244		
20	236	569,6339	1986,04	1990,909	43	1400	1382,326		
21	392	571,0021	2080,83	2072,462	44	1620	1355,067		
22	268	593,1554	2175,62	2091,373	45	1560	1300,375		

Elde edilen klasik doğrusal regresyon ve bulanık doğrusal regresyon tahminlerinin performanslarını ölçmek için MATLAB'da kodlar yazılmış [26] ve çeşitli performans ölçüm kriterleri kullanılarak aşağıdaki sonuçlar oluşturulmuştur:

#### klasik Doğrusal Regresyon Analizi Sonucu

	KLASİK	BULANIK
	DOĞRUSAL REGRESYON	DOĞRUSAL REGRESYON
Adj- $R^2$	0.6268	0.9959
$R^2$	0.6438	0.9963
MAE	306.62	30.74
MAPE	43.89	3.577
MSE	151739.5	1729.55
RMSE	403.21	44.33

Burada kullanılan performans ölçüm kriterleri Düzeltilmiş Belirlilik Katsayısı (Adjusted  $R^2$  (Adj- $R^2$ )), Belirlilik Katsayısı ( $R^2$ ), Ortalama Mutlak Hata (Mean Absolute Error (MAE)), Ortalama Mutlak Yüzde Hata (Mean Absolute Percentage Error (MAPE)), Hata Kareler Ortalaması (Mean Square Error (MSE)) ve Hata Kareler Ortalaması Köküdür (Root Mean Square Error (RMSE)).

#### Tartışma ve sonuç

Bulanık doğrusal regresyon üzerine yaptığımız bu çalışma, yıllardır birçok alanda kullanılan bulanık doğrusal regresyon analizine yeni bir katkı sağlamaktadır. Belirli kısıtlamalar nedeniyle klasik doğrusal regresyon uygulanması yerine bulanık doğrusal regresyon uygulanması daha uygun olan veriler için, veri bulanık değişkenler içermiyorsa olasılıksal modelin kullanılması zorunlu hale gelebilmektedir. Oysaki olasılıksal modelin de belli kısıtlamalar nedeniyle kullanılmasının uygun olmadığı durumlar vardır. İşte böyle zamanlarda çalışmamızda, yeni bir yöntem olan SMRGT ile değişkenlerin bulanıklaştırılıp Bulanık EKK modelinin kullanılabilceği önerilmektedir.

Bu çalışma ayrıca, önerilen tekniğin ne kadar etkili olduğunu ölçmek için çeşitli metottan bağımsız performans ölçüm kriterleri sonuçlarını içermektedir. Görüldüğü gibi 6 performans kriteri göz önüne alındığında bulanık doğrusal regresyon, klasik doğrusal regresyondan daha iyi sonuçlar vermiştir. Böylece tekniğimizin sağladığı katkı gözler önüne serilmiştir.

#### Kaynaklar

Zadeh, L. A., (1965). *Fuzzy Sets*, Information and Control, 8: 338-353.  
 Şen,Z., (2001). *Bulanık Mantık ve Modelleme İlkeleri*, Bilge Yayıncılık, İstanbul  
 Bardossy, A., (1990). Notes on fuzzy regression. *Fuzzy Sets and Systems* 37, 65-75.  
 Yu Xue, I.S. Kim, J.S. Son, C.E. Park, H.H. Kim, B.S. Sung, I.J. Kim, B.Y. Kang., (2005). *Fuzzy regression method for prediction and control the bead width in the robotic arc-welding process*, Journal of Materials Processing Technology 164-165, 1134-1139.  
 Shapiro, A.F., (2004). *Fuzzy regression and the term structure of interest rates revisited*. In:

6. Diamond, P., (1988). *Fuzzy least squares*, Information Sciences 46(3), 141-157.  
 7. Celmins, A., (1987a). *Multidimensional least-squares fitting of fuzzy models*, Mathematical Modelling 9 (9), 669-690.  
 8. Celmins, A., (1987b). *Least-squares model fitting to fuzzy vector data*, Fuzzy Sets and Systems 22, 245-269.  
 9. Tanaka, H., Uejima, S., Asai, K., (1982). *Linear regression analysis with fuzzy model*, IEEE Systems, Trans. Systems Man Cybernet. SMC-2, 903-907.  
 10. Yurtçu, Ş., İcağa, Y., (2007). *Bulanık doğrusal regresyona genel bir bakış*, Yapı Teknolojileri Elektronik Dergisi, (2) 37-43.  
 11. Deniz, E., (2006). *Bulanık Mantık Tabanlı Tahmin Modeli ve Uygulaması*, Yüksek Lisans Tezi, Muğla Üniversitesi, F.B.E., s. 41, ss.97.  
 12. Düzyurt S., (2008). *Bulanık Regresyon İle Tahmin Ve Bir Uygulama*, Yayınlanmamış Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü Ankara  
 13. Chang, Y.H.O., Ayyub, B.M., (2001). *Fuzzy regression methods - a comparative assessment*, Fuzzy Set and Systems, 119(2), 187-203.  
 14. Uras, Y., (1998). *Bulanık Mantığın Doğrusal Regresyon Analizinde Kullanılmasına İlişkin Bir Uygulama*, Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi, S.B.E., ss.111.  
 15. Peters, G., (1994). *Fuzzy linear regression with fuzzy intervals*, Fuzzy Sets and Systems, 63(1) 45-55.  
 16. Wang, H.F., Tsaur, R.C., (2000). *Insight of a fuzzy regression model*, Fuzzy Sets and Systems, 112(3), 355-369  
 17. Savic, D.A., W. Pedrycz, (1991). *Evaluation of fuzzy linear regression models*, Fuzzy Sets and Systems, 39(1), 51-63  
 18. Kim, K.J., Moskowitz, H., Koksalan, M., (1996). *Fuzzy versus statistical linear regression*, European Journal of Operational Research, 92(2) 417-434  
 19. Hojati, M., Bector, C.R., Smimou, K., (2004). *A simple method for computation of fuzzy linear regression*, European Journal of Operational Research (forthcoming)  
 20. Kao, C., Chyu, C.L., (2003). *Least-squares estimates in fuzzy regression analysis*, European Journal of Operational Research 148, 426-435  
 21. Shapiro, A.F., *Fuzzy Regression Models*, <https://www.soa.org/.../arch06v40n1-ii.pdf>, Erişim tarihi: 17.02.2014.  
 22. Arabpour, A.R., Tata, M., (2008). *Estimating the parameters of a fuzzy linear regression model*, Iranian Journal of Fuzzy Systems Vol. 5, No. 2, pp. 1-19.  
 23. Toprak, Z.F., (2009). *Flow Discharge Modeling in Open Canals Using a New Fuzzy Modeling Technique (SMRGT)*, CLEAN-Soil. Air. Water, 37 (9), 742-752.  
 24. Rawlings, O.J., Pantula, S.G., Dickey, D.A., (1998), *Applied Regression Analysis*, Springer, USA.  
 25. MATLAB, <http://www.mathworks.com/products/matlab/tryit.html>.  
 26. Toprak, S., (2011). *Çok Değişkenli Uyarlamalı Regresyon Eğrileri ve Konik Programlama ile Zaman Serilerinin Modellenmesi*, Yüksek lisans tezi, Fen Bilimleri Enstitüsü, Diyarbakır. Proceedings of the 14th International AFIR Colloquium, vol. 1. pp.29-45.