SAU Fen Bilimleri Enstitüsü Dergisi
8.Cilt, 1.Sayı (Mart 2004)

PARÇALI REGRESYON YARDIMI İLE BİTKİ BOYU-ZAMAN İLİŞKİSİ PARAMETRELERİNİN TAHMİNİ

Ufuk KARADAVUT, Aşır GENÇ, Abdurrahman TOZLUCA, Alper SİNAN
Aydın KARAKOCA, Şeref AKSOYAK, Çetin PALTA

Özet - Bu çalışma Bahri Dağdaş Uluslararası Tarımsal Araştırma Enstitüsü Deneme alanlarında yürütülmüştür. Çalışmada Dağdaş-94 buğday çeşiti kullanılmıştır. Bitkilerde 10'ar gün aralıklarla düzenli olarak 15 kez boy ölçümleri yapılmıştır. Her parselden 5 bitki tesadüfen alınmış ve üç tekrarlamalı olarak yapılan ölçümlerde, her ölçüm için toplam 15 bitki kullanılmıştır.

Buğday bitkisinin boyunun zamana göre modellenmesi ve model parametrelerinin tahmini uygulama yönünden önem arzetmektedir. Bu amaçla, doğrusal olmayan bir model gösteren bitkinin boy uzunluğu ile zaman ilişkisi, belli noktalarda zıplamalar gösterirken bazı noktalarda kesilmeler gösterebilmektedir. Bu çalışmamızda buğday bitkisinin boy uzunluğu ve zaman ilişkisi açısından parçalı regresyon modelinin parametre tahmini ve zıplama noktalarının tahmin edilmesi ele alınmıştır.

Anahtar Kelimeler - Buğday, Parçalı regresyon,Bitki Boyu-zaman ilişkisi, Parametre tahmini

Abstract - This study was carry out in experimental areas of Bahri Dağdaş International Agricultural Research Institute . In this study was used Dağdaş-94 bread wheat cultivar. The plants was measurement interval 10 days. And as total 15 times. Every plot was gotten randomized 5 plants and 3 replications, and for every measurement was used 15 plants.

According to time of modelling of plant high that show a non linear models between plant high and time relationships show in some points ascends, some points descends. In this study wheat plants of high and time relation was investigated spline regression models and estimation of ascend points.

Key Words - Wheat, Spline regression, Plant High-time relation, Parameters estimation

U. Karadavut, B.D.Uluslararası Tarımsal Arş. Enst., Karatay-Konya
A. Genç, S.Ü. Fen Edebiyat Fak. İstatistik Böl., Kampus-Konya
A., Tozluca S.Ü. Ziraat Fak. Zooteknik Böl.,Kampus-Konya

I. GİRİŞ

Buğday, kültür bitkileri içerisinde yaklaşık 220 milyon *ha* ekim alanı ile oldukça geniş bir alanda tarımı yapılmaktadır. Sahip olduğu büyük uyum yeteneği ile her türlü iklim ve bölgede yetişebilmektedir. Buğday, insanlığın bir numaralı gıdası durumundadır. Dünya genelinde besinlerden sağlanan kalorinin %20'si buğdaya aittir [1]. Ülkemizde yaklaşık olarak 9.5 milyon *ha* alanda buğday tarımı yapılmaktadır [2]. Ülkemizdeki buğday tarımı %70 oranında kuru tarım alanlarında yapılmaktadır. Bu durum yıllık buğday üretimimizi büyük oranda yağışa bağımlı kılmaktadır [3]. Yağışa bağlı olan bitkiler, yağışın düzensizliği ile büyüme ritimlerinde de değişiklikler göstermektedirler. Bu değişiklik nedeniyle büyüme ve farklılaşmada da değişiklikler olabilmektedir. Bilindiği gibi bir canlının gelişmesi başlıca iki tip değişimle gerçekleşmektedir. Birincisi, niceliğe ait değişiklikler; uzunluk, ve hacim olarak artışı ifade eden boyutlardaki değişmedir. Bunlar zamana bağlı değişikliklerdir ve büyümeyi oluşturmaktadırlar. İkincisi ise, niteliğe ait değişikliklerdir; bunlar morfolojik ve fonksiyonel yeni özelliklerin kazanılması ya da daha genel bir ifade ile farklılaşmayı belirtmektedirler [4].

Buğday bitkisinin gelişme dönemlerinin dikkatli bir şekilde izlenmesi ve incelenmesi erken dönemde verimle ilgili tahminlerin yapılması bakımından yararlı olduğu gibi, ilaç ve gübre uygulama zamanlarının daha doğru bir şekilde belirlenmesi yönünden önemlidir. Elbette bu süreçlerde buğday bitkisinin boyunun zamana göre modellenmesi ve model parametrelerinin tahmini uygulama yönünden önem arzetmektedir. Bu amaçla, doğrusal olmayan bir model gösteren bitkinin boy uzunluğu ile zaman ilişkisi, belli noktalarda zıplamalar gösterirken bazı noktalarda kesilmeler gösterebilmektedir [5]. Buğdayın büyüme periyodu incelendiğinde çıkış ile rozet oluşumuna kadar geçen sürede hızlı bir büyüme ve gelişme göstermektedir. Kış boyunca büyüme ve gelişmesi durmakta ve yalnızca bitki yaşamsal faaliyetleri için gayret sarfetmektedir. Baharla birlikte havaların ısınması ile bitki yeniden büyüme ve gelişmeye başlamakta kardeşlenme ve sapa kalma ile bu süreç hızlanmaktadır. Bitki boyuna ait verilerin değerlendirilerek bu verilere bir tek eğri ile yaklaşmak büyük kolaylıklar sağlasa da bazı hallerde bu durum büyük hatalara neden olabilir [6]. Bu yüzden peş peşe gelen iki veri arasında birinci, ikinci yada üçüncü dereceden fonksiyonlarla yaklaşımın yapıldığı parçalı fonksiyonlar yöntemi kullanımı uygun olmaktadır [7]. Son zamanlarda deneysel verilerin istatistiksel analizinde parçalı fonksiyonlar oldukça yoğun bir şekilde kullanılmaktadır [8]. Bu çalışmamızda buğday bitkisinin boy uzunluğu ve zaman ilişkisi açısından parçalı regresyon modelinin parametre tahmini ve zıplama noktalarının tahmin edilmesi ele alınmıştır.

II. MATERYAL VE YÖNTEM

Bu çalışma Bahri Dağdaş Uluslararası Tarımsal Araştırma Enstitüsü Deneme alanlarında yürütülmüştür. Yapılan çalışmada deneme alanı 50 m^2 lik üç eşit parsele bölünmüş ve bu parsellerden bitkiler tesadüfen alınmıştır. Çalışmada **Dağdaş-94** ekmeklik buğday çeşiti kullanılmıştır. Bitkilerde 10'ar gün aralıklarla düzenli olarak 15 kez boy ölçümleri yapılmıştır. Her parselden 5 bitki tesadüfen alınmış ve üç tekrarlamalı olarak yapılan ölçümlerde, her ölçüm için toplam 15 bitki kullanılmıştır. Ekim işlemleri ekim ayı içerisinde yapılmıştır. Ekim işlemi m^2 ye 550 bitki gelecek şekilde mibzerle yapılmıştır. Ekimle birlikte 15 kg/da DAP (Diamonyun fosfat), üst gübre olarak ta 7 kg/da Amonyum Nitrat (%33) gübresi verilmiştir.

Parçalı Regresyon: $\{(x_i, y_i) : i = 1,2,\ldots, n\}$ veri kümesi verilmiş olsun. Çok sayıda *(x,y)* veri noktalarından tek bir eğri geçirmek her ne kadar uygun düşse de bu durum büyük hatalara neden olabilmektedir. Bu durumda nokta sayısı arttıkça polinomların derecesi artış göstereceğinden, peş peşe gelen iki veri arasında parçalı fonksiyonlarla yaklaşım yöntemi önerilmektedir. Parçalı interpolasyonu veri noktalarını çeşitli aralıklara bölerek, her bir aralıkta daha küçük dereceden polinomlarla yaklaşım yapma esasına dayanır [9].

Veri noktaları $(\underline{x}_1, y_1), (\underline{x}_2, y_2), \ldots, (\underline{x}_n, y_n)$ biçiminde ifade edilsin. $\underline{X}$ ile $\underline{Y}$ arasında farklı aralıklar üzerinde belirlenmiş

$$Y_t = E[\underline{Y}/\underline{X} = \underline{x}_t] + \varepsilon_t \qquad (1)$$

biçiminde bir regresyon fonksiyonu ele alınsın. Burada $E[\underline{Y}/\underline{X}] = f(\underline{x}, \underline{\theta})$ fonksiyonu farklı aralıklarda tanımlanmış farklı fonksiyonlardır ve

$$f(\underline{x}, \underline{\theta}, \underline{\alpha}) = \begin{cases} f_1(\underline{x}, \underline{\theta}_1) & , \quad x \le \alpha_1 \\ f_2(\underline{x}, \underline{\theta}_2) & , \quad \alpha_1 < x \le \alpha_2 \\ \vdots & , \quad \vdots \\ f_D(\underline{x}, \underline{\theta}_D) & , \quad \alpha_{D-1} < x \end{cases} \qquad (2)$$

biçimindedir. α ile belirlenmiş aralıkların uç noktaları bilinmemekte ve bu noktalara düğüm noktaları denilmektedir. Bu noktalar fonksiyonun sürekliliğini bozmadan sıçrama yada ani değişim noktalarıdır. Gözlemlere bağlı olarak bu noktaların tahmin edilmesi gerekir. Bu çalışmada bu noktalar sabit olarak alınacaktır. D alt modeli ortak noktaların yada düğüm noktalarının uç noktalarında modelin safhasını belirtir. Bu problem aynı zamanda parçalı regresyon problemi olarak ta bilinir. Bu tür modeller;

a) sistemdeki parçalı modellerin sayısının küçük olması
b) sistemin davranışının bir doğru fonksiyonu olarak iyi tanımlanmış olması
c) gözlemlerde ani değişimler olması

durumları için amaçlanır [10]. (2) ile verilen model parçalı regresyon modeli olarak bilinir. Burada, düzgün regresyon ilişkisinden daha ziyade regresyondaki yapısal değişimler yada ani değişimleri tanımlayan model üzerinde durulmaktadır. Parçalı regresyonda $f_d(\underline{x}, \underline{\theta}_d)$, $d = 1,2,\ldots, D$ modellerinin polinom olmaları ve α_d düğüm noktalarında ikinci türevlerinin sürekli olduğu varsayılır.

Sıralı $x_1 < x_2 < \ldots < x_n$ verileri için

$$y_i = E[y/x_i] = \begin{cases} \beta_{10} + \beta_{11}x + \beta_{12}x^2 & , \quad x \le \alpha \\ \beta_{20} + \beta_{21}x + \beta_{22}x^2 & , \quad x > \alpha \end{cases} \qquad (3)$$

modeli gözönüne alınsın. Eğer $E[y/x_i]$ sürekli olması kısıtlaması yoksa α değişim noktasının tahmini beklenmez. Bu durumda *i=K* model sayısı olmak üzere model,

$$y_i = \begin{cases} \beta_{10} + \beta_{11}x_i + \beta_{12}x_i^2 + \varepsilon_{1i} & , \quad i = 1,2,\ldots, K \\ \beta_{20} + \beta_{21}x_i + \beta_{22}x_i^2 + \varepsilon_{2i} & , \quad i = K+1,\ldots, n \end{cases} \qquad (4)$$

şeklinde olur. Eğer *K* parametre değeri önceden biliniyorsa problem adi lineer regresyon problemine dönüşür. Eğer *K* bilinmiyorsa sonuç çıkarım işlemlerinde bir çok sıkıntılar ortaya çıkabilir. Biz bu çalışmada *K* değerini biliniyor olarak kabul edeceğiz. Ayrıca burada $\varepsilon_{di} \sim N(0, \sigma_d^2)$, $\sigma_1^2 = \sigma_2^2$ ve ε_1 ve ε_2 bağımsız olduklarını varsayacağız. Bu durumda

$\underline{\beta}_1$ ve $\underline{\beta}_2$, σ_1^2 ve σ_2^2 bilinmeyen parametreleri tahmin edilecektir.

Matris formunda, $\underline{y}_1^{(K)} : K \times 1$ ve $\underline{y}_2^{(K)} : (n-K) \times 1$ olsun. Ayrıca,

$$\underline{y} = \begin{pmatrix} \underline{y}_1^{(K)} \\ \underline{y}_2^{(K)} \end{pmatrix}, \quad \chi = \begin{pmatrix} X_1^{(K)} & 0 \\ 0 & X_2^{(K)} \end{pmatrix}$$

$$\underline{\beta} = \begin{pmatrix} \underline{\beta}_1 \\ \underline{\beta}_2 \end{pmatrix}, \quad \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1^{(K)} \\ \varepsilon_2^{(K)} \end{pmatrix}$$

olmak üzere regresyon modeli,

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} \qquad (5)$$

biçiminde olur. Model parametresi $\underline{\beta}$'nın en küçük kareler tahmin edicisi

$$\hat{\underline{\beta}}^{(K)} = (X'X)^{-1} X'\underline{y} \qquad (6)$$

ile tahmin edilir. Buna göre,

$$\hat{\underline{\beta}}_d^{(K)} = (X_d^{(K)'} X_d^{(K)})^{-1} X_d^{(K)'} \underline{y}_d^{(K)}, d = 1,2 \qquad (7)$$

$$\hat{\sigma}_1^2(K) = \frac{\hat{S}_1^{(K)}}{K}$$

$$\hat{\sigma}_2^2(K) = \frac{\hat{S}_2^{(K)}}{n-K}$$

ve

$$\hat{S}_d^{(K)} = (\underline{y}_d^{(K)} - X_d^{(K)} \hat{\underline{\beta}}_d^{(K)})'(\underline{y}_d^{(K)} - X_d^{(K)} \hat{\underline{\beta}}_d^{(K)}) \qquad (8)$$

modelin hata kareler toplamıdır. $\sigma_1^2 = \sigma_2^2$ olduğundan

$$\hat{\sigma}^2(K) = \frac{\hat{S}_1^{(K)} + \hat{S}_2^{(K)}}{n}$$

dır. Bu durumda,

$$\hat{\underline{\theta}}_d \sim N_{P_d}(\underline{\theta}_d, \hat{\sigma}^2 (X_d' X_d)^{-1}) \qquad (9)$$

şeklinde bir dağılım gösterir ve

$$\hat{\sigma}^2 = \frac{S(\hat{\underline{\theta}})}{n-p}$$

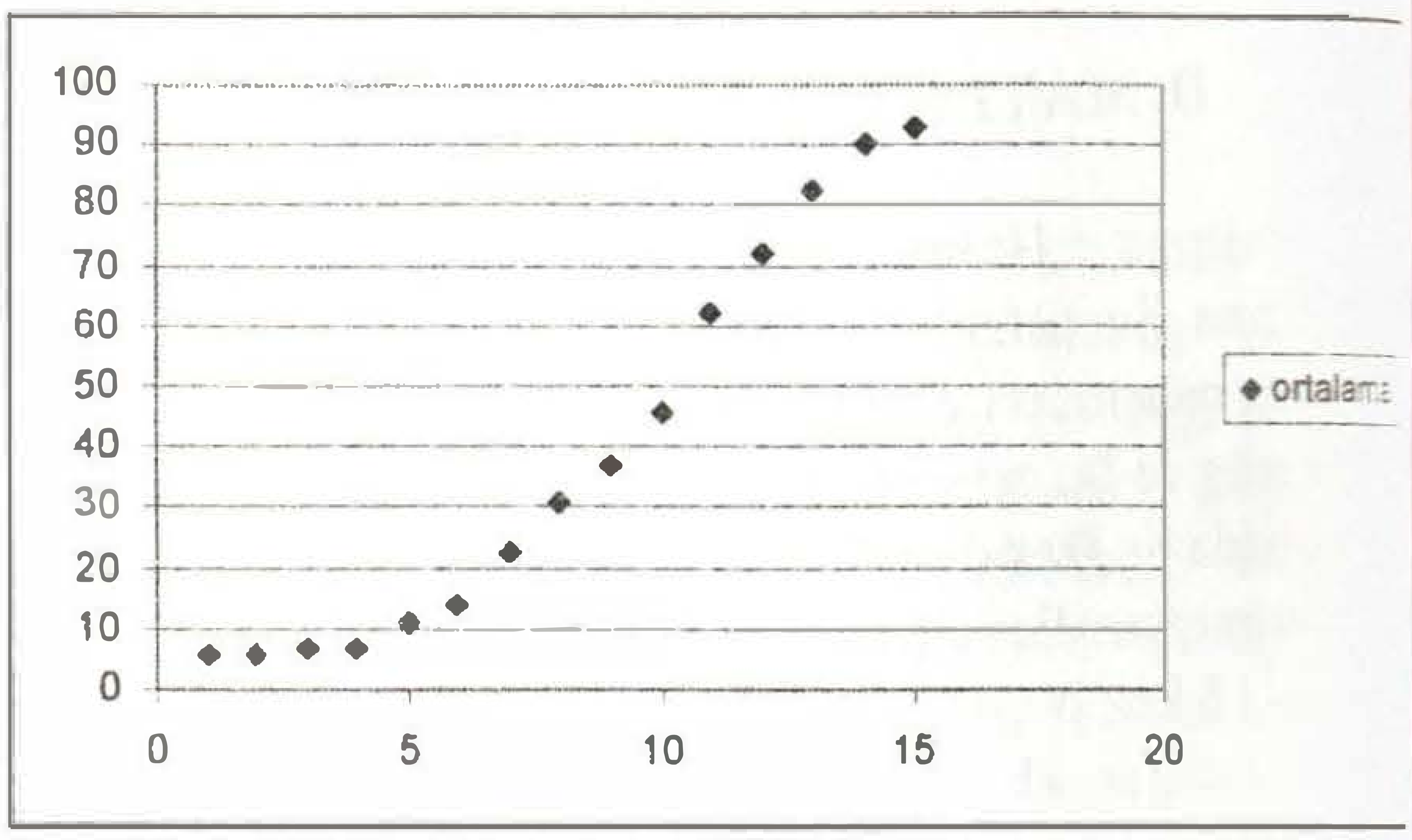
olarak elde edilir. Buradan,

$$\sqrt{n}(\hat{\underline{\theta}} - \underline{\theta}^*) \sim N_p(\underline{0}, \hat{\sigma}^2 (X'X)^{-1}) \qquad (10)$$

yazılabilir. Bu sonuç, parametreler ve model hakkında bir istatistiksel sonuç çıkarımı yapmak için yeterlidir.

III. BULGULAR VE TARTIŞMA

Bu çalışmada Konya ilinde yetiştirilen Dağdaş buğd çeşidinin büyümesi incelenmiştir. Bunun için 10' günlük aralıklarla ölçülen bitki boyu zamana kar değişmesi ile ilgili bir matematiksel model kurmak ve b model üzerinde, elde edilen ölçüm değerlerini istatistiksel açıdan geçerliliğinin kontrol edilmes hedeflenmektedir. Bunun için de elde edilen ölçüm değerlerinin zamana karşı davranışına bakarak bir takı regresyon modelleri karşılaştırılmıştır. Yaptığımız çalışmada elde edilen ölçüm sonuçlarına ilişkin ortalama



değerlere ilişkin dağılımlar Şekil 1' görülmektedir.

Şekil 1. Bitkilerin zamana göre büyüme grafiği

Şekil 1' in incelenmesinden de anlaşılacağı gibi ilk zamanlar büyüme oldukça yavaş olmuş ve 6 ölçüm kadar böyle devam etmiştir. Ancak bu noktadan sonra büyüme hız kazanmıştır. Şekil 1 dikkate alındığında bu veriler için aşağıdaki modeller öne sürülebilir.

$\underline{\theta} = (\theta_1, \theta_2, \Lambda, \theta_p) \in \Theta \subset IR^p$ bilinmeyen model parametresi; $x=1,2,...,15$ 10'ar günlük aralarla yapılan ölçümleri göstermek üzere açıklayıcı (bağımsız) değişken, $\underline{Y'} = (Y_1, Y_2, \Lambda, Y_n)$ açıklanan (bağımlı) değişken olmak üzere,

$$Y_i = f(x_i, \underline{\theta}) + \varepsilon_i \quad , \quad i = 1,2,...,15 \qquad (11)$$

Model 1.

$$f(\underline{x}, \underline{\theta}) = \theta_0 + \theta_1 x_i + \theta_2 x_i^2 + \theta_3 x_i^3 \quad , \quad i = 1,2,...,15$$

Model 2.

$$f(\underline{x}, \underline{\theta}) = \begin{cases} \theta_0 + \theta_1 x_i^2 & , \quad i = 1,2,...,6 \\ \theta_2 + \theta_3 x_i + \theta_4 x_i^2 + \theta_5 x_i^3 & , \quad i = 7,8,...,15 \end{cases}$$

Burada ε_i hata terimidir. ε_i hata terimleri üzerinde $i=1,2,...,15$ için : Beklenen değerleri sıfır, eşit sabi varyanslı, bağımsız ve aynı dağılımlı oldukları varsayılacaktır. Bu durumda gözlemler yardımıyla Model 1 ve Model 2'deki $f(\underline{x},\underline{\theta})$ matematiksel modelinin bilinmeyen $\underline{\theta}$ parametreleri en küçük kareler (EKK) yöntemi ile elde edilen tahmin değerleri Tablo1 ve Tablo 2'de verildiği gibi elde edilmiştir.

Tablo 1. Model 1'e ilişkin tahmin sonuçları

	Katsayılar	Standart hata	T	P
θ_0	14,505	3,782	3,84	0,003
θ_1	-7,983	1,981	-4,03	0,002
θ_2	1,665	0,283	5,88	0,000
θ_3	-0,051	0,012	-4,41	0,001
S = 2,789			R-Sq = 99,4%	

Tablo 1'in incelenmesinden de görüleceği gibi, S değeri 2,789 ve R^2 değeri ise %99,4 bulunmuştur. Buna göre büyümeyi tek bir parça olarak ele alan 1 numaralı modelimizde bulunan açıklayıcı değişkenlerimiz açıklanan değişkenin % 99,4 lük bir kısmını açıkladığı anlaşılmaktadır.

Tablo 2. Model 2'ye ilişkin tahmin sonuçları

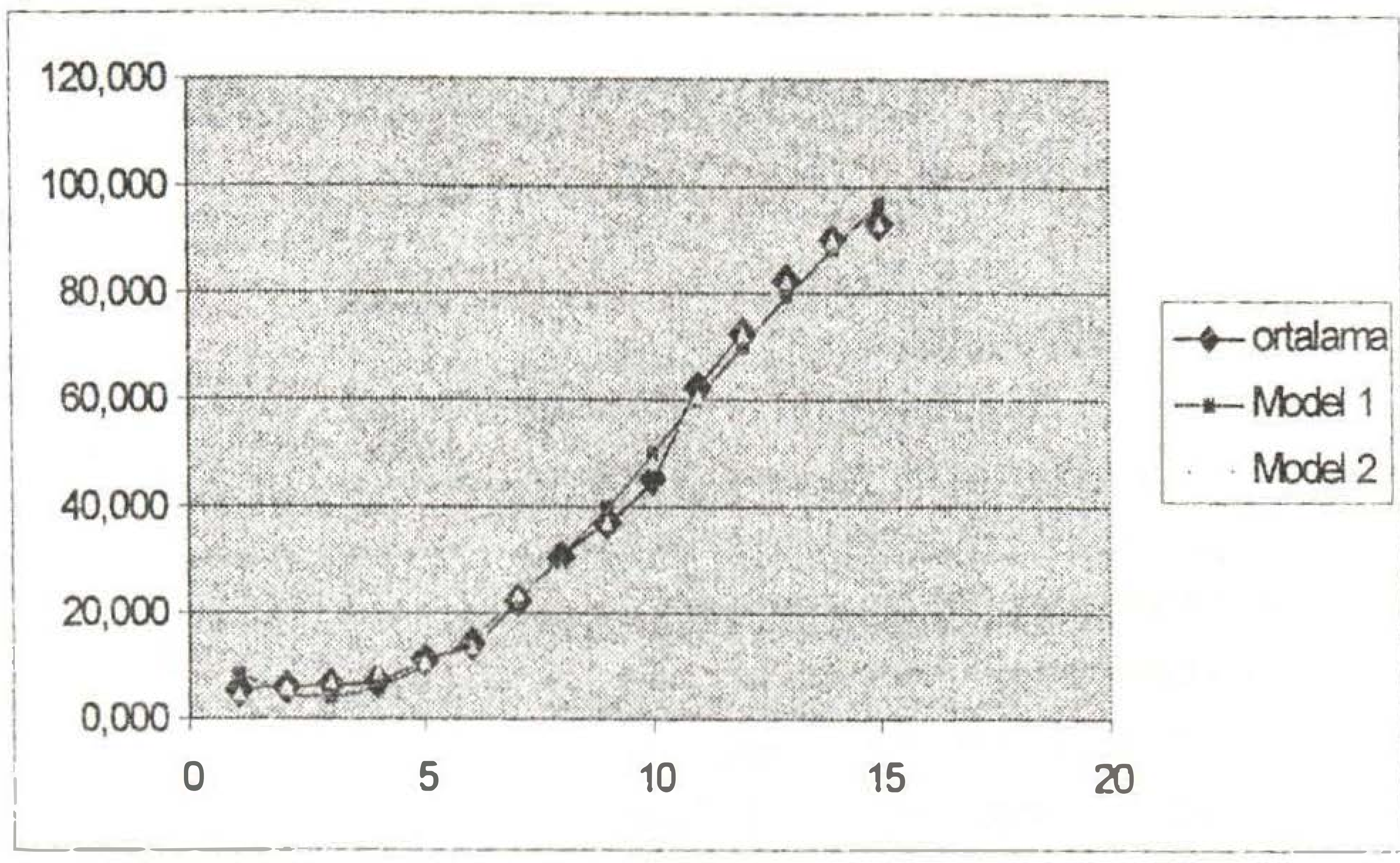
	Katsayılar	Standart hata	T	P
β_0	4,638	0,971	4,78	0,001
β_1	0,246	0,050	4,93	0,001
β_2	197,64	48,480	4,08	0,003
β_3	-63,40	14,000	-4,53	0,001
β_4	7,02	1,306	5,37	0,000
β_5	-0,217	0,039	-5,50	0,000
S = 1,491			R-Sq = 99,9%	

Tablo 2'ye bakıldığında Model 2 ye ilişkin hem R^2 değeri %99,9 ve S değeri ise 1,491 ile oldukça düşük olmuştur. Bitki büyümesini belli dönemlerde hızlı büyüme gösterirken, bazı dönemlerde duraksama göstermektedir. Burada sapa kalkmaya kadarki dönemde bitki hızlı bir gelişme göstermekte kış döneminde büyüme durma noktasına gelmektedir. Baharla birlikte özellikle havaların ısınması ile birlikte bitki yeniden ve hızlı bir şekilde büyümeye başlamaktadır. Büyümeyi iki kısma ayırırken bu noktayı dikkate alarak parçalama işlemini yaptık. Bu işlemin uygun olduğu elde edilen R^2 ve S değerleri bize göstermektedir. Dolayısıyla bu bitki boyu verileri için modellerin R^2 ve S değerlerine bakarak, parçalı regresyon modelinin daha uygun olduğunu söyleyebiliriz. Bu durum Şekil 1'de daha iyi görülmektedir.

Model 1 ve Model 2'nin parametrelerinin anlamlılığını α=0.05 anlam düzeyinde;

$$\begin{aligned} H_0 &: \theta_i = 0 \\ H_1 &: \theta_i \neq 0 \end{aligned} \tag{12}$$

hipotezi ile test ettiğimizde, tüm θ_i parametrelerine ilişkin p değerleri α=0.05 değerinden daha küçük olduğu için H_1 hipotezi kabul edilmektedir. Yani her iki modeldeki parametreler istatistiksel olarak anlamlıdır.



Şekil 2. Gerçek değerler ile tahmin edilen değerler

IV. SONUÇ

Yapılan çalışma, bitki büyümesinin zamana göre değişimini incelemenin tek bir model ile yapmak yerine büyüme dönemlerini ayrı ayrı incelemek sureti ile büyümeyi daha iyi tanımlama imkanı araştırılmıştır. Sonuç olarak büyümeyi dönemler halinde incelemenin bütün olarak incelemeye göre daha uygun olduğu sonucuna varılmıştır. İleride bu konuda yapılacak çalışmalarda bu konunun dikkate alınması çalışmanın başarısını artıracaktır.

V. KAYNAKLAR

[1]. Akkaya, A. 1994. Buğday Yetiştiriciliği. Kahramanmaraş Sütçü İmam Üniversitesi Genel Yayın No:1, Ziraat Fak Yay. No:1, Ders Kit. Yay. No:1. Kahramanmaraş.

[2].DİE, 2001. Tarımsal Yapı ve Üretim.TC.Başbakanlık DİE. Ankara.

[3]. Kün, E. 1988. Serin İklim Tahılları. Ankara Ünv. Ziraat Fak. Yay. 1032, Ders Kitabı No: 299. Ankara

[4]. Akman, Y; Darıcı, C.1998. Bitki Fizyolojisi (Beslenme ve Gelişme Fizyolojisi). Kariyer Matbaacılık. Ankara.

[5]. Karadavut, U., 2002.Bitkilerde Büyüme Modellemesi Teknikleri. S.Ü. Fen Bilimleri Enstitüsü. Konya.

[6]. Ericson, R.O., Modelling of Plant Growth. Annual Rewiev of Plant Physiology. 27:407-434.

[7]. Genç, A.; Kınacı, İ.; Kurnaz, A.; Oturanç, G., 2002. Statistical Analysis of Solar Radiation Using Cubic Spline Function. Energy Sources, September. Vol24,113-128

[8]. Genç, A.; Kınacı, İ.; Kurnaz, A.; Oturanç, G., 2000. Parçalı Regresyon Kullanarak Güneş Radyasyon Verilerinin Değerlendirilmesi. 6. Uluslararası Türk-Alman Enerji Kongresi. Proceeding CD'si.

[9]. Scheid, F., 1989, Theory and Problems of Numerical Analysis, Mc Graw-Hill, Boston University.

[10]. Seber, G.A.F., Wild, C.J., 1988, Nonlinear Regression, John Wiley&Sons, New York