

Finans Kurumları Yazılım Projelerinde Efor Tahmini İçin Parametrik Yöntemlerin Analizi

Ökkeş Emin Balçıcek, Mevlüt Güçlü, Özgür Şahin

Kuveyt Türk Katılım Bankası, Ar-Ge Merkezi, Konya

emin.balcicek@kuveytturk.com.tr, mevlut.guclu@kuveytturk.com.tr, ozgur.sahin@kuveytturk.com.tr

Özet : Yazılım projelerinde, doğru efor tahmininde bulunmak, diğer sektörlerde olduğu gibi finans kurumlarında da oldukça kritik öneme sahiptir. Günümüze kadar bu ihtiyacı karşılamak amacı ile birçok yazılım efor tahmin etme metodu aracı ve modeli geliştirilmiştir. Bu tahmin teknikleri yardımıyla yazılım projelerinin efor, süre ve teslim hızı gibi büyüklükleri hesaplanarak projenin planlanması ve yürütülmesi gerçekleştirilmektedir. Bu makalede en iyi tahmin yöntemini araştırmak amacı ile Bailey-Basili, Doty, Walston-Felix, Halstead, Boehm ve COCOMO gibi parametrik tahmin modelleri kullanılarak, yaklaşık 300 şubeli bir katılım bankasında gerçekleşen 20 yazılım projenin efor tahminlemesi yapılmaktadır. Çeşitli projelerin uzman tahmini yöntemi kullanılarak yapılan tahminlemeye göre hedeflenen ve proje teslimi sonrasında kesin olarak tespit edilebilen, gerçekleşen adam-gün verileri kullanılmıştır. Bu projelerin kaynak kodlarının satır sayıları KLOC cinsinden hesaplanmış 10 farklı hesaplama yöntemi ile kestirimler gerçekleştirilmiştir. Bulunan sonuçlar arasında en iyi yöntemin tespiti için Ortalama Karekök Hata, Ortalama Yüzdesel Hata ve korelasyon katsayısı istatistiksel formülleri kullanılmıştır. Bunlara ek olarak, bu istatistiksel fonksiyonların yardımı ile varolan veri kümesine ait gerçekleşen adam-gün değerlerine yakınsamak amacı ile yapılan iterasyon sonucunda yeni bir denklem elde edilmiş ve önerilmiştir. Yapılan bu analizler ve çıkarımları sonuç bölümünde paylaşılmıştır.

Anahtar Sözcükler : Parametrik Tahminleme, Yazılım Tahminleme, Efor Tahminleme, Yazılım Büyüklüğü Tahminleme, Kod Satır Sayısı (Lines Of Code – LOC), Uzman Tahminleme, Yazılım Tahmini Hata Analizi, RMSE (Ortalama Karekök Hata) , MMRE (Ortalama Yüzde Hata), Korelasyon Katsayısı.

Abstract: Estimating the effort correctly in software projects is a critical issue for financial enterprises as well as other sectors. Until today, plenty of models and tools have been developed to assess and interpret the accuracy of software development effort estimates. With these estimation techniques, software projects are planned and implemented by evaluating the scalars like effort, duration and delivery speed. In this research, twenty software projects of a 300-branched participation bank is estimated using parametric estimating models like Bailey-Basili, Doty, Walston-Felix, Halstead, Boehm and COCOMO in order to find out the best estimation technique. These estimations are based on the person-day data of projects that were implemented in the estimated time in consideration of professional estimation methods. Source code lines of these selected projects are processed as a KLOC and estimations are fulfilled with ten different calculation methods. On behalf of determining the best approach for result set, statistical formulas (mean square error, mean percentage error, coefficient of correlation) are evaluated with calculated parameters. In addition to these calculations, a new approach is obtained and suggested as a result of iteration through the data set which is collected by using statistical formulas in order to converge actual person-day values. Analyses and inferences are shared in result section.

Keywords : Parametric Estimating, Software Estimating, Effort Estimating, Software Size Estimating, Lines Of Code – LOC, Professional Estimating, RMSE (Root Mean Square Error) , MMRE (Mean Magnitude of Relative Error), Correlation

1) Giriş

Yazılım efor tahminleme, maliyetlerin yönetimi ve kaynakların planlaması için oldukça önemli ve gerekli bir ihtiyaçtır. Bu nedenle birçok kantitatif ve kalitatif yazılım efor tahminleme, hedef maliyet yönetim modeli geliştirilmiştir. Bu modelleri kategorize edersek,

- Uzman tahmini modeli
- Analoji tabanlı tahmin metotları
- Parametrik modeller
- Bulanık mantık tahmin şemaları
- Yapay sinir ağ tabanlı metotlar
- Karar ağacı yöntemi
- Deneysel yöntemler olarak sıralanabilir [1].

Bu çalışmada parametrik tahmin modelleri kullanılarak, finans kurumlarında yazılım projelerinin kod satır sayısı ve uzman tahmini tabanlı efor tahmin kıyaslaması yapılmaktadır. Elde edilen efor tahminlerine göre finans kurumları yazılım projelerine uygun model analizi ve uygun model önerisinde bulunulmuş ve uzman tahminin modeller ile kıyaslanması elde edilmiştir. Elde edilen veriler ile önerilen modelin hata analizler ve korelasyon ilişkisi çıkartılmıştır.

Yazılım efor tahmininde kullanılan tipik parametrik ana modeller aşağıda maddeler şeklinde sıralanabilir.

- Walston-Felix Modeli
- Bailey-Basili Modeli
- Doty
- Halstead
- Barry Boehm
- COCOMO

Bu modeller, çeşitli kuruluşlarda tamamlanan çok sayıda yazılım projesi ve uygulamalarından elde edilmiştir [2].

2) Yazılım Projelerinde Parametrik Efor Tahmini Yöntemleri

Yazılım projelerinde efor ve zamanlama tahmini, yazılım boyutuna göre parametrik modeller kullanarak da hesaplanabilir. Yazılım proje efor ve zamanlama tahminlerinde parametrik modelde hesaplamak için dört temel adım vardır. Bu adımlar müteakip maddelerde sıralandığı gibi özetlenebilir [3].

- Kod Satır Sayısı veya İşlev Puanı (Function Points) yöntemi yardımıyla yazılım projelerinde büyüklük tahmini.
- Adam-Gün veya adam-saat olarak yazılım efor tahmini.
- Gün olarak proje zamanlama tahmini.
- Proje efor ve zamanlamasından proje maliyet tahmini.

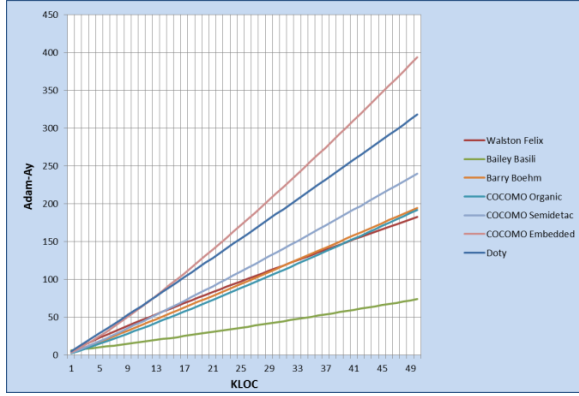
Yazılım projelerinin efor tahmininde birçok parametrik model kullanılmaktadır. Tablo 1’de tipik parametrik modeller denklemleriyle birlikte sunulmaktadır. Denklemlerde kullanılan KLOC parametresi, “Bin Kod Satırı (Kilo Lines of Code)” olarak kullanılmaktadır.

Tablo 1: Tipik Parametrik Modeller ve Denklemleri

S.N.	Model Adı	Denklem
1	Walston –Felix Modeli	$5.2 \times (KLOC)^{0.91}$
2	Bailey – Basili Modeli	$5.5 + 0.73 \times (KLOC)^{1.16}$
3	Halstead Modeli	$5.2 \times (KLOC)^{1.5}$
4	Doty Modeli	$5.288 \times (KLOC)^{1.047}$
5	Barry Boehm Modeli	$3.2 \times (KLOC)^{1.05}$
6	COCOMO Organic Modeli	$2.4 \times (KLOC)^{1.05}$
7	COCOMO Semidetac Modeli	$3.0 \times (KLOC)^{1.12}$
8	COCOMO Embedded Modeli	$3.6 \times (KLOC)^{1.2}$

C.E. Waltson ve C.P. Felix [5] tarafından 1977 yılında IBM’de geliştirilen Waltson-Felix Model, efor ile kaynak kod satır sayısı arasında ilişki sağlar. Bailey ve Basili [6], verilen geliştirme ortamına en iyi uyum sağlayan efor tahmin denklemlerin geliştirilmesine izin veren bir meta-model olarak 1981 yılında tanımlanmıştır. Oluşan tahmin modeli, projelerin arasındaki farklar ve onların çevresel faktörleri gibi veriler toplanmasına dayalı COCOMO ile benzerdir. M.H. Halstead [8], programlama yapısında derinlemesine analize gerek olmadan hata oranını öngören modeli 1977 yılında önermiştir. Bu model yazılımın kod uzunluğu ve hacim metrikleri önermek için kullanılır. 1977 yılında yayımlanan Doty Modeli [9] kod satır sayısı için eforları tahmin etmede kullanılır. Bu model kullanıcı katılımlı, hafıza sınırlamalı ve müşteri odaklı değişiklik yeteneğine sahiplik gibi yazılım geliştirme ortamlarını çeşitli yönleriyle oluşturur.

Bu modeller, çeşitli kuruluşlarda tamamlanan çok sayıda yazılım projesinden türetilmiştir. Söz konusu modellerin yazılım projelerinin tipine göre farklı sonuçlar vermesi beklenmektedir [10]. Bu modellerin 0-50 Bin Kod Satır Sayısı (Kilo Lines of Code-KLOC) arasındaki efor değerleri Şekil.1'de gösterilmektedir.



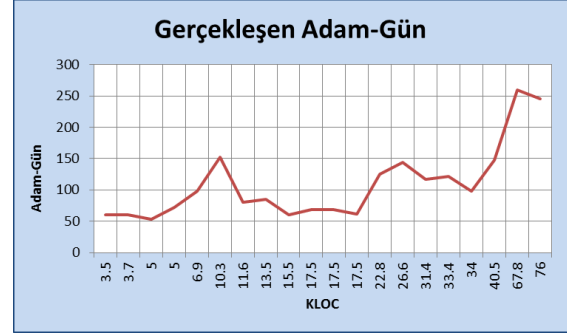
Şekil 1 : 0-50 KLOC Satır Sayısının Modellere Göre Adam-Ay Çıktıları

Parametrik modelde KLOC girdi, x adam-ay ise çıktı olarak modellenmiştir.

Hesaplama yöntemine ve analize, uzman tahmini de eklenerek parametrik model sonucu oluşan tahminlemeler ile kıyaslanması sağlanmış ve modele bağlı tahminleme, önerilen model ve uzman tahmini arasında hata analizleri ortaya konulmuştur.

3) Finansal Kurumlarda Yazılım Efor Tahmini

Bu bölümde, yaklaşık 300 şubeli bir katılım bankasında planlanıp, yürütülmüş ve tamamlanmış 20 farklı yazılım projesinin hedeflenen ve gerçekleşen adam-gün tahminleri ile teslim sonrasında oluşan satır sayıları bir veri kümesi olarak kabul edilmiş ve üzerinden ilgili analizler gerçekleştirilmiştir. Burada asıl amaç finansal kurumlarda yazılım efor tahminini en iyi tespit edecek yöntemin tespit edilmesi ve geliştirilmesidir. Veri kümesi üzerinde iki değişken mevcut olup bu değişkenler Kod Satır Sayısı (LOC), Bin Kod Satır Sayısı (KLOC) olarak tanımlanmıştır. Eforun birimi ise adam-gündür.



Şekil 2 : Gerçekleşen KLOC / Adam-Gün Grafiği

Tablo 2 : Finansal Kurumda gerçekleştirilen 20 Farklı Projenin Veri Kümesi

Proje No	KLOC	Gerçekleşen Adam-Gün
1	3.5	60
2	3.7	60
3	5	53
4	5	72
5	6.9	98
6	10.3	152
7	11.6	80
8	13.5	85
9	15.5	60
10	17.5	68
11	17.5	68
12	17.5	62
13	22.8	125
14	26.6	144
15	31.4	117
16	33.4	122
17	34	98
18	40.5	148
19	67.8	260
20	76	245

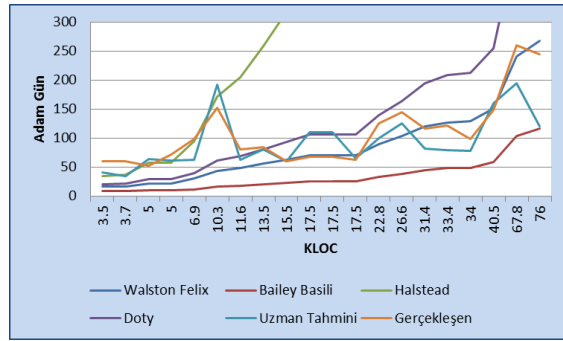
Tablo 2'de Finansal kurumda yapılan 20 farklı yazılım projesinin verisi bulunmaktadır. Her bir yazılım projesinin uygulamaya dayalı gerçek eforları tespit edilmiştir. Projelerin kod satır sayısına bağlı gerçek efor değerleri Şekil 2'deki grafikte görsel olarak gösterilmiştir [6]. Projeler seçilirken kurumda gerçekleştirilen farklı modüllere ait ve farklı tipteki geliştirme projeleri seçilmiştir. Düzeltme ve hata ayıklama projeleri kapsam dışında tutulmuştur.

Finansal kurumda yazılım projeleri veri seti kullanılarak farklı parametrik modeller ve uzman tahmini efor değeri olarak karşılaştırılmıştır. Bu modeller için ayrı ayrı efor değerleri hesaplanmıştır. Parametrik modellerde model çıktı birimi adam-ay olarak tasarlanmıştır. Çıktı birimini Adam-ay olarak kullandığımızda proje sürelerinin gerçekleşen sürelerden çok farklı olduğu görülmüş ve analizde çıktı birimi Adam-Gün olarak kabul edilmiştir. Her

bir modelin tahmini efor değeri finansal kurum yazılım projelerinin gerçek efor değeri ile kıyaslaması Tablo III'de ve Şekil 3'de sunulmaktadır.

Tablo 3A : Waltson Felix, Bailey Basili, Halstead, Doty Modelleri ve Uzman Tahmini Yöntemi Tahmin Değerlerinin Gerçek Efor (Adam-Gün) Karşılaştırma Veri Seti

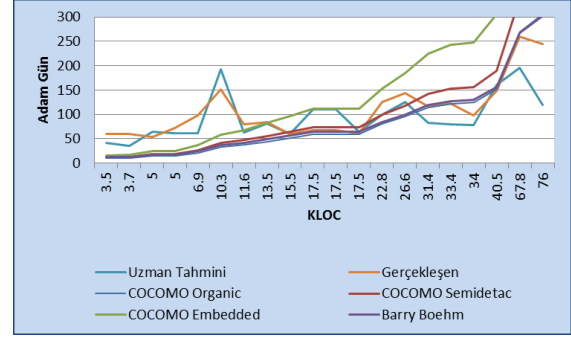
Proje	KLOC	Walston Felix	Bailey Basili	Halstead	Doty	Uzman Tahmini	Gerçekleşen
1	3.5	16	9	34	20	41	60
2	3.7	17	9	37	21	35	60
3	5	22	10	58	29	64	53
4	5	22	10	58	29	61	72
5	6.9	30	12	94	40	62	98
6	10.3	43	16	172	61	192	152
7	11.6	48	18	205	69	63	80
8	13.5	56	20	258	81	81	85
9	15.5	63	23	317	93	60	60
10	17.5	70	26	381	106	110	68
11	17.5	70	26	381	106	110	68
12	17.5	70	26	381	106	65	62
13	22.8	89	33	566	140	100	125
14	26.6	103	38	713	164	125	144
15	31.4	120	45	915	195	82	117
16	33.4	127	48	1004	208	79	122
17	34	129	49	1031	212	78	98
18	40.5	151	59	1340	255	160	148
19	67.8	241	103	2903	437	195	260
20	76	268	116	3445	493	120	245



Şekil 3A : Waltson Felix, Bailey Basili, Halstead, Doty Modelleri ve Uzman Tahmini Yöntemi Tahmin Değerlerinin Gerçek Efor (Adam-Gün) Karşılaştırma Grafiği

Tablo 3B : COCOMO, Barry Boehm Modelleri ve Uzman Tahmini Yöntemi Tahmin Değerlerinin Gerçek Efor (Adam-Gün) Karşılaştırma Veri Seti

Proje	KLOC	COCOMO Organic	COCOMO Semidetac	COCOMO Embedded	Barry Boehm	Uzman Tahmini	Gerçekleşen
1	3.5	10	12	16	12	41	60
2	3.7	10	13	17	13	35	60
3	5	15	18	25	17	64	53
4	5	15	18	25	17	61	72
5	6.9	21	26	37	24	62	98
6	10.3	33	41	59	37	192	152
7	11.6	37	47	68	42	63	80
8	13.5	44	55	82	49	81	85
9	15.5	52	65	97	57	60	60
10	17.5	59	74	112	65	110	68
11	17.5	59	74	112	65	110	68
12	17.5	59	74	112	65	65	62
13	22.8	80	100	153	85	100	125
14	26.6	95	118	185	100	125	144
15	31.4	114	142	225	119	82	117
16	33.4	122	153	243	127	79	122
17	34	125	156	248	130	78	98
18	40.5	152	189	306	156	160	148
19	67.8	270	337	567	268	195	260
20	76	307	383	651	302	120	245



Şekil 3B : COCOMO, Barry Boehm Modelleri ve Uzman Tahmini Yöntemi Tahmin Değerlerinin Gerçek Efor (Adam-Gün) Karşılaştırma Grafiği

4) Parametrik Modeller İçinde En Uygun Modelin Belirlenmesi

Verilerin analizinde, iki değişik hata ölçütü kullanılarak parametrik efor tahmin yöntemleri sonuçları ile gerçek efor değerleri arasındaki farklar değerlendirilmektedir. Hata modelleri, Ortalama Karekök Hata (Root Mean Square Error -RMSE), ve Ortalama Yüzde Hata (Mean Magnitude of Relative Error - MMRE)'dir. Ortalama Kare Hata bir model tarafından tahmin edilen değerler arasındaki farklılıkları ölçmek için sıkça kullanılır. Ortalama Karekök Hata aşağıda sunulan denklemle gösterilmektedir [11].

Denklem (1)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_{tahmin} - E_{gerçek})^2}$$

Literatürde Ortalama Yüzde Hata modeli temel performans ölçütü olarak kabul edilir. Ortalama Yüzde Hata bağıl hataların mutlak değerinin yüzdesidir [12]. Şu şekilde yazılmaktadır:

Denklem (2)

$$MMRE = \frac{1}{N} \sum_{i=1}^N \frac{|E_{tahmin} - E_{gerçek}|}{E_{gerçek}}$$

Denklem (1) ve denklem (2) kullanılarak parametrik modellerin tahmin ettiği eforlar ile gerçek eforlar arasındaki hatalar hesaplanmıştır. Her bir model için bulunan hata değerleri Tablo 4'de sunulmaktadır.

Tablo 4 : Hesaplanan Eforların Hataları

		Hata Değerleri	
		MMRE	RMSE
Parametrik Modeller	Walston Felix	0.31	39.00
	Bailey Basili	0.70	82
	Halstead	4.29	1045.00
	Doty	0.57	393
	Barry Boehm	0.35	44.00
	COCOMO Organic	0.38	46
	COCOMO Semidetac	0.41	55.00
	COCOMO Embedded	0.73	135

Bu kapsamda MMRE ve RMSE olmak üzere iki farklı hata ölçütü kullanılmış, parametrik tahmin modellerinin hata değerleri bulunmuştur. Tablo 4'deki veriler incelendiğinde Walston Felix modelinin gerçekleşen efora en yakın tahmin ürettiği görülmektedir.

5) Yeni Model Önerilmesi

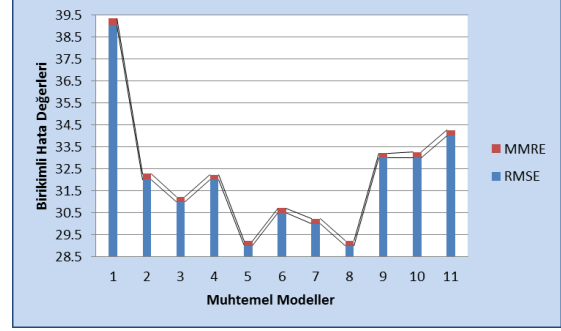
Daha doğru tahminleme çıktıları üreten ve hata oranları daha küçük çıkacak şekilde bir KLOC değişkenine bağlı parametrik model önerilebilir. Model üretilirken modelin çıktıları gerçekleşen veri kümesi ile karşılaştırılıp hata oranı en az olan nokta olan alan seçilmiştir.

$$EFOR = SD + \text{Çarpan} \times (KLOC)^{\text{Üstel Değer}}$$

Formülde SD : Sabit değer olarak alınır ve sabit değer, çarpan ve üstel değerlere değerler atanarak, iterasyon yöntemi ile tahminler üretilir. Üretilen tahminler gerçekleşen değerler ile karşılaştırılarak hata analizi değerleri çıkartılır. Hata analizi değerlerinin minimum olduğu nokta veri kümemiz için en uygun model olarak kabul edilmektedir. Tablo 5 ve Şekil 3'de görüldüğü üzere en uygun değerler 5.satırdaki değer olarak görünmektedir.

Tablo 5 : Değişen Sabit ve Çarpanlara Göre Hata Değerleri

S.N.	Sabit Değer	Çarpan	Üstel Değer	RMSE	MMRE
1	15	6.2	0.9	39	0.33
2	15	6.1	0.85	32	0.27
3	25	6.1	0.8	31	0.23
4	25	6	0.8	32	0.23
5	35	5.8	0.8	29	0.21
6	35	5.5	0.8	30.5	0.21
7	35	5.6	0.8	30	0.22
8	45	5.6	0.8	29	0.23
9	45	5.6	0.75	33	0.21
10	55	5.4	0.75	33	0.25
11	55	5.2	0.75	34	0.24



Şekil 4 : Önerilen Modelin RMSE ve MMRE Hata değerleri Birikimli Grafiği

En uygun hata değerlerine sahip sabit ve çarpanları formülde yerine koyarsak Önerilen Model :

$$EFOR = 35 + 5,8 \times (KLOC)^{0,8}$$

şeklinde olacaktır.

Burada 35 sabit değeri, kurumda 1 satır bile kod yazıldığında SDLC (Software Development Life Cycle) sürecinde geçireceği minimum zaman olması sebebiyle dikkat çekmektedir.

6) En Uygun Parametrik Model, Uzman Tahmin ve Önerilen Modellerin Karşılaştırılması

Çalışmada bu noktada parametrik modellerden en uygun olan model (Walston Felix), Uzman Tahmini ve önerilen model sonucu üretilen tahminlemeler bulunmaktadır. Tahminleme doğruluğu açısından bu üç tahminleme sonuçlarında MMRE, RMSE ve korelasyon katsayısı incelenmektedir. MMRE ve RMSE hata analiz denklemlerinden yukarıda bahsedilmişti.

Korelasyon katsayısı ise iki ya da daha çok değişken arasındaki ilişkiyi gösterir. Korelasyon çözümlenmesinin amacı, değişkenler arasındaki ilişkinin derecesini ve yönünü belirlemektir.

Denklem (3)

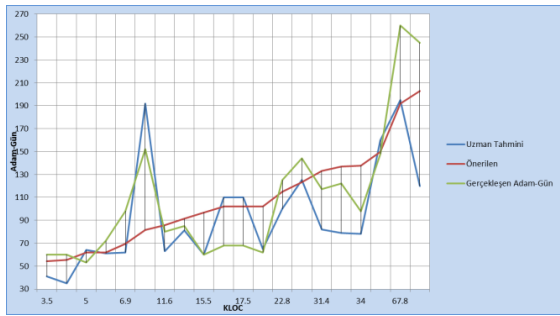
$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}}$$

Korelasyon Katsayısı r'nin Özellikleri

- 1) $-1 \leq r \leq 1$
- 2) Mükemmel pozitif doğrusal ilişki olduğunda $r = 1$ olur.
- 3) Mükemmel negatif doğrusal ilişki olduğunda $r = -1$ olur.
- 4) Doğrusal ilişki yok ise $r = 0$ olur.

Tablo 6 : Önerilen Model, Uzman Tahmini ve Uygun Model Veri Seti

Proje No	KLOC	Gerçekleşen	Önerilen Model	Uzman Tahmini	Walston Felix
1	3.5	60	50.80	41	16
2	3.7	60	51.52	35	17
3	5	53	56.02	64	22
4	5	72	56.02	61	22
5	6.9	98	62.20	62	30
6	10.3	152	72.47	192	43
7	11.6	80	76.21	63	48
8	13.5	85	81.53	81	56
9	15.5	60	86.96	60	63
10	17.5	68	92.26	110	70
11	17.5	68	92.26	110	70
12	17.5	62	92.26	65	70
13	22.8	125	105.76	100	89
14	26.6	144	115.04	125	103
15	31.4	117	126.41	82	120
16	33.4	122	131.03	79	127
17	34	98	132.41	78	129
18	40.5	148	147.04	160	151
19	67.8	260	204.20	195	241
20	76	245	220.39	120	268



Şekil 5 : Önerilen Model-Uzman Tahmini-Gerçekleşen Adanmış Gün Grafiği

Tablo 6'daki veri seti ile denlem(1) ve denlem(2)'de verilen formüllerin uygulanması sonucu hata analiz değerleri elde ederiz. Hata analiz değerleri Tablo 7'de verilmiştir.

Tablo 7 : Önerilen Model, Uzman Tahmini ve Uygun Model Hata Analizi Değerleri

		Hata Değerleri	
		MMRE	RMSE
Modeller	Önerilen Model	0.22	29.00
	Uzman Tahmini	0.26	40
	Walston Felix	0.31	39.00

Hata analizi değerlerinde önerilen modelin değerleri en küçük olduğu görülmektedir. Önerilen modelin ürettiği tahmin değerleri gerçekleşen değerlere en yakındır çıkarımını yapabiliriz.

Tablo 6'daki veri seti ile denlem(3)'de verilen formüllerin uygulanması sonucu korelasyon katsayısı değerini elde ederiz. Korelasyon katsayısı matrisi tablosu Tablo 8 de verilmiştir.

Tablo 8 : Önerilen Model, Uzman Tahmini ve Uygun Model Arasındaki Korelasyon Katsayıları

	KLOC	Gerçekleşen	Önerilen Model	Uzman Tahmini	Walston Felix
KLOC	1				
Gerçekleşen	0.88	1			
Önerilen Model	0.99	0.87	1		
Uzman Tahmini	0.58	0.75	0.59	1	
Walston Felix	0.99	0.87	0.99	0.58	1

Tabloda veriler incelendiğinde en kuvvetli ilişkilerin formülün doğası gereği KLOC-Önerilen Model-Walston Felix modellerinde olduğu görülmektedir. Çünkü modelde prosesi gerçekleştiren formülün, girdisinin (KLOC) çıktısı ile doğrudan ilişkisi olması kaçınılmazdır.

Bu tabloda önerilen model ile gerçekleşen arasındaki ilişki katsayısı (0,87) hem yüksek bir değer hem de uzman tahmini ile gerçekleşen arasındaki korelasyon katsayısından(0,75) büyüktür. Bu sonuç önerilen modelin bağımsız değişken olan gerçekleşen değer ile ilişkisinin kuvvetini, tahminlemede daha başarılı olduğunu göstermektedir.

Uzman tahmini ile KLOC arasındaki ilişkinin korelasyon katsayısı en küçük değerlerden biridir. Bu sonuçtan uzman tahmini yapan kişinin projeyi tahminlerken KLOC değerini baz almadığı veya doğru tahminleme yapamadığı çıkarımını yapabiliriz.

7) Sonuç

Bu çalışmada finans kurumlarında gerçekleştirilen yazılım projelerinin veri seti kullanılarak farklı parametrik modellerin efor tahmin başarısı karşılaştırılması yapılmıştır.

Bailey-Basili, Waltson-Felix, Halstead, Doty, Barry Boehm, COCOMO gibi tahmin model denklemleri yardımıyla aynı yazılım kod satır sayısı baz alınmasına karşın çok farklı efor değerleri tespit edilmiştir. Finans kurumları veri setinde bulunan 20 yazılım projesinin gerçek efor değerleri ile söz konusu modellerle tahmin edilen efor değerleri kıyaslanmıştır. Bu kıyaslamayla veri setindeki Finans kurumu yazılım projeleri için hangi efor tahmin modelinin en iyi sonuçlar verdiği tespit edilmeye çalışılmıştır. Parametrik tahmin modellerin hata değerleri hesaplanmış ve finans kurumu yazılım projeleri için en iyi sonuç veren model tespit edilmiş ve tahmin modelleri ile ilgili bir kıyaslama yapılmıştır.

Çalışmada minimum hata değerini veren, iterasyon yöntemi kullanılarak bir model bulunmuş ve model kullanılarak tahmin üretilmiştir. Bu modelde bulunan 35 sabit değerinin, kurumda 1 satır kod bile yazılsa minimum 35 Adam-Gün'de biteceği görülmektedir. Finans kurumunda uygulanan SDLC (Software Development Life Cycle) sürecinin uygulanmasından dolayı bu sabitin tam da gerçeği yansıtması bulunan önemli bir bulgudur.

Çalışmanın devamında uzman tahmini verileri de analize dahil edilerek, KLOC-Uzman Tahmini-Önerilen Model-En Uygun Parametrik Model'lerin korelasyon katsayıları ve hata değerleri hesaplanmıştır.

Çalışmanın başında yazıldığı üzere tahminleme adımlarından ilki KLOC tahmini yapmaktır. KLOC tahminin doğruluğu ve en uygun modelin kullanılması ile doğru tahminleme yapılabileceği ortaya çıkmaktadır. Korelasyon katsayısı tablosu analiz edildiğinde uzmanın doğru KLOC tahmini yapmadığı veya KLOC değerini baz almadığı görülmektedir. Sonuç olarak KLOC tahminini zorluğu ve KLOC ile gerçekleşen eforun ilişkili olmadığı sonucu ortaya çıktığı için finansal projelerde bu yöntemlerin uygulanabilir olmadığı sonucuna ulaşılmıştır.

Gelecekteki çalışmalarda, bu veri kümesi genişletilecek, bulanık mantık tahmin şemaları,

yapar sinir ağ tabanlı metodlar, karar ağacı yöntemi ve deneysel yöntemler kullanılacaktır. Bunlara ek olarak çalışmanın özgünlüğünü artırmak ve hata oranı daha düşük tahminler yapabilmek amacıyla üretilecek yazılımın analizine ve takım üyelerinin profiline göre hazırlanacak bir modele göre tahminler yapılacaktır.

Ayrıca yazılım projelerinde KLOC değerinin önceden kestirimi zor olduğu için problemin planlanmasında deterministik modelleri ele almak, elde edilen sonuçların uygulanabilirliği açısından zafiyet göstermektedir. Bu kapsamda hazırlanacak alternatif modelde, risk analizi yapabilmek adına Program Değerlendirme ve İnceleme Tekniği (Program Evaluation and Review Technique – PERT) yaklaşımına benzer bir şekilde en kötümser, en olurlu ve en iyimser değerler varsayarak öneride bulunmak amaçlanmaktadır.

KAYNAKLAR

- [1] S. Basha, P. Dhavachelvan, "Analysis of Empirical Software Effort Estimation Models", International Journal of Computer Science and Information Security, vol. 7, no. 3, pp. 68-72, December 2010.
- [2] J. Kaur, S. Singh, K.S. Kahlon, "Comparative Analysis of the Software Effort Models", Proceeding of World Academy of Science, Engineering and Technology, vol. 36, pp. 485-487, December 2008.
- [3] B.K. Singh, A.K. Misra, "Software Effort Estimation by Genetic Algorithm Tuned Parameters of Modified Constructive Cost Model for NASA Software Projects", International Journal of Computer Applications, vol.59, no. 9, pp. 22-26 December 2012.
- [4] L.M. Laird, M.C. Brennan, "Software Measurement and Estimation: A Practical Approach" pp. 103.[5] C. E. Walston, C. P. Felix, "A Method of Programming Measurement and Estimation," IBM Systems Journal, vol. 16, No. 1, pp. 54-73, 1977.
- [6] J. W. Bailey, V. R. Basili, "A meta Model for Software Development
- [7] Resource Expenditure", Proceedings of the International Conference on Software Engineering, pp. 107-115, 1981.

[8] H. Duggal, P. Singh, "Comparative Study of the Performance of M5-Rules Algorithm with Different Algorithms", *Journal of Software Engineering and Applications*, vol.5, pp. 270-276, 2012

[9] M. H. Halstead, "Elements of Software Science," Elsevier, New York, 1977.

[10] Doty Associates, Inc., "Software Cost Estimates Study," vol. 1, pp. 77-220, 1977.

[11] A. Sheta, D. Rine, A. Ayesh, "Development of Software Effort and Schedule Estimation Models Using Soft Computing Techniques",

IEEE Congress on Evolutionary Computation., pp. 1283-1289, 2008.

[12] O. Benediktsson, D. Dalcher, K. Reed, M. Woodman, "COCOMO based effort estimation for iterative and incremental software development", *Software Quality Journal*, vol. 11, pp. 265-281, 2003.