

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ BİLİŞİM ENSTİTÜSÜ

**DALGACIK DÖNÜŞÜMÜ KULLANILARAK ZİRAİ-METEOROLOJİK
VERİLERİN HATA TEŞHİS VE TAMİRİ**

YÜKSEK LİSANS TEZİ

Nigar Tuğbagül ALTAN

Bilgi Teknolojileri Anabilim Dalı

Bilgisayar Bilimleri Programı

AĞUSTOS 2012

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ BİLİŞİM ENSTİTÜSÜ

**DALGACIK DÖNÜŞÜMÜ KULLANILARAK ZİRAİ-METEOROLOJİK
VERİLERİN HATA TEŞHİS VE TAMİRİ**

YÜKSEK LİSANS TEZİ

**N. Tuğbagül ALTAN
(704091027)**

Bilgi Teknolojileri Anabilim Dalı

Bilgisayar Bilimleri Programı

Tez Danışmanı: Yrd. Doç. Dr. B. Berk ÜSTÜNDAĞ

AĞUSTOS 2012

İTÜ, Bilişim Enstitüsü'nün ... **704091027**... numaralı Yüksek Lisans Öğrencisi **N. Tuğbagül ALTAN**'ın, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı "**DALGACIK DÖNÜŞÜMÜ KULLANILARAK ZİRAİ-METEOROLOJİK VERİLERİN HATA TEŞHİS VE TAMİRİ**" başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Yrd Doç. Dr. B. Berk ÜSTÜNDAĞ**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Doç. Dr. Murvet KIRCI**
İstanbul Teknik Üniversitesi

Yrd. Doç Dr. Osman Kaan EROL
İstanbul Teknik Üniversitesi

Teslim Tarihi : **04 Mayıs 2012**
Savunma Tarihi : **17 Ağustos 2012**

Bilime ve İnanlığa,

ÖNSÖZ

Bu çalışmada bana yol gösteren, zamanını ayıran, her zaman pozitif enerjisi ve ilgisi ile bana çalışma enerjisi veren gerçek bir bilim adamı değerli hocam Yard. Doc. Dr. B. Berk ÜSTÜNDAĞ'a sonsuz teşekkürler.. Ayrıca çalışmalarına büyük bir katkısı olan sayın Prof. Dr. Ercan KAHYA hocama sonsuz teşekkürlerimle, çalışmamda bana yardımcı olan arkadaşım Serdar Bağış ve çalışmamın her evresinde bana destek olan değerli bir bilim ve düşünce insanı sevgili Kadir AKIN'a teşekkürlerimle.

Annem, babam, ablam ve kardeşime pazılın büyük parçaları oldukları için sonsuz teşekkürlerim ve sevgilerimle.

Ağustos 2012

N. Tuğbagül ALTAN
(Bilgisayar Mühendisi)

İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	vii
İÇİNDEKİLER	ix
KISALTMALAR	xi
ÇİZELGE LİSTESİ.....	xiii
ŞEKİL LİSTESİ.....	xv
ÖZET.....	xvii
SUMMARY	xxi
1. GİRİŞ	1
1.1 Matematiksel Modelleme	8
1.2 Daha Önceden Yapılan Çalışmalar	9
1.3 Tezin Amacı	12
2. ÇALIŞMADA KULLANILAN VERİLER.....	15
2.1 Çalışmada Kullanılan Verilerin Kullanım Amacı	15
2.1.1 Yersel akıllı istasyonlardan alınan meteorolojik veriler	15
3. ÇALIŞMA ALANININ COĞRAFİ KONUMU	17
4. YÖNTEM.....	19
4.1 Yersel Akıllı İstasyonlar Kullanılarak Eksik Veri Tamamlama Yöntemi	19
4.1.1 Regresyon analizi	19
4.1.2 En küçük kareler yöntemi	22
4.1.3 Kareler toplamı	24
4.2 Zaman Serisi.....	25
4.3 Matematiksel Dönüşüm.....	26
4.3.1 Fourier dönüşümü	26
4.3.2 Kısa süreli fourier dönüşümü (KSFD)	28
4.3.3 Dalgacık dönüşümü.....	28
4.3.3.1 Dalgacık kavramı.....	29
4.3.3.2 Sürekli dalgacık dönüşümü (SDD).....	30
4.3.3.3 Ayık dalgacık dönüşümü (ADD).....	31
5. YAPILAN ARAŞTIRMALAR VE SONUÇLARI.....	35
5.1 Birinci Durum	36
5.1.1 Birinci durum için belirlenen regresyon modeli	40
5.1.2 Birinci durum regresyon ve dalgacık dönüşümü modeli	42
5.2 İkinci Durum	45
5.2.1 Rumi istasyonu ve Aydoğdu istasyon verilerinin Beyazkule istasyonu verileri yerine kullanılması	45
5.2.2 İkinci durum için regresyon modeli	47
5.2.3 İkinci durum için regresyon ve dalgacık dönüşümü modeli	47
5.3 Üçüncü Durum	52
5.3.1 Akçatat istasyonu ve Poyralı istasyon verilerinin Yakacık istasyonu verileri yerine kullanılması	53
5.4 Dördüncü Durum.....	56
5.5 Beşinci Durum.....	57

5.6-Altıncı Durum.....	58
6. SONUÇ.....	61
KAYNAKLAR.....	65
ÖZGEÇMİŞ.....	69

KISALTMALAR

TARBİL	: Tarım Bilişim
DPT	: Devlet Planlama Teşkilatı
FD	: Fourier Dönüşümü
FFT	: Fast Fourier Transform
ADD	: Ayrık Dalgacık Dönüşümü
SDD	: Sürekli Dalgacık Dönüşümü
VERİS	: TARİT Verim İzleme Sistemi
PARİS	: TARİT Parsel izleme Sistemi
KURİS	: TARİT Kuraklık İzleme Sistemi
ZEUS	: TARİT Zirai Erken Uyarı Sistemi
UHUZAM	: İTÜ Uydu Haberleşmesi ve Uzaktan Algılama Merkezi
TAGEM	: Tarımsal Araştırmalar Genel Müdürlüğü
İTÜ	: İstanbul Teknik Üniversitesi
YSA	: Yapay Sinir Ağı
İBGYSA	: İleri Beslemeli Geriye Yayılmalı Yapay Sinir
RTYSA	: Radyal Tabanlı Yapay Sinir Ağları
GRYSA	: Genelleştirilmiş Regresyon Sinir Ağları
ÇDR	: Çok Değişkenli Regresyon Modeli

ÇİZELGE LİSTESİ

Sayfa

Çizelge 5.1 : 5 km çap içinde kalan istasyonlar.....	37
Çizelge 5.2 : İstasyon15 ile istasyon16, istasyon17 ve istasyon18 arasındaki korelasyon katsayıları.....	39
Çizelge 5.3 : İkinci durum için OKH ve OMH değerleri.....	51
Çizelge 5.4 : Akçatat ve Poyralı istasyonlarının Yakacık istasyonu ile olan korelasyon katsayıları.....	52
Çizelge 5.5 : Üçüncü durum için OKH ve OMH değerleri.....	56
Çizelge 5.6 : Dördüncü durum için OKH ve OMH değerleri.....	57
Çizelge 5.7 : Beşinci durum için OKH ve OMH değerleri.....	58
Çizelge 5.8 : Sonbahar, ilkbahar, yaz dönemlerinden alınan sıcaklık serilerinin kış dönemine ait sıcaklık serisi ile olan korelasyon katsayıları.....	59
Çizelge 5.9 : Altıncı durum için OKH ve OMH değerleri.....	60

ŞEKİL LİSTESİ

Sayfa

Şekil 1.1 : Kara-kutu modeli.....	9
Şekil 3.1 : Çalışma alanının coğrafi konumu.	18
Şekil 4.1 : Bir işaretin farklı frekanslardaki fourier dönüşümü.	27
Şekil 4.2 : Kısa süreli fourier dönüşümüne uğrayan bir sinyalin zaman- frekans ekseninde gösterimi.....	28
Şekil 4.3 : Daubechies dalgacığının 2-10 seviyelerinde gösterimi.....	30
Şekil 4.4 : Ayrık dalgacık dönüşümü ve çok çözünürlüklü ayrıştırma örneği.	33
Şekil 5.1 : Yakacık istasyonuna ait sıcaklık zaman grafiği.	37
Şekil 5.2 : Rumi istasyonuna ait sıcaklık zaman grafiği.....	38
Şekil 5.3 : Aydoğdu istasyonuna ait sıcaklık zaman grafiği.....	38
Şekil 5.4 : Demirören istasyonuna ait sıcaklık zaman grafiği.	38
Şekil 5.5 : İstasyonların konumlarının google earth kullanılarak gösterilmesi.	40
Şekil 5.6 : İstasyonların konumlarının tarit istasyonlar haritası kullanılarak gösterilmesi.	40
Şekil 5.7 : Eğitim kümesinin gerçek sıcaklık değerleri ve modelleme sonucu bulunmuş sıcaklık değerlerine ait grafik.	42
Şekil 5.8 : Test kümesinin gerçek sıcaklık değerleri ve modelleme sonucu bulunmuş sıcaklık değerlerine ait grafik.	42
Şekil 5.9 : Eğitim kümesi üzerinde bulunan ve denenen yaklaşık, detay ve orjinal sinyallere ait regresyon+dalgacık dönüşümü modeline ilişkin grafikler.	44
Şekil 5.10 : Eğitim kümesi üzerinde bulunan ve test kümesi üzerinde denenen yaklaşık, detay ve orjinal sinyallere ait regresyon+dalgacık dönüşümü modeline ilişkin grafikler.	44
Şekil 5.11 : Beyazkule ve Rumi istasyonlarının aynı zamana ait sıcaklık serileri. ...	46
Şekil 5.12 : Beyazkule ve Rumi istasyonlarının aynı zamana ait sıcaklık serileri. ...	46
Şekil 5.13 : Regresyon modeli için eğitim kümesinde bulunan değerlerin test kümesinde test edilmesine ilişkin grafik.	47
Şekil 5.14 : Eğitim kümesi üzerinde bulunan ve denenen yaklaşık, detay, orjinal sinyallere ait regresyon+dalgacık dönüşümü modeline ilişkin grafikler, (a)A6 ana bileşen, (b) D1, (c) D2, (d) D3, (e) D4, (f)D5, (g)D6, detay bileşenler (h) Gözlenlenmiş ve hesaplanmış sıcaklık serileri.	50
Şekil 5.15 : Regresyon+ dalgacık dönüşümü modeli için eğitim kümesinde bulunan değerlerin test kümesinde test edilmesine ilişkin grafik.	51
Şekil 5.16 : Yakacık, Akçatat ve Poyralı istasyonları arasındaki mesafeler (km).	53
Şekil 5.17 : Yakacık ve Akçatat istasyonlarının aynı zamana ait sıcaklık serileri. ...	54
Şekil 5.18 : Yakacık ve Akçatat istasyonlarının aynı zamana ait sıcaklık serileri. ...	54
Şekil 5.19 : Regresyon modeli için eğitim kümesinde bulunan değerlerin test kümesinde test edilmesine ilişkin grafik.	55
Şekil 5.20 : Regresyon+ dalgacık dönüşümü modeli için eğitim kümesinde bulunan değerlerin test kümesinde test edilmesine ilişkin grafik.	55

- Şekil 5.21** : Altıncı durum için doğrusal regresyon modeli kullanılarak hesaplanan ve bulunan sıcaklık serisine ait grafik. 59
- Şekil 5.22** : Altıncı durum için doğrusal regresyon model+ dalgacık dönüşümü modeli kullanılarak hesaplanan ve bulunan sıcaklık serisine ait grafik. 60

DALGACIK DÖNÜŞÜMÜ KULLANILARAK ZİRAİ-METEOROLOJİK VERİLERİN HATA TEŞHİS VE TAMİRİ

ÖZET

Toplumların gelişim süreçlerinde yüzyıllar boyu çok önemli yer tutan bir üretim alanı olan tarım, hemen hemen tüm ülkelerin toplumsal üretimlerinin en temel yanını oluşturur. Bu, yalnızca insanlar için yaşamsal önemi olan besin, giyim v.b. maddelerin üretildiği sektör olmasından değil, nüfusun büyük bölümünün yaşadığı ve geçimini sağladığı alan olmasından dolayı da önemlidir. Kırsal alanda toprak üzerinde oluşan üretim ilişkileri tarımın yapısını belirler. Ülkedeki ekonominin yapısı, gelişen koşul ve ilişkilere uygun olarak, bilim ve teknolojik gelişmelerin sunduğu yeni olanaklar çerçevesinde dönüşümlere uğrarken, tarımın, bu gelişmelerin dışında kalması düşünülemez [1]. Gerek sulama faaliyetleri, gerek tarlaların işlenmesi gibi çok çeşitli alanlarda teknolojik gelişimlerden faydalanmak, tarımda verimi etkileyen etkenlerin başında gelir. Tarımı etkileyen en önemli etkenlerden birisi ise, sulamadır. Ülkemizde toplam tarım alanının yaklaşık 1/3'ü ekonomik olarak sulanabilecek potansiyele sahiptir. Resmi kayıtlar bunun şu anda ancak yarısının sulamaya açıldığını ve sulama oranlarında dikkate alındığında sulamaya açılan alanların yaklaşık %65'nin fiilen sulandığını göstermektedir. İzlenmesi gereken sulama politikalarının başında tarımda verimliliğin artırılması için sulanabilecek alanların tamamının sulanması ve suyun etkin kullanımının sağlanması gelir. Bilindiği gibi ülkemizde suyun yaklaşık %70'i tarımsal amaçla tüketilmektedir. Önümüzdeki yıllarda sanayi ve hizmetler sektörünün artan su ihtiyacının karşılanabilmesi su yönetiminin önemini daha da arttırmıştır. Su yönetiminde karşılaşılan sorunlar fiziksel altyapı yetersizliği, su kalitesinin bozulması ve su kirliliği, su iletim ve dağıtım sistemlerinin yetersizliği, aşırı su kayıpları ve organizasyon ve yönetim sorunları olarak sıralanabilir. Bu amaçla sulamada atık suların kullanımı, yüzey sularının suyun kıt olduğu alanlara yönlendirilmesi, çiftçilerin su tasarrufu sağlayan modern sulama yöntemlerini uygulamalarını teşvik etmek ve çiftçi eğitimleriyle bunun yaygınlaştırılmasını sağlamak gibi çalışmalarını kapsayan politikalara önem verilmelidir [2]. Sulama amaçlı projelerin ülkenin yalnız tarımında değil kırsal alanların, sosyal ve ekonomik gelişmesindeki katkısı nedeniyle de, Türkiye'nin ekonomisinde önemli bir yeri vardır. Hatta sulama yatırımları dünyada ve ülkemizde kırsal kalkınmanın en temel araçları olarak görülmektedir. Tarımsal üretimin ve verimin artırılması, sağladığı net gelir artışı, sürdürülebilir tarıma olanak vermesi, ek istihdam yaratması, sosyal yapıda değişime neden olması gibi sayılabilecek birçok yararları nedeniyle sadece kırsal kalkınmanın değil ekonomik gelişme ve kalkınmanın en temel ve en uygun/etkin araçları olarak görülmektedir. Bu işlevleri nedeniyle de sürekli ülkeyi yönetenlerin, politikacıların hep ilgi odağı olmuştur [3]. Öyle ki bir toplumun en temel ihtiyacı olan besin ve yiyecek maddesinin en iyi şekilde temin edilmesi ülke ekonomisi için oldukça önemlidir. Ayrıca bu ürünlerin sağlanması ülke gelişmişliği açısından ilk sıralarda gözlemlenen değerlerdir. Bir ülkenin iklimi ve dolayısıyla su potansiyeli, bulunduğu

yerin enlemine, boylamına, jeolojik, topoğrafik yapısına ve bitki örtüsüne göre oluşmaktadır. Aynı şekilde ülkemizin jeolojik, topoğrafik ve iklim (yağış) özelliklerinin yöre ve bölgelere göre farklı olması nedeniyle yeraltı ve yüzey suyu potansiyelinde farklılıklar görülmektedir (Bilgin, 1997:18; Özgüler,1997:59). Buna ek olarak, tarım, başta iklim ve toprak olmak üzere coğrafi faktörlerin etkisi altındadır (Tanoglu, 1968). Ülkemizde en yüksek sıcaklıklara Güneydoğu Anadolu'da rastlanır. Şanlıurfa, Ceylanpınar, Gaziantep en sıcak sahalarımızı oluşturmaktadır. Bu sahalarda nisan ayında başlayarak Temmuz ve Ağustos aylarında en yüksek seviyeye çıkmaktadır. Şanlıurfa ve çevresinde en uzun devre olarak beliren kuraklık, ekim ayına kadar şiddetli kurak devre olarak devam etmektedir (Tümertekin, 1957). Sulama faaliyeti, iklim ve diğer etkenler gözönünde bulundurularak, ileride geliştirilmesi planlanan tarla sulama çiftçi uyarı sistemi, modern sulama aracı olarak düşünülebilir. Tasarlanması planlanan bu sistemin doğru bir şekilde çalıştırılabilmesi diğer bir ifade ile, tarla sulama zamanının doğru bir şekilde tespit edilmesi gibi hesaplamalar, birçok etkene bağlıdır. Bu etkenlerden en önemlisi şüphesiz meteorolojik olaylardır. Doğru hava tahminlerinin yapılması, tasarlanması planlanan sulama sisteminin de doğru çıktılar vermesi anlamına gelmektedir. Doğru çıktılar kullanılarak yapılan tarım, iklime bağlı olmaktan büyük oranda azalır. Bir tarladan yılda birden fazla ürün alma olasılığı artar. Böylece hem yetiştirilen ürün miktarındaki verimlilikten hem de tarlanın yılda birden fazla kullanılmasından dolayı üretim artışı gözlenir. Üretim miktarındaki dalgalanmalar azalır. Ülkemizde üretilen senelik tarımsal ürün miktarının hesaplanması daha doğru sonuçlara ulaşır. Tarımsal fiyat politikaları, gerçekleşen ve beklenen değerlere göre en doğru biçimde dinamik olarak sağlanır. Bu durum, ülke tarım bütçesinin kararlaştırılmasında büyük bir öneme sahiptir. Bu çalışma, tarımsal desteklerin optimizasyonunda, tarımsal politikaların yönlendirilmesinde, gerçek zamanlı nitelikli veri oluşturulması, hangi ürünün ne kadar ekilmesi konusunda yani arz/talep dengesinin kararlı tutulması için beklenen ve gerçekleşen rekolteye göre coğrafik olarak ekim önceliklerinin belirlenebilmesi, böylelikle talep fazlası ya da arz fazlası oluşumlarının engellenmesi için geliştirilmesi planlanan sistemin doğru şekilde çalışması için bir çözüm önerisinde bulunmaktadır. Diğer bir ifade ile, bu çalışma, her bitki türü için üretim yeri de gözönünde bulundurularak, en ideal sulama zamanları ve en ideal sulama miktarının belirlenmesi için yapılacak çalışmalar için gerekli ön çalışma özelliği göstermektedir. Tarımsal faaliyetlerin izlenebilmesi için bir çok modeller geliştirilmiş ve geliştirilmektedir. Zirai-meteorolojik verilerin doğru tahmini, bitki gelişim modellerininin doğru sonuç verebilmesi için oldukça önem arz etmektedir. Doğru zamanda her bitki türü için doğru sulama sistemi, ürün verimliliğini artırdığı gibi, emek ve su israfını da büyük bir oranda azaltmaktadır. Sulama faaliyetini etkileyen meteorolojik olayların tahmini, bu sorunu büyük bir olasılıkla giderebilecektir. Dalgacık Dönüşümü Kullanarak Zirai-Meteorolojik Verilerin Hata Teşhis ve Tamiri çalışması ile hedeflenen, var olan sistem ile herhangi bir nedenden ötürü, doğru sağlanamayan meteorolojik verileri tespit etmek ve meteorolojik veri sağlanan istasyonlardan olası arıza durumunda alınamayan verileri en doğru şekilde tahmin etmektir. Böylece meteorolojik verilere dayalı çalışan sistemlerin en doğru sonuçları üretmeleri sağlanarak, tarımsal ürün verimliliği artırılarak, su ve emek istafi önlenmiş olacaktır. Bu amaç ile gerekli alt yapı sistemleri ve sistem girdi parametrelerinin kararlaştırılması için araştırılan çeşitli algoritmalar üzerinde durulacak, tasarlanan sistemler arasından karşılaştırma yapılarak, yapılan karşılaştırma sonucunda modellenen sistemler arasından optimum sistem önerisinde bulunulacaktır. Bu çalışmada teorik ve yapılan çeşitli veri

kümeleri üzerinde denenmiş deneysel sonuçların karşılaştırılmalı olarak çözümlenmeleri yer almıştır. Ayrıca konu ile ilgili geçmiş dönemlerde yapılmış olan çalışmalar hakkında da bilgiye yer verilmiştir. Yapılan çalışma uzaktan algılama teknikleri kullanılarak hayata geçirilen TARBİL projesinin altında gerçekleştirilmiştir.

Dalgacık Dönüşümü Kullanarak Zirai-Meteorolojik Verilerin Hata Teşhis ve Tamiri çalışmasının kapsamında gerçekleştirilecek olan eksik verilerin tamamlanması işlemi için dalgacık dönüşümü tekniği kullanılarak, regresyon analizi ile modellemeler yapıp sonuçları karşılaştırmalı olarak açıklanmıştır. Meteorolojik verilerin temin edilmesinde gerekli desteği veren TARBİL projesi ile ilgili daha fazla bilgi edinmek için, projesinin resmi sitesi olan www.tarit.org.tr adresi ziyaret edilebilir.

Dalgacık Dönüşümü Kullanılarak Zira-Meteorolojik Verilen Hata Teşhis ve Tamiri çalışmasının, gerçekleştirilmesi için, tasarlanan sistem için sistem girdilerinin kararlaştırılmasına ihtiyaç vardır. Bu sebep ile TARBİL kapsamında, belirli bölgelerde kurulan, on dakikada bir meteorolojik veri alımı sağlayan akıllı yersel istasyonlardan faydalanılmış ve bu istasyonlardan, tasarlanan model için gerekli veri kümesi temin edilmiştir.

Bu çalışmada, zirai meteorolojik verilerden olan sıcaklık ve nem için, eksik veri tamamlama işlemleri yapılmış, kayıp olan verinin zamanı, kendinden önceki değerleri, nem ve benzeri değişkenler ile olan ilişkisinin ne derecede olduğu saptanmaya çalışılmıştır. Bu sebep ile, belirlenen değişkenler ile sıcaklığın ilişkisini tespit etmek amaçlı korelasyon analizi gerçekleştirilmiştir. Korelasyon sonucu çıkan güçlü, zayıf ilişkili parametreler sınıflandırılmıştır. Sıcaklık ve nem ile ilişkisi güçlü bulunan meteorolojik veriler kullanılarak, eksik sıcaklık değerlerinin tespitine yönelik tasarımlar oluşturulmaya çalışılmıştır. Bu nedenle, tasarlanan sistem için girdi olmasına karar verilen veriler ile ilk önce regresyon analizi yapılmış, ardından bu veriler dalgacık bileşenlerine ayrılmış elde edilen katsayılar ile tekrar regresyon analizi yapılmış elde edilen sonuçlar hem grafik üzerinde hem de matematiksel yöntemler ile karşılaştırılmaya çalışılmıştır. Dalgacık dönüşümü tekniği kullanılarak, ilgili meteorolojik veri serisinin tahmini için regresyon yöntemine göre gerçeğe en yakın seri elde edilmiştir. Modellemede kullanılan verilerin temininde, eksik veya hatalı verisi bulunan istasyonun verisini bulmak için komşu istasyonlardan faydalanılmıştır.

Ayrıca tasarlanan modeller ile, zirai-meteorolojik verilerin tahmin başarımını tespit etmek için, bulunması amaçlanan veriler için, hesaplanan değerler ile verilerin gerçek değerleri arasındaki hata oranı bulunmuştur. Buna ek olarak korelasyonu yüksek çıkan iki istasyonun verileri ayrı ayrı, eksik verileri bulunan diğer istasyonun eksik verileri yerine koyulmuş, böylece gerçek değerlerin serisi ile diğer istasyonların değerleri kullanılarak oluşturulan seriler arasındaki hata oranları ayrı ayrı tekrar hesaplanmıştır. Sıcaklık verileri üzerinde test edilen modellerden elde edilen sonuçlar, açıklamalı olarak tekrar karşılaştırılmıştır. Yapılan araştırmalar, geliştirilen modelin karşılaştırılan diğer model ile elde edilen hata oranları sonuçları doğrultusunda, mantıklı ve kullanışlı bir model olduğunu göstermiştir.

FAULT DIAGNOSIS AND REPAIR OF AGRICULTURAL METEOROLOGICAL DATA USING WAVELET TRANSFORM

SUMMARY

Agriculture has a great importance in our country as the world. Cultivated plant growth and development depends on several factors. There is no doubt that the most important factors are meteorological factors for agricultural systems. In addition, one of the biggest problems in our country is related with agricultural irrigation. Prediction of meteorological phenomena that affects the activity, probably decreases the problems. For each plant species, irrigation at the right time increases the efficiency of the irrigation system, on the other side, decreases the wastage of water and labor.

A range of modelling is being carried out with an aim to improve efficiency of the agricultural products. As a result of a number of reasons, there are problems involved in obtaining meteorological data series and deficiencies occur in these meteorological data series. In this study, an approach has been put forward for this problem by means of using the method of wavelet transform for the reconstruction of the missing data related to the data series, which is a type of meteorological data. Based on the observations, successful results have been obtained by the method developed for the estimation of the missing data in the data series. Previous studies were carried out focusing on the use of the wavelet transform method in meteorological data estimation. One of the most significant studies on the use of the wavelet transform method in relation to meteorological data was carried out in 1990s by Kumar and F. Georgiou. In the study, precipitation data was analyzed and these precipitation data was decomposed into its components by means of using multi resolution method. Then, the satellite pictures obtained were interpreted for precipitation condition.

In another study, the modelling of the hydraulic current series was intended to be achieved using wavelet transform method. In the study, hydraulic current series were intended to be estimated by means of using wavelet transform method. Stream-flow series of two independent measuring stations were separated into their components by means of applying discrete wavelet transform. Modelling was based on a model of regression type developed by these components. As a result of modelling, it was found out that the models established by means of using the components selected in line with the climatic characteristics of each different regions included in this framework were more successful with regard to a number of error criteria.

This study is one of the first of its kind regarding the applications of the wavelet transform method for the reconstruction of the missing data to the data series. The aim of the Fault Diagnosis And Repair of Agricultural-Meteorological Data Using Wavelet Transform study is to increase the efficiency of products, prevent the water and human labor waste, detect and correct errors, fill the gap or reconstruct the data in the measured meteorological data class provided by meteorological stations. To estimate the most accurate data is the main purpose of this study. With this purpose, this study focuses on various algorithms to decide the necessary infrastructure

systems, systems parameters and inputs. Optimum system design is selected comparatively. Theoretical and comparative analysis of experimental results are made. The results are tested on a variety of data sets. Also information about the past to the present related studies is given with their success.

In this study which carried out under Project TARBIL, necessary information about TARBIL will be given. The important parts of the TARBIL Project scope, method and aim will be highlighted. For the completion of missing data process which is the subtitle of “Fault Diagnosis And Repair Of Agricultural-Meteorological Data Using Wavelet Transform” by using the technique of wavelet transform, regression analysis, statistical methods will be performed and the results will be described in comparatively.

The data is supplied by TARBIL. The stations are located in the field and allows for the reception of the temperature and other meteorological data. For more information about the TARBIL project, the official web site of the project www.tarit.org.tr which supports to provide the data, can be visited.

Several modelling are developed for monitoring of agricultural activities. Models of agricultural production systems are developed to search answers to various problems serve purposes such as; analysing of events that occur during plant development (irrigation, soil and air temperature changes, drought, reduced soil moisture content, etc.), crop yield forecasting, determining the effect of soil, plant and meteorological factors to plant growth, training work on reactions in the plant system which is extremely complex, taking the agricultural policy decisions (Ritchie and ark. 1998). It is very important to make accurate estimation of agro-meteorological data for correct modelling of plant development. In order to make progress in agricultural areas not being away from technological developments is inevitable. One of the biggest agricultural problem in our countries, irrigation problem as we know. When the irrigation problem solved the production increases. Agriculture dependance on climate greatly reduced. Having more than one product in a year is increased for a farm. Fluctuations in the amount of production is reduced. Calculating the amount of annual agricultural product produced in our country reach more accurate conclusions. Agricultural price policies, realized and expected values provided the most accurate way by dynamically. This situation has a great importance in deciding the agricultural budget of the country.

In this research, wavelet is used because of wavelet being the most powerful tool. It is especially suitable for non-stational signal processing. In 1990s, it has been successfully used in signal processing, such as image coding, compressing and edge detection, mechanical fault diagnosis. In this study, wavelet is used for non-stational signal processing.

In this thesis, the method for reconstruction of temperature data is explained. We developed a new model that is related with the regression and wavelet transform techniques. Firstly, we used pure regression model than regression and wavelet model was used to reconstruct the missing meteorological data. In this study, the difference known as mean square error and mean absolute error are founded between the measured and calculated data patterns of the station. This station’s missing data series that are calculated with using neighboring stations’ data series. Also we compared the two model’s result. The regression and wavelet model shows more feasible error rate than the pure regression model.

In the study, wavelet decomposition approach is used, Daubechies wavelet is preferred because of its orthonormal property. The experiments shows that the wavelet transform with the different scaling factors of the time series is a reasonable approach that can be used for reconstruction of the missing temperature or humidity data. As a result, this research brings us the new useful approach to the missing data problem.

1. GİRİŞ

Toplumların gelişim süreçlerinde yüzyıllar boyu çok önemli yer tutan bir üretim alanı olan tarım, hemen hemen tüm ülkelerin toplumsal üretimlerinin en temel yanını oluşturur. Bu, yalnızca insanlar için yaşamsal önemi olan besin, giyim v.b. maddelerin üretildiği sektör olmasından değil, nüfusun büyük bölümünün yaşadığı ve geçimini sağladığı alan olmasından dolayı da önemlidir. Kırsal alanda toprak üzerinde oluşan üretim ilişkileri tarımın yapısını belirler. Ülkedeki ekonominin yapısı, gelişen koşul ve ilişkilere uygun olarak, bilim ve teknolojik gelişmelerin sunduğu yeni olanaklar çerçevesinde dönüşümlere uğrarken, tarımın, bu gelişmelerin dışında kalması düşünülemez [1]. Gerek sulama faaliyetleri, gerek tarlaların işlenmesi gibi çok çeşitli alanlarda teknolojik gelişimlerden faydalanmak, tarımda verimliliği arttırmaktadır. Tarımı etkileyen en önemli etkenlerden birisi , sulamadır. Ülkemizde toplam tarım alanının yaklaşık 1/3'ü ekonomik olarak sulanabilecek potansiyele sahiptir. Resmi kayıtlar bunun şu anda ancak yarısının sulamaya açıldığını ve sulama oranlarında dikkate alındığında sulamaya açılan alanların yaklaşık %65'nin fiilen sulandığını göstermektedir. İzlenmesi gereken sulama politikalarının başında tarımda verimliliğin artırılması için sulanabilecek alanların tamamının sulanması ve suyun etkin kullanımının sağlanmasıdır. Bilindiği gibi ülkemizde suyun yaklaşık %70'i tarımsal amaçla tüketilmektedir. Önümüzdeki yıllarda sanayi ve hizmetler sektörünün artan su ihtiyacının karşılanabilmesi su yönetiminin önemini daha da arttırmıştır. Su yönetiminde karşılaşılan sorunlar fiziksel altyapı yetersizliği, su kalitesinin bozulması ve su kirliliği, su iletim ve dağıtım sistemlerinin yetersizliği, aşırı su kayıpları ve organizasyon ve yönetim sorunları olarak sıralanabilir. Bu amaçla sulamada atık suların kullanımı, yüzey sularının suyun kıt olduğu alanlara yönlendirilmesi, çiftçilerin su tasarrufu sağlayan modern sulama yöntemlerini uygulamalarını teşvik etmek ve çiftçi eğitimleriyle bunun yaygınlaştırılmasını sağlamak gibi çalışmalarını kapsayan politikalara önem verilmelidir [2]. Sulama amaçlı projelerin ülkenin yalnız tarımında değil kırsal alanların, sosyal ve ekonomik gelişmesindeki katkısı nedeniyle de, Türkiye'nin ekonomisinde önemli bir yeri vardır. Hatta sulama yatırımları dünyada ve ülkemizde

kırsal kalkınmanın en temel araçları olarak görülmektedir. Tarımsal üretimin ve verimin artırılması, sağladığı net gelir artışı, sürdürülebilir tarıma olanak vermesi, ek istihdam yaratması, sosyal yapıda değişime neden olması gibi sayılabilecek birçok yararları nedeniyle sadece kırsal kalkınmanın değil ekonomik gelişme ve kalkınmanın en temel ve en uygun / etkin araçları olarak görülmektedir. Bu işlevleri nedeniyle de sürekli ülkeyi yönetenlerin, politikacıların hep ilgi odağı olmuştur [3]. Öyle ki bir toplumun en temel ihtiyacı olan besin ve yiyecek maddesinin en iyi şekilde temin edilmesi, ülke ekonomisi için oldukça önemlidir. Ayrıca bu ürünlerin sağlanması ülke gelişmişliği açısından ilk sıralarda gözlemlenen değerlerdir. Bir ülkenin iklimi ve dolayısıyla su potansiyeli, bulunduğu yerin enlemine, boylamına, jeolojik, topoğrafik yapısına ve bitki örtüsüne göre oluşmaktadır. Aynı şekilde ülkemizin jeolojik, topoğrafik ve iklim (yağış) özelliklerinin yöre ve bölgelere göre farklı olması nedeniyle yeraltı ve yüzey suyu potansiyelinde farklılıklar görülmektedir (Bilgin, 1997:18; Özgüler,1997:59). Tarımsal alanda yaşanan başlıca sorunlardan en önemlilerinden birisi sulama sorunudur. Ayrıca bu sorun, ülkemizde tarım ile ilgili yaşanan sorunların en önemlilerinden biridir. Tarımsal faaliyetlerin izlenebilmesi için bir çok modellemeler geliştirilmiş ve geliştirilmektedir. Tarımsal ürün yetiştiricilikteki olası birçok soruna yanıt aramak için oluşturulan modeller; bitki gelişimi sırasında meydana gelen olayların analizi (sulama, hava ve toprak sıcaklığındaki değişimler, kuraklık, toprak nem içeriğindeki azalma vb), bitki veriminin tahmini, toprak, bitki ve meteorolojik faktörlerin bitki gelişimine olan etkisinin belirlenmesi, son derece karmaşık olan bitki sistemi ve reaksiyonları ile ilgili eğitim çalışmaları, tarımsal politik kararların alınması gibi amaçlara hizmet eder (Ritchie ve ark. 1998). Zirai-Meteorolojik verilerin doğru tahmini, bitki gelişim modellemelerinin doğru sonuç verebilmesi için oldukça önem arz etmektedir. Doğru zamanda her bitki türü için doğru sulama sistemi, ürün verimliliğini artırdığı gibi, emek ve su israfını da büyük bir oranda azaltmaktadır. Sulama faaliyetini etkileyen meteorolojik olayların tahmini, bu sorunu büyük bir olasılıkla giderebilecektir. Sulama faaliyetleri, iklim ve diğer etkenler gözönünde bulundurularak, ileride geliştirilmesi planlanan tarla sulama çiftçi uyarı sistemi, modern sulama aracı olarak düşünülebilir. Tasarlanması planlanan bu sisteminin, doğru bir şekilde çalıştırılabilmesi diğer bir ifade ile, tarla sulama zamanının doğru bir şekilde tespit edilmesi gibi hesaplamalar, birçok etkene bağlıdır. Bu etkenlerden en önemlisi şüphesiz meteorolojik olaylardır. Doğru hava tahminlerinin yapılması tasarlanması

planlanan sulama sisteminde doğru çıktılar vermesi anlamına gelmektedir. Doğru çıktılar kullanılarak yapılan tarım, iklime bağlı olmaktan büyük oranda çıkar. Bir tarladan yılda birden fazla ürün alma olasılığı artar. Böylece hem yetiştirilen ürün miktarındaki verimlilikten hem de tarlanın yılda birden fazla kullanılmasından dolayı üretim artışı gözlenir. Üretim miktarındaki dalgalanmalar azalır. Ülkemizde üretilen senelik tarımsal ürün miktarının hesaplanması daha doğru sonuçlara ulaşır. Tarımsal fiyat politikaları, gerçekleşen ve beklenen değerlere göre en doğru biçimde dinamik olarak sağlanır. Bu durum, ülke tarım bütçesinin kararlaştırılmasında büyük bir öneme sahiptir. Bu çalışma, tarımsal desteklerin optimizasyonunda, tarımsal politikaların yönlendirilmesinde, gerçek zamanlı nitelikli veri oluşturulması, hangi ürünün ne kadar ekilmesi konusunda yani arz/talep dengesinin kararlı tutulması için beklenen ve gerçekleşen rekolteye göre coğrafik olarak ekim önceliklerinin belirlenebilmesi, böylelikle talep fazlası ya da arz fazlası oluşumlarının engellenmesi için geliştirilmesi planlanan sistemin doğru şekilde çalışması için bir çözüm önerisinde bulunmaktadır. Diğer bir ifade ile, bu çalışma, her bitki türü için üretim yeri de gözönünde bulundurularak, en ideal sulama zamanları ve en ideal sulama miktarının belirlenmesi için yapılacak çalışmalar için gerekli ön çalışma özelliği göstermektedir. Dalgacık Dönüşümü Kullanarak Zirai-Meteorolojik Verilerin Hata Teşhis ve Tamiri çalışması ile hedeflenen, var olan sistem ile herhangi bir nedenden ötürü, doğru ölçülemeyen meteorolojik verileri tespit etmek ve meteorolojik veri sağlanan istasyonlardan olası arıza durumunda alınamayan yada iletilemeyen verileri en doğru şekilde tahmin etmektir. Böylece meteorolojik verilere dayalı çalışan sistemlerin en doğru sonuçları üretmeleri sağlanarak, tarımsal ürün verimliliği arttırılarak, su ve emek istafi önlenmiş olacaktır. Bu amaç ile gerekli alt yapı sistemleri ve sistem girdi parametrelerinin kararlaştırılması için araştırılan çeşitli algoritmalar üzerinde durulacak, tasarlanan sistemler arasında karşılaştırma yapılarak, yapılan karşılaştırma sonucunda modellenen sistemler arasında optimum sistem önerisinde bulunulacaktır.

Bu çalışmada teorik ve yapılan çeşitli veri kümeleri üzerinde denenmiş deneysel sonuçların karşılaştırılmalı olarak çözümlemeleri yer almıştır. Ayrıca konu ile ilgili geçmiş dönemlerde yapılmış olan çalışmalar hakkında da bilgiye yer verilmiştir.

Ülkemizde tarımsal faaliyetlerdeki gelişim yöntemleri, tam anlamıyla istenilen seviyeye ulaşmamış, bilimsel temellere tam anlamıyla otutturulamamıştır.

TARBİL (Tarımsal Rekolte Tahmin ve Kuraklık İzleme Projesi) altında yürülen bu çalışmada, TARBİL projesi kapsamında uzaktan algılama yöntemleri ile izlenen, her birinin bulunduğu alandaki meteorolojik verilerin de alınmasına olanak sağlayan Türkiye'nin çeşitli coğrafik bölgelerinde bulunan birçok farklı yerde kurulan akıllı yersel istasyonlardan faydalanılarak, yapılan çalışma için gerekli olan sistem girdileri sağlanmış ve tarımda verimliliği etkileyen faktörlerden olan sıcaklık, nem gibi meteorolojik olayların kullanılması ile geçmişe dönük eksik datalarının tahmin edilmesi ve gelecek zamanlara ait hava olaylarının da yine tahmin edilmesi amaçlanmıştır. Dalgacık Dönüşümü Kullanarak Zirai-Meteorolojik Verilerin Hata Teşhis ve Tamiri çalışmasının kapsamında olan eksik verilerin tamamlanması işlemi için dalgacık dönüşümü tekniği kullanılarak, regresyon analizi gibi istatistiksel yöntemler ile modellemeler yapıp, elde edilen sonuçlar karşılaştırmalı olarak açıklanmıştır. Verilerin sağlanmasında destek veren TARBİL projesi ile ilgili daha fazla bilgi edinmek için, TARBİL projesinin resmi sitesi www.tarit.org.tr adresi ziyaret edilebilir.

Çalışmaların üst başlığı olan TARBİL tarımsal rekoltenin ürünler bazında ölçümlendirilmesini ve gelecek hasat rekoltesinin tahmin edilmesini sağlayacak geniş kapsamlı bir projedir. TARBİL projesi ile, Türkiye'nin çeşitli bölgelerindeki tarlalar, uzaktan algılama yöntemleri kullanılarak izlenmektedir. Geniş kapsamlı TARBİL projesi, bir çok altbaşlıktan oluşmaktadır: Uydudan çok spektrumlu görüntülerin elde edilmesi, bu projenin alt başlıklarından biridir. TARBİL Projesinin, uydudan tarlaların görüntülerinin alınmasına benzer çalışmalar daha önceden yapılmıştır ancak TARBİL projesinde uydu görüntülerinin elde edilmesinde kullanılan bant genişlikleri ve teknik yapılar benzer çalışmalar arasında önemli ölçüde fark oluşturmaktadır.

Tarımdaki teknolojik gelişmeler ile ilgili olarak, günümüzde bitkisel ekim alanlarının üretim ve verimlerinin belirlenmesinde bilgisayar destekli Coğrafi Bilgi Sistemleri ve bilinen diğer veri toplama yöntemleri ile karşılaştırıldıkları zaman daha güvenilir bulunmakta ve mali açıdan çok daha ucuza temin edilmektedir. Uzaktan algılamanın tarımda kullanımı ile ilgili olarak birçok çalışma yapılmıştır. Bu konuda çalışan Russel ve ark.,(1992); Gonzales ve ark.(1992); Miller ve ark.(1992); Brisco ve Brown (1995)'un yaptıkları çalışmalar örnek olarak gösterilebilir. Çalışmaların ortak

noktası, yüksek çözünürlüklü görüntülerin kullanılması ile arazilerin sınıflandırılmasının mümkün olduğunu ortaya koymalarıdır.

Uydudan tarlaların görüntülerin alınması ile ilgili örnek bir çalışma, 1998, 1999 yıllarında Bulgaristan'da yapılmıştır. İyi bir çözünürlüğe sahip IKONOS uydusu kullanılarak alınan görüntüler ile 1:5000, 1:25000, 1:50000 gibi oranlarda ölçeklendirme ile test alanları haritalandırılmaya çalışılmış, en başarı sonuçlar 1:50000 ölçeklendirme ile sağlanmış, çalışmada 5600 km² test alanı için haritalandırma yapılmıştır. Bu çalışmalar sonucu elde edilen veriler Bulgasitan'da bir ilk olan, coğrafi bilgi sistemi veritabanına kayıt edilmiş ve elde edilen bu veri tabanı, bazı tarım faaliyetleri, çeşitli tahminler ve kentsel yaşamın planlaması konularındaki çeşitli çalışmalarda kullanılmıştır [4].

Yer ağından belirli mesafelerdeki tarım alanlarında bir örnekleme yöntemiyle toprak ve ürün görüntü özelliklerinin (boy, renk vb metrikler) ölçümlenmesi ve ürün fenolojik gelişiminin izlenmesine yönelik çalışmalarda TARBİL'in ikinci bir altbaşlığını oluşturmaktadır. Yine buna benzer bir çalışma, 2005 yılında Çin'de, Yuzhong'un kuzeyinde Lanzhou'da yapılmıştır. Yapılan çalışmada, TARBİL'de olduğu gibi dijital bir kamera ile gözlemlenen, alandan alınan görüntüler üzerinde işlemler yapılmış, RGB uzayından HSI uzayına geçilerek, elde edilen görüntüler kullanılarak, buğday verimi üzerine bir tahmin sistemi geliştirilmiştir. Dijital kamera görüntüleri kullanılarak bitki örtüsünün tahmini ve yer tabakasından ayrılma noktasının belirlenmesi üzerine başarılı sonuçlar alınmıştır [5].

Ancak TARBİL projesi altında, İTÜ-UHUZAM, 1 Temmuz 2009 tarihinden bu yana SPOT-5 uydu sisteminden veri indirmektedir [6]. Oldukça iyi bir çözünürlüğe sahip, ileri teknoloji sistemi ile kurulmuş SPOT5-uydu sisteminden elde edilen görüntüler ile tarım arazisinin parselleri belirlenmektedir. Ayrıca dijital bir kamera ile izlenen her bir tarlaya ait yakın ve uzak görüntüler olmak üzere iki adet görüntü kullanılarak, bitkinin fenolojik dönemine ait gelişiminin gözlenmesi amaç edinmiştir. Bunun için kamera görüntülerindeki Kırmızı, Yeşil, Mavi bileşenleri Fourier Dönüşümü ile spektral olarak alınır. Ekim öncesi, Fenoloji Başlangıcı, Maksimum Vejetasyon ve Hasat öncesi fazlar, hesaplamalarda gün sayısı olarak kullanılmak üzere teşhis edilir [7]. Burada R, G, B kamera görüntü çerçevesindeki her bir pikselin sırasıyla Kırmızı, Yeşil, Mavi renk bileşen değerleridir. TARBİL kamera görüntülerinde HSI uzayında analizlerin yapılması tercih edilmiştir. Bunun başlıca

nedeni yeşil renk dağılımına göre vejetasyonun özellikle de insan gözü açısından izlenmesi için iyi bir ayrıştırıcılık sağlamasıdır [8].

Verinin bir merkezde toplanarak kayıtlı verilerle birleştirilmesi ve gelişmiş yöntemlerle işlenmesi ise TARBİL'in üçüncü altbaşlığını oluşturmaktadır. Verilerin bir merkezde toplanıp, kullanılması ile, oluşturulan karar mekanizması sistemi bitkinin günlük su kullanım miktarının tespit edilmesine ve tarlaların ne kadar sulama miktarları ile sulanması gerektiğinin bir uyarı mesajı ile çiftçilerin uyarılması için imkan sağlamaktadır. Ayrıca merkezde toplanan veriler ile zirai erken uyarı sistemi tasarlanarak, zararlı canlılara, meteorolojik ve bitkisel hastalıklara karşı ilçe düzeyinde uyarı üretilebilmektedir. Tarla sulama yöntemi, sürekli bir sistemle, intranet üzerinden seçilen, ürün coğrafya ve zaman aralığına ilişkin canlı rapor alınması ve karar destek sistemlerinin oluşturulması ise TARBİL'in dördüncü ve son temel altbaşlığını oluşturmaktadır. Bu başlık altında yer alan VERİS ve KURİS ile hasat sonrası 200 m. çözünürlükle verim dağılım haritası günlük olarak belirlenmekte ve sezon boyunca bitki isteklerine göre sulama istekleri belirlenerek, girilen günlük kuraklık endekslerine göre günlük kuraklık değişim haritası çıkartılabilmektedir.

Özetle, geliştirilen patentli yöntemde, koordinatı hassas şekilde belirlenmiş sabit yer noktalarındaki istasyonlarda toprak nemi, sıcaklığı, dış sıcaklık, nem, rüzgar hızı, rüzgar yönü, basınç, yağış miktarı, tarlanın görüntü özellikleri, izlenen ürünün özellikleri (görüntüsünün işlenmesi ile) ölçülerek, merkeze iletilmektedir. Merkez aynı noktalardan aldığı çok spektrumlu uydu görüntüleri ile yakın çevresinden aldığı görüntüleri kalibre ederek bölgesel ürün gelişim dağılımlarını çıkartmaktadır. Geçmiş verileri oluşturduğunda, gerçekleşen rekolte ile gelecek dönemdeki rekoltenin giderek daha hassas tahmini yapılabilecektir. Tarit bu anlamda büyük bir veritabanı sisteminin oluşmasına imkan sağlayacaktır. Bu projenin çok fazla sayıda altbaşlığı ve yan ürünlerinin olacağı belirtilmiştir. TARBİL projesinin, bundan sonraki kısımlarda sadece yapılan çalışma ile ilgili olan bölümlerinden bahsedilecektir.

Yapılan çalışmanın temel çıkış noktasında proje kapsamında kurulan istasyonlardan düzenli olarak alınan verilerden yapılan çalışmanın temel çıkış noktasında proje kapsamında kurulan istasyonlardan düzenli olarak alınan verilerden, bazen veri iletimi sorunu ya da benzeri nedenler ile belirli zaman aralıklarında alınması gereken verilerde sorun yaşanabilir. Bazen de geçici veya kalıcı bir şekilde veri kaybı yaşanabilir. Çalışmada veriler üçe ayrılarak ele alınır. Verilerin üç gruba

ayrılmasında tasarlanan hata çeşidi teşhisi koyma sistemi görev yapar. Tasarlanan bu sistem verinin var olması ancak ağ sorunlarına bağlı olarak iletilmemesi, verinin yanlış değer olarak alınması, verinin hiç sağlanamaması gibi çıktılar üretir. Bu çıktıların çeşidine göre tasarlanan dalgacık dönüşümü onarım aracı devreye girer. Hata teşhisi koyma aracının ürettiği çıktıya göre, veri iyileştirilir, geçici olarak veriye değer atanır yada kalıcı olarak veriye yeni değer verilir. Tüm bunlar, tasarlanması planlanan tarla sulama çiftçi uyarı sisteminin düzenli bir şekilde çalışması için, gerekli olan eksiksiz veriyi sağlamak içindir. Ayrıca ileriye dönük de bir karar mekanizmasının oluşturulabilmesi için ilgili girdi değerlerinin tahmin edilebilmesi gerekmektedir. Bu problemin çözümünde, zirai meteorolojik verilerden olan sıcaklık ve nem için, eksik veri tamamlama işlemleri yapılmış, kayıp olan verinin zamanı, kendinden önceki değerleri, nem ve benzeri meteorolojik veriler ile olan ilişkisinin ne derecede olduğu saptanmaya çalışılmıştır. Bu sebep ile, belirlenen değişkenler ile sıcaklığın ilişkisini tespit etmek amaçlı korelasyon analizi gerçekleştirilmiştir. Korelasyon sonucu çıkan güçlü, zayıf ilişkili parametreler sınıflandırılmıştır. Sıcaklık ve nem ile ilişkisi güçlü bulunan meteorolojik veriler kullanılarak, eksik sıcaklık değerlerinin tespitine yönelik tasarımlar oluşturulmaya çalışılmıştır. Bu nedenle, tasarlanan sistem için girdi olmasına karar verilen veriler ile ilk önce regresyon analizi yapılmış, ardından bu veriler dalgacık bileşenlerine ayrılmış elde edilen katsayılar ile tekrar regresyon analizi yapılmış, çıkan sonuçlar hem grafik üzerinde hem de matematiksel yöntemler ile karşılaştırılarak yorumlanmaya çalışılmıştır. Dalgacık dönüşümü tekniği kullanılarak, ilgili meteorolojik veri serisinin tahmini için regresyon yöntemine göre gerçeğe en yakın seri elde edilmiştir. Modellemede kullanılan verilerin temininde, eksik veya hatalı veriye sahip olan istasyonun verisini bulmak için komşu istasyonlardan faydalanılmıştır.

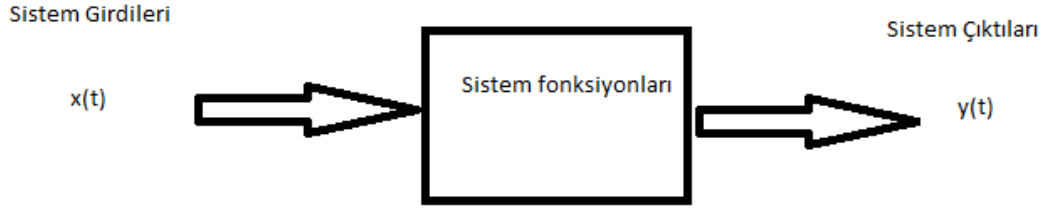
Ayrıca tasarlanan modeller ile, zirai-meteorolojik verilerin tahmin başarımını tespit etmek için, bulunması amaçlanan veriler için, hesaplanan değerler ile verilerin gerçek değerleri arasındaki hata oranı hesaplanmıştır. Buna ek olarak korelasyonu yüksek çıkan iki istasyonun ayrı ayrı verileri alınarak, eksik verileri bulunan diğer istasyonun eksik verileri yerine koyulmuş, böylece gerçek değerlerin serisi ile diğer istasyonların değerleri kullanılarak oluşturulan seriler arasındaki hata oranları ayrı ayrı tekrar hesaplanmıştır. Sıcaklık verileri üzerinde test edilen modellerden elde edilen sonuçlar, açıklamalı olarak karşılaştırılmıştır. Yapılan araştırmalar, geliştirilen

modelin gerek tespit edilen hata oranı, gerekse karşılaştırılan diğer model ile elde edilen sonuçlar doğrultusunda, mantıklı ve kullanışlı bir model olduğunu göstermiştir.

Dalgacık dönüşümü ile yapılan modellemede en uygun pencereleme genişliği tespit edilmeye çalışılmıştır. Ayrıca tasarladığım modeller ile zirai-meteorolojik verilerin en iyi başarım ile tahmini ve yanlış ölçümlü veriler ile gerçek değerleri arasındaki hata oranı tespiti yapılmış, sıcaklık ve nem verileri üzerinde test edilen modeller ile sonuçlar açıklamalı olarak tekrar karşılaştırılmıştır. Dalgacık Dönüşümü Kullanılarak zira-meteorolojik Verilen Hata Teşhis ve Tamiri çalışmasının, gerçekleştirilmesi için, tasarlanan sistemler için sistem girdilerinin sağlanmasında, belirli bölgelerde akıllı yersel istasyonlardan faydalanılmıştır.

1.1 Matematiksel Modelleme

Matematik model, kısaca gözlenen analiz edilen veya tahmin edilen bir olayın bağıntılarla ifade edilmesidir (Küçük, 2004). Matematiksel modeller içerisinde , değişkenler (ölçümler, gözlemler, deney sonuçları v.s.) ve parametreler (katsayılar) barındırırlar. Matematiksel modellemeler sistemler ile birlikte anılırlar. Bu çalışmada, belirlenen sistem girdileri ile oluşturulan modeldeki denkleme ait katsayılar çıktı olarak üretilmektedir. Belirlenen girdiler, çeşitli fonksiyonlardan geçirilerek elde edilen katsayılar ile bir kara kutu modelini oluşturur. Oluşturulan ikinci model için iki kara kutu modeli çizilebilir. Ancak genel olarak, her iki modelde de model çıktıları, modeldeki denklemlere ait katsayılar olduğu için modelleme için tek bir şekil çizimi yeterli olacaktır. Kara-kutu modelleri, yapısal özellikleri ve çalışma tarzı tam olarak belirlenemeyen ve dolayısıyla modellendirilmeleri oldukça güç olan sistemlerin modellenmesinde başarılı olan yöntemlerdir (Partal, 2007) . Kara-kutu modeli Şekil 1.1’de gösterilmiştir.



Şekil 1.1 : Kara-kutu modeli.

1.2 Daha Önceden Yapılan Çalışmalar

Regresyon yönteminin ilk şekli en küçük kareler prensibidir ve ilk olarak Adrien Marie Legendre tarafından 1805 yılında ortaya atılmıştır. Hemen takiben 1809'da C.F. Gauss aynı yöntemi açıklamıştır. En küçük kareler terimi Legendre tarafından moindres carrés olarak kullanılmış, ancak Gauss aynı yöntemi 1795'den beri kullandığını iddia etmiştir. Regresyon terimi 19. yüzyılda İngiliz istatistikçisi Francis Galton tarafından bir biyolojik inceleme için ortaya atılmıştır [9]. Regresyon terimi, 19.yüzyıldan günümüze kadar geliştirilmeye çalışılmış, birçok çalışmada kullanılmıştır.

Dalgacık dönüşümü tekniği, özellikle son yıllarda oldukça revaçta olan bir konudur. Elektronikten biyomedikale, görüntü işlemeden veri sıkıştırılmaya kadar birçok alanda kullanılmakta ve oldukça iyi sonuçlar veren çalışmalarda dalgacık dönüşümü tekniği kullanıldığı tespit edilmektedir. Dalgacık dönüşümü tekniğinin, meteorolojik veriler ile kullanımına ilişkin ilk çalışmalardan en önemlilerinden biri, 90'lı yıllarda yapılan Kumar ve F. Georgiou'ya aittir. (Kumar ve Foufoula-Georgiou, 1993). Yapılan çalışmada yağış verileri incelenmiş ve yağış verileri çok çözünürlüklü teknik ile bileşenlerine ayrıştırılmıştır. Daha sonra elde edilen uydu görüntüleri, yağış durumu için yorumlanmaya çalışılmıştır.

Dalgacık Dönüşümü ve yapay sinir ağları kullanılarak yapılan bir çalışmada, dalgacık dönüşümü ile EEG işaretlerinden çıkarılan öznitelik vektörleri üzerinde istatistiksel işlemlerin gerçekleştirilmesi amaçlanmıştır (Übeyli ve Güler, 2004). Teşhis sistemleri için dalgacık dönüşümünün kullanımı ile işaretlerden öznitelik çıkarılmakta ve işaretlerin zaman-frekans dağılımı istatistiksel özellikler ile gösterilerek, çalışmada farklı kişilerden kayıt edilen elektroensefalogram (EEG) işaretlerinin spektral analizi, dalgacık dönüşümü ile gerçekleştirilmiş ve işaret

hakkında önemli bilgi içeren dalgacık katsayıları işareti temsil eden öznitelik vektörleri olarak kullanılmıştır. Dalgacık katsayıları üzerinde istatistiksel işlemler, öznitelik vektörlerinin boyutlarının azaltılabilmesi için çok katmanlı perseptron sinir ağı sınıflama için kullanılmış ve giriş öznitelik vektörleri seçilmiş, toplam sınıflama doğruluğu dalgacık katsayılarının giriş olarak kullanıldığı çok katmanlı perseptron sinir ağının, EEG işaretlerinin (sağlıklı kişilerden, epilepsi nöbetinin olmadığı durumda epilepsi hastalarından ve epilepsi nöbeti sırasında epilepsi hastalarından kayıt edilen EEG işaretleri) sınıflandırılmasında kullanılabileceğini göstermiştir. Çalışmalarının sonunda, dalgacık katsayılarının giriş öznitelik vektörleri olarak kullanıldığı çok katmanlı perseptron sinir ağı ile EEG işaretlerinin %89,67 doğrulukla sınıflandırıldığı tespit etmişlerdir.

Bir doktora tez çalışmasında, dalgacık dönüşüm tekniği kullanılarak hidrolojik akım serilerinin modellenmesi gerçekleştirilmeye çalışılmıştır (Küçük, 2004). Çalışmada dalgacık dönüşümü tekniği ile hidrolojik akım serileri tahmin edilmeye çalışılmıştır. Öncelikle sürekli dalgacık dönüşümü ve global spektrum yardımı ile ölçüm serilerinin analizi yapılmış, iki ayrı ölçüm istasyonuna ait akarsu akım serileri ayrık dalgacık dönüşümü uygulanarak bileşenlerine ayrılmıştır. Bu bileşenler ile geliştirilen regresyon tipi bir model yardımı ile modelleme yapılmıştır. Modelde tahmin eden ve tahmin ettirici değişkenler yerine akım serilerinin ayrık dalgacık dönüşümü bileşenleri girdi olarak sisteme verilmiştir. Hata kriterleri ile sonuçlar değerlendirilmiştir. Modelleme sonuçlarında çalışma alanındaki her farklı bölgenin iklim karakteristiğine uygun olarak seçilen bileşenler kullanılarak kurulan modellerin diğer modellere göre bir çok hata kriteri bakımından daha başarılı olduğu sonucuna ulaşılmıştır.

Yine bir doktora tez çalışmasında, yağış verilerinin yapay sinir ağları ve dalgacık dönüşümü yöntemleri ile tahmini üzerine yapılmıştır (Partal, 2007). Çalışmada, meydana gelen yağışın doğru tahmin edilmesi için, yine doğrusal olmayan sistemlerin davranışında başarıyla kullanılabilen bir kara-kutu modeli olan yapay sinir ağları kullanılmış, yapay sinir ağları yönteminin tahmindeki başarısını arttırmak için dalgacık dönüşümü kullanılmıştır. Bu çalışmada Yapay Sinir Ağları (YSA) ve dalgacık dönüşümü yöntemleri ile günlük yağış tahmini yapılmıştır. Bu amaçla Türkiye'ye ait 3 istasyonun günlük meteorolojik verileri kullanılmış, YSA yönteminin, İleri Beslemeli Geriye Yayılmalı Yapay Sinir Ağları (İBGYSA) ve

Radyal Tabanlı Yapay Sinir Ağları (RTYSA) yöntemleri yağış tahmini amacıyla kullanılmıştır. Farklı girdi kombinasyonları denenerek her istasyon için en uygun model bulunmaya çalışılmıştır. Sonuçlarda ileri beslemeli geriye yayımlı yapay sinir ağları algoritmasının kullanıldığı yöntem en iyi başarımları sağlamıştır. Dalgacık dönüşümü-YSA yönteminin tahmin sonuçları, çoklu lineer regresyon yönteminin sonuçları ile kıyaslanmış ve başarımları kriterlerine göre daha iyi olduğu gösterilmiştir.

Bir başka çalışmada, Karadeniz Bölgesi'ndeki yağış- akış ilişkisinin farklı yapay sinir ağları metodlarıyla belirlenmesi üzerine gerçekleştirilmiştir (Şahin, 2007). İleri Beslemeli Geriye Yayımlı Yapay Sinir Ağları (İBGYSA) ve Radyal Tabanlı Yapay Sinir Ağları (RTYSA) ve Genelleştirilmiş Regresyon Sinir Ağları (GRSA) olarak farklı yapay sinir ağları metodu ile Çok Değişkenli Regresyon Modeli (ÇDR) ile yağış, sıcaklık, geçmiş akım değerleri kullanılarak, akım tahminleri yapılmış, Sonuçlar karşılaştırılmış ve yapay sinir ağı ile elde edilen sonuçların daha iyi olduğu açıklanmıştır. Ekstrapolasyon için İBGYSA ve RTYSA daha iyi sonuç verirken, düşük akım değerleri için GRYSA'nın daha iyi sonuçlar verdiği açıklanmıştır.

Başka bir çalışmada, Kalp hızının değişiminin dalgacık dönüşümü ve yapay sinir ağları kullanılarak analizi yapılmıştır (Bilgin, 2008). Kalp hızı değişkenliği (KHD) işaretlerinin süreksizlik veya farklı frekanslar içermesinden dolayı dalgacık dönüşümü tekniği ile analizi yapılmaya çalışılmıştır. Çalışmada, ayrık dalgacık dönüşümü (ADD) kullanılarak, ektopikli ve ektopiksiz KHD'lerin analizi ve SD (Sempatik/parasempatik Denge)'nin hesaplanması, ADD ve dalgacık paket dönüşümü (DPD) ile frekans bantlarındaki enerji değerlerinin tespiti ve sonuçların karşılaştırmalı analizi yapılmıştır. Ayrıca bu çalışma, DPD ile ADD uygulaması sırasında meydana gelen frekans bant kaymalarının ortadan kaldırılarak frekans bantlarının literatüre uygun hale getirilmesi, ADD ve yapay sinir ağları (YSA) ile SD'nin otomatik tespiti, DPD ve YSA kullanımı ile alt bantlardaki baskın bileşenlerin belirlenerek frekans bantlarının daraltılması, ÇAF(Çok Alçak Frekans) bandının analizi ve tüm elde edilen sonuçların ayrıntılı olarak değerlendirilmesini kapsar. KHD'lerin DPD ile analizi, ÇAF bölgesinin değerlendirilmesi, SD'nin otomatik tespiti ve baskın frekans alt-bantlarının belirlenmesiyle ilgili, yapılan ilk çalışma olma özelliğindeki bu disiplinler arası tez çalışması ile, önerilen güncel metodlar oldukça önem taşımakla birlikte başka çalışmaların yapılması için bir kapı aralamaktadır.

2009 Yılında yayınlanan bir çalışmada ise, dalgacık dönüşümü ile EKG sinyallerinin özellik çıkarımı ve yapay sinir ağları ile sınıflandırılması yapılmak istenilmiştir (Erdoğan ve Pekçakar, 2009). Bu çalışmada, MIT-BIH veri tabanından alınan Normal, Pace, Rbbb ve Lbbb hastalarından elde edilen ham biyolojik EKG sinyallerine Ayrık Dalgacık Dönüşümü (ADD) uygulanmış ve Daubechies-2 (db2) dalgacıkları kullanılmıştır. Özellik çıkarımında 4. seviyedeki CD4 katsayıları kullanılmıştır. Normal, Pace, Rbbb ve Lbbb hastalarından oluşturulan dört grup özellik katsayıları Matlab programı ile yapay sinir ağlarına (YSA) öğretilmiştir. Daha sonra dört grup için tanımlanan özellik katsayıları test işlemine tabi tutulmuştur. Test sonuçlarında normalizasyon ve qrs taraması yapılmadığı halde, sınıflandırma başarısının, Hızlı Fourier Dönüşümüne göre %90 daha yüksek çıktığı görülmüştür.

Bir başka çalışma, dalgacık dönüşümü ve yapay sinir ağları ile rüzgâr hızı tahmini üzerinedir (Taşcıkaraoğlu ve Uzunoğlu, 2011). Değişken olan rüzgâr hızları, dalgacık dönüşümü yardımıyla daha kararlı karakteristiklere sahip olan bileşenlere ayrılmış ve her bir bileşen ayrı ayrı tahmin edilmeye çalışılmıştır. Böylece tahmin doğruluğu veya aynı doğruluk sınırları içerisinde tahmin süresi arttırılabilmektedir. Bileşenlerin tahminleri için rüzgâr hızları gibi doğrusal olmayan dizileri modellemedeki başarısı nedeniyle Yapay Sinir Ağları (YSA) tercih edilmiştir. Ayrıca yüksek bir frekans aralığına sahip olan ve yalnızca anlık olarak meydana gelen rüzgâr hızı bileşeni filtrelenerek, bu durumun tahmin performansı üzerindeki etkileri araştırılmıştır.

1.3 Tezin Amacı

TARBİL Projesi altında yürülen bu çalışmada, TARBİL projesi kapsamında uzaktan algılama yöntemleri kullanılarak, Türkiye'nin çeşitli coğrafik bölgelerinde kurulan akıllı yersel istasyonlardan faydalanılarak, elde edilen meteorolojik veriler kullanılarak, meteorolojik verilerin eksik olmasından doğan sorunların çözülmesi amaçlı bir tahmin modeli geliştirilmiş ayrıca geleceğe ait meteorolojik verilerin alacağı değerlerin de dalgacık dönüşümü tekniği kullanılarak, tahmin edilmesi amaçlanmıştır. Dalgacık Dönüşümü Kullanarak Zirai-Meteorolojik Verilerin Hata Teşhis ve Tamiri çalışmasının, eksik verilerin tamamlanması bölümü için dalgacık dönüşümü tekniği, regresyon analizi, istatistiksel yöntemler ile modellemeler yapıлып,

elde edilen sonuçlar karşılaştırmalı olarak açıklanarak başarımı en yüksek olan yönetimin tespit edilmesi hedeflenmiştir. Meteorolojik verilerden eksik dataların giderilmesi sorununa ilişkin tasarlanan yöntemlerin karşılaştırılması için modellerdeki hata oranlarının belirlenmesi öngörülmüştür. Ölçülen değer ile gerçek değer arasındaki oluşabilecek hata miktarları ve iki model arasında değişen hata oranları, geliştirilen modellemedeki başarı kriteri olarak kullanılmıştır.

2. ÇALIŞMADA KULLANILAN VERİLER

2.1 Çalışmada Kullanılan Verilerin Kullanım Amacı

Bu çalışma ile amaçlanan, çeşitli nedenler ile elde edilemeyen eksik verilerin en uygun değerler ile doldurulması ve geleceğe yönelik elde edilecek olan zirai-meteorolojik verilerin en doğru şekilde tahminini sağlamaktır.

2.1.1 Yersel akıllı istasyonlardan alınan meteorolojik veriler

Uzaktan algılama teknikleri kullanılarak, elde edilen yersel akıllı istasyonlardan anlık olarak alınan meteorolojik veriler, coğrafi bilgi sistemleri yardımı ile bir merkezde bulunan veritabanına belirli bir düzende ve on dakikada bir zaman aralıkları ile kayıt edilmektedir. Sistemden elde edilen, aralarında on dakika zaman farkları bulunan verileri içeren veri kümeleri, eğitim ve test kümeleri olmak üzere ikiye ayrılır. Geliştirilen modeller önce eğitim kümesi verileri kullanılarak eğitilmiş ardından eğitim ve test kümeleri üzerinde denenerek modellemelere ait başarımlar belirlenmeye çalışılmıştır.

3. ÇALIŞMA ALANININ COĞRAFİ KONUMU

Çalışma alanı olarak, Güney Doğu Anadolu Bölgesi'nde yer alan tarımsal üretimde en yüksek paya sahip Şanlıurfa'daki yersel istasyonların bulunduğu bölgeler seçilmiştir. Seçilen alanlar; Şanlıurfa'nın kuzeyinde ve güneyinde yer alan ilçelerinde bulunan yersel istasyonlar ve çevresini kapsamaktadır . Yapılan çalışmalarda, Yakacık, Poyralı, Akçatat, Rumi, Aydoğdu istasyonları kullanılmıştır. Bahsi geçen istasyonlara ait enlem ve boylam bilgileri TARBİL veri tabanından elde edilmiş ve veri tabanında bulunan her istasyonun birbirine olan Dünya yüzeyinden en kısa kuş uçuşu uzaklıkları (3.1), (3.2), (3.3) denklemleri kullanılarak hesaplanmıştır. Yapılan çalışmalarda belirlenen bir istasyona belirli bir uzaklık referans edilmiş ve buna bağlı olarak, istasyonlar arasında korelasyon katsayıları da baz alınarak, çalışmada kullanılacak olan istasyonlara karar verilmiştir. İlgili açıklayıcı bilgiler, tezin yapılan çalışmalar bölümünde bulunmaktadır.

İki istasyon arası uzaklığı hesaplamak için kullanılan uzaklık formülü aşağıda verilmiştir.

$$a = \sin^2(\Delta\text{lat}/2) + \cos(\text{lat}_1).\cos(\text{lat}_2).\sin^2(\Delta\text{long}/2) \quad (3.1)$$

$$c = 2.\text{atan2}(\sqrt{a}, \sqrt{1-a}) \quad (3.2)$$

$$\text{uzaklık} = R.c \quad (3.3)$$

Burada R dünyanın ortalama yarı çapını göstermektedir ve değeri 6372.797 olarak alınmıştır. Formülde kullanılan enlem ve boylam değerleri, trigonometrik fonksiyonlardan dolayı radyana çevrilerek kullanılmıştır.

Şekil 3.1 tez çalışmasının yapılmış olduğu alanı ve tez çalışmasında kullanılan istasyonları göstermektedir.



Şekil 3.1 : Çalışma alanının coğrafi konumu.

4. YÖNTEM

4.1 Yersel Akıllı İstasyonlar Kullanılarak Eksik Veri Tamamlama Yöntemi

Bitkilerin gereksinim duydukları suyun yağışlarla karşılanamayan kısmı sulama ile toprağa verilir. Bitkilere verilmesi gereken sulama suyu miktarını belirleyebilmek için ihtiyaç duydukları ve tükettikleri su miktarının (evapotranspirasyon) bilinmesi gerekir [10].

Evapotranspirasyon (ET), ya doğrudan ölçülmekte ya da bitki ve iklim verilerinden yararlanarak belirlenmektedir (Kanber, 1997). Doğrudan ölçme yöntemleri sağlıklı sonuç vermesine karşın hem oldukça pahalı, hem de zaman alıcıdır. Bu nedenle, ET yaygın olarak iklim verilerine dayalı tahmin eşitlikleri kullanılarak belirlenmektedir (Güngör ve ark., 1996).

İklim verilerine dayalı ET nin hesaplanabilmesi için gerekli olan günlük ortalama rüzgâr hızı, güneşlenme süresi, oransal nem, minimum ve maksimum sıcaklık gibi meteorolojik veriler Tarit Projesi kapsamında kurulmuş olan yersel akıllı istasyonlardan sağlanmaktadır. Ancak istasyonlarda oluşabilecek çeşitli teknik sorunlardan ötürü, istenen gerekli veriler her zaman elde edilememektedir. Buna ek olarak sulama sisteminin tasarlanmasında ileriye yönelik bir karar mekanizmasının tasarlanması için ilgili parametrelerin var olan verilerinden ileriye dönük değerlerinin de hesaplanması gerekmektedir. Bu nedenler ile, bu bölümde eksik verilerin tamamlanması için tasarlanan yöntemlerden bahsedilecek ve tasarlanan bu yöntemler arasında karşılaştırılma yapılacaktır.

Tasarlanan yöntemler, sıcaklık veri kümesi üzerinde denenmiş ve bu deneme sonucu alınan sonuçlar karşılaştırılmalı olarak incelenmiştir.

4.1.1 Regresyon analizi

Regresyon analizi, iki ya da daha çok değişken arasındaki ilişkiyi ölçmek için kullanılan analiz metodudur. Regresyon analizi ile değişkenler arasındaki ilişkinin varlığı, eğer ilişki var ise bunun gücü hakkında bilgi edinilebilir [11].

Regresyon analizi, bir veya birden çok deęişken arasındaki ilişkilerin deterministik olarak bulunmasını sağlar. Başka bir ifadeyle, ilişkinin modellenmesidir. Regresyon analizi basit ve çok deęişkenli olabilir. Eğer tek bir deęişken kullanılarak analiz yapılıyorsa buna tek deęişkenli regresyon, birden çok deęişken kullanılıyorsa buna da çok deęişkenli regresyon analizi denilir.

Yersel Akıllı istasyonlardan on dakikada bir alınan verilerin oluşturduğu haftalık, aylık veya yıllık veri serilerinde eksiklikler gözlemlenmiştir. Bu eksik sıcaklık verilerinin tamamlanması için önerilen ilk yöntem, regresyon analizi olmuştur.

Eksik verilerin tamamlanması için önerilen regresyon analizinin tek deęişkenli ve çok deęişkenli olarak matematiksel eşitlikleri aşağıda belirtildięi gibidir.

Basit Doğrusal Regresyon Yöntemi:

$$Y = \alpha_0 + X\alpha_1 + e_j \quad (4.1)$$

Çoklu Regresyon Modeli:

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \dots + \alpha_n X_n + \varepsilon_j \quad (4.2)$$

Denklemlerde geçen;

Y : Verisi eksik ve tamamlanacak sıcaklık deęerini,

$X_1, X_2, X_3, \dots, X_n$: Regresyon tahmini için kullanılacak olan deęerleri

$\alpha_0, \alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$: Regresyon katsayılarını,

ε_j : Hata terimini göstermektedir.

Çoklu Regresyon Denklemi,

$$Y = XA + E \quad (4.3)$$

şeklinde de yazılabilir.

Burada X, regresyon için kullanılacak olan deęerler matrisi,

Y ise sonuçlar matrisini gösterir.

X matrisinin ilk sütunun deęerleri 1 olmak koşulu ile, denklem aşağıdaki gibi yazılabilir.

\hat{A} , A'nın en küçük kareler yöntemine göre tahminin vektörü ise,

$$(X'X)\hat{A}=X'Y \quad (4.4)$$

$$\hat{A}=(X'X)^{-1}X'Y \quad (4.5)$$

şeklinde ifade edilir.

Böylece çözüm vektörü \hat{A} 'in nasıl hesaplanacağı bulunmuş olur.

Çoklu Regresyon Modelinin Standard Varsayımları:

1. x_i değerleri sabit ya da ε_i 'den bağımsız X_i rassal değişkenlerinin gerçekleşmiş değerleridir.

2. Hata terimleri, ε_i ortalaması sıfır olan ve varyansları aynı olan rassal değişkenlerdir.

$$E[\varepsilon_i]=0 \quad \text{ve} \quad E[\varepsilon_i^2]=\sigma^2 \quad (i=1, K, n \text{ için})$$

3. Hata terimleri birbirleri ile ilişkili değildir.

$$E[\varepsilon_i\varepsilon_j]=0 \quad (i \neq j \text{ için})$$

4. Her $i=1,2, \dots, n$ için

$$c_0 + c_1x_{1i} + c_2x_{2i} + \dots + c_kx_{ki} = 0$$

eşitliğini sağlayan bir $c_0, c_1, c_2, \dots, c_k$ sayılar kümesinin bulunması olanaksızdır. (x_i 'ler arası lineer bir ilişki olmaması özelliğinden bu durum oluşur.)

Regresyon yönteminde amaç, bilinen bağımsız değişkenler kullanılarak, bilinmeyen bağımlı değişkenin bulunması için katsayı değerinin (basit doğrusal regresyon için), veya birden çok katsayı değerini (çok değişkenli regresyon için) hesaplanmasını ve regresyon denkleminin oluşturulmasını sağlamaktır.

Bu denklemin hesaplanması ile gerçekleştirilmesi istenen amaç, bağımsız değişkenin hesaplanmasında kullanılan denklem katsayılarının, her değişkene en iyi uyumu gösterecek şekilde belirlenmesidir.

4.1.2 En küçük kareler yöntemi

$$SSE = \sum e_i^2 = \sum (y_i - \hat{y}_i)^2 \quad (4.6)$$

$$SSE = \sum e_i^2 = \sum (y_i - (b_0 + b_1x_{1i} + b_2x_{2i} + \dots b_kx_{ki}))^2 \quad (4.7)$$

$(x_{11}, x_{21}, \dots, x_{k1}, y_1), (x_{12}, x_{22}, \dots, x_{k2}, y_2), \dots, (x_{1n}, x_{2n}, \dots, x_{kn}, y_n)$ noktaları,

$$Y_i = \beta_0 + \beta_1x_{1i} + \beta_2x_{2i} + \dots + \beta_kx_{ki} + \varepsilon_i$$

çoklu ana kütle regresyonu sürecindeki n gözlem kümesinden oluşan bir örneklem olsun. $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ katsayılarının en küçük kareler tahminleri olan $b_0, b_1, b_2, \dots, b_k$ değerleri fark kareleri toplamını en küçüğe indirir.

$SSE = \sum (y_i - (b_0 + b_1x_{1i} + b_2x_{2i} + \dots b_kx_{ki}))^2$ fark kareleri toplamını en küçüğe indirir.

\hat{y}_i, b_0, b_1, b_2 değerlerinin hesaplanması için kullanılan denklemler Denklem (4.8), (4.9), (4.10) ve (4.11) olarak belirtilmiştir.

$$\hat{y}_i = b_0 + b_1x_{1i} + b_2x_{2i} \quad (4.8)$$

$$b_1 = \frac{s_Y(r_{x_1y} - r_{x_1x_2}r_{x_2y})}{s_{X_1}(1 - r_{x_1x_2}^2)} \quad (4.9)$$

$$b_2 = \frac{s_Y(r_{x_2y} - r_{x_1x_2}r_{x_1y})}{s_{X_2}(1 - r_{x_1x_2}^2)} \quad (4.10)$$

$$b_0 = \bar{Y} - b_1\bar{X}_1 - b_2\bar{X}_2 \quad (4.11)$$

Denklem katsayılarının optimum düzeyde belirlenmesi kadar, denklemdeki bağımsız değişkenlerin de hesaplanacak olan bağımlı değişken ile olan ilişkilerinin ne derece, olduğu oldukça önemlidir. Regresyonda tahmini yapılacak bağımlı değişken ile, regresyon denklemini oluşturacak her bağımsız değişken arasında birebir güçlü bir ilişki bulunması gerekir.

Böylece tahmini yapılacak bağımlı değişkenin en doğru şekilde hesaplanan katsayı değerleri, en doğru bağımsız değişken sayısı ve değeri kullanılarak, denklemi oluşturulmuş olur.

İki değer arasındaki ilişkinin türünü ve kuvvetini ölçebilmek için korelasyon modeli kullanılır. Sık olarak kullanılan korelasyon modeli, normal korelasyon modelidir. Bu model ile (x,y) olmak üzere iki değişkenin, iki değişkenli normal dağılım olarak neye sahip olduğunu gösterir ve beş parametre ile tanımlanır. Bunlar

x ve y'nin ortalamaları,

x ve y'nin varyansı ve corelasyon katsayısı, p' dir.

Korelasyon katsayısı, iki değişken arasındaki doğrusal ilişkiyi ölçer.

Korelasyon katsayısının özellikleri ise aşağıda sıralandığı gibidir:

- Korelasyon katsayısının değeri -1 ile +1 arasında değerler alır. Ayrıca -1 ve +1 değerlerine de eşit olabilir.
- +1 ve -1 değerlerinin işaretleri, doğru yada ters orantılı ilişkiyi gösterir. Yani - işareti bir değer artarken diğer değer azalması veya değerden biri azalırken diğer değer artmasını gösterirken, bir değer artarken diğerinin de artması ve değerlerden biri azalırken diğerinin de azalması + işareti ile gösterilir. Eğer korelasyon katsayısı değeri -1'e eşit ise mükemmel zıtlık, eğer korelasyon katsayısı değeri +1 ise değişkenler arasında mükemmel uyum vardır.
- İki değer arasındaki korelasyon sayısının, sıfır çıkması iki değer arasında doğrusal bir ilişki olmadığını gösterir.
- Korelasyon katsayısı iki değere göre simetriktir.
- Korelasyon katsayısı değeri, değerlerden herhangi biri için, ölçümün değerine göre değişmez.

Normal korelasyon modelinin değeri p, n gözlem çifti için Denklem (4.12)'de olduğu gibi hesaplanabilir:

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n(\sum X^2) - (\sum X)^2][n(\sum Y^2) - (\sum Y)^2]}} \quad (4.12)$$

(\bar{x}, \bar{y}) verilerin orta noktasını gösterir ise r , Denklem (4.13) 'teki gibi olur.

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{(n-1)S_x S_y} \quad (4.13)$$

4.1.3 Kareler toplamı

Regresyon denklemi iki kısımdan oluşur:

1. Açıklanabilen kısım (Açıklanabilen değişkenlik)

2. Açıklanamayan kısım (Açıklanabilen değişkenlik)

Açıklanan kısım, regresyona dayalı kareler toplamını ifade eder. Yani tahmin edilmiş her değerden, gözlemlenen değerlerin ortalamasının çıkartılmasının karelerinin toplamı ile hesaplanır.

Açıklanamayan kısım ise, üst bölümde açıklanan, hataya dayalı kareler toplamını ifade eder. Yani gözlemlenmiş her değerden, o değer için tahmin edilmiş her değerinin çıkartılmasının karelerinin toplamıdır.

Regresyon denklemini, Denklem (4.14)'te olduğu gibi ifade edebiliriz.

$$\text{Örneklemin toplam değişkenliği} = \text{Açıklanabilen değişkenlik} + \text{Açıklanamayan değişkenlik} \quad (4.14)$$

SST: Toplam sapma (Ortalamadan ayrıştırılan karelerinin toplamı, toplam değişkenlik ölçüsü)

SSR: Açıklanan sapma (Regresyona dayalı kareler toplamı)

SSE: Açıklanamayan sapma olmak üzere (Hataya dayalı kareler toplamı) olmak üzere,

$$SST = SSR + SSE \quad (4.15)$$

regresyon denkleminde çıkartılabilir. Yani sonuç denklemi, Denklem (4.16)'da gösterildiği gibidir.

$$\sum(y_i - \bar{y})^2 = \sum(\hat{y}_i - \bar{y})^2 + \sum(y_i - \hat{y}_i)^2 \quad (4.16)$$

Modellenen regresyon denkleminin başarım gücünün ölçülmesi gerekir. Yani sonuçlar bağımsız değişkenlerin (tahmin değişkenleri) bağımlı değişkendeki (yazma puanı) değişimin yüzde kaçını açıkladığını tespit etmek için, R^2 ile gösterilen belirlilik katsayısı, regresyon modelinin gücünü açıklar. Diğer bir ifade ile, belirlilik katsayısı bağımlı değişkendeki değişimin ne kadarının bağımsız değişkence açıklanabileceğini gösterir. Korelasyon katsayısının karesine eşittir. Denklemdaki açıklanan sapmanın, toplam sapmaya bölünmesi ile bulunur. R^2 değerinin hesaplanmasına ilişkin denklem, Denklem (4.17)'de olduğu gibidir.

$$R^2 = \frac{SSR}{SST} = \frac{\text{Regresyon kareler toplamı(Açıklanan Sapma)}}{\text{Regresyon kareler toplamı(genel kareler toplamı)}} \quad (4.17)$$

R^2 , Belirlilik katsayısı, diğer bir ifade ile determinasyon katsayısı sıfır ile 1 arasında değerler alır.

$0 \leq R^2 \leq 1$, R^2 sıfır değerine eşit ise, bu durum regresyon modelinin oluşmadığını, bir değerine eşit ise de modelde, mükemmel bir lineer bir ilişki kurulduğunu gösterir.

$SST = SSR + SSE$ olduğundan Denklem (4.18)'deki ifade gibi yazılabilir.

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \quad (4.18)$$

Ve R denklem (4.19)'da belirtildiği gibi hesaplanır.

$$R = \sqrt{\frac{SSR}{SST}} \quad (4.19)$$

R, korelasyon katsayısına eşittir. Denklem (4.20) bu ilişkiyi göstermektedir.

$$R = \text{corr}(\hat{y}, y) = \sqrt{R^2} \quad (4.20)$$

4.2 Zaman Serisi

Zaman serisi veya serileri, değişkenin zamana bağlı olarak aldığı değerleri içeren, diğer bir ifade ile zaman ilerledikçe yeni değerler alarak birikimli olarak çoğalan serilerdir. Örneğin sıcaklık serisi, yağış serisi birer zaman serileridir. İstatiksel olarak

kullanılan bu veriler ile matematiksel modellemeler yapılarak, gelecek veya eksik verilerin takmini yapılır.

Zaman serisi, ekonomi, finans, işletme gibi bölümlerde kullanıldığı gibi mühendislik uygulamaları içinde vazgeçilmez bir öneme sahiptir. Zaman serilerini, zaman içerisindeki durumlarına göre iki sınıfa ayırabiliriz. Bunlardan biri zamanla değişmeyen yani durağan işaretler diğeri ise zamanla değişen diğeri bir ifade ile durağan olmayan işaretlerdir. Bu tarz işaretler zaman içerisinde değişen spektral bileşenlerden oluşurlar. Zira meteorolojik olayların inceleme sonucu ortaya çıkan işaretler, değişen yani durağan olmayan bir karakter gösterir. Bu nedenle, doğa olayları sonucu oluşan işaretleri incelemek karmaşıktır. Tabiat olayları arasında zamana bağlı olarak bir modelleme yapılması, çeşitli meteorolojik faaliyetlerin başlangıcı, periyodu ve bitişi gibi bilgilerin en iyi şekilde tahmin edilmesi, zira meteorolojik olaylarındaki bu karmaşıklığı azaltmaktadır. Bu sebeple çeşitli matematiksel modellemeler geliştirilmiştir ve günümüzde daha iyi modellemeler geliştirmek için araştırmalar devam etmektedir.

4.3 Matematiksel Dönüşüm

Dönüşüm, diziler veya fonksiyonlar arasında birinin diğeri haritalanmasını sağlayan matematiksel işlemdir. Dönüşümler yararlıdır çünkü bir fonksiyonun dönüşümü orjinal fonksiyon hakkında başka türlü elde edilemeyen ilave ve gizli bilgileri verebilir. Bir denklemin dönüşümünü çözmek orjinal denkleme çözmekten daha kolay olabilir. Bir fonksiyonun veya dizinin dönüşümü daha az yer tutabilir (Küçük,2004). Bir fonksiyona dönüşüm işleminin uygulanması birçok yönden fayda sağlar. İşareti farklı uzaya taşıyarak, işaret üzerinde yapacağımız işlemler kolaylaştırılmış olur. Ayrıca işaretin sahip olduğu olası farklı özelliklerin de keşfedilebilmesi için imkan sağlanır. Başlıca dönüşüm türleri: Fourier Dönüşümü, Laplace dönüşümü, Z-dönüşümü, Kısa Süreli Fourier Dönüşümü ve Dalgacık Dönüşümü gibi sıralanabilir.

4.3.1 Fourier dönüşümü

Periyodiklik gösteren tüm fonksiyonlar, birbirine dik açı oluşturacak şekilde iki farklı periyodik fonksiyonun, artan frekanslarındaki değerlerinin toplamı şeklinde gösterilebilir. 19. Yüzyıl Fransız matematikçilerinden Jean-Baptiste Fourier belirli

periyota sahip fonksiyonun, sinüs ve kosinüslerin toplamı şeklinde ifade edilebildiğini göstermiştir. Daha sonradan Euler bağıntısıyla, fouriyer dönüşümünde sinüs ve kosinüs fonksiyonlarını kullanmak yerine karmaşık üslü sayıların kullanılması tercih edilmiştir. Ayırık ve sürekli olmak üzere Fourier Dönüşümü başlıca iki bölüme ayrılır. Fourier ve Ters Fourier Dönüşümü'ne ait denklemler aşağıda verilmiştir.

Fourier Dönüşümü:

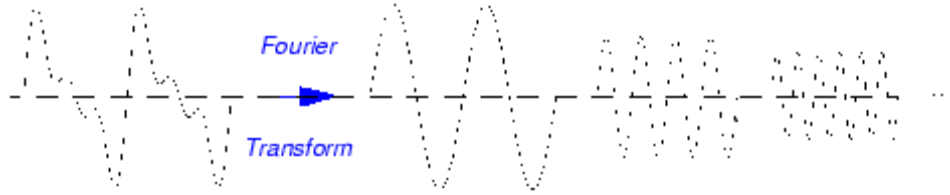
$$X(f) = \int_{-\infty}^{\infty} x(t) \cdot e^{-2j\pi ft} dt \quad (4.21)$$

Ters Fourier Dönüşümü:

$$x(t) = \int_{-\infty}^{\infty} X(f) \cdot e^{2j\pi ft} df \quad (4.22)$$

$$e^{2j\pi ft} = \cos(2\pi ft) + j\sin(2\pi ft) \quad (4.23)$$

Denklemlerde geçen f , frekansı, t ise zamanı belirtir. $X(f)$, $x(t)$ işaretinin fourier dönüşümü sonucunda elde edilen katsayısıdır. Fourier Dönüşümü'ne uğrayan sinyale ait şekil, Şekil 4.1'de gösterildiği gibidir.



Şekil 4.1 : Bir işaretin farklı frekanslardaki fourier dönüşümü.

Denklemlerde geçen f , frekansı, t ise zamanı belirtir. $X(f)$, $x(t)$ işaretinin Fourier Dönüşümü sonucunda elde edilen katsayısıdır. Ters Fourier Dönüşümü ise işaretin zaman ortamında yeniden elde edilmesi için kullanılır. Yani $x(t)$, $X(f)$ katsayılarının üssel fonksiyonun eşleniği ile çarpılıp toplanması sonucu elde edilir. Denklem (4.23)'te ise karmaşık üssel fonksiyonun, gerçekte ve sanal olmak üzere iki fonksiyonun toplamına eşit olduğunu göstermektedir. İfadenin gerçekte kısmı kosinüs dalgası ile, sanal kısım ise f frekansına sahip sinüs fonksiyonu ile temsil edilmektedir.

Fourier Dönüşüm periyodik olan yani durağan işaretlerin çözümlemesi için kullanışlıdır. Fourier Dönüşümü'nde, belirli bir nokta ele alınmak istendiğinde, bu nokta, zaman uzayında bir zaman aralığına karşılık gelmektedir. Dolayısıyla ile durağan

olmayan sinyallerin, zamanına ait bilgiyi tam olarak veremez. Özetle, fourier dönüşümü durağan olmayan işaretler için de kullanılır ancak bu işaretlerin değişimleri hakkındaki bilgiyi veremez. Bu nedenle, Fourier Dönüşümü geliştirilmeye çalışılmış ve durağan olmayan sinyallerin ifade edilebilmesi için, kısa süreli fourier dönüşümü (KSFD) geliştirilmiştir.

4.3.2 Kısa süreli fourier dönüşümü (KSFD)

Diğer bir ifade ile, pencerelenmiş fourier dönüşümü durağan olmayan işaretler için geliştirilmiş bir dönüşüm olup, işaretin frekans bilgisine ek olarak zaman bilgisini de içerir. İşarete Fourier Dönüşümü uygulamadan önce genişliği tanımlanmış pencereden işaret geçirilerek frekanslar elde edilir. KSFD'ye uğrayan bir sinyalin frekans ekseninde gösterimi Şekil 4.2'de gösterilmiştir.



Şekil 4.2 : Kısa süreli fourier dönüşümüne uğrayan bir sinyalin zaman-frekans ekseninde gösterimi.

KSFD'de işaret pencere olarak adlandırılan bir fonksiyon ile çarpılarak zaman ve frekans alanlarında işaret analiz edilmeye çalışılmıştır. Ancak pencereleme fonksiyonu sabit gelişikte bir fonksiyondur. Kısa süreli fourier dönüşümünde kullanılan pencere fonksiyonunun, tarama sırasında sabit genişikte olması, işaretin hızlı değişen yüksek frekanslı değişimlerin zaman uzayında tam olarak bölgelemlenememesine neden olur (Küçük,2004). KSFD'nin eksik kalan bu yönü, dalgacık dönüşümü tekniği ile giredilmeye çalışılmıştır.

4.3.3 Dalgacık dönüşümü

Dalgacık Dönüşümü özellikle son yıllarda bir çok disiplin için sinyallerin yada serilerin incelenmesinde oldukça önemli bir çözüm yolu olmuştur. Dalgacık Dönüşümü, elektronikten biyomedikale, görüntü işlemeden veri sıkıştırılmaya kadar birçok alanda kullanılmış, Fourier ve Kısa Süreli Fourier dönüşümlerinin yetersiz kaldığı noktaları gidermek amacı ile geliştirilmiş bir matematiksel dönüşüm tekniğidir.

4.3.3.1 Dalgacık kavramı

Dalgacık(“Wavelet”), dalganın küçük bir parçası olarak tanımlanabilir. Bu anlamda dalgacık, süresi sınırlı bir titreşim işaretidir (Daubechies, 1996). Dalgacıklar, veriyi farklı frekans bileşenlerine ayıran ve sonra kendi ölçekleriyle eşleştirilmiş bir çözünürlüğe sahip bileşenler üzerinde çalışan matematiksel fonksiyonlardır. Dalgacık algoritmaları, veriyi farklı ölçek veya çözünürlüklerde işler. Eğer geniş bir pencereden bir işarete bakarsak, büyük özellikleri fark ederiz. Benzer şekilde, küçük bir pencereden bir işarete bakarsak, küçük özellikleri fark ederiz. Dalgacıkları ilginç ve aynı zamanda yararlı kılan bu özelliktir (Partal,2007).

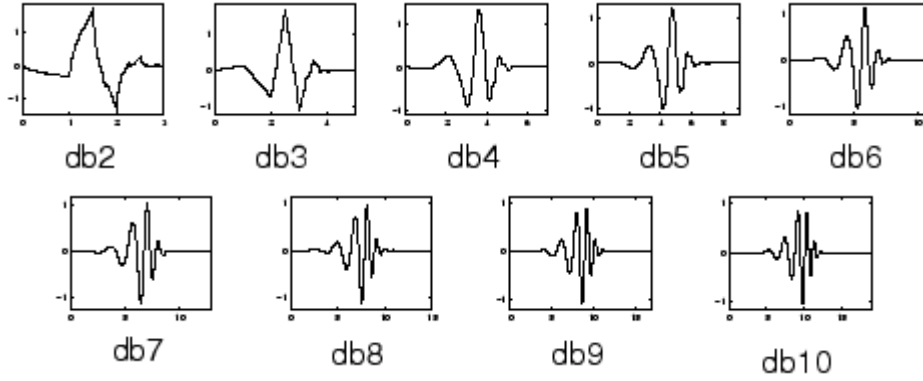
Dalgacık dönüşümü tekniği, bir işaretin aynı zamanda farklı boyutlarda incelemesine imkan sağlamış olur. Dalgacık adı verilen fonksiyonlar zamanda ötelenerek ve genişlikleri değiştirilerek, incelenen işaret ile çarpılırlar. Bu özellikleri ile, işaret üzerinde gezdirilerek dalgacık katsayılarının hesaplanması ile, işaret ve dalgacık fonksiyonu arasındaki benzerlik ilişkisi bulunmuş olur.

Doğada bulunan sinyalleri incelemek amacı ile birçok ana dalgacık fonksiyonu bulunmuştur. Bunlardan, Daubechies, Meksika şapka tipi, Morlet, Meyer, Coiflets, Symlets ve Haar başlıca dalgacık fonksiyonu türleri olarak sıralanabilir. En bilinen dalgacık türü “haar”dir. Basamak şeklindeki bu dalgacık, süreksiz ve ortogonaldır. Zaman işaret grafiğinde işarete ait alanların toplamından hesaplanarak oluşturulmuştur. Diğer dalgacık türü ise Daubechies tarafından geliştirilen “Daubechies” dalgacıdır. Bir çok derecesi olan bu dalgacık çeşidinden Daubechies¹, “Haar” dalgacık türüne eşittir. Ortonormal özelliğe sahip olan Daubechies, işaretin tekrar elde edilmesine imkan sağlar.

Bu çalışmada modele uygunluğu açısından Daubechies dalgacık çeşidi uygulanmış ve çeşitli seviyelerde model üzerinde uygulanan bu dalgacık çeşidinin en iyi sonucu veren Daubechies dalgacı seviyelerinin kullanılmasına karar verilmiştir.

Şekil 4.3’te Daubechies Dalgacı ailesine ait 2, 3, 4, 5, 6, 7, 8, 9, 10 derecelerindeki Daubechies dalgacığının şekilleri verilmiştir.

Dalgacık dönüşümü sürekli ve ayırık olmak üzere iki alt başlığa ayrılır. Sürekli dalgacık dönüşümünde, seçilen dalgacık türüne göre, belirli zaman aralığında işaret ve dalgacık fonksiyonu arasındaki katsayı hesaplanır. Bu katsayı dalgacık fonksiyonu ile işaret arasındaki korelasyonu yani benzerlik ilişkisini gösterir. Dalgacık fonksiyonu işaret üzerinde ötelenerek, her ölçek değerinde C katsayıları hesaplanır. Böylece bir çok katsayı elde edilmiş olur.



Şekil 4.3 : Daubechies dalgacığının 2-10 seviyelerinde gösterimi.

Yüksek ölçeklendirilmelerde dalgacık fonksiyonu esnekleştirilirken, düşük ölçeklerde dalgacık fonksiyonu sıkıştırılarak, orjinal işaret üzerinde gezdirilir. Orjinal işarette frekansın yüksek olduğu yerler sıkıştırılmış olan dalgacık fonksiyonu ile iyi uyum gösterirken, frekansın düşük olduğu yerler ise genişletilmiş dalgacık fonksiyonu ile iyi uyum gösterir.

Ölçeklendirmenin yüksek olduğu yer, işaretin yavaş değiştiği yerlerdir ve işaretin frekansı düşüktür. Ölçeklendirmenin düşük olduğu yer, işaretin hızlı değiştiği yerlerdir ve frekans büyüktür.

Dalgacık dönüşümünün içerdiği değiştirilebilir pencereleme fonksiyonu sayesinde, işaretin hızlı değiştiği yerlerde dar pencereleme ve işaretin yavaş değiştiği yerlerde ise geniş pencereleme fonksiyonları kullanılarak işarete ait özellikler kaybolmadan, işaret analiz edilmeye çalışılır. Yani dalgacık dönüşümü ile işaret, düşük frekanslarda büyük zaman diliminde, yüksek frekanslarda ise küçük zaman diliminde incelenir. Bu özellik dalgacık dönüşümü tekniğini, benzer tekniklerden ayıran en önemli özelliklerinden biridir.

4.3.3.2 Sürekli dalgacık dönüşümü (SDD)

Sürekli Dalgacık Dönüşümü, sürekli ancak periyodik olmayan işaretlerin özelliklerini incelemeye kullanılan bir dönüşüm tekniğidir. Sürekli dalgacık dönüşümüne ait eşitlik aşağıda verilmiştir:

$$SDD_x^\psi(\tau, s) = \psi_x^\psi(\tau, S) = \frac{1}{\sqrt{|s|}} \int x(t) \psi^*\left(\frac{t-\tau}{s}\right) dt \quad (4.24)$$

$$\psi_{\tau, s} = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) \quad (4.25)$$

Eşitliklerde geçen, SDD, sürekli dalgacık dönüşümünde:

T: Öteleme miktarını,

s: Ölçek parametresini,

x(t): incelenecek işareti,

$\psi()$ ise ana dalgacık fonksiyonunu belirtir.

$|s| > 1$ ise zaman ekseninde fonksiyon genişler ve genliği düşer.

$|s| < 1$ ise zaman ekseninde daralır ve genlik büyür.

$|s| < 0$ ise $t=0$ noktasına göre simetriği alınır.

$\tau > 0$ ise zaman ekseninde sağa doğru kaydırma yapılır.

$\tau < 0$ ise zaman ekseninde sola doğru kaydırma yapılır.

SDD ile düşük ve yüksek frekanslı kısımlar çok rahat bir şekilde gözlemlenebilmektedir.

4.3.3.3 Ayık dalgacık dönüşümü (ADD)

Ayrık dalgacık dönüşümü için kullanılan dalgacık fonksiyonuna ait denklem aşağıda ifade edilmiştir.

$$\psi_{m,n}\left(\frac{t-\tau}{s}\right) = s_0^{-m/2} \psi_{m,n}\left(\frac{t-n\tau_0 s_0^m}{s_0^m}\right) \quad (4.26)$$

Denklem (4.26)'da,

m : Dalgacığın ölçek eksenindeki ötelenme parametresi

n : Dalgacığın zaman eksenindeki ötelenme parametresi

s_0 : Sabit bir öteleme adımı ($s_0 > 1$)

τ_o : Zaman eksenindeki ötelenme aralığı değerini gösterir.

Sıklıkla kullanılan s_o ve τ_o değerleri sırasıyla 1 ve 2 dir. Bu çalışmada da bu değerler kullanılmış ve dalgacık fonksiyonu denklemi aşağıdaki gibi düzenlenmiştir.

$$\psi_{m,n}(t) = 2^{-m/2} \psi(2^{-m}t - n) \quad (4.27)$$

Ayrık dalgacık dönüşümü ise Denklem (4.28)'de olduğu gibi ifade edilir:

$$W_{m,n} = 2^{-m/2} \sum_{k=1}^{N-1} x_k \psi(2^{-m}k - n) \quad (4.28)$$

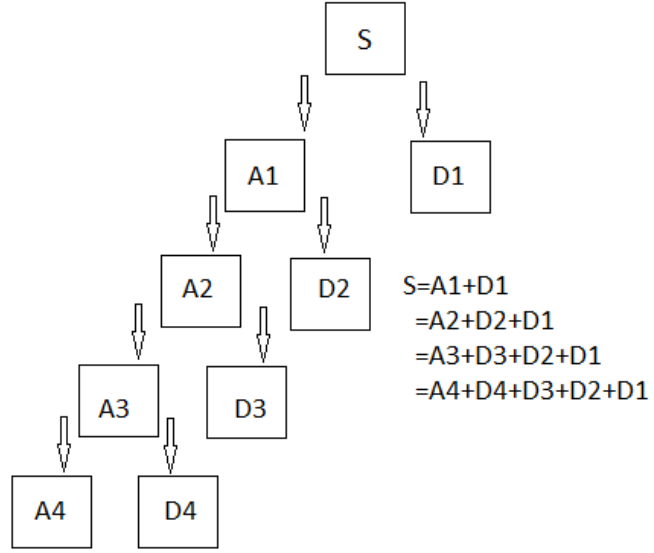
$W_{m,n}$, 2^m ölçeğinde (s), $2^{-m}n$ zaman(τ) değerine sahip dalgacık dönüşümü katsayılarıdır. N ise 2 nin tam katı olan bir sayıdır.

Ayrık dalgacık dönüşümü(ADD), sürekli dalgacık dönüşümünden farklı olarak, dönüşüm katsayılarını, önceden belirlenmiş bir ölçek ve konum kümesini temel alarak hesaplar (Küçük,2004). Ayrık dalgacık dönüşümü, işaretin belirli sayıda ölçeklere ayrılabilmesine olanak sağlar. Ayrık dalgacık dönüşümü ile seçilebilen ölçek ve zaman aralıkları, işaretin incelenmesi için kolaylık sağlar. Böylece sürekli dalgacık dönüşümündeki gibi uzun işlemler azaltılmış olur. Bu çalışmada uygulamada kolaylığı açısından ayrık dalgacık dönüşümü kullanılmıştır.

Mallat tarafından geliştirilen, çok çözünürlüklü analiz (ÇÇA) ile işaret belli sayıda ölçeklere ayrılır. Bu yöntem ile ADD kullanılarak ayrık ve detay bileşenlerine ayrılan sinyal, ardışık olarak tekrar ayrık ve detay bileşenlerine ayrılabilir. ÇÇA için alçak ve yüksek geçiren filtreleme yapan ölçek ve dalgacık fonksiyonları kullanılır.

Ayrık dalgacık dönüşümü ve çok çözünürlüklü analiz ile, her ayrışma işlemi sonunda örnek sayısı yarıya iner. Böylece zamanda çözünürlük azalır ancak frekansta çözünürlük artmış olur çünkü her alt seviyede frekans bandı genişliği yarıya inmiş olur. Böylece frekanstaki belirsizlik azalmış olur. Çok çözünürlüklü ardışık ayrıştırma işlemi işaretin farklı özelliklerinin ortaya çıkartılmasına imkan sağlar. Böylece bir işaretin farklı özellikleri ile birlikte incelenmesine imkan sağlanmış olur.

Ardışık ayrıştırma işlemene ait örnek şekil, Şekil 4.4'te gösterilmiştir.



Şekil 4.4 : Ayrık dalgacık dönüşümü ve çok çözünürlüklü ayrıştırma örneği

,

5. YAPILAN ARAŞTIRMALAR VE SONUÇLARI

Yapılan çalışmalarda, TARİT projesi kapsamında Türkiye'nin çeşitli yerlerine kurulan yersel akıllı istasyonlardan faydalanılarak, enlem ve boylam dereceleri bilinen istasyonlar arasındaki uzaklık hesaplatılarak, karşılaştırılmış olan belirli istasyonlardan edilen aynı on dakika zaman aralıklarına sahip, sıcaklık, nem serileri, eğitim ve test kümesini oluşturmak üzere iki parçaya ayrılmıştır. Seriler hesaplanacak olan istasyon ile, belirlenmiş olan bu merkez istasyonun çevresinde oluşturulan hayali bir r yarıçapı içinde yer alan istasyonlar belirlenmiştir. Bu istasyonlar ile merkez istasyon arasında teker teker korelasyon analizi yapılmış, r yarı çapı içinde kalan istasyonlardan, en iyi korelasyonu gösteren ilk iki istasyonun verileri kullanılarak, üçüncü istasyona ait sıcaklık yada nem zaman serileri regresyon modeli ve dalgacık dönüşümü tekniği kullanılarak hesaplanmaya çalışılmıştır.

Tez çalışmasında örnek olması açısından özellikle meteorolojik verilerden sıcaklık ve nem seçilmiştir. Bunun birçok sebebi vardır. İlk olarak yağış verileri üzerine çok fazla araştırma yapılmıştır ancak sıcaklık üzerine dalgacık dönüşümü kullanılarak yapılan araştırma sayısı oldukça kısıtlıdır. İkinci olarak, sulama faaliyetlerinde kullanılmak üzere tasarlanması planlanan sistem için sıcaklık değerleri oldukça önem arz etmektedir. Tasarlanması planlanan sistem için sıcaklık verileri oldukça doğrudan sistemin çalışmasında etkilidir. Üçüncü olarak, geliştirilen modelin sadece sıcaklık verileri üzerinde değil nem gibi başka meteorolojik veriler üzerinde de etkili olduğu gösterilmek istenilmiştir. Modelin denenmiş olduğu meteorolojik verilerden bir tanesi ise basınç verileridir. Basınç verileri, normal doğa olayları akışında, kısa zamanda çok fazla değişiklik gösteren veriler değildir. Bu sebep ile çalışmalar içerisinde örnek olarak verilmemişlerdir ancak yapılan araştırmalar sonucunda basınç verileri üzerinde de modellemenin iyi sonuçlar verdiği gözlemlenmiştir.

Yapılan çalışmalar sonucunda, bulunan meteorolojik veri serileri test ve eğitim kümesinde ayrı ayrı denenmiştir. Her iki veri kümesi için de ortalama karesel hata, ortalama mutlak hata değerleri hesaplanmış, sonuçlar karşılaştırmalı olarak açıklanmıştır.

İlk olarak, iki istasyonun verileri kullanılarak üçüncü istasyona ait verilerin tahmini için doğrusal regresyon yöntemi kullanılmış ve yönteme ait hesaplanan regresyon katsayıları hem eğitim hem de test kümesinde kullanılarak, üçüncü istasyona ait veriler bulunmuştur

- İkinci olarak, en iyi sonucu veren Daubechies dalgacığı kullanılarak eğitim ve test veri kümeleri üzerinde ayrık dalgacık dönüşümü kullanılarak işaret, ana ve detay bileşenlerine ayrıştırılmıştır.
- Her istasyon için aynı seviyedeki ana ve detay bileşenleri arasında doğrusal regresyon modeli kurularak, eğitim kümesi kullanılarak elde edilen her seviyedeki katsayılar kullanılarak tahmini sıcaklık yada nem serilerine, ters dalgacık dönüşümü uygulanarak toplam meteorolojik veri serisine ulaşılmaya çalışılmıştır. Bu işlem ayrı ayrı eğitim ve test kümesi içinde yapılarak hata oranları elde edilerek oluşturulan yöntemin başarımı analiz edilmiştir. Yapılan karşılaştırmalı çözümleme sonucu dalgacık dönüşümü tekniğinin kullanılması ile geliştirilen yöntemin sonucu, doğrusal regresyon yöntemine göre iyileştirdiği ispatlanmıştır.

Çalışmalarda MATLAB R20011b sürümü kullanılmış, çalışmanın genelinde gerekli fonksiyonlar matlab ile, kodlanmıştır.

5.1 Birinci Durum

Modelin geçerliliğini göstermek amacı ile, gerçekleştirilen modelleme çeşitli mesafeler için ve çeşitli meteorolojik veriler kullanılarak tekrarlanmıştır. İlk çalışma TARBİL sisteminde 36,90184 K enlemi ve 38,92288D boylamında bulunan Şanlıurfa'da kurulan Yakacık istasyonu için yapılmıştır. Bu istasyon için r uzaklığı 5 km olarak belirlenmiştir. Yani Yakacık istasyonuna en fazla 5 km uzaklıkta bulunan istasyonlar tespit edilmiştir.

Yakacık istasyonu çevresinde oluşturulan 5 km çapı içinde kalan istasyonların enlem, boylam ve Yakacık istasyonuna olan uzaklıkları km cinsinden Çizelge 5.1'de verilmiştir.

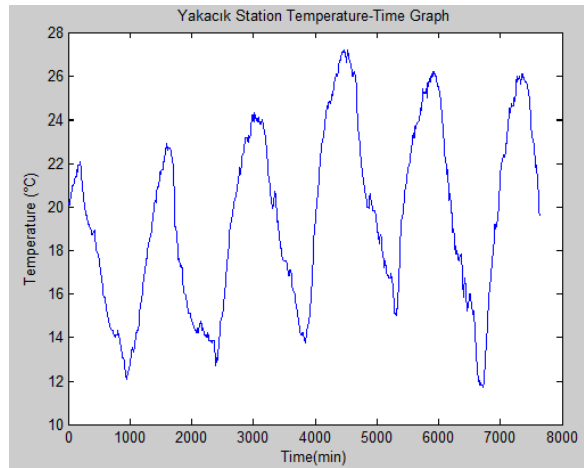
Çizelge 5.1'den anlaşıldığı gibi, Yakacık istasyonuna en yakın istasyon Aydoğdu ve en uzak istasyon Demirören istasyonlarıdır.

Çizelge 5.1 : 5 km çap içinde kalan istasyonlar.

İstasyon Numarası	İstasyon Adı	İstasyon Enlem	İstasyon Boylam	Yakacık İst. olan uzaklığı (km)
16	Rumi	36,90289K	38,91981D	0,296724448
17	Aydoğdu	36,90120K	38,92067D	0,2091637
18	Demirören	36,90324K	38,91585D	0,644245165

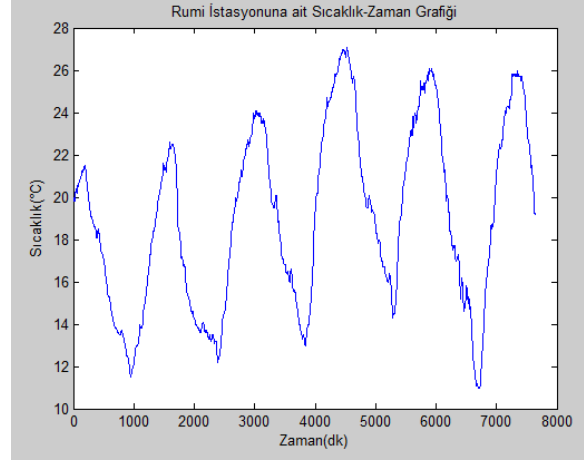
Yakacık istasyonu için yapılan gözlemde, her on dakikada bir alınan kesintisiz sıcaklık verilerinden oluşan sıcaklık zaman serisi 766 adet veri içermektedir. Bu verilerden ilk 388 adeti eğitim, diğer 388 adeti test kümesi olarak ayrılmıştır. Dolayısı ile eğitim ve test kümeleri için yaklaşık 2.7 günlük veri üzerinden çalışma yapılmıştır. Kullanılan verilerin ilki 01.05.2011 saat 13:50:00, sonuncusu ise 06.05.2011 saat:21:20:00 tarihlidir.

İstasyonlara ait 766 adet veri kullanılarak hesaplanan sıcaklık-zaman grafiği Şekil 5.1, Şekil 5.2, Şekil 5.3 ve Şekil 5.4’de gösterilmiştir.

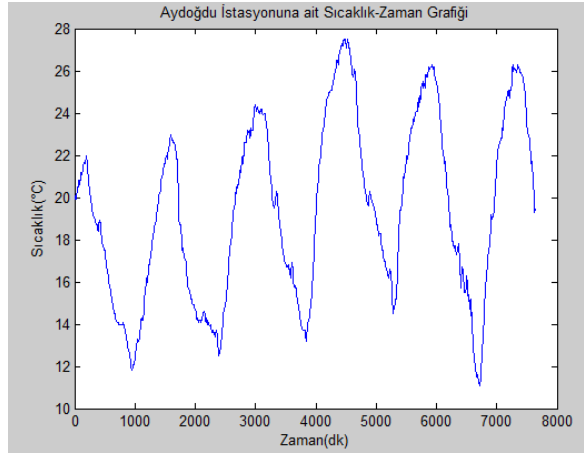


Şekil 5.1 : Yakacık istasyonuna ait sıcaklık zaman grafiği.

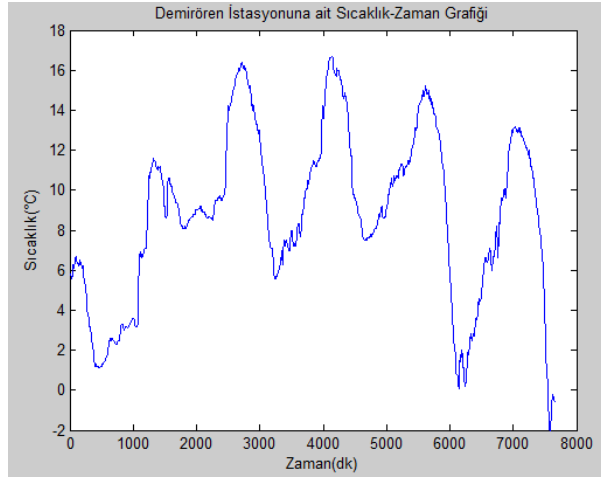
Rumi, Aydoğdu ve Demirören istasyonlarına ait sıcaklık zaman grafikleri ise sırası ile aşağıda belirtilmiştir. Grafiklerden görüldüğü gibi bir istasyonun dışındaki tüm istasyonların grafikleri birbirine benzerdir. Özellikle Güney Doğu Anadolu Bölgesi'nin yer şekillerinin düz olması bu durumun oluşmasında en büyük etkidir.



Şekil 5.2 : Rumi istasyonuna ait sıcaklık zaman grafiği.



Şekil 5.3 : Aydoğdu istasyonuna ait sıcaklık zaman grafiği.



Şekil 5.4 : Demirören istasyonuna ait sıcaklık zaman grafiği.

Rumi, Aydoğdu ve Demirören istasyonlarından alınan sıcaklık zaman serileri, eğitim ve test olarak ikiye ayırmadan önce, yine eğitim ve test verileri olarak ayrılmayan Yakacık istasyonuna ait sıcaklık zaman serisi ile olan ilişkileri gözlemlenmeye

çalışılmıştır. 5 km çap içerisinde kalan her komşu istasyonun ayrı ayrı, Yakacık istasyonu ile olan korelasyon katsayıları hesaplanmıştır. Elde edilen korelasyon katsayıları Çizelge 5.2’de verilmiştir.

Çizelge 5.2 : İstasyon15 ile istasyon16, istasyon17 ve istasyon18 arasındaki korelasyon katsayıları.

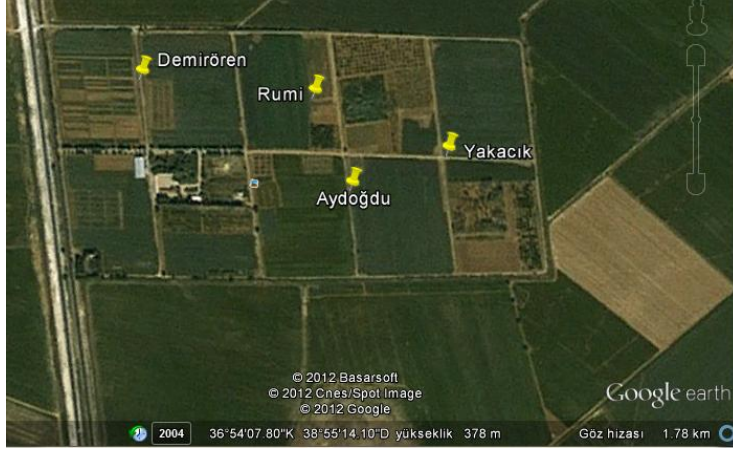
İstasyon Numarası	İstasyon Adı	Yakacık İstasyonu ile olan korelasyon katsayısı
16	Rumi	0,9992
17	Aydoğdu	0,9988
18	Demirören	0,3254

Yakacık istasyonuna en yakın olan Aydoğdu İstasyonu olduğu halde, Rumi istasyonu ile Yakacık istasyonu arasındaki korelasyon katsayısı en yüksek çıkmıştır. Demirören istasyonu ile olan korelasyon katsayısının düşük olmasında istasyonların kuruldukları yerlerdeki yer şekilleri, bakı gibi etkenler büyük rol oynamaktadırlar.

Genel anlamda Demirören İstasyonları hariç, görüldüğü gibi korelasyon katsayıları oldukça yüksek çıkmıştır. Bilindiği gibi korelasyon iki değer arasındaki doğrusal ilişkiyi gösterir. Bunun anlamı değerlerin tıpa tıp aynı olduğu anlamına kesinlikle gelmez. Ayrıca korelasyonun yüksek çıkması da her zaman için iyi uyum olduğunu ifade etmeyebilir. Her iki değer bağımlı olduğu, üçüncü bir değer de olabilir. Ancak bu çalışmada yer alan seriler arasındaki ilişki, iki değer arasında doğrusal bir ilişki olduğunu gösterebilir. Çünkü kullanılan değişkenler benzer değişkenler ve dolayısı ile bu değişkenleri etkileyen üçüncü değişkenler de aynı değişkenlerdir. Korelasyon değerlerinin yüksek çıkması bu değerlerin birbirleri yerlerine kullanılabilirliklerini düşündürebilir. Bu yaklaşım hatalı değildir ancak bize gerçek sonucu vermez. Bu ve diğer çalışmalarda da amaç gerçeğe en yakın veriyi bulmak olduğu için, korelasyon bu çalışmada en yakın veriyi içeren istasyonun tespiti için kullanılmıştır. Buradan hareket ile, seçilen iki istasyon, iki ayı model için kullanılmış ve geliştirilen modelin daha iyi sonuç verdiği, diğer modelin ürettiği sonuç kullanılarak, modelin başarımı desteklenilmek istenilmiştir.

Şekil 5.5’te google earthten istasyonların enlem ve boylam dereceleri girilerek birinci durum için belirlenen yerler gösterilmektedir.

Ayrıca www.tarit.org adresinden edinilen haritadan da istasyonların konumları Şekil 5.6’da gösterildiği gibidir.



Şekil 5.5 : İstasyonların konumlarının google earth kullanılarak gösterilmesi.



Şekil 5.6 : İstasyonların konumlarının tarit istasyonlar haritası kullanılarak gösterilmesi.

Her iki şekilden de anlaşılacağı gibi, yer şekilleri benzer olan istasyonlar arasında güneş ışınlarını alma açıları farklılık göstermektedir. İstasyonların alınan eş zamanlı veriler incelendiğinde de diğer istasyonlara göre daha kuzeyde bulunan Demirören istasyonuna ait verilerin en düşük ikinci sırada ise en düşük Rumi istasyonuna ait veriler gözlemlenmiştir.

Yapılan korelasyon analizi sonucunda, Yakacık istasyonuna ait verilerin tahmini için korelasyon katsayıları oldukça yüksek çıkan Rumi ve Aydoğdu istasyonlarının verilerinin kullanılmasına karar verilmiştir.

5.1.1 Birinci durum için belirlenen regresyon modeli

Farklı durumlar için aynı açıklamalar tekrar tekrar yapılmayacaktır. Örnek olması açısından 1. Durum detayları ile birlikte açıklanmıştır.

Bu çalışmada, Yakacık istasyonu sıcaklık verilerinin tahmini için iki bağımsız değişkenli doğrusal regresyon modeli kullanılmıştır. Eğitim kümesindeki Yakacık istasyonuna ait veriler bağımlı değişken, Rumi ve Aydoğdu istasyonlarına ait veriler

ise bağımsız değişkenler olarak kullanılarak, regresyon denklemi katsayıları hesaplanmıştır.

y: Yakacık İstasyonuna ait sıcaklık verileri

x1: Rumi İstasyonuna ait sıcaklık verileri

x2: Aydoğdu İstasyonuna ait sıcaklık verileri olmak üzere,

$$y = a + b_1 \cdot x_1 + b_2 \cdot x_2 \quad (5.1)$$

(5.1) Doğrusal Regresyon denkleminde a, b1 ve b2 katsayıları değerleri sırası ile 0.7871, 0.7582, 0.2200 olarak bulunmuştur. Dolayısı ile elde edilen doğrusal regresyon denklemi, Denklem (5.2) gibidir.

$$y = 0.7871 + 0.7582 \cdot x_1 + 0.2200 \cdot x_2 \quad (5.2)$$

elde edilen bu denklem hem eğitim hem de test kümeleri üzerinde denenecek, bulunan değerler ile gerçek değerler arasındaki ortalama karesel hata (OKH) hesaplanmıştır.

OKH için denklem aşağıda verilmiştir.

$$OKH = \frac{1}{n} \sum_{i=1}^n ((Y_{i(\text{gerçek değer})} - Y_{i(\text{bulunan değer})}))^2 \quad (5.3)$$

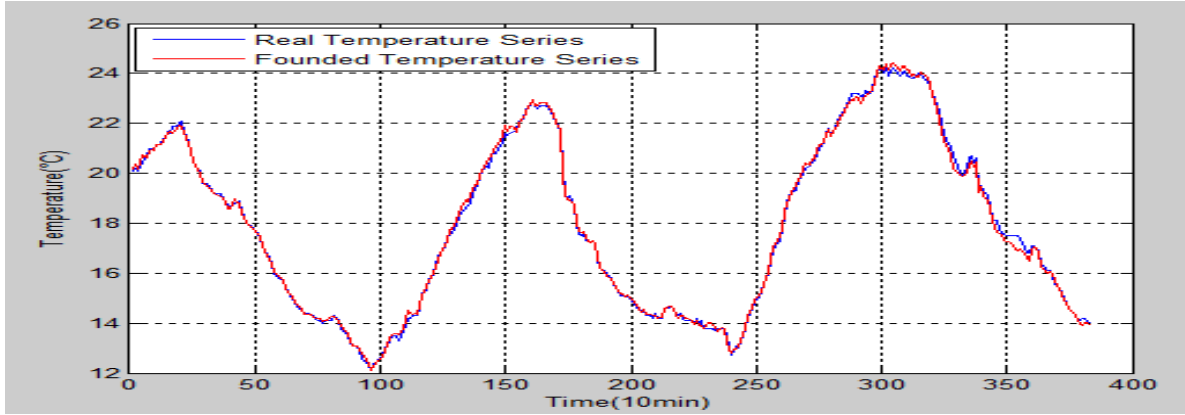
Eğitim kümesi üzerinde hesaplanan ortalama karesel hata (OKH) değeri 0.0202 olarak bulunmuştur.

OKH değeri karesel olarak, hatayı ifade ettiği için, grafikte pik değerlerde ve ani sıcaklık değer düşüklüğüne geçişlerin olduğu bölümlerde, hata oranının yüksek hesaplanmasına neden olmaktadır. Bu nedenle çalışmada gerçek değer ve ölçülen değerlerin mutlak farkının veri sayısına oranını veren Ortalama Mutlak Hata (OMH) kriterinde göz önünde bulundurulmuştur. OMH'yı ifade eden denklem aşağıdaki gibidir.

$$OMH = \frac{1}{n} \sum_{i=1}^n |Y_{i(\text{gerçek değer})} - Y_{i(\text{bulunan değer})}| \quad (5.4)$$

Eğitim kümesi üzerinde hesaplanan Ortalama Mutlak Hata (OMH) değeri de 0.1064 olarak bulunmuştur.

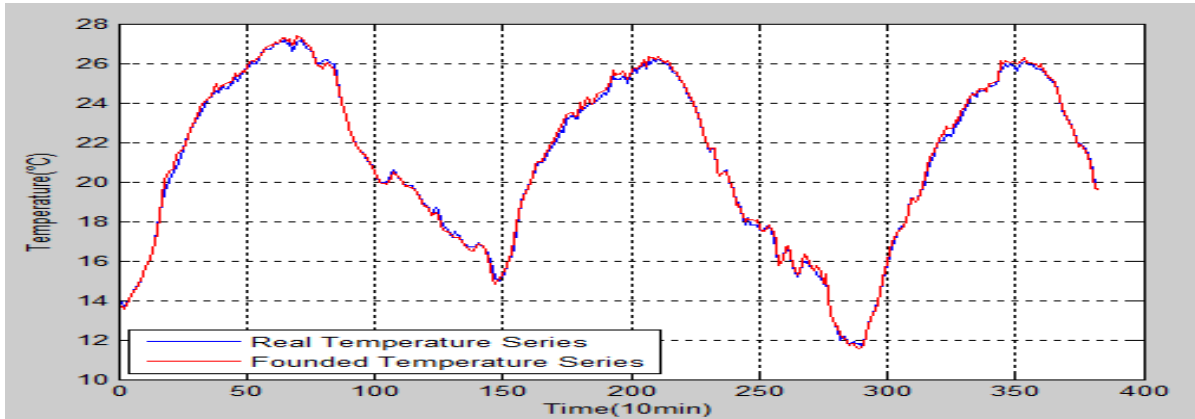
Eđitim kümesinin gerek deęerleri ve modelleme sonucu bulunmuř deęerlerine ait grafik Őekil 5.7’de gsterilmiřtir.



Őekil 5.7 : Eđitim kümesinin gerek sıcaklık deęerleri ve modelleme sonucu bulunmuř sıcaklık deęerlerine ait grafik.

Test kümesi üzerinde hesaplanan OKH deęeri 0.0321 olarak bulunurken, OMH deęeri 0.1337 olarak bulunmuřtur.

Test kümesinin gerek deęerleri ve modelleme sonucu bulunmuř deęerlere ait grafik Őekil 5.8’de gsterilmiřtir.



Őekil 5.8 : Test kümesinin gerek sıcaklık deęerleri ve modelleme sonucu bulunmuř sıcaklık deęerlerine ait grafik.

5.1.2 Birinci durum regresyon ve dalgacık dnřm modeli

Oluřturulan ikinci modelde ise, modelde yer alan tm istasyonlara ayrıık dalgacık dnřm uygulamada kolaylık saęlayan Ingrid Daubechies tarafından bulunan, 1, 2, 3 ve 4 seviyelerinde ‘‘Daubechies’’ dalgacıęı kullanılarak, istasyonlardan elde edilen sıcaklık- zaman serilerine ayrıık dalgacık dnřm uygulanmıřtır. Her seviyede bir nceki seviyeye gre sahip olduęu veri sayısı yarıya dřerek ayrıřtırılan sıcaklık

zaman serisi için dört seviyeli ayrışım modeli diğer seviyeli ayrışım modellerine kıyasla en iyi sonucu vermiştir. Bu kıyaslamada yine her seviye için Ayrık dalgacık dönüşümüne uğrayan Yakacık istasyonuna ait sıcaklık- zaman serisi ile diğer istasyonlara ait ayrık dalgacık dönüşümü uygulanan sıcaklık-zaman serileri arasındaki korelasyon çözümlemesi baz alındı. Eğitim kümesinden alınan bağımlı değişken (Yakacık İstasyonuna ait veriler) ve bağımsız değişkenler (Rumi ve Aydoğdu istasyonlarına ait veriler) kullanılarak, Yakacık, Rumi, Aydoğdu istasyonlarının ana bileşen (A4) ve D1, D2, D3, D4 detay bileşenleri için 5 ayrı regresyon denklemi kuruldu ve regresyon denklemine ait katsayıları hesaplandı.

Oluşturulan modele ait denklemler aşağıda belirtilmiştir.

Yakacık İstasyonuna ait sıcaklık Değerleri= Yakacık _A4+ Yakacık _D4+ Yakacık _D1+ Yakacık _D2+ Yakacık _D3+ Yakacık _D4 olarak,

Rumi İstasyonuna ait sıcaklık Değerleri=Rumi_A4+ Rumi_D4+ Rumi_D1+ Rumi_D2+ Rumi_D3+ Rumi_D4 olarak,

Aydoğdu İstasyonuna ait sıcaklık Değerleri= Aydoğdu _A4+ Aydoğdu _D4+ Aydoğdu _D1+ Aydoğdu _D2+ Aydoğdu _D3+ Aydoğdu _D4 olarak ayrıştırıldı.

Daha sonra ayrışım ve detay bileşenleri arasında aşağıda belirtilen regresyon modelleri oluşturuldu.

Yakacık_A4(Bulunan) =a + b1 . Rumi_A4 + b2 . Aydoğdu_A4

Yakacık_D1(Bulunan) =a + b1 . Rumi_D1 + b2 . Aydoğdu_D1

Yakacık_D2(Bulunan) =a + b1 . Rumi_D2 + b2 . Aydoğdu_D2

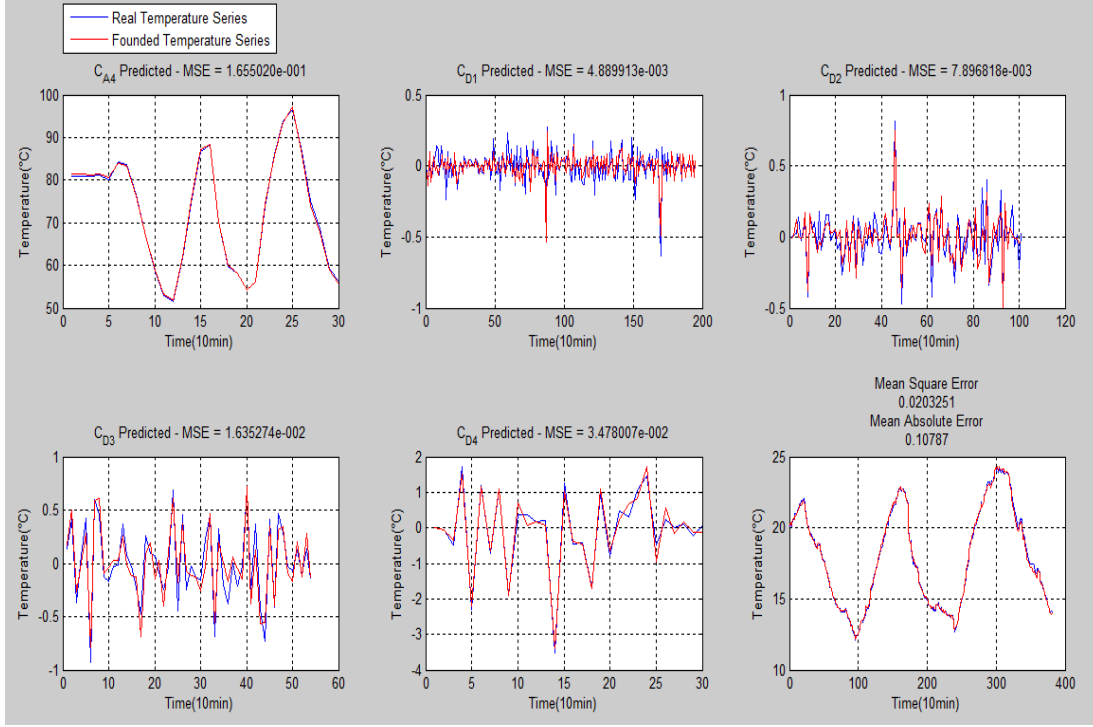
Yakacık_D3(Bulunan) =a + b1 . Rumi_D3 + b2 . Aydoğdu_D3

Yakacık_D4(Bulunan) =a + b1 . Rumi_D4 + b2 . Aydoğdu_D4

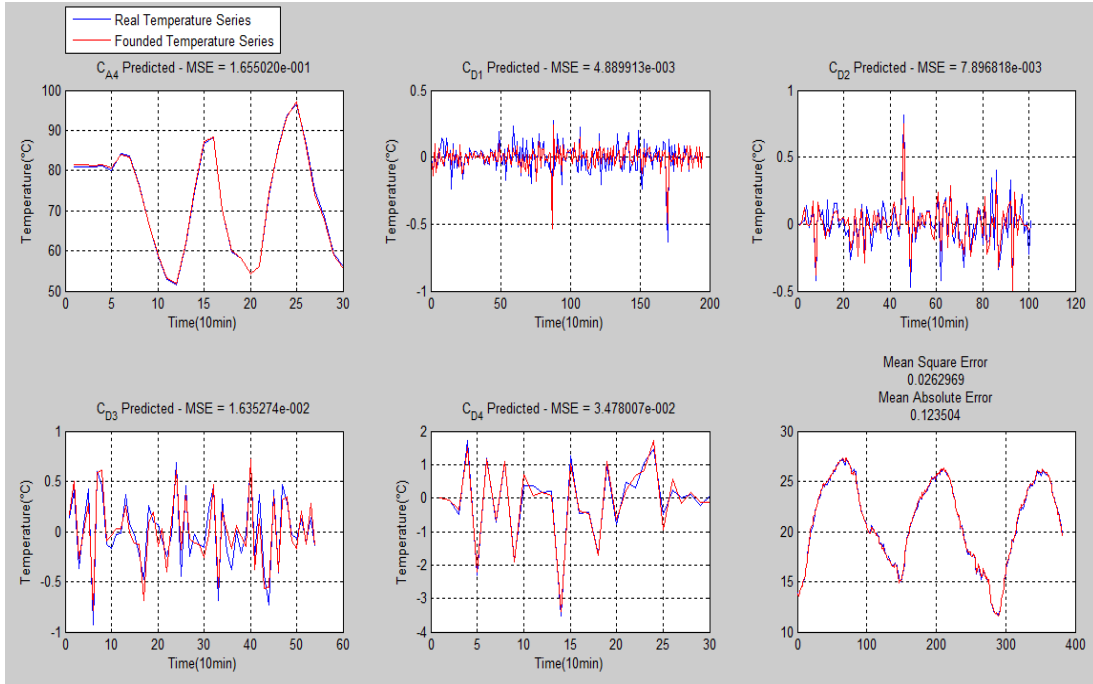
Her seviyedeki a, b1, b2 katsayıları bulundu. Regresyon sonucu bulunan değerlere ters dalgacık dönüşümü uygulanarak ayrık ve detay bileşenler toplandı ve tahmini sıcaklık zaman serisine ait veriler böylece bulunmuş oldu.

Eğitim kümesi ve test kümesi db4 dalgacığı kullanılarak yaklaşık (A4) ve detay bileşenlerine (D1, D2, D3, D4) ayrılmıştır. Modelleme sonucu oluşan bileşenler ile

eđitim ve test kmelerinin yaklařık ve detay bileřenlerinin karřılařtırmasına ait grafikler Őekil 5.9'da ve Őekil 5.10'da gsterilmiřtir.



Őekil 5.9 : Eđitim kmesi zerinde bulunan ve denenen yaklařık, detay ve orjinal sinyallere ait regresyon+dalgacık dnřm modeline iliřkin grafikler.



Őekil 5.10 : Eđitim kmesi zerinde bulunan ve test kmesi zerinde denenen yaklařık, detay ve orjinal sinyallere ait regresyon+dalgacık dnřm modeline iliřkin grafikler.

Eđitim kümesindeki veriler kullanılarak hesaplanan katsayılar öncelikle eğitim kümesi üzerinde daha sonra test kümesi üzerinde denenmiştir. Sırası ile bulunan OKH değerleri, 0.0203, 0.0263' iken sırası ile OMH değerleri ise, 0.1078 ve 0.1235'dir. Bir dalgacık dönüşümü kullanılmayarak oluşturulan regresyon modeli ile bulunan OKH ve OMH değerleri ile dalgacık dönüşümü kullanılarak hesaplanan OKH ve OMH değerler kıyaslanacak olursa, her iki modelde de eğitim kümesi üzerindeki OKH ve OMH değerleri yaklaşık olarak aynı çıkarken test kümesi üzerinde hesaplanan OKH ve OMH değerlerinde dalgacık dönüşümü tekniğinin kullanılması ile sonucun iyileştiđi görülmüştür. Korelasyon sayılarının çok yüksek olmasına rağmen geliştirilen modelin saf regresyon modeline göre test kümesi üzerinde, sonucu iyileştirdiđi gözlemlenmiştir.

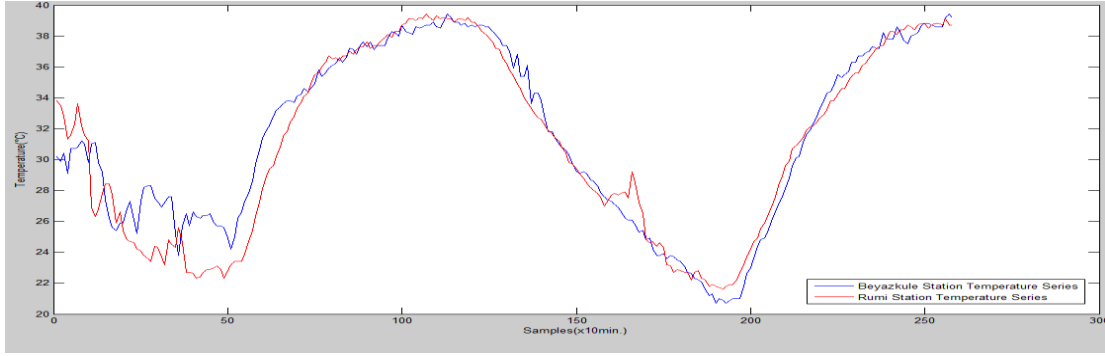
5.2 İkinci Durum

Bu çalışmada da Beyazkule istasyonu sıcaklık verilerinin tahmini için iki bağımsız deđişkenli doğrusal regresyon modeli kullanılmıştır. Eğitim kümesindeki Yakacık istasyonuna ait veriler bağımlı deđişken, Rumi ve Aydođdu istasyonlarına ait veriler ise bağımsız deđişkenler olarak kullanılarak, regresyon denklemi katsayıları hesaplanmıştır. Bu çalışmada, Beyazkule İstasyonu'na olan uzaklıkları en az 95 km en fazla 105 km olacak şekilde belirlenen bir r mesafesi içinde kalan Rumi ve Aydođdu istasyonları seçilmiştir. Beyazkule ve diđer istasyonlar arasındaki korelasyon katsayıları hesaplanarak, Beyazkule ve Rumi İstasyonları arasındaki korelasyon sayısı 0,9576 ve Beyazkule Aydođdu İstasyonları arası korelasyon katsayısı ise 0.0974 olarak bulunmuştur. Korelasyonun yüksek çıkması yerşekilleri ve diđer faktörler ile ilişkilidir. Karşılaştırılan iki model sonucunda, korelasyonun yüksek çıkması sadece istasyon seçimleri için önem taşıdığı anlaşılmaktadır. Yapılan çalışmada, üç istasyon için alınan 516 adet veri kümesi, eğitim ve test kümesi olmak üzere iki bölüme ayrılmıştır.

5.2.1 Rumi istasyonu ve Aydođdu istasyon verilerinin Beyazkule istasyonu verileri yerine kullanılması

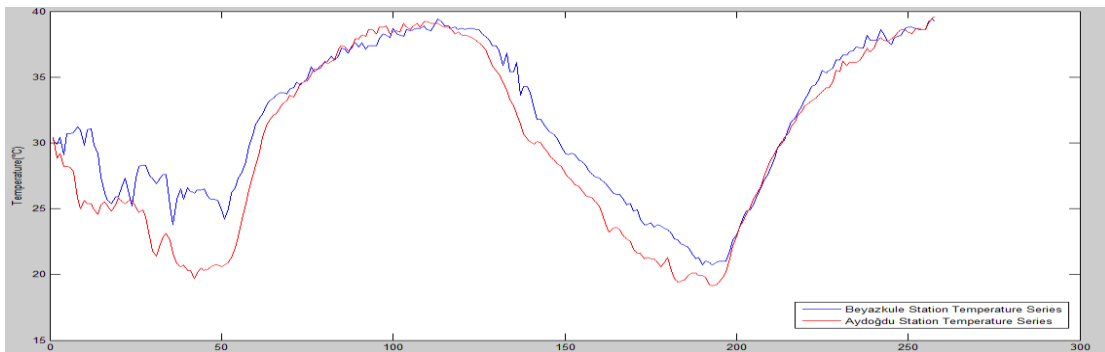
İlk olarak test kümesinde yer alan, korelasyon değerleri yüksek çıkan Rumi ve Aydođdu istasyonlarına ait sıcaklık değerleri ayrı ayrı Beyazkule istasyonunun test değerlerinin yerine yazılmıştır. Beyazkule istasyonuna ait test kümesindeki gerçek değerler ile Rumi istasyonuna ait değerler arasındaki OKH ve OMH değerleri

hesaplanmıştır. Öncelikle, Rumi isyonu test kümesi değerleri Yakacık istasyonu test kümesi değerleri yerine yazılmıştır. Elde edilen değerler: OKH için 2.6206 iken, OMH için 0.1.1500'dır. Daha sonra, aynı işlem Aydoğdu ve Beyazkule istasyonları arasında tekrarlanmıştır. Bu durumda, OKH için hesaplanan değer 5.7060 iken, OMH için hesaplanan değer 1.7271 olarak bulunmuştur. Bu işlemin yapılmasının sebebi, eksik datası bulunan istasyonun değerleri ile korelasyon katsayısı değeri yüksek olan diğer iki istasyonun değerlerinin doğrudan kullanılması ve korelasyon katsayısı değeri yüksek çıkan istasyonların geliştirilen modelde kullanılması sonucu elde edilen OKH ve OMH değerlerindeki değişimi göstermektir. Elde edilen bu sonuçlar ileriki sayfalarda geliştirilen model ile karşılaştırılmalı olarak daha net açıklanacaktır. Şekil 5.11 'de Beyazkule ve Rumi istasyonlarının aynı zamana ait sıcaklık serileri gösterilmiştir.



Şekil 5.11 : Beyazkule ve Rumi istasyonlarının aynı zamana ait sıcaklık serileri.

Şekil 5.12'de Beyazkule ve Rumi istasyonlarının aynı zamana ait sıcaklık serileri gösterilmiştir.



Şekil 5.12 : Beyazkule ve Rumi istasyonlarının aynı zamana ait sıcaklık serileri.

5.2.2 İkinci durum için regresyon modeli

y = Beyazkule İstasyonu Sıcaklık verisi

x_1 = Rumi İstasyonu Sıcaklık verisi

x_2 = Aydoğdu İstasyonu Sıcaklık verisi

i = Veri indeksi olmak üzere,

Beyazkule İstasyonu için kullanılan denklem, Denklem (5.5)'te belirtilmiştir.

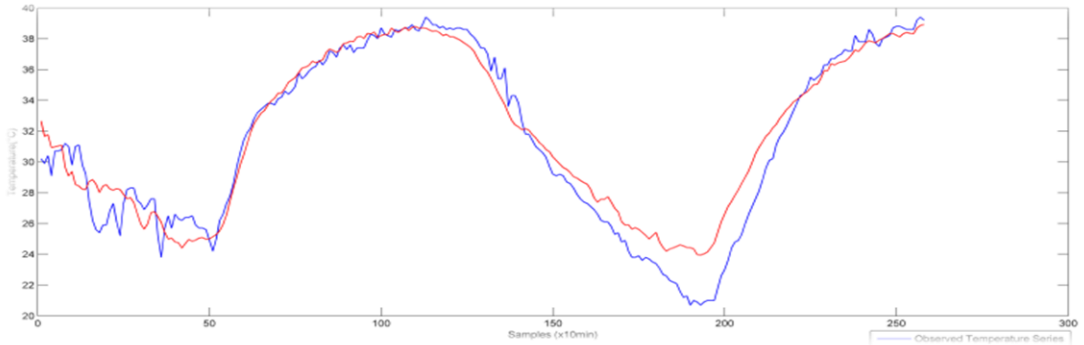
$$y_i = a + b_1 \cdot x_{i1} + b_2 \cdot x_{i2} \quad (5.5)$$

Doğrusal regresyon denklemi için hesaplanan katsayılar:

$a = 8.9958$, $b_1 = 0.1675$ and $b_2 = 0.5920$ olmak üzere, Denklem (5.6), hesaplanan denklemi göstermektedir.

$$y_i = 8.9958 + 0.1675 x_{i1} + 0.5920 x_{i2} \quad (5.6)$$

Şekil 5.13'de regresyon modeli için eğitim kümesinde bulunan değerlerin test kümesinde test edilmesine ilişkin grafik gösterilmiştir.



Şekil 5.13 : Regresyon modeli için eğitim kümesinde bulunan değerlerin test kümesinde test edilmesine ilişkin grafik.

5.2.3 İkinci durum için regresyon ve dalgacık dönüşümü modeli

Modelde yer alan tüm istasyonlara ayrık dalgacık dönüşümü uygulamada kolaylık sağlayan Ingrid Daubechies tarafından bulunan, çeşitli seviyelerde “Daubechies” dalgacığı kullanılarak, istasyonlardan elde edilen sıcaklık serilerine ayrık dalgacık dönüşümü uygulanmıştır. Her seviyede bir önceki seviyeye göre sahip olduğu veri sayısı yarıya düşerek ayrıştırılan sıcaklık serisi için Db6 ayrışım modeli diğer seviyeli ayrışım modellerine kıyasla en iyi sonucu vermiştir.

Oluşturulan modele ait denklemler aşağıda belirtilmiştir.

T: Sıcaklık Değeri

A: Ayrık Bileşen

D: Detay Bileşen

Genelleştirilmiş model denklemleri;

BS: Beyazkule İstasyonu

$$T_{BS} = A6_{BS} + D6_{BS} + D5_{BS} + D4_{BS} + D3_{BS} + D2_{BS} + D1_{BS} \quad (5.7)$$

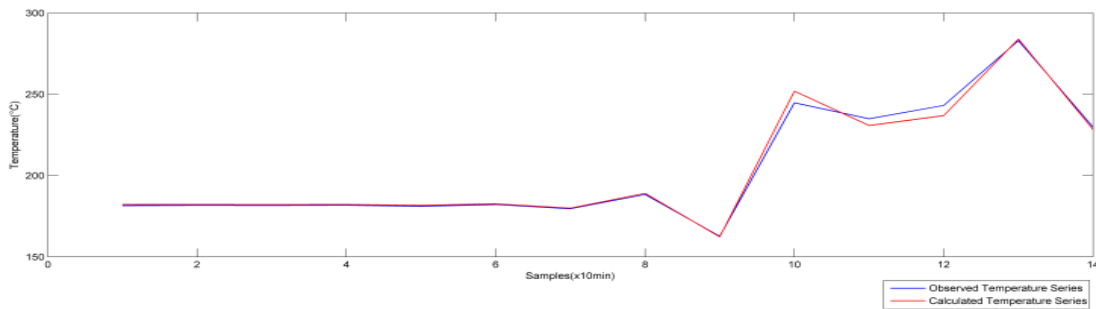
RS: Rumi İstasyonu

$$T_{RS} = A6_{RS} + D6_{RS} + D5_{RS} + D4_{RS} + D3_{RS} + D2_{RS} + D1_{RS} \quad (5.8)$$

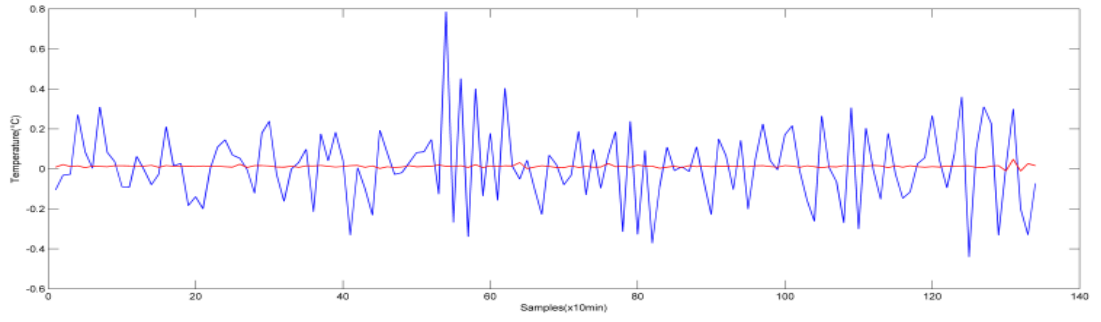
AS: Aydoğdu İstasyonu

$$T_{AS} = A6_{AS} + D6_{AS} + D5_{AS} + D4_{AS} + D3_{AS} + D2_{AS} + D1_{AS} \quad (5.9)$$

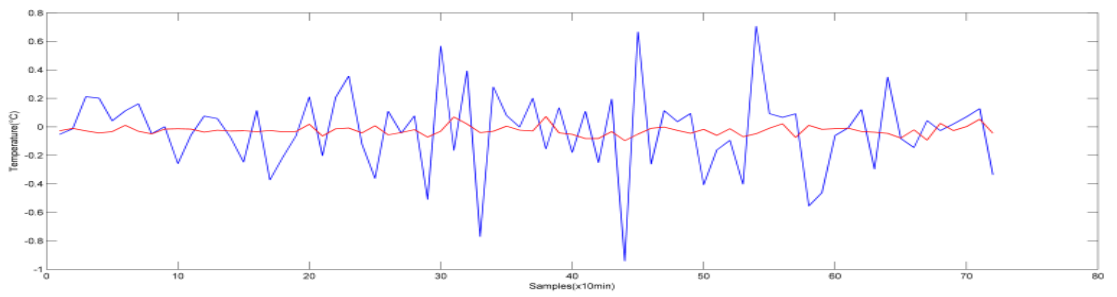
Her seviyedeki a, b1, b2 katsayıları bulundu. Regresyon sonucu bulunan değerlere ters dalgacık dönüşümü uygulanarak ayrık ve detay bileşenler toplandı ve tahmini sıcaklık zaman serisine ait veriler böylece bulunmuştur. Şekil 5.14.a, 5.14.b, 5.14.c, 5.14.d, 5.14.e, 5.14.f, 5.14.g, 5.14.h Eğitim kümesi üzerinde bulunan ve denenen yaklaşık, detay, orjinal sinyallere ait regresyon+dalgacık dönüşümü modeline ilişkin grafikleri göstermektedir.



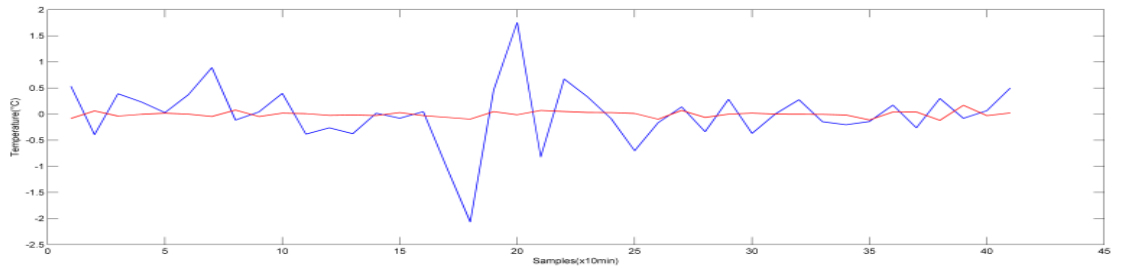
(a)



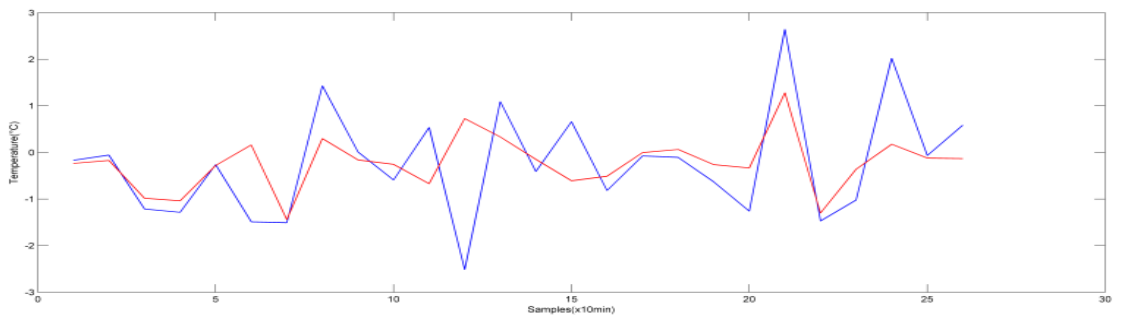
(b)



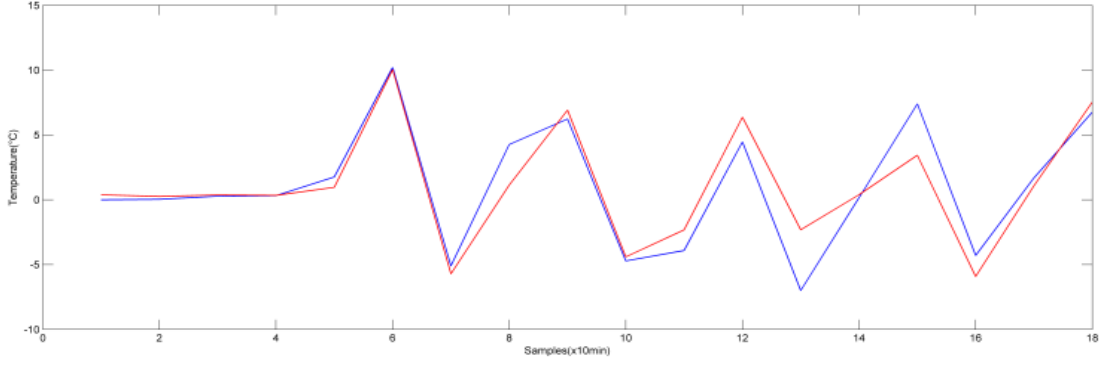
(c)



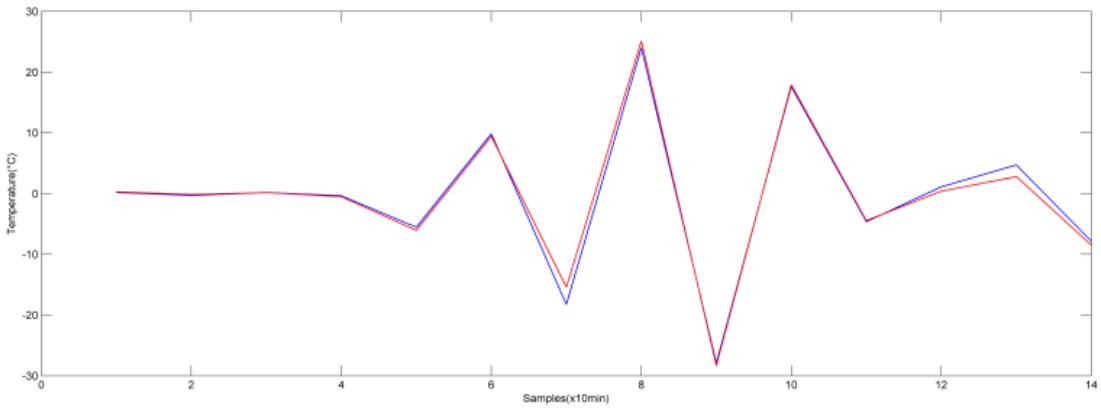
(d)



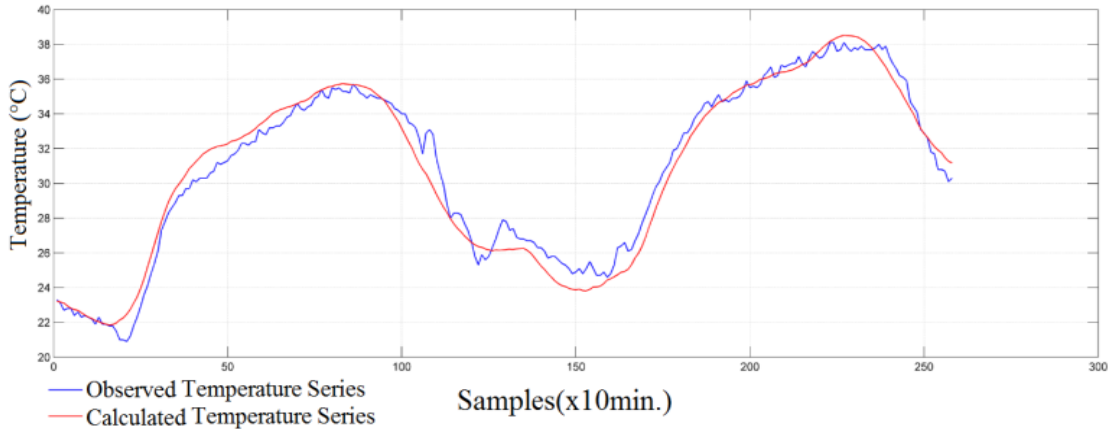
(e)



(f)



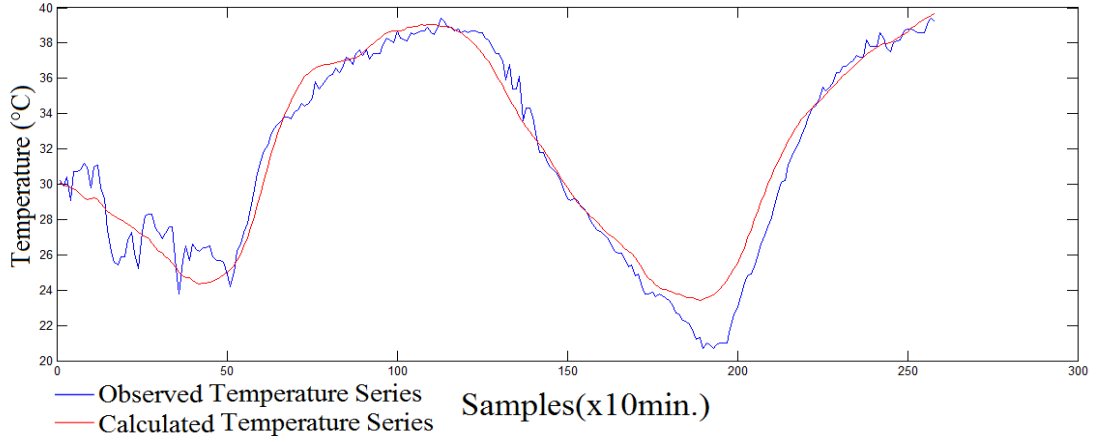
(g)



(h)

Şekil 5.14 : Eğitim kümesi üzerinde bulunan ve denenen yaklaşık, detay, orjinal sinyallere ait regresyon+dalgacık dönüşümü modeline ilişkin grafikler, (a)A6 ana bileşen, (b) D1, (c) D2, (d) D3, (e) D4, (f)D5, (g)D6, detay bileşenler (h) Gözlenlenmiş ve hesaplanmış sıcaklık serileri.

Şekil 5.15'te regresyon+ dalgacık dönüşümü modeli için eğitim kümesinde bulunan değerlerin test kümesinde test edilmesine ilişkin grafik gösterilmektedir.



Şekil 5.15 : Regresyon+ dalgacık dönüşümü modeli için eğitim kümesinde bulunan değerlerin test kümesinde test edilmesine ilişkin grafik.

Yapılan çalışma hem eğitim kümesinde hem de test kümesinde olmak üzere iki farklı kümede gerçekleştirilmiştir. Her iki küme için sonuçlar ayrı ayrı hesaplanmıştır. Eğitim kümesindeki veriler kullanılarak hesaplanan katsayılar öncelikle eğitim kümesi üzerinde daha sonra test kümesi üzerinde denenmiştir. Sırası ile bulunan OKH ve OMH değerleri Çizelge 5.3'te verilmiştir. Elde edilen sonuçlar doğrultusunda genişletilen modelde OKH ve OMH değerlerinde düşme olduğu gözlemlenmiştir.

Çizelge 5.3 : İkinci durum için OKH ve OMH değerleri.

Doğrusal Regresyon Modeli	OKH	OMH
Eğitim Kümesinde	0,962	0,7523
Test Kümesinde	2,0984	1,0923
Doğrusal Regresyon Modeli+ Dalgacık Dönüşümü Modeli	OKH	OMH
Eğitim Kümesinde	0,7291	0,6995
Test Kümesinde	1,5323	0,9526

İki bağımsız değişkenli bir doğrusal regresyon modeli ile bulunan OKH ve OMH değerleri ile dalgacık dönüşümü kullanılarak hesaplanan OKH ve OMH değerler karşılaştırılacak olursa, her iki modelde de eğitim ve test kümeleri üzerinde hesaplanan OKH ve OMH değerlerinde dalgacık dönüşümü tekniğinin kullanılması ile sonucun iyileştiği gözlemlenmiştir. Böylece, %26.98 oranında OKH'da, %12.79 oranında ise OMH'da iyileşme gözlemlenmiştir. Ayrıca her iki sonuçta elde edilen OKH ve OMH değerleri, Akçatat ve Poyralı istasyon değerlerinin Yakacık istasyon değerleri yerine

kullanılması sonucu hesaplanan ayrı ayrı OKH ve OMH değerlerinden çok daha düşük olduğu çok açık bir şekilde gözükmektedir.

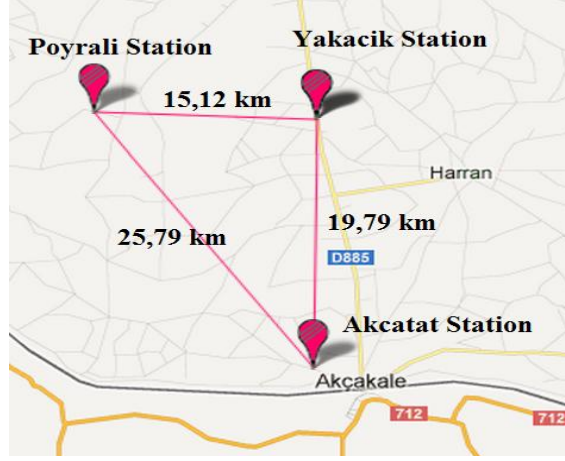
5.3 Üçüncü Durum

Oluşturulan üçüncü modelde ise serisi tahmin edilecek istasyon olarak, Yakacık İstasyonu seçilmiştir. Bu istasyona göre komşu istasyon mesafesi 15 km ile 20 km olarak belirlenmiştir. Yani en yakın 15 km, en uzak 20 km olacak şekilde komşu istasyonlardan, Yakacık istasyonu ile korelasyonu değerine bakılarak, en iyi korelasyonu gösteren iki istasyon, komşu istasyon olarak, Akçatat ve Poyralı istasyonları belirlenmiştir. Yakacık ile komşu istasyonlar arasındaki korelasyon sayıları Çizelge 5.4'te gösterilmiştir. Toplamda 1432 adet sıcaklık verisi kullanılırken, bunlardan ilk 716 adeti eğitim, kalan 716 adeti ise test kümesi için ayrılmıştır. İlk kullanılan verinin tarihi, 01.05.2011 saat 13:50 iken, son kullanılan verinin tarihi 11.05.2011 ve saati 12:20' olarak alınmıştır. Modelde yer alan tüm istasyonlara ayrık dalgacık dönüşümü uygulamada kolaylık sağlayan Ingrid Daubechies tarafından bulunan, çeşitli seviyelerde "Daubechies" dalgacığı kullanılarak, istasyonlardan elde edilen sıcaklık serilerine ayrık dalgacık dönüşümü uygulanmıştır. Her seviyede bir önceki seviyeye göre sahip olduğu veri sayısı yarıya düşerek ayrıştırılan sıcaklık serisi için Db3 ayrışım modeli diğer seviyeli ayrışım modellerine kıyasla en iyi sonucu vermiştir.

Çizelge 5.4 : Akçatat ve Poyralı istasyonlarının Yakacık istasyonu ile olan korelasyon katsayıları.

İstasyon Numarası	İstasyon Adı	Yakacık İstasyonu ile olan Korelasyon Katsayısı
20	Akçatat	0,9834
25	Poyralı	0,9619

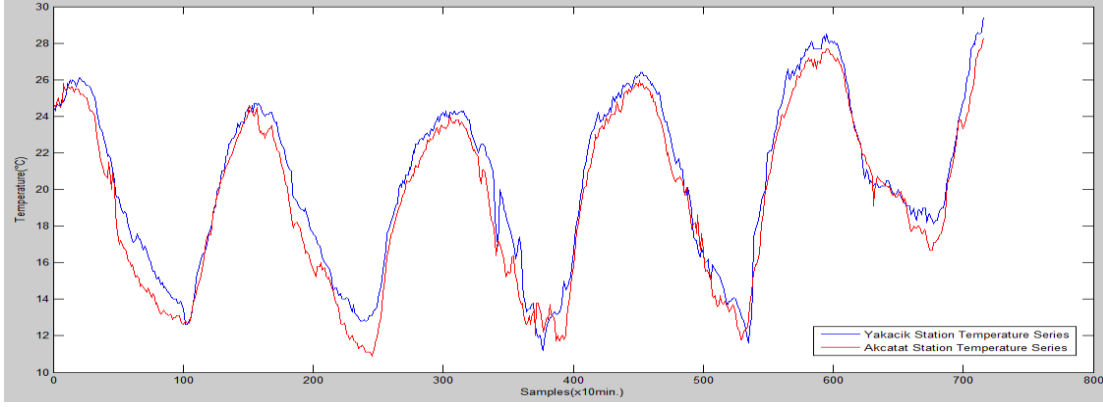
Yakacık, Akçatat ve Poyralı istasyonları arasındaki uzaklık Şekil 5.16'te gösterildiği gibidir.



Şekil 5.16 : Yakacık, Akçatat ve Poyralı istasyonları arasındaki mesafeler (km).

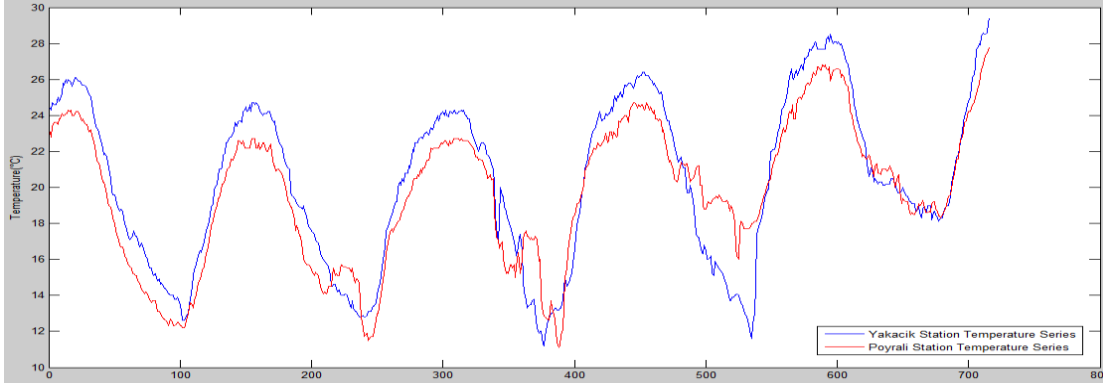
5.3.1 Akçatat istasyonu ve Poyralı istasyon verilerinin Yakacık istasyonu verileri yerine kullanılması

İkinci duruma benzer şekilde, ilk olarak test kümesinde yer alan, korelasyon değerleri yüksek çıkan Akçatat ve Poyralı istasyonlarına ait sıcaklık değerleri ayrı ayrı Yakacık istasyonunun test değerlerinin yerine yazılmıştır. Yakacık istasyonuna ait test kümesindeki gerçek değerler ile Akçatat istasyonuna ait değerler arasındaki OKH ve OMH değerleri hesaplanmıştır. Öncelikle, Akçatat istasyonu test kümesi değerleri Yakacık istasyonu test kümesi değerleri yerine yazılmıştır. Elde edilen değerler: OKH için 1.304 iken, OMH için 0.9620'dir. Daha sonra, aynı işlem Poyralı ve Yakacık istasyonları arasında tekrarlanmıştır. Bu durumda, OKH için hesaplanan değer 3.4486 iken, OMH için hesaplanan değer 1.5696 olarak bulunmuştur. Bu işlemin yapılmasının sebebi, eksik datası bulunan istasyonun değerleri ile korelasyon katsayısı değeri yüksek olan diğer iki istasyonun değerlerinin doğrudan kullanılması ve korelasyon katsayısı değeri yüksek çıkan istasyonların geliştirilen modelde kullanılması sonucu elde edilen OKH ve OMH değerlerindeki değişimi göstermektir. Korelasyon katsayısı değerleri yüksek çıkan istasyonların aynı verilere sahip olduğunu düşünmek hatalı bir yaklaşımdır. Elde edilen bu sonuçlar ileriki sayfalarda geliştirilen model ile karşılaştırılmalı olarak daha net açıklanacaktır. Şekil 5.17 'de Yakacık ve Akçatat istasyonlarının aynı zamana ait sıcaklık serileri gösterilmiştir.



Şekil 5.17 : Yakacık ve Akçatat istasyonlarının aynı zamana ait sıcaklık serileri.

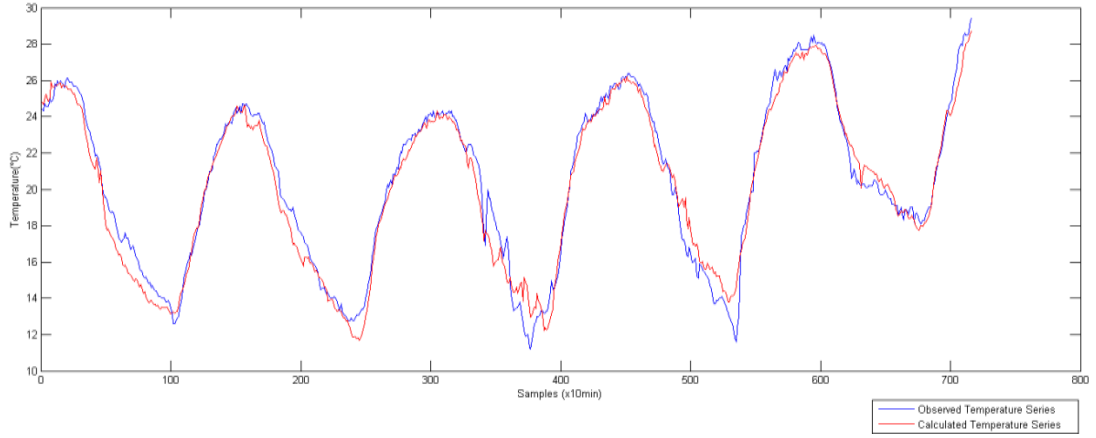
Şekil 5.18 'de Yakacık ve Poyralı istasyonlarının aynı zamana ait sıcaklık serileri gösterilmiştir.



Şekil 5.18 : Yakacık ve Akçatat istasyonlarının aynı zamana ait sıcaklık serileri.

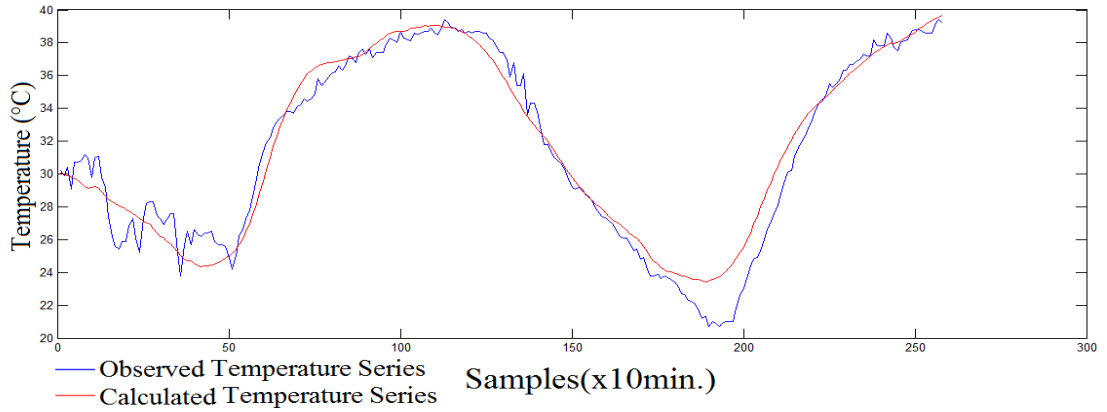
İkinci olarak, Akçatat ve Poyralı İstasyonlarına ait sıcaklık serileri doğrusal regresyon denklemi için bağımsız değişkenleri oluşturup, Yakacık istasyonuna ait sıcaklık serisi tahmin edilmeye çalışılmıştır.

Eğitim kümesindeki sıcaklık serileri kullanılarak bulunan, regresyon denklemine ait katsayı ve parametreler test kümesi üzerindeki veriler kullanılarak test edilmiştir. Hesaplanan ve gözlemlenen sıcaklık serilerine ait doğrusal regresyon denklemine ait grafik Şekil 5.19'da gösterilmiştir.



Şekil 5.19 : Regresyon modeli için eğitim kümesinde bulunan değerlerin test kümesinde test edilmesine ilişkin grafik.

Aynı şekilde eğitim kümesindeki sıcaklık serileri kullanılarak bulunan, regresyon+ dalgacık dönüşümü modeli kullanılarak hesaplanan katsayı ve parametreler test kümesi üzerindeki veriler kullanılarak test edilmiştir. Hesaplanan ve gözlemlenen sıcaklık serilerine ait regresyon+ dalgacık dönüşümü modeline ait grafik Şekil 5.20’de gösterilmiştir.



Şekil 5.20 : Regresyon+ dalgacık dönüşümü modeli için eğitim kümesinde bulunan değerlerin test kümesinde test edilmesine ilişkin grafik.

Üçüncü duruma ilişkin OKH ve OMH değerleri Çizelge 5.5’de verilmiştir.

Görüldüğü gibi gerek eğitim kümesinde gerekte test kümesinde OKH ve OMH değerlerinin her ikisinde birden düşme olmuştur. Bu değeri yüzde olarak ifade edecek olursak %7.82 oranında OKH değerinde, %2.53 oranında da OMH değerinde iyileşme olduğu söylenilebilir.

Çizelge 5.5 : Üçüncü durum için OKH ve OMH değerleri.

Doğrusal Regresyon Modeli	OKH	OMH
Eğitim Kümesinde	0,962	0,752
Test Kümesinde	2,0984	1,092

Doğrusal Regresyon Modeli+ Dalgacık Dönüşümü Modeli	OKH	OMH
Eğitim Kümesinde	0,7291	0,7
Test Kümesinde	1,5323	0,953

Yine ikinci duruma benzer şekilde, gözlemlediğimiz gibi her iki sonuçta elde edilen OKH ve OMH değerleri, Akçatat ve Poyralı istasyon değerlerinin Yakacık istasyon değerleri yerine kullanılması sonucu hesaplanan ayrı ayrı OKH ve OMH değerlerinden çok daha düşüktür. Bu durum geliştirilen modelin gerçeğe çok daha yakın sonuçlar ortaya koyduğunu, açıkça göstermektedir.

5.4 Dördüncü Durum

Oluşturulan dördüncü modelde ise, iki komşu istasyonun nem serileri, Yakacık İstasyon nem serisini tahmin etmek için kullanılmıştır. Bu istasyona göre komşu istasyon mesafesi yine 15 km ile 20 km olarak belirlenmiştir. Yani en yakın 15 km, en uzak 20 km olacak şekilde komşu istasyonlardan, Yakacık istasyonu ile korelasyonu değerine bakılarak, en iyi korelasyonu gösteren iki istasyon, komşu istasyon olarak, Akçatat ve Poyralı istasyonları belirlenmiştir. Toplamda 2396 adet sıcaklık verisi kullanılırken, bunlardan ilk 1198 adeti eğitim, kalan 1198 adeti ise test kümesi için ayrılmıştır. İlk kullanılan verinin tarihi, 09.05.2011, 21:30 iken, son kullanılan verinin tarihi 18.05.2011 ve saati 05:00 olarak alınmıştır. Yapılan çalışma sonucunda, Rumi and Aydogdu istasyonları komşu istasyonlar olarak seçilmiştir. Modelde yer alan tüm istasyonlara ayrık dalgacık dönüşümü uygulamada kolaylık sağlayan Ingrid Daubechies tarafından bulunan, çeşitli seviyelerde “Daubechies” dalgacığı kullanılarak, istasyonlardan elde edilen sıcaklık serilerine ayrık dalgacık dönüşümü uygulanmıştır. Her seviyede bir önceki seviyeye göre sahip olduğu veri sayısı yarıya düşerek ayrıştırılan sıcaklık serisi için Db5 ayrışım modeli diğer seviyeli ayrışım modellerine kıyasla en iyi sonucu vermiştir. Dördüncü durum için OKH ve OMH değerleri Çizelge 5.6’da gösterildiği gibidir.

Çizelge 5.6 : Dördüncü durum için OKH ve OMH değerleri.

Doğrusal Regresyon Modeli	OKH	OMH
Eğitim Kümesinde	20,5028	3,4997
Test Kümesinde	23,308	3,4757
Doğrusal Regresyon +Dalgacık Dönüşümü Modeli	OKH	OMH
Eğitim Kümesinde	17,9043	3,2996
Test Kümesinde	21,51	3,3353

Yine Çizelge 5.6'dan gözlemlendiği gibi OKH ve OMH değerlerinde azalma olmuştur. 7.72% oranında OKH ve 4.04% oranında OMH değerlerinde iyileşme olduğu gözlemlenmektedir.

5.5 Beşinci Durum

Bu çalışmada, Poyrali ve Akçatat İstasyonlarının nem ve sıcaklık değerleri kullanılarak Yakacık İstasyonunun nem serisi tahmin edilmeye çalışılmıştır. Toplamda 1198 adet sıcaklık verisi kullanılırken, bunlardan ilk 599 adeti eğitim, kalan 599 adeti ise test kümesi için ayrılmıştır. İlk kullanılan verinin tarihi 05.05.2011, 17:40 iken, son kullanılan verinin tarihi 18.05.2011 ve saati 05:00 olarak alınmıştır. Modelde yer alan tüm istasyonlara ayrık dalgacık dönüşümü uygulamada kolaylık sağlayan Ingrid Daubechies tarafından bulunan, çeşitli seviyelerde “Daubechies” dalgacığı kullanılarak, istasyonlardan elde edilen sıcaklık serilerine ayrık dalgacık dönüşümü uygulanmıştır. Her seviyede bir önceki seviyeye göre sahip olduğu veri sayısı yarıya düşerek ayrıştırılan sıcaklık serisi için Db3 ayrışım modeli diğer seviyeli ayrışım modellerine kıyasla en iyi sonucu vermiştir.

Bu bölümde kullanılan doğrusal regresyon denklemi dört bağımsız değişkene sahiptir (Denklem 5.7). Bu değişkenler komşu istasyonların nem ve sıcaklık serilerini ifade etmektedir. Buna bağlı olarak doğrusal regresyon + dalgacık dönüşümü modelinde de dört bağımsız değişkenli doğrusal regresyon modeli kullanılmıştır. Bu şekilde doğrusal regresyon denkleminin bağımsız değişken sayısında değişiklik yapmanın nedeni, doğrusal değişkenlerde farklı bir kombinasyon deneyerek başarılı bir sonuç elde edilebileceğini göstermektir.

Beşinci durumda kullanılan doğrusal regresyon denklemi, Denklem (5.10)'da gösterilmiştir.

$$y_i = a + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \beta_3 \cdot x_{i3} + \beta_4 \cdot x_{i4} \quad (5.10)$$

Db3, dalgacığı kullanılarak, bir ana bileşen 3 detay bileşenler için toplam 4 adet doğrusal regresyon denklemi dört bağımsız değişken içerecek şekilde kurulur. Kurulan denklemler sonucu hesaplanan ana ve detay bileşenler, ters dalgacık dönüşümü kullanılarak toplanır. Böylece tahmini nem serisi elde edilmiş olur.

Çizelge 5.7 : Beşinci durum için OKH ve OMH değerleri.

Doğrusal Regresyon Modeli	OKH	OMH
Eğitim Kümesinde	16,7777	3,1564
Test Kümesinde	38,6789	4,9781
Doğrusal Regresyon Modeli + Dalgacık Dönüşümü Modeli	OKH	OMH
Eğitim Kümesinde	15,6041	3,0174
Test Kümesinde	37,6228	4,9302

Çizelge 5.7’de gösterildiği gibi, OKH ve OMH değerlerinde gerek eğitim kümesinde gerekse test kümesinde iyileşme olduğu gözlemlenmiştir. İyileşme yüzdeleri ise, OKH için %2,73 iken OMH için %0,96’dır. Veri kümemizdeki nem değerlerinin yüzde olarak alındığını düşünürsek bu değerler oldukça iyi bir iyileşme yüzdesini vermektedir.

5.6 Altıncı Durum

Son olarak bu bölümde ise Yakacık istasyonunun sahip olduğu sonbahar, kış, ilkbahar, yaz sıcaklık serileri ile olan ilişkisi gözlemlenmiş, alınan örneklem aralıklarında sıcaklıklar arasındaki korelasyon katsayıları tespit edilmiştir. Kış dönemine ait sıcaklık serisi ile en yüksek korelasyon gösteren iki sıcaklık serisi kullanılarak yine kış dönemine ait sıcaklık serisi tahmin edilmeye çalışılmıştır. Bu çalışmada Db9 kullanılmış, geliştirilen modelin sonuçları, doğrusal regresyon modelinin sonuçları ile karşılaştırılmıştır.

Ayrıca her dönemin diğer dönemler ile olan korelasyon katsayıları da ayrı ayrı hesaplanmıştır. Sonbahar dönemi ile ilkbahar dönemi sıcaklık serileri arasında olan korelasyon katsayısı -0.2332, sonbahar dönemi ile yaz dönemi sıcaklık serileri arasında olan korelasyon katsayısı 0.3306, ilkbahar dönemi ile yaz dönemi sıcaklık serileri arasında olan korelasyon katsayısı 0.2506 olarak bulunmuştur.

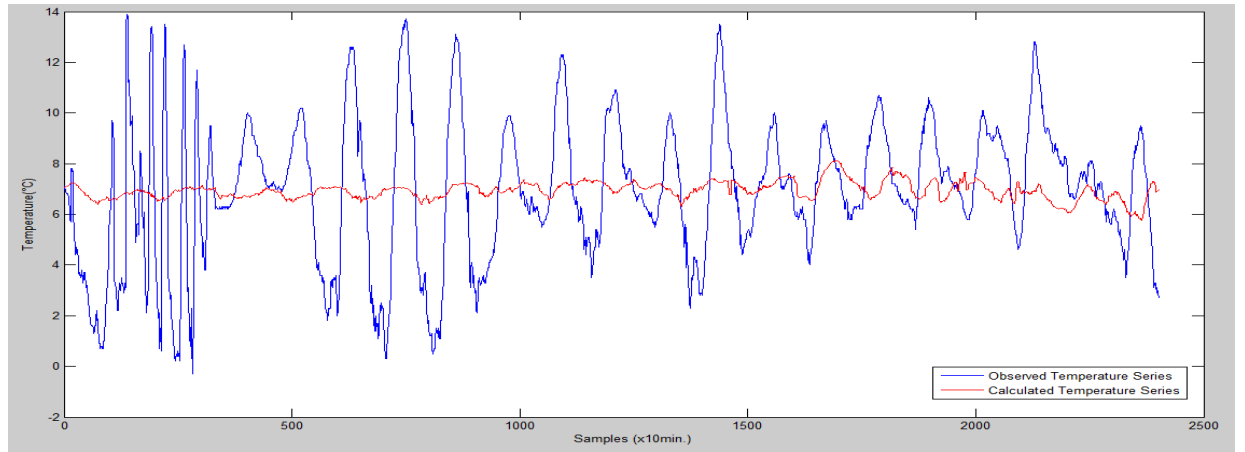
Çizelge 5.8’de sonbahar, ilkbahar, yaz dönemlerinden alınan 2402 adet örneklemin kış dönemine ait sıcaklık serisi ile aralarındaki korelasyon katsayıları gösterilmiştir.

Çizelge 5.8 : Sonbahar, ilkbahar, yaz dönemlerinden alınan sıcaklık serilerinin kış dönemine ait sıcaklık serisi ile olan korelasyon katsayıları.

	Kış Dönemi Sıcaklık Serisi ile olan Korelasyon Katsayıları
Sonbahar Dönemi Sıcaklık Serisi	0,0714
İlkbahar Dönemi Sıcaklık Serisi	0,0817
Yaz Dönemi Sıcaklık Serisi	0,2506

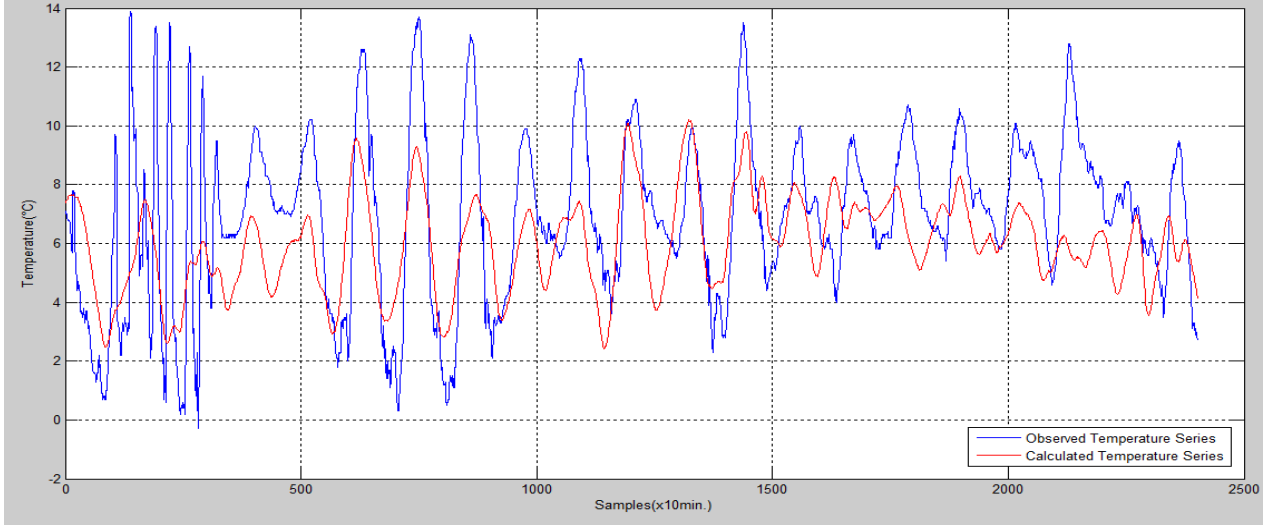
Doğrusal regresyon denkleminde ait a regresyon katsayısı, b_1 ve b_2 parametreleri değerleri ise sırası ile 5.0785, 0.0523 ve 0.0698 olarak bulunmuştur.

Şekil 5.21 altıncı durum için doğrusal regresyon kullanılarak hesaplanan ve bulunan sıcaklık serisine ait grafiği göstermektedir.



Şekil 5.21 : Altıncı durum için doğrusal regresyon modeli kullanılarak hesaplanan ve bulunan sıcaklık serisine ait grafik.

Şekil 5.22 altıncı durum için doğrusal regresyon model+ dalgacık dönüşümü modeli kullanılarak hesaplanan ve bulunan sıcaklık serisine ait grafiği göstermektedir.



Şekil 5.22 : Altıncı durum için doğrusal regresyon model+ dalgacık dönüşümü modeli kullanılarak hesaplanan ve bulunan sıcaklık serisine ait grafik.

Doğrusal regresyon modeli ve dalgacık dönüşümü+ regresyon modeline ait OKH ve OMH değerleri ise Çizelge 5.9’da verilmiştir.

Çizelge 5.9 : Altıncı durum için OKH ve OMH değerleri.

	OKH	OMH
Doğrusal Regresyon Modeli	7,6701	2,2038
Doğrusal Regresyon Modeli + Dalgacık Dönüşümü Modeli	6,1788	2,0187

Şekil 5.21 ve şekil 5.22 den de gözlemlendiği gibi geliştirilen model sonuçları oldukça iyileştirmektedir. Buna ek olarak Çizelge 5.8 OKH ve OMH değerlerindeki azalmaları açık bir şekilde göstermektedir.

6. SONUÇ

Çalışmada bir istasyonda bulunması amaçlanan eksik sıcaklık değerleri ve hatalı değerler için hem istasyona ait zirai meteorolojik veriler kullanılarak hem de çevresinde bulunan istasyonlarca alınan veriler kullanılarak öncelikle bir regresyon modeli oluşturmak amaçlanmıştır. Elde edilen hata katsayıları ve regresyon modelinden elde edilen değerler ile dalgacık dönüşümü kullanılarak elde edilen değerlerin karşılaştırılmıştır. Yapılan çalışmada, bir istasyon için, model olarak kullanılan merkez istasyona ait her 10 dk bir alınan zirai-meteorolojik verilerin, arasında zirai-meteorolojik veri ile en iyi korelasyonu gösteren veriler arasında korelasyon çalışması yapılmış, en yüksek korelasyonu gösteren veriler tespit edilmiş ve eksik veri tahmin edilmeye çalışılmıştır.

Elde edilen regresyon katsayıları ile oluşturulan sıcaklık değerlerine ait grafik ve gerçek zira-meteorolojik veri değerlerine ait grafik, matlab programı kullanılarak çizdirilmiştir. Geliştirilen model ile dalgacık dönüşümü tekniği, incelenen işaretlerin belirli zaman ve frekans aralıklarında gözlemlenen özelliklerinin hesaba katılmasını sağlayarak, veri tahmin sonuçlarının iyileşmesine olanak sağlamıştır.

Farklı veri kümeleri ve istasyonlar kullanılarak elde edilen sonuçlar doğrultusunda, Doğrusal Regresyon Modeli+ Dalgacık Dönüşümü Modelinin, Doğrusal Regresyon Modeline oran ile daha iyi sonuç verdiği gözlemlenmiştir. Birinci durumda iyileşme oranı test kümesinde iyi bir sonuç verirken, ikinci durumda, OKH iyileştirilmesi %26.98 iken OMH oranında iyileşme %12.79 olarak hesaplanılmıştır. Üçüncü durumda, OKH iyileştirilmesi %7.82, OMH iyileştirilmesi %2.53'tür. Dördüncü durumda OKH'da iyileştirme oranı %7.72 ve 4.04% OMH'da iyileştirme %4.04 tür. Beşinci durumda, OKH'da iyileştirme oranı, 2.73% iken OMH değeri 0.96% olarak belirlenmiştir. Yüzde oranlarında hesaplanan iyileşmeler gösteriyor ki doğrusal regresyon tekniği+dalgacık dönüşümü tekniği, doğrusal regresyon tekniğine oranla oldukça başarılı sonuçlar vermektedir. Ayrıca korelasyonun yüksek çıktığı istasyonların verilerinin eksik verisi bulunan istasyonun değerleri yerine yazıldığında hesaplanan hata oranlarının geliştirilen her iki modelden de oldukça yüksek olduğu

gösterilmiştir. Buna ek olarak, altıncı durumda farklı mevsimlere ait örneklemeler üzerinden gerçekleştirilen modelleme sonuçları yine geliştirilen modelin, farklı veri kümelerinde denendiğinde de hata oranlarını düşürdüğünü ispatlamıştır. Geliştirilen bu model, eksik yada bozuk dataaların yeniden yapılandırılması ile ilgili hesaplanan hata değerlerindeki gözönünde bulundurularak, kabul edilebilir bir yaklaşım sunar. Dalgacık dönüşümü tekniğinin doğrusal regresyon modeli üzerindeki iyileştirme etkisi, farklı araştırma konuları için bir kapı aralamış bulunmaktadır. Doğrusal regresyon+dalgacık dönüşümü modellemesi, özellikle uzun yıllar gözlemlenebilmiş veri serilerinin, gelecek yada geçmiş eksik değerlerinin tahmini, iyileştirilmesi, yada ekonomide, borsa dalgalanmalarının kestirilmesi gibi farklı alanlarda uygulanabilir.

Farklı istasyonlardan eş zamanlı eşit aralıklı meteorolojik verilerin alınarak, kullanılması ile gerçekleştirilen bu model benzer çalışmalardan farklı olarak:

- Dalgacık dönüşümü kullanılarak, zirai-meteorolojik verilerin hata teşhis ve tamiri için regresyon tekniğinin iyileştirilmesi ile başarılı bir sonuç elde edilmiştir.
- Geliştirilen model, otomatik bir sistem haline getirilerek, meteorolojik verilerin tahmini ve tekrar yapılandırılmasında kullanılabilir.
- Ülkemizin çeşitli yerlerinde bulunan tarımsal alanlara sanal istasyonlar kurularak, tarım bölgelerine ait meteorolojik veriler, kurulan tek bir merkez ile tahmin edilebilir ve sistemin işleyişi kontrol edilebilir. Böylece, insan gücüne gerek kalmaksızın, uzak ve insan barınımı için elverişli olmayan doğal ortamlarda, meydana gelen hava olaylarının takip edilmesi çok daha kolay bir hal almış olacaktır.
- Sonuç olarak, bu çalışma ile gerçekleştirilen amaçlar, yeni araştırmaların yapılması ve tarım alanında ilerleme kaydedilmesi için oldukça önemli birer adım olmaktadır. Ayrıca geliştirilen model, geliştirilmesi planlanan tarla sulama çiftçi uyarı sistemi için ilgili meteorolojik verilerin en doğru şekilde yapılandırılması, sulama sistemi çıktılarının da en doğru şekilde hesaplanmasına imkan tanıyacaktır. Bu durum, dolaylı olarak ülkemizde gerçekleştirilmesi planlanan tarımsal ürün miktarındaki verimin artırılması, rekolte tahmini, ürün çeşitliliğini artırılması, en uygun yer için en uygun ürün çeşidinin seçilmesi gibi insan yaşamını birebir etkileyen konularda

önemli bir rol oynayacaktır. Özetle, tez çalışması olarak yapılan bu çalışmanın, hayata geçirilmesi ve ilgili tarımsal alanlardaki çalışmalar için kullanılması planlanmaktadır.

KAYNAKLAR

- [1] **YILDIZ, Nazmi ve KÖKTÜRK, Erol**, 1984, “Türkiye’de Kırsal Toprak Düzenlemeleri”, Harita ve KadastroMühendisliği, Sayı: 48-49, İlkbahar-Yaz, s: 98-119.
- [2] **Çakmak, B.ve Aküzüm, T.**, 2006, Türkiye’de Tarımda Su Yönetimi, Sorunlar ve Çözüm Önerileri, TMMOB Su Politikaları Kongresi, 21-23 Mart 2006, Ankara, Türkiye.
- [3] **Akar, M., Silay Erdenir, A., Akkaya, H., Tomar, Ahmet**, 2009, Sulama Araç, Yöntem ve Organizasyonlarının Geliştirilmesi, Türkiye Ziraat Mühendisliği VII. Teknik Kongresi, 7-10 Ocak 2009, Ankara, Türkiye.
- [4] **Travaglia, C., Milenova, L.,FAO, Nedkov, R.,Vassilev, V., Milenov, P., Radkov, R., Pironkova, Z., BASA**, 2001, Preparation of Land Cover Databae of Bulgaria through Remote Sensing and GIS, Rome.
- [5] **Pan, G.,Li, F.,and Sun, G.**, 2007, Camera Based Measurement of Crop Cover for Wheat Yield Prediction,Geoscience and Remote Sensing Symposium 2007. IGARSS 2007.
- [6] **İTÜ-UHUZAM**, Tarım Rekolte Tahmini ve Kuraklık İzleme Projesi, TARİT HİZMET Raporu-8s.
- [7] **İTÜ-UHUZAM**, Tarım Rekolte Tahmini ve Kuraklık İzleme Projesi, TARİT HİZMET Raporu-52s.
- [8] **İTÜ-UHUZAM**, Tarım Rekolte Tahmini ve Kuraklık İzleme Projesi, TARİT HİZMET Raporu-54s
- [9] http://tr.wikipedia.org/wiki/Regresyon_analizi
- [10] **Karataş, B.S., Akkuzu, E. ve Avcı, M.**, 2006, *Uzaktan Algılama Tekniği ile Evapolasyonun*, 4. Coğrafi Bilgi Sistemleri Bilişim Günleri, 13 – 16 Eylül 2006 ,Fatih Üniversitesi / ,İstanbul,Türkiye.
- [11] http://tr.wikipedia.org/wiki/Regresyon_analizi.
- Bilgin, S.**, 2008, Kalp Hızı Değişkenliğinin Dalgacık Dönüşümü ve Yapay Sinir Ağları Kullanarak Analizi, Doktora Tezi, Sakarya Üniversitesi, 2008.
- Bilgin, Recai.**, 1997. “Türkiye’de su sorunları ve çözüm önerileri”. Meteoroloji Mühendisliği TMMOB Meteoroloji Mühendisleri Odası Yayın Organı. Sayı 2: 18.
- Brisco, B., Brown, R. J.**, 1995. Multidate SAR/TM Synergism for Crop Classification in Western Canada. Photogrammetric Engineering & Remote Sensing. Pp. 1009-1014.

- Erdođmuş, P.ve Pekçakar, A.,** 2009, Dalgacık Dönüşümü ile EKG Sinyallerinin Özellik Çıkarımı ve Yapay Sinir Ağları ile Sınıflandırılması, 5. Uluslararası İleri Teknoloji Sempozyumu(IATS'09),13-15 Mayıs 2009, Karabük, Türkiye.
- EROL, H.,** 2006. ISB539 Bilgisayar Destekli İstatistiksel Yöntemler I. Ders Notları,Ç.Ü. Fen ve Edebiyat Fakültesi, İstatistik ABD, Adana. (yayınlanmamış).
- Gonzales, J., Barry, M., Johnson, J., Lackowski, H., Landrum, V., Maus, P.,** 1992. Vegetation Classification and Old-Growth Modelling in the Jemez Mountains. USDA Forest Service Nationwide Forestry Applications Program. Salt Lake City, Utah U.S.A.
- Güngör, Y., Erözel, A.Z. ve Yıldırım, O.,** 1996. Sulama, Ankara Üniv. Ziraat Fak Yayınları No: 1443, Ankara, 295s.
- Kanber, R.,** 1997. Sulama, Çukurova Üniv. Ziraat Fak Yayınları No: 174, Adana, 530s.
- Kumar, P., Fofoula- Georgion, E.,** 1193, A Multi Component Decomposition of Spatial Rainfall Fields, 1. Segregation of Large and Small Scale Features Using Wavelet Transform Water Resources Research 29(8), 2515- 2532.
- Küçük, M.,** 2004, Dalgacık Dönüşüm Tekniđi Kullanarak Akım Serilerinin Modellenmesi, Doktora Tezi, İstanbul Teknik Üniversitesi 2004.
- ÖZGÜLER, H.,** 1997, "Su, su kaynakları ve çevresel konular" Meteoroloji Mühendisliđi. *TMMOB Meteoroloji Mühendisleri Odası Yayın Organı* Sayı 2: 57-63.
- Partal T.,** 2007, Yađış Verilerinin Yapay Sinir Ağları ve Dalgacık Dönüşümü Yöntemleri ile Tahmini, Doktora Tezi, İstanbul Teknik Üniversitesi 2008.
- Ritchie, J. T., U. Singh, D.C. Godwin, W.T. Bowen.,** 1998. Cereal Growth, Development and Yield. In Understanding Options for Agricultural Production. G. Y. Tsuji, G. Hoogenboom, and P. K. Thornton (Editors), Kluwer Academic Publishers, Dordrecht, The Netherlands, p. 79-98.
- Russel, G., Ballogh, M., Bell, C., Green, C., Milliken, J. A., Ottoman, R.,** 1998. Mapping and Monitoring Agricultural Crops and other Landcover in the Lower Colorado River Basin. Photogrammetric Engineering & Remote Sensing. Vol. 64. No.11. Pp 1107- 1113.
- Şahin, M.,** 2007, Karadeniz Bölgesindeki Yađış-Akış İlişkinin Farklı Yapay Sinir Ağları Metotlarıyla Belirlenmesi, Lisans Tezi, İstanbul Teknik Üniversitesi 2007.
- Tanođlu, A.,** 1942. Ziraat Hayatı I, Orta %klim Bölgelerinde Ziraat. İst. Üniv., Ed. Fak. Cog. Enst. Nes.
- Taşcıkaraođlu, A. ve Uzunođlu, M.,** 2011, DALGACIK DÖNÜŞÜMÜ VE YAPAY SİNİR AĞLARI İLE RÜZGAR HIZI TAHMİNİ, Fırat Üniversitesi-Elazığ.

Tümertekin, E., 1957. Kurak Bölgelerde Ziraat. İst. Üniv. Yay. No: 713, İstanbul

Übeyli, E., Güler, İ., 2004, Dalgacık Dönüşümü ile EEG İşaretlerinden Çıkarılan İöznitelik Vektörleri Üzerinde İstatiksel İşlemlerin Gerçekleştirilmesi, ELECO'2004 Elektrik - Elektronik – Bilgisayar Mühendisliği Sempozyumu ve Fuarı Bildirileri,8-12 Aralık 2004.

ÖZGEÇMİŞ

Ad Soyad: N. Tuğbagül ALTAN

Doğum Yeri ve Tarihi: Erzurum 29.01.1985

Lise: Bursa Anadolu Lisesi

Lisans Üniversitesi: Doğu Üniversitesi Bilgisayar Mühendisliği Bölümü

Yüksek Lisans Üniversitesi: İstanbul Teknik Üniversitesi Bilgisayar Bilimleri Bölümü