WAVELET DECOMPOSITION AND ITS APPLICATIONS IN AVIATION INDUSTRY

A Master's Thesis

by
MELİH AKİF GÜRBÜZ

Department of Economics
Bilkent University
Ankara
September 2010

WAVELET DECOMPOSITION AND ITS APPLICATIONS IN AVIATION INDUSTRY

The Institute of Economics and Social Sciences
of
Bilkent University

by

MELİH AKİF GÜRBÜZ

In Partial Fulfillment of the Requirements for the Degree of
MASTER OF ARTS

in

THE DEPARTMENT OF
ECONOMICS
BILKENT UNIVERSITY
ANKARA

September 2010

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Assoc. Prof. Dr. Çağla Ökten
Supervisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Assoc. Prof. Dr. Ümit Özlale
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Prof. Dr. Hakan Berument
Examining Committee Member

Approval of the Institute of Economics and Social Sciences

Prof. Dr. Erdal Erel
Director

ABSTRACT

WAVELET DECOMPOSITION AND ITS APPLICATIONS IN AVIATION INDUSTRY

Melih Akif Gürbüz

M.A., Department of Economics

September 2010

This thesis analyzes the characteristic air traffic time series data of European aviation industry by using wavelet decomposition. Firstly, time frequency representation and discrete time wavelet decomposition is introduced by giving some theoretical background and basic definitions. Then, the attributes of AEA data, i.e. load factor, available seat kilometer and revenue passenger kilometer is analyzed both geographically and based on prominent Airline operators. Furthermore, their relative comparisons have led to unravel some underlying implications and their distinctive interpretations for the airline industry.

Keywords: Wavelet Decomposition, Aviation Industry, AEA, Load Factor

ÖZET

DALGACIK DÖNÜŞÜMÜ VE HAVACILIK ENDÜSTRİSİNDEKİ UYGULAMALARI

Melih Akif Gürbüz

Yüksek Lisans, İktisat Bölümü

Eylül 2010

Bu tez, Avrupa Havacılık Endüstrisi'nin zaman serilerini kullanarak hava trafiğinin özelliklerini analiz etmektedir. Öncelikle, zaman frekans gösterimi ve aralıklı zaman dalgacık dönüşümü teorik altyapısı ve bazı basit tanımlamalarla tanıtılmaktadır. Daha sonra, AEA veri özellikleri incelenerek, doluluk oranı, arz edilen koltuk kilometre ve ücretli yolcu kilometre verileri analiz edilmekte ve bu değerler farklı coğrafi bölgeler ve havayolları arasında karşılaştırılmaktadır. Ayrıca, bu karşılaştırmalar havayolu işletmeleri için bazı temel ilişkileri ve ayırt edici özellikleri ortaya çıkarmıştır.

Anahtar Kelimeler: Dalgacık Dönüşümü, Havayolu Endüstrisi, AEA, Doluluk Oranı

ACKNOWLEDGEMENTS

I would like to express gratitude to my supervisors Professor Ümit Özlale and Çağla Ökten for their guidance, encouragement and support.

I thank Professor Hakan Berument for valuable suggestions and criticism.

I am grateful to Aytaç Erdemir and Çağrı Akkoyun for their friendship, assistance and support.

I would like to convey thanks to TUBİTAK-BİDEB for providing financial support during my studies in M.A. program.

I am also thankful to Meltem Sağtürk for her kindly assistance during M.A. program.

I am thankful to my dear family, for their support and trust.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The everlasting evolution of Airline industry compels us to analyze and disclose the patterns, trends, and changing dynamics of the industry in order to optimize and enhance the business strategies and discover the hidden correlations and/or contradictions among different factors.

Today, the airline industry is a highly competitive market which is subject to shocks such as 9/11 and SARS epidemic etc., and changes in macroeconomic and business alterations, such as GDP growth in different regions, oil price fluctuations, or mergers between Airline operators, along with inherent seasonality patterns dominating the decision making processes.

Up until the mid 1980s, the airline industry was dominated by extensive government control and intervention. Alterations in fares, supervising routes and entry into market were heavily regulated by governments with typical characteristic of state-owned flag carriers. Although, bilateral agreements still characterize air transport between EU and non-EU destinations, deregulation of the European airline industry and extensive privatizations stimulated the growth of the airline industry, and boosted the dynamism of the market globally. According to ICAO, in terms of total international passenger kilometers by region, Europe is the largest region with 40%.

The competition in the airline industry, which have grown with the destinations offered and total passengers since the initiation of deregulation and liberalization of the European market in the mid 1980s, have led to an increase in global air traffic significantly, along with the seat capacity, the total market shares, low cost carrier competition, and trends to create larger units via mergers and acquisitions.

For airline companies, fuel is the primary component of costs, affecting the profitability of the carriers. Sudden exogenous shocks, such as the terrorist attacks of September 11th in 2001, have led to instant extensive changes in airline capacity usage, and mostly resulted with lowering fleet size and reducing employment. Therefore, characterizing these exogenous shocks, analyzing their effects, constitute a vital importance for reactive and pre-emptive decisions, strategies and tactics.

The conventional methods aiming to investigate how industry evolves along with the changing parameters, cannot become fully effective or efficient due to the fact that the patterns and what these patterns or time series in reality represent and how they are related with the corresponding parameters or factors cannot be completely solved because of the non-stationary and complex attributes of the time series data in various forms.

Wavelets are mathematical functions that help to map data from the time domain into different layers of frequency levels. Wavelets have the advantage of being localized both in time and in the frequency domain, and enable to observe and analyze data at different scales compared to the standard Fourier analyses. Their theoretical basis was completed by the late 1980s, and in the 1990s they spread rapidly to a wide range of applied sciences. A number of successful applications show that wavelets are on the edge of entering mainstream econometrics.

In the aviation industry, there is large amount of non stationary and complex time series data to deal with in order to analyze the performance of the airline companies to make strategic decisions. Wavelet analyses are powerful tool to decompose the time series data of aviation industry in order to filter seasonality, determine the shocks and measure their impacts to the airlines.

CHAPTER 2

WAVELET DECOMPOSITION

2.1. Introduction

Mathematically, wavelets are local orthogonal bases consisting of small waves that decompose a function into layers of different scale. Wavelet theory is shares the roots with the Fourier analysis, but there are essential differences. In order to represent a given function, the Fourier transformation uses a combination of sine and cosine functions at different frequencies. However, sine and cosine functions are periodic functions that have the stationary characteristics through the real line. Hence, any change at a particular point of the time domain has an impact that is felt over the entire real line. In practice, this structure of the Fourier transform indicates the assumption of the frequency content of the function is stationary along the time axis. To overcome this problem, physicists invented the windowed Fourier transform. The data are cut up into several intervals along the time axis and the Fourier transform is taken for each interval separately.

With the introduction of windowed Fourier transform, the era of time frequency representations (TFR) has began. The logic behind the TFR is to transform data into time frequency domain to observe time varying frequency characteristics. Different types of representations are used for different purposes. However, perfect resolution cannot be obtained both in time and frequency domain simultaneously. This restriction is formulized by Heisenberg's uncertainty principle. Then researchers tried to optimize this tradeoff between time and resolution.

Unlike the Fourier transform, wavelets are defined over a finite domain; they are localized both in time and in scale (see Figure 1). They provide an efficient and appropriate way of representing complex signals. More importantly, wavelets can sort out data into different frequency components for individual analysis. Wavelets enable us to see both the whole of a digital image and the pixels of it. This scale decomposition opens a whole new way of processing data.

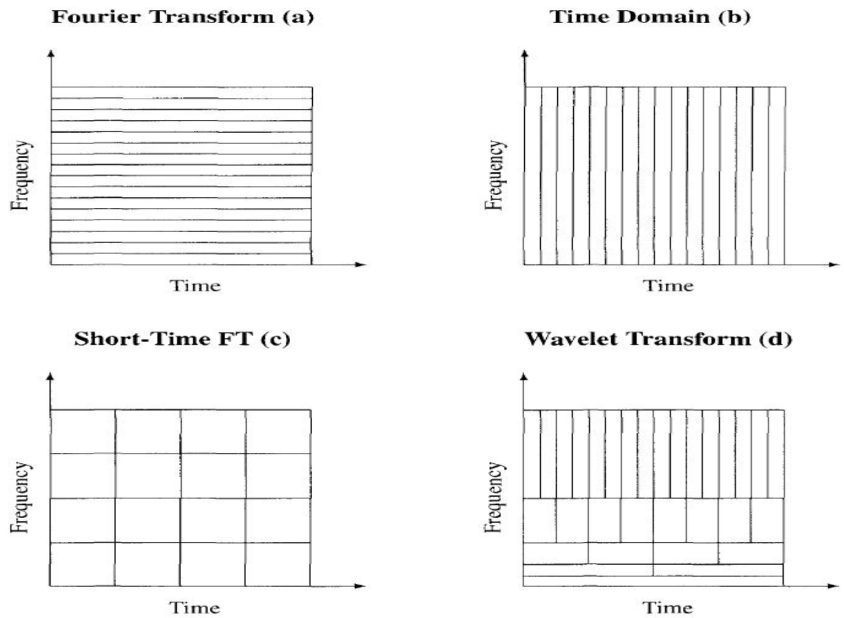


Figure 1.: Partitioning of the time-frequency plane by different techniques.
(a)The frequency domain after computing the Fourier transform, representing perfect frequency resolution and no time resolution. (b) The time domain representation of the observed time series, representing perfect time resolution and no frequency resolution. (c) Represents balanced resolution between time and frequency by using the short-time Fourier transforms (Gabor transform). (d) The wavelet transform adaptively partitions the time-frequency plane. (Gencay et al.)

A wavelet basis consists two main components; a father wavelet that represents the smooth baseline trend and a mother wavelet that is dilated and shifted to construct different levels of detail. This is similar to the structure of a natural organism that is based on self-similarity. At low scales, wavelets capture long-run phenomena, on the other side, at high scales; the wavelet has a small time support, enabling it to zoom in on details like spikes, cusps, and on temporary appearances such as error terms. Their ability to adapt their scale and time support enables them to optimize subject to Heisenberg's uncertainty principle.

By the late 1980s, theoretical basis of wavelets were completed, and wavelets began to enter the applied sciences. One of their first applications was in earthquake prediction. Wavelets provided a time dimension to non-stationary seismic signals that Fourier analysis lacked. Wavelets are now applied in a wide range of fields, from fractals and partial differential equations in mathematics to signal and image processing, speech recognition, software design, engineering, meteorology, and statistics.

The similarities of wavelets to the Fourier transform make them an ideal candidate for frequency domain analysis in time-series econometrics. Additionally, their ability to capture long term movements and high-frequency details simultaneously is very practical when dealing with non-stationary and complex functions.

2.2. Theoretical Expression of Wavelets

2.2.1. Fourier Transform

In 1807, the French mathematician Joseph Fourier developed a representation of series that can produce the generalized solutions to differential equations. These series are called Fourier series. The main idea behind the construction of these series is, any 2π periodic function could be represented by a combination of sines and cosines. The Fourier series of *f(x)* and its coefficients are given by;

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cdot \cos(kx) + b_k \cdot \sin(kx))$$

$$a_k = \int_0^{2\pi} f(x)\cos(kx)dx \text{ and } b_k = \int_0^{2\pi} f(x)\sin(kx)dx.$$

By using the Euler's formula, Fourier transform is given by;

$$F(\omega) = \int_{-\infty}^{\infty} f(x)e^{-iwx}$$

Where ω is the angular frequency, and *exp(iwx)= cos(wx)+ isin(wx).*

Fourier transforms plays a critical role in many fields of applied and pure science. For example HP filter, Baxter and King Filter. However, a quest for new mathematical structures that enables continuous functions to localize in both the time and the frequency domain. Although economy deals with discrete time signals, literature on the continuous time signals is reviewed to provide compact understanding of time frequency approach.

2.2.2. Short Time Fourier Transform (STFT)

The formula for the short time Fourier transforms:

$$F_x(t,f;h) = \int_{-\infty}^{+\infty} x(u)h^*(u-t)e^{-j2\pi fu}du$$

Where h(t) is a short time analysis window. The STFT calculates the frequency transformation of the windowed signal for all time instants. If the length of the window is increased i.e. time resolution is decreased, then frequency resolution increases too.

The formula for the inverse short time Fourier transforms:

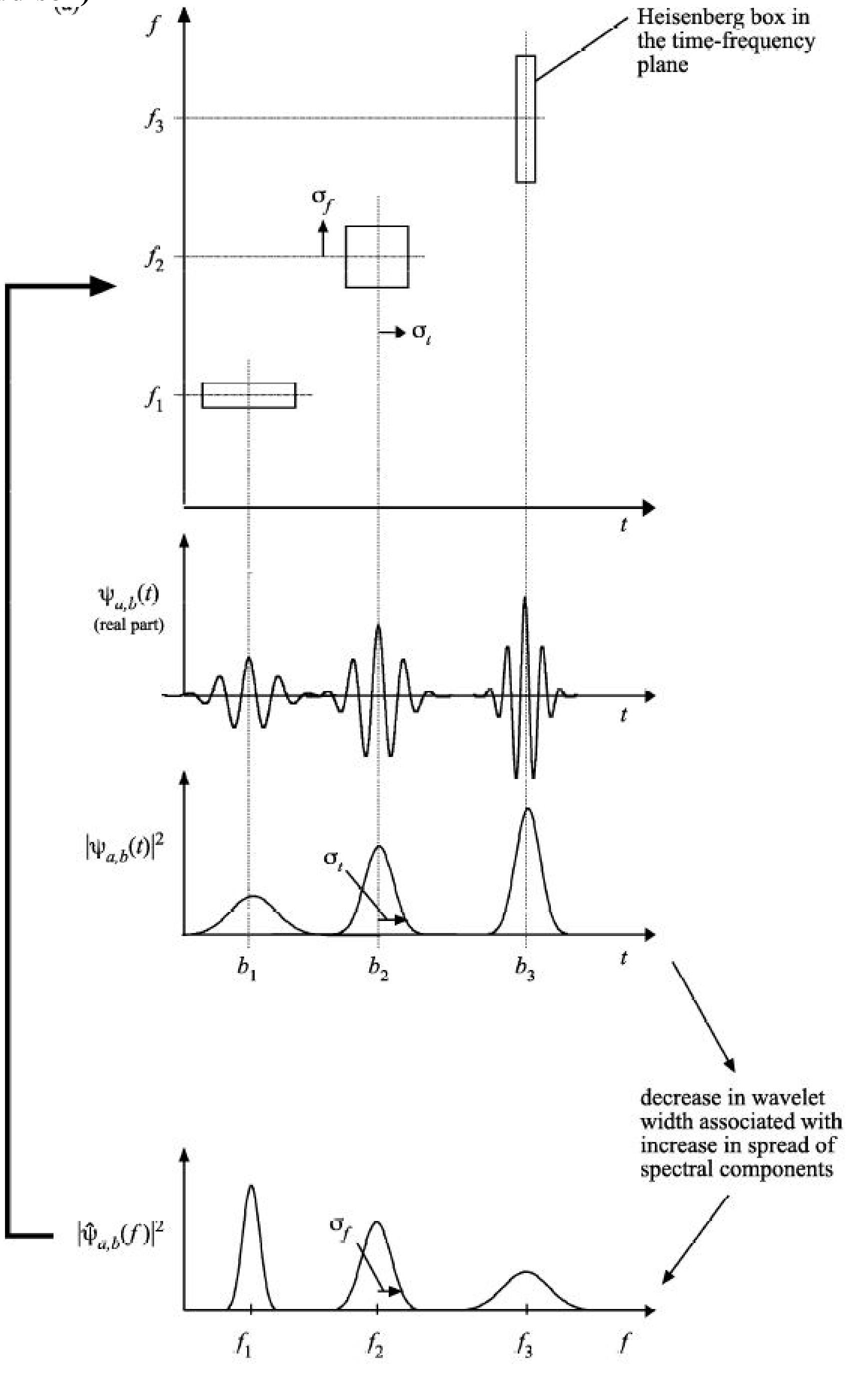
$$x(t) = \frac{1}{E_h} \iint_{-\infty}^{+\infty} F_x(u, \xi; h) h(t-u) e^{j2\pi t \xi} du d\xi$$

Where $E_h = \int_{-\infty}^{+\infty} |h(t)|^2 \, dt$ is the energy of h(t). This is a useful property since; one can filter the noise and unwanted components in time frequency domain and transfer the remaining part from time frequency domain to time domain.

2.2.3. Wavelet Transform (WT)

Unlike the Fourier transform and the STFT which are functions of frequency, the wavelet transform is a function of scale and in a wavelet transform the scale depends on frequency. A new set of base functions is necessary to deal with the fixed time-frequency partitioning. The wavelet transform uses a mother wavelet function to capture features that are local in time and local in frequency, by dilating and transforming it. To overcome the problem hypothesized by Heisenberg uncertainty principle, the wavelet transform reasonably adapts itself to represent attributes across a wide range of frequencies. Roughly, scale and frequency interval are inversely proportional. When the scale parameter increases, the wavelet basis gets longer in the time domain, shorter in the frequency domain, and closer to lower frequencies. On the contrary, when the scale parameter decreases, the time support declines, the number of frequencies captured grows, and the wavelet basis shifts toward higher frequencies. The wavelet transform enables to capture events that are local in time by giving up some frequency resolution. The wavelet transform maintains the balance between time and frequency, which makes them an ideal tool while studying non-stationary and complex time series.

Figure 2. Heisenberg Boxes in the time-frequency Plane for a wavelet at various scales. (Addison)



Wavelets in general are functions like;

$$\psi_{a,b}(t) = |a|^{-1/p}\,\psi\left(\frac{t-b}{a}\right),\ p>0,\ a,b\in\mathbf{R},\ a\neq 0$$

With 2 being the most frequently used value of p. So the orthonormal basis form of;

$$\psi_{j,k}(t) = 2^{j/2}\psi\left(2^{j}t-k\right),\ j,k\in\mathbf{Z}.$$

2.2.4. Basic Properties of Wavelet Transform

Some of the basic properties of the wavelet transform are as follows; the continuous wavelet transform of a function f (t) is defined as the convolution of f(t) with an analyzing function ψ(η). The wavelet function must be localized both in time and frequency space and, for an integrable function, its average should be zero. For η = (t'− t)/s, the wavelet transform is given by;

$$W(t,s) = \int_{-\infty}^{\infty} f(t')\frac{1}{\sqrt{s}}\psi^{*}\left[\frac{(t'-t)}{s}\right]\mathrm{d}t',$$

Where s is the wavelet and t is time scale. The normalization condition needs to be satisfied by the factor 1/√s;

$$\int_{-\infty}^{\infty}\psi\left[\frac{(t'-t)}{s}\right]\psi^{*}\left[\frac{(t'-t)}{s}\right]\mathrm{d}t' = s.$$

As for the discrete time series (which by the way is the type of all the data we have been dealing with), let *Xn*, of *N* observations with a time modulation of *δt*, the integral of the first equation is replaced by a summation over the *N* time steps. Thus, wavelet transform of a discrete time series is given by;

$$W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \sqrt{\frac{\delta}{s}} \psi^* \left[\frac{(n'-n)\delta t}{s} \right]$$

As s is the wavelet scale and n provides the time variation. The wavelet transform is constructed on alteration of s and n, providing both the existence of the frequency component and their related time localization. (Moortel at al)

2.2.5. Choosing a Wavelet

As stated by Mallat, choosing a wavelet depends on the support size of the mother wavelet function and the number of vanishing moments. Because efficient approximation, noise removal and fast running algorithms depends on this property. Orthogonal wavelets imply that a mother wavelet (ψ) has p vanishing moments then its support is at least of size 2p – 1. In this respect Daubachies and Symmlet wavelets are optimum due to their minimum size support for the relative number of vanishing moments. Hence, for optimal time series representation, Symmlet (8) wavelet is used in the analysis.

CHAPTER 3

ANALYSIS OF AVIATION DATA SERIES

3.1. Association of European Airlines (AEA)

The Association of European Airlines (AEA) is the representative body of 35 European airline operators which are mostly traditional large carriers (National Flag Carriers) operate both in Europe and intercontinental level. Most of the low cost airlines are not members of AEA resulting that they are not adequately represented in this organism. AEA follows the relevant EU developments and informs its members about the consequences, and coordinates lobbying and communication efforts related to EU regulations. AEA also provides market intelligence and analyses based on submissions from member airlines' data covers scheduled passenger and capacity traffic and freight flows by major geographical area centered in Europe.

They have had to adjust to a new commercial environment with the implementation of competition and deregulation directives resulting from the European Union's Single Market Program (SMP), which was established in 1992. They also have to face the low cost revolution from the mid 1990 onwards. In 1997, the EU introduced competition into the previously protected national airlines market, promoting efficiency through competition. Moreover, the rising price of aviation fuel has increased the pressure on airlines to achieve more efficient procedures.

3.2. Data Definitions

Airline payload, in other words revenue-earning traffic, essentially consists of passengers, freight and mail. For AEA airlines, revenue from these three sources amounted in 2005 to 86.7%, 12.7% and 0.6% respectively of total operating revenue. Hence passenger traffic can be taken as a main indicator for the performance of airline industry.

Passenger traffic is measured in passenger boardings, or more commonly passenger-kilometers, which are calculated by multiplying the number of passengers by the distance each one flies. This statistic is normally referred to as Revenue Passenger-Kilometers, or RPK, since non-revenue traffic, for example airline staff travelling on duty, is not included.

Passenger capacity is normally measured either as Available Seat-Kilometers (ASK), calculated by multiplying the number of saleable seats by the sector distance. An aircraft's overall carrying capacity is measured in Available Ton-Kilometers. It is not normal practice to measure cargo capacity because this will vary from flight to flight depending on the passenger load. It is also very sensitive to cargo density – passenger aircraft will normally run out of cargo space before they reach their weight limits.

The relationship between traffic and capacity is called load factor (LF) and is of critical importance to airlines. For the AEA airlines collectively, a single point of load factor gained or lost, over the course of a year, is worth about $1.2 billion.

Freight and mail traffic, similarly, may be measured in tones, but more usually ton-kilometers. Freight and mail together constitute cargo, which is sometimes presented as a single statistic.

The entire payload may be expressed as Revenue Ton-Kilometers by multiplying the number of passengers a notional weight (which includes by their baggage) and adding it to the cargo traffic before making the distance calculation.

3.3. Seasonality in Airline Business

Business travel, short-haul leisure travel and long-haul leisure travel all have different seasonal profiles.

Airlines operate according to distinct summer and winter timetables, between April-October and November-March respectively, although there is fine-tuning of operations within each season. Measured in seat-km, the level of activity in summer is about 10% above winter levels.

Air traffic is more volatile than the available seat kilometers. As an example, when 2005 data is analyzed, passenger boardings in the busiest month (September), was almost 37% higher than the number in January which is the quietest month. In terms of passenger-km, the busiest month (July) was 29% higher than the quietest (February).

3.4. AEA Data Analyses

There are two types of data sets. First of them is the monthly data starting from 01/1991 and ends at 02/2010 consisting of 230 observations for PLF, ASK, RPK, TFTK. Monthly data is available for different geographical regions given in the Table 1, there is also categorization for international and domestic traffic.

Table 1. Geographical Codes for the Data

DO - Domestic
Domestic traffic is defined as traffic carried on routes originating and terminating within the boundaries of a State by an air carrier whose principal place of business is in that State, or on routes between the State and territories belonging to it, or ;
EU - Cross-border Europe
Includes all cross border/ international routes originating and terminating within Europe (including Turkey and Russia up to 55ºE), Azores, Canary Islands, Madeira and Cyprus.
ET - Europe Total
The sum of Cross-Border Europe and Domestic.
IE - International Short/Medium Haul
The sum of Cross-Border Europe (no Domestic), North Africa and Middle East.
SA - South Atlantic
Any scheduled service between Europe and North, Central or South America via gateways in, or South of, Brazil (i.e. Argentina including the Falkland Islands, Brazil, Chile, Paraguay and Uruguay).
AE - Europe-Far East Australasia
Services between Europe and points east of the Middle East region, including Trans-Polar and Trans-Siberian flights. All traffic on such services is reported, including local traffic in Europe and the Middle East region.
IC - Total Longhaul
The sum of North-, Mid- and South Atlantic, Far East/Australasia, Sub Saharan Africa and Other. OT (Other) is not shown in Traffic Update publication.
IT - Total International
The sum of Cross-Border Europe, North Africa, Middle East, North Atlantic, Europe Far East-Australasia, Europe-Sub Saharan Africa, South Atlantic and Mid Atlantic.
TO - Total Scheduled
The sum of Total International and Domestic

3.4.1. Passenger Load Factor

In the data set of AEA since there are different types of data such as available seat kilometers (ASK), revenue passenger kilometers (RPK), passenger load factor (PLF) total freight ton kilometers (TFTK), available ton kilometers (ATK), and revenue ton kilometers (RTK), PLF is the most critical indicator for the airlines.

Some of the graphs for PLF are given in figures 3 and 4. Level of PLF is a critical indicator for airlines, because PLF is highly related to efficiency, i.e. due to the fixed costs of each flight, as PLF increases, cost per passenger decreases. Hence profitability and number of passengers rise together. There is a breakeven point of PLF that, for a specified level of ticket prices, total cost for a flight (sum of variable and fixed costs per flight) is equal to the total revenue (PLF*Capacity* Ticket Price). PLF is the dependent variable in this equation. All the other variables (costs, capacity, and ticket price) are predetermined by the airlines in the short term. Therefore PLF is the outcome of inputs determined by the airline and determines profitability.

On the other side high volatility in PLF is undesirable for the airlines. Sudden or seasonal drops in PLF cause some of the flights to lose money. In the long term, optimizing the capacity and ticket prices is dependent to the PLF estimation. With accurate PLF estimation, it is possible to make preemptive moves by changing capacity and price in order to keep PLF at a specified level. Therefore volatility of PLF will decrease by these strategic moves. Hence in a time series graph of PLF, increasing trend and decreasing volatility indicate a wisely managed capacity.

Figure 3. Load Factor - Total Scheduled - Monthly

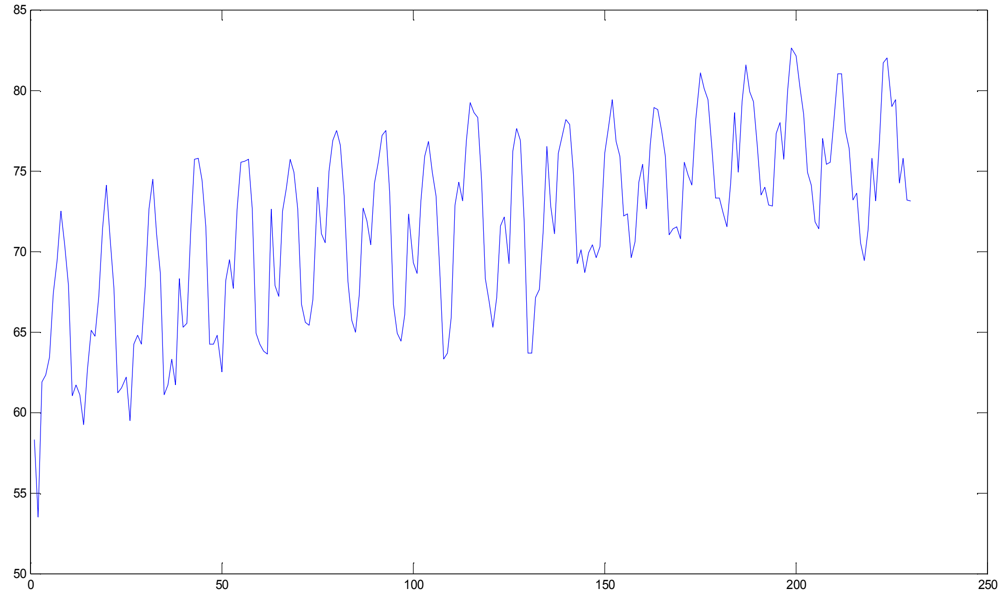
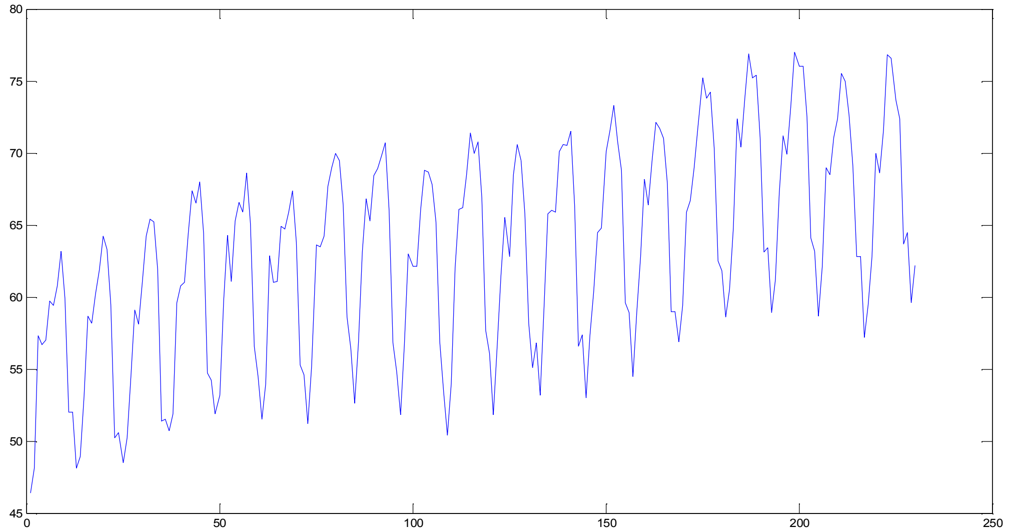


Figure 4. Load Factor - Crossborder EU - Monthly



AEA data is available in airline specific from 07/1999 to 02/2010. For this reason, in comparisons that include monthly airline specific data, data range is narrowed to 128 points.

Figure 5. Load Factor - Lufthansa - Monthly

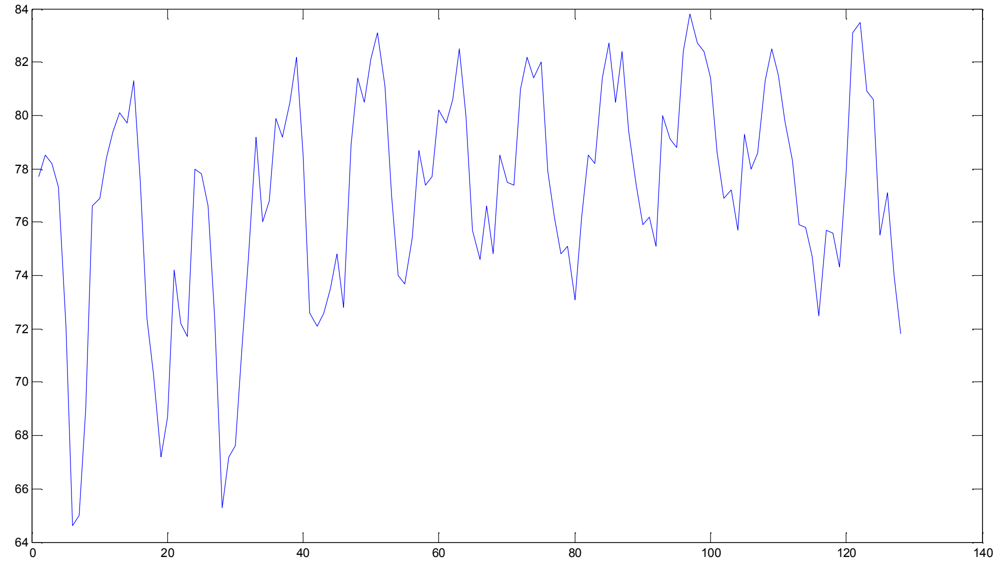
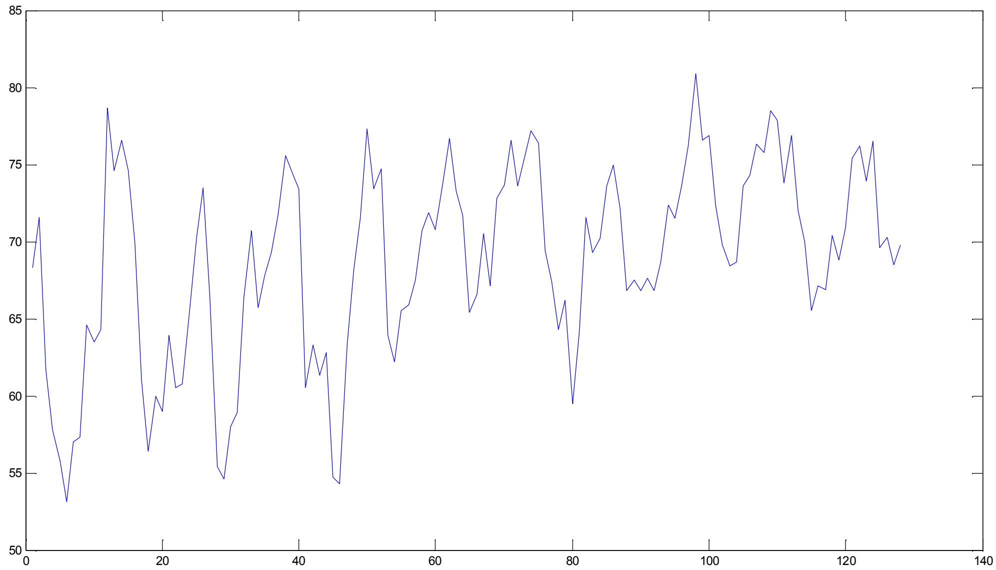


Figure 6. Load Factor - Turkish Airlines - Monthly



On the other hand, there is weekly data that is specified for both airline and geographical basis. This data is available from 2005 to 2010. This is a shorter period but since it has a weekly basis, it gives a better resolution than the monthly data. Thus weekly data will be useful in the identification of structural breaks. Repetitive cycles in figures 7 and 8 are clearly visible.

Figure 7. Load Factor – Total Scheduled - Weekly

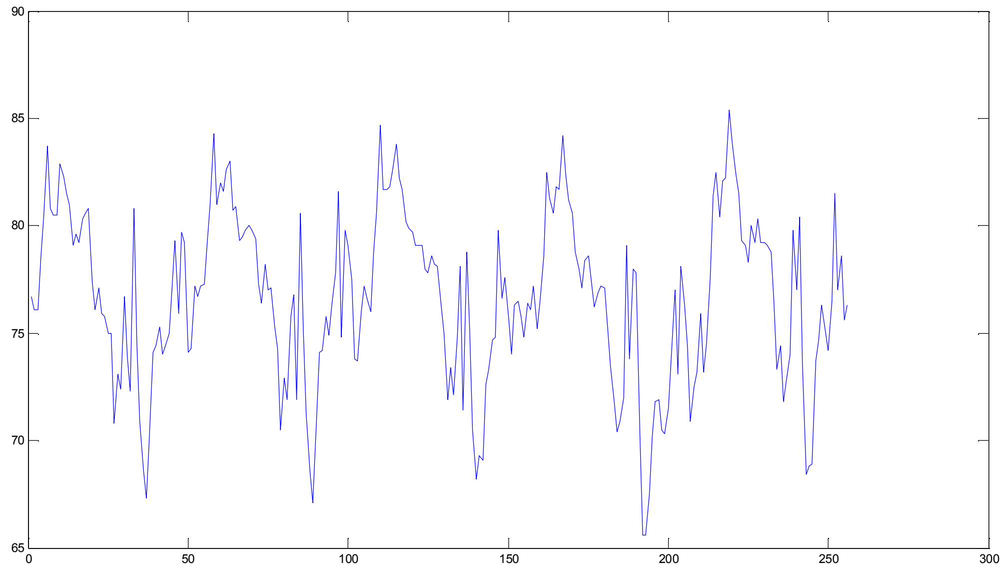
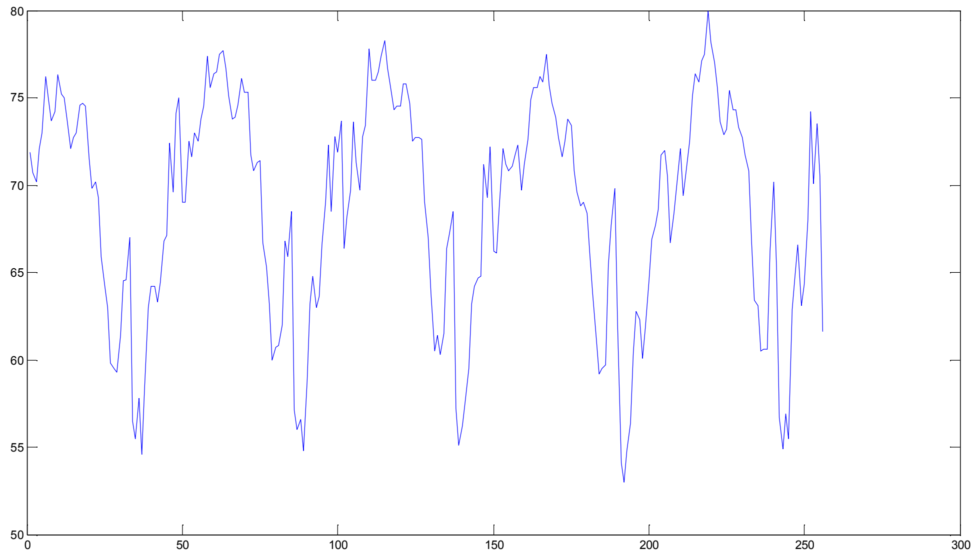


Figure 8. Load Factor - EU - Weekly

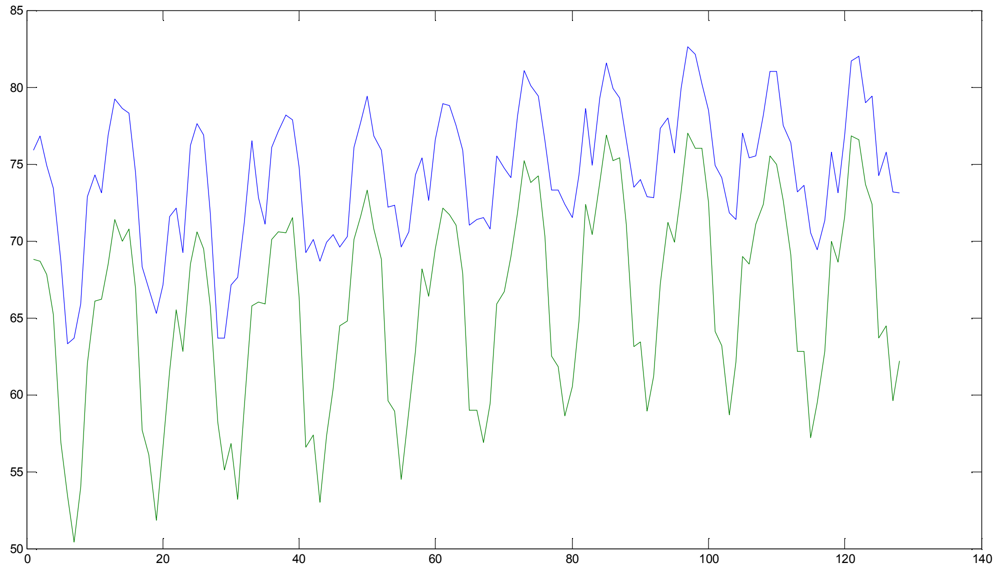


3.4.2. PLF Comparisons and Analyses

In order to understand the fluctuations of PLF, data series are compared both geographically and airline based. The purpose of this comparison is to uncover the hidden patterns and correlations in these series.

When PLF data is analyzed and compared in its original form, a restricted range of observation is possible due to the non-stationary and complex characteristics of data series.

Figure 9. Load Factor Total vs Load Factor EU



When the total load factor series is divided into two parts, the following results are obtained;

For part 1, mean 68.7957, variance 28.5639.

For part 2, mean 74.6435, variance 17.1581.

The change in variance corresponds to the non-stationarity.

And when the EU load factor series is divided into two parts, the following results are obtained;

For part 1 mean 60.1791, variance 40.1394.

For part 2 mean 66.2061, variance 41.0901.

This corresponds that also PLF for EU is non-stationary.

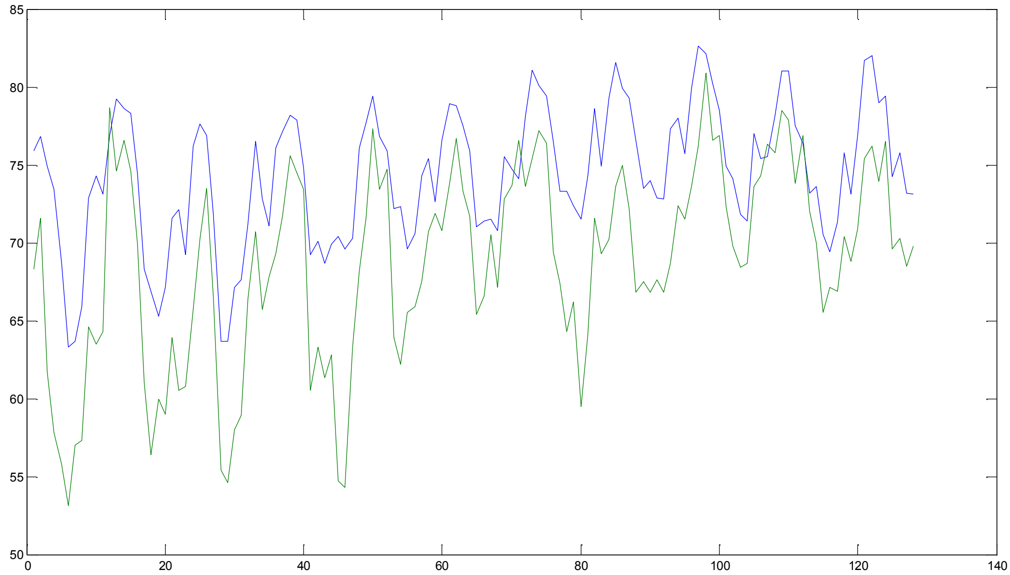
And when the Lufthansa load factor series is divided into two parts, the following results are obtained;

For part 1 mean 75.8531, variance 22.0711.

For part 2 mean 78.2781, variance 9.2455.

Corresponding that also PLF for Lufthansa is non-stationary.

Figure 10. Load Factor Total vs. Load Factor THY



On the other hand, as the time series data for the load factors (namely Total load factor and Load factor for Turkish Airlines) are compared;

The correlation coefficient R = 0.8696. The usual statistical methods can account for up to a certain extent and does not give precise results.

However, if we use the wavelet decomposition, the wavelet coefficients on each scale will allow us to make a comparison and analysis on different time spans.

When both of the series are decomposed with Symmlet 8 mother wavelet;

Figure 11. Wavelet Decomposition of Load Factor – Total Scheduled

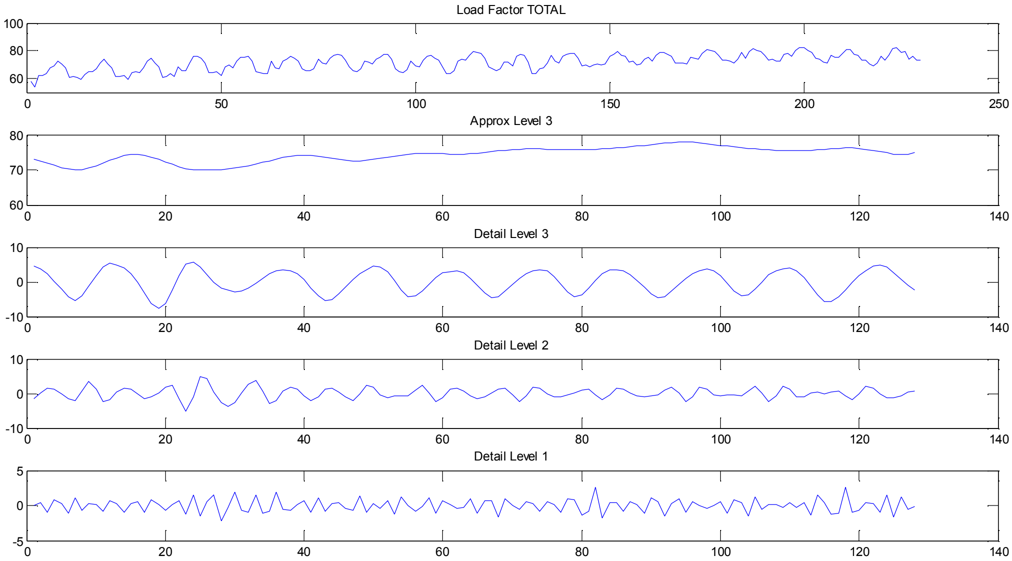
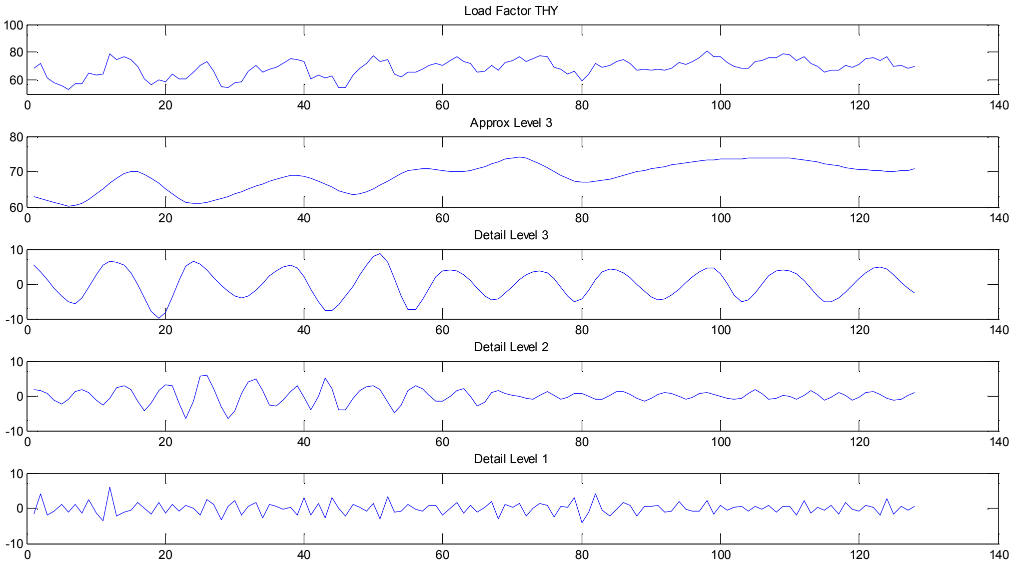


Figure 12. Wavelet Decomposition of Load Factor - Turkish Airlines

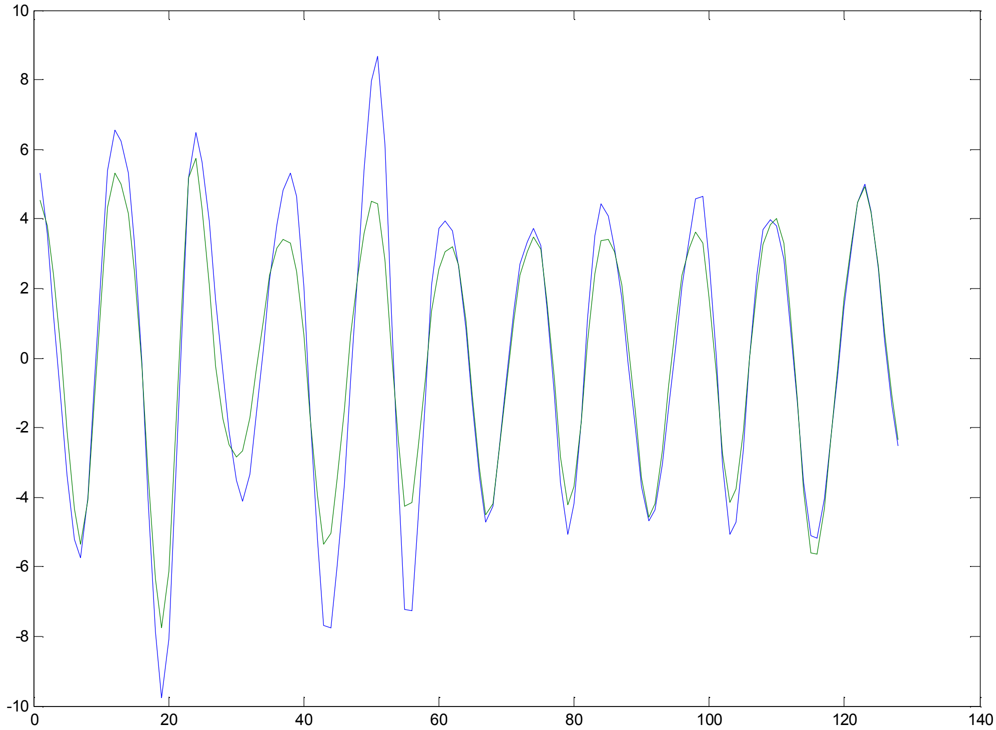


After wavelet decomposition, the corresponding energies at each level can be calculated by means of estimating the sum of the variances. And as a result of this, for the Total load factor, the 83% of all changes are emerging from the level 3 detail band.

(For the for the THY load factor, the 81% of all changes)

At level 3 on frequency band, which corresponds to a period of 12 months, the high correlation between the total load factor and the load factor of the THY can be seen much more clearly in Figure 13.

Figure 13. Load Factor - Detail Level 3 - Total vs THY



It indicates that, most the alteration in load factors are a global phenomenon, and for low-frequency (i.e. long period) band, the load factor of an airline company are very highly correlated.

At, detail level 3, the correlation coefficient between the total load factor and the load factor of the THY R= 0.9720. This is a clear and meaningful result achieved by the use of wavelet decomposition.

Figure 14. Wavelet Decomposition of Load Factor - Lufthansa

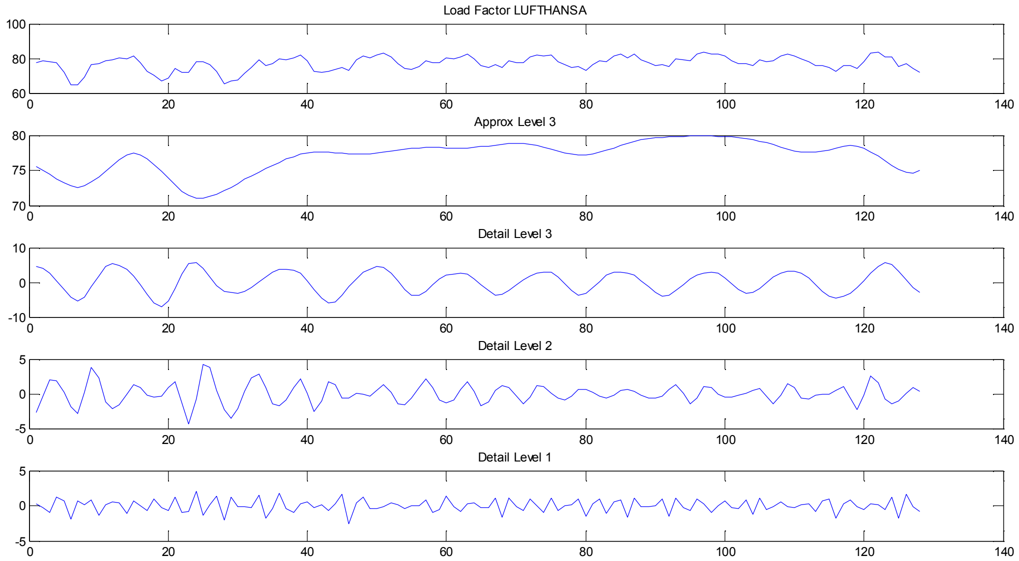
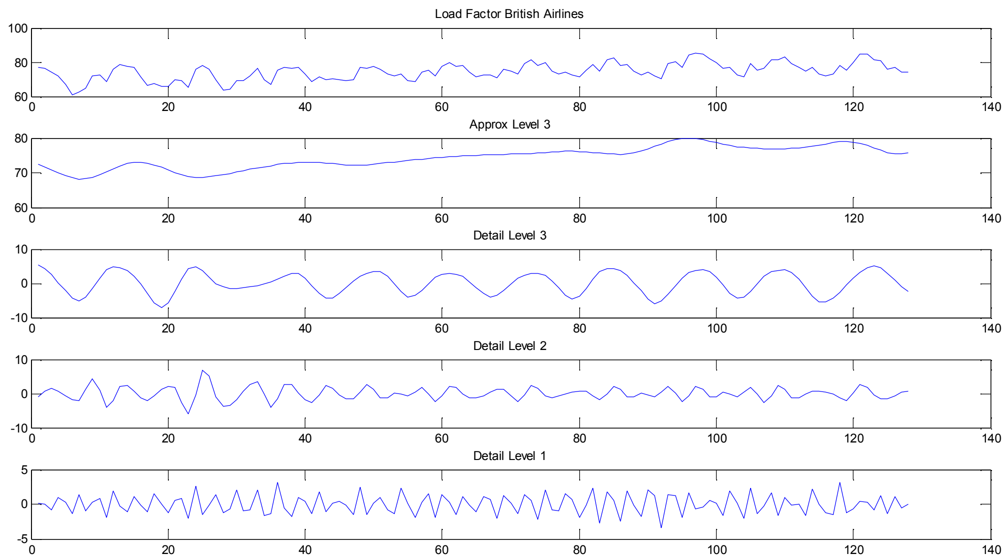


Figure 15. Wavelet Decomposition of Load Factor - British Airways



For British Airways, energy distributions for 3 detail levels are;

dl 1	dl 2	dl 3
(6%)	(14%)	80%

Hence for BA, decomposition level 3 has the highest energy (80% of all energy) too. PLF comparison between two major airlines, Lufthansa and British Airways will be an interesting matchup.

Figure 16. Load Factor Lufthansa vs. Load Factor British Airways

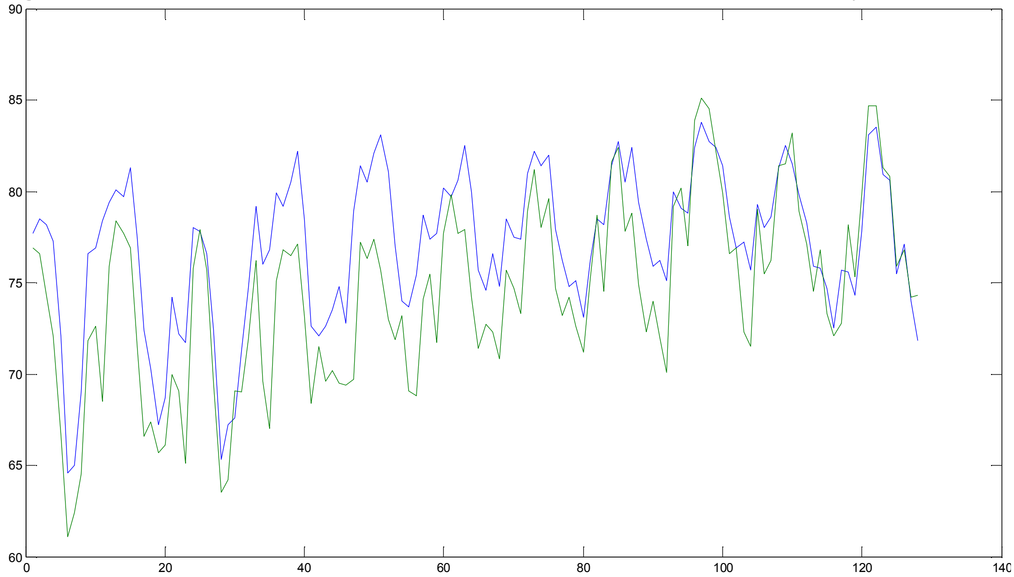
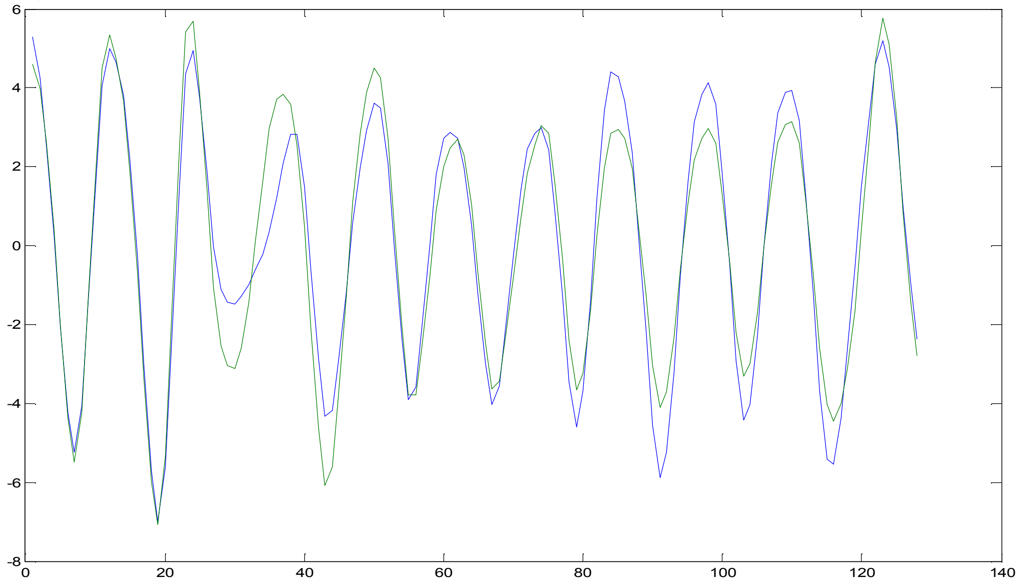
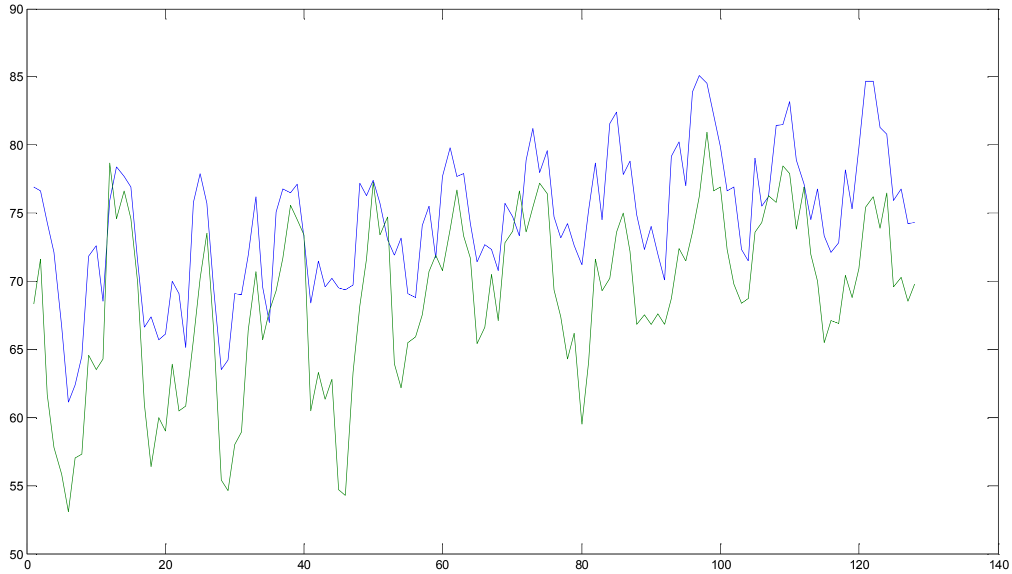


Figure 17. Load Factor - Detail Level 3 - Lufthansa vs. British Airlines



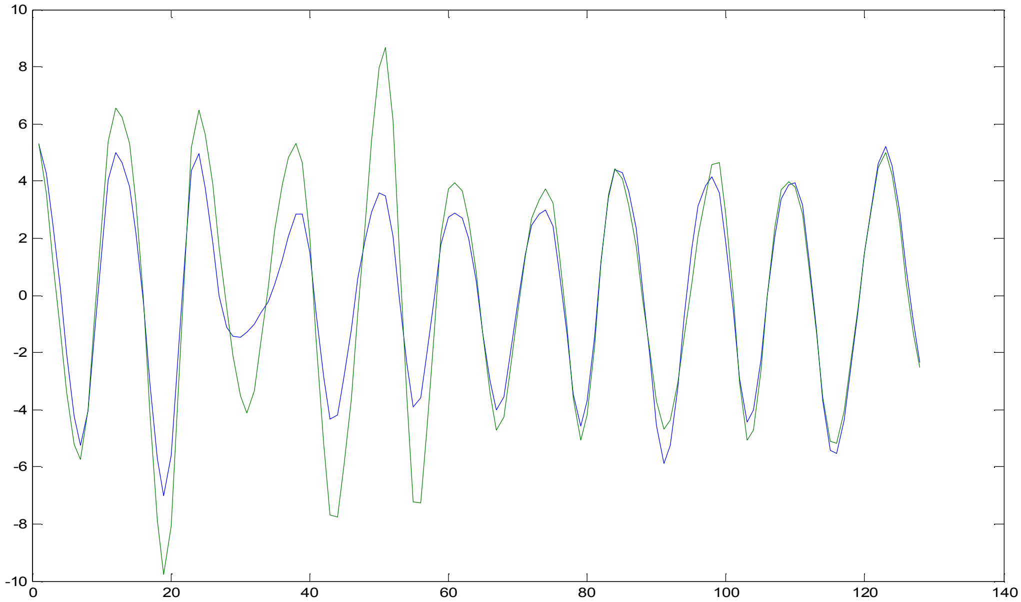
At, detail level 3, the correlation coefficient between the LF level 3 detail coefficients of Lufthansa and British Airways is R= 0.9602. It means that in yearly scale, Lufthansa and British Airways are affected by the same trend in a close way.

Figure 18. Load Factor British vs. Load Factor THY



Load factors of British Airways and THY can be seen at figure 18 and the correlation coefficient R = 0.8342. So there is a strong relation as it is also observable in the graph. Since for both British Airways and THY decomposition level 3 represents most of the fluctuations, they are compared.

Figure 19. Load Factor – Detail Level 3 - British Airlines vs. THY



From figures 13 and 19, one can observe that LF of THY fluctuates more than LF of the total and LF of BA around 2003-2004 periods. It corresponds to the policy change in THY, in 2003. THY implemented a new policy of aggressive growth after 2003. And in aviation sudden growth decreases LF, since it takes time for people to get used to the new flight schedule and destination for new routes. Another inference from figures 13 and 19 is for the last two years, LF of THY has almost the same fluctuation with total LF and LF of BA. The correlation coefficient R = 0.9472 for LF of THY and BA on detail level 3.

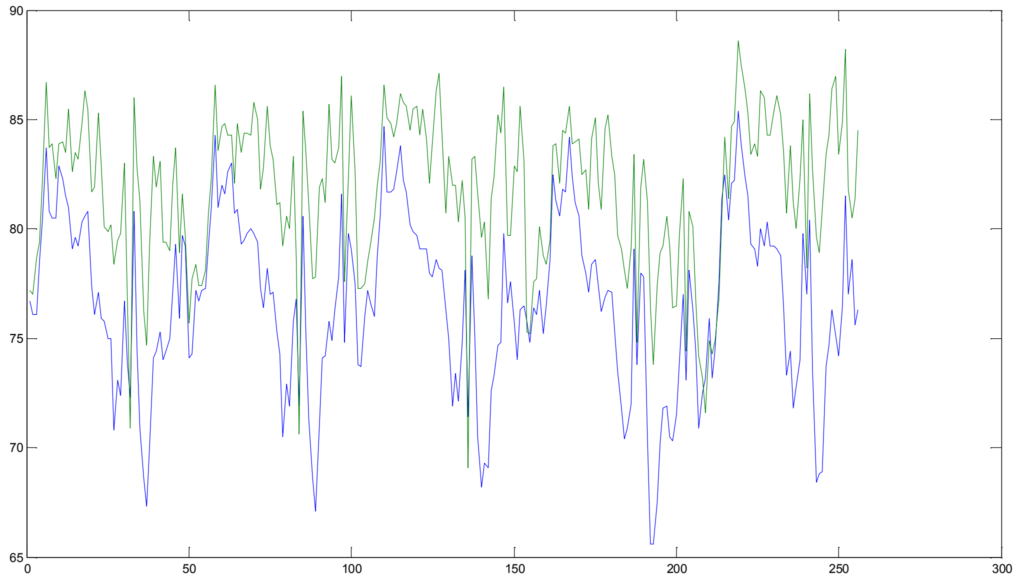
3.4.3. Identification of Structural Breaks

When analyzing the time series data with the wavelet-based approach, detecting the deviations from stationary at specific point is possible. LF data in weekly basis is used for identification of structural breaks since it has more resolution and its low-level components are much more representative.

As the structural break is a sudden change in variance, the low-level wavelet coefficients (which are associated with high-frequency content of the time series) retain the sudden shift in variability while the high-level coefficients are stationary. As the structural break is a probable change in the long-range dependence of the series, all levels of wavelet coefficients may exhibit a structural change since long memory is associated with all scales-especially the low-frequency ones.

(Gencay et al, 2002)

Figure 20. Load Factor Total vs. Load Factor AE - Weekly



The scale-by-scale wavelet decomposition of a time series data has the ability to unveil the structure at different time horizons. Decompositions of LF Total and LF AE in weekly basis are shown in figures 21 and 23. The correlation coefficient for LF Total and LF AE R=0.6675. Energy distributions for 3 detail levels are;

LF total weekly;

dl 1	dl 2	dl 3
(20%)	(26%)	(54%)

LF AE weekly;

dl 1	dl 2	dl 3
(36%)	(28%)	(36%)

Figure 21. Wavelet Decomposition of Total - Weekly

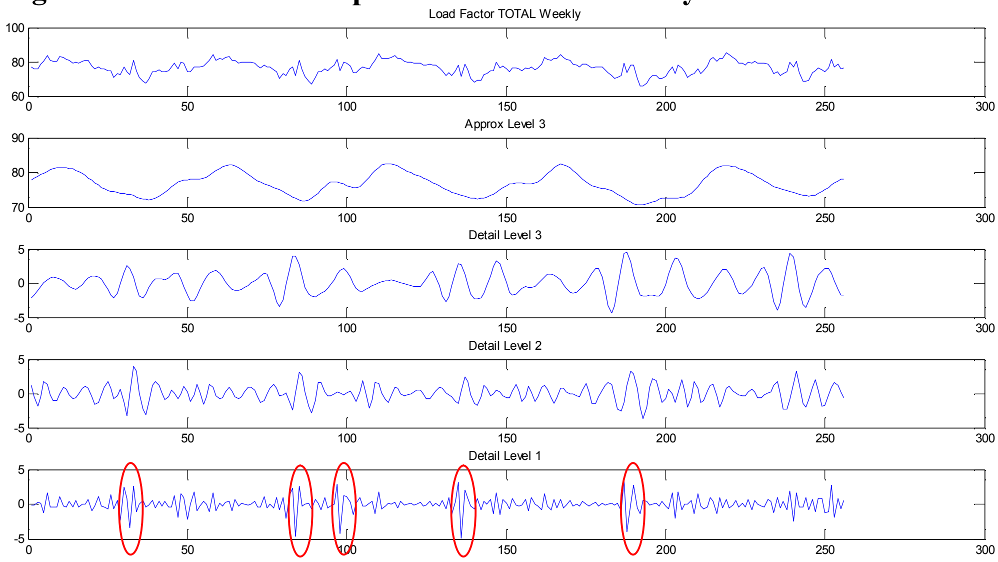


Figure 22. Load Factor Detail Level 1 - Total - Weekly

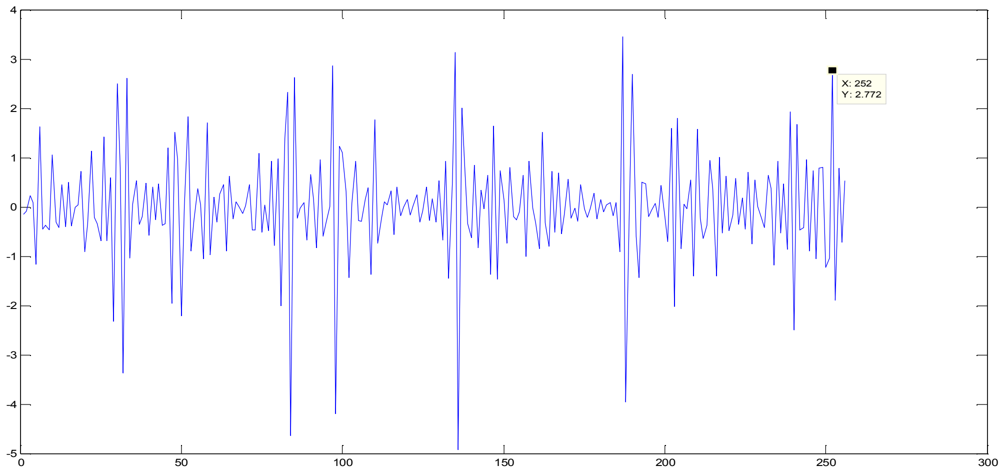


Figure 23. Wavelet Decomposition of AE - Weekly

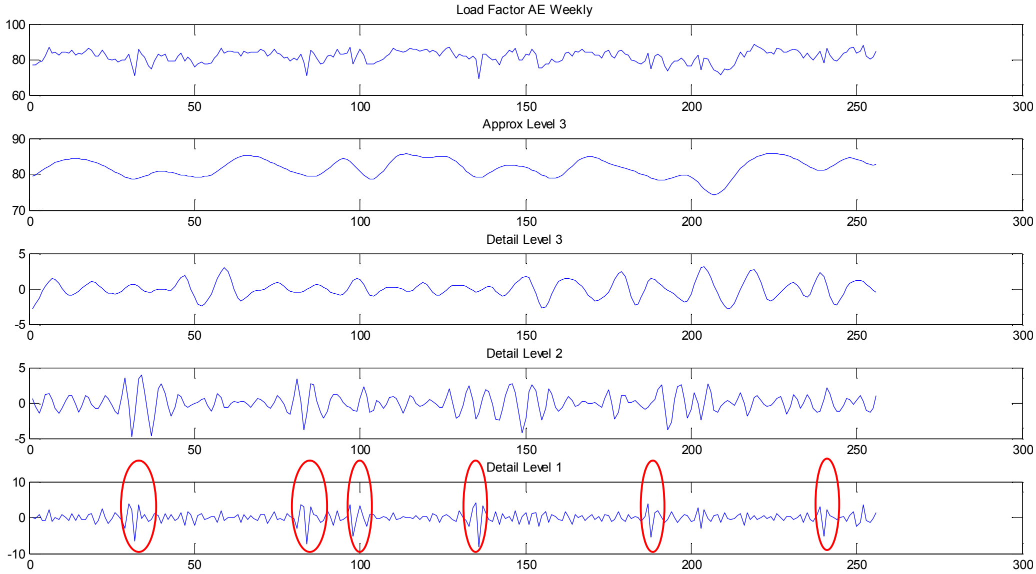
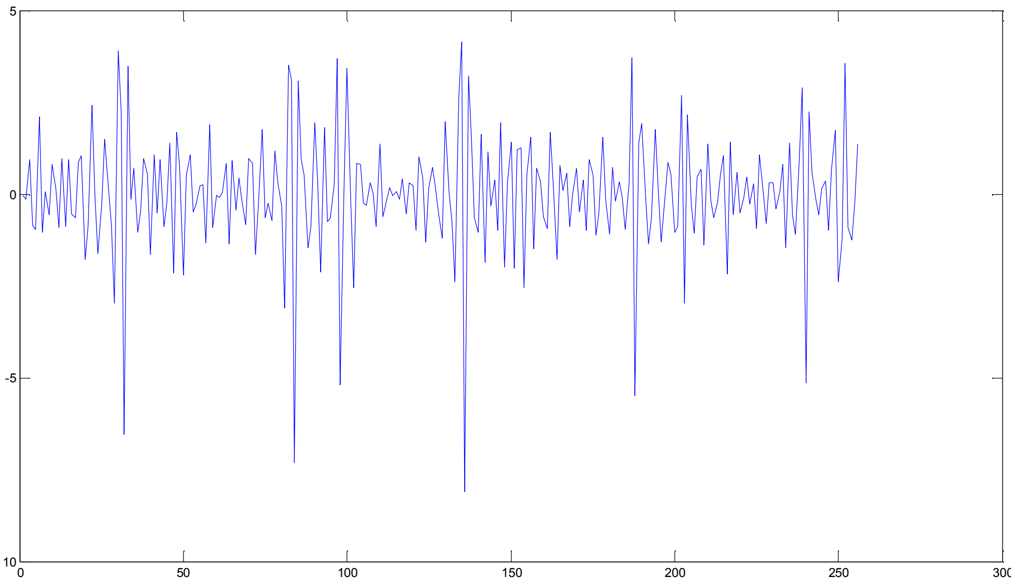


Figure 24. Load Factor Detail Level 1 - Total - Weekly



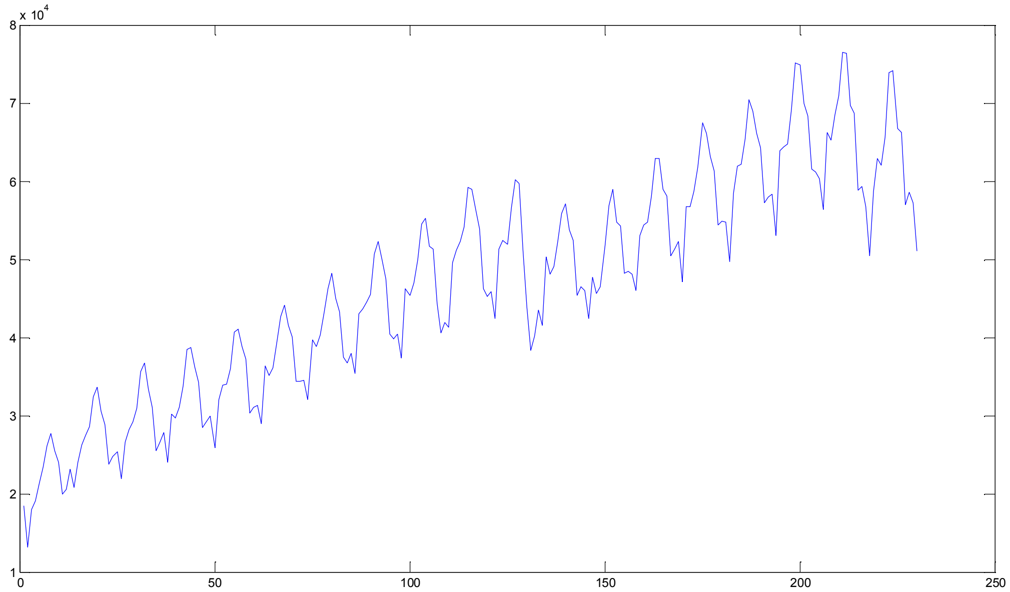
Mother wavelet will allow us to localize the points of structural breaks. Indeed high-frequency coefficients for total load factor, point out the data points corresponding to the dates of Christmas weeks and widespread epidemics such as Avian Influenza and Swine Influenza.

All these analyses prove that wavelet decomposition is a highly efficient tool in identifying structural breaks, discontinuities and comparing their relative importance in irregular time series data.

3.4.4. RPK Analyses and Comparison

Revenue passenger kilometer is an indicator that consists both efficiency and magnitude of airline operations, since it is a product of LF and ASK. With this property RPK is an important data for an outsider to understand the performance and magnitude of an airline company.

Figure 25. RPK - Total - Monthly



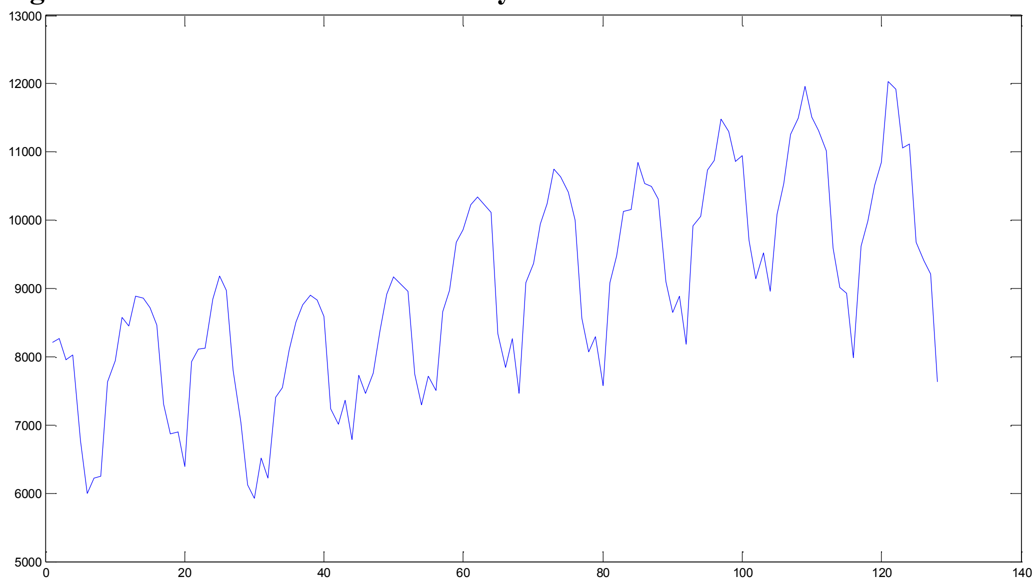
When the RPK total data series is divided into two parts, the following results are obtained;

For part 1, mean 3.5644e+004 variance 9.3088e+007

For part 2, mean 5.7315e+004 variance 7.5035e+007.

The change in variance corresponds to the non-stationarity.

Figure 26. RPK - Lufthansa - Monthly



When the RPK total data series is divided into two parts, the following results are obtained;

For part 1, mean 7.7896e+003 variance 8.6578e+005

For part 2, mean 9.8611e+003 variance 1.2925e+006

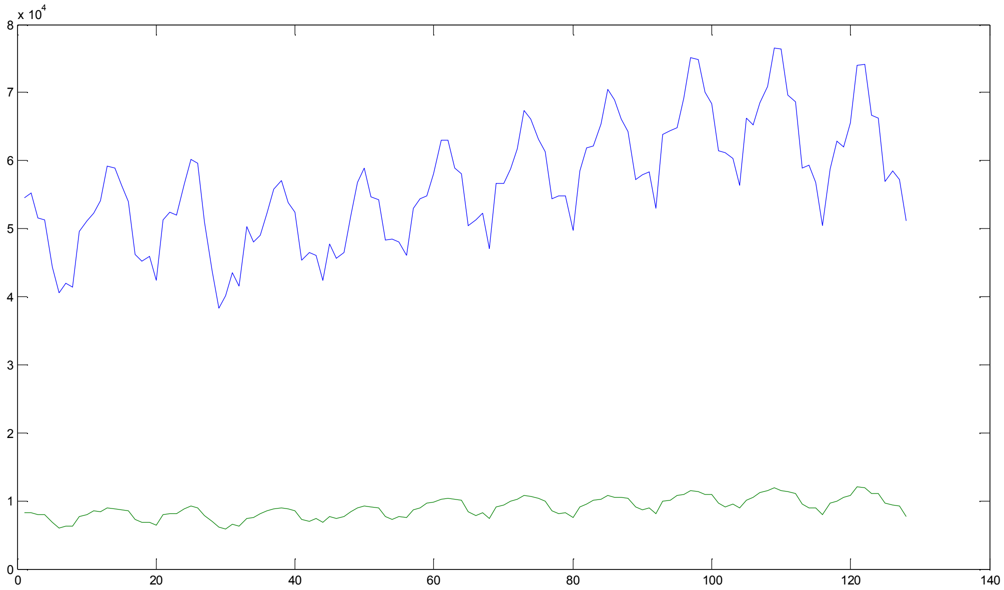
The change in variance corresponds to the non-stationarity.

By decomposing RPK of Lufthansa, energy distributions for 3 detail levels are;

dl 1	dl 2	dl 3
0.0495 (5%)	0.0428 (4%)	0.9212 (91%)

As in LF data, for RPK of Lufthansa detail level 3 represents most of the fluctuations. Hence, in order to clarify RPK series focusing on detail level 3 is necessary.

Figure 27. RPK - Total vs. Lufthansa - Monthly



Comparison of RPK Total and Lufthansa by the graph of figure 27 is not easy, because of the different magnitudes of these two data series. But after wavelet decomposition, comparison for detail level 3 can be seen in figure 28. The correlation coefficient of these two data R = 0.9945, which is interestingly high.

Figure 28. RPK Detail Level 3 - Total vs. Lufthansa - Monthly

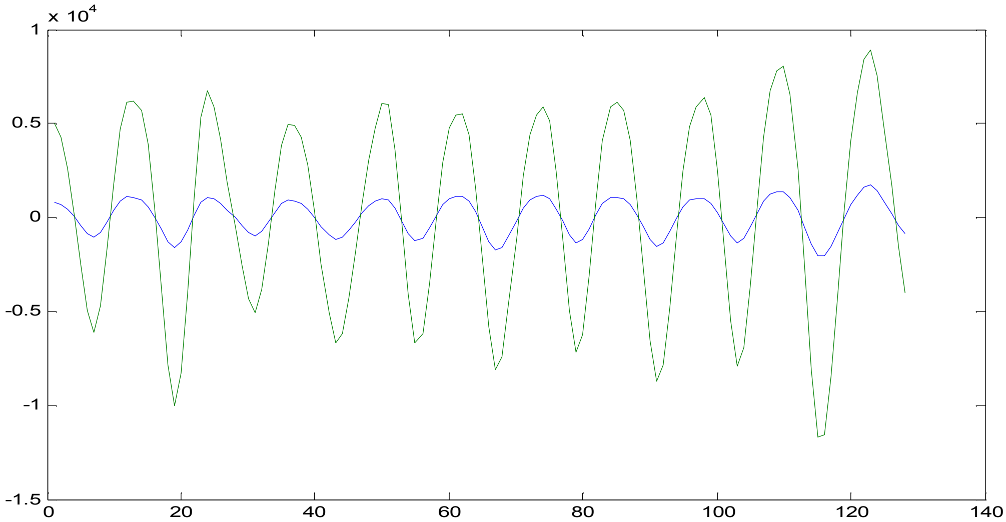
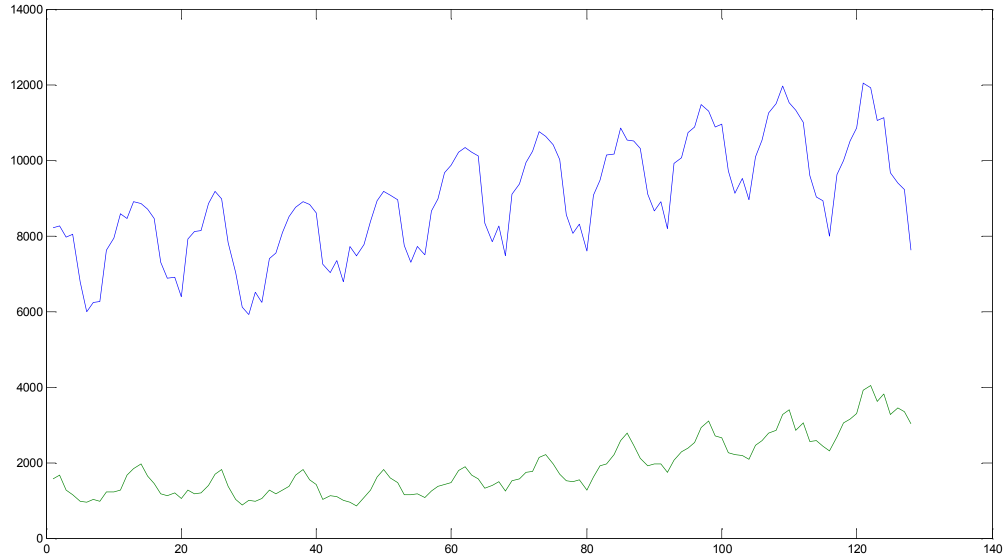


Figure 29. RPK - THY vs. Lufthansa - Monthly

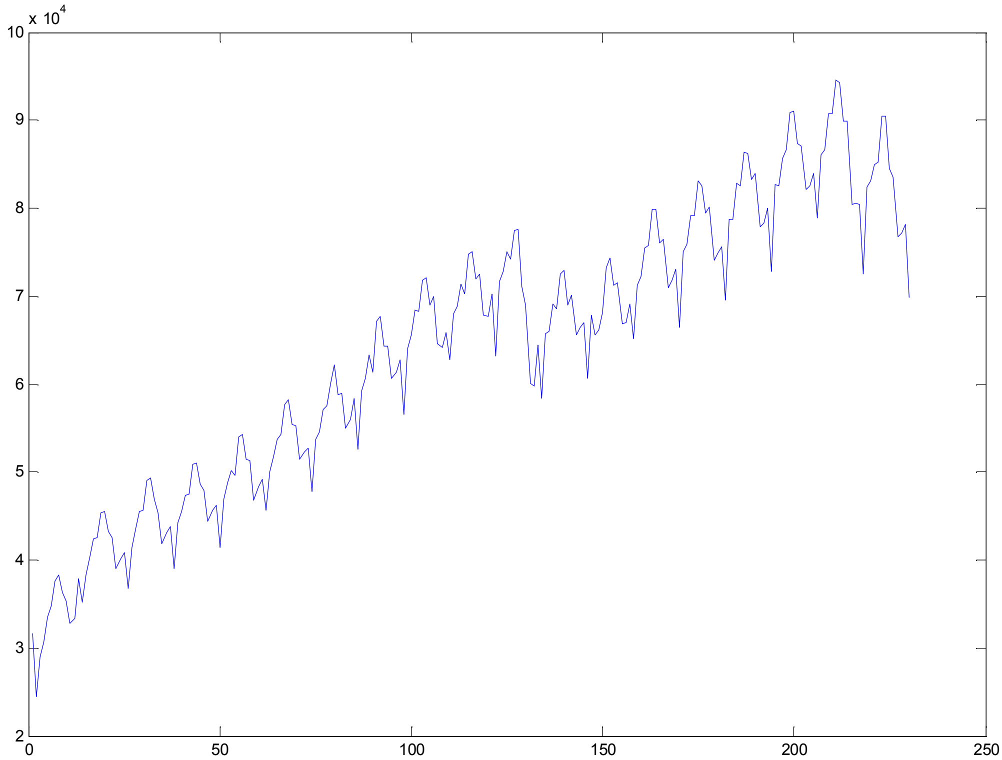


Comparison of RPK between THY and Lufthansa is in figure 29. Correlation coefficient of these two data is R = 0.8174. After decomposition, comparison for detail level 3 gives correlation coefficient of R = 0.9678.

3.4.5. ASK Analyses and Comparison

Available seat kilometer is the supply side of the aviation industry. In the short term it is not optimal for airlines to make sharp changes in ASK, because of fixed costs. For example, an aircraft in the fleet of an airline has fixed costs (rental, owning and maintenance costs) and in order to decrease ASK utilization of aircraft decreases, hence cost per flight increases due to the fixed costs. For airlines, it is possible to increase or decrease the number of aircrafts in its fleet in order to control ASK, but in aircraft market it takes time to purchase or sell. For instance, delivery of aircrafts takes 1 to 3 years of time after the order. On the other side, in the long run airlines have direct control on ASK.

Figure 30. Total ASK - Monthly



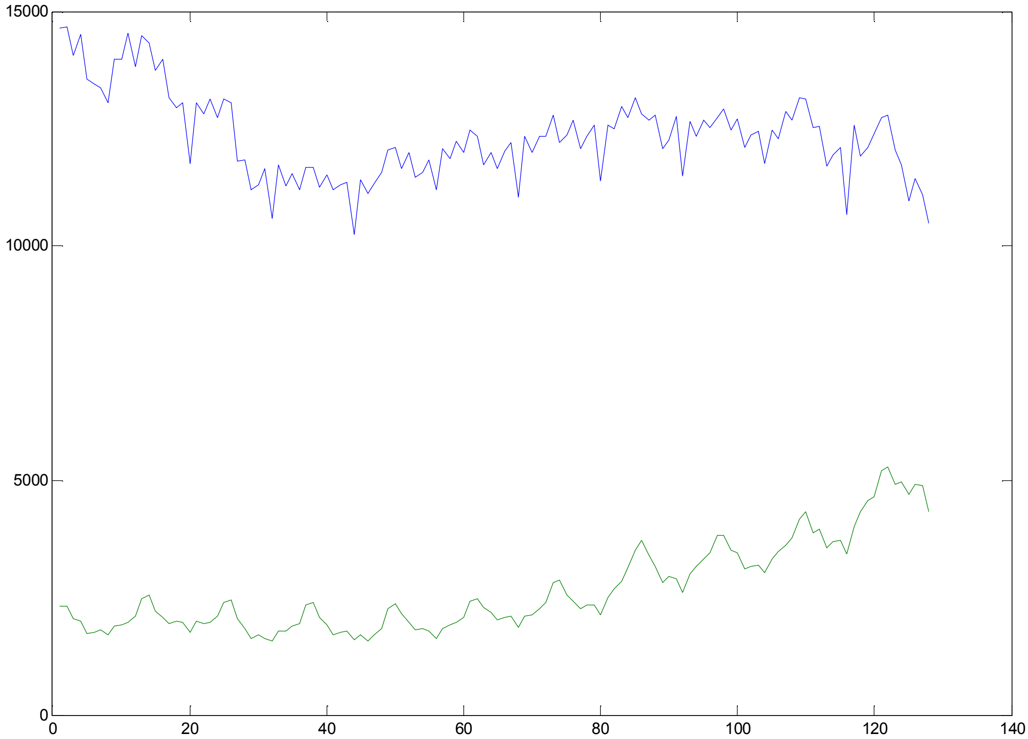
When the ASK total data series is divided into two parts, the following results are obtained;

For part 1, variance 1.2470e+008

For part 2, variance 6.6945e+007

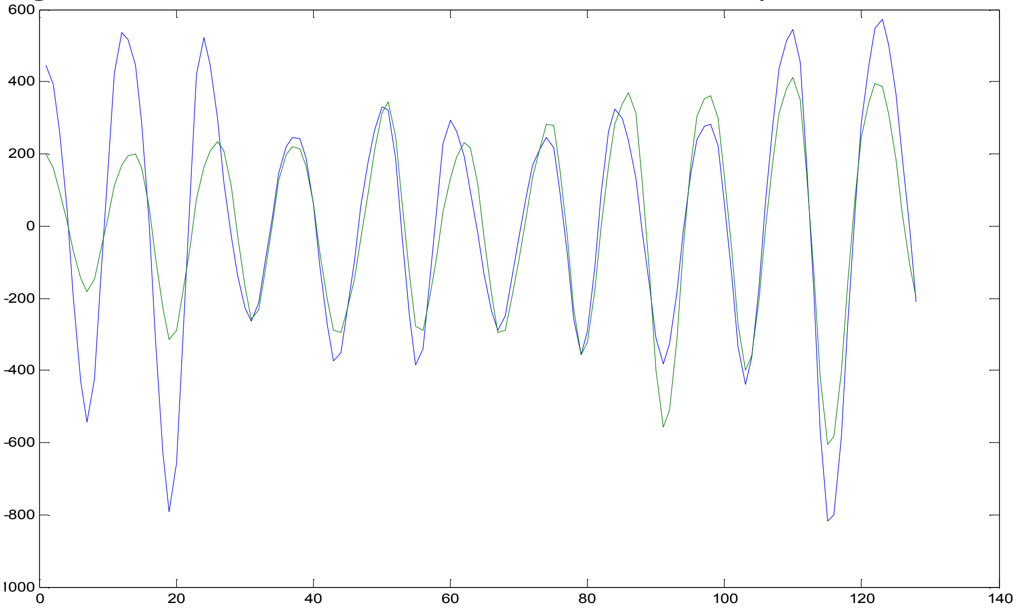
The change in variance corresponds to the non-stationarity.

Figure 31. ASK - British vs. THY - Monthly



Comparison of ASK British and THY by the graph of figure 31 does not give clear results. The correlation coefficient of these two data R = 0.0225. It means that there is no significant correlation. But after wavelet decomposition, comparison for detail level 3 can be seen in figure 32. The correlation coefficient of these two data R =0.9022, which indicates a strong relation in the 12 months time scale.

Figure 32. ASK Detail Level 3 - British vs. THY - Monthly



By using wavelet decomposition hidden relation between two data can be easily clarified. With the help of these analyses it is possible for airlines to unveil the annual seasonality in the supply side of the aviation industry. In this case, optimizing ASK for an airline will be more precise.

CHAPTER 4

CONCLUSION

Wavelet decomposition is a novel means for analyzing time series data. The most distinctive and the most beneficial property of wavelets are their ability to analyze irregular, non stationary time series without any manipulation or loss in data. And, the decomposition on different orthogonal scales corresponding to either high frequencies (prevalent in short periods) or low frequencies (prevalent for longer periods), is in optimal consistence with the Heisenberg-Gabor inequality principle. In result, this particular mathematical technique is more useful in handling such data and unraveling its characteristic behavior both in time and frequency scales.

In terms of aviation industry, airlines have objective to optimize their supply in a stochastic environment with complex and non stationary data sets. Hence, in practice decision makers mostly use their intuition and simple statistics with raw data for airline performance to make strategic and important choices rather than rigorous and comparative analysis of time series data. On the other hand, a composition of seasonality for various time spans such as fluctuations inside the weeks, differences between winter and summer schedules, repeated trends in leisure and holiday traffic makes life harder for aviation professionals. Wavelet decomposition is an efficient tool to deal with seasonality, because different frequency levels in the decomposed data naturally represent them individually, so it enables to clarify the connections, impacts and magnitudes of seasonal components. Moreover, analysis of PLF, ASK, and RPK for AEA airlines with wavelet decomposition have some remarkable outcomes, such as exposing hidden strong correlations of specific detail levels for comparisons between airlines or geographic regions.

The analyses and comparisons of PLF shows that, about 80-85% of changes in load factors are due to the fluctuations in detail level 3 (i.e. 12 months period). Comparisons of decomposed components (mostly detail level 3) revealed strong relationships between load factors of various airlines, enabling to investigate and measure performances of airlines comparatively. On the other hand, the analyses of lowest detail level for the decomposed weekly PLF series, revealed the time and impacts of sudden changes in the airlines industry.

Moreover, the analyses of RPK resulted that, over 90% of chances in RPK is due to the fluctuations in detail level 3. The comparison of decomposed RPK between Total and Lufthansa, resulted with a strong correlation with $R = 0.9945$. Also ASK comparison lead us to interesting results. While correlation coefficient between ASK of British Airways and THY is $R = 0.0225$, after wavelet decomposition at detail level 3, the same comparison resulted with $R = 0.9022$ implying a strong correlation.

Beneath the analysis of AEA data, the approach of wavelet decomposition has promising aspects for aviation industry. Since airlines have a wider range of data available to keep track of their performance, such as future reservations (load factors of future flights due to reservations), price levels, excess luggage etc., the illustrated properties of wavelet decomposition enables them to investigate various real life problems practically. Additionally, airlines have data series in daily basis enabling much better data resolution compared to the monthly and weekly data analysis; therefore they can achieve clearer results.

SELECT BIBLIOGRAPHY

Addison, Paul S. *The Illustrated Wavelet Transform Handbook Introductorv Theorv and Applications in Science, Engineering, Medicine and Finance*. Bristol and Philadelphia: Institute of Physics Publishing, 2002.

Barros Carlos Pestana, Peypoch Nicolas. "An Evaluation of European Airlines' Operational Performances." . *Int. J. Production Economics 122*. 2009. pp 525–533

De Moortel I, Hood A.W, and Munday S. A. "Wavelet analysis: the effect of varying basic wavelet parameters". *Solar Physics* 222. 2004. 203–228.

Gencay Ramazan, Selcuk Faruk and Whitcher Brandon. *An Introduction to Wavelets and Other Filtering Methods in Finance And Economics*. U.S.A.:Academic Press, 2002.

Goetz Andrew R. and Graham Brian. "Air transport globalization, liberalization and sustainability: post-2001 policy dynamics in the United States and Europe." *Journal of Transport Geography 12*. 2004. pp 265–276

Janic Milan. "Liberalization of European aviation: analysis and modeling of the airline behavior." *Journal of Air Transport Management*, Vol. 3, No 4. 1997. 167-180.

Mallat, S.G.: *A Wavelet tour of signal processing, 2nd Ed.*, U.S.A: Academic Press, 1999.

Massart D.L, Pasti L., Reschiglian P., Walczak B.. " Optimization of signal denoising in discrete wavelet transform." *Chemometrics and Intelligent Laboratory Systems 48*. 1999. 21–34.

Modi Jatan K, Nanavati Sachin P, Phadke Amit S and Panigrahi Prasanta K. "Wavelet Transforms: Application to Data Analysis – I". *Resonance I.* 2004. 10-22.

Schleicher, Christoph. *An Introduction to Wavelets for Economists.* Bank of Canada Working Paper, 2002-3.