Is Fuzzy Time Series a fallacy?
Potentials and limitations of fuzzy set approach in time series analysis

Okan Duru
Istanbul Technical University
Department of Maritime Transportation and Management Eng.
duruokan@yahoo.com
What is the FTS?

- Time series clustering
- Computer Intelligence
- Rule-Based forecasting
- An educated guess method
- Rule-of-thumb solution
- Pattern recognition
Fuzzy Set Theory

Developed by Zadeh (1965) to execute uncertain environment and to convert linguistic particulars into executable forms.

There is no widely accepted rules for the selection of the type of fuzzy set and its particulars. There are several studies using many different structures.

Triangular Sets are used frequently

Ex. Low Moderate High Very High

Temperature

Ex. Low Moderate High Very High

Volatility

Trapezoidal Sets are also applied in many studies

Data

Membership degree

Fuzzy Time Series

Okan Duru © 2011 | Laboratory of Transport Economics & Management
According to the literature, FTS is first introduced by Song and Chissom (1993a,b,1994). Later, several improvements are presented by many scholars. Time-variant and time-invariant versions; Univariate and Multi-variate versions; First order and High order versions etc developed.

According to the literature, Fuzzy Time Series (FTS) can work even some of the data is not available. FTS can work under the linguistic time series. FTS model can be established under the time space or state space conditions. FTS can be built under the multi-variate conditions.

<table>
<thead>
<tr>
<th>Arguments</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Conditions such as normality, stationarity are not required.</td>
</tr>
<tr>
<td>2</td>
<td>FTS can work even some of the data is not available.</td>
</tr>
<tr>
<td>3</td>
<td>FTS can work under the linguistic time series</td>
</tr>
<tr>
<td>4</td>
<td>FTS model can be established under the time space or state space conditions</td>
</tr>
<tr>
<td>5</td>
<td>FTS can be built under the multi-variate conditions</td>
</tr>
</tbody>
</table>
Fuzzy Time Series
How it works

- Data is classified into intervals (sets)
- Relationship between classes is identified
- Rules of extrapolation is defined
- Extrapolation is generated by using the rules
- The crisp prediction is calculated by a valuation formula
- Accuracy Control
FTS vs. Fuzzy aided TS

Raw Data

Data Clustering

Expert Oriented

Applying Classical TS under the fuzzy arithmetics

Computer Intelligence

Pattern Recognition & Extrapolation

Building Rules Of Extrapolation

Extrapolation & Defuzzification

Accuracy Check For FTS Forecast
Process of FTS

Data prep

Clustering

Algorithm

Learning

Prediction

Declustering

Data collection, transformations etc.

Defining the no. of cluster & Fuzzification

Design/ Selection of Rule structure (i.e. Univariate)

Estimation, Pattern Recognition by rules

Extrapolation by historical pattern of fuzzy sets

Transforming fuzzy forecast to crisp result

Fuzzy Time Series
Data transformation is usually not used in FTS and stationarity is ignored.

Non-stationary data is frequently used in FTS papers even it is recognized as the advantage of FTS.

However, the principals of spurious regression theory apply for the FTS studies too.
There is no linear regression in FTS

If the series move up more, FTS can not extrapolate and the result is just a Naïve solution.

If the levels of data is compiled in the FTS, many algorithms just suggest to extrapolate the same fuzzy set in previous period.

Non-stationarity can not be declared as an advantage. It is a drawback of a-theoretical supervision.

Data transformation is strongly recommended.

Duru, 2010

Expert Systems with Appl.

Conditions such as normality, stationarity are not required.
FTS can work even some of the data is not available.

FTS model can be established under the time space or state space conditions.
Fuzzy Time Series
Fuzzy Time Series

Clustering

(a) Triangular fuzzy set design

(b) Trapezoidal fuzzy set design

(c) Gaussian fuzzy set design

(d) C-means fuzzy set design (density clustering)
Optimized sets
FTS can work under the linguistic time series

If the data is pure linguistic records, then these linguistic representations can be defined as fuzzy sets without any numerical response. In case of numerical dataset, data space is converted into intervals with linguistic labels. If it is naturally linguistic, such an operation is not needed.

It is quite similar to Rule-based forecasting approach.

Chen (1996)
There are several recommendations for defining the number of clusters. The gap still exists in the literature. How to optimize it? Equal length clusters may ignore the effects of high kurtosis. How to define clusters among the recession (low average deviation) and upturn (high average deviation) periods? A multiple length clustering can be more accurate. An initial number of clusters is required for C-Means clustering.
The development of rules for extrapolation is still an existing gap in the literature. Many scholars apply different rule structures without any theoretical and/or parametric basis.

The most useful criteria for selection is accuracy level.

There are many different structures such as univariate, multivariate with/without theoretical base etc.
Fuzzy Time Series

Algorithm

Univariate FTS model

Data A

Data B

Bivariate FTS model

Second Order FTS
Fuzzy Time Series

Objective Data → Data Clustering → Estimation → Univariate FTS model → In-Sample

Random Data → Data Clustering → Estimation → Bivariate FTS model → Post-Sample

Spurious Modelling Causality
For the simulation, a random dataset is generated normally distributed with a mean (5,000) and standard deviation (1,000).

Objective data is the Baltic Dry Index which is a number issued by Baltic Exchange, London as a combined value of dry bulk shipping freight rates (2009:1-2010:12, last three months as the testing period.

The simulation is based on the first order FTS design.
### In-Sample (21 months)

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>uFTS</td>
<td>233</td>
</tr>
<tr>
<td>bFTS</td>
<td>232*</td>
</tr>
</tbody>
</table>

### Post-Sample (3 months)

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>uFTS</td>
<td>207*</td>
</tr>
<tr>
<td>bFTS</td>
<td>218</td>
</tr>
</tbody>
</table>

Even a random dataset may increase the accuracy of the FTS. If a dataset which is assumed improving to the forecasting accuracy has particular pattern, then the picture may strongly support the bFTS for in-sample period.

However, it is not mean that the proposed model is robust and useful for the forecasting practice. Post-sample accuracy control is frequently ignored in the FTS studies!
FTS can be build under the multi-variate conditions

However, theoretical framework has critical importance. Post-sample control is critical. The causality link between multiple dataset should also be investigated theoretically.

Duru, 2010 – IAME 2010 Conference
A Simple Univariate FTS model

Fuzzy Arithmetics (Zadeh, 1969)

Fuzzy Time Series
Fuzzy Time Series

Prediction

Group of Factor C
C → n.a.

Group of Factor B
B → C

Group of Factor A
A → B
A → A

Data

Estimation
Testing or Prediction

C → ? = C
Naïve Solution

B → ? = C

A → ? = (B+A)/2

Example

Example

Fuzzy Time Series

Okan Duru | Laboratory of Transport Economics & Management
Prediction

Fuzzy Time Series

Model estimation (Learning period)  Forecasting

Crisp Forecast

Okan Duru | Laboratory of Transport Economics & Management
Fuzzy Prediction

Triangular Sets

Crisp Prediction

m (midpoint)

\[ g = \frac{l+m+u}{3} \]

Crisp result can be defined as peak points

C-means Sets

\[ g = \frac{l+a+b+u}{4} \]

Crisp result can be defined as center of gravity

Fuzzy Time Series
<table>
<thead>
<tr>
<th></th>
<th>Conditions such as normality, stationarity are not required.</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>FTS can work even some of the data is not available.</td>
</tr>
<tr>
<td>✓</td>
<td>FTS can work under the linguistic time series</td>
</tr>
<tr>
<td>✓</td>
<td>FTS model can be established under the time space or state space conditions</td>
</tr>
<tr>
<td>✓</td>
<td>FTS can be build under the multi-variate conditions <em>with special care on causality</em></td>
</tr>
</tbody>
</table>

Theoretical foundations have critical importance. Computer may find many causal links by data mining. Post-sample control is strongly recommended.
Is Fuzzy Time Series a fallacy?
Potentials and limitations of fuzzy set approach in time series analysis

Thanks for listening
ISF 2011, Prague
OkanDuru©2011

Okan Duru
Istanbul Technical University
Department of Maritime Transportation and Management Eng.
duruokan@yahoo.com